PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

A STUDY ON SOFTWARE DEFECT ORIGINS AND REMOVAL METHODS

Dr. M. Suresh Babu[1] D. Asha Devi[2]

ABSTRACT

The cost of finding and fixing bugs or defects is the largest single expense element in the history of software. Bug repairs start with requirements and continue through development. After release, bug repairs and related customer support costs continue until the last user signs off. Over a 25-year life expectancy of a large software system in the 10,000-function point size range, almost 50 cents out of every dollar will go to finding and fixing bugs. Given the fact that bug repairs are the most expensive element in the history of software, it might be expected that these costs would be measured carefully and accurately. They are not. Most companies do not measure defect repair costs, and when they do, they often use metrics that violate standard economic assumptions such as "lines of code" and "cost per defect" neither of which measure the value of software quality. Both of these measures distort quality economics. Lines of code penalize high-level languages. Cost per defect penalizes quality.

Poor measurement practices have led to the fact that a majority of companies do not know that achieving high levels of software quality will shorten schedules and lower costs at the same time. However, testing alone is insufficient. A synergistic combination of defect prevention, pre-test defect removal, and formal testing using mathematical methods all need to be part of the quality technology stack.

KEYWORDS

Bugs, Formal Testing, Lines of Code, Removal Methods etc.

INTRODUCTION

The software industry spends about $0.50 out of every $1.00 expended for development and maintenance on finding and fixing bugs. Most forms of testing are below 35% in defect removal efficiency or remove only about one bug out of three. All tests together seldom top 85% in defect removal efficiency. About 7% of bug repairs include new bugs. About 6% of test cases have bugs of their own. These topics need to be measured, controlled, and improved. Security flaws are leading to major new costs for recovery after attacks. Better security is a major subset of software quality. A synergistic combination of defect prevention, pre-test defect removal, and formal testing by certified personnel can top 99% in defect removal efficiency while simultaneously lowering costs and shortening schedules. For companies that know how to achieve it, high quality software is faster and cheaper than low quality software.

SOFTWARE DEFECT ORIGINS

Software defects originate in multiple origins. The approximate U.S. total for defects in requirements, design, code, documents, and bad fixes is 5.00 per function point. Best in class projects are below 2.00 per function point. Projects in litigation for poor quality can top 7.00 defects per function point.

Defect potentials circa 2012 for the United States average:

- Requirements 1.00 defects per function point,
- Design 1.25 defects per function point,
- Code 1.75 defects per function point,
- Documents 0.60 defects per function point,
- Bad fixes 0.40 defects per function point,
- Totals 5.00 defects per function point.

The major defect origins include:

- Functional requirements,
- Non-functional requirements,
- Architecture,
- Design,

[1]Professor & Head, Department of Computer Applications, Madanapalle Institute of Technology & Science, Andhra Pradesh, India, mcahod@mits.ac.in

[2]Professor, Department of ECE, Madanapalle Institute of Technology & Science, Andhra Pradesh, India, drashadevid@mits.ac.in

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

- New source code,
- Uncertified reused code from external sources,
- Uncertified reused code from legacy applications,
- Uncertified reused designs, architecture,
- Uncertified reused test cases,
- Documents (user manuals, HELP text etc.,
- Bad fixes or secondary defects in defect repairs (7% is U.S. average),
- Defects due to creeping requirements that bypass full quality controls,
- Bad test cases with defects in them (6% is U.S. average),
- Data defects in data bases and web sites,
- Security flaws that are invisible until exploited.

Far too much of the software literature concentrates on code defects and ignores the more numerous defects found in requirements and design. It is also interesting that many of the companies selling quality tools such as static analysis tools and test tools focus only on code defects.

Unless requirement and design defects are prevented or removed before coding starts, they will eventually find their way into the code where it may be difficult to remove them. It should not be forgotten that the famous "Y2K" problem ended up in code, but originated as a corporate requirement to save storage space.

Some of the more annoying Windows 8 problems, such as the hidden and arcane method needed to shut down Windows 8, did not originate in the code, but rather in questionable upstream requirements and design decisions.

PROVEN METHODS FOR PREVENTING AND REMOVING SOFTWARE DEFECTS

Defect Prevention

The set of defect prevention methods can lower defect potentials from U.S. averages of about 5.00 per function point down below 2.00 per function point. Certified reusable materials are the most effective known method of defect prevention. A number of Japanese quality methods are beginning to spread to other countries and are producing good results. Defect prevention methods include:

- Joint Application Design (JAD),
- Quality function deployment (QFD),
- Certified reusable requirements, architecture, and design segments,
- Certified reusable code,
- Certified reusable test plans and test cases (regression tests),
- Kanban for software (mainly in Japan),
- Kaizen for software (mainly in Japan),
- Poka-yoke for software (mainly in Japan),
- Quality circles for software (mainly in Japan),
- Six Sigma for Software,
- Achieving CMMI levels => 3 for critical projects,
- Using quality-strong methodologies such as RUP and TSP,
- Embedded users for small projects < 500 function points,
- Formal estimates of defect potentials and defect removal before starting projects,
- Formal estimates of cost of quality (COQ) and technical debt (TD) before starting,
- Quality targets such as > 97% defect removal efficiency (DRE) in all contracts,
- Function points for normalizing quality data,
- Analysis of user-group requests or customer suggestions for improvement.

Analysis of software defect prevention requires measurement of similar projects that use and do not use specific approaches such as JAD or QFD of necessity studying defect prevention needs large numbers of projects and full measures of their methods and results.

Note that two common metrics for quality analysis, "lines of code" and "cost per defect" have serious flaws and violate standard economic assumptions. These two measures conceal, rather than reveal, the true economic value of high software quality.

Function point metrics are the best choice for quality economic studies. The new SNAP non-functional size metric has recently been released, but little quality data is available because that metric is too new.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Pre-Test Defect Removal

The most effective known methods of eliminating defects circa 2012 include requirements models, automated proofs, formal inspections of requirements, design, and code; and static analysis of code and text. These methods have been measured to top 85% in defect removal efficiency individually. Methods such as inspections also raise testing defect removal efficiency by more than 5% for each major test stage. The major forms of pre-test defect removal include:

- Desk checking by developers,
- Debugging tools (automated),
- Pair programming (with caution),
- Quality Assurance (QA) reviews of major documents and plans,
- Formal inspections of requirements, design, code, UML, and other deliverables,
- Formal inspections of requirements changes,
- Informal peer reviews of requirements, design, code,
- Editing and proof reading critical requirements and documents,
- Text static analysis of requirements, design,
- Code static analysis of new, reused, and repaired code,
- Running FOG and FLESCH readability tools on text documents,
- Requirements modeling (automated),
- Automated correctness proofs,
- Refactoring,
- Independent verification and validation (IV&V).

Pre-test inspections have more than 40 years of empirical data available and rank as the top method of removing software defects, consistently topping 85% in defect removal efficiency (DRE).

Static analysis is a newer method that is also high in DRE, frequently toping 65%. Requirements modeling are another new and effective method that has proved itself on complex software such as that operating the Mars Rover. Requirements modeling and inspections can both top 85% in defect removal efficiency (DRE).

One of the more unusual off shoots of some of the agile methods such as extreme programming (XP) is "pair programming." The pair programming approach is included in the set of pre-test defect removal activities.

With pair programming two individuals share an office, work station, and take turns coding while the other observes.

This should have been an interesting experiment, but due to poor measurement practices, it has started into actual use, with expensive results. Individual programmers who use static analysis and inspections have better quality at about half the cost and 75% of the schedule of a pair.

If two top guns are paired the results will be good, but the costs about 40% higher than either one working alone. Since there is a severe shortage of top-gun software engineers, it is not cost effective to have two of them working on the same project. It would be better for each of them to tackle a separate important project. Top-guns only comprise about 5% of the overall software engineering population.

If a top gun is paired with an average programmer, the results will be better than the average team member might product, but about 50% more expensive. The quality is no better than the more experienced pair member working alone. If pairs are considered a form of mentoring there is some value for improving the performance of the weaker team member.

If two average programmers are paired the results will still be average, and the costs will be about 80% higher than either one alone.

If a marginal or unqualified person is paired with anyone, the results will be suboptimal and the costs about 100% higher than the work of the better team member working alone. This is because the unqualified person is a drag on the performance of the pair.

Since there are not enough qualified top-gun programmers to handle all of the normal work in many companies, doubling them up adds costs but subtracts from the available work force. There are also 115 occupation groups associated with software. If programmers are to be paired, why not pair architects, designers, testers, and project managers?

Pairing should have been measured and studied prior to becoming an accepted methodology, but instead it was put into production with little or no empirical data. This phenomenon of rushing to use the latest fad without any proof that it works is far too common for software.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Most of the studies of pair programming do not include the use of inspections or static analysis. They merely take a pair of programmers and compare the results against one unaided programmer who does not use modern pre-test removal methods such as static analysis and peer reviews. By excluding other forms of pre-test defect removal such as inspections and static analysis, the studies of pair programming are biased and incomplete.

Test Defect Removal

Testing has been the primary software defect removal method for more than 50 years. Unfortunately, most forms of testing are only about 35% efficient or find only one bug out of three.

Defects in test cases themselves and duplicate test cases lower test defect removal efficiency. About 6% of test cases have bugs in the test cases themselves. In some large companies as many as 20% of regression, test libraries are duplicates, which add to testing costs but not to testing rigor.

Due to low defect removal efficiency, at least eight forms of testing are needed to achieve reasonably efficient defect removal efficiency. Pre-test inspections and static analysis are synergistic with testing and raise testing efficiency.

Tests by certified test personnel using test cases designed with formal mathematical methods have the highest levels of test defect removal efficiency and can top 65%. The major forms of test-related factors for defect removal include:

- Certified test personnel,
- Formal test plans published and reviewed prior to testing,
- Certified reusable test plans and test cases for regression testing,
- Mathematically based test case design such as using design of experiments,
- Test coverage tools for requirements, code, data, etc.,
- Automated test tools,
- Cyclomatic complexity tools for all new and changed code segments,
- Test library control tools,
- Capture-recapture testing,
- Defect tracking and routing tools,
- Inspections of major code changes prior to testing,
- Inspections of test libraries to remove bad and duplicate test cases,
- Special tests for special defects; performance, security, etc.
- Full suite of test stages including:
    - Subroutine test,
    - Unit test,
    - New function test,
    - Regression test,
    - Component test,
    - Performance test,
    - Usability test,
    - Security test,
    - System test,
    - Supply-chain test,
    - Cloud test,
    - Data migration test,
    - ERP link test,
    - External beta test,
    - Customer acceptance test,
    - Independent test (primarily military projects).

Testing by itself without any pre-test inspections or static analysis is not sufficient to achieve high quality levels. The poor estimation and measurement practices of the software industry have long slowed progress on achieving high quality in a cost-effective fashion. However modern risk-based testing by certified test personnel with automated test tools who also use mathematically-derived test case designs and also tools for measuring test coverage and cyclomatic complexity can do a very good job and top 65% in defect removal efficiency for the test stages of new function test, component test, and system test.

Untrained amateur personnel such as developers themselves seldom top 35% for any form of testing. In addition, the "bad fix injection" rate or new bugs added while fixing older bugs tops 7% for repairs by ordinary development personnel. Bad fixes are inversely proportional to cyclomatic complexity, and inversely proportional to experience. Bad fixes by a top-gun software engineer working with software with low cyclomatic complexity can be only a fraction of 1%.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

At the other end of the spectrum, bad fixes by a novice trying to fix a bug in an error-prone module with high cyclomatic complexity can top 25%. Another issue that is seldom discussed in the literature is that of bugs or errors in the test cases themselves. On average about 6% of test cases, contain errors. Running defective test cases adds costs to testing but nothing to defect removal efficiency. In fact defective test cases lower DRE.

National Defect Removal Efficiency and Software Quality

Some countries such as Japan, India, and South Korea place a strong emphasis on quality in all manufactured products including software. Other countries, such a China and Russia, apparently have less interest and less understanding of quality economics and seem to lag in quality estimation and measurement. Among the quality-strong countries Japan, for example, had more than 93% DRE on the projects.

All of the countries in the top 20 can produce excellent software and often do. Countries with significant amounts of systems and embedded software and defense software are more likely to have good quality control than countries producing mainly information technology packages. Almost all countries in 2012 produce software in significant volumes. More than 150 countries produce millions of function points per year. The preliminary ranks shown here indicate that more studies are needed on international software quality initiatives.

Quality measures and predictive quality estimation are necessary precursors to achieving top quality status. Defect prevention and pre-test defect removal must be added to testing to achieve top-rank status.

Industry Defect Removal Efficiency and Software Quality

In general, the industries that produce complex physical devices such as airplanes, computers, medical devices, and telephone switching systems have the highest levels of defect removal efficiency, the best quality measures, and the best quality estimation capabilities.

This is a necessity because these complex devices will not operate unless quality approaches zero defects. In addition, the manufacturers of such devices have major liabilities in case of failures including possible criminal charges. The top 20 industries in terms of defect removal efficiency are:

Industries

- Government – Intelligence Agencies,
- Manufacturing – Medical Devices,
- Manufacturing – Aircraft,
- Manufacturing – Mainframe Computers,
- Manufacturing – Telecommunications Switching Systems,
- Telecommunications – Operations,
- Manufacturing – Defense Weapons Systems,
- Manufacturing – Electronic Devices and Smart Appliances,
- Government – Military Services,
- Entertainment – Film and Television Production,
- Manufacturing – Pharmaceuticals,
- Transportation – Airlines,
- Manufacturing – Tablets And PC's,
- Software – Commercial,
- Manufacturing – Chemicals and Process Control,
- Banks – Commercial,
- Banks – Investment,
- Health Care – Medical Records,
- Software – Open Source,
- Finance – Credit Unions.

As of 2012, more than 300 industries produce software in significant volumes. Some of the lagging industries that are near the bottom in terms of software defect removal efficiency levels are those of state governments, municipal governments, wholesale chains, retail chains, public utilities, cable television billing and finance, and some but not all insurance companies in the areas of billing and accounting software.

Companies that are considering outsourcing may be curious as to the placement of software outsource vendors. From the author's studies of various industries, outsource vendors rank as number 25 out of 75 industries for ordinary information technology

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

outsourcing. For embedded and systems software outsourcing the outsource vendors are approximately equal to industry averages for aircraft, medical device, and electronic software packages.

Another interesting question is how good do quality companies themselves such as static analysis companies, automated test tool companies, independent testing companies, and defect tracking tool companies practice the defect removal methods?

Interestingly, these companies publish no data about their own results and seem to avoid having outside consulting studies done that would identify their own defect removal efficiency levels. No doubt, the static analysis companies use their own tools on their own software, but they do not publish accurate data on the measured effectiveness of these tools.

All of the test and static analysis companies should publish annual reports that show ranges of defect removal efficiency (DRE) results using their tools, but none is known to do this.

Software Development Methods

Some software development methods such as IBM's Rational Unified Process (RUP) and Watts Humphrey's Team Software Process (TSP) can be termed "quality strong" because they lower defect potentials and elevate defect removal efficiency levels.

Other methods such as Waterfall and Cowboy development can be termed "quality weak" because they raise defect potentials and have low levels of defect removal efficiency. The 30 methods shown here are ranked in approximate order of quality strength. The list is not absolute and some methods are better than others are for specific sizes and types of projects. Development methods in rank order of defect prevention include:

- Mashup (construction from certified reusable components),
- Hybrid,
- IntegraNova,
- TSP/PSP,
- RUP,
- T-VEC,
- Extreme Programming (XP),
- Agile/Scrum,
- Data state design (DSD),
- Information Engineering (IE),
- Object-Oriented (OO),
- Rapid Application Development (RAD),
- Evolutionary Development (EVO),
- Jackson development,
- Structured Analysis and Design Technique (SADT),
- Spiral development,
- Structured systems analysis and design method (SSADM),
- Iterative development,
- Flow-based development,
- V-Model development,
- Prince2,
- Merise,
- Data state design method (DSDM),
- Clean-room development,
- ISO/IEC,
- Waterfall,
- Pair programming,
- DoD 2167A,
- Proofs of correctness (manual),
- Cowboy.

Once again, this list is not absolute and situations change. Since agile development is so popular, it should be noted that Agile is strong in quality but not the best in quality. Agile projects frequently achieve DRE in the low 90% range, which is better than average but not top-ranked.

Agile lags many leading methods in having very poor quality measurement practices. The poor measurement practices associated with Agile for both quality and productivity will eventually lead CIO's, CTO's, CFO's, and CEO's to ask if actual agile results are as good as being claimed.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Until **Agile** projects publish productivity data using function point metrics and quality data using function points and defect removal efficiency (DRE) the effectiveness of Agile remains ambiguous and uncertain. We found Agile to be superior in both quality and productivity to waterfall development, but not as good for quality as either RUP or TSP.

In addition, a Google search using phrases such as "Agile failures" and "Agile successes" turns up about as many discussions of failure as success. A new occupation of "Agile coach" has emerged to help reduce the instances of getting off track when implementing Agile.

Overall Quality Control

Successful quality control stems from a synergistic combination of defect prevention, pre-test defect removal, and test stages. The best projects in the industry circa 2012-combined defect potentials in the range of 2.0 defects per function point with cumulative defect removal efficiency levels that top 99%. The U.S. average circa 2012 is about 5.0 bugs per function point and only about 85% defect removal efficiency.

The major forms of overall quality control include:

- Formal software quality assurance (SQA) teams for critical projects,
- Measuring defect detection efficiency (DDE),
- Measuring defect removal efficiency (DRE),
- Targets for topping 97% in DRE for all projects,
- Targets for topping 99% in DRE for critical projects,
- Inclusion of DRE criteria in all outsource contracts (> 97% is suggested),
- Formal measurement of cost of quality (COQ),
- Measures of "technical debt" but augmented to fill major gaps,
- Measures of total cost of ownership (TCO) for critical projects,
- Monthly quality reports to executives for on-going and released software,
- Production of an annual corporate software status and quality report,
- Achieving > CMMI level 3.

IBM started to measure defect origins, defect potentials, and defect removal efficiency (DRE) levels in the early 1970's. These measures were among the reasons for IBM's market success in both hardware and software. High quality products are usually cheaper to produce, are much cheaper to maintain, and bring high levels of customer loyalty.

The original IBM DRE studies used six months after release for calculating DRE, but due to updates that occur before six months, that interval was difficult to use and control. The switch from six month to 90-day DRE intervals occurred in 1984.

Defect removal efficiency is measured by accumulating data on all bugs found prior to release and on bugs reported by clients in the first 90 days of use. If developers found 90 bugs and users reported 10 bugs then DRE is clearly 90%.

The International Software Benchmark Standards Group (ISBSG) uses only a 30-day interval after release for measuring DRE. The author measures both 30-day and 90-day intervals.

Unfortunately, the 90-day defect counts average about four to five times larger than the 30-day defect counts, due to installation and learning curves of software, which delay normal usage until late in the first month.

A typical 30-day ISBSG count of DRE might show 90 bugs found internally and 2 bugs found in 30 days, for a DRE of 97.82%.

A full 90-day count of DRE would still show 90 bugs found internally but 10 bugs found in three months for a lower DRE of only 90.00%.

Although a fixed time interval is needed to calculate DRE that does not mean that all bugs are found in only 90 days. In fact, the 90-day DRE window usually finds less than 50% of the bugs reported by clients in one calendar year.

Bug reports correlate strongly with numbers of production users of software applications. Unless a software package is something like Windows 8 with more than 1,000,000 users on the first day, it usually takes at least a month to install complex applications, train users, and get them started on production use. If there are less than 10 users the first month, there will be very few bug reports. Therefore, in addition to measuring DRE, it is also significant to record the numbers of users for the first three months of the application's production runs. If we assume an ordinary information technology application the following table shows the probable numbers of reported bugs after one, two, and three months for 10, 100, and 1000 users:

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Table-1: Defects by Users for Three Months

Month	10 Users	100 Users	1000 Users
1	1	3	6
2	3	9	18
3	6	12	24

Sources: Authors Compilation

As it happens, the central column of 100 users for three months is a relatively common pattern.

Note that for purposes of measuring defect removal efficiency a single month of usage tends to yield artificially high levels of DRE due to a normal lack of early users.

Companies such as IBM with continuous quality data are able to find out many interesting and useful facts about defects that escape and are delivered to clients. For example for financial software, there will be extra bug reports at the end of standard fiscal years, due to exercising annual routines. Also of interest is the fact that about 15% of bug reports are "invalid" and not true bugs at all. Some are user errors, some are hardware errors, and some are bugs against other software packages that were mistakenly reported to the wrong place. It is very common to confuse bugs in operating systems with bugs in applications.

As an example of an invalid defect report, the author's company once received a bug report against a competitive product, sent to us by mistake. Even though this was not a bug against our software, we routed it to the correct company and sent a note back to the originator as a courtesy. It took about an hour to handle a bug against a competitive software package. Invalid defects such as this do not count as technical debt or cost of quality (COQ). However, they do count as overhead costs.

An interesting new metaphor called "technical debt" was created by Ward Cunningham and is now widely deployed, although most companies do not deploy it the same way. Several software quality companies such as OptiMyth in Spain, CAST Software, and SmartBear feature technical debt discussions on their web sites.

The concept of technical debt is intuitively appealing. Shortcuts made during development that lead to complex code structures or to delivered defects will have to be fixed at some point in the future. When the time comes to fix these problems downstream, the costs will be higher and the schedules longer than if they had been avoided in the first place. The essential concept of technical debt is that questionable design and code decisions have increasing repair costs over time. As a metaphor or interesting concept, technical debt has much to recommend it.

However, the software industry is far from sophisticated in understanding finance and economic topics. In fact, for more than 50 years the software industry has tried to measure quality costs with "lines of code" and "cost per defect" which are so inaccurate as to be viewed as professional malpractice for quality economics. In addition, many companies only measure about 37% of software project effort and 38% of software defects. Omitting unpaid overtime, managers, and specialists are common gaps. Omitting bugs found in requirements, design, and by unit testing are common quality omissions.

Technical debt runs head on into the general ineptness of the software world in understanding and measuring the older cost of quality (COQ) in a fashion that matches standard economic assumptions. Cost per defect penalizes quality. Lines of code penalize modern high-level languages and of course make requirements and design defects invisible. Defect repair costs per function point provide the best economic indicator. However, the new SNAP metric for non-functional requirements needs to be incorporated.

The main issues with technical debt as widely deployed by the author's clients are that it does not include or measure some of the largest quality costs in all of software history. About 35% of large software systems are cancelled and never delivered at all. The most common reason for cancellation is poor quality. However, since the cancelled projects are not delivered, there are no downstream costs and hence no technical debt either. The costs of cancelled projects are much too large to ignore and just leave out of technical debt.

The second issue involves software that is delivered and indeed accumulates technical debt in the form of changes that need to be repaired. However, some software applications have such bad quality that clients sue the developers for damages. The costs of litigation and the costs of any damages that the court orders software vendors to pay should be part of technical debt.

For financial debt, there is a standard set of principles and practices called the "Generally Accepted Accounting Principles" or GAAP. The software industry in general, and technical debt in particular, need a similar set of "Software Generally Accepted Accounting Principles" or SGAAP that would allow software projects and software costs to be compared in a uniform fashion.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Software engineers interested in technical debt are urged to read the GAAP and IFRS accounting standards and familiarize themselves with normal cost accounting as a precursor to applying technical debt. The major GAAP principles are relevant to software measures and to technical debt:

- Principle of regularity,
- Principle of consistency,
- Principle of sincerity,
- Principle of permanence of methods,
- Principle of non-compensation or not replacing a debt with an asset,
- Principle of prudence,
- Principle of continuity,
- Principle of periodicity,
- Principle of full disclosure,
- Principle of utmost good faith.

The major software metric associations such as the International Function Point User Group (IFPUG) and the Common Software Metric International Consortium (COSMIC) should both be participating in establishing common financial principles for measuring software costs, including cost of quality and technical debt. However, neither group has done much outside of basic sizing of applications. Financial reporting is still ambiguous for the software industry as a whole. Many companies are trying to use technical debt because it is an intriguing metaphor that is appealing to CFO's and CEO's. However, without some form of SIFRS or standardized accounting principles every company in every country is likely to use technical debt with random rules that would not allow cross-country, cross-company, or cross-project comparisons.

Harmful Practices to be Avoided

Some of the observations of harmful practices stem from lawsuits where the author has worked as an expert witness. Discovery documents and depositions reveal quality flaws that are not ordinarily visible or accessible to standard measurements. In every case where poor quality was alleged by the plaintiff and proven in court, there was evidence that defect prevention was lax, pre-test defect removal such as inspections and static analysis were bypassed, and testing was either perfunctory or truncated to meet arbitrary schedule targets. These poor practices were unfortunate because a synergistic combination of defect prevention, pre-test defect removal, and formal testing leads to short schedules, low costs, and high quality at the same time. The most severe forms of schedule slips are due to starting testing with excessive numbers of latent defects, which stretch out testing intervals by several hundred percent compared to original plans. Harmful and dangerous practices to be avoided are:

- Bypassing pre-test inspections,
- Bypassing static analysis,
- Testing by untrained, uncertified amateurs,
- Truncating testing for arbitrary reasons of schedule,
- The "good enough" quality fallacy,
- Using "lines of code" for data normalization (professional malpractice),
- Using "cost per defect" for data normalization (professional malpractice),
- Failure to measure bugs at all,
- Failure to measure bugs before release,
- Failure to measure defect removal efficiency (DRE),
- Error-prone modules (EPM) with high defect densities,
- High cyclomatic complexity of critical modules,
- Low test coverage of critical modules,
- Bad-fix injections or new bugs in bug repairs themselves,
- Outsource contracts that do not include quality criteria and DRE,
- Duplicate test cases that add costs but not test thoroughness,
- Defective test cases with bugs of their own.

It is an unfortunate fact that poor measurement practices, failure to use effective quality predictions before starting key projects, and bypassing defect prevention and pre-test defect removal methods have been endemic problems of the software industry for more than 40 years.

Poor software quality is like the medical condition of whooping cough. That condition can be prevented via vaccination and in today's world treated effectively. Poor software quality can be eliminated by the "vaccination" of early estimation and effective defect prevention. Pre-test defect removal such as inspections and static analysis are effective therapies. Poor software quality is a completely treatable and curable condition.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

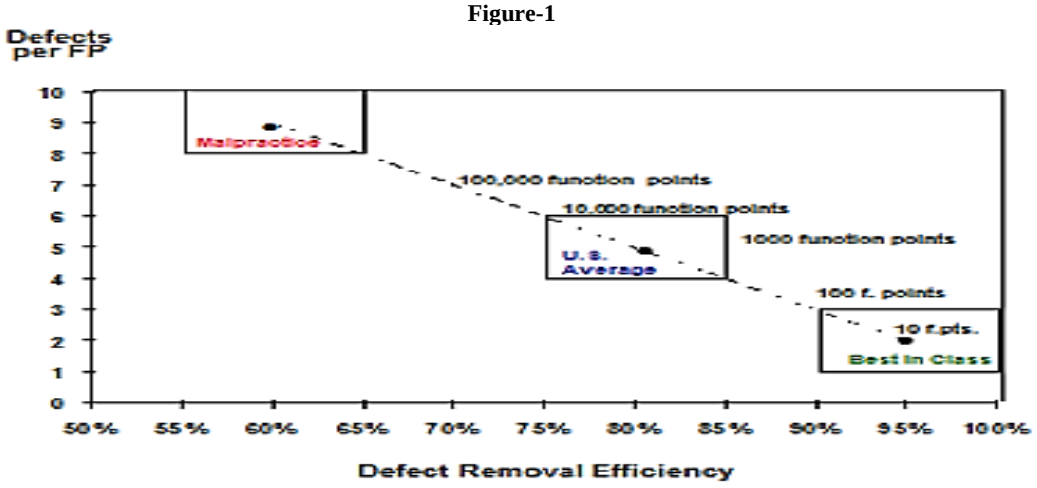
It is technically possible to lower defect potential from around 5.00 per function point to below 2.00 per function point. It is also technically possible to raise defect removal efficiency (DRE) from today's average of about 85% to at least 99%. These changes would also shorten schedules and reduce costs.

Illustrating Software Defect Potentials and Defect Removal Efficiency (DRE).

Figure one shows overall software industry results in terms of two dimensions. The vertical dimension shows defect potentials or the probable total number of bugs that will occur in requirements, design, code, documents, and bad fixes. Note that large systems have much higher defect potentials than small applications. It is also harder to remove defects from large systems.

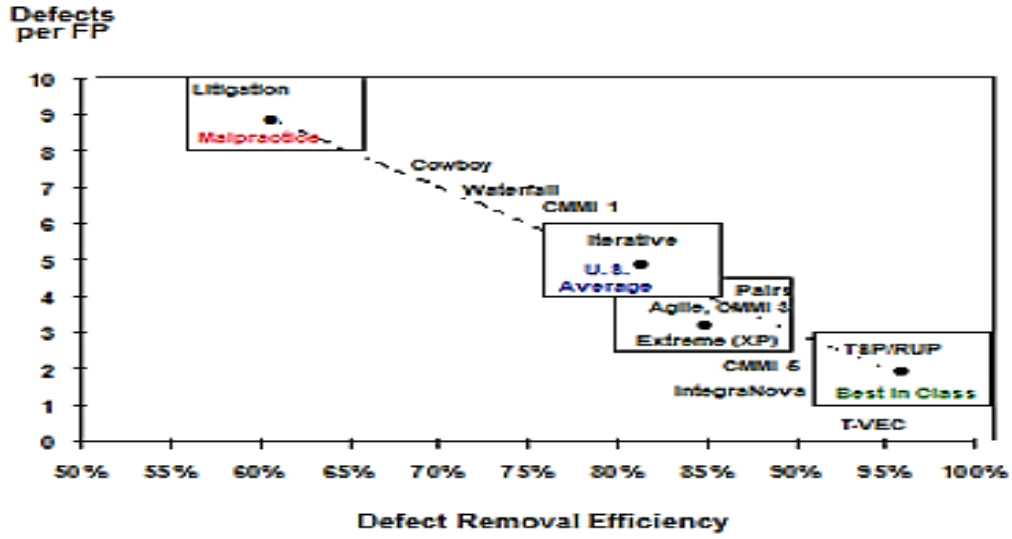
Figure-1 shows the relationship between software methodologies and software defect potentials and defect removal:

Figure-1



Sources: Authors Compilation

Figure 2: Methodologies and Software Defect Potentials and Removal



Sources: Authors Compilation

SUMMARY AND CONCLUSIONS ON SOFTWARE QUALITY

The software industry spends more money on finding and fixing bugs than for any other known cost driver. This should not be the case. A synergistic combination of defect prevention, pre-test defect removal, and formal testing can lower software defect removal costs by more than 50% compared to 2012 averages. These same synergistic combinations can raise defect removal efficiency (DRE) from the current average of about 85% to more than 99%.

Any company or government group that averages below 95% in cumulative defect removal efficiency (DRE) is not adequate in software quality methods and needs immediate improvements. Any company or government group that does not measure DRE and does not know how efficient they are in finding software bugs prior to release is in urgent need of remedial quality improvements. Inadequate defect prevention and inadequate pre-test defect removal are strongly correlated with failure to measure defect removal efficiency.

REFERENCES

1. Beck, Ken. (2002). ***Test-Driven Development.*** Addison Wesley. MA: Boston. ISBN 10: 0321146530; 240 pages.

2. Black, Rex (2009). ***Managing the Testing Process: Practical Tools and Techniques for Managing Hardware and Software Testing,*** 672 pages**.** Wiley. ISBN-10 0470404159.

3. Chelf, Ben, & Jetley, Raoul. (2008). ***Diagnosing Medical Device Software Defects Using Static Analysis*** (Coverity Technical Report). CA: San Francisco.

4. Chess, Brian, & West, Jacob. (2007). ***Secure Programming with Static Analysis,*** 624 pages**.** Boston. MA: Addison Wesley. ISBN 13: 978-0321424778.

5. Cohen, Lou. ***Quality Function Deployment – How to Make QFD Work for You***, 368 pages. Upper Saddle River. NJ: Prentice Hall. ISBN 10: 0201633302.

6. Everett, Gerald D., & McLeod, Raymond. (2007). ***Software Testing,*** 261 pages. Hoboken. NJ: John Wiley & Sons. ISBN 978-0-471-79371-7.

7. Gack, Gary. (2010). ***Managing the Black Hole: The Executives Guide to Software Project Risk.*** Thomson. GA: Business Expert Publishing. ISBN10: 1-935602-01-9.

8. Gack, Gary. ***Applying Six Sigma to Software Implementation Projects.*** Retrieved from http://software.isixsigma.com/library/content/c040915b.asp .

9. Gilb, Tom, & Graham, Dorothy. (1993). ***Software Inspections***. Addison Wesley. MA: Reading, ISBN 10: 0201631814.

10. Hallowell, David L. ***Six Sigma Software Metrics, Part 1.*** Retrieved from http://software.isixsigma.com/library/content/03910a.asp .

11. ***International Organization for Standards.*** ISO 9000 / ISO 14000. Retrieved from http://www.iso.org/iso/en/iso9000-14000/index.html .

12. Jones, Capers, & Bonsignour, Olivier. (2011). ***The Economics of Software Quality,*** 587 pages. Addison Wesley. MA; Boston. ISBN 978-0-13-258220-9.

13. Retrieved from http://insights.cermacademy.com/2012/12/6-software-defect-origins-and-removal-methods-c-capers-jones.. .

14. Retrieved from http://www.ifpug.org/Documents/Jones-SoftwareDefectOriginsAndRemovalMethodsDraft5.pdf

15. Retrieved from http://www.itmpi.org/Portals/10/PDF/Software%20Defect%20Origins%20and%20Removal%20Methods2013.pdf

*****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

IDENTIFICATION AND VERIFICATION OF SPEAKER USING MEL FREQUENCY CEPSTRAL COEFFICIENT

Ashish Sharma[3] Viplav Gautam[4] Saurabh Sharma[5] Swapnil Gautam[6] Gaurav Sharma[7]

ABSTRACT

Speech processing is emerged as one of the important application area of digital signal processing. Various fields for research in speech processing are speech recognition, speaker recognition, speech synthesis, speech coding etc. Feature extraction is the most important step for speaker recognition. In this work, the Mel Frequency Cepstrum Coefficient (MFCC) feature has been used for designing a text dependent speaker identification system. MFCC is based on the human peripheral auditory System. Generally, MFCC for feature extraction is used to improve the efficiency of speaker recognition.

KEYWORDS

Feature Extraction, Feature Matching, Mel Frequency Cepstral Coefficients (MFCC), Speaker Recognition etc.

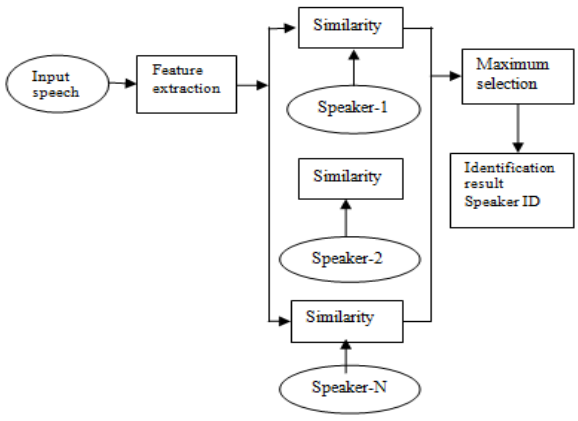
INTRODUCTION

As human beings, we are able to recognize someone just by hearing him or her talk. Usually, a few seconds of speech are sufficient to identify a familiar voice. Speech contains significant energy from zero frequency up to around 5 kHz. The objective of speaker recognition is to Extract, characterize and recognize the information about Speaker identity. At the primary level, speech conveys a message via words. However, at other levels speech conveys information about the Language being spoken and the emotion, gender and, generally, the identity of the speaker. To study the spectral properties of speech signal the concept of time varying Fourier representation is used.

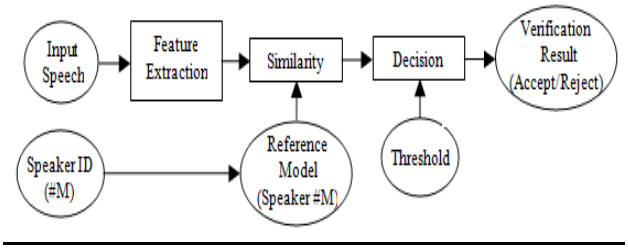
Speaker recognition is divided into 2 classifications: speaker recognition and speaker identification and it is the method of automatically identify who is speaking based on individual information integrated in speech waves. Speaker identification is the task of determining who is talking from a set of known voices or speakers and Speaker verification is the task of determining whether a person is who he/she claims to be. The main aim of project is speaker identification, which consists of comparing a speech signal from an unknown speaker to a database of known speaker. The system can recognize the speaker, which has been trained with number of speakers. Below figure shows the fundamental formation of speaker identification & verification systems.

Figure-1: Basic Structures of Speaker Recognition Systems

a) Speaker Identification



b) Speaker Verification



Sources: Authors Compilation

[3]Associate Professor & H.O.D., Department of Computer Sciences, Maharishi Arvind International Institute of Technology, Rajasthan, ashishsharma.pce@gmail.com
[4]Student, ECE Department, Global Institute of technology, Rajasthan, India, viplovecgautam@gmail.com
[5]Student, Engineering Physics Department, IIT, Delhi, India, sgnsharma123@gmail.com
[6]Student, Engineering Physics Department, IIT, Delhi, India, swapnilgautamiitd@gmail.com
[7]Student, Metallurgy Department, MNIT, Rajasthan, gauravsharma000005@yahoo.in

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Speaker recognition can also divide into two methods, text- dependent and text independent methods. In text dependent method the speaker to say key words or sentences having the same text for both training and recognition trials. Whereas in the text independent does not rely on a specific texts being speak. Formerly text dependent methods were widely in application.

Like any other pattern recognition systems, speaker recognition systems also involve two phases namely, training and testing. Training is the process to upload the system with the voice characteristics of the speakers registering. Testing is the actual recognition task. The block diagram of training phase is shown in Figure below. In training phase, the voice characteristics of the speaker are extracted from the training utterances and are used for building the reference models. During testing, similar feature vectors are extracted from the test utterance, and the degree of their match with the reference is obtained using some matching technique. The level of match is used to arrive at the decision.

Figure-2: Block Diagram of Training Phase

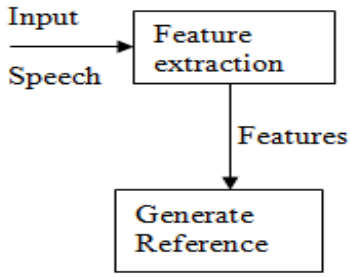
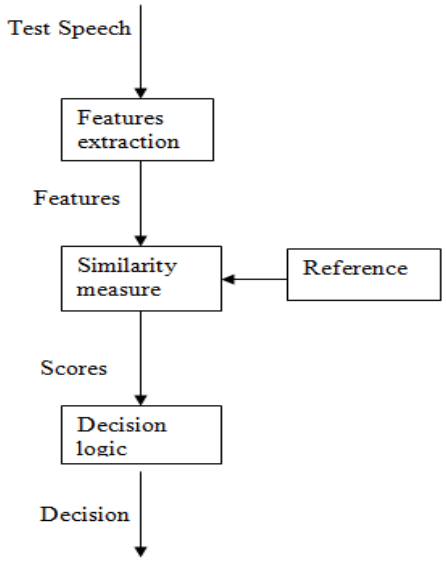


Figure-3: Block Diagram of Testing Phase



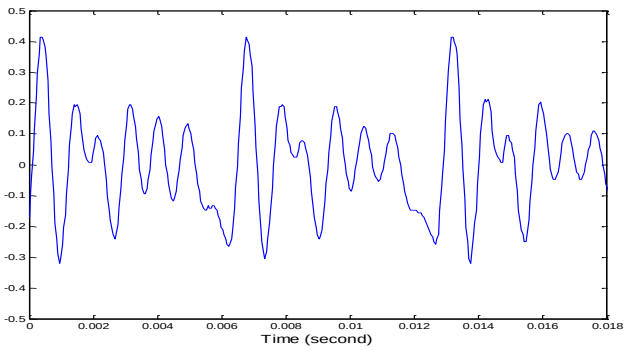
Sources: Authors Compilation

SPEECH FEATURE EXTRACTION

Introduction

The purpose of this module is to convert the speech waveform, using digital signal processing (DSP) tools, to a set of features for further analysis. This is often referred as the signal-processing front end. The speech signal is a slowly timed varying signal. An example of speech signal is shown in Figure below. When examined over a sufficiently short period (between 5 and 100 msec), its characteristics are stationary. However, over long period (about 1/5 seconds or more) the signal characteristic change to reflect the different speech sounds being spoken. Therefore, short-time spectral analysis is the most common way to characterize the speech signal.

Figure-4: Example of Speech Signal



Sources: Authors Compilation

A wide range of possibilities exists for parametrically representing the speech signal for the speaker recognition task, such as Linear Prediction Coding (LPC), Mel-Frequency Cepstrum Coefficients (MFCC), and others. MFCC will be discussed in this paper, because MFCC is perhaps the best known and most popular and also, it shows high accuracy results for clean speech and experiments show that the parameterization of the Mel frequency Cepstral coefficients is best for discriminating speakers and is different from the one usually used for speech recognition applications.

Steps of MFCC

Step 1 - ***Frame Blocking***

In this step the continuous speech signal is blocked into frames of *N* samples, with adjacent frames being separated by *M* (*M* < *N*). The first frame consists of the first *N* samples. The second frame begins *M* samples after the first frame, and overlaps it by *N* - *M* samples and. Similarly, the third frame begins 2M samples after the first frame (or M samples after the second frame) and overlaps it by N - 2M samples. This process continues until all the speech is accounted for within one or more frames. Typical values for *N* and *M* are *N* = 256 and *M* = 100.

Step 2 - ***Windowing***

The next step in the processing is to window each individual frame to minimize the signal discontinuities at the beginning and end of each frame. The concept here is to minimize the spectral distortion by using the window to taper the signal to zero at the beginning and end of each frame. If we define the window as w (n), 0<n<N-1, where *N* is the number of samples in each frame, then the result of windowing is the signal.

$$y_l(n) = x_l(n)w(n), \quad 0 \le n \le N-1$$

Typically, the *Hamming* window is used, which has the form:

$$w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right), \quad 0 \le n \le N-1$$

Step 3 - ***Fast Fourier transform***

The next processing step is the Fast Fourier Transform, which converts each frame of *N* samples from the time domain into the frequency domain. The FFT is a fast algorithm to implement the Discrete Fourier Transform (DFT), which is defined on the set of *N* samples $\{x_n\}$, as follow:

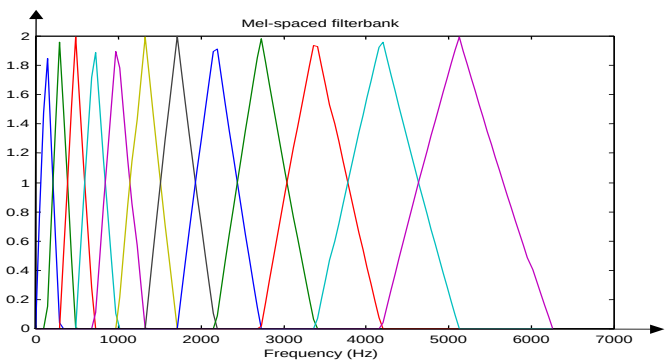
$$X_k = \sum_{n=0}^{N-1} x_n e^{-j2\pi kn/N}, \qquad k = 0,1,2,\ldots, N-1$$

Step 4- ***Mel-frequency Wrapping***

In sound processing, MFCC's are based on the known variation of the human ear's critical bandwidths. It is derived from the Fourier Transform of the audio clip. In this technique the frequency bands are positioned logarithmically, whereas in the Fourier Transform the frequency bands are not positioned logarithmically. As the frequency bands are positioned logarithmically in MFCC, it approximates the human system response more closely than any other system. These coefficients allow better processing of data. Each tone with an actual frequency t measured in Hz, a subjective pitch is measured on a scale called the 'Mel Scale'. The Mel frequency scale is linear frequency spacing below 1000 Hz and logarithmic spacing above 1 kHz. As a reference point, the pitch of a 1 kHz tone, 40 dB above the perceptual hearing threshold, is defined as 1000 Mels. Therefore we can use the following formula to determine the Mels for a given frequency f in Hz. Mel (*f*) = 2595*log10 (1 + *f*/700).

To obtain the subjective spectrum we use a filter bank, which is spaced uniformly on the Mel scale is described on the figure below. That filter bank has a triangular band pass frequency response, and the spacing as well as the bandwidth is determined by a constant Mel frequency interval.

Figure-5: Example of Mel-spaced Filter Bank



Sources: Authors Compilation

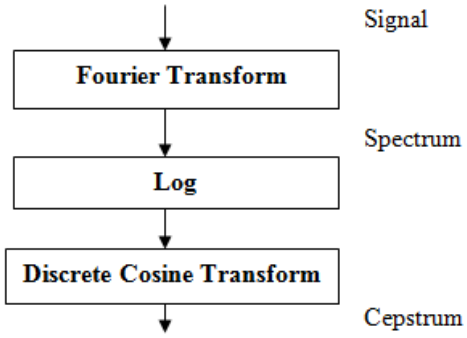
Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Step 5-***Cepstrum***
Cepstrum name was derived from the spectrum by reversing the first four letters of spectrum. We can say Cepstrum is the Fourier Transformer of the log with unwrapped phase of the Fourier Transformer.

Mathematically we can say Cepstrum of signal = FT (log (FT (the signal)) +j6.28m)
Where m is the integer required to properly unwrap the angle or imaginary part of the complex log function.
Algorithmically we can say – Signal - FT - log - phase unwrapping - FT -Cepstrum.

We can calculate the Cepstrum by many ways. Some of them need a phase-warping algorithm, others do not. Figure below shows the pipeline from signal to Cepstrum.

Figure-6: Signal to Cepstrum Pipeline



Sources: Authors Compilation

In this final step log Mel spectrum is converted back to time. The result is called the Mel Frequency Cepstrum Coefficients (MFCC).The discrete cosine transform is done for transforming the Mel coefficients back to time domain.

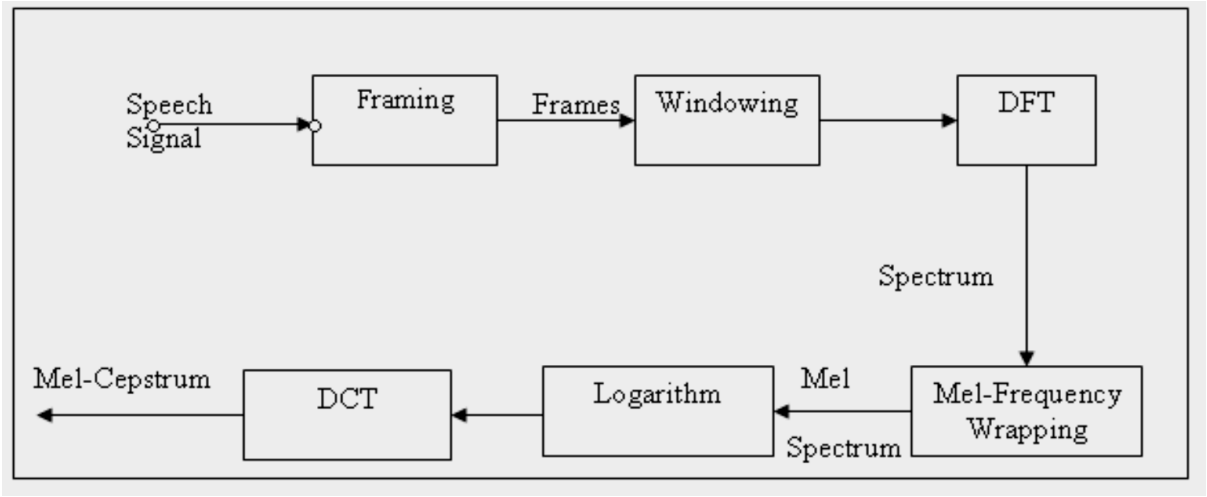
$$\tilde{c}_n = \sum_{k=1}^{K} (\text{Log}\tilde{S}_k) \cos\left[n\left(k - \frac{1}{2} \right) \frac{\pi}{K} \right], \qquad n = 0,1,\ldots, K-1$$

Where $\tilde{S}_0, k = 0,2,\ldots, K-1$

$\tilde{c}_0$, is excluded from DCT because it represents the mean value of the input signal, which carried little speaker specific information.

The complete figure, which shows the calculation of the Mel frequency Cepstrum coefficient, is shown below:

Figure-7: Complete Pipeline for MFCC



Sources: Authors Compilation

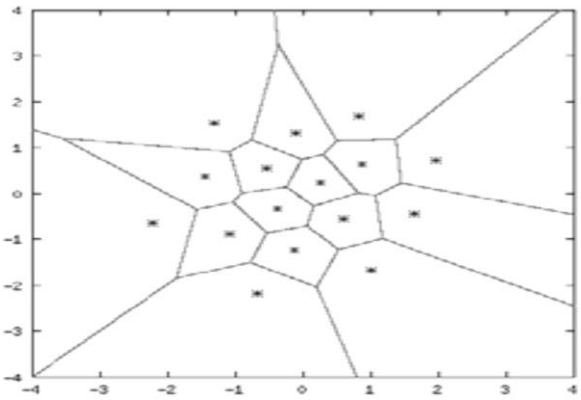
Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

SPEECH FEATURE MATCHING

Vector Quantization (VQ)

VQ is a process of mapping vectors from a large vector space to a finite number of regions in that space. Each region is called a cluster and can be represented by its center called a codeword. The collection of all codeword's is called a codebook. The density matching property of vector quantization is powerful, especially for identifying the density of large and high-dimensioned data. Since data points are represented by the index of their closest centroid, commonly occurring data have low error, and rare data high error. Hence, Vector Quantization is also suitable for lossy data compression. It is a fixed-to-fixed length algorithm. VQ may be thought as an aproximator. Figure shows an example of a 2-dimensional VQ.

Figure-8: An Example of a 2-dimensional VQ



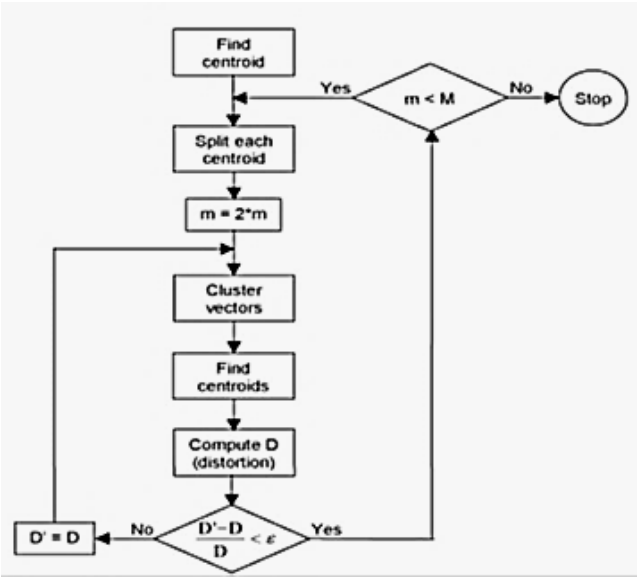
Sources: Authors Compilation

Here, every pair of numbers falling in a particular region is approximated by a star associated with that region. In Figure 2, the stars are called codevectors and the regions defined by the borders Frame Blocking Windowing FFT Cepstrum Mel-frequency Wrapping 567 are called encoding regions. The set of all codevectors is called the codeboo*k* and the set of all encoding regions is called the partition of the space.

LBG Design Algorithm

The LBG VQ design algorithm is an iterative Algorithm (as proposed by Y. Linde, A. Buzo & R. Gray) which alternatively solves optimality criteria. The algorithm requires an initial codebook. The Initial codebook is obtained by the splitting method. In this method, an initial codevector is set as the Average of the entire training sequence. This codevector is then split into two. The iterative algorithm is run with these two vectors as the initial Codebook. The final two codevectors are split into four and the process is repeated until the desired number of codevectors is obtained. The algorithm is summarized in flowchart of Figure-9.

Figure-9: Flow diagram of the LBG algorithm



Sources: Authors Compilation

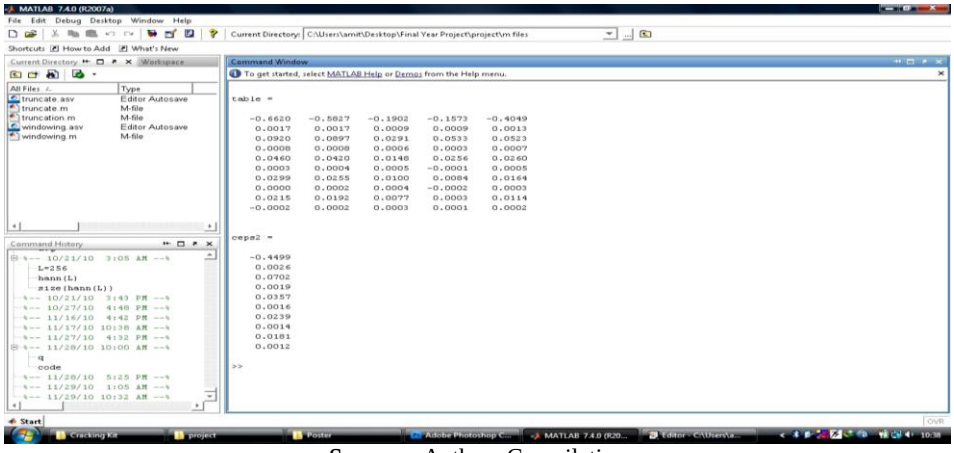
CONCLUSION

The main idea of this paper was to discuss a speaker recognition system that could be applied to a speech of an unknown speaker. By determining the extracted features of the unknown speech and then comparing them to the stored extracted features for each different speaker in order to identify the unknown speaker. The feature extraction was done by using MFCC (Mel Frequency Cepstral Coefficients).

The figure below shows the result of Cepstral Coefficient Calculations for five users and first ten DCT coefficients are Cepstral coefficients. Each user having five vocalization of the word "hello". Then it was averaged and represented in tabular form named

"table". Each column corresponds to a given speaker. The next column denoted as "ceps2" is Cepstral coefficient of 2nd speaker. We can clearly see its resemblance to 2nd column of "table". The result obtained is shown in the next page.

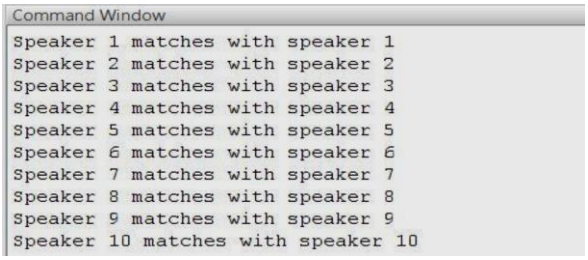
Figure-10: Result of Cepstral Coefficient Calculation



Sources: Authors Compilation

The speaker was building up using Vector Quantization (VQ). By clustering the training feature of each speaker we produce the VQ codebook and then stored in the speaker database. In this method, the K-means algorithm was used for clustering purpose. In the recognition stage, a distortion measure which based on the minimizing the Euclidean distance was used when matching an unknown speaker with the speaker database. VQ based clustering approach is best as it provides us with faster speaker identification process.

Figure-11: Result of Speaker Recognition



Sources: Authors Compilation

ACKNOWLEDGEMENT

To complete this work, we have got valuable suggestions and guidance from Dr. Rajesh Kumar (Department of electrical engineering, MNIT, Jaipur). I am thankful to all researcher and authors of manuscript from where I got valuable information for completing the work.

REFERENCES

1. Dash, Kshamamayee, Padhi, Debananda, Panda, Bhoomika, & Mohanty, Sanghamitra. (2012, April). Speaker Identification using Mel Frequency Cepstral Coefficient and BPNN. ***International Journal of Advanced Research in Computer Science and Software Engineering***, 2(4).

2. Bhalla, Anand Vardhan, Khaparkar, Shailesh, Bhalla, & Mudit Ratana. (2012, March) Performance Improvement of Speaker Recognition System. ***International Journal of Advanced Research in Computer Science and Software Engineering,*** 2(3).

3. Revathi, A., Ganapathy, R., & Venkataramani, Y. (2009, November). Text Independent Speaker Recognition and Speaker Independent Speech Recognition Using Iterative Clustering Approach. ***International Journal of Computer science & Information Technology (IJCSIT),*** 1(2).

4. Patel, Ibrahim, Rao, Y. Srinivas. (2010, December). Speech Recognition Using Hmm With Mfcc- An Analysis Using Frequency Specral Decomposion Technique. ***Signal & Image Processing: An International Journal (SIPIJ)***, 1(2).

5. Hasan, Md. Rashidul, Jamil, Mustafa, Rabbani, Md. Golam, Rahman, Md. Saifur. (2004, December 28-30). Speaker Identification Using Mel Frequency Cepstral Coefficients. ***In Proceedings of the 3rd International Conference on Electrical & Computer Engineering ICECE 2004.*** Bangladesh: Dhaka.

6. Gill, Manjot Kaur, Kaur, Reetkamal, & Kaur, Jagdev J. (2010, July). Vector Quantization based Speaker Identification. ***International Journal of Computer Applications,*** 4(2). (0975 – 8887).

7. Stolcke, A., Shriberg, E., Ferrer, L., Kajarekar, S., Sonmez, K., & Tur, G. ***Speech Recognition As Feature Extraction For Speaker Recognition.*** SRI International: Speech Technology and Research Laboratory. USA. CA: Menlo Park.

8. ***An Automatic Speaker Recognition System*** (Project Report). Retrieved from http://www.ifp.uiuc.edu/~minhdo/teaching/speaker_recognition**.**

9. ***Speaker Recognition*** (Project Report). Retrieved from http://speakerrecognition.googlecode.com/files/Finally_version1.pdf**.**

10. Retrieved from http://www.ifp.illinois.edu/~minhdo/teaching/speaker_recognition/speaker_recognition.html

11. Retrieved from http://www.researchtrend.net/ijet/4_Vibha.pdf

12. Retrieved from http://powerpointpresentationon.blogspot.in/2011_11_23_archive.html

13. Retrieved from http://www.slideshare.net/HiraShaukat/speaker-recognition-using-mffc

14. Retrieved from http://www.ll.mit.edu/mission/communications/publications/publication-files/full_papers/020513_Reyno.. .

15. Retrieved from http://www.slideshare.net/PhanDuy/speaker-identification-using-mel-frequency

16. Retrieved from http://www.researchtrend.net/ijet/4_Vibha.pdf

17. Retrieved from http://www.intechopen.com/books/advances-in-data-mining-knowledge-discovery-and-applications/the-per.. .

18. Retrieved from http://www.ijsr.net/archive/v2i3/IJSRON2013583.pdf

19. Retrieved from http://www.linkedin.com/pub/mukthar-yousuf/47/1b7/887

20. Retrieved from https://ccrma.stanford.edu/~unjung/mylec/mfcc.html

21. Retrieved from http://citeseerx.ist.psu.edu/showciting?doi=10.1.1.16.5590

22. Retrieved from https://www.erpublication.org/admin/vol_issue2/upload%20Image/IJMCTR011103.pdf

23. Retrieved from https://www.clear.rice.edu/elec532/PROJECTS98/speech/cepstrum/cepstrum.html

24. Retrieved from http://en.wikipedia.org/wiki/Cepstrum?oldid=0

25. Retrieved from http://www.springerreference.com/docs/html/chapterdbid/72072.html

26. Retrieved from http://www.ijsce.org/attachments/File/v3i4/D1797093413.pdf

27. Retrieved from http://www.maxwellsci.com/print/rjaset/v4-33-40.pdf

28. Retrieved from http://www.matlabcodes.com/2010/08/speaker-recognition-matlab.html

29. Retrieved from http://www.ijcaonline.org/archives/volume14/number3/1821-2393

****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

A STUDY ON LEARNERS PERSPECTIVE ON LEARNING MANAGEMENT SYSTEM IN INFORMATION TECHNOLOGY INDUSTRY

Dr. Suruchi Pandey[8] Megha Pathak[9]

ABSTRACT

Learning Management System is an online repository of information in the form of e-books, videos, online tutorials, online libraries, etc. Organizations are increasingly focusing on trying to improve their LMS in order to provide the latest and best in class knowledge to their employees. LMS enables employees to learn at their own pace, place and time. Courses can be uploaded in the LMS or downloaded from there as required. There are also options to assign courses to employees. Users can login to the LMS and keep a complete track of all the courses they have completed so far. Online reports can be generated for each registered user. Options for assessments and surveys can also be provided if the user is interested in taking one. It has multiple advantages to employees and organizations.

