Global Journal of Computer Science and Technology

discovering thoughts and inventing future

Volume 10 Issue 15 Version 1.0 Online ISSN: 0975-4172 Print ISSN: 0975-4350

12 Technology Reforming Ideas

highlights

- CDMA Wireless Communication
- Expression Cancer Diagnosis
- Software Risk Management
- Computer Assisted Instruction

December2010

ENG

Global Journal of Computer Science and Technology

Global Journal of Computer Science and Technology
Volume 10 Issue 15 (Ver.1.0)

Global Association of Research

Incorporation No.: 0423089
License No.: 42125/022010/1186
Registration No.: 430374
Import-Export Code: 1109007027
Employer Identification Number (EIN):
USA Tax ID: 98-0673427

Global Journals Inc.

(A Delaware USA Incorporation with "Good Standing"; Reg. Number: 0423089)

Sponsors:Global Association of Research

Open Scientific Standards

Publisher's Headquarters office

Global Journals Inc., Headquarters Corporate Office, Cambridge Office Center, II Canal Park, Floor No. 5th, ***Cambridge (Massachusetts)***, Pin: MA 02141 United States
USA Toll Free: +001-888-839-7392
USA Toll Free Fax: +001-888-839-7392

Offset Typesetting

Global Journals Inc., City Center Office, 25200 Carlos Bee Blvd. #495, Hayward Pin: CA 94542 United States

Packaging & Continental Dispatching

Global Journals, India

Find a correspondence nodal officer near you

To find nodal officer of your country, please email us at *local@globaljournals.org*

eContacts

Press Inquiries: *press@globaljournals.org*
Investor Inquiries: *investers@globaljournals.org*
Technical Support: *technology@globaljournals.org*
Media & Releases: *media@globaljournals.org*

Pricing (Including by Air Parcel Charges):

For Authors:
22 USD (B/W) & 50 USD (Color)
Yearly Subscription (Personal & Institutional):
200 USD (B/W) & 500 USD (Color)

Editorial Board Members (HON.)

Dr. Bart Lambrecht
Director of Research in Accounting and FinanceProfessor of Finance
Lancaster University Management School
BA (Antwerp); MPhil, MA, PhD (Cambridge)

Dr. Carlos García Pont
Associate Professor of Marketing
IESE Business School, University of Navarra
Doctor of Philosophy (Management), Massachusetts Institute of Technology (MIT)
Master in Business Administration, IESE, University of Navarra
Degree in Industrial Engineering, Universitat Politècnica de Catalunya

Dr. Fotini Labropulu
Mathematics - Luther College
University of ReginaPh.D., M.Sc. in Mathematics
B.A. (Honors) in Mathematics
University of Windso

Dr. Lynn Lim
Reader in Business and Marketing
Roehampton University, London
BCom, PGDip, MBA (Distinction), PhD, FHEA

Dr. Mihaly Mezei
ASSOCIATE PROFESSOR
Department of Structural and Chemical Biology, Mount Sinai School of Medical Center
Ph.D., Etvs Lornd University
Postdoctoral Training,
New York University

Dr. Söhnke M. Bartram
Department of Accounting and FinanceLancaster University Management SchoolPh.D. (WHU Koblenz)
MBA/BBA (University of Saarbrücken)

Dr. Miguel Angel Ariño
Professor of Decision Sciences
IESE Business School
Barcelona, Spain (Universidad de Navarra)
CEIBS (China Europe International Business School).
Beijing, Shanghai and Shenzhen
Ph.D. in Mathematics
University of Barcelona
BA in Mathematics (Licenciatura)
University of Barcelona

Philip G. Moscoso
Technology and Operations Management
IESE Business School, University of Navarra
Ph.D in Industrial Engineering and Management, ETH Zurich
M.Sc. in Chemical Engineering, ETH Zurich

Dr. Sanjay Dixit, M.D.
Director, EP Laboratories, Philadelphia VA Medical Center
Cardiovascular Medicine - Cardiac Arrhythmia
Univ of Penn School of Medicine

Dr. Han-Xiang Deng
MD., Ph.D
Associate Professor and Research
Department Division of Neuromuscular Medicine
Davee