Companies spend a huge part of their budgets in implementing and deploying LMS. It is important to analyses if LMS is being used effectively or not. An even broader question can be whether LMS is beneficial for the organization as a whole or not. If not, then what are the loopholes and how can they be rectified. Essentially, two main questions are who all are the employees who are being benefited by LMS and what percentage of employees are interested in using LMS. The aim of the study is to collect data from employees, who are the ultimate users and beneficiaries of LMS and record their opinions regarding LMS and its effectiveness. The researchers also tried to highlight the best practices existing in the industry.

KEYWORDS

Learning Management System, Learning, Development, E- Learning etc.

INTRODUCTION

A study on the Learners Perspective on Learning Management System in the IT Sector operational in India was chosen keeping in mind the growing popularity of Learning Management System in organizations today. Employees in most organizations do not have sufficient time after their daily jobs to use LMS for learning. Even if they do have time, they do not want to invest it in learning. So, what is it that still makes it so important for organizations to have an LMS in place? Is it branding, competition, or just the fact that they have not done enough research to find out whether employees are interested in using LMS or not.

Online LMS is a relatively new concept in the history of organizations, however, learning is not. Earlier, learning used to be by organizing classrooms and seminars for employees. Now, employees have to themselves show interest in learning and enroll for courses. Therefore, one dimension that was highlighted through this research was that employees have to be self-motivated to show a keen interest in learning.

Constant Learning has become an important part of any business today. Learning supports innovation and Innovation supports growth. Hence, it is imperative to make the employees aware of the need for learning, thus making them realize the importance of the LMS in the organization. The report gives an insight into the nuances for the success of LMS in an organization and looks at certain organizations specifically in order to understand the LMS and its effectiveness better.

LMS is a software application for the administration, documentation, tracking and reporting of training programs, classroom and online events, e-learning programs, and learning content. The motto of LMS is non-stop learning at your time, place and pace. E-learning can be briefly defined as electronically supported learning and teaching. LMS also uses computer and network-enabled transfer of skills and knowledge. Through LMS, the e-learning offerings can include online courses, recorded sessions, classroom sessions, online videos, virtual classrooms and webinars.

OFFERINGS OF LMS FOR THE LEARNER

Online Courses

Online courses provide a very interesting approach to learning, their greatest advantage being their 24x7 availability, anytime and anywhere. Unlike Classroom sessions, online courses do not require an approval and the learners do not have to wait for anybody. The learner can keep going through the course at his own pace and place. A big advantage of online courses is that if users leave an online course in between, then next time when they login, they can choose to start either from where they left or from the beginning. The learner only has to search for the course he wants to go through and enroll / register himself.

[8]Associate Professor, Symbiosis Institute of Management Studies (SIMS), Maharashtra, India, suruchi.p@sims.edu
[9]Talent Scout, Adobe Systems, Noida, Uttar Pradesh, India, meghapathak10288@yahoo.com

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Classroom Courses
Classroom Training Sessions are developed keeping in mind the Training and Development needs of the learners. These are generally decided according to the Training Calendar of the respective organizations. Individual learning needs are identified and accordingly a training program is planned. Not all classroom sessions can be made available to everybody due to the limitation of space and resources. Hence, the Leaning & Development team according to the job requirements of each learner generally prepares a carefully designed training program.

Curriculum
Curriculum can be defined as *Blended Learning.* The learner is provided with different varieties of material for thorough understanding of the topic. It is an amalgamation of various forms of training like online learning, classroom training and other forms like discussions, interviews and plant visits as the need may be.

Library
It has a large collection of e-books i.e. online version of books. The users can read an entire book online and read it again, whenever they want to.

Posting
It consists of material that is uploaded by other users of the LMS.

Online Videos
Videos are available online, which the learners can view whenever they like. These can be lectures from great managers or renowned people or recorded classroom sessions.

USERS

LMS for Managers
Managers can use LMS as a reporting, tracking, allocating and approving tool i.e. in order to keep a track of which employees have completed certain courses, they can also allocate courses to their subordinates if they feel it is necessary and they can approve/decline requests for courses for which their subordinates have applied.

LMS for Administrators
Administrators can use LMS for Creating and Managing Events and sessions, managing enrolments, managing vendors and trainers, creating and managing Groups, sending mailers, etc.

LMS in the IT SECTOR

The easiest and earliest implementation of LMS was in the IT sector. This can be partly credited to the fact that in the IT sector, LMS is extremely important because of the dynamic nature of this field. There are newer and better technologies coming in every day. There are updations in the old technology. Hence, it is very important to stay up to date with the latest technology. This calls for constant learning and innovation.

Secondly, IT is the first sector where LMS can be made easily available. This can be attributed to the fact that online LMS requires an internet connection and a hardware device, which are both integral parts of an IT company. Intranet and laptops/desktops are available in IT companies.

The core employees in IT are the ones who make/ work on software and they have complete access to LMS. Hence, the very purpose of LMS doesn't fail when we talk about LMS in IT unlike in any other sector like manufacturing or so, where the heart and soul of the company are the labourers in the factories, who do not have any access to LMS and the employees who actually have access to LMS are the management alone.

All these factors together make it interesting to track the usage of LMS in the IT sector.

RESEARCH METHODOLOGY

The Methodology adopted for the present was through:

- Survey from around 50 employees working in IT and user of LMS. In addition, Discussions with various organizations L&D heads were done.
- Questionnaire comprised of 17 questions to seek feedback on LMS.
- Secondary data about the organization is LMS.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

FINDINGS OF STUDY

The responses clearly depict the views of employees about LMS and what are the areas of improvement according to them. Following are the major findings of the study.

The presuppositions of the study 'LMS is being used effectively in various companies in the IT sector' has been discussed in depth in this report and it can be concluded that LMS is still in its nascent stages and has a long way to go before it starts being used effectively.

Currently, LMS is gaining momentum and becoming a popular concept with each passing day. However, employees are not as comfortable logging in to LMS, browsing for courses and completing them as much they are in asking their fellow colleagues for help. LMS is a brilliant initiative by any organization. The bright future of LMS calls for a lot of innovation and constant improvement. The conclusions reached through this report are:

1. **IT Sector**
   It is the most promising sector for the growth of LMS, considering the availability of the required resources and the dynamic nature of the field. 98% employees said that their organization had an LMS in place.

2. **Look and Feel**
   While deciding to use LMS or not, employees take into account the look and feel of the portal also apart from the content in the portal. The ease with which employees can browse through the portal matters to them. Even in the survey, employees have been asked that 'if their LMS was made more user friendly, would they want to use it more often' and 64% employees replied Yes, clearly indicating that the unfriendliness of the page was one of the reasons for the low learner turnover of LMS.

Table-1: Views of Employees about Friendliness of LMS

Views of Employees about the Friendliness of their LMS	Percentage of Employees
Dissatisfied	**2%**
Scope for Improvement	**68%**
Satisfied	**28%**

Sources: Authors Compilation

3. **Ease of Applying for / Completing a Course**
   Ease of applying for / completing a course is important for the employees. If they find the entire task of finding or completing a course too laborious, they might as well lose interest.

Table-2: How Frequently Employees Enroll for Courses in LMS Annually

How frequently employees enrol for courses in LMS annually	Percentage of employees (Job tenure< 2 years)	Percentage of employees (Job tenure> 2 years)
Less than 5 times	**8%**	**4%**
5-10 times	**40%**	**24%**
More than 10 times	**4%**	**20%**

Sources: Authors Compilation

4. **Features**
   The availability of certain features like bookmarking courses, giving / reading feedback, versioning of courses matters to the success of the LMS.

Table-3: How Frequently Employees Login to LMS

How Frequently Employees Login to LMS	Percentage of Employees
Never	**14%**
Once a week	**6%**
Once in two weeks	**16%**
Once in a month	**62%**

Sources: Authors Compilation

5. **Relevance to Job**
   If the employees do not find courses relevant to their job on LMS, the might lose interest and never log in to LMS again. Hence, it is important that courses relevant to the employee's job profile should be clearly visible to him, over and

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

above other courses. Unless learner finds that the course has added certain value to his knowledge, then only he will be motivated to enroll for a course again. Only 10% employees said that they did not find courses in LMS relevant to their job.

Table-4: Employees Feeling that the Course Added to their Knowledge and Helped in Performing Job Better

Employees feeling that the course added to their knowledge and helped in performing their job better	Percentage of Employees
Not at all	**4%**
To some extent	**74%**
To a large extent	**22%**

Sources: Authors Compilation

6. **Frequent Mailers**
Constant touch with all the learners is essential. Out of sight, out of mind can greatly affect the popularity of LMS. Hence, employees should be constantly reminded of the advantages of LMS and the latest additions/ upgradations in the portal. 84% employees said that they received frequent mails from the Learning and Development team. 64% replied that they received monthly notifications about the completion level of their courses in LMS, thus serving as a reminder also in case some course was left.

7. **Mobile Interfacing**
Many employees are not even familiar with the concept of mobile interfacing. If mobile LMS is introduced; firstly, employees will be curious to know how it works and then it can motivate employees to use LMS while they are on the go. 84% employees said that their LMS did not have a mobile interface.

8. **Certifications**
Since it has been clearly seen from the data that certifications are a big attraction for learners. So, more and more certifications should be introduced in order to motivate people to log in, use LMS, and enhance their knowledge. Companies can tie up with accredited organizations and certifying authorities in order to make their LMS a win-win deal. From the research, it is clear that employees who do not have certifications in their LMS said that if certifications are introduced, and then they would definitely enroll for them. 64% employees said that they had certifications in their LMS. The rest 36% said that they would enroll for certifications if provided with any.

9. **Evaluations**
Employees are interested in knowing the results the courses that they take and hence, in some cases are interested in writing tests. For the organization as well, evaluations will help to know if the learner has been able to effectively grasp the course. Two types of evaluations can be conducted- pre course test and post course test. Pre course test tells us about the performance of the learner before taking the course and post course will help to compare the performance of the learner after the course. Hence, it can be judged how much the employee has been able to gain from the course. Kirkpatrick's model can also be applied here. The Reaction, Learning, Behavior and Results are measured for calculating overall performance. Grades, marks or ranks can be allocated for evaluation. In addition, the option for evaluation can be left to the discretion of the learner, whichever way he feels comfortable. Satisfaction of employees with evaluations was also surveyed and it was found major chunk of the employees were satisfied with it. For the ones that were not, evaluations can be made more transparent and the method for evaluating can be made clearer to the learners. 74% employees said that their LMS had the system of evaluations. Out of these, 5 employees were not satisfied with the evaluation system

10. **Support System**
LMS support system includes a telephonic helpline, email support, user manuals and a LMS support helpdesk. For telephonic help and email support, learners are given a telephone number or email id to contact in case of any queries. User manuals are available on the portal. These can be divided into different manuals for different kinds of users like learners, managers, administrators, etc. LMS support helpdesk is a portal wherein the employee can log his query and the corresponding support assistant will reply to the issue within the slated time (according to the priority time of the query as set by the learner/complainant). 74% employees said that their LMS had a support system.

Employee Recommendations

- Evaluation to include measures to calculate how effective the course was for the candidate in the end.
- It can be improved by developing some learning skills & styles.
- By giving some experience somehow, reviewing it, concluding it & then planning for the next step.
- It can also be improved by providing competency based learning.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

- After evaluation, instead of just displaying the score alone, the user can be fed with the correct answers for the ones he had incorrectly marked.
- It can be made more popular among the employees. Currently, it only caters to courses for achieving organizational targets. There can be courses suited to the user's interest only as well.
- By supporting portability and standards: Courses in LMS shall motivate streamlining throughout organization. Then, there shall be benchmarking of information.
- By delivering webinars and Online Seminars.
- Better visibility: This deals with the look and feel of the portal, or in simpler terms, the appeal of the page to attract more number of learners.
- Constant persuasion: Employees should be constantly reminded of the advantages of LMS, by always staying in touch with employees and informing them about the various courses relevant to their job and interest available in LMS.
- Rewards: This technique can be followed to lure employees into using LMS, especially first time users. In order to attract learners, they can be lured into logging in to LMS by offering certain rewards and organizing games, etc.
- By bringing new and better courses- Many learners felt that the courses they were currently being offered were irrelevant or insufficient. Hence, newer and better courses according to the user's needs should be introduced. Users stressed that Videos should be made popular in order to give practical explanation of a running process. Employees believed that LMS exists more or less for organization's mandatory course requirements and there was no charm for employee's personal interests.
- Better collection of e-books- Learners demanded a better collection of e-books. The advantages of e-books outweigh many other forms of learning tools because e-books can be read anywhere, any time and at the learner's pace.

Table-5: Courses most preferred by Employees

Courses Most Preferred By Employees	Percentage of Employees
Online Courses	**32%**
Classroom Training	**48%**
Seminars	**8%**
E-Books	**10%**
Online Videos	**2%**

Sources: Authors Compilation

ANALYSIS OF DATA

Correlation between Job Tenure and LMS Enroll Frequency

Variable 1- job_tenure: How long the employee has been working in the organization.
Variable 2- LMS_enroll_frequency: How frequently does the employee enroll for a course on LMS.

Table-1: Correlation between Job Tenure and LMS Enroll Frequency

		job_tenure	**LMS_enroll_frequency**
job_tenure	Pearson Correlation	1	.343*
	Sig. (2-tailed)		.015
	N	50	50
LMS_enroll_frequency	Pearson Correlation	.343*	1
	Sig. (2-tailed)	.015	
	N	50	50
Note: *Correlation is significant at the 0.05 level (2-tailed).			

Sources: Authors Compilation

Hence, it becomes clear that the number of years an employee has been in a job is directly proportional to the number of times he enrolls for a course on LMS. The correlation here is 0.343, which is not very strong though, but is significant.

Correlation between Personal Takeaway from LMS and Improvement in Job Performance

Variable 1: personal_takeaway- How much the employee believes the courses in LMS have added to his knowledge and skills personally

Variable 2: job_performance_improvement- How much employee believes the courses in LMS have helped him in performing better in his job

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Table 2: Correlation between personal takeaway and improvement in job performance

		personal_takeaway	job_performance_improvement
personal_takeaway	Pearson Correlation	1	.805**
	Sig. (2-tailed)		.000
	N	50	50
job_performance_improvement	Pearson Correlation	.805**	1
	Sig. (2-tailed)	.000	
	N	50	50

Note: **Correlation is significant at the 0.01 level (2-tailed).
Sources: Authors Compilation

The correlation here is very high i.e. 0.805, thus indicating that when an employee is able to add to his knowledge and skills with the help of LMS, then it also in turn impacts his performance in his job positively. Thus, the two variables are directly proportional and the correlation is significant.

Areas of Improvement

Following are certain suggestions for better implementation of LMS and increasing its popularity. Few of these points also are demonstrated best practices by some of the companies under study:

1) **Versioning of Courses**
The courses in LMS need to be versioned. When a user searches for a course, he should be presented with the list of the latest courses on that topic. The old courses should be automatically pushed back. The first page shall list out the latest courses followed by the next page with the older courses. Also, on the first page, should be the list of courses with maximum hits.

2) **Blogs by Users**
Users shall be given a page to write blogs. Here, they can write down about courses and their experience with LMS and read the experiences of other users or their views and opinions about certain courses.

3) **Review Pages by Users**
Users shall be allowed to write reviews and there should be a page completely dedicated to reviews by users.

4) **Chat Forums**
There can be chat forums where users can convene online and discuss their needs, problems, experiences, issues and feedback about courses and LMS.

5) **Certifications**
There can be certifications that the user can be given an opportunity to complete on LMS. Certain certifications can be made mandatory for employees of a particular department according to the need of that department. The company can tie up with accredited organizations and certifying authorities to introduce certifications in LMS.

6) **Online Tutoring**
Option for online tutoring can be made available wherein experts can guide learners online and take sessions online.

7) **Monthly Reports**
Monthly reports should be sent to learners about the courses that they completed in that month. These way learners shall be able to keep a track of the courses they have completed and the ones they are yet to complete.

8) **Mobile Interfacing**
LMS can be made mobile. LMS can be interfaced with mobile networking. Thus, users shall be able to access LMS anytime and anywhere. Users can get notifications about new courses in LMS on their mobile.

9) **Feedback Forms**
With the advancement of LMS, feedback forms shall also be implemented to get a detailed view of how much the users have been able to understand from the course from application point of view. Feedback can be taken from the learners about the course, its relevance, the facilitator, the resources, etc.

10) **Support Team**
A support team is very essential for the success of the system. This team shall be responsible for providing support to the users at all stages. Support shall comprise of a support helpline, email support and a support helpdesk.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

11) **Increase Awareness:** Awareness about LMS can be increased by:

a) **Mailers:** Group mails can be sent out for rolling out communication about courses in order to make employees acquainted with the latest offerings of LMS and the upcoming courses.

b) **Events:** Events like get-togethers, team parties; etc., can be conducted to highlight the launch of a new program.

c) **Campaigns:** Campaigns can be conducted by using banners, posters and standees.

d) **Games:** Making people learn through the entertainment way interests them. So, use games like online crossword puzzles, etc. to motivate people to login to LMS to find the answers and win the game.

e) **Quiz:** This is another entertaining way of spreading awareness. Send quiz in mails or update on portal and ask employees to login to LMS to find the answers and win prizes.

f) **Offers:** Offers can be used e.g., first three employees to give the correct answers will win some goodie.

12) Arouse Interest

a) **Connecting with Users:** For any purpose, be it spreading awareness or knowing what the users want to learn, it is important to connect with the users. The users must feel that they are being taken care of and their needs and requirements are being looked into.

b) **Know what the Users Want to Learn:** It is important to design a course according to the user's learning needs and requirements. Therefore, first step is to connect to the users and know what they want to learn. Then only the users will feel interested in learning the course.

c) **Create a Bucket of Courses Catering to the Needs of Specific User Groups:** Nobody would be interested in going through the long list of about 20,000 courses. Hence, it is important to segregate courses into groups according to the needs to users. A group of users can be made acquainted with the courses that will be helpful for them, thus arousing their interest in LMS.

d) An online page can be created where users can rate courses, thus helping other users to know the effectiveness of courses beforehand.

13) How to Connect to Users

a) **Surveys:** In order to know the needs of users, surveys can be carried out. Users can be asked to fill out surveys asking questions about their expectations from a course, their learning needs or may be their experience of a course or their feedback for a course.

b) **Questionnaires:** Similarly, users can be asked to fill out questionnaires.

c) **Meetings / Events** e.g. Flash Mob-Meetings and events can be conducted in which the LMS team comes together with the users of LMS. The requirements and learning needs of users can be discussed in these meetings.
The flash mob that was conducted across different locations in Mumbai was a huge success and very well received. The number of hits for the LMS page increased drastically after the flash mob.

14) Larger Team

a) **Innovation and Development:** As the number of users of LMS increases, a larger team will be required to maintain it. Constant innovation will be required to add on new features to LMS. Web developers will be required for adding new functionalities.

b) **Support**: A support team needs to be in place for LMS in order to provide support to the users whenever they need and at every stage of learning.

15) Infrastructure

a) **Laptops:** LMS is an online tool. Hence, the availability of state of the art IT infrastructure is essential. Desktops/Laptops are necessary for using LMS.

b) **Internet:** Since LMS is an online portal, so internet is essential. It is important that a good speed and bandwidth is provided for best usage of LMS.

CONCLUSIONS

Based on the interaction with the employees and findings from the data, it can be concluded that LMS is being practiced effectively in IT industry. It is serving to be a great utility and learning tool for employees as well as managers. Organizations under study have leveraged technology to facilitate learning and development. However, above recommendations would help improve its effectiveness. Future is of E-learning and organizations are leveraging it to the best possible way.

REFERENCES

1. Bob, Brotherton, & Roy, C. Wood. (2008). ***Handbook of Hospitality management***, 273-290. Sage Publication.
2. Bhattacharjee, C. ***Service Sector Management an Indian Perspective***, 423-436. Jaico Publishing House.
3. Chakravarti, B. K. ***Hotel Management Theory,*** 2, 79-100. New Delhi: APH Publishing Corporation.
4. Claus, Pahl. (2008). ***Architecture Solutions for E-Learning Systems,*** 154-174, 320-339. IGI Global Snippet.
5. David, Mallon. (2010). High-Impact Learning Practices, 99-143. Bersin & Associates.
6. David, L. Vance. (2010). ***The Business of Learning***, 425-447.
7. Dugan, Laird, Elwood, F. Holton, & Sharon, S. Naquin**.** (2003). ***Approaches to Training and Development*** (3rd ed.). Basic Books, Perseus Books Group. ISBN 0-7382-0698-9.
8. Eli, Hustad, & Aurilla, Bechina Arntzen. (2013). Facilitating Teaching and Learning Capabilities in Social Learning Management Systems: Challenges, Issues, and Implications for Design. ***Journal of Integrated Design and Process Science***, 17 (1), 17-35.
9. Eurich, Nell P. ***The learning Industry education for Adult workers,*** pp. 30-40. Princeton: Carnegie Foundation for the Advancement of Teaching.
10. Gary, Kroehnert. (1995). ***Basic Training for Trainers, A handbook for new trainers*** (2nd ed.), pp. (206, 225). Sydney: McGraw-Hill Book Company.
11. (2013, June/July). ***Harnessing Technology for Better Human Capital Management***. Network World Asia
12. Juli, Purwanti, Fadillah, Rizky, & Wahyudi, Handriyanto. (2013). Managing Learning Style across Generation in Workplace. ***In Proceedings of ICICKM 2013, The 10th International Conference on Intellectual Capital, Knowledge Management and Organisational Learning***, (***pp. 655-663).***The George Washington University.
13. Kamla, Ali Al-Busaidi. (2012, January). ***Learners' Perspective on Critical Factors to LMS Success in blended Learning: An Empirical Investigation***, 30, 11-34.
14. Liping, Shi, Liming, Sheng, & Qiang, Liu. (2013). Study on Relationships among High performance Work practices Systems, Quality Management Practices and organizational Learning. ***Journal of Applied Sciences***, 3310-3314.
15. Pandey, S., & Bapat, S. (2011, February 24). A Study of Learning Systems at Pharmaceutical Industry Unit in India. ***In Proceedings of the International Conference on Human Resource Management – GCeMP 2K11***. Gujarat: Ganpat University.
16. Pandey, Suruchi. (2011). ***Training Practices in Hospitality Industry – Indian Perspective,*** 257-275. Lab Publishing.
17. Thongchai, Kaewkiriya, Ryosuke, Saga, & Hiroshi, Tsuji. (2013, July). Framework for Distributed e-Learning Management System. ***Journal of Computers,*** 8(7), 1637-1647.
18. Retrieved on on 2013, February from http://www.elearnspace.org/Articles/ldc.html
19. Retrieved on 2013, March from http://www.valuelabs.com/careers/learning-and-development

20. Retrieved on 2013, February from http://searchcio.techtarget.com/definition/learning-management-system

21. Retrieved on 2013, December from http://en.wikipedia.org/wiki/Learning_management_system

22. Retrieved on 2013, February from http://weblearning.atrixware.com/

23. Retrieved on 2013, February from http://www.wipro.com/industries/media-segments/education.aspx

24. Retrieved on 2014, January from http://www.infosys.com/talentedge/offerings/learning-management/Pages/index.aspx

25. Retrieved on 2014, January from https://wikis.uit.tufts.edu/confluence/display/TuftsLMS/Research+on+the+Effectiveness+of+Learning+Management+Systems+(LMS)+Use+in+Higher+Education

26. Retrieved from http://www.elearninglearning.com/lms/standards/

27. Retrieved from http://www.cs.osakafu-u.ac.jp/mis/pub/paper11.html

*****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

LEVERAGING SOCIAL MEDIA TO ENABLE OPERATIONAL EFFICIENCY FOR CLOUD COMPUTING USED BY SOFTWARE PROVIDERS

Shiva Goyal[10] Havish Madhvapaty[11]

ABSTRACT

Owing to high – capacity networks, low cost computers and storage devices there has been a growth in cloud computing. Cloud computing is being looked at as being one of the most often quoted trends expected to define I.T. in the future. Cloud computing has brought with it massive advantages – cost efficiency, near – unlimited storage, recovery ease, automatic software integration etc. The three fundamental models on which cloud computing providers are offering their services are – Infrastructure as a Service (Iaas), Platform as a Serve ice (PaaS), and Software as a service (Saas). There is a need to be extremely vigilant as well. Increased use of cloud computing has brought to light the issue of privacy.

This paper makes an effort to understand what these issues are and how software providers are integrating cloud computing into their services. The social media angle looks at how the companies and the regulatory framework can both utilize social media including pure – networking, semi – networking and non – networking sites to both keep a check on cloud computing practices as well as improve efficiency.

The paper evaluates the performance and extent of cloud computing integration of several companies. The final suggestion is that a regulatory framework is of utmost importance, and immediate changes need to be made to the I.T. Act. It is also concluded that social media in terms of forums and discussion boards can be used to improve intra – company and company – customer interaction and social networking can be used to gauge customer feedback and market feedback.

KEYWORDS

Cloud Computing, Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a service (SaaS), Social Media, Information Technology Act etc.

INTRODUCTION

Cloud computing is a boon of today and has many innovations unrevealed for tomorrow's world. Technically speaking , cloud computing can be defined as "a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (such as networks, servers, storage, applications, and services) that can be quickly provisioned and released with minimal management effort or service provider interaction". In essence, the concept of "the cloud" is as a metaphor for the Internet as an operational environment where applications are utilized over the Internet rather than through more traditional means such as a desktop. The emergence of cloud computing is transforming the way organizations purchase and manage computing resources, providing a fundamentally different concept in which a cloud provider might be responsible for a range of information technology related activities, including hardware and software installation, upgrades, maintenance, backup, data storage, and security. The users are not bound by the limitations of a single computing device and are free to experience a multitude of devices, platforms, and mobility (both socially and physically).

The result is that organizations can lower their IT capital expenditures. The emergence of cloud computing also facilitates the progression of IT standardization and commoditization, which refers to the phenomenon that user companies can use IT resources (especially infrastructure resources, e.g. servers, storage, and networks) as standardized commodities without the need for being uniquely designed, installed and maintained. Thus cloud computing can be defined as an I.T. service model, which delivers a set of convenient, on-demand, and configurable computing services and resources to clients "over a network in a self-service fashion - independent of device and location service provider interaction". Not only PCs but also mobile devices, such as smart phones and tablets, can access these cloud applications and services. Since the emergence of the concept, I.T. providers have developed a wide range of cloud computing services.

Categories of Cloud Computing

These cloud services can be divided into three main categories/models:

a) Software as a Service (SaaS)

In the SaaS model, software applications (e.g. organizational email systems, office applications, sales/accounting systems, and even Enterprise Resource Planning (ERP) systems are run on a vendor-managed and controlled infrastructure are made available

[10]Assistant Professor, Amity Global Business School, Noida, Uttar Pradesh, India, sgoyal1@amity.edu
[11]Assistant Professor, Amity Global Business School, Noida, Uttar Pradesh, India, havish.madhvapaty@gmail.com

to clients through web browsers. Software-as-a-Service (SaaS) allows users to run a variety of software applications on the internet without having possession or managing applications (e.g. Salesforce.com, Gmail, Microsoft Online) and social media sites (e.g. Facebook, LinkedIn, Twitter).

b) Platform as a Service (PaaS)

In the PaaS model, computing platforms are provided as a service to deploy and run user applications. It offers a programmable environment and middleware to support IT application development and deployment in user companies. Platform-as-a-Service (PaaS) provides a computing platform to support building of web applications and services completely residing on the Internet (e.g. Google Apps, Force.com, and 3Tera AppLogic). For example, Facebook allows third parties to build and distribute applications within its service. The main factor distinguishing the three models is the level of control the subscriber retains over its data. Cloud computing provides significant benefits for the development and growth of businesses. Companies that embrace this technology and venture into the cloud must be careful and thoughtful. Companies should implement data classification schemes and scrutinize what they put into the cloud, along with selecting reliable and security conscience cloud providers. Well-drafted agreements and policies with both providers and employees can help reduce the risk of the disclosure of trade secrets in the cloud. A comprehensive cloud computing strategy can help companies realize the cost savings and financial opportunities in cloud computing, including social media; while ensuring that these benefits are not outweighed by the potential legal and business risks.

c) Infrastructure as a Service (IaaS)

In the IaaS model, hardware and IT infrastructure resources (e.g. CPUs, hard discs, databases, and servers) are provided as a service to companies through the virtualized cloud environment. Infrastructure-as-a-Service (IaaS) allows the use of computer hardware and system software, including operating systems and communication networks in which the cloud provider is responsible for hardware installation, system configuration, and maintenance (e.g., Amazon EC2, Citrix Cloud Center). The main factor distinguishing the three models is the level of control the subscriber retains over its data. Cloud computing provides significant benefits for the development and growth of businesses.

Social Media Usage by Companies

Computer Software companies such as Adobe Systems, Inc. Autodesk, Inc. concentrate on the non – networking / semi – networking side of social media. Since these applications are used by processionals who of artist, designers, architects etc. who are a very active community – discussion boards, forums etc. are very active on these websites.

These internet forums, or message boards, are an online discussion site where people can hold conversations in form of posted messages. These forums have a specific jargon associated with them where a single conversation is called a 'thread'. The users for these forums can be:

- Company employees,
- Non-Company employees (Anonymous),
- Non-Company employees (Registered).

Companies started these forums to foster an active community and these thread enabled them to post either a query in the form of a thread or a general FAQs (Frequently Asked Questions) list. It should also be noted that an equal, if not more, active community is also present in non – company owned (unofficial) forums.

Digital Millennium Copyright Act 2000

The Digital Millennium Copyright Act (DMCA) is a United States copyright law that implements two 1996 treaties of the World Intellectual Property Organization (WIPO). It criminalizes production and dissemination of technology, devices, or services intended to circumvent measures (commonly known as digital rights management or DRM) that control access to copyrighted works. It also criminalizes the act of circumventing an access control, whether or not there is actual infringement of copyright itself.

With regard to cloud computing – things are a little unclear. The question is whether data, which might be violating a copyright, can be saved on the cloud; and whether such a cloud server provider is eligible for DMCA safe harbors. Some cloud services can indeed fall under the ambit of DMCA safe harbors.

OBJECTIVES OF STUDY

- Evaluate the performance and extent of cloud computing integration of companies.
- Understand how an understanding of social media traits can help cloud-computing providers.
- Conduct a critique of cloud computing providers and consumer response.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

ANALYSIS

The companies chosen for the study are:

Table-1

Company Name	Industry
Adobe Systems, Inc.	Computer Software
Autodesk, Inc.	Computer Software
Facebook	Internet
Google	Internet / Computer Software / Telecoms Equipment
Microsoft	Computer Software / Computer Hardware

Sources: Authors Compilation

Adobe Systems, Inc.

Adobe Systems is an American multinational computer company. The company sells multimedia and creativity software products.

Criticism: In May 2013, Adobe Systems, Inc. announced plans to end perpetual software licensing for its Creative Suite products in favor of a subscription-based service model called Creative Cloud. By shifting to the software as a service (SaaS) model, they announced more frequent feature updates to its products and thereby eschewing their traditional release cycles. They also forced customers to pay a monthly subscription fee – which when stopped will mean that the customers lose access to all the work saved in a proprietary format. These files will also not be backward compatible with the Creative Suite, which Adobe has admitted is a valid concern. The subscription – only pricing in particular drew strong criticism.

Autodesk, Inc.

Autodesk, Inc. is an American multinational software corporation that focuses on acquiring 3D design software. It is not a developer in the strictest sense because its fundamental motive is not innovating software, but rather to generate revenue streams on the back of existing innovation.

Criticism: Since the company monopolizes computer graphics industry in nearly every sector, it receives a lot of flak for their practices. In addition to the fact that the company has a questionable history regarding customer support, the recently announced plans of the company to transition most of its software into cloud services is seen as a worrying development because people feel that too much control will be in the hands of the company.

Google

Google is an American multinational corporation specializing in Internet-related services and products, which include search, cloud-computing, software, and online advertising technologies. Google is pushing cloud computing with devices like the Chrome OS.

Criticism: Google has been criticized regularly in the past for its alleged manipulation and misuse of search results. The concern is that Google is spreading its branches into several Internet-based services and products and much of the criticism that Google gets pertains to issues that have not been clearly been addressed by Cyber Law.

Microsoft

Microsoft Corporation is an American multinational corporation that develops, manufactures, licenses, supports and sells computer software, consumer electronics and personal computers and services. Its best-known software products are the Microsoft Windows line of operating systems, Microsoft Office suite and Internet Explorer web browser.

Criticism: The Company has been the subject of several lawsuits for unlawful monopolistic practices regularly in the past. EULAs (End User License Agreements) for Microsoft programs are often criticized as being too restrictive.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Facebook

Facebook is an online social networking service. It is the most used social networking service by worldwide monthly active users.

Criticism: Facebook has been criticized often for the way it treats privacy of its users. In a Greenpeace report on cloud computing, it was found that the data centers that Facebook is operating are relying mostly on coal for electricity.

FUTURE OF CLOUD COMPUTING

Cloud computing has become a core part of how CIOs operate their IT environments today. With the perspective of computer software companies, cloud computing represents a conscious and deliberate move to transition traditional processes to the cloud.

The way the integration with social media will help is of immense importance.

An analysis of social media can bring with it possible advantages:

- Trends in customer demand and tastes can be easily identified by keeping a check on the sort of feedback and voice it gets on social media – especially social networking sites.
- Prior to launch of new releases - and to assist with alpha and beta testing; these software providers can use social media effectively to reach out to their target audience.
- Several cloud computing websites offering storage services like Amazon, Google, and Microsoft etc. have integration with social media both to ease the process of sharing information and ease of access of data for customers.

FINDINGS AND SUGGESTIONS

- Cloud Computing would help immensely to keep piracy in check.
- Moving to the cloud would help in dispersion of updates seamlessly.
- There are massive cost savings for the company because the intermediaries are removed and there is direct contact with the customers.
- Cloud computing enables the concept of BYOD (Bring Your Own Device) which allows employees to get their own devices to work. Since all the software and data is in the cloud there is no risk of data theft.
- Some services require a strong broadband connection – and some countries might be lagging behind thereby affecting performance.
- With cloud computing, systems become more prone to factors such as network outages, denial of services attacks etc.
- Large internet and technology companies like Google, Microsoft and Amazon are pushing forward their plans to deliver information and software over the internet, which are forcing people to buy into locked proprietary systems, which will escalate substantially in cost over time.
- Trends can be seen from how social media talks about the software provider in forums and other social networking sites.
- Data privacy would be an extremely important concept. The development in security especially when dealing with sensitive company information or personal information needs to keep pace with the intended growth, which the software companies seek.
- The I.T. Act in India in particular needs to be strengthened and requisite amendments need to be made.

REFERENCES

1. Choudhary, V., & Vithayathil, J. (2013). The Impact of Cloud Computing: Should the IT Department Be Organized as a Cost Center or a Profit Center? ***Journal of Management Information Systems.*** 30(2), pp. 67-100.

2. Jansen, W., & Granee, T. (2011). ***Guidelines on Security and Privacy in Public Cloud Computing***, pp. 800-144. NIST Special Publication.

3. Kalyvas, J. R., Overly, M. R., & Karlyn, M. A. (2013). Cloud Computing: A Practical Framework for Managing Cloud Computing Risk. ***Intellectual Property & Technology Law Journal,*** 25(3), 7-18.

4. Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J., & Ghalsasi, A. (2011). Cloud Computing - The Business Perspective. ***Decision Support Systems***, 51(1), 176-189.

5. Nanath, K., & Pillai, R. (2013). A Model for Cost-Benefit Analysis of Cloud Computing. ***Journal of International Technology & Information Management,*** 22(3), 93-117.

6. Garrison, G., Kim, S. W., & Robin, L. (2012). Success Factors for Deploying Cloud Computing. ***Communications of the ACM,*** 5(9), 62-68.

7. Stieninger, M., & Nedbal, D. (2014). Characteristics of Cloud Computing in the Business Context: A Systematic Literature Review. ***Global Journal of Flexible Systems Management***, 15(1), 59-68.

8. Budrienė, D., & Zalieckaitė, L. (2012). Cloud Computing Application in Small and Medium Sized Enterprises. ***Issues of Business & Law***, 4(1), 199-130.

9. Hayes, B. (2008). Cloud Computing. ***Communications of the ACM***, 51(7), 9-11.

10. Parekh, S. (2013). How Finance Teams Can Use Cloud Technology To Save Their Organizations Time And Money. ***Financial Management***, 57-59.

11. Retrieved from http://samuel-warde.com/2013/09/intellectual-property-rights-dummies/

12. Retrieved from http://www.nist.gov/itl/cloud/

13. Retrieved from http://www.netugc.com/digital-millennium-copyright-act-dmca

14. Retrieved from http://cloud-computing.findthebest.com

15. Retrieved from http://cloudcomputingtutorials.org

16. Retrieved from http://www.unicc.org/Pages/ICC-Services.aspx

17. Retrieved from https://answers.yahoo.com/question/index?qid=20140315023856AAKbSIr

18. Retrieved from http://en.wikipedia.org/wiki/Adobe_Systems

19. Retrieved from http://www.thejobportal.in/aboutgoogle.php

20. Retrieved from http://techservicehub.com

*****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

A NEW DECISION TREE APPROACH TO IMAGE DATA MINING AND SEGMENTATION

Dr. V. Mani Sarma[12] Mohammad Imran[13] M. Srivenkatram Reddy[14] Dr. P. Narahari Sastri[15]

ABSTRACT

In this paper, a general mining approach based on decision trees for segmenting image data is proposed. Pixel - wise image features are extracted and transformed into a database like table that allows existing data mining algorithms to dig out useful information. Each tuple in the table has a feature descriptor consisting of a set of feature values for a given pixel along with its label. With the feature label, we can employ the decision tree to (1) discover relationship between the attributes of pixels and their target labels, (2) build a model for image processing by using the training data set. Both experiments and theoretical analysis are performed in our research. The results show that the proposed model is very efficient and effective for image mining and image segmentation. It can also be used to develop new image processing algorithms, refine existing algorithms, or act as an effective filter.

KEYWORDS

Data Mining, Decision Tree, Association Rule, Image Indexing, Image Segmentation, Pixel, MIST etc.

I. INTRODUCTION

The term: knowledge discovery in image databases as *image mining*. The main goal of data mining is to discover previously unknown knowledge from a huge amount of historical data that can help us initiate proper actions. "Knowledge mining from data" is another name for the term "data mining", which is more appropriate but somewhat too long. Many people treat data mining as a synonym for another popular term, Knowledge Discovery in Databases (KDD). Although plenty of knowledge can be hidden in image data, very few literatures discuss KDD in this type of data. Issues of image mining have classified as four classes. They were *associations*, *classification*, *sequential patterns*, and *time series patterns.* However, only the prototype of *finding associations* has been proposed.

Image segmentation is an important procedure to crop useful information from images. Knowledge can be more easily recognized when presented in the form of images. For example, geophysical and environmental data from satellite photos, Web pages containing images, medical imaging including Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and Ultrasound Imaging (UI), are sources of useful information used in our daily life. They are conformed to various standard image protocols. Although many image segmentation algorithms have been proposed, only few of them can be applied to image mining.

Mining non-standardized data and multimedia data is the trend in the future. However, most existing data mining techniques have been designed for mining numerical data and are thus not well suited for image mining. In this paper, we solve this problem by presenting a new approach based on decision trees for both of image data mining and segmentation. Decision tree induction is a well-known methodology used widely on various kinds of domain, such as artificial intelligence, machine learning, data mining, and pattern recognition. A decision tree is a flow-chart-like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and leaf nodes represent classes or class distributions. An advantage of decision trees over other methodologies, such as neural network, is that it could provide understandable English-like rules or logic statements, for example, *"If a pixel's gray level ranges from 180 to 240 and its local variation is greater than 80 and its slope variation is greater than 0.5, then it is the pixel we wanted."* This basic idea of simple and easily understandable is also the main principle of our approach.

In an image mining method that works at a higher generality level for mining image associations is proposed. In contrast to that, our proposed model works on a relative low generality level for image pixel classification. Pixel-wise image classification is an essential part of many image segmentation methods, for example, determining pixels of an edge (corner) in edge (corner) detection methods pixels of a particular object in objects segmentation based methods], pixels of abnormal tissue of medical image processing , and pixel classes in thresholding, etc.

The proposed model can be used to mine hidden relationships between an image's pixel and its class label, and determine the interrelated features. Besides, the created model can be applied to perform pixel-wise segmentation on input images.

[12] Assistant Professor, Department of IT&SC, Addis Ababa Institute of Technology, AAU, Ethiopia, manisharma.vittapu@gmail.com

[13]Assistant Professor, Department of CSE, Muffakham Jah College of Engineering and Technology, Andhra Pradesh, India, imran.quba@gmail.com

[14]Associate Professor, Department of ECE, Hyderabad Institute of Technology and Management, Andhra Pradesh, India, venkatram99@yahoo.com

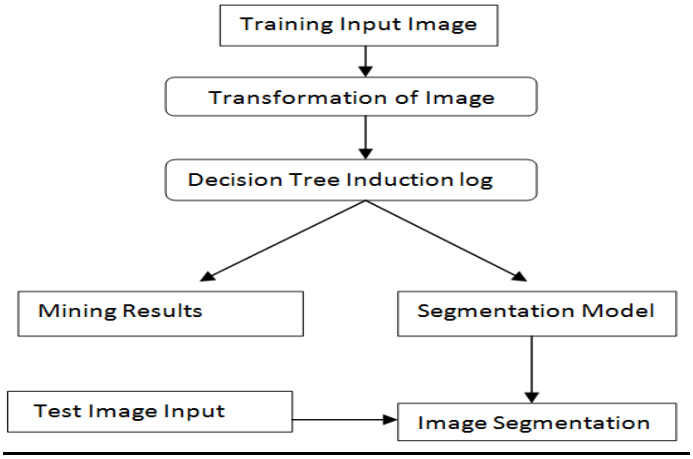
[15]Associate Professor, Department of ECE, CBIT, Andhra Pradesh, India, ananditahari@yahoo.com

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

II. OVERVIEW

The general processing flow of the proposed model is depicted in Figure-1. The data we used for input is formatted as a set of *raw* and *label* image pair. Each pixel's value of the *label* image is a class label with respect to the pixel in the raw image at the same position. The label of a pixel could indicate the type of a pixel, its frequency, etc.

Figure-1: The Proposed Image Segmentation Model



Sources: Authors Compilation

Once a set of interested *raw* and *label* image pair has been obtained, they are transformed and stored in a database-like table. Each row of the transformed table represents a given pixel, and each column of such table represents an encoded feature associated with that pixel. After obtaining such a database-like table from the images we are interested in, we can then begin to dig on it. In this paper, we have chosen the decision tree methodology for this purpose. Based on the decision tree technology, our proposed model is able to generalize rules between the label of pixels and their features.

Image Transformation and Feature Extraction

As mentioned, the input data of the proposed model is formatted as a set of equal sized *raw* and *label* image pairs. The transformation of the input image dataset into a database like table and subsuming of the related features is described in this subsection. For the sake of clarity, various terms used for this process are defined below. In addition, we propose three kinds of input data sources.

Definition 1 The **raw image** is a *d*-dimensional light-intensity function, denoted by R ($c1$, $c2$, cd), where the amplitude (or value) of R at spatial coordinates ($c1$, $c2$, cd) gives the intensity of the raw image at that point (or pixel).

Definition 2 The **label image** is a *d*-dimensional light-intensity function, denoted by L ($c1$, $c2$, cd), where the value of L at spatial coordinates ($c1$, $c2$, cd) gives the class identifier of the pixel at same spatial coordinates of its corresponding raw image.

Definition 3 The **database-like table** $X = \{x1, x2, \ldots, xt\}$ is a set of records, where each record $xr \in \Re k$ is a vector with elements $<a1, a2, \ldots, ak>$ being the value of attributes (or features) of X

Figure-2: An Example of Input Image Dataset

(a) Raw Image

7	9	9	9	7
5	7	9	7	5
0	7	9	7	0
5	7	9	7	5
7	9	9	9	7

(b) Label Image

1	0	0	0	1
1	1	0	1	1
0	1	0	1	0
1	1	0	1	1
1	0	0	0	1

Sources: Authors Compilation

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

In this work, only $d = 2$ is considered, *i.e*., images with dimensionality of 2. An example of the input image dataset is shown in Fig. 2. Each pixel value of the *raw* image represents the gray level of a pixel. Each pixel value of the *label* image represents the class label of the pixel. Both pixel values are in the same position. In this example, the *raw* image contains the capital English letter "I" with certain degree of blur. Thus, the inside pixels of the letter are darker and the outside pixels are brighter. If a pixel in the *label* image has the value "1", the pixel in the same position of the *raw* image is a pixel of *outside contour*. It is assumed to be a *pixel of interest* (POI).

In practice, the pixel value of the *label* image is not limited to the binary form but could take any kind of form. In addition, we can have as many *raw* and *label* image pairs at the same time as required for the input. In order to mine useful information from a set of *raw* and *label* images, we propose a methodology to transform them into a database-like table and allow any data mining algorithms to work on top of the table. This process is simple and straightforward as shown in Fig. 3.

Fig. 4 shows a part of the results of this transformation process according to the data in Fig. 2. Each row of such result table stands for a pixel. Hence, its *cardinality* (number of rows) equals the number of total pixels in the *raw* image. In addition, each column of such table represents a feature associated with the given pixels.

In Fig. 4, F*eature*1 represents the gray level and *feature*2 the local variation. In order to simplify this demonstration, the local variation in this case is replaced with the average difference of a pixel to its 4-neighbors. Other pixel-wised features [17, 18] such as entropy, contrast, mean, *etc*. can also be encoded into the table as long as they might have affection on the collected dataset.

Various encoding strategies such as normalization (*e.g*., adjusting the value ranging from 0 to 1) or generalization (*e.g*., transforming the value to high, medium, or low) can be applied when generating the desired features. Moreover, the *label* image was included as a column in that table. With the presence of the label feature, hidden relationships between these two kinds of images can be mined.

Figure-3: Pseudo Code of the Image Transformation Algorithm

```
procedure img2tab(image: raw, label);
begin
set feature_generated_functions[1..n];
set label_generated_function;
initiate table, pixel;
while pixel exists do
{pixel scanning process}
insert into table value :=
feature_generated1(raw, pixel),
…,
feature_generatedn(raw, pixel),
label_generated(label, pixel);
continue to scan on the next pixel;
end while
return table;
end
```

Sources: Authors Compilation

Figure-4: Result Table of Image Transformation According to the Input in Figure-2

	Feature 1	**Feature2**	**…**	**Feature n**	**Label**
Pixel 1	7	2	…	Value 1,n	1
Pixel2	9	1.25	…	…	0
Pixel 3	9	0	…	…	0
…	..	…	…	…	…
Pixel 4	7	2	…	Value 25,n	1

Sources: Authors Compilation

Data Reduction

Because of the image characteristics, pixels from a neighboring area will generate similar feature vectors in the transformation process. Under some circumstances, it will cause remarkable redundant information in the result table; for example, an image with a large portion of background. Here we present some basic types of redundancy and show how they can be eliminated while converting the input image set.

Definition 1 The **feature scope** of a pixel *M* with spatial coordinates (*c*1, *c*2) is an $n \times n$ pixel area with center at *M*, from which all the desired features of *M* can be generated. Usually *n* is an odd number, and the sub-image within the feature scope, *i.e*., pixels within spatial coordinates $(c1 \pm n - 1\ /2$, (*c2* ±n-1/2), is called the **root space** of the pixel *M*, denoted as {*RSM*}.

Definition 2 Two root spaces {*RSN*}, {*RSO*} are **rotation reachable** if {*RSN*} = {*RSO*}*R*, where {.}*R* stands for a root space after rotating the angle once by 90°, 180°, or 270°.

Definition 3 Two root spaces {*RSN*}, {*RSO*} are **mirror reachable** if {*RSN*} = {*RSO*}*F*, where {.}*F* stands for a root space after flipping horizontally or vertically.

Given two pixels *P* and *Q* at different *spatial coordinates* of an image *I*, they are said to be:

- **equivalent redundant**, if {*RSP*} is equal to {*RSQ*},
- **rotation redundant**, if {*RSP*} and {*RSQ*} are rotation reachable,
- **mirror redundant**, if {*RSP*} and {*RSQ*} are mirror reachable,
- **conflict redundant**, if {*RSP*} and {*RSQ*} satisfy any one of the first three conditions, but the label information of pixels *P* and *Q* is not equal to each other.

Figure-5: Pseudocode of Redundancy Reduction Algorithm

function *RR (image: raw, label; pixel: C);*
begin
apply quantization on {RSC} if necessary;
if *{RSC} can be matched in Ξ* ***do*** *{redundant pixel} discard {RSC} for further record generation;*
if *the label information of the two matched entries are not equal* ***do*** *{conflict redundant pixel}*
update the corresponding information in Ξ; retrieve or update previously generated record if necessary;
else
{non-redundant pixel} record all characterized redundancies of {RSC} and the corresponding label information in Ξ;
end

Sources: Authors Compilation

Users could characterize other types of redundancy according to the image problem they wish to solve. In order to pinch more redundancies, quantization techniques can be applied on the root space. The pseudo code regarding the function of redundancy reduction is shown in Fig. 5. This function can be added to the pixel scanning process of the image transformation algorithm in Fig. 3.

Fig. 6 shows the results of this reduction process according to the images in Fig. 2. The number of pixels for transformation after reduction has reduced from 25 to 9.

Mining Results and their Applications

After having obtained such a database-like table in accordance to the desired input image dataset, mining algorithms can then be used on it. In this study, we have chosen the decision tree for this purpose. An advantage of the decision tree over other methodologies, such as neural networks, is that it can provide understandable English-like rules or logic statements. For instance, *if the gray level of a given pixel ranges between 180 and 240 and its entropy is greater than 0.5*, *then it is a pixel of interest, POI.*

This basic idea of simplicity and easy understandability is also the main principle of our approach. The results of such a mining process may help us to better understand the image properties and relate to real world instances. The results can also be used to process new images of the same domain.

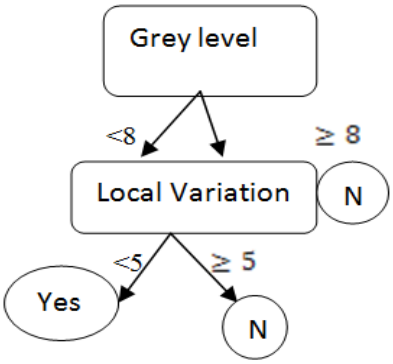
The result of the proposed model is a decision-tree classifier. Fig. 7 depicts a classifier derived from the data shown in Fig. 4 by using CART. A result classifier can be further straightforwardly translated into a set of human readable *if-then* rules. For instance, from the three leaf nodes in Fig. 7, we can obtain the following three rules: − ***If*** *the gray level of a given pixel is less than 8 and its local variation is less than 5,* ***then*** *it is a pixel of outside contour.*

− ***If*** *the gray level of a given pixel is less than 8 and its local variation is greater than or equal to 5,* ***then*** *it is not a pixel of outside contour.*

− ***If*** *the gray level of a given pixel is greater than or equal to 8,* ***then*** *it is not a pixel of outside contour.*

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Figure-6: A Decision Tree for the Concept *Is Outside Contour*

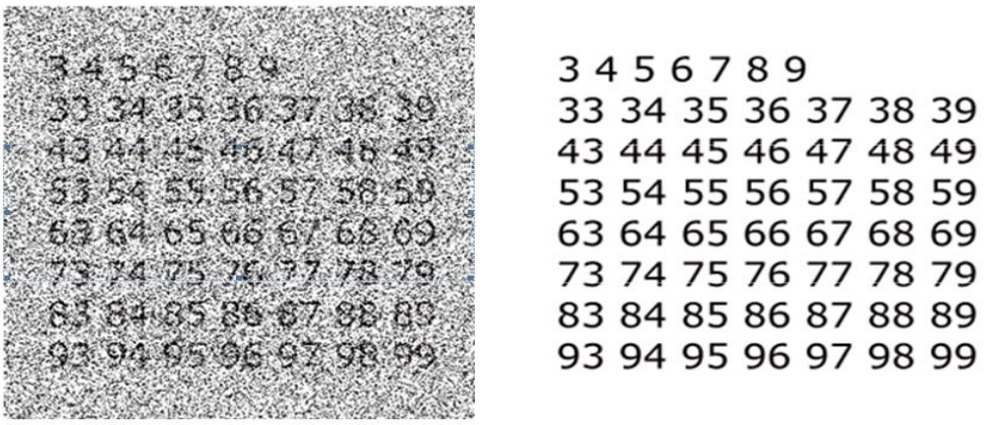


Sources: Authors Compilation

Derived from Fig. 6, indicating whether a pixel is a pixel of outside contour. These rules can provide useful information about the training image. Besides, in order to obtain a higher level of appearance and meet the different information granularity requirements, the rules can be post-processed by rule induction algorithms. More prominently, they can be used to process new images from the same domain. The practical image processing capabilities include image restoration, image enhancement, image segmentation, *etc*. Both experimental and theoretical analyses were performed in this study to examine the proposed model. The built classifier can also be used to select important features. Features used at higher tree levels for the splitting criteria show a higher significant influence on the pixel class. The selected features can reflect the characteristics of the *label* image and help design or refine other image processing algorithms.

Figure-7

Raw Image **Label Image**



Sources: Authors Compilation

Similarly, in the English alphabet training dataset, the distorted and the original images synthesized by the letters "*F*" to "*Z*" and their two combinations (*i.e.*, *FF*, *FG*, …, *YZ*, *ZZ*) were used for the *raw* and *label* images, respectively. The other letters (*i.e.*, *A*, *B*, …, *E*) were used to synthesize the testing image dataset.

For image transformation, a feature scope of size 5×5 was used and the selected features included gray level, local variation, mean, local minimum, local maximum, and entropy. The label of a given pixel in the experiments of image restoration with enhancement was set to its gray level in the *label* image. We did not apply any encoding strategies on the features to simplify the demonstrations. However, in practice, we can use any encoding strategy if required. In the image segmentation experiments, the label feature was transformed to 0 or 1 according to the threshold *label* image. In this way, the segmentation nature was imitated to distinguish between "background" or "object". After we have settled the transformation details, a database-like table can be derived. By applying a classification algorithm on the database-like table, a classifier for label prediction can be obtained. Under the same way, testing images can be transformed into a database-like table to predict the label attributes. These predicted labels can moreover be visualized in a natural form of the input data, *i.e.*, image. As we are proposing a general image mining and image processing framework and any existing decision tree algorithms can be used to do the job, we show only the testing result to simplify the demonstration. For the other results regarding the constructed classifier or the corresponding rules, if interested, examples can be found in our previous work

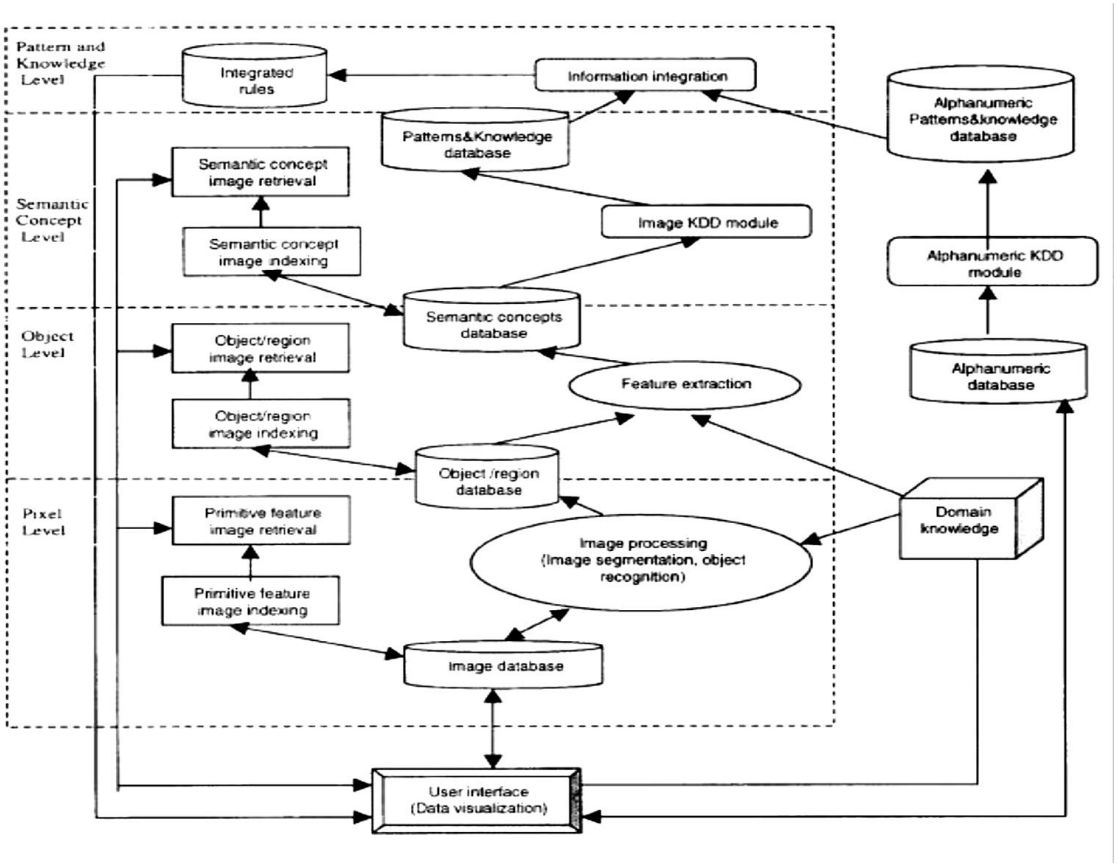
III. IMAGE MINING AND SEGMENTATION TECHNIQUE

Besides investigating suitable frameworks for image mining, early image miners have attempted to use existing techniques to mine for image information. The techniques frequently used include object recognition, image indexing and retrieval, image classification and clustering, association rules mining, and neural network.

Image Segmentation Approach

Image Segmentation is a key task in image processing aiming at partitioning a digital image into multiple objects, which share some common properties. Image segmentation is a critical issue as the quality of its outcomes has a strong influence on the posterior image-understanding task. Among its practical applications are medical imaging (where it is employed for tasks such as tumor location, computer guided surgery, and diagnosis); traffic control systems; object location in satellite images (roads, forests, etc.); and machine vision. Segmentation is one of the most important techniques for image processing]. The purpose of segmentation is to partition an image into distinct, semantically meaningful entities by defining boundaries between features and objects in an image based on some constraint, or homogeneity predicate. Specifically, the segmentation problem is defined as sufficiently partitioning an image into non-overlapping regions.

Figure-8: An Information-Driven Image Mining



Sources: Authors Compilation

Segmentation can therefore be formally defined as follows:

If F is the set of all pixels and P( ) is a homogeneity predicate defined on groups of connected pixels, then segmentation is a partitioning of the set F into a set of connected subsets or regions (S1, S2, …, Sn) such that:

U*ni S F* =1 = *with* I = Æ, *i j S S i* ¹ *j . The homogeneity predicate*
P(Si) = true for all regions (Si) and P(Si U *Sj)=false , when Si is adjacent to Sj.*

Homogeneity predicates are usually based on image intensity, color, texture, etc. According to Harlick and Shapiro, image segmentation can be classified into these categories: spatial clustering split and merge schemes, and region growing schemes.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Spatial Clustering

Haralick and Shapiro present that the difference between clustering and segmentation is that in image segmentation, grouping is done in the spatial domain of the image, while clustering is done in measurement space. It is also possible for clustering to result in overlapping regions, while that is not the case for segmentation results. Clustering and spatial segmentation can be combined to form spatial clustering, which combine histogram techniques with spatial linkage techniques for better results.

Split and Merge Segmentation

Regions in an image are a group of connected pixels with similar properties]. The split method begins with the entire image, and repeatedly splits each segment into quarters if the homogeneity criterion is not satisfied. These splits can sometimes divide portions of one object. The merge method joins adjacent segments of the same object. In intensity based segmentation, the boundaries that separate regions may need to be redefined due to under or over-segmentation of regions. Split and merge segmentation can also handle this task. Under-segmented regions are corrected by adding boundaries to, or splitting, certain regions that contain parts of different objects. Over segmented regions are corrected by eliminating false boundaries and merging adjacent regions if they belong to the same object or feature.

Region Growing

The focus of the remainder of this thesis will be with this class of segmentation. Region growing has shown to be a very useful and efficient segmentation technique in image processing. Region growing in its simplest sense is the process of joining neighboring points into larger regions based on some condition or selection of a threshold value. Seeded region-growing starts with one or more seed points and then grows the region to form a larger region satisfying some homogeneity constraint. The homogeneity of a region can be dependent upon any characteristic of the region in the image: texture, color or average intensity.

THE EXPERIMENT ON SEGMENTATION TECHNIQUES

Description of Test Images

The 10 test images used in the following experiments are taken; full color anatomical images are from the thorax and abdomen regions of the Male dataset. The images are stored as 24-bit 2046x1214 pixel RGB images in RAW format. Color images can be separated into color components based on a specific model. Some of the common color models include red, green, blue (RGB), luminance, chrominance (YUV) and hue, saturation, intensity (HSI). The images are decomposed into three parts representing each of the three components (i.e. red, green, blue for the RGB color model). Our application resizes the image proportionally to an 8-bit 512x302-pixel resolution image.

The resolution is reduced so that more image slices can be kept in memory. The reduction from a 24-bit image to 8-bit image results in utilizing only the red component of the original RGB image to retain the color information, since information is lost when converting color images to grey-scale images. We expect to retrieve more detailed edge information than that retrieved from performing operations on grey-scale images. The original 10 test images are shown in Figure-10. For each of the experiments in this section we attempt to segment the liver from the images.

Figure-9: Volume-Growing Algorithm

Volume_Grow();
Let W be the set of dataset voxels and t be a threshold on
magnitude difference
S = {};
Choose seed voxel w0 with intensity |w0|.
Determine median intensity I in the window of voxels about w0.
Recursive_Region_Grow(W, S, w0, I, t).
Remove Isolated Interior Voxels; Close.
Expand Region Boundary; Dilate One Voxel.
Recursive_Volume_Grow(W, S, w0, I, t):
S = S+W.
If wi Î W adjacent to w0, wi Ï S, and
if |I| - |wi| < t then
Recursive_Region_Grow(W, S, wi, I, t).

Sources: Authors Compilation

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Reconstruction Tool

The overall goal of this proposed paper is to use an appropriate segmentation technique to segment 2D regions to form one 3D object. The 3D objects are rendered using the free source toolkit, ImageJ, using the stack of 2D segmentations as input into the application. In addition to 3D projections, ImageJ can display, edit, analyze, process, save and print 8-bit, 16-bit and 32-bit images. It can read many image formats including TIFF, GIF, JPEG, BMP, DICOM, FITS and "raw". It supports "stacks", a series of images that share a single window. ImageJ was designed with an open architecture that provides extensibility via Java plug-in that can be written with its built in editor and Java compiler. User-written plug-in make it possible to solve almost any image processing or analysis problem.

The Experiments using MIST

This section is dedicated to showing how the MIST algorithm evolved into its final state. The experiments conducted in this section use a seeded region-growing algorithm. The region is grown using the threshold, equal to the standard deviation of each input image in the sequence. The idea is to segment the same anatomical feature from each of the sequential 2D images.

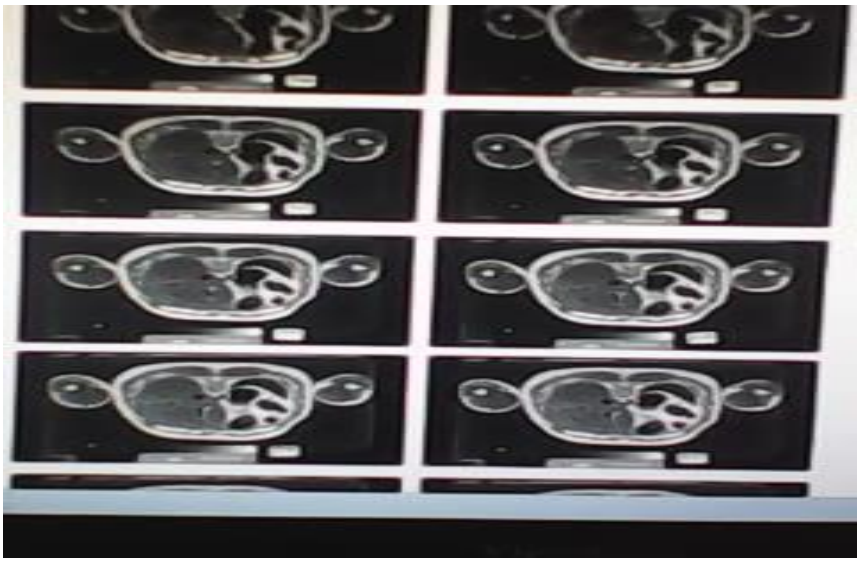
To accomplish this task, the center of mass of a segmented region is used as the seed point for the next image in the sequence. These 2D segmentations are joined together using ImageJ, to create a 3D visualization of the object of interest. Do not alter the size of our segmentations by a significant amount. This ensures us that the segmentations from the modified MIST algorithm produced segmentations with consistent sizes as the segmentations resulting from the MIST algorithm in Experiment 3.

Table-1: Comparing Area of Segmented Regions from Experiments

Image Number	Seed Point for Experiment One	Area of Region After Experiment One, in pixels	Seed Point for Experiment Two	Area of Region After Experiment Two, in pixels
1	(198, 160)	11,647	(198, 160)	12,984
2	(214, 143)	11,176	(215, 143)	12,315
3	(213,144)	11,381	(213, 145)	15,675
4	(210, 130)	1	(209, 129)	12,597
5	(212,145)	11,623	(209, 129)	12,597
6	(212, 145)	11,098	(212, 145)	12,961
7	(211, 143)	11,415	(212, 143)	12,383
8	(211, 143)	3	(213, 143)	13,038
9	(211, 142)	1	(213, 143)	12,966
10	(211, 142)	10,942	(213, 141)	12,623

Sources: Authors Compilation

Figure-10: The Set of 10 Sequential 2D Test Images Used for Experiments



Sources: Authors Compilation

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Table-2: Area of Segmented Regions in Experiments

Image Number	Area of Region After Experiment Three, in pixels	Area of Region After Experiment Four, in pixels
1	12,984	13,603
2	12,315	13,866
3	12,466	13,033
4	12,597	13,269
5	12,961	13,662
6	12,383	13,001
7	12,038	13,631
8	12,966	13,589
9	12,733	13,287
10	12,623	13,153

Sources: Authors Compilation

CONCLUSION

The MIST algorithm corrects the issues faced with Newman's algorithm. Region growing guarantees the segmentation of a connected closed contour, while the use of the contour filling operation ensures that the segmented region of interest is free of gaps and whole artifacts unlike the segmentations produced by Newman et al. In this chapter, we have presented results from the segmentation results produced by Newman's algorithm as well as in each step of our proposed MIST algorithm.

Experimental results show that our MIST method performs better for whole organ and tissue segmentations. Segmentation of the Visible Human Dataset offers many additions to the original goal of a three-dimensional representation of a computer generated anatomical model of the human body and to the general study of human anatomy. In this paper, we have presented a new automatic region-growing algorithm called the Medical Image Segmentation Technique (MIST) that improves image segmentation of 2D contours for reconstructing 3D anatomical structures. It is our first attempt to address the issue of segmenting organs, tissue and other structures from color anatomical images. Seeded region growing offers several advantages over conventional segmentation techniques.

Unlike gradient and Laplacian based edge detection methods, a region found by region growing is guaranteed to be connected and consist of a one pixel thick boundary, since we only add pixels to the exterior of our region. MIST addresses the *adjacency problem*, therefore the segmented region will never contain too much of adjacent tissues, as long as the parameters are defined correctly. In addition, our technique guarantees that he seed is contained in the region by addressing what we call the *centroid problem*, unlike the method presented in [10]. We have compared the results from MIST with papers attempting to achieve the same goals. In our experiments, our method proved to perform better and produce better 3D visualizations.

REFERENCES

1. Liu, H., & Motoda, h. (1998). ***Feature Selection for Knowledge Discovery and Data Mining.*** Kluwer Academic Publisher.

2. Haralick, R. M., & Shapiro, L. G. (1985). SURVEY: Image Segmentation Techniques. ***Computer Vision Graphics Image Processing,*** 1985, 29, pp. 100-132.

3. Revol, C., & Jourlin, M. (1997). A new minimum variance region growing algorithm for image segmentation", Pattern Recognition Letters, 1997, 18, pp. 249-258.

4. Pal, N., & Pal, S. (1993). "A Review on Image Segmentation Techniques", Pattern Recognition, 1993, 26, pp. 1277-1294.

5. Han, Jiawei, & Kamber, Micheline. (2001). ***Data Mining: Concepts and Techniques***. Simon Fraser University.

6. Gonzalez, Rafael C., & Woods, Richard E. (1992). ***Digital Image Processing.***

7. Retrieved from http://www.ijerd.com/paper/vol10-issue5/Version_2/D1051625.pdf

8. Retrieved from http://www.researchgate.net/publication/229010876_Decision_trees_based_image_data_mining_and_its_app.. .

9. Retrieved from http://www.iis.sinica.edu.tw/page/jise/2009/200907_02.pdf
10. Retrieved from http://www.ijerd.com/paper/vol10-issue5/Version_2/D1051625.pdf
11. Retrieved from http://www.researchgate.net/publication/229010876_Decision_trees_based_image_data_mining_and_its_app.. .
12. Retrieved from http://www.iis.sinica.edu.tw/page/jise/2009/200907_02.pdf
13. Retrieved from http://www.computervisiononline.com/software/imagej
14. Retrieved from http://rsb.info.nih.gov/ij/docs/intro.html
15. Retrieved from http://www.ijerd.com/paper/vol10-issue5/Version_2/D1051625.pdf
16. Retrieved from http://arbor.ee.ntu.edu.tw/pakdd02/paper/P0198.pdf

*****

BANKS & ACCOUNT DETAILS

Bank Details for Online Transactions
Important Instructions to remember in case of:
NEFT Transfers / Online payments:

Please forward us the '**Automatic Receipt / Acknowledgement Receipt**' generated, soon after you make online (NEFT) transfer in any of below mentioned banks. Forward the slip on callandinvitation@pezzottaitejournals.net, callandinvitations@pezzottaitejournals.net, callforpapers@pezzottaitejournals.net

- Cash Deposit:
  Please forward us the scanned copy of bank's deposit slip, received after depositing the cash in our account / or send us the photocopy of the same along with Declaration & Copyright Form;
- Demand Draft:
  Please forward us the scanned copy of demand draft. You are directed to keep a photocopy of the Demand Draft with you for future references and to liaison with us.

Note: We don not accepts 'Cheques' in any conditions from researchers and paper submitters.

The said information is needed to complete formalities against your submission.

Name of Bank: UCO Bank
Title of Account: Pezzottaite Journals,
Current Account Number: 07540210000878,
District: Jammu**,**
State: Jammu & Kashmir [India],
Branch: Talab Tillo,
IFSC Code: UCBA0002502 (used for RTGS & NEFT transactions),
Contact: +91-(0191)-2100737.

Name of Bank: Oriental Bank of Commerce
Title of Account: Pezzottaite Journals,
Current Account Number: 12821011000033,
District: Jammu,
State: Jammu & Kashmir [India],
Branch: Trikuta Nagar,
IFSC Code: ORBC0100681 (used for RTGS & NEFT transactions),
Contact: +91-(0191)-2472173.

Details for Demand Drafts

All the Demand Drafts must be made in favour of '**Pezzottaite Journals**' payable at '**Jammu, India**' and to be send at:

Editor-In-Chief
Pezzottaite Journals,
64/2, Trikuta Nagar, K. K. Gupta Lane, Jammu Tawi, Jammu & Kashmir -180012, INDIA
(M): +91-09419216270 – 71

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

ROLE OF ICT IN WOMEN EMPOWERMENT: A STUDY WITH A FOCUS ON 'KUDUMBASHREE' PROGRAMME IN KERALA STATE OF INDIA

Dr. Manoj P. K.[16]

ABSTRACT

Proliferation of Information and Communication Technology (ICT) has notably been playing a cardinal role in women empowerment, financial inclusion and poverty alleviation in India in the ongoing regime of economic reforms. ICT has proven to be a vital tool for enhancement of one's productivity and hence earnings, reduction of manual efforts, higher standard of living, and above all, quicker and faster alignment with the developments in the rest of the world. In fact, the developmental role of ICT has been more prominent in various Governmental schemes for the empowerment of the women, the poor and other disadvantaged sections of the population. In the above context, with special reference to the case of 'Kudumbashree' – the poverty alleviation programme of the Government of Kerala, this paper makes a look into the role of ICT initiatives in empowering women, the problems and prospects of such initiatives, and finally offers suggestions for meaningful use of ICT for women empowerment and poverty alleviation.

KEYWORDS

Inclusive Growth, Microfinance, Women Empowerment, ICT, Productivity etc.

INTRODUCTION

The cardinal significance of Information and Communication Technology (ICT) in bringing about productivity enhancement in business operations as well as socio-economic changes in the society is well documented in literature. ICT has emerged as an important developmental tool right from the 1980s both in the developing and developed countries. It has been pointed out by many experts that the world over rapid developments in the field of ICT and its fast proliferation into all walks of life have brought about remarkable social and economic changes (Barton and Bear, 1999; Fullantelli & Allergera, 2003; and Liu and Luo, 2003). Further, it has been observed that apart from the development of information industry (Barton & Bear, 1999; and Bhasker, 2003), ICT has greatly enabled the development of traditional industries like manufacturing and service in a competitive market; locally, nationally and globally (Barton & Bear, 1999; Wang & Hou, 2003; and Pease & Rowe, 2003). Moreover, it has been observed that focus on IT-enabled operations can help organizations to improve quality on the one hand and provide various channels of delivery to meet customer expectations (Mohammed, 2002); by transforming the marketing and production activities that are considered necessary to directly produce a new product.

Of late, many, particularly with reference to developing nations like India, highlight the role of ICT as a tool for financial inclusion, women empowerment and poverty alleviation. John Paul, Robert Katz and Sean Gallagher (2004) have observed, "ICTs have great potential to catalyze development in some of the poorest regions of the world". In short, it may be noted that ICT is largely recognized as the vital ingredient and most important catalyst that facilitates the fast economic development of any economy; whether in its totality or parts thereof. The developmental role of ICT may be noted to be more prominent in various Governmental schemes for the empowerment of women, the poor and other disadvantaged sections. A number of studies have pointed out the significant potential of ICT and digital technology in women empowerment (Phiphitkul, W. & Sodarak, J. 2002; Agarwal, R. 2009).

RELEVANCE AND SIGNIFICANCE OF STUDY

Sustainable and equitable economic development calls for reducing divides of all kinds, particularly gender divide. In respect of India in particular, empowerment of women is of vital significance because of the low social status for women and significantly low opportunities for upward mobility. As it is widely recognized that ICT is a very powerful tool for fast and sustainable economic development, this paper seeks to look into the 'Kudumbashree' experience in Kerala in women empowerment, and to suggest strategies for more effective use of ICT for women empowerment.

OBJECTIVES OF STUDY

- To study the global experiences in respect of microfinance initiatives on women empowerment with special reference to India in general and Kerala in particular;
- To study the major issues associated with ICT interventions in microfinance initiatives for women empowerment by way of ICT-supported initiatives, with special reference to the 'Kudumbashree' experience in Kerala state in India; and

[16]Assistant Professor, Department of Applied Economics, Cochin University Science and Technology (CUSAT), Kerala, India, manoj_p_k2004@yahoo.co.in

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

- To suggest strategies for more effective use of ICT for women empowerment and poverty alleviation.

MICROFINANCE FOR WOMEN EMPOWERMENT: GLOBAL AND INDIAN EXPERIENCES

There are different perspectives regarding the very need for empowerment. For instance, authors as if Mosedale (2003) has stated that if women were to be empowered, then it should be understood that currently they are disempowered, i.e. disadvantaged by the way power relations shape their choices, opportunities and well-being. According to the author, empowerment cannot be bestowed by a third party but must be claimed by those seeking empowerment through an ongoing process of reflection, analysis and action. But, Kabeer (1999) points out that women need empowerment as they are constrained by "the norms, beliefs, customs and values through which societies differentiate between women and men, and accordingly, it is the process by which those who have been denied the ability to make strategic life choices acquire such an ability", where strategic life choices are "critical for people to live the lives they want".

The definition of the term 'Microfinance' varies from expert to expert. However, a few popular ones are as follows. According to Otero (1999), it is "the provision of financial services to low income poor and very poor self-employed people". Ledgerwood (1999) considers microfinance as one, which generally includes savings and credit and other financial services such as insurance and payment services. For Schreiner and Colombet (2001), it is "the attempt to improve access to small deposits and small loans for poor households neglected by banks".

Regarding the microfinance initiatives in Bangladesh with special reference to the initiatives of BRAC (Bangladesh Rural Advancement Committee), Zaman (2001) has observed that self-help groups intermediated by microcredit have been shown to have positive effects on women, and some of these effects have been reported to have ripple effects as well. It has been reported that such initiatives have played valuable roles in reducing the vulnerability of the poor, through creation of assets, smoothening income and consumption patterns, provision of emergency services, and after all empowering and emboldening women by giving them control over assets and increased self-esteem and knowledge. Another positive report in the Bangladesh context is that of Yunus (1998) with regard to the world-renowned Grameen Bank experience in Bangladesh. The Grameen Bank provides housing loans to members with a record of accomplishment of three loan cycles and with title deeds to the land on which the house is built. As most Grameen members are women, one of the results is that women have had title deeds transferred to them, often from their husbands, to obtain these loans. It has been reported that incidence of divorce since women are owners of their homes cannot be easily evicted (Yunus, 1998). Many impact assessment studies relating to the impact of microfinance on women empowerment have also reported such positive effects. For instance, the 2000 United Nations Common Country Assessment for Bangladesh has reported the lessened severity of poverty and that total income per household raised by 29 percent. Another study undertaken by the World Bank has reported that a 10 percent increase in borrowing has led to an increase in women's non-land assets by 2 percent for loans from the Grameen Bank and 1.2 percent for loans from BRAC (World Bank, 1998).

Discussing the development of microfinance in China, Ruomei (2003) has reported that microfinance has played a critical role in poverty alleviation in the country and has become quite popular in the development front. However, it has been noted that organizational weaknesses prevent its growth. Besides, there are problems in the products and delivery system as well. Llanto, G. M. (2004) has reported that in the context of Philippines, microfinance has promised to provide and improve access to financing by for small farmers, fishermen and micro entrepreneurs. Besides, it has become a credible replacement to the government programmes, which provided finance to the poor. It has been pointed out that in case a systematic and sustained programme devoid of the current limitations (like, populist but unsustainable programmes because of political interference) is developed, microfinance could be made more successful.

Swain and Wallentin (2007) has reported that a study of women in rural Kenya has revealed that in microcredit programs, women do not gain much in terms of decision-making power within the household. However, when loans are channeled through women's groups and are combined with more investment in social intermediation, substantial shifts in decision-making patterns is observed. This involves a remarkable shift in norm-following and male decision-making towards more bargaining and sole female decision-making within the household.

Thus, not all microfinance initiatives have been successful in women empowerment. For instance, a study by UNDP has observed that microcredit will not work in locations that do not have sufficient cash based market activity, are isolated and with low population densities, or are largely self-contained with few outside ties, such as in some Pacific island countries (UNDP, 1997). Accordingly, microfinance is not a panacea in itself, for meeting challenges in economic, social and political empowerment of women. Such schemes have limitations to the extent that they cannot transform social relations and the structural causes of poverty.

As in the case of UNDP (1997) Mayoux (1997) has also reported that the impact of microfinance programmes on women has not always been positive. Increase in income for women can come at the cost of heavier workloads and repayment pressures. The loans availed by women are used by men in the family to set up enterprises, or sometimes women end up being employed as unpaid family workers with virtually no returns. Moreover, it may result in the men withdrawing their support to the family thus reducing the male contribution to the family expenditure. Similar conclusion has been arrived at by Rahaman (1999) who used an

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

anthropological approach with in-depth interviews, participant observations, case studies and a household survey in a village for assessing the influence of microfinance for women empowerment. His study has revealed that their spouses and that tension use about 40 to 70 percent of the loans disbursed to women and violence within the household have increased.

In the Indian scenario, the Task Force on Microfinance (1999) has recognized that microfinance is much more than microcredit and has defined microfinance in a very comprehensive manner. Accordingly, it is "provision of thrift, credit and other financial services and products of very small amounts to the poor in rural, semi-urban and urban areas for enabling them to raise their income levels and improve living standards". The promoters of Self Help Group (SHGs) point out that the first building block of financial services is that of mobilizing savings. In India, the national budget and other policy documents of the Government have been almost equating microfinance with promoting linkage between SHGs and banks. There has been appreciable growth in SHG-Bank linkage in India over the last on decade or more. The policy decision of the central bank, viz. Reserve Bank of India (RBI), to count lending to microfinance institutions (MFIs) towards meeting the priority sector lending targets shows the policy flexibility of the Government towards MFIs.

The term, 'empowerment' particularly when it refers to that of women, involves undoing negative social constructions, so that people affected can perceive themselves as having the capacity and right to act and have influence (Rowlands, 1995). According to Rao and Keller (1995), women empowerment is defined as "the capacity of women to be economically self-sufficient and self-reliant with control over decisions affecting their life options and freedom from violence". The definition of empowerment as per the United Nations (UNDP 2001) is the process by which women take control and ownership of their lives through expansion of their choices. Accordingly, empowerment refers to acquiring the ability to make strategic life choices in an environment where such ability has previously been denied or not available. Two vital processes are observed to be vital for empowerment of women. The first one is social mobilization and collective agency, because poor women often lack the basic capabilities and self-confidence that equip them to counter and challenge the disparities and barriers against them. The second process is economic security, because as long as economic deprivation and livelihood insecurity prevail women would not be able to mobilize.

Microfinance model emerged largely because of the failure of the various subsidized rural banking schemes to reach the poorer sections. During the early 1980s, the All India Debt and Investment Survey (Government of India, 1981) revealed that the share of non-institutional (informal) agencies in outstanding cash dues of the rural households was 38 percent. The main deterrent to the formal system in India in reaching the poor has been the high transaction costs and lack of appropriate credit and savings products. Accordingly, the Self Help Group (SHG)-Bank linkage model evolved from the pioneering efforts of NABARD and two NGOs, MYRADA and PRADHAN, which had persuaded the government in 1992 to remove the obstacles so that NGOs could take loans from banks. It was in the above situation that microfinance programs started being aggressively promoted in India too, as is the case of many of the South East Asian countries, in view of their perceived positive economic impact, and the belief that they empower women. In India, the microfinance started picking up momentum in the 1980s with the formation of pockets of informal SHGs engaging in micro activities financed by microfinance. However, India's first microfinance institution SEWA (Shri Mahila SEWA Sahakari Bank) was set up way back in 1974, by the Self Employed Women's Association (SEWA), soon after SEWA got its registration in 1972.

In the Indian context also, there are several instances of microfinance initiatives having successful track record of women empowerment. Most microfinance programs in India target women with the explicit goal of empowering them, though their underlying premises may be different. Some experts are of the view that women are amongst the poorest and the most vulnerable among the underprivileged sections. Some others, however, believe that investing in women's capabilities empowers them to make choices, which in turn will contribute to greater economic growth and development. A third group argues that an increase in woman's resources results in higher well-being of the family, especially children. In spite of the above conceptual and definitional issues, it is meaningful to look into the major empirical findings on the role of microfinance on women empowerment in India. SEWA has been functioning quite successfully since the early 1970s in India and it is the flagship microfinance movement that has been so successful in women's empowerment- economic, political and social.

Apart from SEWA head quartered in Ahmedabad, there are quite a large number of successful microfinance institutions in India, devoted to the cause of women. DHAN foundation in Madurai (Tamil Nadu state) is one such organization. Their microfinance programme called, "Kalanjiam Community Banking Programme" is one of the largest microfinance programmes based on promotion of community institutions reaching out to over 2,50,000 women, benefiting over a million poor people in the process. A case study by Rajagopalan, S (2006) has reported significant empowerment of women through the participation in 'DHAN' activities by way of improvement in their economic, social and political status. Another study in the context of Andhra Pradesh by Reddy, C. S. et. al. (2006) regarding the activities of APMAS (Andhra Pradesh Mahila Abhivrudhi Society), a society founded in 2001. Functions of APMAS act as the support organization for the self–help movement. APMAS has begun its operations towards promoting livelihood, taking advantage of its experience of quality assessment and enhancement. It has been reported that the unique business model of APMAS as the support organization to help develop SHGs has helped it to become a role model for similar organizations in other states, within a short span of time. The case of "Kudumbashree" in Kerala is another unique model, a Government sponsored project; its case is explained in detail in the next section (Part – III) of this paper.

Swain and Wallentin (2007) have observed that on average there is a significant increase in women's empowerment in respect of the SHG members' group. No significant change is noted on average for the members of the control group. It is noted that even

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

though the degree of change and the pace of empowering women are likely to vary, the SHG members' group has experienced a significant and higher empowerment.

Of late, there are growing instances of Indian women, particularly the poor, taking up micro-enterprise development, in spite of the various limitations that they suffer in terms of inadequate capital, lack of requisite technical expertise etc. The experience of having worked with microfinance initiatives acts as a confidence-booster. For instance, Agarwal, R (2009) has pointed out that Indian IT sector "has not only created a large number of jobs, but has also resulted in new types of challenging careers" and that "in the IT sector, emphasis is on intellectual rather than physical resources". Accordingly, the author has observed that in the IT sector as the emphasis is on knowledge "the industry is considered to be non-discriminating" and "an equal opportunity employer for men and women, minorities and handicapped all alike". Thus, ICT provides, better opportunities for women empowerment than any other industry.

ICT FOR ECONOMIC DEVELOPMENT: AN OVERVIEW OF THE MAJOR ICT INITIATIVES OF THE GOVERNMENT OF KERALA IN THE INDIAN UNION

Kerala is one of the most developed states of Indian Union. It has got to its credit an array of unparalleled achievements, which include inter alia, universal literacy, physical quality of life index (PQLI) on par with advanced nations, international standards in social development, service delivery points in health care, education and public distribution up to the village level and a very industrious manpower which is well-known the world over. Now Kerala is in the forefront of the global trend towards a new 'knowledge-based' economy. The very peculiar socio-economic climate of Kerala is, in fact, quite conducive for becoming a 'knowledge society', which it is striving to achieve over the years. Few regions in the developing world are as ready to embrace IT technology at the grass-root level, as is Kerala. This is because of its potential to generate opportunities and employment with little pressure on land, environment and other resources. Further, this is one of the most people-friendly and environment-friendly industries of modern times.

The Government of the State has realized the enormous potential of Information Technology. Firstly, it has been recognized as a very effective tool for improving governance and creating more jobs. Secondly, and probably more significantly, it has been recognized as a means for greatly enhancing the standard of living of the people. Over the years, the use of ICT in enhancing the delivery of Government services has proved to be quite responsive. Besides, ICT initiatives have been successful in offering a transparent administration, which can facilitate the empowerment of people and satisfy their right to information.

ICT Initiatives: The Kerala Advantage

The specific advantages of the state of Kerala in embracing ICT initiatives in comparison with developing nations in general and other states of the Indian union in particular are as follows:

- Kerala is a highly advanced society. The wide mass base which the media enjoys in the State and the penetration that communication technologies have been able to make in Kerala, will see Kerala emerging soon as a 100% internalized State - very truly, an Information Society.

- Kerala has the highest literacy among the different states of India. Besides, it has the highest exposure to the different media. Thus, what the industry needs most is what the State has the most of, namely educated men and women seeking white-collar jobs.

- Kerala has the highest proportion of citizens living outside the State. Being exposed to the best from the latest in terms of products, services, technologies and life styles, Keralites can very effectively impart strong 'demonstration effects' on those living in the State.

- Keralites by nature have high degree of awareness of and willingness to adopt the latest. Because of the peculiar cultural heritage of the state, with very diverse but quite harmonious religious and communal groups, Keralites have got very high level of tolerance for cultural diversities.

- Kerala society gives the highest premium to education and is willing to invest heavily for acquiring marketable skills. It has a vast pool of English speaking technical and science graduates.

- Kerala has abundance of most of the infrastructural facilities required for the proliferation of Internet - the next frontier of growth in the IT sector. None of the other major Indian States has such an extensive fibre optic network, as does Kerala. This network which reaches right down to the block level, along with the digital exchanges most of which have ISDN capability represent a formidable backbone for the making of a fully networked intelligent State. With the commissioning of the two submarine cable landings at Cochin, Kerala has emerged as a major telecom gateway for the country.

- Kochi, the coastal city in the central part of Kerala and one of India's second-tier cosmopolitan cities and is fast emerging as a unique IT destination; because of its locational, infrastructural and other advantages. Kochi was ranked number two in a recent report by NASSCOM regarding the 'Super ITES Destinations' of the country. Besides, it was ranked number three by NASSCOM in a study of cities ideal to do business. Moreover, this city is directly connected by two submarine cables and satellite gateways, which in turn are used to support cities including Bangalore. Thus, the clear telecom infrastructural advantage offers Kochi better reliability as well as reduced tariffs. Apart from the major attraction of Kochi because of the presence of 'Infopark' which has been conferred the status of a SEZ (Special Economic Zone) by the Central Government in September 2006. Moreover, it is now all set to host another mega IT project named 'Smart City' – a collaborative project of the State with M/s. Dubai Internet City (DIC), Dubai, UAE. It is estimated that 'Smart City' project can offer over 33,000 employment opportunities within the next ten years.

- Kerala's development experience is characterized by significant progress in 'ICT for Development' sector in terms of investments, infrastructure development and employment generated in the sector through focused initiatives. The remarkable initiative in this regard has been the project 'Akshaya', which is set to be rolled – out to reach all 64 lakh families of the State (presently being rolled – out in seven districts). The first 'Citizen's Call Centre' in the entire country is another glaring example of Kerala unique position in respect of ICT for Development.

- There has been a high spurt in the use of computers and Internet in the State and in the sale of PCs and accessories in the State (assembled PC segment, Indian brands and MNC brands) with the increased emphasis given to the e-governance sector. Household users are now increasingly considering the PC as an aid for education and entertainment. Besides, as per the findings of a recent survey conducted by MAIT (Manufacturers' Association in IT), the sale of hardware is growing in the State. At present as high as 7 Internet Service Providers are functioning in the State – about double of what was available about 4 years' back.

ICT FOR WOMEN EMPOWERMENT: THE CASE OF "KUDUMBASHREE" IN KERALA

Kerala has an appreciable record of accomplishment of women empowerment through the microfinance route in the entire country. Apart from improving the socio-economic status of women, there is significant political empowerment also as they become participants in the decision-making process as members of various committees. Likewise, ownership of land, house or other assets in due course, also makes them empowered. A recent study done by Bina Agarwal in Kerala has found that women's risk of physical violence from husbands is dramatically less if they own hand or a house. The incidence of violence is 49 per cent among women without property, but 18 per cent among land owning women and 7 per cent of they own both land and house. "Kudumbashree" members have undertaken ICT-based Enterprises (IBEs) in a big way over the years. The number of IBEs is on the rise vis-à-vis traditional units, as more and more units become equipped to take up such initiatives.

"Kudumbashree": the Poverty Alleviation Programme of Kerala Government

It was as part the decentralization, measures initiated by the Government of Kerala in October 2005 and August 2006, that the idea to universalize the Anti-Poverty Programme of the State under the name of "Kudumbashree" took shape. Accordingly, Government of Kerala launched Kudumbashree, the State Poverty Eradication Mission (SPEM), in 1998 with the active support of Government of India and NABARD for wiping out absolute poverty within a period of 10 years. The project is implemented through Local Self Governments empowered by the 73rd and 74th Constitutional amendments.

The slogan of the "Kudumbashree" Mission is "Reaching out to families through Women and reaching out to community through Families". Kudumbashree is a holistic, participatory, women oriented innovative overarching poverty reduction approach. 'Kudumbashree' envisages prosperity of the economically backward families in the state with multiple programmes that will provide them information, create awareness, build up their capability and capacity, enhance their confidence and show them opportunity for better social security and empower them physically, socially, economically and politically.

Kudumbashree has altered lives of economically backward women in the state, changed their perception, built their confidence, boosted their morale, rediscovered their dignity and honor, and empowered them economically, socially and politically. As of 2009, over 3.6 million women participate in the Kudumbashree movement in the state cutting across political ideologies and religious faiths. Within a short span of one decade or more, Kudumbashree could bring about considerable change in the lives of women of Kerala by converging resources, ideas and programmes of various departments through the CDS system. Over 36 lakh women of State have been organized into 1,82,969 grass root level Neighborhood groups (NHGs). Apart from thrift mobilization and informal banking, the CDS structure has given birth to 29,436 vibrant Micro Enterprises making around 54949 women owners of these units.

"Kudumbashree": the Strategy

- Identification of poor families using a non-monitory poverty index,
- Organizing the poor to a 3-tier Community Based Organization (CBO),

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

- Empowerment of women through Community Based Organizations,
- Formation of Informal Bank of Poor women operating round the clock throughout the year, starting from thrift & credit operations,
- Formation of micro-enterprises,
- Convergent Community Action,
- Strong liaison with LSGs (Local Self Governments),
- Intervention in Anti-Poverty Sub Plan, Women Component Plan, Local Economic Development of local self-governments,
- Community monitoring mechanism for local development.

ACTIVITIES OF "KUDUMBASHREE"

Economic Development Activities

Thrift & Credit

Kudumbashree promotes Thrift mobilization by setting up Thrift & Credit Societies at NHG level to encourage the poor to save and to provide them cost effective and easy credit. Thrift and credit Societies facilitate easy and timely credit to the un-reached. The amount of loan and the priority of disbursement are decided by the NHG. The repayment is collected weekly during the routine NHG meetings. The total thrift collected by NHGs in the state comes to Rs 844.95 crore and the internal loans generated are to the tune of Rs.2255.59 crore.

Linkage Banking

The Bank Linkage programme has helped the NHGs to augment their existing resources collected through thrift. The efficiency and effectiveness of the NHGs are verified based on some objectively verifiable and easily identifiable parameters. NABARD has developed a 15-point index for rating NHGs because of which they will be allowed to link with various banks under the Linkage Banking Scheme. The total amount, which has been mobilized under linkage banking, is Rs 461.81 and 87715 NHGs have availed of the loans.

Bhavanashree

The micro housing loan scheme for the poor called "Bhavanshree" has been designed by the Kudumbashree mission to meet the housing needs of the poor. Since the starting of the scheme 41284 loans worth Rs 168.11 have been dispersed by the various banks under the Bhavanshree programme.

Micro Enterprise

Kudumbashree views Micro Enterprise Development as an opportunity for providing gainful employment to the people below poverty line and thereby improving their income and living standard. Kudumbashree considers micro enterprise development as an emerging process which will start with low capital low risk and low profit at the initial stage which will gain momentum and later switch to low to medium capital and then too low to medium risk. Micro enterprises facilitated by Kudumbashree fall under the banner of Rural Micro Enterprise (which takes after the SGSY pattern), Yuvasree (which caters to educated youth), and the SJSRY programme (which is implemented through the agency of Kudumbashree).

Women Empowerment and Community Development

Asraya

Kudumbashree designed a project called Asraya - Destitute Identification Rehabilitation and Monitoring Project for the rehabilitation of destitute families. The State Government's assistance of Rs 10 Lakh per LSG was a Challenge fund for extending support to local governments that take up the implementation of the project. The project envisages addressing the various deprivations faced by destitute families such as lack of food, health problems including chronic illness, treatment for life threatening diseases, pension, educational facilities to children, land for home, shelter and shelter up gradation, drinking water, safe sanitation facilities, skill development, employment opportunities, etc. Altogether 49,080 destitute families have been brought under the purview of the scheme so far.

Buds

This is a Special school for mentally & physically challenged children .It caters the developmental, social, and emotional needs of children whose needs were formerly were often ignored or neglected. These young children battle with disabilities including Autism, Cerebral Palsy, blindness, deafness, and speech impairments due to cleft palates. They are being provided with necessary

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

medical attention, physical and mental therapy, mobility equipment, hearing aids, individual vocational and educational training, and skills to enable them to be independent. There are 10 such schools.

Bala Sabha

Balasabhas are grass root level groups of the children from NHG families. The concept came from experiences in the field when children used to accompany their mothers to the NHG meetings and they began to have their informal children NHG meetings. The mission organizes the children into Balasabhas as a part of its holistic approach to provide opportunities for personal growth and realization of talent. Capacity building, cultivation of creative potentials, ensuring of opportunities to express, learn and develop, offering opportunity to identify and nurture innate talents.

Finishing School

Capacity building to improve the employability of educated youth from poor families is a major challenge that has to be undertaken by the State. The Government of India through the special SGSY stream sanctioned an innovative project for this purpose. Dr. Reddy is Foundation, Hyderabad, designs this project.

Gender Resource Center

This focuses on addressing gender concerns, and capacitating Kudumbashree NHGs and LSG representatives to develop gender sensitive policies and to prioritize gender sensitive action. This would require constant collation and analysis of local gender issues and supporting a resource pool of gender experts to facilitate equitable policy setting and action in the field.

Social Audit

In Kerala, the process of social audit has not really taken root. In its 11th Five Year Plan, the state government chose the Kudumbashree NHGs as the main forum for social audit. This will be started in the districts where the NREGA programmes are being implemented. The quality of the work, the labour card distribution system, the wages and the time etc. would be audited and the report would be presented in Grama Sabhas. A dimension of this intervention is that women are being empowered through a democratic process.

Training

Kudumbashree gives prime importance to training and initiated a series of efforts in this regard. There are mainly two broad categories of training viz. (1) for team building and social mobilization and (2) for skill development. Major training areas fall into 4 types viz, enterprise related, micro finance related, social development related and gender development related. It may be noted that "Kudumbashree" has an impeccable record of accomplishment of women empowerment and all allied development activities. Over the limited period of over one decade, it has become a role model for the entire country. Some of the other state governments in India are in the process of replicating the "Kudumbashree" model in their own states.

ICT-BASED ENTERPRISES (IBES) BY "KUDUMBASHREE" AND THEIR POTENTIAL FOR WOMEN EMPOWERMENT: AN ASSESSMENT

The Kudumbashree poverty eradication initiative is a unique scheme that has created a statewide series of cooperative microenterprises, which are all owned, managed and operated by women from "below poverty line" families. In addition to more "traditional" enterprise, Kudumbashree has facilitated the creation of more than 100 ICT-based enterprises (IBEs) units that provide jobs for more than 1,000 poor women. Each unit employs ten women from its immediate neighborhood, plus others on a piece-rate basis. The units are focused around three main activities: data entry/digitization, IT training, and hardware assembly. Women's IBEs have been observed to be bringing benefits to both women (its members) and the agencies (like, NOGOs) that support such IBEs and these are shown in Table (I) below:

Table-(I): Different Benefits from Different Viewpoints Viewpoint

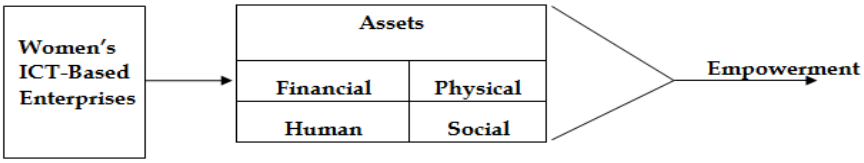
Perspectives (View Points)	Benefits Focus	Benefit Example
Enterprise	Performance of the business.	Growth of a sustainable enterprise.
Livelihood	Changing livelihood assets of individuals or families.	Generating stable employment and income.
Gender	Approach to gender equity.	Changes in opportunities for women or cultural attitude changes to the choices open to women and the role they can play.

Sources: Authors Compilation

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Empirical studies on real cases of women entrepreneurs working in IBEs by Duncombe, Heeks et. al (2005) has revealed that there are many benefits accruing to women undertaking such initiatives. Accordingly, the major benefits that lead to empowerment of women are as shown in Figure (I) below:

Figure-(I): Categories of Livelihood/Gender Benefit for Women's IBEs



Sources: Authors Compilation

The major benefits from women's IBEs are as noted by Duncombe, Heeks et. al (2005) based on their empirical study are summarized hereunder:

✓ **Improved Financial Assets**

Regular income from work in ICT-based enterprises has enabled women to contribute to their family welfare (such as healthcare, education or payment for marriage of siblings) and even to add regularly to savings, when previously they might be unemployed with no income.

✓ **Improved Physical Assets**

Regular income from ICT-based enterprises has enabled women to purchase land, housing, gold or physical goods for their family and to purchase hardware and software equipment for use at work where previously that would not have been possible.

✓ **Improved Human Assets**

Women develop personally and professionally through work in an ICT-based enterprise, particularly in terms of technical skill development and in personal confidence. Many become involved with management activities and decision-making, and some develop entrepreneurial skills such as an understanding of cash flow, customer service, etc.

✓ **Improved Social Assets**

Women working in an ICT-based enterprise see improvements in three main areas of social relations: links to customers and suppliers (business linkages); links to support agencies plus banks or credit unions (other institutional linkages); and links to other women working in the enterprise or in similar/nearby enterprises (social and community linkages).

✓ **Empowerment**

Women working in ICT-based enterprises seem to talk about this more than anything else. They talk about gaining confidence to apply new skills, to tackle problems, to deal with businesses and agencies. They talk about new "respect", "recognition" and "acceptance" within their communities. As a result, they can make some inroads into traditional gender biases: taking on management roles traditionally seen as "men's work"; hiring and managing men as employees; taking on traditionally male activities like working late or travelling with their work; and having a different role in their families due to their new income and status.

Additionally, potential gender-related benefits for women involved in ICT-based enterprises can include the following:

- Employment and financial independence particularly in the absence of social safety nets;
- The opportunity for skills development in a women-only environment;
- Work suitable for disabled women who might otherwise have limited options;
- The opportunity to increase competitiveness, both within the economy as well as in relation to men; and
- A means to involve women in ICT policy making.

There may also be broader benefits. These women are likely to act as role models for others. Their enterprises may create a "business node" that encourages other businesses – including women-run businesses – to set up. They may start to create a critical mass of skilled women, making migration of those women to other areas and other sectors less likely. More generally, they start to close the gender digital divide.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Table-(II): Women's ICT-Based Enterprises under "Kudumbashree"

Units	Number of Units	Location		Average Number of Core Group Members	Average Number of Supporting Staffs
		Urban	Rural		
IT @ School Units	151	72	79	6	1021
Data entry Units	72	42	30	10	3200
Hardware Assembly Units	5	4	1	6	15

Sources: Authors Compilation

There are 228 Kudumbashree ICT enterprises spread across 14 districts in Kerala. Based on their activities, these units can be classified into three viz. (i) Data processing units, (ii) IT @ School, and (iii) Hardware assembling units as shown in Table (II). Based on a study of the data processing units alone by taking a sample of 50% (i.e. 36 out of 72), Chandrasekhar & Siva Prakash (2010) have made the following observations:

- Women ICT enterprises are significantly helping empowerment of women,
- Women ICT enterprises are facing difficulty in obtaining input resources,
- Women ICT enterprises are providing job satisfaction to its members,
- The sustainability of women ICT enterprises in Kerala is difficult at present.

Thus, Chandrasekhar & Siva Prakash (2010) have concluded that women ICT initiative under "Kudumbashree" has vast potential for empowering poor women. The socio-economic barriers of women could be effectively overcome by ICT initiatives.

MAJOR PROBLEMS OF ICT-BASED ENTERPRISES (IBES) OF WOMEN

Let us consider some of the major problems faced by women's ICT-Based enterprises (IBEs). Studies have revealed that the following are the major challenges of IBEs:

- Getting the payments for the work done is one of the worst problems, which results in unsustainable cash flows. The problem is chronic in respect of Government assignments. Cash crunch results in payment delays to workers of IBEs, who have to be paid promptly irrespective of the receipt of contract money.
- In respect of most micro-enterprises started by women, the worst problem is relating to marketing. Lack of sufficient knowledge regarding the market and the potential profitability are quite common. Often problems are there in pricing of products, apart from marketing.
- Credit policies that can gradually ruin their business (many customers cannot pay cash; on the other hand, suppliers are very harsh towards women).
- Long-term sustainability of women's IBEs is often a question mark, given the multifarious problems that they face.

SUGGESTIONS FOR MORE EFFECTIVE FUNCTIONING OF WOMEN'S IBES

- Though credit is required for women to take up income generating activities, like IBEs, credit alone will not lead to their empowerment. Supporting measures are also required (like, marketing, prompt payment for work done etc.).
- Additional services like training, awareness raising workshops and other activities are also an important determinant of the degree of its impact on the empowerment process of women through ICT based initiatives.
- It is quite advisable to bring together the different players in the ICT sector to draft coherent policies and procedures for more effective functioning of IBEs. Thus, ultimately, the focus should be on developing diversified women's IBEs where different organizations work coherently, collaborate and work together to make a significant contribution to gender equality and pro-poor development.

CONCLUDING REMARKS

In view of the foregoing, it may be noted that while there is vast potential for women's ICT-based enterprises (IBEs) for empowering poor women in various fronts (like, economic, social and political), this potential is yet to be effectively tapped. Thus, encouraging of IBEs is of utmost significance in the present scenario, because of the multiplicity of credit needs for the poor and the capacity of ICT to bring about fast and equitable economic development. Equally important is the need to design flexible credit instruments for the poor women to enable them to go ahead with ICT-based initiatives.

REFERENCES

1. Chandrasekhar, K. S., & Siva, Prakash C. S. (2010, November). The Role of Information and Communication Technology in Women Empowerment and Poverty Eradication in Kerala. ***APJRBM,*** I (2).

2. Kabeer, N. (2001). Resources, Agency, Achievements: Reflections on the Measurement of Women Empowerment. ***In Discussing Women's Empowerment – Theory and Practice, (***SIDA Studies No. 3*).*

3. Llanto, G. M. (2004). Microfinance in Philippines. In Ghosh, R. (Ed.), ***Microcredit in Asia, Concepts and Cases.*** Hyderabad: ICFAI University Press, 2005.

4. Mayoux, L. (1997). ***The Magic Ingredient? Microfinance and Women Empowerment, a Briefing***, Paper contributed to the Microcredit Summit, Washington, USA.

5. Rajagopalan, S. (2006). DHAN Foundation: Microfinance through Community Banking. In Bhargava, H., and Sharma, R. (Eds), ***Managing Microfinance, a Corporate Approach***, 233-245.Hyderabad: ICFAI University Press.

6. Reddy, C. S., Livingstone, J., & Manak, S. (2006). Institutionalizing Support to SHGs: The APMAS Experience. In Bhargava, H., and Sharma, R. (Eds), ***Managing Microfinance, a Corporate Approach***. Hyderabad: ICFAI University Press.

7. Richard, D. (*et. Al*.). (2005). ***Supporting Women's ICT Based Enterprises***. DFID.

8. Ruomei, S. (2003). The Development of Microfinance in China. In Ghosh, R. (Ed.), ***Microcredit in Asia, Concepts and Cases.*** Hyderabad: ICFAI University Press. 2005.

9. Swain, R. Bali, & Wallentin, F. Yang. (2007, December). Empowering Women through Microfinance. ***Poverty in Focus.*** UNDP.

10. (2001). ***Participatory Governance, People's Empowerment and Poverty Reduction (***UNDP Conf. Paper Series). United Nations Development Programme.

11. Yunus, Muhammed, & Alan, Jolis (1998). ***Banker to the Poor: The Autobiography of Muhammed Yunus***. London: Aurum Press.

12. Zaman, H. (2001). ***Assessing the Poverty and Vulnerability Impact of Microcredit in Bangladesh: A Case Study of BRAC*** (Background paper for World Bank, World Development Report 2000/2001). Washington: World Bank.

13. Retrieved from http://indiagovernance.gov.in/files/kudumbashree-casestudy.doc

14. Retrieved from Retrieved from http://iosrjournals.org/iosr-jhss/papers/Vol3-issue6/B0360612.pdf

15. Retrieved from http://www.lsg.kerala.gov.in/htm/Aboutus.asp

16. Retrieved from http://iosrjournals.org/iosr-jhss/papers/Vol3-issue6/B0360612.pdf

17. Retrieved from http://www.ourkeralam.com/usefulinformation/itinkerala.htm

18. Retrieved from http://www.expert-eyes.org/archive/itpol98f.html

19. Retrieved from http://www.shividevelopmentsociety.org/women-human-rights-defenders.html

20. Retrieved from http://www.managementparadise.com/kundanhaha/documents/12324/working-paper-for-financial-inclusion-e...

21. Retrieved from http://www.kudumbashree.org/?q=thrift

22. Retrieved from http://www.kudumbashree.org/?q=buds2

23. Retrieved from http://news.entecity.com/corporation-to-open-its-first-buds-school

24. Retrieved from http://www.womenictenterprise.org/project.html

25. Retrieved from http://iosrjournals.org/iosr-jhss/pages/v3i6.html

****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

COMPARATIVE STUDY ON NUMBER OF TEST CASES USING DIFFERENT SOFTWARE TECHNIQUES

Ashima Wadhwa[17] Havish Madhvapaty[18]

ABSTRACT

Software development growth has been stupendous over the last few decades. Software testing which is aimed at reducing errors, cutting maintenance and overall software costs - has also been keeping pace. The core purpose of software testing is aimed at enhancing software quality. Testing typically consumes 40–50% of development efforts and exacts very high effort from systems that require higher levels of reliability. Hence, the importance of software testing as a part of software engineering cannot be undermined. One of the major problems within software testing area is generating a suitable set of cases to test a software system. The set should be generated in such a way so that it assures maximum effectiveness with the least possible number of test cases. There are numerous testing techniques available for generating test cases broadly grouped under White Box and Black Box techniques. This paper gives a comparative study on number of test cases using two White Box techniques (Basis Path testing; Graph Matrix) and two Black Box techniques (Worst Case Analysis; Boundary Value Analysis).

The objective of the paper is to use the study of these representative techniques to highlight the future implications for several areas of both industry and academia. As a representation of the way, the industry can use these techniques in ways so far unthought-of, primary sales data collected from a primary survey by students as a part of a Live Project conducted by Wipro Consumer Care has been used. Black Box testing has been used to show how the numbers can be inputted into a system. It is to be noted that the results can easily be extrapolated and scaled to huge sets of data as well.

KEYWORDS

Software Testing, Basis Path, White Box Testing, Black Box Testing, Graph Matrix, Worst Case Analysis, Boundary Value Analysis etc.

INTRODUCTION

Software testing is a set of activities conducted with the intent of finding errors in software. It also verifies and validates whether the program is working correctly; and whether there is the presence of any bugs. Software testing is not just used for finding and fixing of bugs but also to ensure that the system is working according to the prescribed specifications. Software testing is a series of processes, which is designed to make sure that the computer code does what it is designed to do. The presence of software bugs cannot be completely negated. Any software module of moderate size will usually have a few bugs. This is not because of a lackadaisical approach of the programmers – but because complexity of software is generally intractable and humans possess only limited ability to manage complexity. It is also true that for any complex systems, design defects can never be completely ruled out. Therefore, software testing is executing a system in order to identify any gaps, errors or missing requirements in contrary to the actual desire or requirements.

Terminology for Software Testing

- **Specification**
  A document that specifies in a complete, precise, and verifiable manner the requirements, design, behavior, or other characteristics of a system or component; and often the procedures for determining whether these provisions have been satisfied. We observe errors, which can often be associated with failures. However, the ultimate cause of the fault is often very hard to find.

- **Mistake**
  A human action that produces an incorrect result.

- **Fault (or Defect)**
  An incorrect step, process, or data definition in a program.

- **Error**
  The difference between a computed, observed, or measured value / condition; and the true, specified, or theoretically correct value / condition.

[17]Assistant Professor, Amity Global Business School, Noida, Uttar Pradesh, India, awadhwa1@amity.edu
[18]Assistant Professor, Amity Global Business School, Noida, Uttar Pradesh, India, havish.madhvapaty@gmail.com

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

- **Failure**
  The inability of a system or component to perform its required function within the specified performance requirement.

Software Testing Techniques

The main aim of software testing techniques is to provide total number of test cases. This way software correctness is provided as only those test cases are chosen which have greatest probability of finding errors. Hence, the testers are not asked to decide which test cases are to be chosen. Test techniques enable them to design testing conditions in a systematic manner. Additionally - if one combines various types of existing test techniques, one will get better results as opposed to using only one testing technique. The software can be tested in following three ways:

White Box Testing
The White Box testing method is that software entity which can be viewed as a "White Box". The generation of test cases is totally based on the execution of the software entity. The process of White Box testing is giving the input to the system and then to check how the system entity processes that input in order to get the required output. It is required for a tester to have complete knowledge of the source code and programming. White Box testing is mostly used at unit integration and system levels of the software testing process. The main aim of White Box testing is to ensure that all parts of test objects are properly and completely executed.

Advantages:
1. This technique reveals hidden errors in the code.
2. It approximates the partitioning of code done by execution equivalence.
3. Apart from providing number of test cases, this technique encourages the developer to give reasons about implementation of code.

Disadvantages:
1. This technique is very expensive.
2. Only the programmers who have the knowledge about code can use this technique.

Black Box Testing
Black Box testing tests the software with input and output requirements. This essentially means that anybody can test the software using this technique even without any prior knowledge of the code and internal structure of the program. Only the tester knows the inputs, expected outcomes, and ensures that program arrives at those outputs. The Black Box tester does not ever examine the coding and need not possess any future knowledge of the code other than its specifications.

Advantages:
1. Black Box tester need not to be a programmer.
2. The process is very simple as compare to White Box method.
3. In this case, the programmer and the tester are independent of each other.

Disadvantages:
1. Test cases are hard to design without clear specifications.
2. This method can be executed with small number of inputs.
3. There is possibility that some parts of the code are not tested at all.

Grey Box Testing
Grey Box testing method combined the testing process of both White Box and Black Box. This method is used for testing a piece of software against its input and output specifications whilst using some knowledge of its code and internal structure as well. Grey Box testing method may also include reverse engineering to get the number of test cases. As a result, this method incorporates the requirements of both White Box and Black Box testing methods. The knowledge of internals of the code in Grey Box testing is more than Black Box testing, but less than White Box testing.

Advantages:
1. This method provides combined advantages of black box and white box testing wherever possible.
2. Grey Box testers do not rely on the source program; instead, they rely on interface definition and functional specifications, which makes the task easier than White Box method.
3. This method needs limited information and designs excellent test cases especially in communication protocols and data type handling.

Disadvantages:
1. Because access to the source program is unavailable, the ability to go over the program and ease of test coverage is limited.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

2. Testing every possible path under this approach is unrealistic, as it would take an unreasonable amount of time; therefore, many program paths will go untested.

LITERATURE REVIEW

Mohd. Ehmer Khan (2010) highlighted the four purposes of software testing: correctness testing, performance testing, reliability testing, and security testing. His paper stated that since it is nearly impossible to find all errors in a program, the fundamental question is to find out the strategy employed in testing.

Mohd. Ehmer Khan (2011) described the working process of White Box technique. He stated that need to understand the White Box techniques that are available to make educated decisions about their use for the specific system.

Kwang Ik Seo and Eun Man Choi (2006) presented the empirical comparison of major Black-Box testing methods and showed the different results by applying them to test a certain software system. They concluded that a combination of efficient test methods, which combine extended use case test method and OCL test method, would be helpful in planning and improving the performance of software testing.

R. Lutz and I.C Mikulski (2003) described the role of requirements discovery during the testing of a safety-critical software system. The results also suggested that "false positive" a problem report from testing (in which the software behaves correctly but unexpectedly) provides a rich source of requirements information that can be used to reduce operational anomalies in critical systems.

OBJECTIVES OF STUDY

The core objective of the paper is an analysis of Black Box and White Box technique.

- To use primary market survey data to show application of Black Box testing to an actual market scenario.
- To illustrate two major techniques of each Black Box and White Box and generate the number of test cases of given problem.

The paper also attempts to highlight the possible uses for specific Black Box and White Box testing techniques for both industry and academia.

RESEARCH METHODOLOGY

Market research data conducted from a primary market survey has been used.

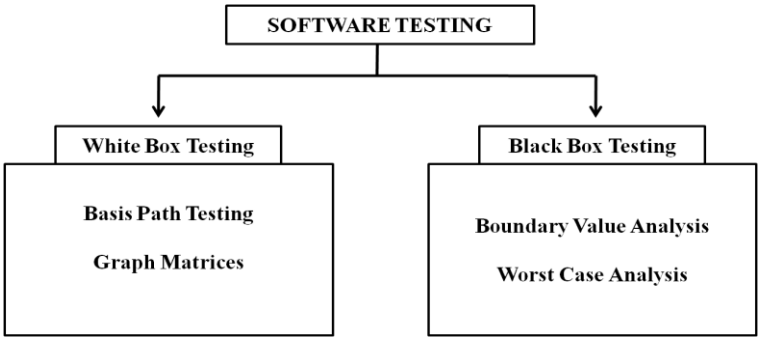
- *Mode of research:* Field survey
- *Survey Tool:* Questionnaire

Students working on a Live Project for Wipro Consumer Care filled the questionnaire. The area covered for the survey was Delhi – NCR region.

TESTING

The testing has been done under 2 different types of White Box and Black Box technique each. This has been shown in Figure 1.

Figure-1: Types of Testing Techniques Used in Paper



Sources: Authors Compilation

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Techniques Under White Box Testing

The two techniques used in this paper are Basis Path testing and Graph Matrices.

Basis Path Testing

Basis Path testing is one of the important White Box testing techniques which was first proposed by Tom McCabe. This technique allows the test case designer to produce a logical complexity measure of procedural design, which is used as the measure for outlining a basic set of execution path. (*Basic set is the set of all the execution of a procedure*). These test cases process basic set and ensure that every statement of code is executed at least once. The basis path testing is done with the help of a control flow graph – which is a graphical representation of flow of control in a program. There are two parts to this – Flow Graph Notation and McCabe's Cyclomatic complexity.

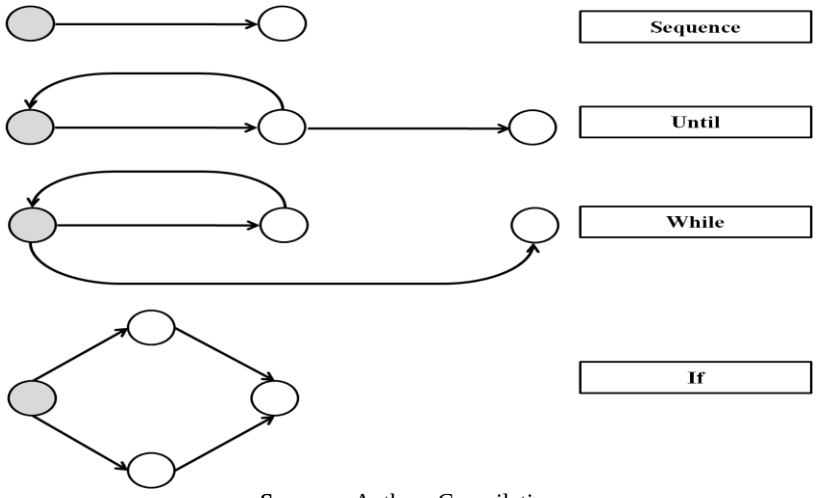
Flow Graph Notation

The flow graph is the essence of basis path testing which is used to get the program control structure. The notations are as follows:

- The circle in a flow graph is called flow graph node.
- The arrows are called edges or links (represent flow of control).
- The Area bounded by nodes and edges are called regions.

The individual notations combine together to make the flow graph for any particular code. These have been given in Figure 2.

Figure-2: Types of Flow Graphs



Sources: Authors Compilation

McCabe's Cyclomatic Complexity

This technique is very important in Graph Theory and provides us with extremely useful software metric to know the number of independent paths. Cyclomatic Complexity is computed in one of the three ways:

i. The number of regions of the flow graph.

ii. Cyclomatic complexity V (G) for a flow graph
V (G) = E – N + 2
Where E = Number of flow graph edges
N - Number of flow graph nodes

iii. V (G) = P + 1
Where P = Number of predicted nodes contained in the flow graph G

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Figure-3: Shows the Program Code for a Procedure Done to Calculate Average

```
PROCEDURE average
t=1;
total.input=total.valid=0;                                  ... [1]
sum=0;
do while value[t] <> - 999 AND total.input<100              ... [2, 3]
increment total.input by 1;                                 ... [4]
IF value[t] >= minimum AND value[t] <= maximum              ... [5, 6]
        THEN increment total.valid by 1;
                SUM = ssum + value[t]                       ... [7]
        else skip
endif                                                       ... [8]
increment i by 1;
endoo                                                       ... [9]
IF total.valid > 0                                          ... [10]
THEN average = sum / total.valid;                           ... [11]
else average = -999;                                        ... [12]
endif                                                       ... [13]
END average
```

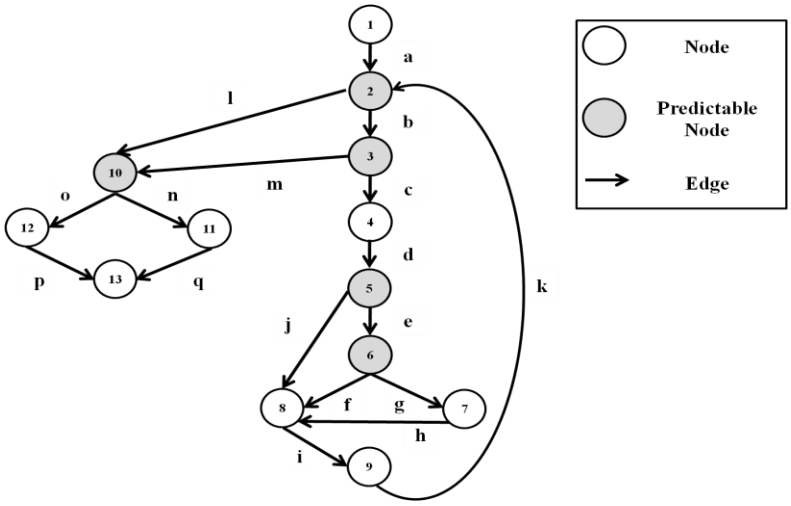
Sources: Authors Compilation

The control flow diagram has been shown in Figure 4.

Independent Paths for the Procedure Average are:

Path 1: 1 – 2 – 10 – 11 - 13
Path 2: 1 – 2 – 10 – 12 – 13
Path 3: 1 – 2 – 3 – 10 –12 –13
Path 4: 1 – 2 – 3 – 4 – 5 – 8 – 9 – 2...
Path 5: 1 – 2 – 3 – 4 – 5 – 6 – 8 – 9 – 2...
Path 6: 1 – 2 – 3 – 4 – 5 – 6 – 7 – 8 – 9 – 2...

Figure-4: Control Flow Graph for Procedure Average



Sources: Authors Compilation

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Using the 3 methods, we see that:

i. V(G) = No. of Regions = 6
ii. V(G) = 17 Edges – 13 Nodes + 2 = 6
iii. V(G) = 5 Predictable Nodes+ 1 = 6

Graph Matrix

A Graph Matrix is a square matrix whose size is equal to the number of nodes in a control flow graph. Each row and column corresponds to an identified node; and matrix entries correspond to connections between nodes. This technique is used to give total independent paths of the procedure with the intent that test case values should be such, which execute each path once. Once the Graph Matrix is prepared, a Connection Matrix is developed taking Graph Matrix as input. Table-1 and 2 show the Graph Matrix and Connection Matrix respectively. As can be seen from Table 2, the final tally adds up to 6.

Table-1: Graph Matrix

Node	1	2	3	4	5	6	7	8	9	10	11	12	13
1		a											
2			b							l			
3				c						m			
4					d								
5						e		j					
6							g	f					
7								h					
8									i				
9		k											
10											n	o	
11													q
12													p
13													

Sources: Authors Compilation

Table-2: Connection Matrix

Node	1	2	3	4	5	6	7	8	9	10	11	12	13	*Connections* Row Sum - 1
1		1												**1 - 1 = 0**
2			1							1				**1 + 1 - 1 = 1**
3				1						1				**1 + 1 - 1 = 1**
4					1									**1 - 1 = 0**
5						1		1						**1 + 1 - 1 = 1**
6							1	1						**1 + 1 - 1 = 1**
7								1						**1 - 1 = 0**
8									1					**1 - 1 = 0**
9		1												**1 - 1 = 0**
10											1	1		**1 + 1 - 1 = 1**
11													1	**1 - 1 = 0**
12													1	**1 - 1 = 0**
13														
Column Sum + 1														**6**

Sources: Authors Compilation

Techniques Under Black Box Testing

The two techniques used in this paper are Boundary Value Analysis and Worst Case testing.

Boundary Value Analysis

It is a common observation that many systems have a tendency to fail on boundaries. So testing boundary values of the application is very important. Boundary Value Analysis (BVA) is a Black Box testing technique where all the extreme boundary values are chosen and tested.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

The values used to test the boundaries are:

- Min : Minimal
- Min + : Just above Minimal
- Nom : Average
- Max - : Just below Maximum
- Max : Maximum

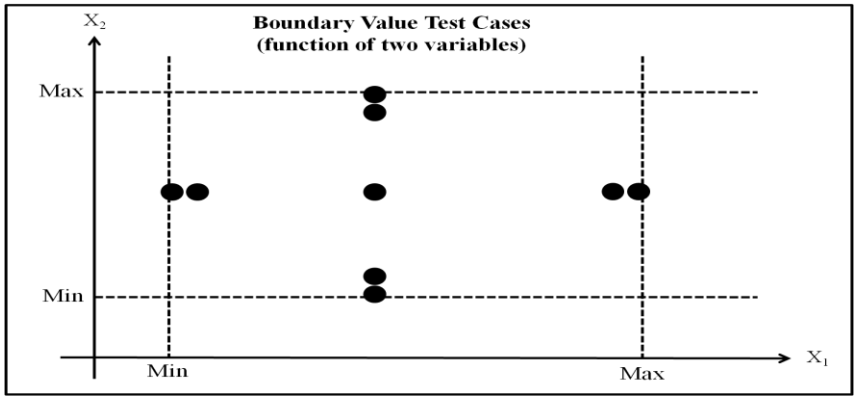
The number of test cases for boundary value analysis is **4n + 1** where **n** is the number of variables in a procedure.

$f = 4n + 1$

where *f* is function calculating node path test cases.

A visual representation showing the possible test cases is shown in Figure-5.

Figure-5: Boundary Value Test Cases



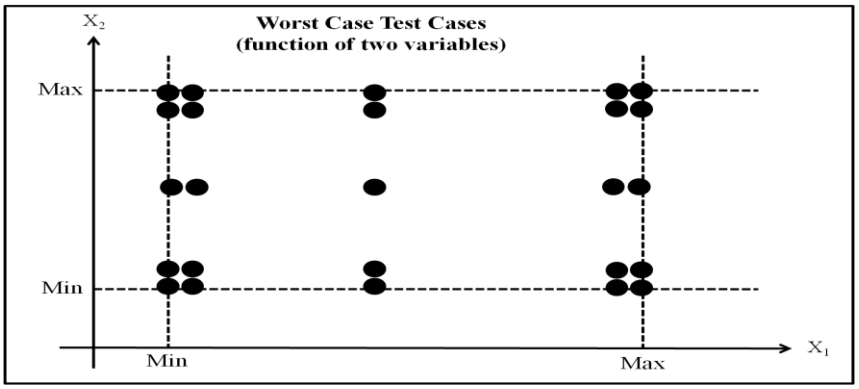
Sources: Authors Compilation

The Boundary Value Analysis test cases for the given problem are shown in Exhibit 1.

Worst Case Testing

Boundary Value analysis is based on the critical fault assumption and therefore it only tests for a single variable at a time assuming its extreme values. In Worst Case testing, we use a process to create test cases, which disregard the theory of critical fault assumption. The Worst Case testing checks for each combination of input value. To generate test cases we take the original 5-tuple set of Boundary Value analysis (min, min +, nom, max -, max) and perform the Cartesian product of these values. The end product generated is a much larger set of results than we have seen we get from Boundary Value analysis. A visual representation showing the possible test cases is shown in Figure 6.

Figure-6: Worst Case Scenario Test Cases



Sources: Authors Compilation

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

ANALYSIS

The data was received for 7 zones in Delhi. The data shows the average sale / month / outlet of 3 products – Godrej Ezee; Wipro Safewash and Godrej Genteel. The retailer quoted sales figures has been shown in Table-3.

Table-3: Sale Figures Data Collected From Survey

ZONE	Average Sales/Month/Outlet *EZEE*	Average Sales/Month/Outlet *SAFEWASH*	Average Sales/Month/Outlet *GENTEEL*
South Delhi	28	18	18
North Delhi	18	11	11
West Delhi	23	11	11
East Delhi	25	11	11
North West Delhi	23	11	11
Noida	23	11	11
Ghaziabad	28	11	18

Sources: Authors Compilation

Wipro intended to feed the numbers into the system such that when the 3 Brand sale figures are entered – the average should correspond to the category of market. The categories are given below:

- Average sale 0 to 10 : Bronze
- Average sale 11 to 20 : Silver
- Average sale > 20 : Gold

The values used to test the boundaries were specified as a minimum of 5 (minimum sale figure required for our data set to be accepted) and a maximum of 30 units sold:

- Min : 5
- Min + : 6
- Nom : 18
- Max - : 29
- Max : 30

The Boundary Value analysis test cases for the given problem are shown in Exhibit 1.

The Worst Case testing cases have been shown in Exhibit 2.

FUTURE SCOPE AND RECOMMENDATIONS

- The Black Box technique used for Wipro for can easily be utilized for a pan – India survey also. The software testing is naturally scalable.
- Black Box technique does not require prior technical knowledge on the part of the user. Therefore, this technique can be used by even smaller organizations with limited resources in terms of technical knowledge.
- The fact that in Black Box testing the programmer and tester are independent of each other means that the task of testing can be given to a separate department or even outsourced.
- Organizations need to decide among the various techniques based on the cost, ease of access of manpower and system (hardware / software) and the amount of importance attached to the code.
- Development of efficient research techniques is becoming an important research area in the future.
- An attempt should be made for creation of partnerships between industry and researchers, which will help, develop new techniques as well as help translate the same into industry practice.

REFERENCES

1. France, R., & Rumpe, B. ***Model-Driven Development of Complex Software: A Research Roadmap.*** In Future of Software Engineering (2007), L. Briand and A. Wolf, Editors. 2007, IEEE - CS Press.

2. Khan, M. E. (2010). Different Forms of Software Testing Techniques for Finding Errors. ***International Journal of Computer Science Issues***, 7(3), 11-16.

3. Alamitos. (1990). ***CA: IEEE Computer Society Press***.

4. Khan, M. E. (2011). Different Approaches to White Box Testing Techniques for Finding Errors. ***International Journal of Software Engineering and Its Applications***, 5 (3), 1-14

5. Kwang, Ik Seo, & Eun, Man Choi. Comparison of five Black-Box testing methods for Object-Oriented software. ***In Proceedings of the Fourth International Conference on Software Engineering Research, Management and Applications*** (SERA'06)

6. Lyu, M. R. (Ed.). (1996). ***Handbook of Software Reliability Engineering***. IEEE Comp. Soc. Press / McGraw Hill.

7. Lutz, R., & Mikulski, I. C. (2003). Requirements Discovery during the Testing Of Safety-Critical Software. ***In Proceedings of the Software Engineering 25th IEEE international conference,*** (pp. 578-583).

8. (2005). ***Software Engineering: A Practitioner's Approach, 6/e; Chapter 14: Software Testing Techniques.*** R.S. Pressman & Associates, Inc.

9. ***White Box Testing***. Retrieved on 23 December 2013, from http://agile.csc.ncsu.edu/SEMaterials/WhiteBox.pdf

10. Retrieved from http://thesai.org/Downloads/Volume3No6/Paper%203-A%20Comparative%20Study%20of%20White%20Box,%20Black.. .

11. Retrieved from http://www.servocity.com/html/header_row_pins.html

12. Retrieved from http://www.ijcsi.org/papers/7-3-1-11-16.pdf

13. Retrieved from http://www.pmdinfo.org.in/hostbeclients/pmdinfo.org.in/sqt.aspx

14. Retrieved from http://agile.csc.ncsu.edu/SEMaterials/IntrotoTesting.pdf

15. Retrieved from http://www.csie.ntut.edu.tw/labsdtl/99-1/ComparisonOfFiveBlack-boxTestingMethodsForObject-OrientedSo.. .

16. Retrieved from http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1691383&abstractAccess=no&userType=inst

17. Retrieved from http://www.softwaretestinghelp.com/black-box-testing

18. Retrieved from http://www.geekinterview.com/question_details/15463

19. Retrieved from http://www.testingexcellence.com/boundary-value-analysis

20. Retrieved from http://cs.swan.ac.uk/~csmarkus/CS339/dissertations/NeateB.pdf

21. Retrieved from http://www.testinganswers.com/2013/02/boundary-value-analysis-6.html

EXHIBITS

Exhibit-1: Boundary Value Analysis Test Cases

Case	Variable 1 *Ezee*	Variable 2 *Safewash*	Variable 3 *Genteel*	Expected Category of Market
1	18	18	5	Silver
2	18	18	6	Silver
3	18	18	18	Silver
4	18	18	29	Gold
5	18	18	30	Gold
6	18	5	18	Silver
7	18	6	18	Silver
8	18	29	18	Gold
9	18	30	18	Gold
10	5	18	18	Silver
11	6	18	18	Silver
12	29	18	18	Gold
13	30	18	18	Gold

Sources: Authors Compilation

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Exhibit-2: Worst Case Scenario Test Cases

Boundary Value Analysis test cases									
Case	**Variable 1** ***Ezee***	**Variable 2** ***Safewash***	**Variable 3** ***Genteel***	**Expected Category of Market**	**Case**	**Variable 1** ***Ezee***	**Variable 2** ***Safewash***	**Variable 3** ***Genteel***	**Expected Category of Market**
1	5	5	5	Bronze	**64**	18	18	29	Gold
2	5	5	6	Bronze	**65**	18	18	30	Gold
3	5	5	18	Bronze	**66**	18	29	5	Silver
4	5	5	29	Silver	**67**	18	29	6	Silver
5	5	5	30	Silver	**68**	18	29	18	Gold
6	5	6	5	Bronze	**69**	18	29	29	Gold
7	5	6	6	Bronze	**70**	18	29	30	Gold
8	5	6	18	Bronze	**71**	18	30	5	Silver
9	5	6	29	Silver	**72**	18	30	6	Silver
10	5	6	30	Silver	**73**	18	30	18	Gold
11	5	18	5	Bronze	**74**	18	30	29	Gold
12	5	18	6	Bronze	**75**	18	30	30	Gold
13	5	18	18	Silver	**76**	29	5	5	Silver
14	5	18	29	Silver	**77**	29	5	6	Silver
15	5	18	30	Silver	**78**	29	5	18	Silver
16	5	29	5	Silver	**79**	29	5	29	Gold
17	5	29	6	Silver	**80**	29	5	30	Gold
18	5	29	18	Silver	**81**	29	6	5	Silver
19	5	29	29	Gold	**82**	29	6	6	Silver
20	5	29	30	Gold	**83**	29	6	18	Silver
21	5	30	5	Silver	**84**	29	6	29	Gold
22	5	30	6	Silver	**85**	29	6	30	Gold
23	5	30	18	Silver	**86**	29	18	5	Silver
24	5	30	29	Gold	**87**	29	18	6	Silver
25	5	30	30	Gold	**88**	29	18	18	Gold
26	6	5	5	Bronze	**89**	29	18	29	Gold
27	6	5	6	Bronze	**90**	29	18	30	Gold
28	6	5	18	Bronze	**91**	29	29	5	Gold
29	6	5	29	Silver	**92**	29	29	6	Gold
30	6	5	30	Silver	**93**	29	29	18	Gold
31	6	6	5	Bronze	**94**	29	29	29	Gold
32	6	6	6	Bronze	**95**	29	29	30	Gold
33	6	6	18	Bronze	**96**	29	30	5	Gold
34	6	6	29	Silver	**97**	29	30	6	Gold
35	6	6	30	Silver	**98**	29	30	18	Gold
36	6	18	5	Bronze	**99**	29	30	29	Gold
37	6	18	6	Bronze	**100**	29	30	30	Gold
38	6	18	18	Silver	**101**	30	5	5	Bronze
39	6	18	29	Silver	**102**	30	5	6	Bronze
40	6	18	30	Silver	**103**	30	5	18	Bronze
41	6	29	5	Silver	**104**	30	5	29	Gold
42	6	29	6	Silver	**105**	30	5	30	Gold
43	6	29	18	Silver	**106**	30	6	5	Silver
44	6	29	29	Gold	**107**	30	6	6	Silver
45	6	29	30	Gold	**108**	30	6	18	Silver
46	6	30	5	Silver	**109**	30	6	29	Gold
47	6	30	6	Silver	**110**	30	6	30	Gold
48	6	30	18	Silver	**111**	30	18	5	Silver
49	6	30	29	Gold	**112**	30	18	6	Silver
50	6	30	30	Gold	**113**	30	18	18	Gold
51	18	5	5	Bronze	**114**	30	18	29	Gold
52	18	5	6	Bronze	**115**	30	18	30	Gold
53	18	5	18	Silver	**116**	30	29	5	Gold
54	18	5	29	Silver	**117**	30	29	6	Gold
55	18	5	30	Silver	**118**	30	29	18	Gold
56	18	6	5	Bronze	**119**	30	29	29	Gold
57	18	6	6	Bronze	**120**	30	29	30	Gold
58	18	6	18	Silver	**121**	30	30	5	Gold
59	18	6	29	Silver	**122**	30	30	6	Gold
60	18	6	30	Silver	**123**	30	30	18	Gold
61	18	18	5	Silver	**124**	30	30	29	Gold
62	18	18	6	Silver	**125**	30	30	30	Gold
63	18	18	18	Silver					

Sources: Authors Compilation

*****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

WEB BASED APPLICATIONS BY USING CLOUD COMPUTING

Sadhana Rana[19] Dr. M. K. Sharma[20]

ABSTRACT

Cloud computing is a new way of delivering computing resources, not a new technology. Computing services ranging from data storage and processing to software, such as email handling, are now available instantly, commitment-free and on-demand. Since we are in a time of belt-tightening, this new economic model for computing has found fertile ground and is seeing massive global investment.

Cloud Computing represents one of the most significant shifts in information technology many of us are likely to see in our lifetimes. Reaching the point where computing functions as a utility has great potential, promising innovations we cannot yet imagine. Customers are both excited and nervous at the prospects of Cloud Computing. They are excited by the opportunities to reduce capital costs. They are excited for a chance to divest them of infrastructure management, and focus on core competencies. Most of all, they are excited by the agility offered by the on-demand provisioning of computing and the ability to align information technology with business strategies and needs more readily. However, customers are also very concerned about the risks of Cloud Computing if not properly secured, and the loss of direct control over systems for which they are nonetheless accountable.

This paper explains, based on concrete scenarios, what cloud computing means for web applications, network and information security, data protection and privacy. We look at the security benefits of cloud computing and its risks. We cover the technical, policy and legal implications. Most importantly, we make concrete recommendations on how to address the risks and maximize the benefits.

KEYWORDS

RAT, Web Risk, Cloud Computing Analysis etc.

INTRODUCTION

Cloud computing can refer to several different service types, including Application/Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). The risks and benefits associated with each model will differ and so will the key considerations in contracting for this type of service. The following sections attempt to make the distinction when the risks or benefits apply differently to different cloud models.

Software as a service (SaaS): It is software offered by a third party provider, available on demand, usually via the Internet configurable remotely. Examples include online word processing and spreadsheet tools, CRM services and web content delivery services (Salesforce CRM, Google Docs, etc).

Platform as a service (PaaS): It allows customers to develop new applications using APIs deployed and configurable remotely. The platforms offered include development tools, configuration management, and deployment platforms. Examples are Microsoft Azure, Force and Google App engine.

Infrastructure as service (IaaS): It provides virtual machines and other abstracted hardware and operating systems, which may be controlled through a service API. Examples include Amazon EC2 and S3, Terremark Enterprise Cloud, Windows Live Skydrive and Rackspace Cloud.

TOP RECOMMENDATIONS

Assurance for Cloud Customers

Cloud customers need assurance that providers are following sound security practices in mitigating the risks facing both the customer and the provider (e.g., DDoS attacks). They need this in order to make sound business decisions and to maintain or obtain security certifications. An early symptom of this need for assurance is that many cloud providers are swamped with requests for audits. For this reason, we have expressed many of the report's recommendations as a standard list of questions, which can be used to provide or obtain assurance. Documents based on the checklist should provide a means for customers to:

[19]Research Scholar, Bhagwant University, Rajasthan, India, sadhanarana@gmail.com
[20]Head & Associate Professor, Computer Science Department, Uttarakhand, India, sharmamkhld@gmail.com

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

- Assess the risk of adopting cloud services;
- Compare different cloud provider offerings;
- Obtain assurance from selected cloud providers;
- Reduce the assurance burden on cloud providers.

The security checklist covers all aspects of security requirements including legal issues, physical security, policy issues and technical issues

LEGAL RECOMMENDATIONS

Most legal issues involved in cloud computing will currently be resolved during contract evaluation (i.e., when making comparisons between different providers) or negotiations. The more common case in cloud computing will be selecting between different contracts on offer in the market (contract evaluation) as opposed to contract negotiations. However, opportunities may exist for prospective customers of cloud services to choose providers whose contracts are negotiable

Unlike traditional Internet services, standard contract clauses may deserve additional review because of the nature of cloud computing. The parties to a contract should pay particular attention to their rights and obligations related to notifications of breaches in security, data transfers, creation of derivative works, change of control, and access to data by law enforcement entities. Because the cloud can be used to outsource critical internal infrastructure, and the interruption of that infrastructure may have wide-ranging effects, the parties should carefully consider whether standard limitations on liability adequately represent allocations of liability, given the parties' use of the cloud, or responsibilities for infrastructure.

SECURITY BENEFITS

Security and Benefits of Scale

Put simply, all kinds of security measures are cheaper when implemented on a larger scale. Therefore, the same amount of investment in security buys better protection. This includes all kinds of defensive measures such as filtering; patch management, hardening of virtual machine instances and hypervisors, etc. Other benefits of scale include multiple locations, edge networks (content delivered or processed closer to its destination), timeliness of response, to incidents, threat management.

Security as a Market Differentiator

Security is a priority concern for many cloud customers; many of them will make buying choices based on the reputation for confidentiality, integrity and resilience of, and the security services offered by, a provider. This is a strong driver for cloud providers to improve security practices.

STANDARDISED INTERFACES FOR MANAGED

Security Services

Large cloud providers can offer a standardized, open interface to managed security services providers. This creates a more open and readily available market for security services.

Rapid, Smart Scaling of Resources

The ability of the cloud provider to dynamically reallocate resources for filtering, traffic shaping, authentication, encryption, etc., to defensive measures (e.g., against DDoS attacks) has obvious advantages for resilience.

Audit and Evidence-Gathering

Cloud computing (when using virtualization) can provide dedicated, pay-per-use forensic images of virtual machines which are accessible without taking infrastructure off-line, leading to less down-time for forensic analysis. It can also provide more cost-effective storage for logs allowing more comprehensive logging without compromising performance.

More Timely, Effective and Efficient Updates and Defaults

Default virtual machine images and software modules used by customers can be pre-hardened and updated with the latest patches and security settings according to fine-tuned processes; IaaS cloud service APIs also allow snapshots of virtual infrastructure to be taken regularly and compared with a baseline. Updates can be rolled out many times more rapidly across a homogenous platform than in traditional client-based systems that rely on the patching model.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Benefits of Resource Concentration

Although the concentration of resources undoubtedly has disadvantages for security [see Risks], it has the obvious advantage of cheaper physical parameterization and physical access control (per unit resource) and the easier and cheaper application of many security-related processes.

SECURITY RISKS

The following points should be noted in relation to the descriptions of risk below:

- Risk should always be understood in relation to overall business opportunity and appetite for risk – sometimes risk is compensated by opportunity.
- Cloud services are not only about convenient storage, accessible by multiple devices, but include important benefits such as more convenient communication and instant multi-point collaboration. Therefore, a comparative analysis needs to compare not only the risks of storing data in different places (on premises v the cloud) but also the risks when on premises-data stored on premises - e.g. a spreadsheet - is emailed to other persons for their contributions, against the security issues of a spreadsheet stored in the cloud and open to collaboration between those persons. Therefore, the risks of using cloud computing should be compared to the risks of staying with traditional solutions, such as desktop-based models.
- The level of risk will in many cases vary significantly with the type of cloud architecture being considered.
- It is possible for the cloud customer to transfer risk to the cloud provider and the risks should be considered against the cost benefit received from the services. However *not all risks can be transferred*: if a risk leads to the failure of a business, serious damage to reputation or legal implications, it is hard or impossible for any other party to compensate for this damage.
- The risk analysis in this paper applies to cloud technology. It does not apply to any specific cloud computing offering or company. This paper is not meant to replace a project-specific organizational risk assessment.
- The level of risks is expressed from the perspective of the cloud customer. Where the cloud provider point of view is considered, this is explicitly stated.

The most important classes of cloud-specific risks are:

Data Protection

Cloud computing poses several data protection risks for cloud customers and providers. In some cases, it may be difficult for the cloud customer (in its role as data controller) to effectively check the data handling practices of the cloud provider and thus to be sure that the data is handled in a lawful way. This problem is exacerbated in cases of multiple transfers of data, e.g., between federated clouds. On the other hand, some cloud providers do provide information on their data handling practices. Some also offer certification summaries on their data processing and data security activities and the data controls they have in place, e.g., SAS70 certification.

Isolation Failure

Multi-tenancy and shared resources are defining characteristics of cloud computing. This risk category covers the failure of mechanisms separating storage, memory, routing and even reputation between different tenants (e.g., so-called guest-hopping attacks). However it should be considered that attacks on resource isolation mechanisms (e.g., against hypervisors) are still less numerous and much more difficult for an attacker to put in practice compared to attacks on traditional OSs.

Malicious Insider

While usually less likely, the damage which may be caused by malicious insiders is often far greater. Cloud architectures necessitate certain roles, which are extremely high-risk. Examples include CP system administrators and managed security service provider

Insecure or Incomplete Data Deletion

When a request to delete a cloud resource is made, as with most operating systems, this may not result in true wiping of the data. Adequate or timely data deletion may also be impossible (or undesirable from a customer perspective), either because extra copies

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

of data are stored but are not available, or because the disk to be destroyed also stores data from other clients. In the case of multiple tenancies and the reuse of hardware resources, this represents a higher risk to the customer than with dedicated hardware.

THREATS

Our goal is to provide a threat identification deliverable that can be quickly updated to reflect the dynamics of Cloud Computing and its rapidly evolving threat environment. We look forward to your participation on subsequent versions of "Top Threats to Cloud Computing".

Threat-1: Abuse and Nefarious Use of Cloud Computing

IaaS providers offer their customers the illusion of unlimited compute, network, and storage capacity often coupled with a 'frictionless' registration process where anyone with a valid credit card can register and immediately begin using cloud services. Some providers even offer free limited trial periods. By abusing the relative anonymity behind these registration and usage models, spammers, malicious code authors, and other criminals have been able to conduct their activities with relative impunity. PaaS providers have traditionally suffered most from this kind of attacks; however, recent evidence shows that hackers have begun to target IaaS vendors as well. Future areas of concern include password and key cracking, DDOS, launching dynamic attack points, hosting malicious data, botnet command and control, building rainbow tables, and CAPTCHA solving farms.

Examples: IaaS offerings have hosted the Zeus botnet, InfoStealer trojan horses, and downloads for Microsoft Office and Adobe PDF exploits. Additionally, botnets have used IaaS servers for command and control functions. Spam continues to be a problem — as a defensive measure, entire blocks of IaaS network addresses have been publicly blacklist.

Threat-2: Insecure Interfaces and APIs

Cloud computing providers expose a set of software interfaces or APIs that customers use to manage and interact with cloud services. Provisioning, management, orchestration, and monitoring are all performed using these interfaces. The security and availability of general cloud services is dependent upon the security of these basic APIs. From authentication and access control to encryption and activity monitoring, these interfaces must be designed to protect against both accidental and malicious attempts to circumvent policy. Furthermore, organizations and third parties often build upon these interfaces to offer value-added services to their customers. This introduces the complexity of the new-layered API; it also increases risk, as organizations may be required to relinquish their credentials to third parties in order to enable their agency.

Examples: Anonymous access and/or reusable tokens or passwords, clear-text authentication or transmission of content, inflexible access controls or improper authorizations, limited monitoring and logging capabilities, unknown service or API dependencies.

Threat-3: Malicious Insiders

The threat of a malicious insider is well known to most organizations. This threat is amplified for consumers of cloud services by the convergence of IT services and customers under a single management domain, combined with a general lack of transparency into provider process and procedure. For example, a provider may not reveal how it grants employees access to physical and virtual assets, how it monitors these employees, or how it analyzes and reports on policy compliance. To complicate matters, there is often little or no visibility into the hiring standards and practices for cloud employees. This kind of situation clearly creates an attractive opportunity for an adversary ranging from the hobbyist hacker, to organized crime, to corporate espionage, or even nation-state sponsored intrusion. The level of access granted could enable such an adversary to harvest confidential data or gain complete control over the cloud services with little or no risk of detection.

Threat- 4: Shared Technology Issues

IaaS vendors deliver their services in a scalable way by sharing infrastructure. Often, the underlying components that make up this infrastructure (*e.g.*, CPU caches, GPUs, etc.) were not designed to offer strong isolation properties for a multi-tenant architecture. To address this gap, a virtualization hypervisor mediates access between guest operating systems and the physical compute resources. Still, even hypervisors have exhibited flaws that have enabled guest operating systems to gain inappropriate levels of control or influence on the underlying platform. A defense in depth strategy is recommended, and should include compute, storage, and network security enforcement and monitoring. Strong compartmentalization should be employed to ensure that individual customers do not affect the operations of other tenants running on the same cloud provider. Customers should not have access to any other tenant's actual or residual data, network traffic, etc.

Threat-5: Data Loss or Leakage

There are many ways to compromise data. Deletion or alteration of records without a backup of the original content is an obvious example. Unlinking a record from a larger context may render it unrecoverable, as can storage on unreliable media. Loss of an

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

encoding key may result in effective destruction. Finally, unauthorized parties must be prevented from gaining access to sensitive data. The threat of data compromise increases in the cloud, due to the number of and interactions between risks and challenges, which are either unique to cloud, or more dangerous because of the architectural or operational characteristics of the cloud environment.

Threat-6: Account or Service Hijacking

Account or service hijacking is not new. Attack methods such as phishing, fraud, and exploitation of software vulnerabilities still achieve results. Credentials and passwords are often reused, which amplifies the impact of such attacks. Cloud solutions add a new threat to the landscape. If an attacker gains access to your credentials, they can eavesdrop on your activities and transactions, manipulate data, return falsified information, and redirect your clients to illegitimate sites. Your account or service instances may become a new base for the attacker. From here, they may leverage the power of your reputation to launch subsequent attacks.

Threat-7: Unknown Risk Profile

One of the tenets of Cloud Computing is the reduction of hardware and software ownership and maintenance to allow companies to focus on their core business strengths. This has clear financial and operational benefits, which must be weighed carefully against the contradictory security concerns complicated by the fact that cloud deployments are driven by anticipated benefits, by groups who may lose track of the security ramifications. Versions of software, code updates, security practices, vulnerability profiles, intrusion attempts, and security design, are all important factors for estimating your company's security posture. Information about who is sharing your infrastructure may be pertinent, in addition to network intrusion logs, redirection attempts and/or successes, and other logs. Security by obscurity may be low effort, but it can result in unknown exposures. It may also impair the in-depth analysis required highly controlled or regulated operational areas.

THE BENEFITS OF CLOUD COMPUTING

As cloud computing begins to take hold, several major benefits have become evident:

Costs

The cloud promises to reduce the cost of acquiring, delivering, and maintaining computing power, a benefit of particular importance in times of fiscal uncertainty. By enabling agencies to purchase only the computing services needed, instead of investing in complex and expensive IT infrastructures, agencies can drive down the costs of developing, testing, and maintaining new and existing systems.

Access

The cloud promises universal access to high-powered computing and storage resources for anyone with a network access device. By providing such capabilities, cloud computing helps to facilitate telework initiatives, as well as bolster an agency's continuity of operations (COOP) demands.

Scalability and Capacity

The cloud is an always-on computing resource that enables users to tailor consumption to their specific needs. Infinitely scalable, cloud computing allows IT infrastructures to be expanded efficiently and expediently without the necessity of making major capital investments. Capacity can be added as resources are needed and completed in a very short period. Thus, agencies can avoid the latency, expense, and risk of purchasing hardware and software that takes up data center space and can reduce the traditional time required to scale up an application in support of the mission. Cloud computing allows agencies to easily move in the other direction as well, removing capacity, and thus expenses, as needed.

Resource Maximization

Cloud computing eases the burden on IT resources already stretched thin, particularly important for agencies facing shortages of qualified IT professionals. The cloud presents an environment where users can develop software-based services that enhances collaboration and fosters greater information sharing, not only within the agency, but also among other government and private entities.

Customization

Cloud computing offers a platform of tremendous potential for creating and amending applications to address a diversity of tasks and challenges. Its inherent agility means that specific processes can be easily altered to meet shifting agency needs, since those processes are typically changeable by making a configuration change, and not by driving redevelopment from the back-end systems (Heyward and Rayport, 2009).

CONCLUSIONS

The conclusion of this paper is that the cloud's economies of scale and flexibility are both a friend and a foe from a security point of view. The massive concentrations of resources and data present a more attractive target to attackers, but cloud-based defenses can be more robust, scalable and cost-effective. This paper allows an informed assessment of the security risks and benefits of using cloud computing - providing security guidance for potential and existing users of cloud computing**.**

REFERENCES

1. Retrieved from http://www.cloudsecurityalliance.org/topthreats
2. Retrieved from http://www.malwaredomainlist.com
3. Retrieved from http://www.coe.int/t/dghl/cooperation/economiccrime/cybercrime/cy-activity-interface-2010/presentations/Outlook/Udo%20Helmbrecht_ENISA_Cloud%20Computing_Outlook.pdf
4. Retrieved from http://blogs.zdnet.com/security/?p=5110
5. Retrieved from http://voices.washingtonpost.com/securityfix/2008/07/amazon_
6. Retrieved from http://www.programmableweb.com
7. Retrieved from http://securitylabs.websense.com/content/Blogs/3402.aspx
8. Danielson, Krissi. (2008, March 26). ***"Distinguishing Cloud Computing from Utility Computing.*** Retrieved 2010, August 22 from http://www.ebizq.net/blogs/saasweek/2008/03/distinguishing_cloud_computing/
9. Gruman, Galen, (2008, April 07). ***What Cloud Computing Really Means*** Retrieved 2009-06-02 from http://www.infoworld.com/d/cloud-computing/what-cloud-computing-really-means-031
10. (2009, October 15). Cloud Computing: Clash of the Clouds. ***The Economist.*** Retrieved on 2009-11-03 from http://www.economist.com/displaystory.cfm?story_id=14637206
11. Buyya, Rajkumar, Chee, Shin Yeo, & Srikumar Venugopal. (PDF). ***Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities***, pp. 9. Australia. University of Melbourne: Department of Computer Science and Software Engineering.
12. ""The Rise of Cloud Computing." Michael Otey. April 2010". windowsITpro.com. 2010-04-26. Retrieved 2010-08-22.
13. ***Google Apps is More Valuable than YouTube and Gmail Combined.*** Retrieved from http://www.windowsitpro.com/article/cloud-computing2/The-Rise-of-Cloud-Computing/2.aspx
14. Retrieved from http://www.slideshare.net/piyushmittalin/top-cloud-threats
15. Retrieved from http://www.coe.int/t/dghl/cooperation/economiccrime/cybercrime/cy-activity-Interface-2010/Presentati..
16. Retrieved from http://www.enisa.europa.eu/media/news-items/cloud-computing-speech
17. Retrieved from http://usingtheclouds.wordpress.com/2011/02/01/where-is-the-real-value-proposition-in-cloud-computin.. .
18. Retrieved from http://www.illuminata.com/?p=3414
19. Retrieved from http://docs.askives.com/4-security-risks-in-using-cloud-computing-as-per-the-enisa-guide.html
20. Retrieved from https://cloudsecurityalliance.org/topthreats/csathreats.v1.0.pdf
21. Retrieved from http://book.itep.ru/depository/cloud/csathreats.v1.0.pdf
22. Retrieved from http://www.cibtech.org/J%20ENGINEERING%20TECHNOLOGY/PUBLICATIONS/2014/Vol-4-No-1/JET-11-023-MANISH-A. ..

23. Retrieved from http://www.ijser.org/researchpaper/A-Survey-on-Security-Threats-and-Vulnerabilities-In-Cloud-Computi.. .
24. Retrieved from http://unicomgov.com/files/2713/7666/8325/cloud-Computing.pdf
25. Retrieved from http://www.ijmer.com/papers/vol2_issue2/BM22380382.pdf
26. Retrieved from http://www.ijceronline.com/papers/Vol3_issue3/G033030035.pdf
27. Retrieved from http://virusremovalexpert.wordpress.com/2010/11/24/dangerous-clouds
28. Retrieved from http://blog.protogenist.com/?p=451
29. Retrieved from http://blog.protogenist.com/?tag=iaas
30. Retrieved from http://www.sunzu.com/articles/abuse-and-nefarious-use-of-cloud-computing-174998
31. Retrieved from http://ttajts0.tripod.com/cloud/apis.htm
32. Retrieved from http://protogenist.wordpress.com/2012/10/10/threat-malicious-insiders
33. Retrieved from http://ttajts0.tripod.com/cloud/dataloss.htm
34. Retrieved from http://ttajts0.tripod.com/cloud/hijacking.htm
35. Retrieved from http://www.ebizq.net/blogs/cloudsoa/2010/09/will-cost-savings-continue-to-be-a-significant-driver-fo..
36. Retrieved from http://www.ijarcsse.com/docs/papers/Volume_3/4_April2013/V3I4-0358.pdf
37. Retrieved from http://en.wikipedia.org/wiki/Cloud_computing
38. Retrieved from http://smart-cloud-computing.blogspot.in/2011/01/benefits-of-cloud-computing.html
39. Retrieved from http://virusremoval-expert.blogspot.in/2010/11/dangerous-clouds_23.html
40. Retrieved from http://itlaw.wikia.com/wiki/Cloud_Computing:_Benefits,_Risks,_and_Recommendations_for_Information_Se.. .

*****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

SEARCH ENGINE OPTIMIZATION IN DESIGN & DEVELOPMENT OF WEB 2.0 SITES

Nitin Deepak[21] Gunjan Mathur[22]

ABSTRACT

Web 2.0, the second phase in the Web's evolution, is drawing an attention of IT professionals, businesses, and users of Web. Whenever there is a need to design and develop the website, it is somehow necessary to present web pages in such a way so that they can be searched easily and interactive also. SEO (Search Engine Optimization) is the first step to think before designing or developing any website (Web 2.0 site). In this paper, we are trying to present a need of the concept of SEO and trying to put forward design methodologies in website optimization strategies for web-development and design.

KEYWORDS

Web 2.0, Search Engine Optimization, Keywords Optimization, Ajax, SEO Strategy etc.

Web 2.0

Web 2.0, the II phase in the Web's evolution, is drawing an attention of IT professionals, businesses, and users of Web. Instead of only informative or dynamic web, the Web 2.0 is also called the Intelligent Web, community-centric Web, participative Web, and read/write Web [9].

WEB 2.0 TECHNOLOGIES AND SERVICES

Blogs

A blog, precision of Web log, is a powerful Web-based communication tool. A blog is a Web site or application where people can enter their thoughts, ideas, suggestions, and comments. Blog entries, also known as *blog posts*, are made in journal style and are usually displayed in reverse chronological order.

Really Simple Syndication

RSS is a family of Web feed formats used for associating the content from blogs or Web pages. RSS is an XML file that summarizes information items and links to the information sources. It informs users of updates to blogs or Web sites they are interested in. Web or blog RSS feeds are typically linked with the word "subscribe," an orange rectangle, or with the letters XML or RSS in an orange box.

Wikis

A wiki is a simple yet powerful Web-based collaborative-authoring (or content-management) system for creating and editing content. It lets anyone add a new article or revise an existing article through a Web browser. Users can also track changes made to an article. The term wiki is derived from the Hawaiian word wikiwiki, which means fast or quick. The user-generated online encyclopedia Wikipedia (http://en.wikipedia.org) is a wiki.

Mashups

A Web mashup is a Web page or Web site that combines information and services from multiple sources on the Web. Similar to music mashups, where artists combine, for example, vocals from one song with the music from another, Web mashups combine information and/or complementary functionality from multiple Web sites or Web applications. Mashups can be grouped into seven categories: mapping, search, mobile, messaging, sports, shopping, and movies.

Tags

Tags are those keywords that are added to articles displayed in blogs or Web pages via social page tag tools such as del.icio.us and Yahoo's My Web. Most blogs and Web publications use tags. Tags are also known as labels, and the process of creating tags is known as tagging.

[21]Research Scholar, Bhagwant University, Rajasthan, India, nitin.d12@gmail.com
[22]Lecturer, Amrapali Institute of Management & Computer Applications, Uttarakhand, India, gunjan.m19@gmail.com

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Folksonomy

Folksonomy refers to user-created taxonomies of information. It is an ad hoc classification scheme that Web users create as they surf the Web to categorize the content they find online.

Social Bookmarking

Users bookmark interesting pages and assign tags to each by the process. It is a great way of capturing contextual knowledge.

DEVELOPMENT APPROACHES

Web 2.0 is a sun shelter term encompassing several new Web technologies, which is discussed in the following sections:

I. Ajax

AJAX is a relatively new approach to creating Web applications. It enriches the user interface, making it highly interactive and more responsive. It is several technologies coming together in powerful new ways—XHTML or HTML, cascading style sheets (CSS), JavaScript, and XML.

II. Adobe Flex

Adobe Flex (http://www.adobe.com/products/flex) is an application development solution for creating and delivering cross-platform rich Internet applications (RIAs) on the Web. Flex is based on Flash and provides a standards-based language and programming model that supports common design patterns. It provides a more productive Eclipse-based development environment; dramatically improves application performance; supports new classes of applications, such as those requiring real-time data push; and provides more fine-grained control over an application's look and feel. Flex and Flash have complementary strengths. While Flash helps users create rich interactive content, Flex leverages the development of data-driven RIAs.

WHAT IS SEO?

SEO (Search Engine Optimization) is the process of following the principle of the search engine such as site structure, webpage language and interaction diplomatic strategies for the rational planning to improve the site search performance in the search engine, and increasing the possibility of customer discovery and access to the website. SEO is a scientific development concept and methodology, which develops along with the development of search engine, and promotes the development of search engines at the same time [1].

Importance of applying SEO in web 2.0 site

I. Why do we need SEO?

The formation of search habits: users have become accustomed to use search engines to search for valuable information on Internet, because there will be no type of website meet needs for all the information in addition to search engines.

Most of users focus on the first page: as hundreds or thousands of results will appear after the key words are given, users prefer to pay more attention to the first several data of the first page. When users click a few pages and could not find the desired results, they generally give up the existing search results, and re-enter additional keywords. Therefore, the more on top of the search results, the more possibility to be clicked [2]

The self-value of SEO: relevant information shows that, the majority of companies introduce SEO techniques to their website. SEO can make the website have a higher rank, get a higher click-through rate and bring greater value ultimately.

II. The Problems Web2.0 Site Facing

Now people become more and more familiar with web2.0. It is said that some even regarded Obama won the President of the United States probably because of his well use of web2.0. Abroad popular web2.0 sites are Myspace, YouTube, Facebook, and OpenSocial and so on, while domestic are osits.in, amrapali.ac.in, nainitalrudrapur.com and so on. We take a popular web2.0 website www.amrapaliinstitute.ac.in as an example.

Amrapali has been established since 1999 and then develop quickly to become famous in short period. The user who has experienced the Amrapali knows it has such a number of advantages like:

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

- Simple and clear pages,
- User friendly environment, which may be easy to use
- New features: User feedbacks and, user recommendations.
- It is necessary to optimize the web2.0 site using SEO because of its importance in the internet and the problems of web2.0 site facing currently.

PROCESS OF DESIGNING WEB 2.0 SITES FOR SEO

Designing Web2.0 site for SEO is not meet the search engine forwardly, but by adhering to the use of web standards to ensure webpage quality can be recognized by search engines and make search engine more friendly. Each search engine's rank algorithms are different, but it is certain that making high quality, structure, clear, well-known site to be top is any of the rank algorithm goal. Therefore, adhering to webpage design of web standards is the basis of SEO [2].

I. Design by Navigation

Navigation plays a very vital role in navigating a web page design. There are many different styles for navigation. There are three basic styles:

Text Navigation

Text is a very good navigation for SEO website, which can be easily design and develop via CSS, DIV and Scripting. DHTML is the fastest medium is reducing the page size greatly; improving the speed of web page while opening in client's machine, which is truly more important and identified by the spider easily

Picture Navigation

Images and Flash animation are the ways through which we can visually present web site information, which is used by many enterprises. However, this type of navigation is not supported by SEO, as the spider such as text cannot identify it. One has to make separate encoding *(Alternative text to images and parameters to Flash animation coding)* to make the spider to recognize the content.

Text Navigation with Picture Backgrounds

It is a good choice as it is recognized by spider because of the existing text and interactive as it has impressive background to communicate visually.

After analyzing all three categories, I recommend to **third** strategy, which will definitely help to improve the web site performance and SEO.

II. Not to use Frame Structure

Many search engines never search the inner content stores in frames as the frame makes the criteria of HTML DOM standards very difficult while searching. Frame structure is therefore not recommended [2].

III. Optimizing Pictures

In using pictures, there are two ways to optimize the same: *firstly;* select proper compression technique and *secondly;* embed the text as a piece of information to make the picture searchable.

i. Gif, jpeg and png are the suitable formats for web pages. Gif and png format is usually used to include low quality or single colored required pictures and to use transparent background; jpeg is used for high quality pictures. Jpeg format loads row by row while gif format displays picture structure first and then clear details, so picture buttons or web structure pictures considering user experiences and SEO commonly use gif format.

ii. Alt and title properties: alt property used to designate the instead text when picture missed and title property used to explain the current object are special for the picture search engine which cannot identify the picture contents themselves, so they are essential for the pictures. Alt property can set the meaning of the picture when title property set Indicative information during the SEO optimization. Therefore, when picture cannot display correctly, alt property helps show instead text or when a picture button link is pointed by mouse its title property will help say "back to the main web page" as be set.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

IV. CSS Optimization

Optimization of CSS mainly including the following aspects:

i. Tag property optimization: when we open dream weaver, click modify - page properties, set the web margin of left, right, top, bottom 0px, then look for the code. It says left Margin=0 top Margin=0. Shortly we can directly write margin: 0px instead. This is one simple example of the CSS optimization.
ii. Grouping selector: we actually make many reduplicate styles, such as width property, height property etc., when we define the style. For example: h1 {color: blue; font-size: 16px; font-weight: bold;} h2 {color: blue; font-size: 16px; font-weight: bold;} can be shortly written as h1,h2 {color: blue; font-size:16px; font-weight: bold;}
iii. Class selector: Class selector can be used repeatedly in html codes once defined in style. So only basic styles are suggested to define in id selector while the other styles are expected to define in class selector when design the webpage layout using CSS. This is one of CSS optimization methods [3].

V. XHTML Structural Optimization

Xhtml structure uses DIV+CSS style, which makes a concise webpage codes and fits well with the internet spider. Its concise codes ensure the spider to go through the whole page easily using least time, thus this structure is more inclined to be accessed by spider. It is also true that most websites that have a high rank in the search engine are designed by xhtml structure.

SEO STRATEGY

Some of the websites rank high on search engine and some can't make place at good rank on search engine, why? How to make website has a high rank? For this, we have to understand the basic optimization strategy:

Page Title

Owing to its greatest exposure in the page of search engine results, the page title is the highest priority in the search engines, which is often taken as the first line of search results. Concerning the website click rate, the page title should be concise, eye-catching, and in order to achieve the best transformation, the page title should be accurately summed up the contents of the page.

There are several attentions when setting the page title:

i. It's better to make a brief introduction for the contents of the current page, and in order to allow the news expressed as clearly as possible, we should put important keywords on the beginning of the title.
ii. Target keywords must be included in the html page title, which is a very important factor in the rank algorithm. However, excessive repetition of keywords in the title is considered as a bad strategy.
iii. Set different title for each page from the respective visual angle.
iv. Do not repeat website navigation information in the title [4].

Keywords Optimization

Appropriate keywords are the first step to build a high-rank website. Keywords density is the ratio between keywords number and webpage words number, which are the most important factors in optimization Strategy. In order to make website rank top in search engines, webpage should contain keywords as much as possible, but once keywords are excessive, it may face the risk of registration cancellation. The prettiest density can be confirmed with the help of first several highest rank websites in search results, meanwhile keywords analysis tools should be used to analyze the position of keywords in search results for adjustment in time [4].

Improve the click popularity

Click popularity is a very important factor for website rank in search engines. The popularity of the page often Clicked will be high. When a visitor clicks website from the search results, search engine will give some certain value. However, do not try to repeat clicking on your own website, because the clicks from a single IP will be calculated only once [5].

Improve link popularity

Page Rank (PR) value is significant for Google to judge the importance of website. If the linked page has a higher PR value, then the page has a higher PR value. Raising the value of PR can effectively enhance the website rank in search engine. Search engine evaluate PR value through external website links quantities and the quality of links to Web sites, that is, it needs to improve the extensive degree of site's Link. This objective can be achieved by exchanging links (between sites by adding links to each other). However, it does not mean that can be to exchange links with any website arbitrarily. It is better to choose the PR value, which is more than four [6].

Website Internal Links

Internal links of website refer to the one another links between the separate pages of website. Reasonable links can make the web pages more visual, help visitors focus on the theme of website and make the theme key words more sensitive to the search engine. In addition, internal links of website can also accelerate the rate of the spider crawling [8].

CONCLUSIONS

Web 2.0 is actually the need for the time of intelligence, which needs communicative websites and web-applications instead of only informative and static sites. As studied for the problems web2.0 sites facing, this paper proposes methodologies for SEO and introduces the basic SEO optimizing strategies, which will help in optimizing websites. As the technology changes day by day in search engines there is always the keen need to research on more and better optimize ways to make one's website more effective and searchable[9].

REFERENCES

1. ***What is SEO?*** Retrieved from http://www.dunsh.org/2006/07/24/seo-definition .
2. Jian, Wang. (2007). ***Be Proficient In the Standards of Building Websites.*** BeiJing: Posts & Telecom Press.
3. ***For SEO, Standards-Based Web Site Design WEB***. Retrieved from http://www.eastsem.com/seo/20081118204516.html
4. Jennifer, Grappone, & Gradiva, Couzin. (2007, July). ***Search Engine Optimization***, pp. 100-147. BeiJing: Tsinghua University Press.
5. ***How to optimize Website?***. Retrieved from http://blog.sina.com.cn/s/blog_4cedb9ba0100085g.html .
6. Dan, Yi. (2006). The Exploration and Realization of SEO. ***Computer Applications of Petroleum,*** 14(1), 35-37.
7. ***Check Page Rank of any web site pages***. Retrieved from http://www.prchecker.info/check_page_rank.php
8. Siyun, Tan, Jiachao, Zhu, Qianhua, Zhou, & Wei, Xiong. (2008, February). Research of SEO about Website. ***China Water Transport***, 8, 139-140.
9. San, Murugesan. (2007). Understanding Web 2.0. ***IEEE Computer Society.***
10. Retrieved from http://www.computer.org/portal/web/buildyourcareer/fa009
11. Retrieved from http://91-592-722.wiki.uml.edu/file/view/understanding_web_20.pdf

*****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

INNOVATIVE IVR SYSTEM FOR FARMERS: ENHANCING ICT ADOPTION

Abhishek Behl[23] Maneesha Vashishtha[24]

ABSTRACT

India is a country with more than half the population depending on agriculture for their livelihood but manages only 14 % contribution in its GDP. The country has fragmented land holdings, which has given rise to considerable amount of independent farmers with a near stagnant productivity. In spite of initiatives by Government of India both at the center and at state levels, there continues a gap between land and lab. Karnataka is no exception with more than 80 % of total land holdings under small and marginal farmers. The Government of Karnataka, with the aim of providing this sect of farmers the latest crop related and information using ICT tools, has launched a service to relay agricultural advisories by means of a 'PUSH' voice-message to farmers in five delta districts. These services will be used to record feedback and assist farmers on various fronts including weather updates, agriculture equipment's etc.

This research paper enlightens ways and methods of delivery of services and improves its efficiency of prevalent services. ICT plays a transformative role and potential, which is supported by both literature of field experience. It empowers farmers to convey proper and real time feedback in accordance with proper guidance on their agricultural needs. The results indicate that feedback from customers should be part of delivery cycle in order to enable better and efficient delivery of services to desired audience.

KEYWORDS

IVR System, ICT, Farmers, Delivery Cycle etc.

INTRODUCTION

The contribution of agriculture to the overall Gross Domestic Product (GDP) of India has been falling rapidly and has gone down from 30 percent in 1990-91 to a mere 13.9 per cent in 2012 (State of Indian Agriculture, 2011-2012). Yet, the agriculture sector remains the backbone of our country as it employs 52% of its workforce. Indian farmers, particularly, small and marginal ones, currently face a variety of challenges – fragmentation of land into small holdings, limited land and water availability, effects of climate change that include but are not limited to new pest and disease outbreaks, new demand and consumption patterns, liberalization of trade as well as a move towards high-value agriculture (Claire J. Glendenning, 2010). Lack of quality information at multiple levels has prevented many of these farmers from increasing their productivity and profitability. Information gaps exist in terms of what and where to buy much-needed inputs, how does one overcome stunted growth and disease outbreak, when and where to sell the produce as well as how and where to get information about financial services, transport and local weather. Karnataka has a major portion of small and marginal farmers accounting for more than 90% of the 8 million total land holdings (Department of Evaluation and Applied Research, 2008-2009). Experts agree that there is a strong need today to align the information that is being delivered through various extension approaches with the needs of farmers, their existing experience and knowledge base, by personalizing the information that is being delivered to them. Considering that there are around 88 million farmers in India, with 98.5 million holdings (operating an average area of 1.1 hectare) (National Sample Survey Office - Survey Finding - 59th Round), speaking some 20 different languages and many more dialects, dispersed geographically in 630,000 villages, it is not surprising that extension approaches fail to do so.

In such a scenario, it was conceived that Information and Communication Technology (ICT) could play a significant role in overcoming the current shortcomings in Indian Agricultural Advisory Systems. (Jhunjhunwala, Umadikar, Prashant, & Canagarajah, 2013) (Chhibber, 2004)

ACCESS TO ICT: THE MOBILE REVOLUTION

Before looking at ICT solutions, it was important to assess what kind of reach ICT had amongst farmers. In 1994, India had less than ten million phones and was not adding more than 10 million phones a year. A combination of wireless telephony, better signal processing and innovations from Indian telecom operators changed the picture dramatically (Jhunjhunwala & Ramamurthi, Wireless in Local Loop: Some Key Issue, 1995). The costs of telephony were brought down significantly and 2G GSM (Global System for Mobile Communications) and GPRS (General Packet Radio Service) networks mushroomed throughout the country, including a large number of villages. With these networks offering robust voice connectivity and with the prices of mobile handsets also dropping significantly, most farmers can now afford a cell phone. India, today, has over 900 million users with close to 300 million rural subscribers (Telecom Regulatory Authority - Highlights on Telecom Subscription Data). Further, farmers are

[23]Junior Research Fellow, Symbiosis International University, Maharashtra, India, abhishekbehl27@gmail.com

[24]Post Graduate Student, Symbiosis Institute of Management Studies, Symbiosis International University, Maharashtra, India, maneesha.0602@gmail.com

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

comfortable with voice telephony, limited data handling (such as SMS), as well as using a camera and sending pictures using MMS (Multimedia Messaging Service). The other challenge before ICT solutions is that most farmers are either illiterate or semi-literate. Though they are capable of taking and sending pictures, they find it difficult to handle and interpret text/data information. They are most comfortable when the rest of the communication takes place in the local language, using voice. It is imperative therefore, that any agricultural advisory provider takes these realities into account.

This paper will demonstrate the power of two-way voice services over mobile phones today, which when combined with innovative methods, can serve to deliver information as well as seek feedback. We will do this in the context of the efforts of the Department of Agriculture, Government of Karnataka to deliver agricultural advisory voice messages directly to the mobile phones of farmers of five delta districts through a 'PUSH' mechanism. The next section will present an overview of some of the current services delivering agricultural advisory messages to farmers after which we present the background to our work followed by the research methodology and summary of findings. We discuss extensively the potential and challenges of using voice technology to obtain and implement feedback, and finally we present the conclusion.

OVERVIEW OF MOBILE AGRICULTURAL ADVISORY SERVICES

In India, mobile phone users have increased from fewer than 37 million users in 2001, to over 900 million users in 2012 with close to 300 million rural subscribers. As mentioned earlier, farmers can now afford a cell phone and are comfortable communicating over the phone in their local language. The potential of mobile technology in serving the more vulnerable, less-empowered and often dispersed users, has been extensively discussed (Behl and Singh, 2014) (Thompson, 2008) (Hellstrom, 2009). In recent years, it has also been recognized that mobile technology makes it possible to deliver timely, relevant, and quality information to farmers with fewer resources than any traditional service. Some agricultural advisory services that have tried to capitalize on the mobile boom are:

- Reuters Market Light (RML) (Reuters Market Light) (Thomson Reuters: Removing Barriers to Growth in Agriculture, 2011): Reuters Market Light (RML) promoted by Thomson Reuters was initiated in 2006. It provides highly customized and localized agricultural information services via mobile phone based Short Message Service (SMS) primarily for RML subscribed farmers, in 8 local languages.

- Nokia Life Tools (IBS Center for Management Research, 2009): Nokia Life Tool was launched by Nokia India Private Limited in 2009, in Karnataka, later expanding to provide agricultural information through SMS, to customers in 18 Indian states, covering 11 Indian languages - English, Hindi, Oriya, Gujarati, Bengali, Marathi, Punjabi, Telugu, Tamil, Kannada and Malayalam. The services are subscription oriented, where one pays a monthly fee and some are chargeable on a per download/request basis.

- Intuit Fasal (Intuit Fasal): Intuit Fasal, started in the year 2009, aims to deliver personalized messages through SMS in the local language to farmers. Currently operational in Gujarat, Andhra Pradesh and Karnataka, the service allows farmers to contact a toll free number and register their profile. Based on this profile, the service attempts to deliver personalized market information to the farmer in an effort to directly connect him/her to a buyer/agent/institution. (A quick point to be made here is that the above-mentioned services use SMS as a medium of information delivery; it remains to be seen how these services will reach out to the largely illiterate or semi-literate farmers.)

- IFFCO Kisan Sanchar Limited (Narula, Sikka, Singh, & Chawla, June 2012): Started in the year 2007, IFFCO Kisan Sanchar Limited offers the following services to farmers from 18 states:
  - 5 free voice messages every day on areas of interest to rural subscribers. Messages are prepared by subject matter experts; Content Managers trim it down to one minute lengths; Panel of eminent scientists monitor the quality,
  - Dedicated helpline for query resolution by Experts-Conference Calling available,

- Avaaj Otalo (Patel, Savani, Klemmer, & Parikh, 2012): This service is designed to target farmers and help them in accessing timely and relevant information related to agriculture over phone. This service kick started with collaboration between Stanford HCI Group, UC Berkeley School of Information, IBM India Research Laboratory and Development Support Center (DSC), an NGO in Gujarat, India. The key features of this service included dialing a phone number and navigating through simple audio prompts. This also had a facility of recording the voice of farmers, browse their menu and get responses of questions and suggestions related to agriculture.

- mKrishi (mKrishi: A Rural Service Delivery Platform) : mKrishi is a mobile platform, launched in 2009, that allows a farmer to send queries in the form of text, voice or images to the system; the expert will analyses the query based on the farmer profile and the advice/information is delivered to the farmer's mobile. Since then, 12 additional mKrishi pilot projects have been introduced in the Indian states of Punjab, Karnataka, Karnataka, Gujarat, Andhra Pradesh, Karnataka and Rajasthan. Today, mKrishi serves about 10,000 farmers in all markets.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

- ITC's Namma Sandesh (ITC launches interactive mobile telephony for tobacco farmers, 2012): This is a new entrant in this industry, which is designed with reference to tobacco and ragi farmers. This interactive mobile telephony based communication offers advisory services as well along with market prices of agriculture equipment's, weather forecast and local news to the farmers. Started with their pilot launch in Mysore, Karnataka the messages were in Kanadda and slowly gathered momentum by diversifying its IVR menu and reaching to a wider audience.

The success of advisory voice messages, which are delivered because of farmers subscribing to the system or in response to a specific request for information, is easy to gauge depending on the number of subscriptions, queries or requests. For systems deploying the 'PUSH' mechanism to deliver personalized messages to farmers, delivery reports can be a trifle deluding and there is no easy way of understanding if messages are reaching the intended beneficiaries or if the information is truly useful.

So how does one gauge the success of these services? Can ICT itself be used to gain feedback on the usefulness of the service regularly and make the advisories more relevant? The paper addresses this question – as a solution, a system that not only delivered advisory voice messages but also incorporated a real-time feedback mechanism, was developed.

BACKGROUND OF STUDY

The Government of Karnataka has undertaken a programme to partially fill the existing information gaps for farmers by using ICT tools. The pilot project of relaying agricultural advisories by means of a 'PUSH' voice-message system to the mobile phones of farmers in five delta districts (Thanjavur, Thiruvarur, Thiruchirapalli, Nagapattinam and Cuddalore) was started in August 2012 and presently covers around 0.2 million farmers. The Cauvery Delta Zone in the state of Karnataka is a major rice-producing zone in the country and the above-mentioned districts cover a large fraction of the delta zone (Agro Climatic Zone Profile, 2009). The Agriculture Department naturally thought it fit to start relaying advisories to the districts where the rice productivity matters most to the state.

The Department first conducted a baseline survey of all farmers in Karnataka over 2010-11 as part of their Farm Crop Management System (FCMS). This comprehensive baseline survey was intended to cover regional details, personal details, mobile details, bank details, land details, details about farmer group memberships, farm animals, farm equipment, plantation crops, wells, micro-irrigation, fodder crops grown, and other activities such as beekeeping, sericulture, fisheries and agroforestry. It also endeavoured to obtain information regarding facilities available to farmers and what their requirements were – such as credit facilities, inputs availability, farm equipment and market information. This data was then used to classify farmers and understand the requirements for specific sets of farmers.

Based on the categorization, the Department then started relaying voice messages to farmers on a monthly basis, targeted block-wise and depending on the farmer-category. The voice messages were of maximum one minute duration. If the farmer does not pick up the phone at the first attempt, another call is made in an hour; if that too fails, a last attempt is made in the evening between 5 pm and 8 pm. Taking feedback from when the maximum calls get connected, the timings were fine-tuned for each region. The messages contain information on subjects that are presumed to be relevant to farmers such as:

- Fertilizer application,
- Pesticide application,
- Pest management,
- Disease management,
- Best Agricultural practices,
- Seed varieties,
- Seed treatment,
- Weeding,
- Government Schemes.

Currently, over 3 million messages have been relayed to farmers with an average call connectivity rate of 87%; it was found that 50% of farmers listen to the entire message, whereas others disconnect the call. The system was initially built only to deliver messages and did not have any mechanism of obtaining feedback from farmers about whether the call was found to be useful. To make the messages truly useful to farmers, it was believed that feedback on the following aspects would be desirable:

- Are the messages timely – do the messages reach the farmers when they need it the most during the crop cycle?
- Are the messages relevant – do the messages contain information that is relevant to farmers or would they like information on other topics?
- Can the messages be more personalized – is there a way in which feedback can be obtained from farmers to understand the crops they are growing and what cropping stage they are at?

Further, if this feedback is to be used to modify the content and delivery of the messages, it had to be obtained and analyzed in real-time.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Considering the importance of gaining feedback, an IVR system was introduced into the same system, to seek feedback from farmers in an innovative way – by including questions at the end of the messages so that farmers could provide instant responses; responses that could be recorded, analyzed and reported to the Department of Agriculture. The questions are designed to be simple, are in the local language and asked at the end of voice message, which the farmer responds to. This response gives the required feedback to make the next message more relevant.

RESEARCH METHODOLOGY

Prior to implementing feedback using the IVR system, an exploratory study was carried out to understand the time, cost and resources needed to seek feedback from farmers using a more traditional method – manual survey over the phone.

Three districts were chosen for the survey as farmers in these districts had received at least five advisory messages from the Department; the total number of farmers included in the survey was 147,117. Out of this, every 200th farmer was contacted on a random basis; the total number of farmers surveyed was thus 167. A pre-determined questionnaire consisting of 19 questions was administered to them. The survey was administered in the local language, i.e. Tamil, over the phone and was conducted from: September 24, 2012 to October 12, 2012. This meant a total of 3 weeks, with each farmer taking about 10 minutes to respond to the questionnaire.

Some of the key questions included in the questionnaire were:

- **Socio Economic Details:** Age, gender, education level, total farming experience and number of family members involved in farming.
- **Sources of Information**: What are the different sources from where farmers currently acquire agricultural information? Options included TV, radio, extension officer, progressive farmers, family members, etc.
- **Implementation of information received by farmers through agricultural advisory voice messages:** Had farmers implemented or planned to implement the information relayed through the voice messages?
- **Perceived usefulness of messages**: How useful did farmers perceive the information to be? Were they hearing any of the information for the first time?
- **Level of satisfaction with the messages**: Were farmers satisfied with various aspects of the message – voice, simplicity and audio quality?
- **Comparison of voice messages with other sources of information**: How did farmers compare the information they received through the voice messages relayed by the Department of Agriculture with information they received from other sources mentioned in (2) above?
- Was the farmer receiving mobile-based SMS or Voice messages agricultural related information from any other sources, in addition to the Department of Agriculture, Government of Karnataka?

The number of farmers contacted to get 167 responses was 774. (A majority of the respondents had their phones switched off, some numbers were not valid and some were from a different district or block than that mentioned in the Government database).

Following this, a more innovative attempt was made to gain an understanding of whether farmers found the information contained in the message useful. In the month of October, two messages containing information on Government schemes were to be relayed to farmers by the Department of Agriculture. The duration of the first message was of 40 seconds and it was decided to deliver this message to all farmers with a feedback question inserted at the end. The question was asked in Tamil, through an IVR prompt and the response recorded by requesting users to select specific keys on their Dual-Tone-Multi-Frequency (DTMF) keypads.

The message along with the question asked was as follows:

Farmers can get free paddy seedlings through community nurseries and get 25% subsidy on fertilizers; to know how to avail these facilities, they can get in touch with officials of the nearest block agricultural extension center.

Was the information contained in the message useful to you? Press '1' for Yes and '2' for No

This message along with the question was relayed to 130,421 farmers from five districts over five days beginning 24 September 2012 to 28 September 2012, with 80338 farmers having heard the entire message.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

SUMMARY OF FINDINGS

Delivering advisory voice messages to farmers using mobile technology had certainly proved to be more cost-effective than other traditional advisory method. Further, obtaining feedback by integrating it with the ICT service itself, also meant less time, less costs, less resources and far greater reach.

- In the traditional survey method, the resource person could contact only 774 persons in 3 weeks and get survey responses from 167. In contrast, using the second evaluation method, 130,421 farmers could be contacted in 5 days and 4216 responses were recorded.

- Though 80338 farmers heard the message and the question at the end of it, 76123 farmers did not provide any response. It is possible that there needs to be greater clarity in communicating how the DTMF response is to be registered or that farmers were not comfortable with the DTMF options at all. However, out of the 4216 farmers who did respond, 3841 (91%) responded 'Yes' (The information was useful) while 375 said 'No.

- In the traditional method, a resource had to manually enter the data into an online form or excel sheet while speaking to farmers, increasing the chances of errors in data entry during the phone survey. Since the responses in the IVRS feedback method were recorded directly by the system, they are likely to be error free.

- A comparison of the costs[2] of administering a phone feedback survey versus that of administering it through the IVR system is also very illuminating.
  - Cost of administering the survey and data entry: INR. 12000 (time of the resource administering the survey)
  - Cost of phone calls: INR.1 per minute (10 minutes for each survey) to 774 farmers – INR. 7740

  Total costs involved in conducting the manual phone survey: INR. 19740.

  When the feedback question was included at the end of the message, no additional costs were incurred to actually administer the question through IVRS. The only time spent was in report generation, the cost of which was INR. 1250. (As we have mentioned, the question was inserted at the end of the message so that no extra costs were borne for the time being, other than that of actually relaying the message; cost of relaying all the messages were being borne by the Department of Agriculture).
  Total cost involved in conducting the survey through IVRS: INR. 1250

- The traditional phone survey provides an opportunity for gathering feedback that is more comprehensive and it is true that only one feedback question could be included in the message during the relay through IVRS. Yet, we believe our findings are sufficient proof of the potential of using IVRS to obtain feedback and aid in improving delivery of services.

Innovative Nature of IVRS to Obtain and Implement Feedback

Potential

The savings in cost, time and the availability of error free data are some of the obvious advantages of using the IVRS feedback method. The study also revealed the tremendous potential of this method as could be seen from some of the aspects listed below:

- **Customized Questions:** The questions can be customized to obtain feedback on the aspects that are most relevant to the Department. For example, the question could be about whether the information is indeed useful to the farmers, if it is reaching farmers in time or whether they would like information on any other issue.

- **Better Targeting**: The study revealed that data collected from farmers during the baseline survey had some errors in them or had changed over the last couple of years. The feedback technique can serve as an excellent mode of data cleansing as well as data updating. This will help the Department understand if the beneficiaries are indeed farmers, has there been a change in the category they belong to and how best can messages be personalized for them.

- **More Relevance**: Since the responses received from farmers can be immediately recorded and analyzed, the feedback thus received can be utilized faster for better targeting of messages. For example, if the Department of Agriculture is made aware of the fact that farmers in a certain block have already finished transplanting their paddy saplings and they are no longer in the nursery stages, rather than continue with relay of nursery related messages; an immediate switch can be made to messages that contain information about the main field stage.

- **Communication in Local Language**: In what is a significant feature, gathering feedback through the mobile ensures that farmers – most of whom are semi-literate or illiterate – are free to communicate through a medium that they are most comfortable with viz., voice conversations in the local language over a mobile. Rather than depend on intermediaries who could misrepresent them, farmers are provided a tool to directly give feedback to the Department.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

- **Extending Reach**: As mentioned before, three attempts are made to relay the message (and the feedback question it carries) to farmers. Through field interactions, it has been observed that farmers are available on their mobile phones either early in the morning or later in the evening. Based on call reports, the timing of the message (plus feedback question) relay can be fine-tuned so that a large fraction of farmers can be reached.

- **Scalability:** In a country like India where there are different languages, voice technology is a powerful tool that can be used to scale the proposed system across the different states very quickly. If the system proves its benefits in the State, it will not take long for it to be replicated in other States as well.

Challenges

It is also important to be aware of the technological challenges facing such a feedback technique. Though the feedback question was relayed to over 80,000 farmers, only about 4000 could actually respond by pressing the right keys. Mention must be made here that farmers were not provided any training or sensitization about this new mode of gathering feedback. The next aspect of study will be to understand if farmers are comfortable using the DTMF options, if additional marketing or sensitization efforts will help or if using an Automatic Speech Recognition (ASR) system in the local language, might work better.

CONCLUSION

In all the efforts mentioned earlier, we have seen both Government and non-Government organizations harnessing ICT, in particular mobile telephony, to deliver messages to farmers. However, in these cases, farmers have either opted in for the service or it is a pre-defined and targeted service delivery; by a Government entity or Contract Farming entity. We have, so far, not seen an effort similar to what has been attempted in this study – to use ICT to 'PUSH' informational messages to a large scale of farmers as well to continuously customize delivery based on the feedback obtained directly from farmers. In fact, a very effective way of determining the success of 'PUSH' services is by integrating these feedback mechanisms within the system so that the end-users have the power to immediately convey feedback to the organizations and organizations can act in time to respond to that feedback.

This study clearly shows that ICT through innovative use of mobile technology can help in implementing the feedback, through what we believe is a unique and so far, under-utilized approach; not to mention dramatically reduce the time and costs involved in carrying out such feedback efforts. Recognizing this potential and making it a part of the Government ICT schemes as they are being conceived and operated will go a long way in increasing the productivity and effectiveness of these schemes.

Though initial attempts have shown that some farmers have not provided any feedback due to varying reasons, one must note that there were no specific marketing efforts or training that was provided to the farmers introducing the innovative channel using which their feedback was to be obtained. As we go forward one can certainly envision that the farmers will be acclimatized to this innovative method of feedback gathering and even welcome it, after they witness the change such an approach will bring to the farming community.

REFERENCES

1. (2009). ***Agro Climatic Zone Profile.*** Government of India: Planning Commission. Retrieved from http://planningcommission.nic.in/reports/sereport/ser/7vgtn/v3_ch3.pdf

2. Behl, A., & Singh, D. M. (2014). Mapping Infrastructure Using Geomatics Model in Mgnrega in India. ***International Journal of Trade & Global Business Perspectives,*** 2(4), 695-700.

3. Chhibber, N. (2004). Enhancing Developmental Opportunities by Promoting ICT Use: Vision for Rural India. ***I-Ways: The Journal of E-Government Policy and Regulation***, 27(3), 190-196.

4. Glendenning, C. J., Babu, S., & Asenso-Okyere, K. (2010). Are farmers information needs being met? ***Review of Agricultural Extension in India.***

5. Hellström, J. (2008). Mobile Phones for Good Governance–Challenges and Way Forward. **In *W3C Workshop, Africa Perspective On the Role of Mobile Technologies in Fostering Social Development, Maputo, Mozambique***. Retrieved on 2010, August 30 from http://www.w3.org/2008/10/mw4d_ws/papers/hellstrom_gov. pdf

6. IBS Center for Management Research. (2009***). Emerging Markets Strategy: Nokia Life Tools for Rural Markets.***

7. (2012). ***Improving Earning Capacity of Farmers: Case of Fasal by Intuit Labs.*** Governance Knowledge Center, Department of Administrative Reforms and Public Grievances, Government of India. Retrieved from http://indiagovernance.gov.in/files/intuit_fasal_edited_intuit_final.pdf

8. (2012, August 16). ***ITC Launches Interactive Mobile Telephony for Tobacco Farmers.*** eGov magazine enjeGov Magazine- Asia's first magazine on e-Governance. Retrieved from http://egov.eletsonline.com/2012/08/itc-launches-interactive-mobile-telephony-for-tobacco-farmers/

9. Jhunjhunwala, A., Jalihal, D., & Giridhar, K. (2000). Wireless in Local Loop-Some Fundamentals. ***Iete Journal of Research,*** 46(6), 421-434.

10. Jhunjhunwala, A., Umadikar, J., & Prashant, S. (2013, April). A New Personalized Agriculture Advisory System. ***In Proceedings of the Wireless Conference (EW), Proceedings of the 2013 19th European (pp. 1-6).*** VDE.

11. mKrishi: A Rural Service Delivery Platform. (n.d.). Retrieved from Tata Consultancy Services: http://www.tcs.com/offerings/technology-products/mKRISHI/Pages/default.aspx

12. Narula, S. A., Sikka, B. K., Singh, S. P., & Chawla, K. L. (2008). ***Empowering Farmers through Mobile Telephony in India: A Case of IFFCO Kisan Sanchar Limited,*** 469**.** New Delhi: IFFCO Kisan Sanchar Limited. Analysis of the Usage of Market Information System.

13. Patel, N., Shah, K., Savani, K., Klemmer, S. R., Dave, P., & Parikh, T. S. (2012, March). Power to the Peers: Authority of Source Effects for a Voice-Based Agricultural Information Service in Rural India. ***In Proceedings of the Fifth International Conference on Information and Communication Technologies and Development (pp. 169-178***). ACM.

14. Sharma, R. (2010). ***Use of ICT in Improving Governance and Social Accountability with a Pro-Poor Approach: Reflections from India***. Government of Andhra Pradesh: Centre for Good Governance.

15. ***Telecom Regulatory Authority - Highlights on Telecom Subscription Data.*** (n.d.). Retrieved from http://www.trai.gov.in/WriteReadData/WhatsNew/Documents/PR-TSD-Sep2012.pdf.

16. Thompson, M. (2008). ICT and development studies: Towards development 2.0. ***Journal of International Development***, *20*(6), 821-835.

17. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2363960

18. Retrieved from http://www.tenet.res.in/Publications/Presentations/pdfs/A_New_Personalized_Agriculture_Advisory_Syst. ..

*****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

A COMPARATIVE STUDY OF WEB MINING IN HEALTH CARE WEBSITES

Nirmala Shinge[25] Dr. Nilesh Mahajan[26]

ABSTRACT

Web mining is application of data mining and its meaning that to find useful information from web data. The information in the form of web document, Hyper links between n number of document, web figure web maps anything information related to search pattern. Web mining has three categories: Web content mining Web structure mining & Web usage mining. Above categories are describe in detail.

This paper provides comparative study of two healthcare website. We believe that web mining will be hot research topic in future. This paper gives overview of web mining importance of web mining in medical and healthcare and compares two healthcare website & gives some suggestion.

KEYWORDS

Web Mining, Data Mining, Web Content Mining, Web Structure Mining, Web Usage Mining, Evaluation Checklist Health Care Websites etc.

INTRODUCTION

The www is a popular and largest source of data. Web is broadcast medium to extract wide range of information and these obtain at very low cost. The number of health related information is available on the website and they provide online service day by day. This web site gives information such as information about hospital in particular area, information about specialists, for particular disease, information about medicine and services. Unlimited information may also be provides by web site of above criteria. Therefore, website is very popular concept for human society. It bring changes in business, organization, educational field in the form of e-commerce, e-learning etc.

These healthcare website are helpful to patient, physicians but some issues are rising. Some issues are patient or user cannot understand medicine vocabularies, as well as language problems means website should provide in regional language that user can understand information.

OBJECTIVES OF STUDY

- To comparative study between two healthcare website in Pune city,
- To identify criteria for the evaluation of healthcare website,
- To evaluate the healthcare websites with identified criteria for the verification of validity reliability and usefulness,
- To rank the selected healthcare websites based on the identified criteria.

SCOPE AND LIMITATIONS

The study is limited to the two select healthcare websites of Pune city and based on the evaluation criteria of the currency, accuracy, relevance, organization and structure, presentation, URL maintenance and features etc. and manual evaluation. The evaluation of select healthcare websites is carried with a purpose of verifying the validity, reliability and the usefulness of identified criteria.

METHODOLOGY AND EVALUATION CHECKLIST

The quantitative rating system (ten-point scale) was designed to evaluation checklist whether to provide a meaningful numerical rating for each individual feature of the select healthcare websites. The rating scale 0-10 with rating 0 is being assigned to the least and 10 to the highest rate by the evaluator of the study [6].

The rating table allows the evaluator to assign numerical value to select healthcare website in Pune city purpose is best served when comparing and selecting sites similar in purpose and scope [6].For privacy we cannot mention website's name.

[25] Assistant Professor, Sinhgad Institute of Business Administration And Computer Application, Maharashtra, India, nirmalashinge@sinhgad.edu

[26]Professor, Institute of Management and Entrepreneurship Development, Maharashtra, India, nileshmahajan2004@yahoo.co.in

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

ANALYSIS OF THE DATA

This study evaluates the health care websites in five areas: currency, accuracy and relevance; organization and structure; presentation; maintenance; and different features of the healthcare website. The comparative statement for ranking of healthcare websites is presented in the Table VII [6].

Currency, Accuracy, and Relevance of Health Care Websites

Currency refers to the timeliness of information. Accuracy generally refers to the information content to get the information source and the correctness of the source of information about health care. In many respect, the need to determine accuracy underpins the whole process of evaluation: it is often the reason for looking critically at any information and relevance being an important part of the evaluation process. Table I shows the evaluation of healthcare websites in term of retrieval efficiency, contact information, updating, copyright status, and online reference desk assistance.

Table-I: Currency, Accuracy and Relevance of Health Care Websites

Currency, Accuracy and Relevance criteria	Health care Websites		Score
	Web Site 1	Web Site 2	
Retrieving all the hyperlinks in the webpage	09	10	19(95%)
Given all the contact information	10	10	20(100%)
Copyright status are clearly stated	08	10	18(90%)
All hyperlinks appropriate and relevant for an online reference desk	08	09	17(85%)
Last time the site was updated	10	05	15(75%)
Total Score (Max. 50)	42(84%)	47(94%)	

Sources: Authors Compilation

It is observed from the table-I that 95% of health care websites retrieve all the hyperlinks in the webpage, followed by 100% provide contact information, 90% provide copyright status, and 85% are providing appropriate and relevant hyperlinks for reference desk and only 75% of study websites have provided the information about the date of last updating the site. It is noted that the date of last revision is very important aspect of the website content otherwise; the number of user will be reduced because of out dated information. Currency is an important consideration because outdated information can become useless as well; inaccurate or misleading Copyright in relation to electronic information is a complex area. However, one consideration in terms of evaluation is the availability of copyright information. User may want to re-use textual. However, as a basic rule, any information, which is published via the Internet, will be covered by copyright, including images, the text of web pages and the contents of e-mail and Usenet messages. It is therefore useful if the authors or webmasters provide a statement of the copyright ownership of materials, details of how materials should be cited in a publication or attributed to an author as well as the individual who should be contacted where copyright permission is required [6].

Organization and Structure of Health Care Websites

The web page content needs or rather demands the hierarchical, linear and randomly interlinked combinations of the two styles. The structure used to organize the contents of a web page is easy for users to navigate. Organization is an important factor, which should be done in such a fashion that each web page will be independent of the other. Each web page has appropriate images, icons and graphics and these should load the home page very fast in any web browser. Proper linking must be maintained so that the user can have provision to come back again to any one of the earlier pages. However, every page should include some information also and not only the links. Table II presents the analysis of organization and structure of study websites in term of accessibility, web browser, loading of images and graphics, and fast retrieval of web pages [6].

Table-II: Organization and Structure of Healthcare Websites

Organization and Structure Criteria	Healthcare Website		Score
	Website 1	Website 2	
Site is accessible from all the Browsers	10	10	20 (100%)
How fast does the web page paint?	10	08	18(90%)
All images, icons and graphics paint when the web page loads	10	07	17(85%)
Total Score (Max. 30)	30(100%)	25(83.3 %)	

Sources: Authors Compilation

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

The data in table-2 is related to organization and structure of the health care websites, 100 % each health care websites are accessible from all the web browsers, "the web page designed should be accessible by a variety of browsers" [13]. 83% sites are properly loading images, icons and graphics on their home page. Not all the homepages of health care websites with image files took more than 25 seconds to appear completely on the computer screen, "for web pages with images files, the loading time was less than 20 seconds" [11]. From the above table it is clear that only web site 1 is good in this criterion,

Therefore, the scores obtained in this table are consolidated with the scores of other tables (means transferred to final ranking table, i.e. Table II score of the ranking table) in the final ranking Table VII.

Presentation of Health Care Websites

The presentation and arrangement of information on the screen very much matter for the ease of assimilation. This includes whether screens are clearly laid out and aesthetically pleasing, whether there is too much information on each screen, whether the text is easy to read and whether heading stands out. Care should be taken that the page is clearly laid out and all the links are easily displayed on one screen. It is possible to navigate from one section to any of the major sections of the service and the resources that are linked to each hyperlink are self-evident. It is also worth considering whether a source is consistently, logically presented and arranged.

Table III explains the graphics, text and easy navigation, and presentation of health care websites [6].

Table-III: Presentation of Health Care Websites

Presentation criteria	Healthcare Websites		Score
	Website 1	Website 2	
The graphics and texts are most clear and easy to read	10	10	20(100%)
Web pages are easy to navigate	07	09	16(80%)
Every page included way to turn the home page for the site	10	10	20(100%)
Graphics are the most appropriate middle group	08	09	17(85%)
Total Score (Max. 40)	35(87.5 %)	38(95 %)	

Sources: Authors Compilation

In addition, "good presentation" is often a matter of personal taste. Some of us will feel a particular feature is essential while others may feel the same feature is redundant. However, a comprehensive perspective should include every aspect. Further considerations are: the number of clicks required to locate relevant information, whether steps are unnecessarily repeated, and whether useful shortcuts are available, such as a "home" icon to take directly to the start of a document or resource. In addition, some web pages are extremely long and users must scroll through them to find what they want. One alternative is providing links between different sections of the same page or splitting a document into parts by providing links between the different sections. However, the opposite extreme might have many separate pages for small sections of the same document, continually forcing users to download different parts but without offering the option to scroll through larger sections. Obviously, the length of a page is a highly subjective issue but it might consider whether a useful balance has been achieved and whether the organization is appropriate to the content [6].

Maintenance of Health Care Websites

Maintenance of the Health Care website is ongoing practice and tedious job of the webmaster. One factor to be considered is the currency of all hyperlinks; some sites include a policy regarding the updating process. If individuals or groups maintain a site voluntarily, they may be more likely to lose interest and therefore fail to maintain the site effectively in the long term. Contact information for site maintainers is also a useful feature and suggests a concern for site maintenance .In fact, the most exciting and useful feature of the web site is the implementation of web forms. Forms provide a way for collecting detailed information from Web users. The suggestions should be the integral part of the website development, especially in the initial stages it helps in correcting the design, as the suggestions are the views and reactions of the end-users. "Web-based forms, which are effective tools for library-user interaction and communication" [12].

Speed of access is of particular concern and factors affecting speed include the location of sources, the number and size of any images. In addition, "sites are faster to access if it is possible to view a text-only version of the information". This ensures that pages are meaningful to any user irrespective of the fact whether the images are viewed or not. The maintenance of the health care websites in term of maintenance, feedback and speed of access is presented in Table IV.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Table-IV: Maintenance of Healthcare Websites

Maintenance Criteria	Healthcare Websites		Score
	Website 1	Website 2	
Persons responsible for the site display	08	10	18(90%)
Allows a user to send feedback or comments to Healthcare	05	10	15(75%)
Server appear to be fast	07	10	17(85%)
Total Score(Max.30)	20(66.6%)	30(100%)	

Sources: Authors Compilation

It is observed data from the Table IV that the website 2 scores (100 percent) highest among the study websites.

Features of Health Care Websites

If the user is looking for specific information this is not reflected in the contents pages of the website, an additional search facility will be of great help. The table V presents the different features of health care websites as prescribed by the evaluation checklist.

Table-V: Features of Health Care Websites

Features of Heath care Websites criteria	Healthcare Website		Score
	Website 1	Website 2	
Hyperlinks to open access resources	08	10	18(90%)
Search Engine functionality	05	10	15(75%)
Hyper links to other web reference Heath care Websites	07	10	17(85%)
Total Score (Max. 30)	20(66.6%)	30(100%)	

Sources: Authors Compilation

A website's primary function is to provide a search engine for gathering and reporting information available on the internet or a portion of the internet. The survey indicates that, out of two healthcare websites under the study only one is incorporated this feature; the same fact is attested by the recent study conducted by [14].

Rating Scale for Health Care Websites

Using data gathered from the study websites analysis and the evaluation checklist, a numeric score was generated for each health care website in the study. The five point rating scale was fixed equally based on the maximum score 180 of currency, accuracy and relevance (Table I and 50 points); organization and structure (Table II and 30 points); presentation (Table III and 40 points); maintenance (Table IV and 30 points); and features of library websites (Table V and 30 points), to rank the study the study libraries, the rating scale was designed to rank "Excellent" to "Needs Improvement" based on five point scale below (Table-VI):

Table-VI: Rating Scale for Health Care Websites

Range (Score)	Rank
150 - 180	Excellent
100 - 149	Above Average
70 - 99	Average
43 –69	Below Average
01 - 42	Needs Improvement

Sources: Authors Compilation

Ranking of Health Care Websites

Ranking of the study websites based on five point rating scale (Table VI) and points taken from the Table I through V is presented in the following Table VII [6].

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Table-VII: Rankings of Health Care Websites

Sr. No.	Healthcare Website	Evaluation Criteria and Points*					Total Points (180)	Rank
		Currency, accuracy and relevance (Table I, 50)	Organization and structure (Table II, 30)	Presentation (Table III, 40)	Maintenance (Table IV, 30)	Features of library websites (Table V, 30)		
1	Web Site 1	42	30	35	20	20	147(81%)	Above Average
2	Web Site 2	47	25	38	30	30	170(94.7%)	Excellent

Sources: Authors Compilation

It is observed from the Table VII that the rankings based on the points assigned to the Health care websites and the table reveals that Web Site 2 got 170 points (94.7 %) and is ranked "excellent", and web site 1 got 147 points(94.7%) ranked as "above average". These ranks clearly indicate that there is a need to develop health care websites in the above features. Only developing their websites and meeting the web challenges to strengthen their vital web-based health care services is necessary.

SUGGESTIONS

The evaluation of study websites and the subsequent analysis of the data and the findings of the study have enabled the investigator to provide some practical suggestions about starting and improving the web-based health care services expected from various websites [6]. The main suggestions for improvements are as follows:

1. Health care websites should be developed into a more dynamic, interactive, and compatible with Web 2.0 and Web 3.0 technologies.
2. Health care provide dynamic contents, RSS feeds, instant messaging (IM) reference services ,online health care and floor maps, FAQ s , bulletin boards, discussion forums .
3. Blogs, or web logs, are versatile platforms for presenting information to health care patrons.
4. Healthcare should make the portal interactive by hosting the announcements about new services, links to the websites.
5. Every effort should be made on a consistent basis to update the heath care website frequently.
6. While uploading any web page to web server the Webmaster must test the web pages thoroughly using different browsers. He/she should ensure that the page should be seen similar in all the browsers.
7. Web forms in each web-based health care service should be provided in the websites to attract the suggestions, comments about the pages. In case of health care websites, the feedback mechanism will enable the owner to know about patient requirements and accordingly improve the patient services.
8. There should be a regular and continuous user-oriented evaluation of websites as to adjust the site depending upon the requirements of the users.
9. Each health care on a regular basis should compare its own website with those of similar websites in order to simulate the continuous development of the site.
10. All health care website should be provided in various languages that each patient understand information of the health care.
11. All the health care websites must and should evaluate their websites with online tools.

CONCLUSION

In this paper, we have discussed three types of web mining in particulars web mining divided into three types such as web content mining, web structure mining and web usage mining. At last, compare two healthcare website structures that means good facilities provide through web pages and give Common suggestion to both Health care website for their interactive web site

REFERENCES

1. Raikwal, J. S., & Saxena, Kanak. (2011, September). Application of Web Mining in Medical and Healthcare: A survey. ***Research Cell: An International Journal of Engineering Sciences***, 4, 255-264. ISSN: 2229-6913

2. Oren, Etzioni. The world wide web: Quagmire or gold mine. ***Communications of the ACM***,39(11):65-68,1996

3. Salton, G., Wong, A., & Yang, C. S. A. (1975). Vector Space Model for Automatic Indexing. ***Communications of the ACM,*** 18, 613-620.

4. Dunham, M. H., & Sridhar S. (2006). ***Data Mining: Introductory and Advanced Topics*** (1st Ed.). New Delhi: Pearson Education. ISBN: 81-7758-785-4.

5. Ellen, Spertus. (1997, April). ParaSite: Mining Structural Information on the Web. ***In Proceedings of 6th International WWW Conference.***

6. Konnur, P. V., Rajani, S., & Madhusudhan, M. (2010). ***Academic Library Websites in Bangalore City, India: An Evaluative Study***. Library Philosophy and Practice (e-journal). Paper 408.

7. Sita Gupta, Vinod Todwal (2012, February). Web Data Mining & Application. ***International Journal of Engineering and Advanced Technology (IJEAT)***, 1(3), 20-24. ISSN: 2249 – 8958.

8. Yan, Wang. (2000, February). ***Web mining and Knowledge Discovery of Usage Patterns. Cs 748T Project*** (Part I).

9. Spiliopoulou, M. Data mining for the web***. In Proceedings of principle of data mining and knowledge Discovery, Third European Conference, PKDD'99, (pp. 588-589)***.

10. Ahmed, S. S. (2002). Managing Change to Enhance Web-Based Services in the Arabian Gulf Libraries. ***Online Information Review*** 26(4), 265-270.

11. Lee, K. H., & Teh, K. H. (2001). Evaluation of Academic Library Websites in Malaysia. ***Malaysian Journal of Library and Information Science,*** *5*(2), 95-108.

12. Prasad, A. R. D. (2001). Designing and Developing Library Websites. ***In Proceedings of First ASSIST seminar on Internet Resources and Librarianship. Asian Society for Social Services, Information Science and Technology,*** 12-14 August 2001, Bangalore): 228-238

13. Detlor, B., & Lewis, V. (2006). Academic library web sites: Current practice and future directions. ***The Journal of Academic Librarianship,*** *32*(3), 251–58.

14. Retrieved from http://www.webpages.uidaho.edu/~mbolin/konnur-rajani-madhusudhan.htm

15. Retrieved from http://www.readperiodicals.com/201007/2125344131.html

16. Retrieved from http://www.webpages.uidaho.edu/~mbolin/konnur-rajani-madhusudhan.pdf

17. Retrieved from http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1419&context=libphilprac

*****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

IMPROVED LEARNING ENVIRONMENT USING ARTIFICIAL INTELLIGENCE AND KNOWLEDGE MANAGEMENT

Sona Kanungo[27]

ABSTRACT

We are steadily increasing reliance on use of Information Technology and on Artificial Intelligence in the educational processes. Artificial Intelligence provides various techniques to represent the knowledge and efficient heuristic search algorithms to store, process and retrieve the knowledge for solving complex problems that cannot be solved by conventional Information Technology applications. This paper addresses relevant Artificial Intelligence (AI) issues and their use for addressing problems in education. We have explored the issue of personalization of contents and tutoring, expert system, web based learning etc.

KEYWORDS

Artificial Intelligence, Knowledge Management, Intelligence, Heuristics, Expert Systems, Collaborative Learning etc.

INTRODUCTION

According to Allan Rich, "Artificial Intelligence (AI) is the study of how to make the computers do things at which, at the moment, people are better" [13]. According to Avron Barr, A. Feigenbaum "Artificial Intelligence is a part of computer science concerned with designing the intelligent computer systems; i.e. the computer system which exhibit the characteristics we associate with intelligence in human behaviour" [3]. Examples of Artificial Intelligence applications include Speech recognition, Smell and Face recognition, inference techniques, learning new skills, decision making etc. [14]. Various Task Domains of AI include:

A. Mundane Tasks

- Perception,
- Vision,
- Speech Synthesis and Recognition,
- Natural Languages:
    - Understanding,
    - Generation,
    - Translation,
- Common sense reasoning,
- Robot Control.

B. Formal Tasks

- Games Playing: Chess, Chinese Checkers, 8-puzzle Problem etc.
- Mathematics: Geometry, Mathematical Logic, Theorem Proving,

C. Expert Tasks

- Engineering (Design, Fault Finding, Manufacturing Planning),
- Scientific Analysis,
- Medical Diagnosis,
- Financial Analysis etc.

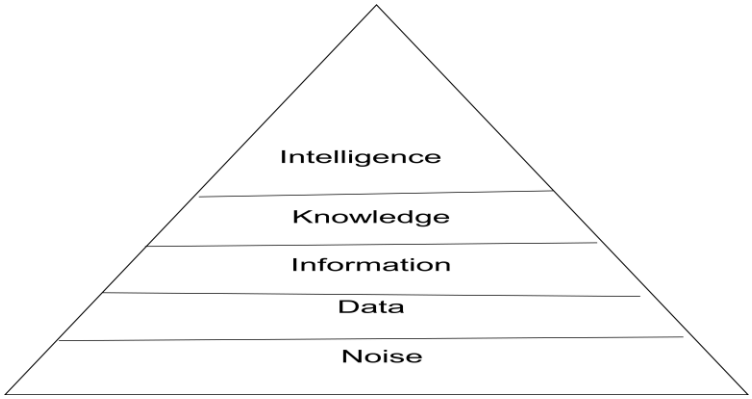
Main components of AI include knowledge and intelligence. The relationship between knowledge and intelligence can be depicted with the help of D-I-K (Data-Information-Knowledge) Pyramid as shown in Figure 1.

Data are the facts and figures in scattered way. Data can be coded and is transferable. From the data, information is extracted. Information is the data endorsed with relevance and purpose. Knowledge is different from and is the body of facts and principles accumulated by human minds. Knowledge needs learning. Knowledge has familiarity with language, concepts, procedures, rules, ideas, abstractions, places, customs, facts and associations coupled with ability to use these notions effectively in modeling different aspects of world.

[27]Lecturer, Shri Vaishnav Institute of Management, Madhya Pradesh, India, sonapkanungo@gmail.com

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Figure-1: D-I-K Pyramid



Sources: Authors Compilation

Knowledge is justified belief, which is known to be true. Belief is any meaningful/logical expression, which can be represented and can be true or false. Hypothesis is justified belief not known to be true and is backed by evidences but still may be false. Knowledge can be declarative or tacit Declarative Knowledge consists of the statement of facts about the world, e.g. Personal data of an employee. Tacit knowledge is personal, context specific, exists is the heads and is difficult to formalize. It is handled with trial and error. Procedural knowledge consists of the methods of doing some task. Heuristic knowledge is the non-algorithmic. This uses experience, judgment, strategies, tricks and rules of thumb. Such knowledge is used to solve complex problems. Knowledge management refers to the integrated approach to identifying, capturing, retrieving, sharing and evaluating enterprise information assets (database, intranets, process documents, policies, procedures etc.). Knowledge-based systems are based on expert knowledge, which has been coded into facts, rules etc.

Intelligence is the possession and capability to use knowledge. Human brain stores knowledge in neurons, which are cells, which comprise our nervous system. There are 1011-1013 neurons in our nervous system. The storage capacity of the brains is 1013 - 1015 bits. Characteristics of Intelligence include [8]:

- Respond flexibly in different situations,
- Find similarity between different situation,
- To draw distinction between situations despite of similarity,
- To make sense out of ambiguous and contradictory messages,
- To recognize relational importance of different elements of a situation,
- Learn from experience,
- Apply knowledge acquired from experience,
- Handle complex situations,
- Solve problems when important information is missing,
- Determine what is important,
- React quickly and correctly to a new situation,
- Understand visual images,
- Process and manipulate symbols,
- Be creative and imaginative,
- Use heuristics.

KNOWLEDGE REPRESENTATION TECHNIQUES

AI deploys ways to efficiently represent the knowledge and use efficient searching algorithms based on heuristics to make use of knowledge to solve complex problems. In fact, Artificial Intelligence deals with symbolic, non-algorithmic methods of problem solving [6]. Knowledge can be represented in following different forms:

- Mental images in one's thoughts,
- Spoken or written word in some language,
- Character strings,
- Graphical or other picture,
- Magnetic spots stored in a computer.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

VARIOUS AREAS OF ARTIFICIAL INTELLIGENCE

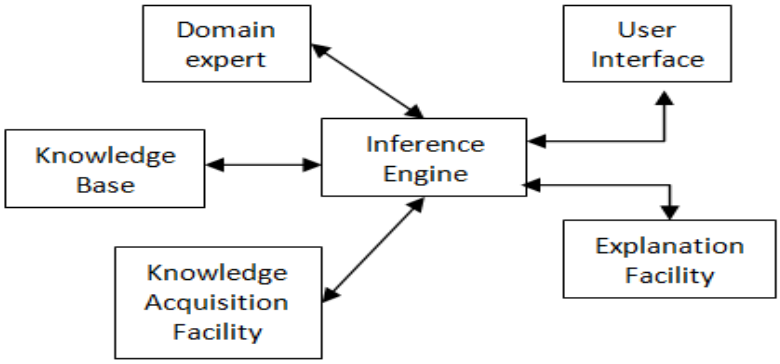
Various knowledge representation techniques available in Artificial Intelligence include Propositional Logic (PL), First Order Predicate Logic (FOPL), Frames, Conceptual Dependencies, and Scripts etc. Choice of representation scheme depends on nature of problem under consideration. Various areas of Artificial Intelligence include [10]:

Intelligent Agents: An intelligent agent is a system that perceives its environment, learns from it, and interacts with it intelligently. Intelligent agents can be divided into two broad categories: software agents and physical agents. A software agent is a set of programs that are designed to do particular tasks. For example, a software agent can check the contents of received e-mails and classify them into different categories (junk, less important, important, very important and so on). Another example of a software agent is a search engine used to search the World Wide Web and find sites that can provide information about a requested subject.

Machine Learning: Machine learning concerns the construction and study of systems that can learn from data. For example, a machine learning system could be trained on email messages to learn to distinguish between spam and non-spam messages. Machine learning programs detect patterns in data and adjust program actions accordingly. For example, Facebook's News Feed changes according to the user's personal interactions with other users. If a user frequently tags a friend in photos, writes on his walls or likes his links, the News Feed will show more of that friend's activity in the user's News Feed due to presumed closeness.

Expert Systems: An expert system is an intelligent computer program that uses human knowledge to solve the complex problems that require significant expertise. An expert system uses knowledge and inference procedures and contains modules like knowledge base, inference engine, user interface etc. as shown in Figure 1 [9].

Figure-1:Components of an Expert System



Sources: Authors Compilation

COMPUTERS IN EDUCATION

Indian higher education system has grown in a remarkable way and has become one of the largest systems in the world. The growth is in terms of number of students' enrolment, number of Universities, institutions and increasing percentage of educational expenditure. However, the country lacks the availability of teachers with appropriate teaching skills, technical skills and business skills. It is widespread belief that the quality of education, which is of paramount importance, is degrading and most of the new Professional Education Institutions are not able to produce employable graduates. The serious concern is to improve the quality of teaching-learning process to increase the percentage of graduates who meet the entry requirements for engineering and management sector [11].

The goal of any education system is to acquire and retain knowledge and skills, understand students acquired knowledge and skills and ability to apply one is learning to solve novel problems [12]. Since the invention of computers, attempts have been made to use them for educational purposes. Information Technology is being used in education in a variety of ways. This includes preparing slides for presentation, online delivery of lectures, online examinations, learning management systems etc. However, in such systems, learners' ability was not taken into account, as instructions could not be individualized as per their needs. It is also possible that specialized examinations will be conducted for a specific candidate, which will cause no additional overhead to the IT based environment. However, in e-Learning systems, although we have only digitalized traditional resources e.g. books, journals, exams etc., they are being delivered in the same traditional manner used in conventional classrooms. Use of AI-based educational systems, which incorporate rich contents of subject domain and pedagogical knowledge, are able to produce better results in learning than less sophisticated e-Learning and conventional classroom based teaching methods [15]. These benefits can be attributed to following features of AI based systems:

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

- Ability to track the "mental steps" of the learner and underlying goal structure of problem solving tasks [1].
- Diagnosis of misconceptions and estimation of the learner understands of the domain [17].
- Providing timely guidance, feedback and explanations [16].
- Promoting productive learning behaviors, such as self-regulation, self-monitoring and self-explanation [12].
- Prescribing learning activities at the right level of difficulty and with the most appropriate content [18].

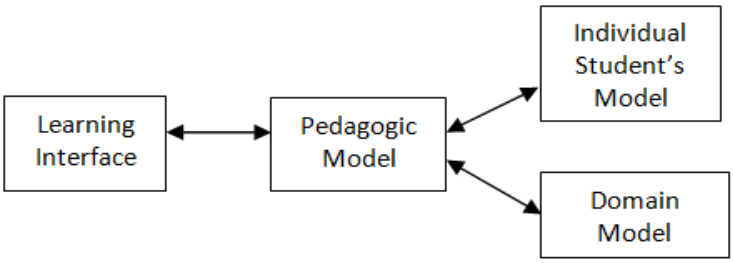
Various AI technique viz. natural language processing, machine learning, intelligent agents, uncertain reasoning, cognitive modeling, case-based reasoning etc. have been applied to achieve the above mentioned goals [19].

a) Intelligent Tutoring System

A good teacher has a knowledge of the subjects, which is delivered as per the specific needs of the students. He has ability to correlate different parts of knowledge and uses this ability to remove the gap between the students' knowledge of the domain to some "ideal" knowledge.

With the proliferation of internet and availability of powerful computers, it has become possible to use Intelligent Tutoring Systems. It is next stage of e-content development after video lectures, animations and interactive simulations. An Intelligent Tutoring Systems (ITS), as shown in Figure 2 allows one-to-one teaching by providing the problem-solving environment. The system keeps track of the students' progress and helps the student whenever required. The knowledge about the students learning characteristics may be collected from data produced by the students' participation in intelligent computer aided learning. The tutoring system is imparting individualized instructions by making minute-to-minute adjustments. Thus, intelligent tutoring systems allow students to learn both faster and better as compared to traditional classroom instruction. The domain model represents subject matter expertise and with the subject knowledge. The student model gives information about the students' knowledge. The pedagogic model enables the system to teach by encoding instructional strategies used by the learning interface.

Figure-3: Intelligent Tutoring System



Sources: Authors Compilation

Intelligent tutoring systems have been shown to be highly effective in increasing students' motivation and learning. ITSs can outperform untrained tutors and are capable of effectiveness of expert tutors. The potential to provide individualized tutoring support to every learner in every situation would revolutionize education. The idea that homework might be viewed not as required drudgery, but as an opportunity to succeed and demonstrate, understanding is compelling. Intelligent Tutoring Systems give better performance as compared to human tutors in some areas of instructions like aeroplane pilot and spaceship pilot. Real life training can put the trainee in dangerous life threatening situation. Computer simulation can help substantial transfer in learning without involving risk [2].

b) Web Based Education for All

There is a need of web based educational system that can adopt to different users from different backgrounds, learning goals and the prior knowledge levels of the relevant subjects. An adaptive textbook can be developed with the material in different forms viz. lectures, examples, tutorials, assignments, quizzes etc. Web based courses are used by wider variety of users as compared to on-campus students who are well prepared and motivated with access to teachers to resolve their doubts. Therefore, web based courses have to be richer and more flexible so that remote students may have personalized contents.

Adaptive Navigation Support (ANS) aims to help students in finding an optimal path through the hyperspace for learning material. The hypermedia interface theoretically makes it possible for every student to find his unique path through the web based learning material. The ANS also provide adaptive guidance mechanism so that educational opportunity is maximized for each type of learner. The domain model contains the pieces of knowledge for the given domain and their interrelationship. This model is also capable of measuring the students' knowledge of different topics in form of an overlay model. All the actions of the students like page visit, problem solving, answering the quizzes etc. are tracked by the system and used to adjust the knowledge level. This model also stores students learning to ensure that student achieves the learning goals in desired sequence. The system provides all

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

the navigation tools like sequential (forward or backward), hierarchical up and down) links and one click transfer to all the sections and a table of contents for clickable links. Glossary is the central part of web based educational system. Each entry in the glossary corresponds to one of the domain concept. The links between domain model concepts constitute the navigation paths between glossary entries. The web based adaptive textbook is indexed into units with domain model concepts. The indexing provides knowledge about the contents of the pages.

The design of web-based systems has grown exponentially in last few years. These systems are independent of location of students and teachers and the hardware platform used [5].

c) Expert systems for Educational Diagnosis

There are students with severe learning difficulties and teachers do not know where they are struggling. Exact difficulty of the student should be determined before planning an instructional program for the student. This system not necessarily tests the students directly. An expert system could be developed to guide the teacher through various stages of diagnosing the learning disabilities. The expert system would provide the summary of the diagnostic findings along with a prescription including appropriate instructional techniques. After the analysis of student's diagnosis, the system will suggest the next step to the teacher [7].

d) Expanding the Methods of Learning

The future concern of the AI applications in education is to implement a wide range of teaching-learning methods [4]. These methods include:

- **Collaborative Learning**: ITS helps in one-on-one coaching and has deep knowledge of both the subject and student. Most of us have had the experience of improving the clarity of our knowledge through discussions with friends who do not understand what we are doing but who can ask questions that cause us to restructure our thoughts. Using networking technology together with ITS, it is possible to develop methods for learning within and across the classrooms. However, powerful network-based technologies are now making possible novel methods of group problem solving and learning by students. The design of the collaborative ITS should be easier as the instruction does not have to be perfect and if a student becomes confused, another student may be able to help without relying on the assistance from ITS. Not all the students in the group have same ability.

- **On-demand Learning:** this method of learning and teaching helps learning in context of a work and makes learning relevant to the task. Learning on demand or "just in time" learning will be increasingly important as future workers will need to continually improve their skills.

- **Case-based Learning**: In case-based learning, students acquire expertise e.g., in military history or telephone communication skills) by accessing cases from a rich library of experiences. Thus, instead of learning abstract rules to apply to situations, students exploit the analogies found in familiar cases to synthesize their own decision rules to resolve a situation.

- **Simulation Based Learning**: One of the most effective training devices is the flight simulator. It provides large amounts of inexpensive time-on-task to perfect skills that are too costly, slow and dangerous to acquire using real aircraft. Building on this insight, researchers are now developing simulation-based learning environments for a wide range subjects.

- **Apprenticeship:** The new computer-based environments that provide some of the desirable features of traditional apprenticeship as a model of learning and teaching can be stimulated. Many of these new ideas and systems move apprenticeship mentoring activities -- modeling, coaching, and fading; from training physical skills, like tailoring, to cognitive skills, like writing, reading, or mathematics.

- **Visualization:** Visualization often permits learners and professionals to get broad pictures of data, to "generate intuitions", and to suggest hypotheses for further testing. Visualization tools exploit the computer's ability to render complex patterns succinctly in graphical representations. In some cases visualization may even be a substitute for symbolic reasoning and proof, not merely supplement it. In these areas, then, visualization tools are providing new perceptual methods to more simply accomplish activities that, previously, required symbolic reasoning. However, in other cases, visualization technologies are opening up completely new fields of study and defining new goals for curricula as well as providing new methods of learning [11].

CONCLUSION

Growth of a country depends on its ability to continuously educate the population and create the skilled manpower. Present content development and e-learning tools are general purpose and do not satisfy the specific learning needs of a student or monitor the progress of individual learner. Use of Artificial Intelligence can enable a positive change in the learning environment leading to

increased involvement of students in learning process. In this paper, some techniques and uses of Artificial Intelligence techniques in teaching-learning process have been suggested.

REFERENCES

1. Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive Tutors: Lessons Learned. The ***Journal of the Learning Sciences***, 4(2), 167-207.

2. Azevedo R. & Hadwin A.F., (2005). Scaffolding Self Regulated Learning and Meta Cognition Implications for Design of Computer [3] Based Scaffolds***. Instructional Science***, Springer Netherlands, 33(5), 367-379.

3. Barr A., Feigenbaum E. A. (1981). ***The Handbook of Artificial Intelligence,*** Vol. 1., Heuris Tech Press.

4. Brooks R.A. ***Cyber Learning: AI applications in Education***. Retrieved from http://www.ai.mit.edu/people/brooks/

5. Brusilovsky P., Eklund J. & Schwarz E. (1998). Web Based Education for All: A tool for Development of Adaptive Courseware. ***Computer Network and ISDN Systems,*** 30. 291-300.

6. Buchanan B. G. & Shortliffe E. H. (1984). ***Rule Based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project (The Addison-Wesley series in artificial intelligence)***. Addison-Wesley Longman Publishing Co., Inc.

7. Hatzilygeroudis I., Chountis P., Christos Giannoulis C. & Koutsojannis C. ***Using Expert Systems Technology for Student Evaluation in a Web Based Educational System***. Retrieved from http://www.hci.gr/files/461-832_hatziligeroudis.pdf

8. Hofstadter D. (1999). ***Gödel, Escher, Bach: An Eternal Golden Braid***. Basic Books, 20 Anv Edition.

9. Khanna S., Kaushik A. & Barnela M. (2010). Expert Systems Advances in Education, National Conference on Computational Instrumentation (NCCI 2010) ***CSIO Chandigarh***, India, 19-20.

10. Marlene J. (1985). Applications of Artificial Intelligence in Education. ***Comp. & Maths with Appl***. 11(5), 517-526.

11. Perkins D. (1992). ***Smart schools: Better thinking and learning for every child***. NY: The Free Press.

12. Rich E. & Knight K. (1999). ***Artificial Intelligence***. Tata McGraw Hills.

13. Saikumar M, (2012). Role of Artificial Intelligence in Enhancing the Effectiveness of Education. ***CSI Communications***, 12-16.

14. Vanderbilt T. (2012). ***How Artificial Intelligence Can Change Higher Education***. Retrieved from Smithsonian http://www.smithsonianmag.com

15. Shute V. J. (2008). VJ Shute Intelligent Tutoring Systems: Past, Present and Future. ***Review of educational research***, 78 (1), 153-189,

16. VanLehn K. (2006). The Behavior of Tutoring Systems. ***International Journal of Artificial Intelligence in Education***, Vol. 16(3), pp. 227-265.

17. Woolf B. P., (2008). Building Intelligent Interactive Tutors: Student-centered strategies for revolutionizing e-Learning. Morgan Kaufmann. First Edition.

18. Retreived From http://www.etc.edu.cn/%E8%AE%BA%E8%91%97%E9%80%89%E6%91%98/%E8%8B%B1%E6%96%87%E6%96%87%E7%AB%A0/AIEA...

19. Retreived From http://cyberlearning.sri.com/w/index.php/Cyberlearning:AI_applications_in_education

20. Retreived From http://dictionary.reference.com/browse/correspond+to

21. Retreived From http://www.csus.edu/indiv/s/scanland/mis101/index_htm_files/AI-ES-Handouts.ppt

22. Retreived From http://www.cse.yorku.ca/~mack/1010/1010-Chapter11.ppt

23. Retreived From http://seminarinfo29.blogspot.in

24. Retreived From http://evosys.biz/used-technology

25. Retreived From http://www.youtube.com/watch?v=mB1cbOv7qco

26. Retreived From http://en.wikipedia.org/wiki/Machine_learning

27. Retreived From http://www.machinelearningtechniques.com/index.php/using-machine-learning/extensions/components/cont...

28. Retreived From http://whatis.techtarget.com/definition/machine-learning

29. Retreived From http://www.ent.mrt.ac.lk/~ekulasek/en577/Applications%20of%20AI%20in%20Education.htm

30. Retreived From http://link.springer.com/article/10.1023/A%3A1021258506583

31. Retreived From http://www.bpminstitute.org/resources/articles/business-rules-going-old-school-expert-systems-0

32. Retreived From http://richardoconnell.wordpress.com/2011/01/27/artificial-intelligence-in-e-learning-systems

****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

CONTROLLING OF INVERTED PENDULUM SYSTEM USING MPC AND PID APPROACH

Manisha Pandey[28] Harshita Joshi[29]

ABSTRACT

Model predictive control has become a major research topic during the last few decades and it has been successfully applied in industry. Inverted Pendulum is a classical example for control engineers to verify and apply different logics in the field of control theory. The pendulum is stable while hanging downwards, but the inverted pendulum is inherently unstable and need to be balanced. The balancing of an inverted pendulum by moving a cart along a horizontal track is a classic problem in the area of control. In this paper, model predictive control is applied to an inverted pendulum system. In this paper PID controller is tuned to control the inverted pendulum system and advance MPC controller is used to control the inverted pendulum system. By comparing both controllers, we can see that advance MPC controller is better as compare to conventional PID controller.

KEYWORDS

Model Predictive Controller, PID Controller, Inverted Pendulum, Tuning etc.

INTRODUCTION

The inverted pendulum is a classical model of how the use of control can be working to stabilize an essentially unstable system. The inverted Pendulum system represents also a perfect model pitch, yaw behaviors of a flying rocket, and can be used as a benchmark for numerous control methodologies. The segway PT is two wheeled, self-balancing vehicle that transports a single person, which uses the properties of Inverted Pendulum [1].

Inverted pendulum is a very good stage for control engineers to prove and apply different logics in the field of control theory. Most of up to date technologies use the necessary idea of Inverted Pendulum, such as attitude control of space satellites and rockets, landing of aircraft, balancing of ship against tide, seismometer etc. [2].

An Inverted Pendulum has its mass on top of the pivoted point, which is mounted on a cart, which can be move horizontally. The Pendulum is stable while hanging downwards, although the inverted pendulum is naturally unstable and need to be balanced. In this case, the system has one input, the force applied on cart and two inputs, position of the cart and angel of Pendulum making it as a SIMO system. There are mainly three ways of balancing an Inverted Pendulum:

-
- By applying torque at the pivoted point
- By moving the cart horizontally
- By rapidly the support up and down.

The Inverted Pendulum systems forever show numerous problems offered in industrial applications, for instance, a variety of nonlinear behaviors under different operations conditions, external disturbances, and physical constraint on some variables. Therefore, the assignment of real stabilization and tracking control of a highly nonlinear unstable inverted pendulum systems have been intensively studies due to the challenging demand of fast accurate performance [3].

It is recognized that an inverted pendulum and cart (IPC) system is a complex nonlinear system. Two different dynamics of the pendulum and the cart are attached mutually. The friction between the cart and rail is also nonlinear. In addition, there exist constraints on the most control actions and the length of the rail. Therefore, an IPC system is often utilized to represent the effectiveness of different class of controllers.

Due to high nonlinearity and critical sensitiveness to the external disturbances, many investigations have been carried out into the control problem of an IPC system [4].

The dynamics of balancing a pendulum at the unstable position know how to be employed in the applications of controlling walking robots, rocket thrusters, etc. Many control design techniques have been used to examine the control behaviors of an inverted pendulum. The successful application of classical controller design techniques required considerable knowledge of the accurate system dynamic model and desired system behaviors with the expression of an objective function. However, the mathematical model derived from physical relationships or identified from experimental results is only an approximated model.

[28] Research Scholar, Department of Electrical & Electronics Engineering, Graphic Era University, Uttarakhand, India, er.manisha.pandey009@gmail.com

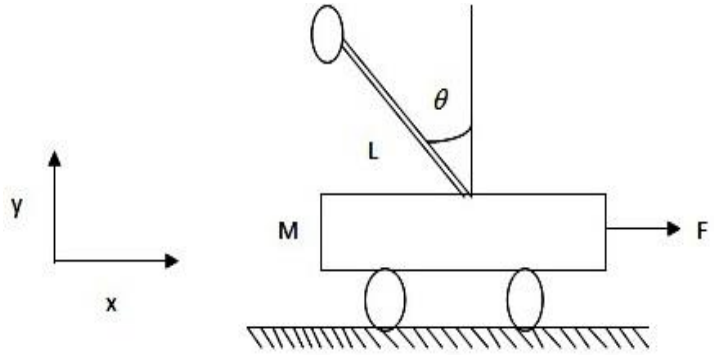
[29] Faculty, Department of Electrical & Electronics Engineering, Graphic Era University, Uttarakhand, India, harshita.joshi0408@gmail.com

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Generally, it is a local linearized result. If the operating range is too large, the control performance of a traditional controller is not satisfactory. Hence, the highly nonlinear and unstable features of an inverted pendulum are the behaviors of the traditional controller, which are difficult to overcome [5].

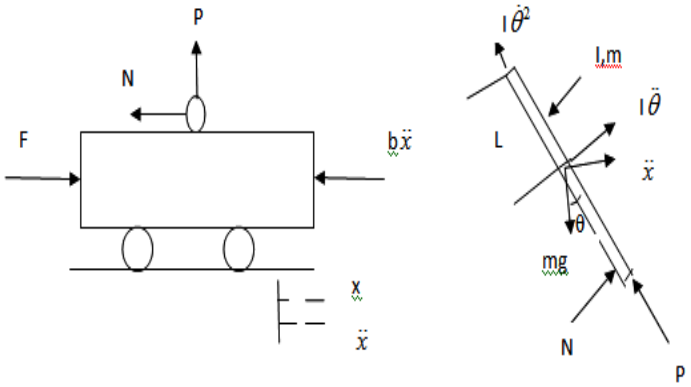
MATHEMATICAL MODELING OF INVERTED PENDULUM

Figure-1: Inverted Pendulum System



Sources: Authors Compilation

Figure-2: Free Body Diagram of Cart and Inverted Pendulum



Sources: Authors Compilation

$$\frac{d^2x}{dt^2} = \frac{1}{M}(F - N - b\frac{dx}{dt})$$

$$\frac{d^2\theta}{dt^2} = \frac{1}{I}(NL\cos\theta + PL\sin\theta)$$

$$m\frac{d^2x_p}{dt^2} = N$$

$$m\frac{d^2y_p}{dt^2} = P - mg$$

$$P = m(\frac{d^2y_p}{dt^2} + g)$$

On the other hand, x_p and y_p are accurate functions of theta. As a result, we can represent their derivatives in terms of the derivatives of theta

$$x_p = x - L\sin\theta$$

$$\frac{dx_p}{dt} = \frac{dx}{dt} - L\cos\theta\frac{d^2\theta}{dt^2}$$

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

$$y_p = L\cos\theta$$

$$\frac{d^2x_p}{dt^2} = \frac{d^2x}{dt^2} + L\sin\theta(\frac{d\theta}{dt}) - L\cos\theta\frac{d^2\theta}{dt^2}$$

$$\frac{dy_p}{dt} = -L\sin\theta\frac{d\theta}{dt}$$

$$\frac{d^2y_p}{dt^2} = -L\cos\theta\frac{d^2\theta}{dt^2} - L\sin\theta\frac{d^2\theta}{dt^2}$$

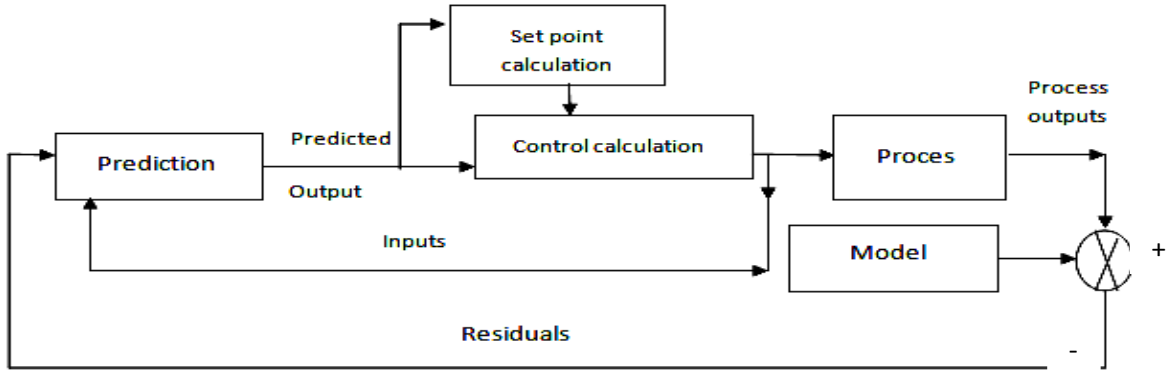
MODEL PREDICTIVE CONTROLLER

Model predictive control (MPC) is an attractive feedback approach, particularly for linear or nonlinear systems subject to input and state constraints, referred to as moving horizon control or receding horizon control. In most cases, linear and nonlinear MPC are distinguished. In the process industries, linear MPC approaches have found successful applications. The achievement of MPC is because it is maybe the most general way of posing the control problem in the time domain. The use a finite-horizon strategy allows the explicit handling of process and operational constraints by the MPC [6].

Three major aspects of model predictive controller are as follows:

- Design formulation.
- The ability of method to handle both soft constraints and hard constraints in a multivariable control system.
- The ability to perform process online optimization.

Figure-3: Block diagram of MPC



Sources: Authors Compilation

In this case, the system has one input, the force applied to the cart and one output, position of the cart making it as a SIMO system.

In MPC, there are two horizon control horizon and predictive horizon. By adjusting these two variables, we get the response of the inverted pendulum.

PID CONTROLLER

In industrial control systems, a proportional –integral –derivative controller (PID) is a general control loop feedback mechanism. PID controllers can calculate an 'error' value, which is the difference between a measured process variable and a desired set point. By adjusting the process, control inputs the controller attempts to minimize the error.

The PID controller calculation algorithm involves three separate constant parameters and is accordingly sometimes called three-term control: the proportional, the integral and derivative values, denoted *P*, *I*, and *D*. Based on current rate of change these value can be interpreted in terms of time: P depends on the *present* error, *I* on the accumulation of *past* errors, and *D* is a prediction of *future* errors. The weighted summation of these three actions is used to adjust the process via a control element such as the position of a control valve, a damper, or the power supplied to a heating element [7].

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

$$u(t) = MV(t) = K_P e(t) + K_i \int_0^t e(\tau)\,\mathrm{d}\tau + K_d \frac{d}{dt} e(t)$$ K_P =Proportional gain

K_i =Integral gain

K_d =Derivative gain

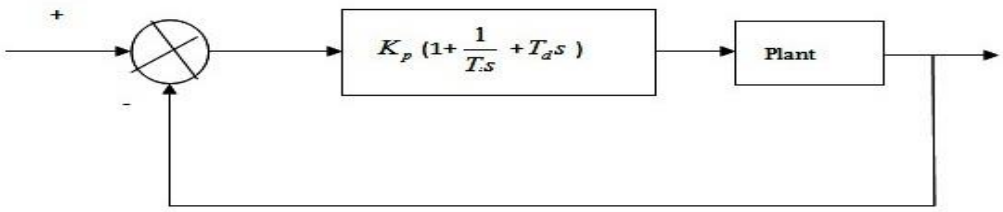
e =Error

t =Time or instantaneous time

τ =Variable of integration

ZIEGLER-NICHOLS TUNING METHOD

Figure-4: PID Controller of Plant



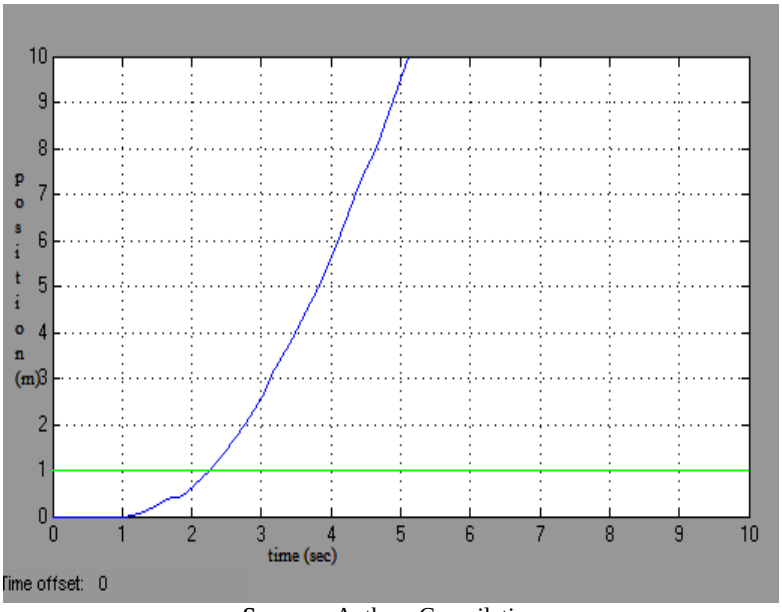
Sources: Authors Compilation

A PID control of a plant shows in figure1.4. To apply various design techniques for determining parameters of the controller that will meet the transient and steady state specifications of the closed-loop system a mathematical model of the plant can be derived. Its mathematical model cannot be easily obtained if the plant is so complicated, then an analytical or computational approach to the design of a PID controller is not possible. Then we must resort to experimental approaches to the tuning of PID controllers.

Type of Controller	K_p	K_i	K_d
PID	60	150	6

SIMULATION RESULTS

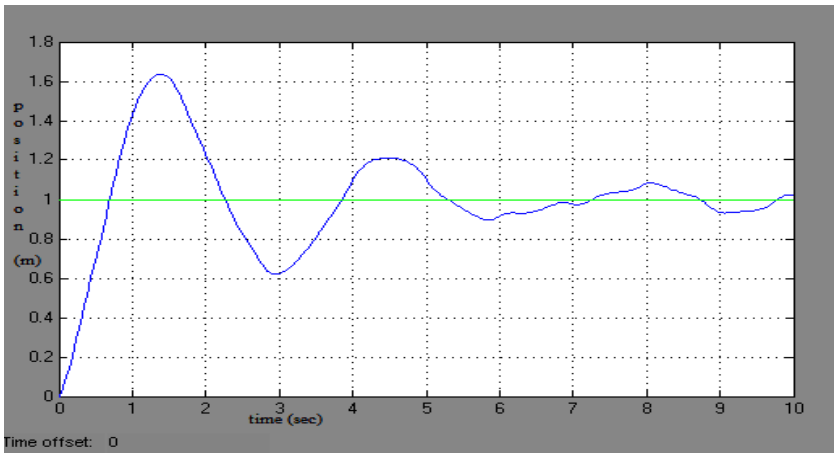
Figure-5: Response of Inverted Pendulum without Controller



Sources: Authors Compilation

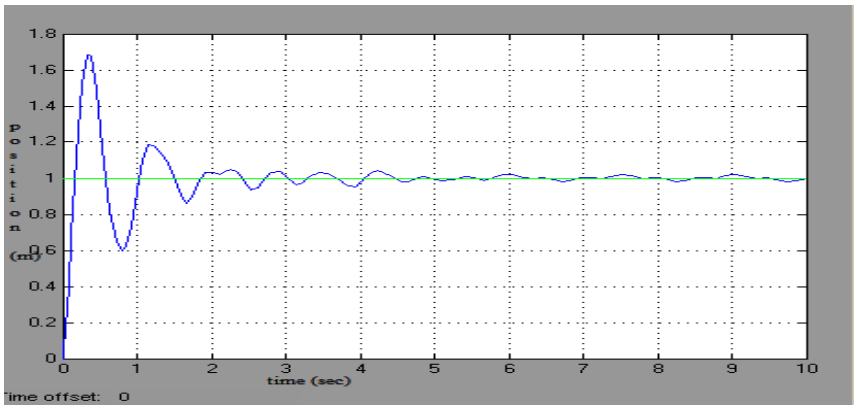
Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Figure-6: Response of Inverted Pendulum with PID Controller



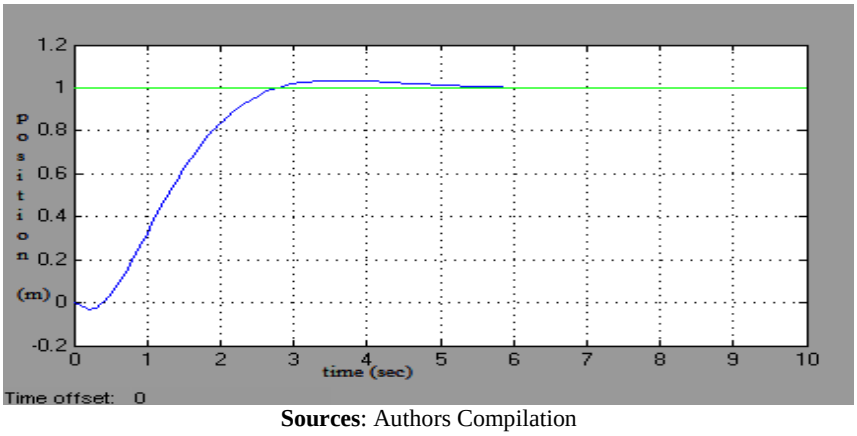
Sources: Authors Compilation

Figure-7: Response of Inverted Pendulum with PID Controller by Tuning Method



Sources: Authors Compilation

Figure-8: Output Response of Inverted Pendulum with MPC Controller



Sources: Authors Compilation

From figure-5 the output response of inverted pendulum is not, stable i.e. system is unstable and it does not follow the reference line. From figure-6, the output response of inverted pendulum is stable but in this case, the oscillation is not dying out. By tuning method from figure-7 we can see that the output response of inverted pendulum is stable and it follow the reference line and oscillation are die out in this case. By using MPC controller, we can see that the output response of inverted pendulum is stable and it follow the reference line.

CONCLUSIONS

By comparing above results, we can conclude that the output response of advance controller i.e. MPC is better as compare to PID controller as by using MPC controller the stable graph was obtained.

REFERENCES

1. Honjo, T., Luo, Z., & Nagano, A. (2008). Parametric Excitation of a Biped Robot as an Inverted Pendulum. ***In IEEE/RSJ International Conference on Intelligent Robots and Systems, (pp. 3408-3413)***.

2. Anderson, C. W. (1989, April). Learning to Control an Inverted Pendulum Using Neural Network. ***IEEE Control syst.Mag.***, 9(3), 31-37.

3. Brock, S. Practical Approach To Fuzzy Control Of Inverter Pendulum. ***In Proc. ICIT, Slovenia, (pp. 31–35)***.

4. Ishida, T., Shiokawa, N., Nagado, T., & Ganeko, S. (1991). Learning control of an inverted pendulum using a neural network. ***In Proc. IEEE Conf. Industrial Electronics, Control and instrumentation, (volume 2, pp. 1401–1404).***

5. Abu-Ayyad, M., & Dubey, R. (2007). Real Time Comparison of a Number of Predictive Controllers. ***ISA Transactions,*** 46, 411-418.

6. Katsuhiko, Ogata. ***A textbook of Modern Control Engineering***. University of Minnesota. ISBN 978-81-317-0311-3.

7. Retrieved from http://ttuadvancedrobotics.wikidot.com/pid-controller

8. Retrieved from http://sca-ipaindia.com/pid.php

9. Retrieved from http://en.wikipedia.org/wiki/Inverted_pendulum

10. Retrieved from http://www.ripublication.com/aeee_spl/aeeev4n4spl_02.pdf

11. Retrieved from http://sca-ipaindia.com/pid.php

12. Retrieved from http://wn.com/Inverted_Pendulum_using_a_PID_controller

13. Retrieved from http://www.scribd.com/doc/35215743/Software-PID-Control-of-an-Inverted-Pendulum-Using

****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

REVIEW OF ROUTING PROTOCOLS

Surbhi Siwach[30]

ABSTRACT

Wireless communication is rapidly growing in our day-to-day life because it is easy to deploy and is more flexible. In particular, the wireless sensor networks (WSN's) is one of the latest innovation in the field of wireless communication. It consists of tiny mote, which hold a small battery, CPU and sensor. Underwater wireless sensor networks (UWSN), similar to the terrestrial sensor networks, have different challenges such as limited bandwidth, low battery power, defective underwater channels, and high variable propagation delay. A crucial problem in UWSN is finding an efficient route between a source and a destination. Consequently, great efforts have been made for designing efficient protocols while considering the unique characteristics of underwater communication. Several routing protocols are proposed for this issue and can be classified into geographic and non-geographic routing protocols.

In this paper, we focus on the geographic routing protocols. We introduce a review and comparison of different algorithms proposed recently in the literature. We also presented a novel taxonomy of these routing in which the protocols are classified into three categories (greedy, restricted directional flooding and hierarchical) according to their forwarding strategies. Major routing protocols are taken into account and each protocol is described in detail. Taxonomy of the routing protocols is presented where the protocols are categorized into four categories i.e. flooding based protocols, multipath based, cluster based and miscellaneous protocols.

KEYWORDS

Underwater Sensor Networks (UWSN), Routing Protocols, Radio Signals, Quality of Service (QoS) etc.

INTRODUCTION

Underwater sensor networks are becoming popular every day due to their unique characteristics that make them easy to deploy in areas where adverse environmental conditions exist. Underwater sensor nodes play a major role in pollution monitoring in environmental systems, remote control in the offshore oil industry, collection of scientific data that is observed at ocean-bottom stations, navigational assistance and surveying ocean floor to search for new resources, disaster prevention and tactical surveillance applications. Under water sensor networks help in undersea explorations like detect underwater oilfields, determine routes for laying undersea cables and environmental monitoring like monitoring of ocean currents and winds, for improved weather forecast, biological monitoring such as tracking of fishes or micro-organisms, disaster prevention, measure seismic activity from remote locations provide tsunami warnings to coastal areas, in distributed tactical surveillance, auvs and fixed underwater sensors collaboratively monitor areas for surveillance, reconnaissance, targeting, and intrusion detection etc.

Example of use of underwater sensor network are distance of epicenter of recent earthquake in japan from sendai = 128.74752 km, to measure the speed i.e. @ 970 km/h (speed of a wave) the tsunami takes 7 minutes 59 seconds to reach the sendai coast, information about the tsunami can be transmitted to coastal warning centers at the speed of sound – takes 1 min 26 seconds. The purpose of underwater sensor network is need for monitoring ocean environment for scientific, commercial and military purposes. Communication between various underwater sensors and vehicles is required. Sensors are deployed in remote locations and collect data over a long period. Amount of data is large and is expected to grow as sensor technology becomes more complex.

Need of acoustic communication is in traditional wireless protocols: use electromagnetic waves; suffer from scattering and attenuation problems underwater. Radio waves propagate at long distances through conductive salty water only at extra low frequencies (30−300hz) requiring large antennae and high transmission power. Berkeley mica2 motes, a popular experimental platform in the sensor networking community, have been reported to reach an underwater transmission range of 120 cm at 433 MHz in experiments performed at the University of Southern California. Optical waves do not suffer from such high attenuation but are affected by scattering. Furthermore, transmitting optical signals requires high precision in pointing the narrow laser beams. Thus, links in underwater networks are typically based on acoustic wireless communications.

The Earth is a water planet. About 2/3 of the Earth is covered by oceans, resulting to a huge amount of resources to discover underwater. Many applications such as Ocean sampling networks, Environmental monitoring, undersea explorations, Disaster prevention, Mine reconnaissance etc. There is largely unexplored area and recently humans are showing interest towards exploring it, Underwater Acoustic Sensor Networks (UW-ASN) consist of a variable number of sensors that are deployed to perform the monitoring tasks over a given area.

[30]M. Tech (Final Year), CBS Group of Institutions, Fatehapuri-Jhajjar, India, surbhisiwach504@gmail.com

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

WHAT IS UNDERWATER SENSOR NETWORK?

We can call underwater sensor network simply a network that's primary portion is built underwater.
It consists of a number of underwater sensor nodes, underwater sink, surface station, Autonomous Underwater Vehicles (AUVs) that are deployed to perform collaborative monitoring and resource exploration tasks over a given area.

UNDERWATER SENSOR NETWORK – WHY?

Underwater sensor networks are used for any purposes: a) Undersea exploration, b) Seismic monitoring, c) Pollution monitoring, d) Equipment monitoring and control, e) Autonomous Underwater Vehicles (AUV), and f) Mine reconnaissance.

CHALLENGES OF UWSN

Challenges of acoustic network are severely limited bandwidth, Underwater channels are severely impaired, especially due to multipath and fading; Propagation delay is five orders of magnitude higher than in Radio Frequency (RF) terrestrial channels, High bit error rates and temporary losses of connectivity (shadow zones) can be experienced; Battery power is limited and usually batteries cannot be recharged, also because solar energy cannot be exploited, Power absorption and attenuation. Underwater sensors are prone to failures because of fouling and corrosion.

- Long and variable propagation delay,
- The available bandwidth is severely limited,
- Doppler Spread,
- Multipath,
- Noise and high path loss,
- Limited Battery power (this is why Low radio frequency is not used in underwater transmission).

Why sound is using as communication medium in UW-SN?

- High Radio wave is extremely attenuated in salt water.

Therefore, we can Use low radio frequency right? However, Low radio frequency requires a very large antenna and high transmission power. Therefore, not practical at water medium. In addition, the optical waves are not efficient in underwater environments because they may be scattered.

2D UNDER WATER SENSOR NETWORK

A group of sensor nodes is anchored to the bottom of the ocean. Underwater sensor nodes are interconnected to one or more underwater sinks (underwater-sink). UW-sink is used to receive data from the ocean bottom network and send to a surface station. Underwater-sinks are equipped with two acoustic transceivers. A vertical and a horizontal transceiver.

3D UNDER WATER SENSOR NETWORK

In three-dimensional underwater networks, sensor nodes float at different depths. The depth of the sensor can then be regulated by adjusting the length of the wire that connects the sensor to the anchor

ROUTING IN UWSN

Conventional Proactive and Reactive routing does not work efficiently in the networks with high propagation delay and high dynamic topology such as UWSN. For example, a path from each node to other nodes is discovered and stored in the route table. This path is expired in a short period due to high movement of nodes. We can divide routing protocols into two categories:

Location-Based (Geographical) Routing

The location information-based routing algorithm uses location information to guide routing discovery and maintenance as well as data forwarding, enabling directional transmission of the information and avoiding information flooding in the entire network. Consequently, the control overhead of the algorithm is reduced, and routing is optimized. Moreover, with network topology based on nodes' location information, network management becomes simple and global network optimization can be easily achieved.

Vector Based Forwarding Protocol (VBF)

In vector Based Forwarding, data packets are forwarded along redundant and interleaved paths from the source to sink. This helps in handling the problems of packet losses and node failures. The routing vector from sender to target nominates forwarding path.

Each of the sender's neighboring nodes determines its candidacy to be the next relay node. Handles node mobility efficiently the location information of each sensor node can be obtained through a localization service. The energy of the network is saved because only the nodes that come across the forwarding path are involved in packet routing. It is assumed that every node already knows its location and each packet carries the location of all nodes involved. The forwarding path is virtually a routing pipe and the nodes inside this pipe are eligible for packet forwarding. In VBF, each packet carries the positions of the sender, the target and the forwarder (i.e., the node that transmits this packet). The forwarding path is specified by the routing vector from the sender to the target. Upon receiving a packet, a node computes its relative position to the forwarder by measuring its distance to the forwarder and the angle of arrival (AOA) of the signal2. Recursively, all the nodes receiving the packet compute their positions. If a node determines that it is close to the routing vector enough (e.g., less than a pre defined distance threshold), it puts its own computed position in the packet and continues forwarding the packet; otherwise, it simply discards the packet.

Focused Beam Routing Protocol (FBR)

Let us assume that node A wants to transmit to node B. To do so, node A will issue a request to send (RTS) to its neighbors. This request is a short control packet that contains the location of the source node (A) and of the final destination (B). Note that this is in fact a multicast request. The initial transaction is performed at the lowest power level and the power is increased only if necessary. Power control is an integral part of routing and medium access control. We assume open loop power control, in which the transmitting node decides which power level to use, rather than being instructed explicitly by a receiving node.

Depth Based Routing (DBR)

DBR well utilizes the general underwater sensor network architecture: data sinks are usually situated at the water surface. In DBR, a data packet has a field that records the depth information of its recent forwarder and is updated at every hop. The basic idea of the DBR is as follows: when a node receives a packet, it forwards the packet if its depth is smaller than that embedded in the packet. Otherwise, it discards the packet. The main advantages of DBR are it does not require full-dimensional location information, it can handle dynamic network with good energy efficiency, and it takes advantages of multiple sink network architecture without introducing extra cost.

Hop-by-hop Vector Based Forwarding (HH-VBF)

HH-VBF is the successor of the VBF protocol, which is a well-known routing protocol proposed for underwater sensor networks. In VBF, the flooding is performed in a constrained virtual routing pipe where the radius of the virtual routing pipe is predetermined (a certain threshold). The radius of the routing pipe covers the area around a vector. A vector is a virtual line from a source node towards the destination. A node participates in the routing when its distance from the vector is smaller than the radius of the virtual routing pipe. In HH-VBF the routing vector from each forwarder/sender towards the destination is computed. Upon the receipt of a packet, a node computes a vector starting from the transmitter of the packet towards the destination. Then the node calculates the distance between the computed vector and itself. In case, the distance between the vector and the node is smaller than the radius of virtual routing pipe, this node becomes eligible for forwarding and becomes a candidate forwarder. The candidate forwarder holds the packet for a particular time (holding time) before forwarding it. The holding time is based on a desirableness factor. A desirableness factor illustrates the suitableness of a node for forwarding. The desirableness factor is based on the distance between the transmitter and the receiver of the packet as well as the angle formed by the vectors: the vector starting from the transmitter towards the destination and the vector from the transmitter towards the receiver of the packet. The HH-VBF scheme, where A, B and C are the source nodes and S is the sink/destination node. The rectangles illustrate the areas covered by the virtual routing pipe of each node.

The need to overcome two problems encountered by the VBF, i.e. small data delivery ratio in sparse networks, and sensitivity to the routing pipe's radius, the HH-VBF (hop-by-hop VBF) was proposed. HH-VBF forms the routing pipe in a hop-by-hop fashion, enhancing the packet delivery ratio significantly. Although it is based on the same concept of routing vector as VBF, instead of using a single virtual pipe from the source to the sink, it defines a different virtual pipe around the per-hop vector from each forwarder to the sink. In that way, each node can adaptively make packet-forwarding decisions based on its current location.

This design can directly bring the following benefits:

- Since each node has its own routing pipe, the maximum pipe radius is the transmission range. In other words, there is no necessity to increase the pipe radius beyond the transmission range in order to enhance the routing performance.
- In sparse networks, though the number of eligible nodes may be small, HH-VBF can find a data delivery path as long as there exists one in the network.

Sector Based Routing with Destination Location Prediction (SBR-DLP)

The SBR-DPL is a location-based routing algorithm, in which a sensor node neither carries any information about its neighboring nodes (including their movements) nor about the network topology. Each node is aware of its own position and the destination

node's pre planned movements. All other nodes are also aware of the destination's fixed location. Still the fact that the destination node may deviate from its schedule due to the ocean currents cannot be neglected. More importantly, a node routes a packet to the destination in a hop-by hop fashion, instead of finding the complete path before sending a packet. This is in contrast to some other applications, where the destination node can be fixed on the water surface acting as a gateway or a sink, and is in turn connected to a high-speed backbone. SBR-DLP is different from both VBF and HH-VBF. Here, it is not the candidate node that decides whether it should relay the packet but the sender who determines its next hop using information received from candidate nodes.

Directional Flooding Based Routing (DFR)

DFR protocol employs scoped flooding for the transmission of the data packets. The packets are transmitted in a restricted flooding zone where the zone area is selected based on an angle formed by the vectors: the vector between the receiver and the sender of the vector between the receiver and the destination node. In the proposed approach the geographical information is available to all nodes i.e. the location information of the neighboring nodes and the destination nodes.

Location-Free (Non-geographical) Routing

Hop-by-Hop Dynamic Addressing-Based Routing (H2-DAB)

It is a location free routing protocol. *H^2-DAB* is composed of two phases: assigning dynamic address to sensor nodes and data delivery. In the first phase, all initial hop ID for sensor is equal to 99. Then sinks start to send hello packet and each node that receives the hello packet should update its hop ID according to the number of hops to the sink. Therefore, closer sensors to sinks have smaller hop ID. For instance, the hop ID of node N11 is equal to 34 that indicate its hop distance from one sink is equal to 3 while its distance to another sink is equal to 4. In the second phase, the data is send to the sinks. Each forwarder node sends an inquiry request message to its neighboring nodes. Nodes in the communication range receive the message and send a reply message contains their node ID and their hop ID. The forwarder node selects a node with smallest hop ID as a next hop node. Hop ID is updated after an interval of time due to movement of nodes. In this way data is send to sink.

The flooding based approach is employed along with the utilization of unique IDs of the sensor nodes. The proposed protocol H2-DAB (hop-by-hop dynamic addressing based protocol) assigns an ID to each sensor node. The unique ID (called a hop ID) illustrates the distance (in terms of hop count) from a sink node towards the sensor node. In H2-DAB protocol, multi sink architecture is taken into account where the transmitted packets are considered delivered to the destination if any of the sinks receives the packet correctly. The proposed approach works as follows. Initially, all the sink nodes broadcast hello packets. The nodes receiving the hello packets are assigned the hop ID. The hop ID of a node consists of a two-digit number where the first digit denotes the distance (in terms of number of hops) from a sink node towards the current node. Similarly, the second digit denotes the number of hops from another sink node. Hence, each intermediate node keeps a record of maximum of two sinks nodes. For the transmission of data packets, the source node broadcasts an inquiry packet. Upon receiving the inquiry packet, neighboring nodes reply with an inquiry reply packet. The inquiry reply packet consists of the assigned hop ID and the address of the neighboring node. A neighboring node with the smallest hop ID is selected as a data forwarder. In case, the source node receives no reply, after the transmission of the inquiry packet, from its upper layer nodes (the nodes existing above the source node), the source node waits for a predefined time and retry.

Distributed Underwater Clustering Scheme (DUCS)

DUCS is an adaptive self-organizing protocol that forms clusters. It is considered that there are always data to be sent to the sink by the underwater sensor nodes and that power control can be used to adjust the transmission power. DUCS tries to be adapted to the intrinsic properties of underwater environments, such as long propagation delays, low data rates and difficulty of synchronization. DUCS compensates the high propagation delays of the underwater medium using a continually adjusted timing advance combined with guard time values to minimize data loss and maintain communication quality.

Table-1: Comparison of Routing Protocols of UWSN

Protocol / Architecture	Hop-By-Hop / End-To-End	Clustered / Single Entity	Single / Multi-Sink	Delivery Ratio	Energy Efficiency	Approach Used	Localization Need
Vbf	End-To-End	Single Entity	Single-Sink	High	High	Vector Based	Yes
Fbr	Hop-By-Hop	Single Entity	Multi-Sink	Medium	Medium	Vector Based	Yes
Dbr	Hop-By-Hop	Single Entity	Multi-Sink	High	Medium	Depth Based	Partially
Hh-Vbf	Hop-By-Hop	Single Entity	Single-Sink	High	Medium	Vector Based	Yes
Sbr-Dlp	Hop-By-Hop	Single Entity	Single-Sink	Medium	High	Vector Based	Yes
Ducs	Hop-By-Hop	Clustered	Single-Sink	Medium	High	Cluster Based	No
Dfr	Hop-By-Hop	Single Entity	Single-Sink	Medium	Medium	Vector Based	Yes
Hh-Dab	Hop-By-Hop	Single Entity	Multi-Sink	Medium	High	Addressing Based	No

Sources: Authors Compilation

CONCLUSION

Various approaches can be used for underwater sensor networks. The reliability and efficiency of an underwater sensor network can be increased by comparing the characteristics of different routing protocols and choosing the best protocol for a specific application in underwater environment. In this paper, different routing protocols are described with their pros and cons. It is important to choose the perfect routing protocol for underwater environment due to which we can find the best results from the particular application.

REFERENCES

1. Bayrakdar, Yonca, Meratnia, Nirvana, & Kantarci, Aylin. (2011). ***A Comparative View Of Routing Protocols For Underwater Wireless Sensor Networks.*** 978-1-61284-4577-0088-0/11/$26.00,2011IEEE.

2. Nicolaout, Nicolas, Cui, Jun-Hong, & Maggiorini, KDario. (2007). ***Improving the Robustness of Location-Based Routing for Underwvater Sensor Networks***. 1-4244-0635-8/07/$20.00, IEEE.

3. Ayaz, Muhammad, & Abdullah, Azween. ***Hop-by-Hop Dynamic Addressing Based (H2 DAB) Routing Protocol for Underwater Wireless Sensor Networks***. 978-0-7695-3922-5/09 $26.00,2009 IEEE,DOI 10.1109/ICIMT 2009.70

4. Domingo, Mari Carmen, & Prior, Rui. (2007). Design and Analysis of a GPS-free Routing Protocol for Underwater Wireless Networks in Deep Water. ***In Proceedings of the International Conference on sensor Technologies and applications***, 0-7695-2988-7/07 $25.00 2007.40

5. Akyildiz, Ian F., Pompili, Dario, & Melodia, Tommaso. (2006, September 25). ***State of the Art in Protocol Research for Underwater Acoustic Sensor Networks***. WUWNet06, Los Angeles, California, USA, ACM 1-59593-484-7/06/0009 ...$5.00.

6. Jae, Hoon Ko, Jongmin, Shin, Soonmok, Kwon, & Cheeha, Kim. ***Localization of Sensor Nodes in Underwater Acoustic Sensor Networks***. Pohang University of Science and Technology (POSTECH) San 31 HyoJa-Dong, Nam-Gu, Pohang, 790-784, Korea.

7. Liang, Wei, Yu, Haibin, Liu, Lin, Li, Bangxiang, & Che, Chang. (2007). Information-Carrying baed Routing Protocol for Underwater Acoustic Sensor Network. ***In Proceedings of the International Conference on Mechatronics and Automation ICMA 2007, (pp.729-734).***

8. Hwang, Daeyoup, & Kim, Dongkyun. (2008). DFR: Directional Flooding-Based Routing Protocol for Underwater Sensor Networks. ***IEEE OCEANS,*** 1-7.

9. Xie, P., Cui, J. H., & Lao, L. (2006). VBF: Vector-Based Forwarding Protocol for Underwater Sensor Networks. ***IFIP Networking,*** 1216-1221.

10. Wang, Pu, Li, Cheng, & Zheng, Jun. (2007). Distributed Minimum-Cost Clustering Protocol for Under Water Sensor Networks (UWSN). ***In Proceedings of the IEEE International Conference on Communications ICC'07, (pp. 3510-3515***).

11. Sharma, Aman, & H. Abdul Gaffar. Survey on Routing Protocol for Underwater Sensor Network. ***International Journal of Computer Science and Communication Network,*** 2(1), 74-82.

12. Heidemann, John, Li, Yuan, Syed, Affan, Wills, Jack, & YE, Wei. ***Underwater Sensor Networking: Research Challenges and Potential Applications***. USC/ISI Technical Report ISI-TR-2005-603.

13. Bayrakdar, Yonca, Meratnia, Nirvana, & Kantarci, Aylin. (2011). ***A Comparative View Of Routing Protocols For Underwater Wireless Sensor Networks***. 978-1-61284-4577 00880/11/$26.00,2011IEEE.

14. Sozer, E. M., Stojanovic, M., & Proakis, J. G. (2000, January). Undersea Acoustic Networks. ***IEEE Journal of Oceanic Engineering. OE,*** 25(1), 72–83.

15. Retrieved from http://www.readperiodicals.com/201012/2198228771.html

16. Retrieved from http://www.ijrte.org/attachments/File/v2i2/B0639052213.pdf

17. Retrieved from http://www.ijera.com/papers/Vol2_issue4/FD24967971.pdf

18. Retrieved from http://arxiv.org/abs/1403.3779

19. Retrieved from http://www.getcited.org/pub/103529715

20. Retrieved from http://www.ece.gatech.edu/research/labs/bwn/UWASN/index.html

21. Retrieved from http://a4academics.com/be-seminar-topics/18-be-ece-electronics-and-communication-seminar-topics/451-.. .

22. Retrieved from http://www.faadooengineers.com/threads/13267-under-water-wireless-communication

23. Retrieved from http://www.ijitee.org/attachments/File/v3i12/L16250531214.pdf

24. Retrieved from http://www.ijarcsse.com/docs/papers/Volume_3/9_September2013/V3I9-0183.pdf

25. Retrieved from http://www.chacha.com/question/can-you-explain-why-the-atomic-radius-of-oxygen-is-smaller-than-the-a. ..

26. Retrieved from http://www.chacha.com/question/to-whom-explain%2C-in-terms-of-electrons%2C-why-the-radius-of-a-calci.. .

27. Retrieved from http://www.computerscijournal.org/dnload/MOHAMMED-NAJM-ABDALLAH-and-AHMED-MOSA-DINAR/OJCSV06I02P85-9.. .

28. Retrieved from http://ojs.academypublisher.com/index.php/jnw/article/view/jnw0802317330

*****

FOR ANY CLARIFICATION OR SUGGESTION, WRITE US:

Editor-In-Chief
Pezzottaite Journals,
64/2, Trikuta Nagar, K. K. Gupta Lane, Jammu Tawi, Jammu & Kashmir - 180012, India.
(Mobile): +91-09419216270 – 71
editorinchief@pezzottaitejournals.net, contactus@pezzottaitejournals.net

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

MONITORING & FEEDBACK ANALYSIS FRAMEWORK FOR A HIGH-PERFORMANCE DISTRIBUTED WEB CRAWLER

Shishir Sarkar[31] Prateeksha Pandey[32]

ABSTRACT

Broad web search engines as well as many more specialized search tools rely on web crawlers to acquire large collections of pages for indexing and analysis. Such a WebCrawler may interact with millions of hosts over a period of weeks or months, and thus issues of robustness, flexibility, and manageability are of major importance. In addition, I/O performance, network resources, and OS limits must be taken into account in order to achieve high performance at a reasonable cost. In this paper, we describe the design the framework of Monitoring & feedback analysis component which is monitor the running distributed crawler and fetch the core information of running CPU and pass to feedback component to analysis on given parameter of core information than take action accordingly.

KEYWORDS

Web Search Engines, Crawling, Monitor, Feedback Analysis etc.

INTRODUCTION

Web crawlers, programs that automatically find and download web pages, have become essential to the fabric of modern society. This strong claim is the result of a chain of reasons: the importance of the web for publishing and finding information; the necessity of using search engines like Google to find information on the web; and the reliance of search engines on web crawlers for the majority of their raw data, as shown in Figure 1 (Brin & Page, 1998; search engines is emphasized by Van Couvering (2004), who argues that they alone, and not the rest of the web, form a genuinely new *mass* media.

Web users do not normally notice crawlers and other programs that automatically download information over the Internet. Yet, in addition to the owners of commercial search engines, they are increasingly used by a widening section of society including casual web users, the creators of email spam lists and others looking for information of commercial value. In addition, many new types of information science research rely upon web crawlers or automatically downloading pages (e.g., Björneborn, 2004; Faba-Perez, Guerrero-Bote, & De Moya-Anegon, 2003; mechanism for stopping crawlers from visiting some or all of the pages in their site.

Suggestions have also been published governing crawling speed and ethics (e.g., Koster, 1993, 1996), but these have not been formally or widely adopted, with the partial exception of the 1993 suggestions. Nevertheless, since network speeds and computing power have increased exponentially, Koster's 1993 guidelines need reappraisal in the current context. Moreover, one of the biggest relevant changes between the early years of the web and 2005 is in the availability of web crawlers. The first crawlers must have been written and used exclusively by computer scientists who would be aware of network characteristics, and could easily understand crawling impact. Today, in contrast, free crawlers are available online. In fact there are site downloaders or offline browsers that are specifically designed for general users to crawl individual sites, (there were 31 free or shareware downloaders listed in tucows.com on March 4, 2005, most of which were also crawlers). A key new problem, then, is the lack of network knowledge by crawler owners. This is compounded by the complexity of the Internet, having broken out of its academic roots, and the difficulty to obtain relevant cost information. In this paper, we review new and established moral issues in order to provide a new set of guidelines for web crawler owners. This is preceded by a wider discussion of ethics, including both computer and research ethics, in order to provide theoretical guidance and examples.

INTRODUCTION TO ETHICS

The word 'ethical' means, 'relating to, or in accord with, approved moral behaviors' (Chambers, 1991). The word 'approved' places this definition firmly in a social context. Behavior can be said to be ethical relative to a particular social group if that group would approve of it. In practice, although humans tend to operate within their own internal moral code, various types of social sanction can be applied to those employing problematic behavior. Formal ethical procedures can be set up to ensure that particular types of recurrent activity are systematically governed and assessed, for example in research using human subjects. Seeking formal ethical approval may then become a legal or professional requirement.

In other situations, ethical reflection may take place without a formal process, perhaps because the possible outcomes of the activity might not be directly harmful, although problematic in other ways. In such cases, it is common to have an agreed written

[31] M.Tech. Scholar, Department of CSE., CSIT Engineering College, CSVTU Technical University, Chhattisgarh, India, shishir.sarkar406@gmail.com

[32] Assistant Professor, Department of CSE., CSIT Engineering College, CSVTU Technical University, Chhattisgarh, India, prateekshapandey@csitdurg.in

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

or unwritten ethical framework, sometimes called a code of practice or a set of guidelines for professional conduct. When ethical frameworks or formal procedures fail to protect society from a certain type of behavior, it has the option to enshrine them in law and apply sanctions to offenders.

The founding of ethical *philosophy* in Western civilization is normally attributed to ancient Greece and Socrates (Arrington, 1998). Many philosophical theories, such as utilitarianism and situation ethics, are relativistic: what is ethical for one person may be unethical for another (Vardy & Grosch, 1999). Others, such as deontological ethics, are based upon absolute right and wrong. Utilitarianism is a system of making ethical decisions, the essence of which is that "an act is right if and only if it brings about at least as much net happiness as any other action the agent could have performed; otherwise it is wrong." (Shaw, 1999, p.10). Different ethical systems can reach opposite conclusions about what is acceptable: from a utilitarian point of view car driving may be considered ethical despite the deaths that car crashes cause but from a deontological point of view, it could be considered unethical. The study of ethics and ethical issues is a branch of philosophy that provides guidance rather than easy answers.

COMPUTER ETHICS

The philosophical field of computer ethics deals primarily with professional issues. One important approach in this field is to use social contract theory to argue that the behavior of computer professionals is self-regulated by their representative organizations, which effectively form a contract with society to use this control for the social good (Johnson, 2004), although the actual debate over moral values seems to take place almost exclusively between the professionals themselves (Davis, 1991). A visible manifestation of self-regulation is the production of a code of conduct, such as that of the Association for Computing Machines (ACM, 1992). The difficulty in giving a highly prescriptive guide for ethical computing can be seen in the following very general important advice, "One way to avoid unintentional harm is to carefully consider potential impacts on all those affected by decisions made during design and implementation" (ACM, 1992).

There seems to be broad agreement that computing technology has spawned genuinely new moral problems that lack clear solutions using exiting frameworks, and require considerable intellectual effort to unravel (Johnson, 2004). Problematic areas include: content control including libel and pornography (Buell, 2000); copyright (Borrull & Oppenheim, 2004); deep linking (Fausett, 2002); privacy and data protection (Carey, 2004; Reiman, 1995; Schneier, 2004); piracy (Calluzzo & Cante, 2004); new social relationships (Rooksby, 2002); and search engine ranking

Technology is never inherently good or bad; its impact depends upon the uses to which it is put as it is assimilated into society (du Gay, Hall, Janes, Mackay, & Negus, 1997). Some technologies, such as medical innovations, may find themselves surrounded at birth by a developed ethical and/or legal framework. Other technologies, like web crawlers, emerge into an unregulated world in which users feel free to experiment and explore their potential, with ethical and/or legal frameworks later evolving to catch up with persistent socially undesirable uses. Two examples below give developed illustrations of the latter case.

The fax machine, which took off in the eighties as a method for document exchange between businesses (Negroponte, 1995), was later used for mass marketing. This practice cost the recipient paper and ink, and was beyond their control. Advertising faxes are now widely viewed as unethical but their use has probably died down not only because of legislation which restricted its use (HMSO, 1999), but because they are counterproductive; as an unethical practice they give the sender a bad reputation.

Email is also used for sending unwanted advertising, known as spam (Wronkiewicz, 1997). Spam may fill a limited inbox, consume the recipient's time, or be offensive (Casey, 2000). Spam is widely considered unethical but has persisted in the hands of criminals and maverick salespeople. Rogue salespeople do not have a reputation to lose nor a need to build a new one and so their main disincentives would presumably be personal morals, campaign failure or legal action. It is the relative ease and ultra-low cost of bulk emailing that allows spam to persist, in contrast to advertising faxes. The persistence of email spam (Stitt, 2004) has forced the hands of legislators in order to protect email as a viable means of communication (www.spamlaws.com). The details of the first successful criminal prosecution for Internet spam show the potential rewards on offer, with the defendant amassing a 24 million dollar fortune (BBCNews, 4/11/2004). The need to resort to legislation may be seen as a failure of both ethical frameworks and technological solutions, although the lack of national boundaries on the Internet is a problem: actions that do not contravene laws in one country may break those of another.

RESEARCH ETHICS

Research ethics are relevant to a discussion of the use of crawlers, to give ideas about what issues may need to be considered, and how guidelines may be implemented. The main considerations for social science ethics tend to be honesty in reporting results and the privacy and well-being of subjects (e.g., Penslar, 1995). In general, it seems to be agreed that researchers should take responsibility for the social consequences of their actions, including the uses to which their research may be put (Holdsworth, 1995). Other methodological-ethical considerations also arise in the way in which the research should be conducted and interpreted, such as the influence of power relationships (Williamson, & Smyth, 2004; Penslar, 1995, ch. 14).

Although many of the ethical issues relating to information technology are of interest to information scientists, it has been argued that the focus has been predominately on professional codes of practice, the teaching of ethics, and professional dilemmas, as

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

opposed to research ethics (Carlin, 2003). The sociology-inspired emerging field of Internet research (Rall, 2004a) has developed guidelines, however, although they are not all relevant since its research methods are typically qualitative (Rall, 2004b). The fact that there are so many different environments (e.g., web pages, chatrooms, email) and that new ones are constantly emerging means that explicit rules are not possible; instead broad guidelines that help researchers to appreciate the potential problems are a practical alternative. The Association of Internet Researchers has put forward a broad set of questions to help researchers come to conclusions about the most ethical way to carry out Internet research (Ess & Committee, 2002), following an earlier similar report from the American Association for the Advancement of Science (Frankel & Siang, 1999). The content of the former mainly relates to privacy and disclosure issues and is based upon considerations of the specific research project and any ethical or legal restrictions in place that may already cover the research. Neither alludes to automatic data collection.

Although important aspects of research are discipline-based, often including the expertise to devise ethical frameworks, the ultimate responsibility for ethical research often lies with universities or other employers of researchers. This manifests itself in the form of university ethics committees (e.g., Jankowski & van Selm, 2001),

PROBLEM IDENTIFICATION

- Mostly web crawler does not have any automated monitoring system.
- It is very crucial to given 25*7 supports from human beings sometime it is possible but very costly.
- Previous crawler there is no provision for fetching raw core information of CPU.
- Previous crawler architecture there is no such mechanism, which will mine the raw core information and take action accordingly.
- Highly dependent on human interference.
- Monitoring system should also take care the network bandwidth & management information.

WEB CRAWLING ISSUES

Having contextualized ethics from general, computing and research perspectives, web crawling can now be discussed. A web crawler is a computer program that is able to download a web page, extract the hyperlinks from that page and add them to its list of URLs to be crawled (Chakrabarti, 2003). This process is recursive, so a web crawler may start with a web site home page URL and then download all of the site's pages by repeatedly fetching pages and following links. Crawling has been put into practice in many different ways and in different forms. For example, commercial search engines run many crawling software processes simultaneously, with a central coordination function to ensure effective web coverage (Chakrabarti, 2003; Brin & Page, 1998). In contrast to the large-scale commercial crawlers, a personal crawler may be a single crawling process or a small number, perhaps tasked to crawl a single web site rather than the 'whole web'.

It is not appropriate to discuss the software engineering and architecture of web crawlers here (see Chakrabarti, 2003; Arasu, Cho, Garcia-Molina, Paepcke, & Raghavan, 2001), but some basic points are important. As computer programs, many crawler operations are under the control of programmers. For example, a programmer may decide to insert code to ensure that the number of URLs visited per second does not exceed a given threshold. Other aspects of a crawler are outside of the programmer's control. For example, the crawler will be constrained by network bandwidth, affecting the maximum speed at which pages can be downloaded.

Since crawlers are no longer the preserve of computer science researchers but are now used by a wider segment of the population, which affects the kinds of issues that are relevant. Table 1 records some user types and the key issues that particularly apply to them, although all of the issues apply to some extent to all users. Note that social contract theory could be applied to the academic and commercial computing users, but perhaps not to non-computing commercial users and not to individuals. These latter two user types would be therefore more difficult to control through informal means.

There are four types of issue that web crawlers may raise for society or individuals: denial of service, cost, privacy and copyright.

ALGORITHMS

- Initialize the configuration parameter.
- Set the frequency to accept the message.
- Take of input of core raw information of CPU cluster of Control program (CP).
- Put it into MAP of key value pair where key is attribute name & value is attribute value.
- Start iterating on MAP.
- Check whether the new CPU added in a cluster or working with previous configuration parameter.
- IF new CPU added take the CPU processor level and RAM level of information do compare with other processor & RAM information.
- Modified the THRESHOLD value
- Get the threshold value from Cluster Performance Calculator.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

- Set as constant until new CPU is added.
- Take decision if current value is goes beyond to critical value than generate alarm or divert the traffic to another cluster.

BENEFITS OF PROPOSED ARCHITECTURE

- Component efficiently automatically monitors the running web crawler.
- Able to fetch the core level raw information.
- Also able to fetch the network related information.
- Efficiently transfer and communicate with feedback analysis component.
- Feedback analysis component having efficient algorithm to check the critical condition according to incoming information.
- Feedback analysis algorithm having dynamic modification ability of threshold value based on availability of hardware CPU information.
- Able to take decision of forwarding the traffic of another route based on availability.

CONCLUSION

This kind of mechanism should be used in large scale of crawler kind of system, so that we can avoid the system shutdown, fail kind of problem generally we can have a scanner system to scan the complete system but does not have a system or mechanism, which will take an action according.

REFERENCES

1. Chen, S., Mulgrew, B., & Grant, P. M. (1993, July). A Clustering Technique for Digital Communications Channel Equalization Using Radial Basis Function Networks. ***IEEE Trans. on Neural Networks,*** 4, 570-578.

2. Duncombe, J. U. (1959, January). Infrared Navigation—Part I: An Assessment Of Feasibility. ***IEEE Trans. Electron Devices***, volume ED-11, 34-39.

3. Lin, C. Y., Wu, M., Bloom, J. A., Cox, I. J., & Miller, M. (2001, May). Rotation, Scale, and Translation Resilient Public Watermarking For Images. ***IEEE Trans. Image Process.*** 10(5), 767-782.

4. Bharat, K., Broder, A., Henzinger, M., Kumar, P. & Venkatasubramanian, S. (1998, May). The Connectivity Server: Fast Access to Linkage Information on the Web. ***In Proceedings of the 7th Int.world Wide Web Conference.***

5. S. Brin and L. Page. The Anatomy of a Large-Scale Hyper textual Web Search Engine. ***In Proc. of the Seventh worldwide.***

6. Burner, M. (1997). Crawling Towards Eternity: Building an Archive of the World Wide Web. ***Webtechniques.***

7. Retrieved from http://citeseerx.ist.psu.edu/showciting?doi=10.1.1.18.4230

8. Retrieved from http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=994750&abstractAccess=no&userType=inst

9. Retrieved from http://syntigrate.blogspot.in

10. Retrieved from http://www.scit.wlv.ac.uk/%7Ecm1993/papers/Web_Crawling_Ethics_preprint.doc

11. Retrieved from http://ijltet.org/wp-content/uploads/2013/08/72.pdf

12. Retrieved from http://ebookily.org/doc/what-yahoo-search-resources-are-available-for-webmasters

13. Retrieved from http://ijarcsee.org/wp-content/uploads/2012/03/IJARCSEE_Paper_Template.doc

14. Retrieved from http://ijarcsee.org/index.php/IJARCSEE/article/view/252

15. Retrieved from http://link.springer.com/chapter/10.1007/3-540-45703-8_30

*****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

OPTIMIZATION IN SENSOR NETWORKS

Surbhi Siwach[33]

ABSTRACT

Two-third of our earth surface is covered with underwater sensor networks. Communication in underwater sensor network is a difficult task as compared with the terrestrial networks. The unique characteristics of limiting access quality are very different from those of the sensor network communication. These networks present very challenging aspects, such as low data rates and large delays, as well as the special propagation characteristics of the underwater medium.

A prototype of underwater communication system constituted by an underwater sensor and a hub buoy that relays data to the mainland is finally presented. Conclusions are drawn in terms of its network life, also in comparison with number of rounds in underwater EM communication systems and solutions.

In this paper, the work is done to improve the network life as in an underwater sensor network the main problem is related to the localization of the node and the formation of the cluster head. There are various problem associated with underwater sensor networks i.e. energy efficiency, frequency identification, node deployment, utilize the physical strength of the network.

KEYWORDS

Underwater Sensor Network, Cluster Head, Energy Efficiency, Dead Nodes, Number of Rounds, RAID Node etc.

INTRODUCTION

The main problem associated with the underwater sensor network is the problem of the localization of the nodes. An underwater sensor network is one of the major emerging technologies that require the data transmission at high rate with higher reliability ratio. These kind of network needs the equal concern for the architectural definitions as well as the algorithmic enhancements.

While defining these kind of network, the concern is required while selecting the sensors based on the type of surface, the type of link, control center, control parameters etc. These networks needs the regular monitoring of network because of continuous change is possible as the sensors are having floating movement and relatively need to analyze the energy definitions, requirement, consumption etc. It also needs to analyze based on type of communication, type of channel etc.

The nodes are localized in such a way that the network consumes less energy. For this purpose, cluster heads are formed in such a way that they consume less energy. The main goal of the network is to improve the network life. RAID node is the most important. The RAID node selected must be in such a way that it consumes less amount of energy. If the energy consumed in the network is less, then the life of the network enhances.

As the number of nodes will be less in a cluster head, the network would be alive for a longer time. Energy efficiency of a network is the main issue in the underwater sensor networks. In this work, we did the parametric localization and the related clustering in same phase. It means while deciding the cluster heads the same localization parameters will be considered such as connected components, energy etc. The work also extended in terms of two-cluster head definition. As one cluster head will die, the second cluster head will replace it and the search will be performed for second cluster head.

The presented work will perform the effective clustering and localization and improve the network life. The analysis will be done using graphs.

SENSOR NODES AND NETWORKING MODEL

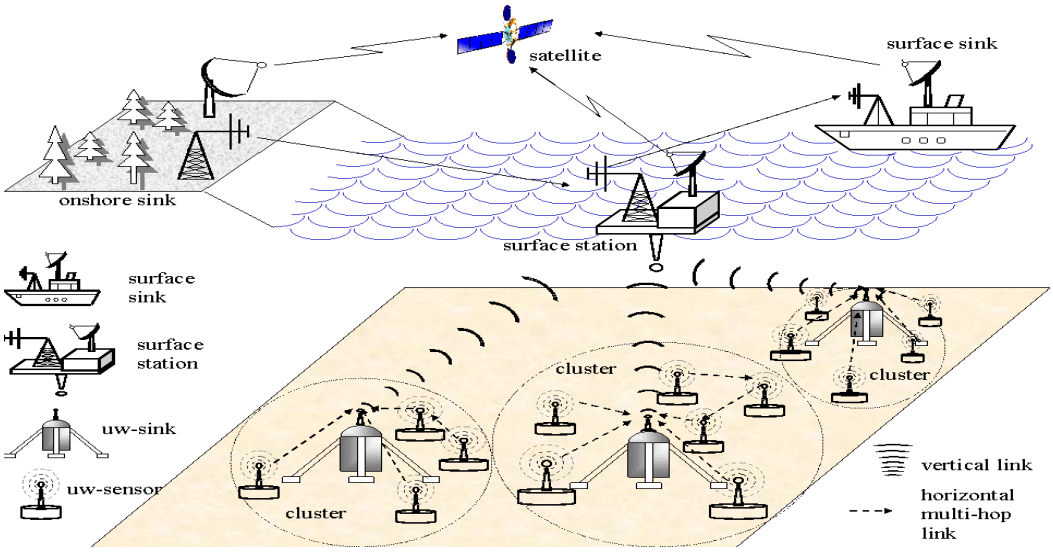
In the underwater sensor environment, the sensor nodes appropriate location is very important. We model the network via a graph plot in a two-dimensional modeling the dead nodes and the number of the rounds. The horizontal x-direction represents the number of the rounds taken for the communication and the vertical direction represents the no of dead nodes.

The decision vector will decide the actual positions of the sensors. Decision vector is the function applied to the sensor nodes, which examine the appropriate location of sensor node in such a way that they consume less amount of energy

[33]M. Tech (Final Year), CBS Group of Institutions, Fatehapuri-Jhajjar, India, surbhisiwach504@gmail.com

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Figure-1: Creation of Sensor Nodes and Cluster Heads



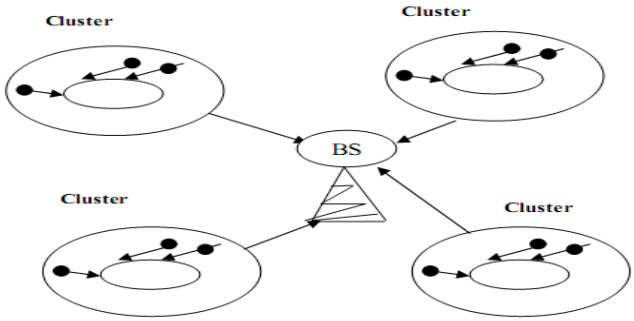
Sources: Authors Compilation

Cluster Head Formation

Cluster is formation of the group of nodes. More the number of nodes in a cluster head more would be the dead nodes. If the dead nodes will go on increasing soon the network life will be ruined because the energy level decreases with the increase of the number of nodes. The life of the network is improved by energy efficiency and it is possible only when the cluster head formation is in such a way that it consume less energy and becomes energy constraint. Cluster head node is the RAID node, which control all the clusters formed with in the network. RAID node is the node, which takes decision that on which cluster is to communicate with the signal such that it is energy constraint with the network. Clusters are formed with the various sensor nodes, which lead to communication with the RAID node.

Figure-2

The clustering formation starts with the procedure of discovery of neighboring sensor nodes (SN) by sending periodic Beacon Signals. After the creation of the clusters, each cluster is coordinated by the cluster head (CH) node, which is responsible for getting the measured value from its cluster's nodes and then aggregate them before send the aggregates to the sink through other cluster head. Control over the signals is done through the cluster head or the RAID node.



Sources: Authors Compilation

Hardware Specification

The equipment used in acoustic undersea networks is understandably much different from those used by terrestrial networks. First, undersea sensor nodes are much more expensive than a typical sensor node. They are generally much larger and more difficult to build, as they have to be able to handle a whole slew of new conditions, such as being in a liquid environment and being susceptible to natural occurrences such as corrosion and fouling.

The cost of the equipment and the general underwater conditions make it difficult to create dense undersea sensor networks, and as a result, the sensor nodes are spread out much farther than they would be in a typical terrestrial environment. The larger sensing and communication range help to offset this, but the sparseness must still be taken into account when developing detection algorithms because it becomes much more difficult to use the degree of aggregation principle without experiencing an increase in negative false alarms. This is because due to the large distances between the sensors, there is very little spatial correlation and so often, the degree of aggregation requirement will not be met since the target will only be within one node's sensing radius. Although in our assumptions we have generally ignored this fact, undersea sensor networks require much higher power. This is because there is added complexity at the receivers, and because communication is over much longer distances. Furthermore, solar

energy cannot be used as a power source. Depending on the duration that the network has to be left unattended, energy conservation techniques may need to be implemented. However, currently, the size of undersea sensor nodes has not been restricted and therefore battery life is not a current concern.

Node Specifications

This kind of network having two types of nodes called sensor node and the master node:

a) Sensor Nodes: These basic information node transfers the information over the network in the form of signals. These work as the transmitters and defined with some energy constrains. These nodes transfer data to the master nodes.

b) Master Nodes: Also node as head node or the collectors that work as the controller node. These nodes work as the gateway that collect data from the nodes and connect the sensor network with outer environment such as with internet etc. These nodes are capable to issue some command to sensor nodes

The sensor nodes are connected to the master node in a hierarchical manner. The number of hops that is required for a sensor node to communicate with the master node determines the level of the node. The available frequency band is divided into sub-bands and each sub-band is assigned to a *cluster* of nodes. A cluster of nodes is deployed in the same general geographical region. The neighboring clusters are assigned different frequency bands to assure low interference. Each cluster communicates with the master node through its first level node.

RELATED WORK

Underwater acoustic transmission has been studied since last decade. Recently, significant advances in MAC and routing protocols for underwater sensor networks have been witnessed. Good surveys reviewing the recent advances and challenges in underwater sensor networks can be found in [17, 8]. New work has been formulated on the problems associated with the underwater sensor networks.

Slim Rekhis (2012) suggested a work to analyze the quality of the underwater sensor network some monitoring mechanism is required such monitoring system. In this work, a water quality monitoring system is combined win sensor network based on Radio Frequency Identification System. The presented approach will reduce the network cost and the energy constraint and will provide improvement in terms of scalability, fault tolerance and reduction of error etc. [11].

Salvador Climent (2011) discussed the work on different routing and scheduling techniques for the underwater routing protocol. The channel characteristics and the nodes deployment is discussed in this paper along with the relative challenges. Besides this the main work is about the scheduling of the request under the fixed physical resources such as the channel bandwidth, radio frequency etc. The work is about to get the better utilization of physical strength of the channel. The range propagation, TDMA, CDMA approaches are discussed in this work. The network delay, throughput analysis are the basic parameters respective to which results and the effectiveness of the techniques is discussed [9].

Umberto M. Cella (2011) presented a case study of the sensor network under coastal marine environment. The paper investigates the experimental results under different scenarios. The paper discusses different aspects that affect the physical parameters of the network such as the temperature, illuminate of data etc. The work is here studied under the real environment and along with realistic problems that any underwater sensor network can face. The deployment analysis, stress factor and the environmental parameters are the important things that are discussed in this paper [12].

Salvador Climent (2011) defined a work for the scheduling and retransmission process for routing protocol in underwater sensor network. The work includes the node deployment with physical layer specification in terms of bandwidth and the power specification of the nodes. While performing the transmission the TDMA is considered with big time-guards. The analysis also performed under CSMA based approach. The work also performed the analysis of differntschduling and retransmission techniques for under water sensor network [10].

Michael Zuba (2011) presented work on security in under water sensor network. Security is one of the major issues for under water sensor network. One of the major security problems is network jamming that gives the denial of service over the network. A work on this problem. The work includes the study of existing DOS and jamming approaches and based on it a hardware-based jammer is proposed is to analyze the network. The analysis is done over OFDM network and the work is to resolve the attack with energy efficient approach [7].

Kenneth P. Hunt (2010) performed a work on the antenna based working on mussel based under water sensor network for the rivers. In this work, the biological sensors are being discussed along with environmental factors. In this work, some gape sensors are discussed along with relative microcontroller and the hardware. In this work, the challenges are discussed for the water proofing of physical devices such as sensors antennas etc. The work has defined in the form of a new antenaa called diploe folded

antenna. The physical characteristics of the devices and relatively the antenna are analyzed under different power factors and the different distance parameters [5].

Muhammad Ayaz (2009) performed a work on Underwater Wireless Sensor Networks: respective to different routing challenges and the issues. Here the different aspects are discussed that can affect the communication parameters, network delay, data rate etc. To get these changes the analysis is performed in terms of propagation delay, topological analysis, bandwidth analysis etc. The paper also explores different routing algorithms along with its comparative analysis under different scenarios and the environments. The presented work includes the relative routing algorithms along with the future work that is possible in each direction [8].

Antonio Caruso (2008) presented the analysis on the mobility model for the underwater sensor network. Here the physical movement of the nodes are analyzed respective to the connectivity, range localization, coverage area etc. The work includes study this network under real scenario and the environment. The protocol based implementation and the analysis is discussed in this chapter and relative to the coverage and the connective of the nodes [1]. The localization problem is also discussed in this paper

L. Badia (2006) performed a work on the optimization of the routing and scheduling mechanism respective the scheduling and the deployment of the nodes. The paper also discussed the intelligent scheduling scheme called integer linear programming. The optimization is been discussed along with link scheduling and the node placement in different scenarios [6].

Ian F. Akyildiz (2005) performed a work on the challenges of Underwater. It is a kind of survey work. The paper has discussed different physical architectures under the functional aspects. In this work, the analytical research is performed for two and three dimensional underwater sensor networks. More over to it the detail study is performed regarding the cross layered approach in under water sensor network. The paper also suggested some relative solution for the problem domain [2].

JiejunKong (2005) defined a work relative to the problem of localization and synchronization in underwater sensor network. The paper discusses the main characteristics of underwater sensor network such as floating mobility, signaling, link capability etc. The paper finds the physical analysis regarding the node placement or the localization respective to propagation delay, GPS signaling etc. The paper also discusses various kinds of attacks that can affect such kind of network. The integration of these all performance parameters and attacks are investigated along with the architectural design [3].

JiejunKong (2005) performed presented a work in 2005 on different kind of attacks respective to the different communication parameters such as packet delivery, time synchronization and the attack over the network. Security is the major concern in this paper. Different kind of security attacks and their impact is discussed in this paper. The integrity and the security constraint of delivery services and the localization are discussed [4].

Zhong Zhou under large network with vast number of nodes. Here all the environmental challenges are discussed. To over all the nodes and the scenario a segmented approach is discussed that is again connected in the form of hieratical architecture. In this work the challenges respective to the problem domain are discussed along with localization parameters. The parameters discussed are the coverage area along with mobility factor. Some other techniques are suggested along with large scale modeling [14].

Performance Task

The process involved in improving the network performance. The work is about the localization of nodes and the identification of cluster head in an effective way that will improve the network life and will give the efficiency as well as the reliability of the network: a) Design the sensor nodes, b) Cluster formation, c) Identify the cluster head, d) Communicate through the signal, e) Check for energy efficiency, f) Define threshold energy, g) Delimit the energy constraint, h) Draw variation on graph, i) Analysis of result, and j) End process.

Threshold energy is the minimum energy below which the node is the dead node i.e. $E<Th_e$= dead node. The maximum limit of the energy is the energy defined over which communication occurs smoothly. The range of energy lying between the threshold energy and the maximum energy is the range over which a proper sensor node works in underwater sensor network.

CONCLUSION

This paper is all about the performance of the network in underwater sensor network. In this paper, a relation is between the round performed and the dead node which is below the threshold energy level. Various clusters are formed with grouping of the sensor node in such a way that enhances cluster effectiveness due to which the system performance enhances.

Localization is the main issue in such communication network. Results will show that network life can be improved with the effective localization of nodes and better clustering techniques. By using the variable, a graph can be drawn between the dead nodes and the count of steps measured. Further work can be carried out for improving the system performance using cluster heading.

REFERENCES

1. Antonio Caruso, (2008), "The Meandering Current Mobility Model and its Impact on Underwater Mobile Sensor Networks", In ***IEEE INFOCOM,*** 978-1-4244-2026-1/08© 2008 IEEE

2. Jiejun Kong. Low-cost ***Attacks against Packet Delivery, Localization and Time Synchronization Services in Under-Water Sensor Networks***", WiSE'05, 1-59593-142-2/05/0009

3. Kenneth P. Hunt," ***Short Paper: Antennas for Mussel-Based Underwater Biological Sensor Networks in Rivers***", WUWNet'10, 978-1-4502-3/09/0010

4. L. Badia," ***An Optimization Framework for Joint Sensor Deployment, Link Scheduling and Routing in Underwater Sensor Networks***", WUWNet'06, 1-59593-484-7/06/0009

5. Michael Zuba," ***Short Paper: Launching Denial-of-Service Jamming Attacks in Underwater Sensor Networks", WUWNet***'11, 978-1-4503-1151-9

6. Muhammad Ayaz, "***Underwater Wireless Sensor Networks: Routing Issues and Future Challenges***", MoMM 2009 ERPAS 978-1-60558-659-5/09/0012.

7. Salvador Climent," ***Short Paper: Impact Analysis of Different Scheduling and Retransmission Techniques on an Underwater Routing Protocol***", WUWNet'11 978-1-4503-1151-9

8. Ibid,7

9. Slim Rekhis, "***A Wireless Sensor Network Based Water Monitoring System***", Q2SWinet'12 978-1-4503-1619-4/12/10

10. Umberto M. Cella, "***Electromagnetic Wave Wireless Communication in Shallow Water Coastal Environment: Theoretical Analysis and Experimental Results***", WUWNet'09 978-1-60558-821-6

11. Umberto M. Cella, "***Wireless Sensor Networks in Coastal Marine Environments: a Study Case Outcome***", WUWNet'09, 978-1-60558-821-6

12. Yogesh Kumar, "***Enhanced Energy Conserving Improved and Balanced LEACH for Heterogeneous Wireless Sensor Networks***", International Journal of Computer Science & Engineering Technology (IJCSET), Vol. 3 No. 9 Sep 2012, pp. 422-430

13. Retrieved from http://www.ijettcs.org/Volume2Issue3/IJETTCS-2013-04-15-030.pdf

14. Retrieved from http://ijircce.com/upload/2014/may/4Q_Energy.pdf

15. Retrieved from http://www.cs.virginia.edu/~son/cs862.f06/arsalan.doc

16. Retrieved from http://www.researchgate.net/publication/220926198_An_optimization_framework_for_joint_sensor_deploym...

17. Retrieved from http://senseslab.di.uniroma1.it/publications?task=show&id=121

18. Retrieved from http://espace.library.uq.edu.au/view/UQ:220239

19. Retrieved from http://www.ijser.in/archives/v1i2/MDExMzEwMDY=.pdf

20. Retrieved from http://www.learningace.com/doc/4253981/cfc1d304bded46742ba98c6db11aa49b/arsalan

21. Retrieved from http://www.ece.neu.edu/groups/tcit/opnet.html

22. Retrieved from http://www.ijettcs.org/Volume2Issue3/IJETTCS-2013-04-15-030.pdf

23. Retrieved from http://dl.acm.org/citation.cfm?id=1347374

****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

VOICE COMPUTING: IMPLEMENTATION AND APPLICATIONS

Amit A. Mokashi[34] Syed Akhter Hussain[35]

ABSTRACT

The present paper deals with the implementation and applications of Voice computing software. Voice computing is the newly proposed system for latest technical generation. Today's computer system is useful if the operating person is even unable to speak but there is no provision for the people who are handicap by other organs. As the name of project indicates, we are going to implement the concept of computing totally based on voice especially for blind or handicap people. Because if anyone is mentally fit then the Weakness of physical organ must not be the limitation for learning or operating computer. To overcome a problem we design a system consists of following facility in a single unit: a) Controlling Operating System, B) Talking Editor, C) People Choice, D) Mail Reader, And E) Security Module. The concept of voice computing is useful in many areas like for blind people, officers, students, etc.

KEYWORDS

Speech Recognition, Voice Computing, Speech, Text etc.

INTRODUCTION

Voice computing is the application specially designed foe physically challenged people or blind person for whom physical disorder is the only obstacle in handling current or traditional system. In our last paper we put forth the concept and application suggested using voice computing. In present paper we presented the implementation and applications in detail as per tested at our local level.

The problem statement can be given, as 'We want to design a system which replaces the current computer system by a strong, protected and efficient system.' The efficiency can be measured in terms of reusability, platform dependability, manpower, various resources etc.

Actually, the basic purpose of this project is to overcome the limitations of use of key board and mouse. In addition, other purpose is to give flexibility to the administrator for handling system and provide all in one system software.

Speech is the primary means of communication between people. Speech recognition, generation of speech waveforms, has been under development for several decades' .Automatic speech Recognition is a process by which a computer takes a speech signal and Converts it into words. A computer recognizes what a person Said by the process. Keyboard, although a popular medium is not very convenient, as it requires a certain amount of skill for effective usage. A mouse on the other hand requires a good hand eye co-ordination.

Physically challenged people find computer difficult to use. Partially blind people find reading from a monitor difficult. All these constraints have to be eliminated. Speech interface help us to tackle these problems. The objective is to trap human voice in a digital computer and decode it into corresponding text.

Speech recognition can be defined as the process of converting an acoustic signal, captured by a microphone or a telephone, to a set of words. When two people speak to one another, they both recognize the words and the meaning behind them. Computers, on the other hand, are only capable of the first thing: they can recognize individual words and phrases, but they do not really understand speech in the same way as humans do. Computer recognizes the command and software tells the computer what to do when that command is recognized. Speech Recognition (is also known as Automatic Speech Recognition (ASR) or computer speech recognition) is the process of converting a speech signal to a sequence of words, by means of an algorithm implemented as a computer program.

Speech recognition technology has made it possible for computer to follow human voice commands and understand human languages. The main goal of speech recognition area is to develop techniques and systems for speech input to machine. Speech is the primary means of communication between humans. These recognizers can be used on the construction of speech-based applications, but with some limitations due to the difficulty of integration with other software applications and possible license restrictions.

[34]P.G. Student, Aditya Engineering College, Maharashtra, India, amitmokashi12@gmail.com
[35]Professor, Aditya Engineering College, Maharashtra, India, akhter.it@gmail.com

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Figure-1: Speech and Text Interconversions

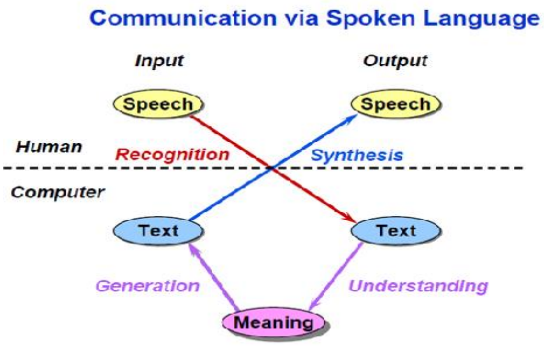
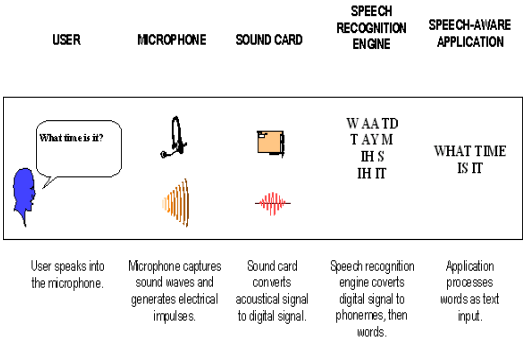


Figure-2: Speech to Text Conversions



Sources: Authors Compilation

IMPLEMENTATION

In literature survey, we got some of the common characteristics in each voice software, which is already available in market. We sucked all the filtered features from each software and consider the most important attributes all in together. The modules provided in our proposed software are described in detail as follow:

Controlling Operating System

In this module, we use the human voice for controlling the operations of operating system. For this facility, we use human voice as an input and the output of this facility is the operations performed as per human voice. Suppose the user give the command of 'START' through his voice that time system show the menus present in 'START' button. For performing this operation, we use the facility provided by speech engine. By considering feature of speech engine, we develop the GUI for controlling this facility, which will easily, managed by targeted people.

Talking Editor

In this module, we provide facility for blind person to listen the words or document present in front of them. Blind people are unable to operate computer only due to their physical disorder but mentally they are able to give the same efforts and output as normal person. This tool is useful in such a situation. In this module, the input is user-selected document and the output is the voice of computer system.

People Choice

By using this facility, any person is able to write the data on editor by using their voice. This tool is useful when someone want to write large amount of data. This module reduces the efforts of normal person but it design concentrating on the persons who are physically challenged by their hand.

Mail Reader

In this tool, we can hear our mail in mailbox automatically without reading them personally. This tool is specially designed for people who are in software profession. The facility is useful for the professional who are busy almost all the time, they get there mail updates after some time of interval atomatically. They can read their mailbox any time they want.

As the Concept Relates to Voice, So Security is the Very Important Issue

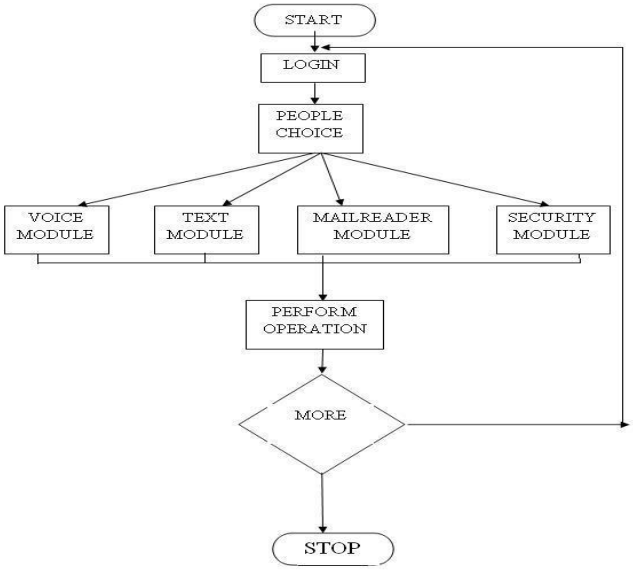
Security Module

To implement the security we provide the option for encryption and Decryption of data. It will perform using AES, RSA algorithm. In this module, we provide the advance security option for strong security. Because password or shoulder sniffing are not comparatively strong. Cryptography is the latest technique and easy to implement.

In encryption, we convert the normal text to cipher text and key is provided to receiver. At receiver end, decryption is done in which cipher text is converted to normal text.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Figure-3: Implementation of Voice Computing



Sources: Authors Compilation

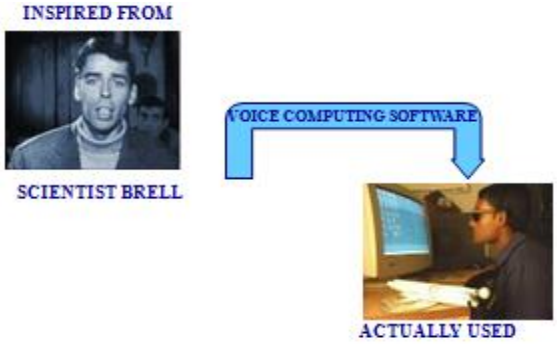
APPLICATIONS

From Medical Perspective

People with disabilities can benefit from speech recognition programs. Speech recognition is especially useful for people who have difficulty using their hands, in such cases speech recognition programs are much beneficial and they can use for operating computers. Speech recognition is used in deaf telephony, such as voicemail to text.

Figure-4: Use of Voice Computing Software for Physically Handicap People

Figure-5: Use of Voice Computing Software for Blind People



Sources: Authors Compilation

From Military Perspective

Speech recognition programs are important from military perspective; in Air Force, speech recognition has definite potential for reducing pilot workload. Beside the Air force, such Programs can also be trained to be used in helicopters, battle management and other applications.

From Educational Perspective

Individuals with learning disabilities who have problems with thought-to-paper communication (essentially they think of an idea but it is processed incorrectly causing it to end up differently on paper) can benefit from the software.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

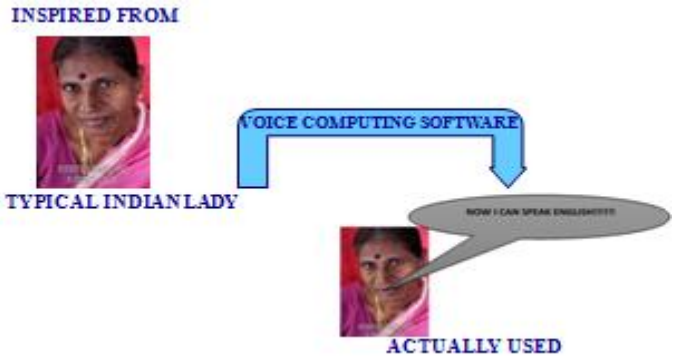
Figure-6: Use of Voice Computing Software for Students

Sources: Authors Compilation

From Domestic Perspective

Individuals like housewives who are unable to go outside and learn something is quite difficult, for them the present software works as a teacher for learning anything. Here we consider the example of typical Indian women who wish to learn English but she is unable to learn due to time, schedule constraint. She can able to learn English easily through E-learning voice software.

Figure-7: Use of Voice Computing Software for Housewives



Sources: Authors Compilation

From Business Perspective

The software is useful for business people during their official work to avoid their wastage of time. The mail reader module automatically reads the mail as there is no need of continues monitoring. It automatically intimates the arrival of new mails through voice.

Figure-8: Use of Voice Computing Software for Business People

Sources: Authors Compilation

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Likewise, the software can be used for:

- Illiterate people,
- Special Childs,
- Ad agency.

ADVANTAGES OF SOFTWARE

- Able to write the text through both keyboard and voice input.
- Voice recognition of different notepad commands such as open save and clear.
- Open different windows soft wares, based on voice input.
- Requires less consumption of time in writing text.
- Provide significant help for the people with disabilities.
- Lower operational costs.

LIMITATIONS OF SOFTWARE

- Comparatively Low accuracy,
- Not good in the noisy environment.
-

TESTING

The accuracy of project is tested by comparing it with the speech recognition feature inbuilt in Windows7 and later version. The software results 80-92% better than that, as it is user friendly and works more efficiently. It also proves better as it contains the extra modules with are join with voice control which are not present in operating system or even in any other present software's.

FUTURE ENHANCEMENTS

This work can be taken into more detail and more work can be done on the project in order to bring modifications and additional features. The current software does not support a large vocabulary, the work will be done in order to accumulate more number of samples and increase the efficiency of the software. By considering Artificial Neural Network (ANN) the proposed system becomes more powerful and efficient. The scope of project can be extended to implement the system on small microchip to make it more popular, cost effective and user friendly.

CONCLUSION

This Project work of voice computing started with a brief introduction of the technology and its applications in different sectors. At the later stage, we discussed different tools for bringing that idea into practical work. After the development of the software, finally, it will test and results will be discussed, few deficiencies factors will bring in front. A word based acoustic model is used. This model can be used only for limited vocabulary. As the size of the vocabulary increases performance of the system decreases. At last, we conclude that this project can be used at very large scale with very little modifications. The system can be extended to continuous word recognition with large vocabulary based on a phone Acoustic model, using the HMM Technique or using other growing techniques like Artificial Neural Network.

REFERENCES

1. Mokashi, A. A. (2013, December). Voice Computing: Technology for Next Technical Era. ***International Journal of Engg. Research & Technology***, 2(12), 2876- 2881. ISSN: 2278-0181.

2. Young, S., Evermann, G., Hain, T., Kershaw, D., Moore, G., Odell, J., Ollason, D., Povey, D., Valtchev, V., & Woodland, P. (2004). ***The HTK Book for HTK V3.2,*** Cambridge: Cambridge University Press.UK.

3. Sadaoki Furui, November 2005, ***ECTI Transactions on Computer and Information Technology, 50 years of Progress in speech and Speaker Recognition Research***, 1(2).

4. Davis, K. H., Biddulph, R., & Balashek, S. (1952). Automatic Recognition of spoken Digits, ***J.Acoust.Soc.Am***, 24(6):637-642.

5. Boersma, P., & Weenink, D. (2005). ***Praat: doing phonetics by computer***-(Version 4.3.33) [Computer program]. Retrieved from http://www.praat.org/

6. Witten, E. Frank. (2005). ***Data Mining: Practical Machine Learning Tools and Techniques (***2nd Ed.). San Francisco: Morgan Kaufmann.

7. Hoyt, R. (2010, January 01). ***Lessons Learned from Implementation of Voice Recognition for Documentation in the Military Electronic Health Record System***. National center for Biotechnology Information. (Online Article). Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2805557/

8. Retrieved from http://en.wikipedia.org/wiki/Speech_recognition

9. Retrieved from http://www.ijpttjournal.org/volume-2/issue-1/IJPTT-V2I1P407.pdf

10. Retrieved from https://developer.lumenvox.com/learn/articles/speech-recognition-overview

11. Retrieved from http://www.speechrecognition.net/index.html

12. Retrieved from https://answers.yahoo.com/question/index?qid=20070812062358AAZSUur

13. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2805557

14. Retrieved from http://egouvea.users.sourceforge.net/paper/smli_tr-2004-139.pdf

References for Images

15. A. Hagen, D. A. Connors, & B. L. Pellom. ***The Analysis and Design of Architecture Systems for Speech Recognition on Modern Handheld-Computing Devices.***

16. Lawrence R. Rabiner. ***Applications of Speech Recognition in the Area of Telecommunications***.

17. Willie Walker, Paul Lamere, Philip Kwok, Bhiksha Raj, Rita Singh, Evandro Gouvea, Peter Wolf, Joe Woelfel. ***Sphinx-4: A Flexible Open Source Framework for Speech Recognition.***

18. ***4-8, Concept put forth first in present paper.***

*****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

SOLAR POWER CONVERSION USING 11-LEVEL CASCADED H-BRIDGE MULTI LEVEL INVERTER

Sk. Tenjim Parveen[36] G. Sai Nandini[37] K. Preethi[38] B. Suneetha[39] V. Dega Rajaji[40]

ABSTRACT

This project presents a single-phase 11-level (5 H-bridges) cascade multilevel DC-AC grid-tied inverter. Each inverter bridge is connected to a 200 W solar panel. OPAL-RT lab was used as the hardware in the loop (HIL) real-time control system platform where a Maximum Power Point Tracking (MPPT) algorithm was implemented based on the inverter output power to assure optimal operation of the inverter when connected to the power grid as well as a Phase Locked Loop (PLL) for phase and frequency match.

A novel SPWM scheme is proposed in this paper to be used with the solar panels that can account for voltage profile fluctuations among the panels during the day. Simulation and experimental results are shown for voltage and current during synchronization mode and power transferring mode to validate the methodology for grid connection of renewable resources.

KEYWORDS

Multilevel Converter, Cascaded H-Bridges, Solar Panel, Photovoltaic, MPPT, PWM etc.

INTRODUCTION

Because energy resources and their utilization will be a prominent issue of this century, the problems of natural resource depletion, environmental impacts, and the rising demand for new energy resources have been discussed fervently in recent years. Several forms of renewable zero-pollution energy resources, including wind, solar, bio, geothermal and so forth, have gained more prominence and are being researched by many scientists and engineers.

Solar cell installations involve the use of multiple solar panels or modules, which can be connected in series or in parallel to provide the desired voltage level to the inverter. The cascaded H-bridge multilevel inverter topology requires a separate DC source for each H-bridge so that high power and/or high voltage that can result from the combination of the multiple modules in a multilevel inverter would favor this topology.

To maximize the energy harvested from each string, a maximum power point tracking (MPPT) strategy is needed. The task of finding the optimum operation point might increase the complexity and component count as the number of isolated DC sources increase. The approach chosen to deal with the number of input sources was to monitor AC output power parameters instead of DC input measurements.

Traditional multilevel inverters include cascaded H-bridge inverter, diode clamped inverter, and flying capacitors inverter. This paper focuses on the single-phase 11-level (5 H- Bridge) cascade multilevel inverter.

MULTILEVEL INVERTER AND PV INTERFACE

An overview of the system is shown in Figure 1. The core component of this inverter design is the four-switch combination shown in Figure 1. By connecting the DC source to the AC output by different combinations of the four switches, Q11, Q12, Q13, and Q14, three different voltage output levels can be generated for each DC source, +Vdc, 0, and -Vdc. A cascade inverter with N input sources will provide (2N+1) levels to synthesize the AC output waveform. The DC source in the inverter comes from the PV arrays, and the switching signals come from the multicarrier sinusoidal pulse width modulation (SPWM) controller.

[36]Student, Department of EEE, Acharya Nagarjuna University College of Engineering & Technology, Acharya Nagarjuna University, Guntur, India, shaik.parveen.3252@gmail.com

[37]Student, Department of EEE, Acharya Nagarjuna University College of Engineering & Technology, Acharya Nagarjuna University, Guntur, India, sainandini.guntupalli@gmail.com

[38]Student, Department of EEE, Acharya Nagarjuna University College of Engineering & Technology, Acharya Nagarjuna University, Guntur, India, preethi555.kommala@gmail.com

[39]Student, Department of EEE, Acharya Nagarjuna University College of Engineering & Technology, Acharya Nagarjuna University, Guntur, India, starsunita.b@gmail.com

[40]Assistant Professor, Department of EEE, Acharya Nagarjuna University College of Engineering & Technology, Acharya Nagarjuna University, Guntur, India, rajajiyadav@gmail.com

The 11-level inverter connects five H-bridges in series and is controlled by five sets of different SPWM signals to generate a near sinusoidal waveform. The connection to the grid is done through a variable transformer to assure that at any time the number of H- Bridges used can be controlled, the grid voltage generated by the inverter is met and to give more flexibility to the experiment since irradiance levels might not be enough. For that reason, an additional fixed 10mH inductance was added as the connection inductance for power transferring mode.

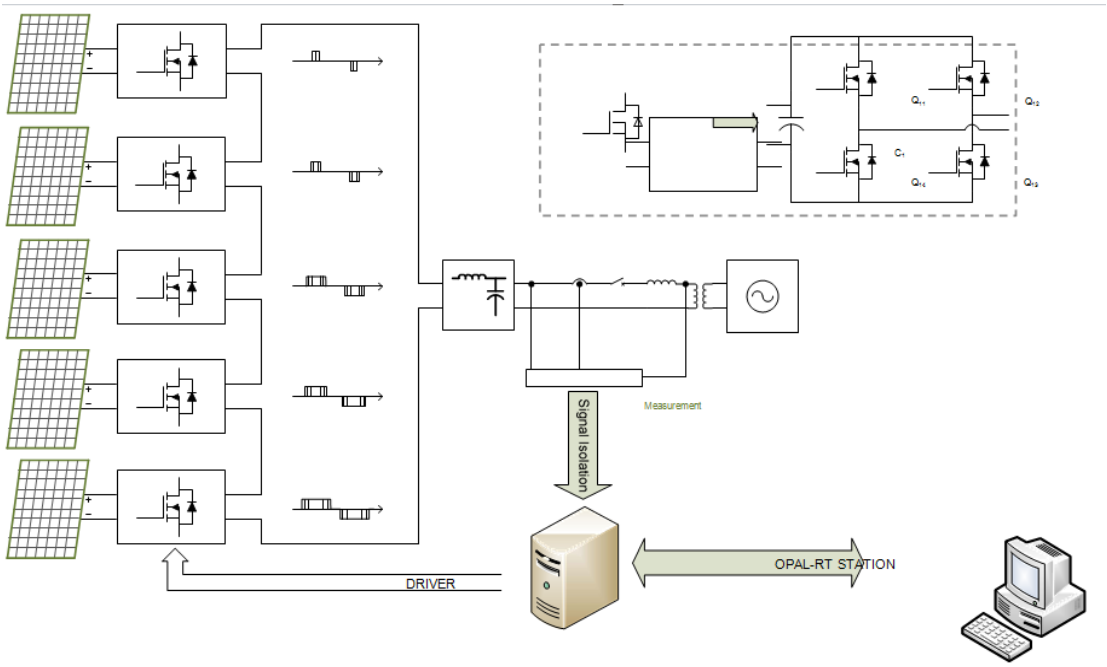
The individual solar panel output power is proportional to solar irradiance variations that occur during the day. The MPPT algorithm will work sensing the output power so no feedback from the individual panels is provided to reduce the number of sensors.

As can be seen in Figure-1, the lower panels, in terms of control signals, will deliver more energy than the upper panels. In order to avoid uneven power to be drawn from the panels by the inverter, a different inverter control approach for the SPWM scheme is proposed here to be used with the solar panels that can account for the voltage profile variation of the panels that occurs during the day. The MPPT and grid synchronization algorithm are fed by output and voltage current signals to generate the gate driver signals as shown in Figure-2.

In Figure-3 are shown the inverter and its cycle-by-cycle SPWM control methodology. The irradiance profile over a day changes a few orders of magnitude than a 60 Hz system i.e. means that a control change action over the modulation index can be taken over a few cycles of the 60 Hz control system. It is desired to get the same amount of power from each string, which cannot be achieved using a conventional SPWM approach. For example, the lower panels in Figure-1 would send more power than the upper panels as they are switching for a longer time. The sinusoidal nature of the current comes as another factor that makes the power drawn from different panels uneven.

The multilevel cascade topology does not require any of the H-bridges to be switched in a determined sequence, as would be the case for a diode clamped multilevel (DCM) converter. This gives freedom to switch the H-bridges in the circuit in any order, which can be used as strategy to equalize the power transferred from individual panels. The control strategy implemented shifts the carrier signal over N cycles in the case of a (2N+1) level inverter to make it possible to draw the same amount of power from each string.

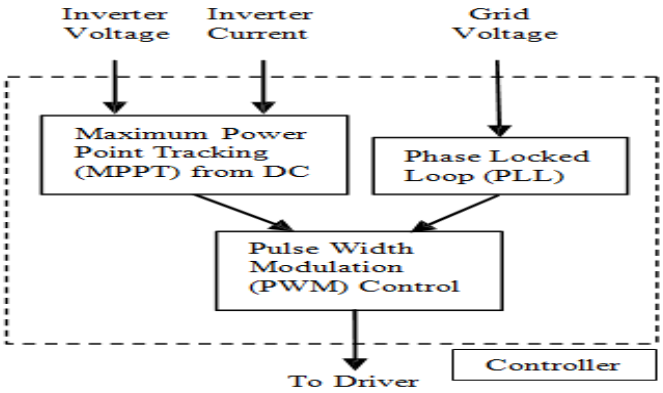
Figure-1: Multilevel Inverter System Overview



Sources: Authors Compilation

Shifting the carrier down for each cycle is the same as physically changing the position of the H-bridges shown in Figure-3(a) for an 11-level inverter. In that figure, five cycles of the fundamental frequency are needed to have each panel switch position with the other four. The energy stored in the capacitor will come to help in this process in a cycle-by-cycle basis to avoid a considerable voltage drop due to its considerable large capacitance (1000 uF).

Figure-2: Control Block Diagram



Sources: Authors Compilation

SYNCHRONIZATION AND TRACKING CONTROL SYSTEM

Synchronization between inverter and grid means that both will have the same phase angle, frequency and amplitude. This can be done noise proof with respect to the grid by sensing the grid voltage in a Phase Locked Loop (PLL). Typical PLL algorithms include inverse Park-based PLL, Hilbert transformer-based PLL, and transport delay-based PLL. The one to be included in this design is the transport delay-basedPLL.

PLL algorithm notice that the delayed angle can be directly controlled at the computer station during the experiment to provide the signal in quadrature with the grid, which is the input to the Park transform block.

PLL output is the actual angle position of grid voltage. This signal is used to generate the sine wave that is used as the reference signal to control system, which will generate the SPWM signals to drive the switches. The time required for synchronization will be dependent on the PI block parameters. PLL synchronization simulation. In that figure, the PLL starts its synchronization at 0.03 second, and it is in synchronization after about 0.13 second. Since the angle is now known, it is possible to control the phase difference between inverter and grid by controlling δ. This allows the power flow to be controlled according to (1).

Figure-3: (A) Inverter Topology, and (B) Carrier Shifting Control Scheme

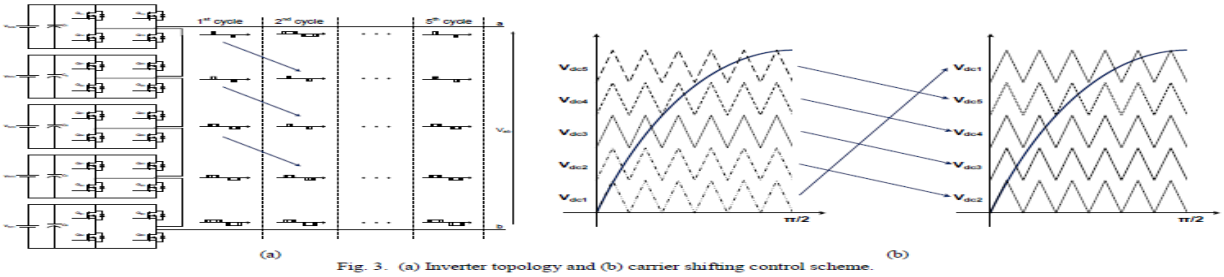


Fig. 3. (a) Inverter topology and (b) carrier shifting control scheme.

Sources: Authors Compilation

$$P = \frac{V_{inv} \cdot V_{grid}}{X_L} \sin \delta \qquad (1)$$

Where V_{inv} is the inverter voltage, V_{grid} is the grid voltage, X_L is the connection impedance, and δ is the angle between grid and inverter.

Instead of sensing the individual panel voltages, the maximum power point tracking (MPPT) algorithm determines the optimal point of operation of the panel by calculating the output power and phase angle variation [18]. It monitors output voltage and current parameters by making small changes on the phase angle and looking at the power variation, as in a hill climbing optimization method to track the maximum power point.

EXPERIMENTAL DEMONSTRATION

Each one of five H-bridges has its own 200 W PV panel connected as an independent source. The panels' specification can be seen in Table-1. The control signals to the bridges are sent by the OPAL-RT workstation where software and I/O boards are installed. The system acquires grid voltage and inverter output current and voltage to the control block (OPAL-RT workstation) to generate the driver signals to the inverter.

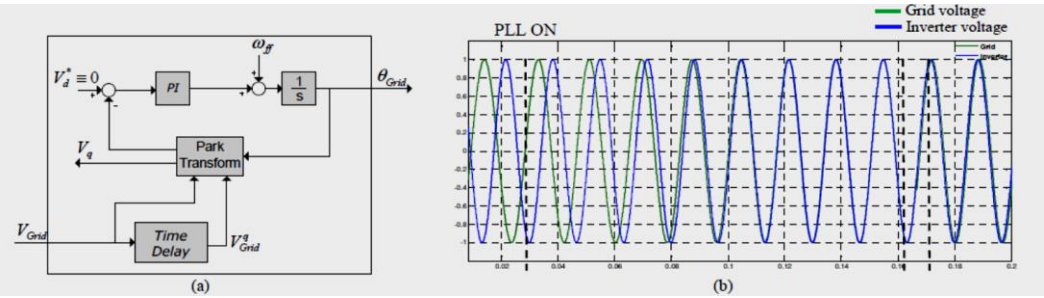
Table-1: Solar Panel Specifications

Sanyo HIP-200DA3	
Rated Power	200 W
Maximum power voltage	56.2 V
Maximum power current	3.56 A
Open circuit voltage	68.8 V
Module efficiency	16.5 %

Sources: Authors Compilation

The RT-Lab control platform, which connects the software (PWM, PLL, MPPT) with the hardware (solar panel, grid, 11- level cascaded H-bridge inverter), to create a real time platform, is the main tool to perform the experiments. Due to hardware limitation, the maximum achievable frequency for the SPWM signals is 2 kHz, which requires bulk-filtering components as shown in Table 3.2. A 2 kHz discrete-time RTLab drives the inverter with five solar panels as DC inputs and provides a waveform to interface with the 60 Hz AC grid.

Figure-4: (A) Transport Delay-Based PLL Algorithm, and (B) Voltage Synchronization Using PLL



Sources: Authors Compilation

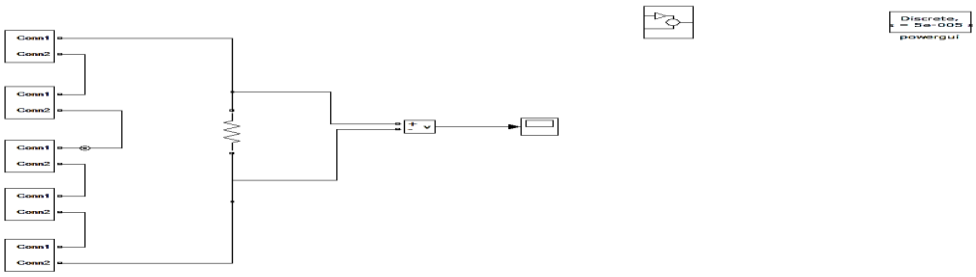
Table-2: Passive Components Specification

Filter Inductance	1 mH
Filter Capacitance	92 uF
Connection inductance	10 mH

Sources: Authors Compilation

Matlab Modeling and Results

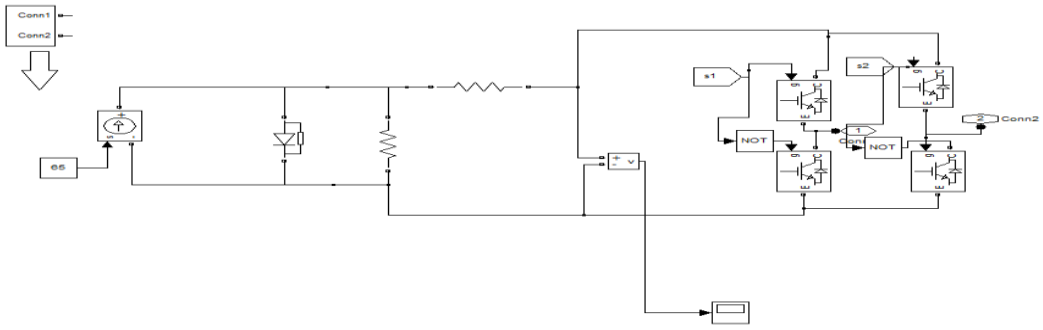
Figure-5



Sources: Authors Compilation

The above circuit represents the solar power conversion using 11 level-cascaded H-Bridge multi-level inverter.

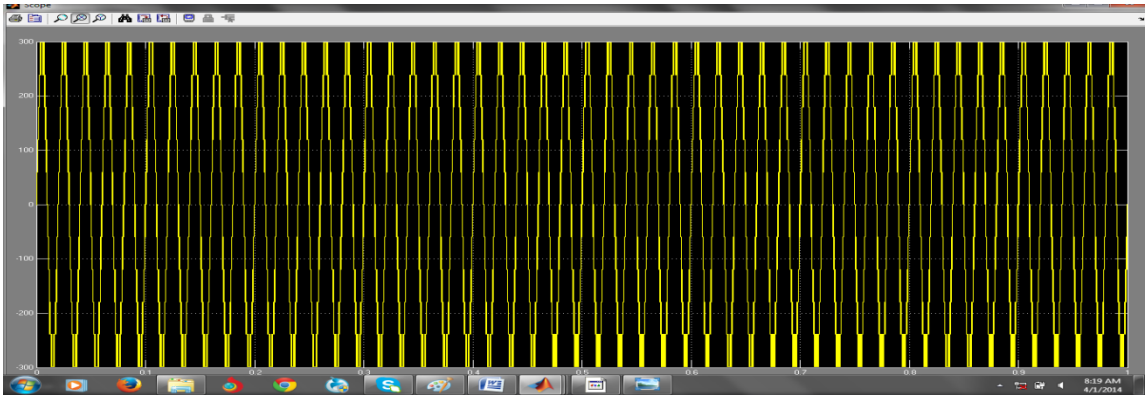
Figure-6



Sources: Authors Compilation

The above circuit represents internal circuit of subsystem1.

Figure-7



Sources: Authors Compilation

The above figure represents the output waveform of the solar power conversion using 11level cascaded H-Bridge multilevel inverter.

CONCLUSION

This project presented an eleven-level cascade H-bridge inverter, which uses PLL and MPPT with separate solar panels as DC sources to interact with the power grid. A SPWM approach was presented to deal with the uneven power transferring characteristics of the conventional SPWM modulation technique. This technique proved to be successful due to the irradiance profile and the use of capacitors to smooth the voltage fluctuation. The system was driven at 2 kHz because of speed constrains of the control platform, which required bulk filter components. Grid connection results were shown using the proposed MPPT algorithm. Future work includes the use of a DSP platform to increase switching frequency and reduce filter requirements. The entire PV system structure and its interaction with the grid through PLL and MPPT algorithms were shown by the simulation and experimental results.

REFERENCES

1. J. M. Carrasco, L. G. Franquelo, J. T. Bialasiewicz, E. Galvan, R. C. P.Guisado, Ma. A. M. Prats, J. I. Leon, & N. Moreno-Alfonso. (2006, June). Power- Electronic Systems for the Grid Integration of Renewable Energy Sources: A Survey. ***In Proceedings of the IEEE Transactions on Industrial Electronics***, 53(4), 1002-1016.

2. Morrison, A. J. (2007, October 25-26). Global Demand Projections for Renewable Energy Resources. ***In Proceedings of the IEEE Canada Electrical Power Conference, (pp. 537-542)***.

3. Rodriguez, J., Bernet, S., Bin Wu, Pontt, J. O., & Kouro, S. (2007, December). Multilevel Voltage-Source Converter Topologies for Industrial Medium-Voltage Drives. ***IEEE Transactions on Industrial Electronics,*** 54(6), 2930-2945.

4. Tolbert, L. M., & Peng, F. Z. (2000, July 15-20). Multilevel Converters as a Utility Interface for Renewable Energy Systems. ***In Proceedings of the IEEE Power Engineering Society Summer Meeting (pp. 1271- 1274).*** Seattle, Washington.

5. Khomfoi, S., & Tolbert, L. M. (2007). Multilevel Power Converters. ***Power Electronics Handbook*** (2nd ed.), Chapter 17, pp. 451-482. Elsevier. ISBN 978-0-12- 088479-7.

6. Busquets-Monge, S., Rocabert, J., Rodriguez, P., Alepuz, S., & Bordonau, J. (2008, July). Multilevel Diode-clamped Converter for Photovoltaic Generators with Independent Voltage Control of Each Solar Array. ***In Proceedings of the IEEE Transactions on Industrial Electronics, (volume 55, pp.2713-2723)***.

7. Ozdemir, E., Ozdemir, S., Tolbert, L. M., & Ozpineci, B. (208, February 24-2). Fundamental Frequency Modulated Multilevel Inverter for Three-phase Stand-alone Photovoltaic Application. ***In Proceedings of the IEEE Applied Power Electronics Conference and Exposition, (pp. 148-153)***.

8. Khajehoddin, S. A., Bakhshai, A., & Jain, P. The Application of the Cascaded Multilevel Converters in Grid Connected Photovoltaic Systems. ***In Proceedings of the IEEE Canada Electrical Power Conference, (pp. 296-301)***.

9. Ozdemir, S., Ozdemir, E., Tolbert, L. M., & Khomfoi, S. (2007, November 27-30). Elimination of Harmonics in a Five-level Diode-clamped Multilevel Inverter Using Fundamental Modulation. ***In Proceedings of the International Conference on Power Electronics and Drive Systems***, ***(pp. 850-854)***.

10. Lai, J. S., & Peng, F. Z. (1996, May/June). Multilevel Converters-A New Breed of Power Converters. ***IEEE Transactions on Industry Applications,*** 32(3), 509-517.

11. Retrieved from http://www.opal-rt.com/sites/default/files/technical_papers/University%20Tenesse_2010.pdf

12. Retrieved from http://www.ijseat.com/index.php/ijseat/article/view/52

13. Retrieved from http://www.researchgate.net/publication/4324761_Elimination_of_Harmonics_in_a_Five-Level_Diode-Clamp.. .

****

CALL TO JOIN AS MEMBER OF EDITORIAL ADVISORY BOARD

We present you an opportunity to join Pezzottaite Journals as member of 'Editorial Advisory Board' and 'Reviewers Board'. Pezzottaite Journals seek academicians and corporate people from around the world who are interested in serving our voluntarily 'Editorial Advisory Board' and 'Reviewers Board'. Your professional involvement will greatly benefit the success of Pezzottaite Journals.

Please forward below stated details at contactus@pezzottaitejournals.net.

- Updated Resume, Scanned Photograph, and Academic Area of Interest.

For Paper Submission & Clarification or Suggestion, Email Us @:

callandinvitation@pezzottaitejournals.net, callandinvitations@pezzottaitejournals.net
callforpapers@pezzottaitejournals.net, editorinchief@pezzottaitejournals.net

Editor-In-Chief
Pezzottaite Journals,
64/2, Trikuta Nagar, K. K. Gupta Lane, Jammu Tawi,
Jammu & Kashmir - 180012, India.
(Mobile): +91-09419216270 – 71

(sd/-)
(Editor-In-Chief)

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

MICROCHIPS: RESERVOIR BASED CONTROLLED DRUG DELIVERY DEVICES

Harshita Singh[41] Meera Chandradatt Singh[42]

ABSTRACT

Microchip is the basic functional unit in the computer science technology, which is finding its application in pharmaceutical sciences as recent development. Microchips as a controlled drug delivery device are an important and novel technology in the field of pharmaceutical sciences. It has tremendous opportunities in the field of pharmaceutical research. Microchips are bio-sensing devices of modern therapeutic techniques, which implies rational design, and highly targeted and controlled delivery of specific drug compounds. Microchip technologies offer alternative approaches to miniaturization. Advances in micro fabrication technology have made this new class of controlled drug delivery possible, such that, programmable devices.

Silicon microchips have the ability to store and release multiple chemicals on demand. Micro fabrication technology has been used to create micro particles for site-specific delivery. These micro particles can be designed with a thickness of 1-50 micrometer and diameter of 1-100 micrometer. Microchips include several reservoirs, substrate and a release system. Each of the reservoir on a single microchip can hold different molecules and in different concentrations which can release independently. Microchips may be used as therapeutic, prophylactic or diagnostic agents providing new treatment option to clinicians in fight against diseases.

KEYWORDS

Microchip, Controlled Drug Delivery, Micro fabrication, Silicon, Reservoir, Micro particles etc.

INTRODUCTION [3,7,8,15,17]

Over the last several years, micro fabrication technology has applied to lead a successful development of new and different drug delivery systems. Majority of research works are focused on miniaturized system so development of micro devices for therapeutic application has become important concern of researchers. Drug delivery becomes problematic when the drugs are too potent for systemic delivery. Therefore, attempts have been made to design and fabricate new delivery devices, capable of controlled, pulsatile or continuous release. The technology is designed to prove a more convenient and accurate method of drug dosing, which could improve patient compliance in taking medicines.

Controlled-release systems for drug delivery first appeared in the 1960s and 1970s. In the past three decades, the number and variety of controlled release drug delivery system has been enhanced tremendously. Many of these use polymer with different physical and chemical characteristics like biodegradability, biocompatibility etc. Recent advances in the field of micro fabrication have led to development of new class of controlled drug delivery systems.

The idea to develop a programmable, wirelessly controlled microchip that would deliver drugs after implantation in patients' body had come 15 years back to MIT professors (Robert Langer and Michael Cima). Which resulted as a first successful test by using programmable implants to deliver an osteoporosis drug called Teriparatide to 7 women aged 65 to 75 (Feb. 16 online edition of *Science Translational Medicine*, 2012).

Microchip as an implantable device can be used for the treatment and diagnosis. Due to its small size, microchip device may be implanted in the body in a variety of locations, including, but not limited to, under the skin and in the peritoneal cavity. The device may also be ingested for drug delivery or content exposure throughout the gastrointestinal tract.

IMPLANTABLE MICROSYSTEMS [16]

Implantable drug delivery devices are desired when prescribed drug regimen is critical. These devices deliver the drug at a specific rate without regular physician or patient intervention. Currently available implantable drugs are divided into two main categories, based on whether the drug is delivered in active or passive manner. Polymer depots are the most common passive drug delivery systems. Microchips as controlled drug delivery devices deliver the drug actively and passively both.

Microelectronic system should be able to:

[41]Student - M. Pharm. (IV Sem), Department of Pharmaceutics, Sinhgad Technical Education Society's Smt. Kashibai Navale College of Pharmacy, University of Pune, Maharashtra, India, harshitasingh1930@gmail.com

[42]Associate Professor, Sinhgad Technical Education Society's Smt Kashibai Navale College of Pharmacy, University of Pune, Maharashtra, India, meera.singh@sinhgad.edu

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

- Monitoring the physiological conditions of the patient body and its conversion to electric signals by means of physical and chemical transducers.
- Receive the electronic signals an analyze them which is properly controlled and regulated by microcontrollers.
- Release the appropriate amount of drug by help of micro actuators.

MICROCHIPS [4,5]

Microchip devices are provided which can accurately deliver precise quantities of molecules at defined rates and times according to the needs of the patient or experimental system. A 'microchip' is defined as a miniaturized device fabricated using forming methods such as compression molding, ejection molding, thermoforming etc.

The microchip devices can be described as "passive devices'' or ''active devices''. Both types control the rate and release of the molecules.

Each microchip device, whether passive or active, includes the following three components:

- Substrate
- Array of Reservoirs,
- A Release System.

These devices typically include hundreds to thousands of reservoirs, or wells, containing the molecules and a release system that controls the rate of release of the molecules. For example, the reservoirs may have caps made of a material that degrades at a known rate or that has a known permeability (passive release). Alternatively, the caps may include a conductive material capable of dissolving upon application of an electrical potential (active release).

MICROCHIP TECHNOLOGIES [1,9,14,15]

Microchip technologies for biological and pharmaceutical research fall into two general categories:

- Micro fabricated arrays of biopolymers - for performing solid-phase binding or hybridization experiments.
- Microfluidic devices - for performing high throughput biochemical or cell - based assays.

Much of the technological development has been focused in the areas of assay automation and miniaturization. Miniaturization reduces reagent consumption and facilitates higher-throughput assay formats by enabling the use of parallel sample-processing and multiplexed detection modes.

Micro fabricated Array Technology

The technology of micro fabricated peptide arrays were first described by Fodor et al in 1991 for use in drug discovery experimentation. This technology includes many biosensors and some DNA chips. Biosensors can be manufactured from silicon by using semiconductor processes. DNA chips uses immobilize materials on the device surface to identify genetic material or other chemicals. A combination of photolithography and solid-phase peptide synthesis was used to create spatially addressable arrays of peptides for use in multiplexed receptor–ligand binding assays.

Microfluidic Device Technology [6,9]

Microfluidic technology has emphasis on the miniaturization of fluid handling systems such as pumps, valves and flow channels. Microfluidic devices are typically designed in such a way to move gases or fluids through a series of various processing steps to perform solution phase chemical analysis.

A micro fabricated gas chromatograph that included a sample injector, separation column and thermal conductivity detector was described by Terry and colleagues in 1979 [Terry, S. *et al*. (1979)]. A gas chromatographic air analyzer fabricated on a silicon wafer.

Microfluidic technology is widely used for chemical or DNA analysis and combinatorial chemistry, bioassay and capillary electrophoresis systems.

MECHANISM OF MICROCHIP DEVICES: FABRICATION AND RELEASE [11,12,18]

Micro fabricated devices (microchips) have the ability to store and release multiple chemical substances on demand by a mechanism devoid of moving part.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

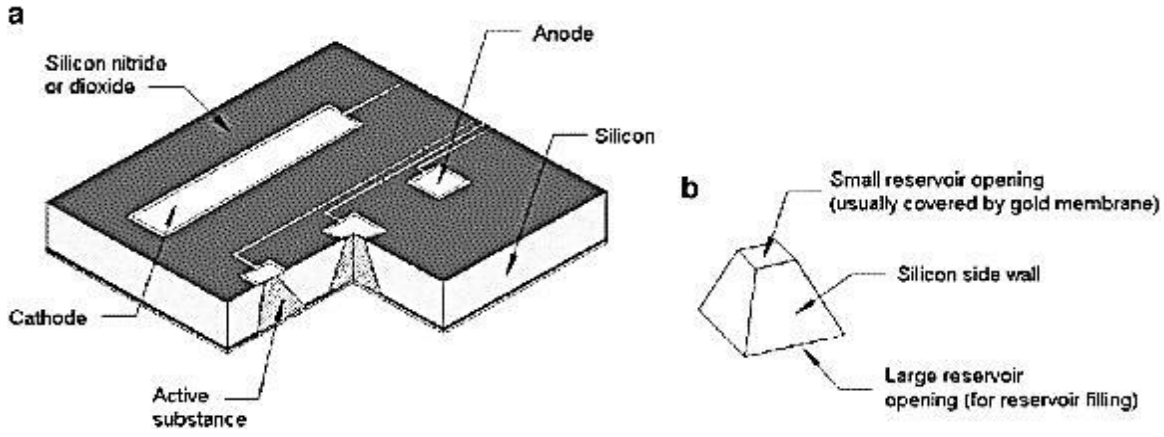
THEORY OF OPERATION [6,13,14,15]

Microchips consist of a series of reservoirs, which extends through substrate, which is impermeable for electrolyte. A material that is impermeable to the drugs or molecules which has to be deliver and the surrounding fluids is used as the substrate. Reservoirs are itched into the substrate. Reservoirs are sealed at one end by a thin membrane of material of semiconductor that serves as an anode. It dissolves when an electric potential is applied to it in an electrolyte. Cathodes are also fabricated on the device. Several reservoirs can be fabricated on a single microchip.

The reservoirs are filled through the open end by inserting a release system, which includes the molecules to be delivered by means of injection. Inkjet printing, spin coating methods, or other means. The reservoir can contain multiple drugs or other molecules in variable dosage. The filled reservoirs can be capped with materials that passively disintegrate, or materials that disintegrate upon application of an electric potential. Release from an active device can be controlled by a preprogrammed microprocessor, remote control or by biosensors.

After filling, the device is submerged in an electrolyte containing ions that forms a soluble complex with the anode material in its ionic form. An electric potential is applied to an anode membrane when release from its corresponding reservoir is desired. This causes oxidation of the anode material, formation of the soluble complex with the electrolyte ions, the complex then dissolves in the electrolyte, and the membrane disappears. The chemical in the newly opened reservoir is now exposed to the surrounding electrolyte and is free to dissolve in the electrolyte and diffuse out of the reservoir. In this way drug, release occurs from the microchip device.

Figure-1: A Solid State Microchip for Controlled Release (from Nature 397, 335–338 (1999) Macmillan Magazines Ltd)



Sources: Authors Compilation

Applications [2,10,15]

Microchip device system has a wide variety of applications in pharmaceutical science. Microchips can be used as therapeutic, prophylactic and diagnostic agent.

The active microchip serves as an implantable device for delivery of drugs in body for treatment or diagnosis purposes. Microchips have very small size, so it can be implanted in the body in a variety of locations like under the skin and in the peritoneal cavity. The device can also be ingested to facilitate the exposure of drug throughout the gastrointestinal tract.

The microchip systems are especially useful for drug therapies in which it is desired to control the exact amount, rate, and/or time of delivery of the drug. Preferred drug delivery application includes the delivery of potent compounds, including both small and large molecules, such as hormones, steroids, chemotherapy medications, vaccines, gene delivery vectors and some strong analgesic agents.

The microchips can be implanted via surgical procedures or injections or swallowed and can deliver many different drugs, at varying rates and varying times.

In one example, the microchip device is adapted for oral administration and the remote controller comprises a radio frequency transmitter. In another example, the microchip device is adapted for implantation onto or in the eye of a human or animal and the remote controller and / or power source comprises an ophthalmic laser.

MARKETED MICROCHIP PRODUCT [2,11,16]

Cylindrical silicon rubber can be used to encapsulate contraceptive steroids and can be implanted into the body.

One such product is currently being marketed under the trade name, **Norplant®,** by Wyeth pharmaceuticals (http://www.wyeth.com/products/all.asp).

Constant steroid release for up to five years has been shown by this. Once released, the drug diffuses through interstitial space of the surrounding tissue before reaching the bloodstream.

The Ohio-based company **ChipRx** is developing an integrated, self-regulating responsive therapeutic device. This device is of small matchstick size and it is designed to sense the physiological levels of metabolites, such as glucose. When a change is sensed or detected, the sensors send a signal to the batteries, which emits an electric charge. This electrical charge triggers the opening of a responsive material and allows the release of the desired therapeutic agent (e.g. insulin) from a reservoir. When glucose levels return to normal, the sensor stops the release of electrical charge from the battery, closing the reservoir and preventing the release of more insulin. This type of self-regulating therapeutic device or delivery system will truly revolutionize individualized patient care.

FUTURE APPROACH[16]

Now–a-days, various attempts are going on by the scientists to incorporate individual fluid handling components into integrated microfluidic circuits for expanding microchip functionality and simplify common laboratory procedures.

Recent applications have included microchip devices designed to perform electrophoretic separations of proteins and DNA biochemical and immunoassays, cell-based assays, cell sorting, PCR reactions, microfiltration, diffusion-based separations, and sample injection for mass spectrometry. Functional integration could be one of the most important requirements for future drug delivery systems.

Although no fully integrated, fully functional drug delivery device has yet been reported, the technology required to build such a system exists. Similar technologies are being studied for a number of other bioMEMS applications. For example, a microfluidic system has been developed for rapid delivery of small sample volumes to biosensors, which can assay samples taken from a bulk flow.

CONCLUSION

Microchips have an important approach in significant drug delivery techniques with higher specificity and efficacy in controlled drug delivery systems. At the micro scale level of this technology, we have the unique ability to engineer control over the cellular environment, leading to novel ways in which we can control molecular delivery and cell / tissue interactions. The micro fabricated device can be the core component of a fully implantable, intelligent system for the controlled release of potent therapeutic compounds by releasing the precise quantity of drug with higher specificity. This technology avoids requirement of frequent injections or complex dosing regimens. It includes pulsatile release, continuous release, or the release of multiple drugs. Microchips have made various technological wonders in the field of information technology, computer science and automation. This fascinating science though yet in infancy definitely offers a big promise and highest technological advance in delivering drugs to the patients for better healthcare.

REFERENCES

1. C. Montemagno, G. Bachand, S. Stelick, M. Bachand. (1999). "Constructing biological motor powered nanomechanical devices", ***Nanotechnology***, 10, 225–231.

2. Edmund Chen, B.Sc. (0T5), Tony Yeung, B.Sc. (0T6). (2003, March). The Controlled-release Microchip as a Drug Delivery Agent, ***Technology Review,*** 80 (2)

3. Elizabeth Lopatto, Microchip Implant makes Remote-Controlled Drug Delivery Possible, Feb 17, 2012

4. John T Santini, Amy C Richards, Rebecca A Scheidt, Michael J Cimaand Langer "Microchip technology in drug delivery", Massachusetts Institute of Technology, Cambridge, MA, USA, 2000, Vol. 32, No. 6 , Pages 377-379.

5. John T. Santini Jr., Michael J. Cima, Scott Albert Uhland, Thermally Activated Microchip Chemical Delivery Devices, US Patent No. US 6,669,683 B2, Date of Patent: Dec.30, 2003

6. John T. Santini, Jr. Jackson, Mich.; Michael J. Cima, Winchester; Robert S. Langer, Newton***, Fabrication of Microchip Drug Delivery Devices***, US Patent No. 6,123,861, Date of Patent: Sep. 26, 2000

7. Langer R., ***Drug delivery and targeting. Nature***, 1998, 392: supp, 5-10.

8. Langer R., ***Drug Delivery. Drugs on Target. Science***, 2001, 293(5527):58-59.

9. M.J. Madou, ***Fundamentals of Micro fabrication: the Science of Miniaturization***, (2nd Ed.) CRC Press, Boca Raton (2002).

10. Norman F. Sheppard, John T. Santini, Stephen J. Herman, Michael J. Cima, Robert S. Langer, Dennis Ausiello. ***Microchip Reservoir Devices using Wireless Transmission of Power and Data***, US Patent No. US 7,226,442,B2, Date of Patent: June 5, 2007

11. Robitzki AA, Kurz R. ***Bio sensing and drug delivery at the micro scale: Novel devices for controlled and responsive drug delivery***, Handb. Exp. Pharmacol, 2010; (197):87-112. doi: 10.1007/978-3-642-00477-3_3

12. Sarah L. Taoa, Tejal A. Desai. ***Micro fabricated drug delivery systems: from particles to pores***. 24 February 2003, Pages 315-328.

13. Santini JT, Cima MJ, and Langer R. (1999). ***A controlled-release microchip***. Nature. 97:335-338.

14. Santini JT, Richards AC, Sheidt RA, et al. (2000). ***The Finnish Medical Society Duodecim***, Ann Med. 32:377-379.

15. Santini Jr Jt, Richards AC, Scheidt R, Cima MJ, Langer R. ***Microchips as Controlled Drug Delivery Devices***, Angew Chem Int Ed Engl., 2000 Jul 17;39(14):2396-2407

16. Steven A. Sundberg, Andrea Chow, Theo Nikiforov, H. Garett Wada. ***Microchip based systems for target validation and HTS***, Drug Discovery Today, Volume 5, Issue 12, Supplement 1, December 1st, 2000, Pages 92-103

17. S. Zafar Razzakia, Prasanna K.Thawra, Ming Yanga, Victor M. Ugaz, Mark A. Burns. ***Integrated microsystems for controlled drug delivery***, Advanced Drug Delivery Reviews, Volume 56; Issue 2, 10 February 2004, Pages 185-198.

18. Terry S.C., Jerman J.H., Angell J.B., ***AGas Chromatographic Air Analyzer fabricated on a Silicon wafer***, Electron Devices, IEEE, Vol. 26, Issue 12, Dec 1979, Page No. 1880-1886.

19. Retrieved from http://www.google.com.mx/patents/US6123861

20. Retrieved from http://www.google.nl/patents/US6123861

21. Retrieved from http://www.faqs.org/patents/app/20080221555

22. Retrieved from http://www.google.com/patents/US6123861

23. Retrieved from http://www.businessweek.com/news/2012-02-22/microchip-implant-makes-remote-controlled-drug-delivery-...

24. Retrieved from http://www.bloomberg.com/news/2012-02-16/microchip-implant-makes-remote-controlled-drug-delivery-pos...

25. Retrieved from http://ihelpothersthrive.com/drugs-by-microchip-implant-how-to-not-empower-your-patients

26. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/1521-3773(20000717)39:14%3C2396::AID-ANIE2396%3E3.0.CO;2-...

27. Retrieved from http://www.google.it/patents/US7226442

28. Retreived from http://www.google.com/patents/US20080083041

29. Retrieved from http://www.igg-light-chain-variable-region.com/2014/06

30. Retreived from http://www.igg-light-chain-variable-region.com/olaparib-the-future-of-soft-lithography-involves-the-...

****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

AUTOMATED RATING OF PRODUCT REVIEWS

Rajat Shah[43] Saurabh P. Chamedia[44]

ABSTRACT

In recent years, there has been huge growth in the field of online reviews of products on the web. These reviews are helpful for potential buyers who are looking for general outlook for others who purchased the same product. Due to the very vague and unstructured format in which these reviews are retrieved, each review might not consider all the features of the product. In addition, with increasing number of reviews for each product, it becomes tedious for a user to go through and identify all the key features to consider while buying a product.

In this paper, we study a number of approaches for each subtask involved in product review sentiment analysis. Furthermore, an end-to-end study of all the steps involved in developing a review processing system is presented. We have mentioned techniques used by us for each subtask during the development of prototype review processing system. Finally, we have embarked upon a promising way for improving currently available sentiment analysis techniques, which outperforms all of them, though by a small margin. Also standardized system performance measure for review processing system does not exist in the literature during the time when this paper was published as per best of our knowledge, which also needs to be looked upon.

KEYWORDS

Text Classification, Feature Extraction, Feature Rating, Sentiment Analysis, Opinion Mining, Semantic Orientation, Polarity Detection, Machine Learning etc.

INTRODUCTION

Day by day technology is advancing and overcoming challenges, which were not tackled by. With the advent of Web 2.0, e-commerce and social web that emphasize the participation of users, Websites such as Amazon (www.amazon.com) encourage users to express opinions on products by posting overall ratings and textual reviews. These ratings are often used by recommender systems to recommend highly rated products, assisting users in making decisions. In addition, users are willing to share their experiences and thoughts about products or services they have used in order to assist others to make better decisions. This has resulted into generation of vast amount of unstructured text data. Manually dealing with this data is not intuitive and possible.

Automatic efficient text processing techniques needs to be developed. This field comes under broad field of Natural Language processing and Data Mining. Information extraction from unstructured text is one of the challenges being tackled by current research in the field of Natural Language Processing and Data Mining.

In the field of information extraction from product reviews, most of the work has focused on finding the opinion words related to predefined attributes. Recently, there has been growing interest in the automated learning of the attributes themselves and then finding the associated opinion words.

Consider you have a random review and you want to determine the class of product to which that review belongs. Class of review can be determined using machine learning classifiers used for text classification in existing literature. After this, we want to extract the features of product about which user has commented on automatically. This task of extracting features automatically involves taking into consideration the natural structure of language like extracting noun phrases and then statistically determining whether this phrase is a product attribute or not.

Once the product attributes are extracted, the next task is to extract sentiment words in the review used to describe those attributes and then determining polarity of those sentiment words. Now we know the product category of review, attributes of product about which user has commented on, sentiment words used for describing those attributes and polarity of sentiment words.

Next task is to summarize the overall result for each product and associated features after examining the complete set of reviews.

We decompose the problem of opinion mining (Sentiment Analysis and Feature Extraction) into following main subtasks:

[43]Student, Department of Computer Science and Engineering, Visvesvaraya National Institute of Technology, Maharashtra, India, rajat.shah200792@gmail.com

[44]Student, Department of Computer Science and Engineering, Visvesvaraya National Institute of Technology, Maharashtra, India, chamedia.saurabh@gmail.com

Review Class Identification

Given an arbitrary product review, our task is to determine particular product to which that reviews belongs. Here we have used & analysed performances of classic machine learning approaches for text classification like Naïve Bayes, Support Vector Machine, Multi-class Logistic Regression.

Feature Extraction

Normally for a given product class, set of attributes are fixed and pre-determined by domain experts. But this is time consuming process and so we would like to automate this as well. Sometimes, it may be the case that reviewer might want to express his/her sentiment about some attributes which were not expected previously. So it's better to extract attributes of product from reviews itself. For this, we take help of POS tagger using which we extract noun phrases. Noun phrases, which are above statistically determined threshold, are considered as product features.

Feature Rating

Reviewer's express their sentiments about product feature using sentiment words like good, bad, excellent, unpleasant, etc. Here our task is to find out which sentiment words are used by reviewer about each product features extracted previously. Our approach for this is we extract adjectives used in neighborhood of noun phrases (product features). After this, we have used SentiWordNet (outcome of gloss classification method [22]) in order to determine the polarity of sentiment words (adjective) used by the reviewer.

Result Summarization

Here we have determined the combined effect of all sentiment words used in complete set of reviews for each feature of every product by summing up polarity of each sentiment word used.

PAPER OUTLINE

In the rest of the paper, we will first review related work on how various Machine Learning techniques for sentiment analysis can be utilized to assist and automate the various sub-tasks we have mentioned previously. We will then describe an approach used by us for each phase and describe the gain achieved with that approach. Then we will perform two sets of evaluation. First, we will present results from a simulation experiment to demonstrate how this system model can outperform various previous systems that utilize supervised learning approach, and how the system can be improved over time as more user-generated labels are collected to improve the classifier in the current system. Finally, we will discuss the implication of those results for the design of systems that rely on user-generated content in general, as well as the future direction of the current research.

RELATED WORK AND PREVIOUS APPROACH

Below we summarize the previous research, which we have found as interesting and helpful:

a) Categorizing Product Reviews

Basic approach for this task is to decide upon list of keywords for each product class and then find out which predetermined keywords exist in the given review and assign that product class to review for which this count of keywords is maximum. From previous research it has been shown that this basic approach is outperformed by well-established text classification approaches using supervised machine learning classifiers such as Naïve Bayes, SVM, etc. This classifiers use Bag of Words approach for classification. We have examined these different supervised machine learning classifiers such as Naïve Bayes, Multi-class Logistic regression and Support Vector Machine (SVM) and found that SVM outperforms other classifiers for our purpose. Below we discuss these classifiers.

1) Naïve Bayes

Assumes sentiment orientation independent of position and neighbor and uses bag of words to compute the category of a new test review [10].

Maximum likelihood estimates:

Prior probability:

$$P(c_j) = \frac{\text{doccount}(C = c_j)}{N_{doc}}$$

Likelihood, after Laplase (add-1) smoothing |V|: Vocabulary:

$$P(w_i \mid c_j) = \frac{count(w_i, c_j) + 1}{\sum_{w \in V} count(w, c_j) + |V|}$$

Classifying a new test review:

$$c_{NB} = \underset{c_j \in C}{\operatorname{argmax}} P(c_j) \prod_{i \in positions} P(w_i \mid c_j)$$

2) Multi-class Logistic Regression

Here we train a multiple binary classifiers one for each product class .In order to determine product class of new review; we run all classifiers on this review. We accept class of review as the class of classifier showing maximum probability.

Logistic regression hypothesis is defined as:

$$h_\theta(x) = g(\theta^T x)$$

Where function g is the sigmoid function. The sigmoid function is defined as:

$$g(z) = \frac{1}{1 + e^{-z}}$$

The cost function for binary logistic classifier is defined as follows:

$$J(\theta) = \frac{1}{m} \sum_{i=1}^{m} [-y^{(i)} \log(h_\theta(x^{(i)})) - (1 - y^{(i)}) \log(1 - h_\theta(x^{(i)}))]$$

and the gradient which will be used to minimize cost function is:

$$\frac{d J(\theta)}{d\theta_j} = \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)} - y^{(i)})) x_j^{(i)}$$

3) Support Vector Machine

Main idea of SVM is to find out the linear separating hyper plane (the optimal separating hyper plane (OSH)) between different categories, which maximizes the margin among these categories. One of the parameter that we need to choose is the selection of kernel. For our purpose, we have used a linear kernel because linear kernel is simplest among different kernels that could be used [11].

b) Feature Extraction

The subtask of extracting features generally involves usage of POS tagged words and taking into consideration only high frequency words (usually nouns).

One of the mostly cited papers in the field of "attribute" (feature) extraction is by *Hu et al [19] which* discuss how to use a frequency based approach in identifying the features of a product from product reviews. They extract the most frequently occurring nouns in the review by POS tagging it and then consider those nouns as a feature of a product to which that reviews belongs.

Ghani et al [20] have shown achievement in extraction of attribute-value pair using Naïve Bayes classifiers and co-EM. Since their work is focused on official product description from merchant sites, rather than on product reviews, these data usually have a very unique way of listing the attributes (like "Operating system: Windows Vista Home"), analysing this data is different from working with unstructured user reviews.

OPINE system developed by *Popescu et al* [21] also uses the dataset provided in [19].In this method, once the noun phrases are extracted based on some frequency cut-off, it computes Point wise Mutual Information (PMI) scores between the phrase and metonymy discriminators associated with the product class. Its novel use of relaxation labelling techniques for determining the semantic orientation of potential opinion words in the context of given features and sentences leads to high precision and recall on the tasks of *opinion phrase extraction* and *opinion phrase polarity extraction*.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

c) Feature Rating

Now, we have a set of product features (noun phrases) and we need to identify the opinion words that describe them. For this, we extract the adjectives that are within some fixed distance from each of the feature words. Thus, we get a full list of adjectives describing each of the features.

The logical output of this step is described below:

Product name: HTC One X
Feature name: Screen
Positive review count: 436
Negative review count: 672

Feature name: Battery Life
Positive review count: 239
Negative review count: 145

Previously used methods for opinion phrase polarity detection are:

***I. Hatzivassiloglou and Mckeown* [15]:** Based on the observation that conjunctions often inhibit a certain property. Example, "and" usually connects adjectives of same orientation whereas, "but" connects adjectives of opposite orientation. Various supervised Machine Learning and clustering algorithms are employed in training and testing the orientation of adjectives.

The system identifies and uses this indirect information in the following stages:

- All conjunctions of adjectives are extracted from the corpus along with relevant morphological relations.
- A log-linear regression model combines information from different conjunctions to determine if each two conjoined adjectives are of same or different orientation. The result is a graph with hypothesized same- or different-orientation links between adjectives.
- A clustering algorithm separates the adjectives into two subsets of different orientation. It places as many words of same orientation as possible into the same subset.
- The average frequencies in each group are compared and the group with the higher frequency is labeled as positive.

II. Turney and Littman [16]: This method involves bootstrapping from an initial small set of "seed" words.

- Positive seeds = {good, positive, nice, excellent}
- Negative seeds = {bad, negative, sad, boring}
- Now PMI is calculated for a term *t* with each seed term *ti.*

Its orientation is calculated as:

$$O(t) = \sum_{t_i \in S_p} PMI(t, t_i) - \sum_{t_i \in S_n} PMI(t, t_i)$$

Where PMI is defined as:

$$\mathrm{PMI}(word_1, word_2) = \log_2 \left(\frac{\mathrm{p}(word_1 \,\&\, word_2)}{\mathrm{p}(word_1)\ \mathrm{p}(word_2)} \right).$$

III. Kamps et al [17]: Lexicon relations defined in WordNet are used as basics for finding orientation of an opinion. A graph is constructed with initial set of words (generally just "good" and "bad") and then recursively, for every synonym, we add an edge between the words. Finally, in order to determine the orientation, we use the following equation.

$$SO(t) = \frac{d(t, \mathrm{bad}) - d(t, \mathrm{good})}{d(\mathrm{good}, \mathrm{bad})}$$

Where d (t, ti) means the shortest distance between term *t* and *ti* in the graph.

If SO (t) >0, we say the adjective belongs to positive.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

For example, using Word-Net the word 'honest' gets assigned the value 1 as follows:

$$EVA(honest) = \frac{d(honest,bad) - d(honest,good)}{d(good,bad)} = \frac{6-2}{4} = 1$$

INNOVATIVE AND EFFICIENT APPROACH

Since the task of product categorization had already good accuracy from previous research, we have not explored it further. We have focused on Feature Extraction and Feature Rating. This are described below:

Automatic Feature Extraction

Though the systems developed by implementing various methods described in previous section are practical enough for most of the purposes, their fundamental assumption that features can only be unigram nouns is not completely correct. There can be multi word features like "optical zoom", "hot shoe flash" where one of the words is an *adjective* or features which are not nouns or *are combination of different POS tags*. Considering these aspects provides significant improvement in the performance when measured in terms of recall, precision and F-Measure.

For constructing the feature list upon which our product will be trained on, if we only include the frequent nouns, with threshold of say 0.5% in Reuter's dataset, we fail to identify the features (such as optical zoom). Hence, we also include the words, which are adjectives, i.e. words POS tagged with (JJ/JJR/JJS) in out feature list. On doing this, we figured out that this approach also adds words such as "good", "bad" which are in fact opinion words. To filter out such terms, we consider only those terms, which occurred as noun at least once in our dataset, or in case of a word phrase, only those phrase whose one of the words appear as noun somewhere in the dataset.

We trained the classifiers using basic approach and this modified approach and obtained accuracy gain of approximately 3.26%.

Feature Rating: Determining Polarity Using Gloss Classification

This unique approach for rating of each feature is to determine sentiment words (mostly adjective) used to describe that feature and then determine polarity of each associated adjectives using dictionary annotations for those adjective. This method is based on intuition that words having similar/related meaning have same kind of dictionary meaning/annotation.

We have utilized logistic classifier with feature set consisting of 1-grams and 2-grams. For determining polarity of adjective using bag of words approach. Polarity of each adjective is on scale of 0-1. We have used positive seed words as {good, excellent, charming, superior, positive and pleasing} with polarity of 1 and negative seed words as {bad, worst, dull, unexpected, unintuitive and inferior} with polarity of 0. We have extracted thesaurus gloss of this seed words and applied preprocessing steps such as tokenization, stemming, stop words removal, etc. and then trained logistic classifier using feature set of 1-grams, 2-grams with binary values (1 if token is present,0 if not).We have used WorldNet for synonyms expansion. We have trained our classifier using seed words and synonyms got using WordNet expansion for each seed words.

Now given an adjective of unknown polarity, we find its gloss from thesaurus and then apply same preprocessing steps used during classifier training. We then convert output of preprocessing step into feature vector (x) for our classifier. Once feature vector is created, we predict polarity of this adjective as:

$$h_\theta(x) = g(\theta^T x).$$

Where 0 <= h(x) <= 1

$$g(z) = \frac{1}{1 + e^{-z}}$$

Where theta is learned parameter after training classifier.

Consider adjective word 'easy'. First step involved is to find synonyms of this seed word using WordNet. So it synonyms comes out to be: {Uncomplicated, undemanding, unexacting, unchallenging, effortless, painless, facile, simple, straightforward ...}. Actually, list is more elaborate than shown. Here, we expand list-using recursion so; synonyms of each synonym are also considered up to recursive depth of five.

Now using meaning of each synonym from thesauras, we merge them. In our example, result of merging meanings of different synonyms look like: {easy; achieved without great effort; presenting few difficul-ties; free from worry or problems; uncomplicated; simple or straightforward; effortless; requiring no effort; achieved with admirable ease; …}

Then we create feature vector after preprocessing this text of merged meanings. After that, we use trained classifier to predict polarity of 'easy'. Our trained classifier predicts value of 0.63. So from this value, we can predict polarity of 'easy' as being more positive as expected.

We consider the value predicted by classifier as measure of positive degree of adjective. If classifier predicts 1, that indicates classifier is sure that given adjective is having positive polarity equal to that of positive seed words and if classifier predicts 0, that indicates classifier is sure that given adjective is having negative polarity equal to that of negative seed words.

Number of total positive & negative seed words used to train our classifier were 3568 and 3895 respectively.

Once polarity of each adjective used as sentiment word is determined, we predict overall rating of each feature as sum of polarity of individual adjectives.

EXPERIMENTS

Before applying any specific techniques to reviews, we have performed preprocessing steps like tokenization, stemming, stop words removal, POS tagging. For our experimentation purpose, we have used python as language of choice with its standard libraries. Data set used to study product category identification is Reuter. For determining semantic orientation of an adjective we have used, Thesaurus for meaning extraction and WordNet for getting synonyms.

Results show that the overall performance in terms of precision and recall increases with usage approach we have used. Since bigram words like *"optical zoom"* can also be a feature, so considering bigrams as a candidate features results in better accuracy as compared to simple single word features.

The associated opinion word, mostly adjective, for the extracted feature is found by looking at vicinity of distance 4, increasing this number decreased the accuracy as the sentence becomes complex and multiple features might exists in the same sentence like *"Even though its service is not bad, I can't believe people like Product1 given its pathetic battery life*".

Use of Gloss classification method in order to determine the polarity of the adjective is so far the best way to assign weightage to each adjective for the corresponding feature.

The combined use of these methods provides the rating for a new review much closer to the one expected.

Also summing up the rating of reviews for each product feature provides the overall rating for the product.

RESULTS

We have demonstrated and compared various approaches and phases in opinion mining and sentiment analysis on raw reviews available on various websites. We have also presented a novel approach for automated feature extraction from raw reviews which not only takes nouns into consideration, but also the adjectives phrases tagged by POS tagger.

Extracted feature list when used for training a classifier, produced much better results as compared to previously employed methods for feature extraction. Though the preprocessing time required in this approach was approximately 25.2% more, yet the improvement in results obtained overcomes this preprocessing time.

The Gloss classification method when used for rating of features on publicly available corpora out performs all the published methods for feature rating due to its unique and more systematic way of determining the polarity of sentiment words.

CONCLUSION AND SCOPE

Various steps involved in information extraction from a set of reviews have been studied and an improvised model for the same has been proposed.

- One possible avenue for improvement here is that domain knowledge can also be incorporated in product feature extraction. For example, Nokia, HTC, etc. can be added to the feature list for higher accuracy in classification.

- The feature rating approach defined by us recursively takes synonym's of seed word into consideration only up to a depth of 5. One possible study could be examining the effect of change in accuracy of rating of each feature by varying this depth.

ACKNOWLEDGMENT

Thanks for the support of professors at Visvesvaraya National Institute of Technology, Nagpur, India especially our project guide Prof. U. A. Deshpande. We are also grateful to ACM Digital Library and IEEEexplore for their enormous collection of crucial research papers.

REFERENCES

1. Ronen, Feldman. ***Communications of the ACM***, 56(4), 82-89. 10.1145/2436256.2436274.
2. Hatzivassiloglou, V., & McKeown, K. R. Predicting the Semantic Orientation Of Adjectives. (1997). ***In Proceedings of ACL-97, 35th Annual Meeting of the Association for Computational Linguistics, (pp. 174–181)***. ES: Madrid.
3. Turney, P. D., & Littman, M. L. (2003). Measuring Praise and Criticism: Inference of Semantic Orientation from Association. ***ACM Transactions on Information Systems***, 21(4), 315–346.
4. Kamps, J., Marx, M., Mokken, R. J., & Rijke, M. D. (2004). Using WordNet to Measure Semantic Orientation Of Adjectives. In Proceedings of LREC-04. ***In Proceedings of the 4th International Conference on Language Resources and Evaluation, volume IV, (pp. 1115–1118).*** PT: Lisbon.
5. Talwar, A., Jurca, R., & Faltings, B. (2007). Understanding User Behaviour in Online Feedback Reporting. ***In Proceedings of the 8th ACM Conference on Electronic Commerce (EC), (pp. 134–142).***
6. Minqing, Hu, & Bing Liu. (2004). Mining and Summarizing Customer Reviews. ***SIGKDD.***
7. Rayid, Ghani, Katharina, Probst, Yan, Liu, Marko, Krema, & Andrew, Fano. (2006). Text Mining for Product Attribute Extraction. ***SIGKDD.***
8. Ana-Maria, Popescu, & Oren, Etzioni. (2005). Extracting Product Features and Opinions from Reviews. ***EMNLP.***
9. Retrieved from http://www.ijsr.net/archive/v2i10/MDIwMTMzMzU=.pdf
10. Retrieved from http://ijltet.org/wp-content/uploads/2013/07/31.pdf

*****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

INTERCONNECTION OF ELECTRICAL POWER SYSTEM GRIDS VIA CLOUD

Uday Arun Deshpande[45]

ABSTRACT

This paper deals with the interconnection of electrical power system grids using cloud. This cloud includes the type of each user interface and the linkages between them.' The cloud provides path for ancillary services. "Cloud" classifies interconnections as, for example, bilateral, private exchange, or pool. Linkages between multiple grids may be implicit transfer relationships caused by arbitrage or explicit transfer activity in another.

Architecture should be specified before rules are written, but it is often necessary to test the architecture during the design process, and this requires a rough specification of the rules. Architectural design must also consider the grid structure in which it is embedded, which may inhibit the proper function. Connecting via cloud provides rigidly compartmentalized output, So, it will surely useful to consider the new methodology of interconnection of grid, as it will overcome the limitations presently prevailing in the interconnections of grids.

KEYWORDS

Power System Grid, Interconnections of Grids, Cloud, Bilateral Network Connection of Grids etc.

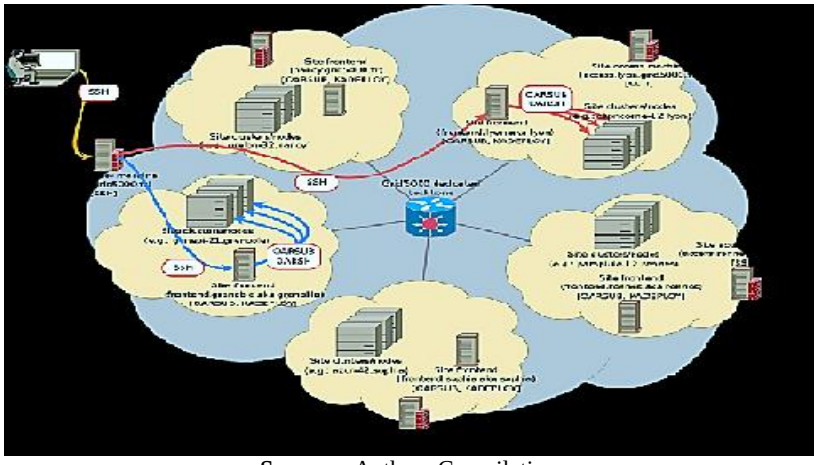
INTRODUCTION

A process has been underway worldwide for interconnection of grids with the use of internet to exchange data between them to maintain a healthy power flow throughout the connected electrical network. This process was usually carried with the help of heavy tie lines in early days then later it started to carry out with the use of optical fiber cables; then as technological advancement and rapid as well as spontaneous requirement and exchange of data between two or more grids led to introduction of transferring of data using radio frequencies; but with the advancement of grids as rapid increase in size and loads these techniques start to fall short so use of internet is introduced in order to cut short the data acquisition and processing time so as to keep the entire power system stable, healthy and running without any disturbances, with advent of smart grids, super smart grids, micro grids GPRS is been introduced so as to keep power system away from malfunctioning.

PROBLEM STATEMENT

The term interconnection is used in many ways and has no strict definition. Tirole (1997) explains that it should not be so sound as to encompass only a specific data produced at a specific location, should exchange nor should it be "the entire data." He concludes that there is no simple recipe. Now days as with the rapid increase of consumers and increase in complication and conjunction, with emerging of smart grids and micro grids the exchange of data between the various grids should be more spontaneous and rapid to meet the demands at consumer ends with keeping the generating plants in stable conditions.

Figure-1: Basic Grid Interconnection Architecture



Sources: Authors Compilation

[45]Student, Department of Electrical Power System, National Institute of Technical Teachers Training and Research, RGTU-Bhopal, Madhya Pradesh, India, udaycj@gmail.com

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

GAP / AREA OF IMPROVEMENT

As with above cited architecture it can be easily understood that the entire interconnection of various grids depends only on one back bone or broadly stated as data center from which the grids are connected via one single tie line because of which there is always possibility of overloading of that tie line which may cause conjunction, between data transfer between two grids and with advent of smart and micro grids with having distributed generation the problem of single tie line and single backbone may give rise to deadlock situation which may cause damages like blackout.

SOLUTION

Even though short-term and long-term data transfer transactions could be in bilateral forms in the electrical grid where connected grids may require spontaneous data for keeping the stability of power system constant. For this purpose, cloud can be employed.

Connecting the grids via cloud will provide ample amount of opportunities to any power system or power quality engineer working on any part of grid to maintain the grid in healthy condition, as cloud can be accessible from each and every point over the entire power system , it provides the data instantaneously as it has recorded at the instant and the rectification can be done real time, It also provides freedom from heavy and hazy architecture of tie lines, optical fiber cable network, and lessens the dependency on single backbone network.

CLOUD

Cloud computing ('cloud') is an evolving term that describes the development of many existing technologies and approaches to computing into something different. Cloud separates application and information resources from the underlying infrastructure, and the mechanisms used to deliver them. Cloud enhances collaboration, agility, scaling, and availability, and provides the potential for cost reduction through optimized and efficient computing.

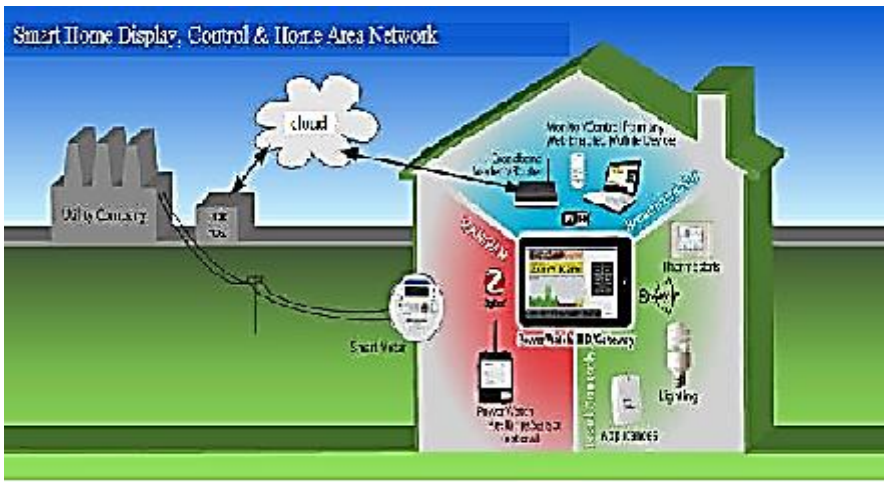
More specifically, cloud describes the use of a collection of services, applications, information, and infrastructure comprised of pools of compute, network, information, and storage resources. These components can be rapidly orchestrated, provisioned, implemented and decommissioned, and scaled up or down; providing for an on-demand utility-like model of allocation and Consumption.

PROPOSED RESEARCH WORK

To prepare a module using cloud as interface for connecting of two grids, after executing it will going to provide opportunities to implement it in interconnections of substations, grids.

EXPECTED OUTPUT STRUCTURE

Figure-2



Sources: Authors Compilation

BENEFITS OF ITREFACING WITH CLOUD

1. Instrumented
   Smart Connected Devices,

2. Interconnected
   Integrated Communication Network,
   System Integration platform,

3. Intelligent
   Application and analytics,

4. Instantaneous
   Speed is increased,

5. Secured
   Eliminates Threat of Physical Damage,
   Eliminates Condition of Conjunction and Dead locks.

REFERENCES

1. Retrieved from https://www.grid5000.fr/mediawiki/images/thumb/Grid5000_SSH_access.png/1000px-Grid5000_SSH_access.png
2. Retrieved from http://www.mahindrasatyam.com/industry-offerings/energy-and-utility-services/smart-grid-articles.asp
3. Retrieved from http://www.igbc.in/site/igbc/publication.jsp
4. Retrieved from https://apps.powergridindia.com/smartgrid/
5. Retrieved from http://www.greenbusinesscentre.org/datacenter/datacenterbook.pdf
6. Retrieved from http://www.rationalsurvivability.com/blog/2009/07/cloud-computing-security-architectural-framework/
7. Retrieved from http://www.iosrjournals.org/iosr-jce/papers/Vol4-issue2/I0424652.pdf
8. Retrieved from http://clean-clouds.com/2012/03/01/top-17-faqs-on-cloud-computing/

*****

PEZZOTTAITE JOURNALS

Volume 3, Number 2, April – June' 2014

ISSN (Print):2319-9016, (Online):2319-9024

THE CHALLENGES AND PROSPECTS IN IMPLEMENTING ERP IN PUBLIC DEVELOPMENT ENTERPRISE: A CASE STUDY OF ETHIO TELECOM ADAMA

Dr. S. M. Murali Krishna[46] Rediat Worku[47] Dr. P. Athma Karan Reddy[48] Dr. Y. Rajendra Prasad[49]

ABSTRACT

The overall objective of this study is to assess and examine the challenges and Prospects in implementing Enterprise Resource Planning in Public Development Enterprise in case of Ethio telecom. The research problem that this Project sought to address is to find out the challenges and Prospects in Implementing ERP system in Ethio telecom Adama Branch. The method of data collection includes semi-structured interview, English Version Questionnaire and Document Review and processed primary and secondary data sources using qualitative and quantitative data analysis techniques.

Accordingly, the key finding is that the most serious challenge is attitude of Employee about computer systems with an average result of 66.7%. Followed by Resistance by the End users is 61.5%, insufficient training of ERP operators 56.4%. Next to this, problem during file conversion is also a serious problem with an average response rate of 82%, size of the enterprise is with an average response rate of 56.4% and finally, complexity of ERP system to use is a serious challenge with an average response rate of 51.3%.

Based on these it is concluded that lack of adequate training about ERP system operation in the Enterprises would hinder the effective implementation and operation of the system. Moreover, when examining the survey data there seemed to be fewer consensuses that transformational goals have in fact been realized. There is more agreement that the relational and operational benefits have achieved because of ERP. Finally, based on the Major findings, conclusions were drawn and Certain Recommendations were forwarded.

KEYWORDS

Benefits, Challenges, Prospects, Critical Success factors, Ethio Telecom, Enterprise Resource Planning (ERP), Public Development Enterprises etc.

INTRODUCTION

An ERP system is an integrated software solution, typically offered by a vendor as a package that supports the seamless integration of all the information flowing through a company, such as financial, accounting, human resources, supply chain, and customer information (Davenport, 1998). ERP implementation is a lengthy and complex process, and there have been many cases of unsuccessful implementations (Parr and Shanks, 2000), which have had major impacts on business performance. As ERP plays a very important role in business, ERP implementation and its critical issues, success factors and implementation challenges have been investigated in the past (Parr and Shanks, 2000; Mated et al., 2003; Soho et al., 2000; Sumner, 2000).

Prior research has shown that conflict with consultants is one of the main managerial problems during the implementation period of ERP system (Themistocleous et al., 2001). Consultants can bring to the organisation specialised skills, experience, and know-how that the organisation needs when it is both time-consuming and expensive for it to build internally (Gable, 2003). They can also offer a firm-wide view, encourage unity between members, and they are usually neutral (Davenport, 1998). ERP implementation is by no means a purely technical system implementation, and will include Business Process Reengineering (BPR). Consultants can perform the role of change facilitator and are involved in very important knowledge transfer. Consulting firms use techniques such as guided learning, formal training and knowledge creation activities to direct clients to the necessary knowledge required for a successful implementation. This guidance saves the client considerable time and effort in knowledge search costs (Gable, 2003).

It has been also found that the mismatch between ERP and organization can have significant impacts on organizational adoption, and this could be the main reason causing the ERP implementation failure (Umbel et al., 2003). The need for greater customization of ERP software will increase in this case, and the risks associated with the ERP implementation will be much higher (Soho et al., 2000). According to Soho et al. 2000, there could be different levels of mismatch, namely business function, data and output. Careful selection and evaluation of ERP systems is required in order to reduce the potential risk of software mismatch.

[46]Professor, Department of Management & Accounting, Hawassa University, Ethiopia, drsmmk777@gmail.com
[47]Research Scholar, Department of Management & Accounting, Hawassa University, Ethiopia, athmakaran@yahoo.com
[48]Lecturer, Department of Management & Accounting, Hawassa University, Ethiopia, drsmmk777@gmail.com
[49]Associate Professor, Department of Management Studies, Krishnaveni Engineering College for Women, Andhra Pradesh, India, yagantirp@gmail.com

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Different ERP implementation phases are associated with specific ERP implementation problems (Markus et al., 2000). The ERP implementation literature has provided a solid theoretical background to ERP research. However, my review of literature suggests that there seems to be insufficient research investigating the failure factors of ERP implementation from planning to post ERP implementation in developing countries. Further in-depth research here seems justified in order to provide useful information for practitioners and a research framework for understanding critical factors and how those factors influence ERP implementation.

ERP implementation is a lengthy and complex process and it plays a very important role in business. Prior research has tended to focus upon large sized enterprises ERP systems. However, increasingly more small and medium sized enterprises (SMEs) are realizing that they would benefit from implementing ERP systems to improve their process efficiencies and enjoy the benefits of using integrated systems.

The size of the organization also has an impact on ERP implementation. For example, smaller companies spend a higher proportion of their budgets on the cost of software. Larger companies tend to spend a higher percentage on their ERP implementation teams. It was found that small companies that "best fit" with current business practices and package flexibility were key criteria in package adoption decisions. However, no further investigations or research has been conducted on how large and small companies select their ERP systems, or on implementation problems encountered during implementation processes.

As ERP implementation is more complex than other types of IS implementation, it would be appropriate to adopt the research framework proposed by Thong et al. with modifications, and with reference to other researchers' measurement of ERP implementation success. Modifications should be made concerning ERP implementation effectiveness. There are both common and distinct risk factors and project characteristics associated with IS and ERP implementation. Therefore, characteristics specific to ERP should be included in the implementation measurement.

Finally, this research study is exploratory in nature, and it is expected that by conducting study on Enterprise implementing ERP system, the Challenges and prospects in ERP implementation it can be identified. Conclusions may then be drawn regarding "what" and "how" organizational factors and other factors influence ERP implementation.

OBJECTIVES OF STUDY

General Objective

The overall objective of the proposed study is to assess and to examine the existing ERP implementation practice in line with Prospects and challenges prevailing in Ethio telecom Adama Branch.

Specific Objectives

Besides, the study will try to address the following specific objectives:

- To identify the major Motivational Factors to Implement ERP
- To identify the confronts and openings that exist in ERP implementation
- To identify the critical success factors for ERP implementation
- To Identify the major benefits of ERP Implementation
- To offer recommendations based on identified good practices and areas of improvement

Major Research Questions

- What are the prospects and challenges that exists in ERP Implementation?
- What are the critical success factors for ERP implementation?
- What are the major motivational factors in Implementation of ERP?
- What are the benefits of ERP Implementation?

INFORMATION COMMUNICATION TECHNOLOGY DEVELOPMENT IN ETHIOPIA

The 1960 has witnessed the advent of information technology in Ethiopia, when the mainframe computers were introduced for the first time in Africa. Despite being installed in the important sectors of military, defense, finance, telecom, electricity & power, and railways, the mainframes had no interactive boundaries within the sector. Later, microcomputers spread widely in the country either by direct purchase or through donor development aid. Like its predecessor, inadequate literacy and patchy understanding of the technology by potential users characterized the initial use of microcomputers. Donor agencies such as UNESCO and others first proposed a centralized approach to building information and communication technologies in Ethiopia as soon as microcomputers were introduced to the county. In 1986 national computing and information center was also established. However, the center due to high control did not realize the intended goals by the military government of the time, inadequate resources, high turnover, and shortage of skilled manpower and insufficient knowledge of information technology by decision makers. After the change of government in 1991, the efforts were resumed with regard to promoting the awareness and growth of

computers. A survey carried out in response to a study on national informatics policy in Ethiopia in 1993 covering 81 major institutions reports 1550 microcomputers, 38 mini computers and 8 mainframe computers. 25% of these machines were owned by international organizations and aid agencies. By the end of 1998, extensive growth in computers in business, academic institutions and households were seen raising the figure to an estimated 15,000 microcomputers. Awareness of the effect of information technology has been growing among users in the field of research, those working for international organizations and those in business. All of which has encouraged the IT sector in the country - proliferation of private companies that provide value added services in information technology, training centers and the establishment of a computer science unit at Addis Ababa University, efforts towards standardizing Ethiopic software, etc.

Public access links were introduced in 1990 and the Ethiopian Telecommunications Corporations (ETC) established the first Internet connection in October 1996 with a T1 connection to the United States. Currently there are over 2500 Internet users but the quality of service leaves much to be desired. Recent move of the government towards increasing the bandwidth to 1Mb line, privatization of energy and telecommunication sector could open further opportunities for growth in information and communication sector. However, surveys show as much as 83% of the capacity of a PC is often underutilized and the quality of services offered is often poor.

After the change of government in 1991, the Ethiopian People's Revolutionary Democratic Front (EPRDF)-led government adopted a new constitution (in 1994), which placed a special emphasis on access to information, freedom of expression and human rights. Under Article 29 of the Constitution, "Everyone shall have the right to freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through other media of his choice." However, internet availability in Ethiopia is still low. Only 360,000 people had internet access in 2008 – a penetration rate of 0.42%, one of the lowest in Africa compared to the sub-Saharan African average of 3.23% and Africa's average of 5.34%.

The government-owned Ethiopian Telecommunications Corporation is the sole internet service provider (ISP) in the country. Internet cafes are a major source of access in urban areas with a rising and active community of bloggers and users of other social networking tools. However, three quarters of the country's internet cafes are in the capital city, and even there access is often slow and unreliable. This report reviews access to information in Ethiopia, with a focus only on online access to information, and with reference to the relevant policy and legislative environments that are in place.

DATA ANALYSIS

Table-1: Responses of Employee towards Motivation of Implementation Issues

Expected Motivations for ERP Implementation	Responses										Mean
	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree		
	Freq	Per	Freq	Per	Freq	Per	Freq	Per	Freq	Per	
To replace the old system							11	28.2	28	71.8	4.72
To simplify and standardize the system							4	10.3	35	89.7	4.90
To keep up the globalization pressure					2	5.1	10	25.6	27	69.2	4.64
Ease of communication with suppliers					4	10.3	17	43.6	18	46.2	4.36
To restructure the organization			1	2.6	6	15.4	16	41	16	41	4.21
To meet management contract			10	25.6	7	17.9	15	38.5	7	17.9	3.49
To fulfill the strategic plan					10	25.6	17	43.6	12	30.8	4.05
To familiarize intl. best practices					3	7.7	16	41.0	20	51.3	4.44
To Improved logistical process							17	43.6	22	56.4	4.56
To Improved Management control					3	7.7	16	41	20	51.3	4.44
Grand Mean											4.38

Sources: Authors Compilation

The respondents were asked to rate the motivational factors for the implementation of Enterprise Resource Planning at Ethio telecom in five stages, strongly Disagree, Disagree, Neutral, Agree, Strongly Agree. As shown in Table-1 it provides data that shows the Motivational factors for ERP Implementation. In this data, most of the respondents indicate almost all Expected Motivations for ERP Implementation are agreed. The highest average response was to simplify and standardize the system with an average response of 4.90. The second highest average response was to replace the old system with an average result of 4.72. To keep up the globalization Pressure was the third highest Motivational Factor with an average response of 4.64. However, there was disagreement with the motivational factor for ERP implementation in the enterprise- To meet management contract specification with an average response of 3.49. This indicates most of the ERP operators and Designers are not aware of the management contract specification that France telecom and Ethio telecom agreed.

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

Table-2: Responses to Major Factors Hindering the Implementation of ERP in an Enterprise

Challenges in ERP Implementation	Responses					
	Most Serious		Serious		Not Serious	
	Freq	Per	Freq	Per	Freq	Per
Before Implementation						
Enterprise BPR Experience standard	14	35.9%	17	43.6%	8	20.5%
Management commitment to ERP system	1	2.6 %	12	30.8%	26	66.7%
Mismatch b/n organization and ERP systems	2	5.1%	10	25.6	27	69.2%
ICT Infrastructure in the enterprise	0	0 %	9	23.1%	30	76.9%
Major organizational structure changes	16	41%	15	38.5%	8	20.5%
During Implementation						
Problems created during file conversion	2	5%	32	82%	5	12%
Resistance by the user	24	61.5 %	9	23.1 %	6	15.4%
Size of the Enterprise	0	0%	22	56.4%	17	43.6%
After Implementation						
ERP led staff layoff	18	46.2%	16	41.0%	5	12.8%
Unrecognized Benefits of the system	3	7.7%	11	28.2%	25	64.1%
Attitude of Employee about computer system	26	66.7%	8	20.5 %	5	12.8 %
Complexity of ERP System to use	16	41 %	20	51.3%	3	7.7%
Training of ERP Users	22	56.4%	12	30.8%	5	12.8%

Sources: Authors Compilation

Most Serious Challenges

The respondents were asked to rate the factors that hinder the implementation of ERP at Public development enterprises in three stages, before implementation, during implementation and Post implementation. As shown in table 5.8, the majority of the respondents rated 'Most Serious' ERP implementation challenges are Attitude of Employee about computer systems 26/39 (66.7%), Resistance by the user 24/39 (61.5%) this indicate that human beings are resistant to change as ERP implementation is Change come from Information Technology, next to this Insufficient Training of ERP Users 22/39 (56.4%), ERP led staff layoff 18/39 (46.2%), organizational structure changes 16/39 (41%), complexity of ERP System to use 16/39 (41%) and Enterprise Business Process Reengineering (BPR) Experience 14/39 (35.9%).

Serious Challenges

Some of the respondents categorize the following challenges as serious problem during file conversion 32/39 (82%), Size of the enterprise 22/39 (56.4%), complexity of ERP Systems to use 20/39 (51.3%), Major organizational Structure Change 15/39 (38.5%), unrecognized benefits of the system 11/39 (28.2%) and Size of the Enterprise (27.2%).

Not Serious Challenges

ICT Infrastructure in enterprise 30/39 (76.9%), Mismatch between organization and ERP systems 27/39 (69.2%), Top Management commitment to ERP system 26/39 (66.7%). As shown in the above descriptions most serious, serious and not serious challenging factors that hinder the proper implementation of Enterprise Resources Planning Systems in Ethio telecom are presented. From the factors presented complexity of ERP System, Attitude of Employee about computer systems-even though, 66% of the ERP users youngsters, Resistance by the user are Highly Related to Insufficient Training of ERP Users. This shows that the characteristics of software, delivering proper and sufficient training are important to successful implementation of ERP.

Table-3: Responses to Major Prospects that Instigate the Implementation of ERP in an Enterprise

Major Prospects in ERP implementation	Responses					
	Not At all		To a Limited Extent		To Great extent	
	Frequency	%	Frequency	%	Frequency	%
Ease of selected ERP system	11	28.2%	28	71.8%	0	0%
Good consultant effectiveness	14	35.9%	23	59.0%	2	5.1%
Existence of expected IT infrastructure	9	23.1%	8	20.5%	22	56.4%
Vendor interest to transfer knowledge	12	30.8%	16	41.0%	11	28.2%
Project management skill of project team	0	0%	12	30.8%	27	69.2%
Top management commitment	0	0%	7	17.9%	32	82.1%
Having structured Enterprise	12	30.8%	18	46.2%	9	23.1%
Existence of good BPR implementation	9	23.1%	16	41.0%	14	35.9%
Government policy in system development	0	0%	13	33.3%	26	66.7%
Extensive top management support	3	7.7%	13	33.3%	23	59.0%

Sources: Authors Compilation

Volume 3, Number 2, April – June' 2014
ISSN (Print):2319-9016, (Online):2319-9024

As can be seen from Table-3 concerning the question about the prospects or opportunity in ERP implementation, the responses are categorized into three: 'Not at all', 'To limited extent', 'To great extent'; out of which Top Management Commitment to the Project Implementation 32/39 (82.1%), Project Management Skill of the Project team 27/39 (69.2%), Government Policy in system Development 26/39 (66.7%). Any system is implementation project can be triggered form three sources, Opportunity, Problem, Directive and these are forced by both internal and External Sources. Therefore, this finding indicates that form the listed prospects Ethio telecom has a positive Scenario in internal environment for system implementation.

RECOMMENDATIONS

The following recommendations are, therefore, basic to ensure that tangible progress could be made in ERP operation. Based on the major findings and the conclusions drawn from the study, the researcher forwards the following recommendations for both Ethio telecom and other Public development enterprises, which are interested in ERP implementation. The challenges faced by Ethio telecom in ERP implementation is categorized in three stages. Based on this, the researcher provides the following recommendations for each stage.

Before Implementation

- An organization, which needs to implement ERP system, should first restructure the business and apply Business Process Reengineering (BPR).
- An organization with an interest should match the ERP and organizational information flow.
- The ICT infrastructure of an organization should in a better position.
- Finally, commitment of Top management to the ERP implementation Project is needed

During Implementation

- Resistance by the user should be alleviated first by giving adequate on the job or off the job training for ERP operators to enhance the service operability.
- The system should be designed as per the file structure of the organization. In addition, this can alleviate the Problems created during file conversion.
- As the data presented, Size of the enterprise is also one success factor for ERP implementation. Therefore, organizations should adopt the phased approach to minimize the risk associated with system failure.

After Implementation

- As it is observed form the data presented, Attitude of Employee about Computer systems is also the challenging factor that hinders successful ERP implementation. Therefore, the organization should change the attitude of Employees towards computer system by delivering attitude-changing training to ERP operators.

CONCLUSION

On a final note, it is recommended that the Enterprise constitute a task force represented by key internal stakeholders that could employ this study as a foundation to embark on an extensive ERP operation audit and train employees for Successful ERP system operation.

The national impact of the System Automation of the Enterprise can only be meaningfully if government and ERP vendors are also committed in alleviating the factors affecting the ERP implementation and operation in Public Development Enterprises in Ethiopia. And finally as the goal of the ICT policy of Ethiopia is by effective utilization of ICT, bring about significant development nation-wide in all sectors and in all citizens walks of life, the government should participate in ERP development Issues and accomplish its goal. Should the above recommendations be implemented?, the researcher believes that the Enterprise will be in a much better position to Transfer Experience towards the empowerment of this nature-blessed country to escape from the poverty marsh by Transferring ICT driven knowledge and power.

REFERENCES

1. Abdinnour-Helm, S., Longneck-Hall, M. L., & C. A. Longneck-Hall. (2003). Pre-Implementation Attitudes And Organizational Readiness For Implementing An Enterprise Resource Planning System. ***European Journal of Operation Research***, 146(2), 258-273.

2. Adam, F., & P. O'Doherty. (2000). Lessons From Enterprise Resource Planning Implementations In Ireland – Towards Smaller And Shorter ERP Projects. ***Journal of Information Technology***, 15, 305-316.

3. Adolph, W. (1996). ***Cash cow in the tar pit: Reengineering a legacy system. IEEE Software***, 41-47.

4. Akkermans, H., & Van, Helden K. (2002) Vicious and Virtuous Cycles in ERP implementation: a Case Study of Interrelations between Critical Success Factors. ***European Journal of Information Systems***, 11(1), pp. 35-46.

5. Aladwani, A. M. (2002). An Integrated Performance Model of Information Systems Projects. ***Journal of Management Information Systems,*** 19(1), 185-210.

6. Al-Masher, M., Al-Mudimigh A, & Zaire, M. (2003). Enterprise Resource Planning Taxonomy of Critical Factors. ***European Journal of Operational Research***, 146, 352-364.

7. Al-Masher, M., & Al-Mudimigh, A. (2003). ERP Implementation: Lessons from a Case Study. ***Information Technology & People,*** 16(1), 21-23.

8. ***Ibid,*** 6.

9. ***Ibid,*** 6.

10. Awazu, Yakima & Desouza, Kevin C. (2003). Knowledge Management. ***HR Magazine***, 48(11), 107.

11. Bobbie, E, & Mouton J. (2001). ***The Practice of Social Research***. Cape Town: Oxford

12. Bancroft, N., Seip, H. & Sprengel, A. (1998). ***Implementing SAP R/3*** (2nd Ed.). Manning Publications, Greenwich.

13. Bergsma, K., & Kingman, S. (2005). Developing a Cultural Perspective on ERP. ***Business Process Management Journal,*** 11(2), 123-136.

14. Bradford, M., & Florin, J. (2003). Examining the Role of Innovation Diffusion Factors on the Implementation Success of Enterprise Resource Planning Systems. ***International Journal of Accounting Information Systems***, 4, 205-225.

15. Brown, C. V., & Vessey, I. (2003). Managing the Next Wave of Enterprise Systems: Leveraging Lessons from ERP. ***MIS Quarterly Executive,*** 2(1), 65-77.

16. Brown, David. (2002). E HR – Victim of Unrealistic Expectations. ***Canadian HR Reporter,*** 15(16), 1, 6.

17. Chung, S., & Snyder, C., (2000). ERP Adoption: A Technological Evolution Approach. ***International Journal of Agile Management Systems,*** 2(1).

18. Cliff, S. (1999). ERP implementation. ***Harvard Business Review,*** 77(1), 16-17.

19. Cohen, T., & South Wood, R. (2004). ***An Overview of VoIP Regulation in Africa: Policy Responses and Proposals.*** London: Commonwealth Telecommunications Organisation (CTO).

20. Collin, E. (1998, August 06). ERP market to reach US $ 52 billion by 2002 (Report says). ***Information Week***.

21. Cotteleer, M. J. (2002). ***ERP: Payoffs and Pitfalls.*** Harvard Business School Working Knowledge. Retrieved from http://hbswk.hbs.edu/item.jhtml?id=3141&t=operations .

22. (2009). ***National Population Abstract***. CSA (Central Statistics Authority).

23. Daniel, Kobo. (2001). ***How Ghanaian Migrates In Amsterdam Assist Relatives At Home*** (Research Report). Netherlands: African Studies Center.

24. Darlington, R. (2009***). How The Internet Could Be Regulated.***

25. Davenport, T. (1998). Putting the Enterprise In to The Enterprise System. ***Harvard Business Review***, 76(4), 121-131.

26. Davenport, T. (2000). ***Mission Critical***. Boston, MA: Harvard Business School Press.

27. Duplaga, E. A., & Astana, M. (2003). Implementing ERP in Manufacturing. ***Information Systems Management***, 20(3), 68-75.

28. (2009). ***ICT Assisted Development Project Monitoring and Evaluation Report on ICT Laws Enacted and ICT Business Status in Major Towns of Ethiopia*** (2008). Addis Ababa: EICTDA (Ethiopian ICT Development Agency).

29. (2008). ***Country Profile 2008.*** Ethiopia. London: EIU (The Economist Intelligence Unit).

30. Estevez, J., Pastor, J., & Casanovas, J. (2003). A Goal/Question/Metric Research Proposal to Monitor User Involvement and Participation ERP Implementation Projects. ***Information Resources Management Association Conference (IRMA),*** Philadelphia (USA) 325-327.

31. Fisher, D. M., Fisher, S. A., Kiang, M. Y., & Chi, R. T. (2004). Evaluating Mid-Level ERP Software. ***Journal of Computer Information Systems***, 45(1), 38-46.

32. Gable, G. (2003) Consultants and Knowledge Management. ***Journal of Global Information Management***, 11(3), 1-4.

33. Gardiner, S. C., Hanna, J. B., & Lamoure, M. S. (2002). ERP and the Re-Engineering Of Industrial Marketing Processes: A Prescriptive Overview for the New-Age Marketing Manager. ***Industrial Marketing Management***, 31, 357-365.

34. Geffen, D., & Ridings, C. M. (2002). Implementation Team Responsiveness and User Evaluation of Customer Relationship Management: A Quasi-Experimental Design Study of Social Exchange Theory. ***Journal of Management Information Systems***, 19(1), 47-69.

35. Haines, M. N., & Goodhue, D. L. (2003). Implementation Partner Involvement and Knowledge Transfer in the Context of ERP Implementations. ***International Journal of Human-Computer Interaction***, 16(1), 23-38.

36. Hong, K. & Y. Kim. (2002). the Critical Success Factors For ERP Implementation: An Organizational Fit Perspective. ***Information & Management,*** 40(1), 25-40.

37. Hunter, P. (2004). ***Survey Analysis. Centre for Survey Research and Analysis***. University of Connecticut. Hartford

38. Iraqi, Z., & Love, P. E. D. (2001). The Propagation of Technology Management Taxonomies for Evaluating Investments in Information Systems. ***Journal of Management Information Systems,*** 17, 161-178.

39. Kumar, R. L. (2002). Managing Risks in IT Projects: An Options Perspective. ***Information and Management,*** 40, 63-74.

40. Laukkanen, S., Scapula, S., & Hallikainen, P. (2005). ERP system adoption {Does the size matter. Pages 1-9 of: Sprague, R.H. (Ed), ***Proceedings of the 38th. Hawaii International Conference on System Sciences***. Computer Society Press.

41. Leopold, E., & Otieno, J. 2005. Critical Success Factors of ERP implementation. Pages 628{633 of: Encyclopedia of Information Science and Technology. IDEA Group.

42. Loha T. C., & Kohl S. C. (2004). Critical Elements for a Successful Enterprise Resource Planning Implementation in Small- And Medium-Sized Enterprises. ***International Journal of Production Research,*** 42(17), 3433–3455.

43. Markus, M. L., & Tanis, C. 2000). The Enterprise Systems Experience (From Adoption to Success). Chap. 10, pages 173{207 of: Mud, R.W. (Ed), ***Framing the Domains of IT Management: Projecting the Future through the Past.*** Cincinnati, OH. Pinna flex Publishing.

44. Mayfield, J., Mayfield, M., & Lance, S. (2003). Human Resource Information Systems: A Review and Model Development. ***Advances in Competitiveness Research***, 11(1), 139-151.

45. Murray, M., & Coffin, G. (2001). A Case Study Analysis of Factors for Success in ERP System Implementations**.** ***In Proceedings of the Seventh Americas Conference on Information Systems***, ***(pp. 1012–1018)***. Boston.

46. Myers, M.D., & Alison, D.E. (Eds). (2002). ***Qualitative Research in Information Systems: A Reader***. London: Sage Publications.

47. Nah, F. F., Zuckweiler, K. M. & J. L. Lau. (2003). ERP Implementation: Chief Information Officers' Perceptions of Critical Success Factors**.** ***International Journal of Human-Computer Interaction***, 16(1), 5-22.

48. Newell, S., Huang, J. C., Galliards, R. D. & S. L. Pan. (2003). Implementing enterprise resource planning and knowledge management systems in tandem: Fostering efficiency and innovation complementarily. Information ***and Organization,*** 13(1), 25-52.

49. Otieno, J. (2008). Enterprise Resource Planning (ERP) Systems Implementation Challenges A Kenyan Case Study. Pages 399{409 of: Abramowitz, W.; Fennel, D. (Ed), ***Lecture Notes in Business Information Processing,*** vol. 7. Springer.

50. Palani, Swamy. (2002). an Innovation-Diffusion View of Implementation of Enterprise Resource Planning (ER) Systems and Development of A Research Model. ***Information & Management,*** 40(2), 87-114.

51. Rebus, U. (2006). Critical Success Factors of Implementing Enterprise Portals. ***In Proceedings of the 39th Hawaii International Conference on System Sciences.***

52. Robbers, P. M. A., & K. School. (2002). Program Management and Complexity of ERP Implementations. ***Engineering Management Journal,*** 14(2), 45-52.

53. Robe, D., Ross, J. W., & M. Boudreau. (2002). Learning to Implement Enterprise Systems: An Exploratory Study Of The Dialectics Of Change. ***Journal of Management Information Systems***, 19(1), 17-46.

54. Shang, S., Sedan, P. B. (2002). Assessing and Managing the Benefits of Enterprise Systems: The Business Manager's Perspective. ***Information Systems Journal,*** 12, 271-299.

55. Somers T. M., & Nelson K. G. (2004). Taxonomy of Players and Activities across the ERP Project Life Cycle. ***Information and Management,*** 41(3), 257–278.

56. Tilley, C. M., Bruce, C. S., & Hall am, G. C. (2007). ***Adaptive Technology for People with Disabilities Using Information and Communications Technology (ICT). Improving Library Services to People with Disabilities,*** pp. 65-86. Chandos Publishing.

57. UNDP (United Nations Development Programme). 2009. Human Development Report. Overcoming Barriers: Human Mobility and Development. United Nations Economic Commission for Africa (2009) Implementing the World Summit on the Information Society Action Lines: Analysis of Country Reports, UNECA, Addis Ababa.

58. Wei, C., & Wang, M. (2004). A Comprehensive Framework for Selecting an ERP System. ***International Journal of Project Management***, 22, 161–169.

59. (2008). Legal and Regulatory Issues in the Information Economy/ Censorship or Content Regulation. ***Wiki books.*** Retrieved ON 5 July 2009, from http://www.en.wikibooks. org/wiki/Legal_and_Regulatory_Issues

60. Yin, R. (2003). ***Case Study Research: Design and Methods***. Sage, London.

61. Henderson, J. C., & Venkatramanm, N. (1993). Strategic alignment: Leveraging information technology for transforming organizations, ***IBM Systems Journal***, 32 (1), 4-16.

62. Zhang L, K. O. Lee K.O, & Banerjee, P. (2002). Critical Success Factors of Enterprise Resource Planning Systems Implementation Success in China. ***In Proceedings of the 36th Hawaii International Conference on System Sciences.***

63. Retrieved from http://www.eitpa.org/

64. Retrieved from http://www.onlinenigeria.com/articles/ad.asp?blurb=554

65. Retrieved from http://blog.trginternational.com/trg-in-the-board-room/bid/192533/Critical-Failure-Factors-in-ERP-Im.. .

66. Retrieved from http://www.studymode.com/essays/Sap-Issues-572827.html

67. Retrieved from http://www.ehow.com/how_5195052_identify-critical-success-factors.html

68. Retrieved from http://www.positivearticles.com/Article/INFORMATION-TECHNOLOGY-GROWTH-IN-ETHIOPIA/25343

69. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-38244-4_11

70. Retrieved from http://www.cibmp.org/journals/index.php/ijib/article/view/8

71. Retrieved from http://thermalscience.vinca.rs/2010/3/7

72. Retrieved from http://www.giswatch.org/sites/default/files/Ethiopia.pdf

73. Retrieved from http://www.biomedcentral.com/1756-0500/6/106

74. Retrieved from http://link.springer.com/article/10.1186/1756-0500-6-106

75. Retrieved from http://tapironline.no/fil/vis/367

76. Retrieved from http://my.safaribooksonline.com/book/software-engineering-and-development/enterprise/9781599048260/t...

*****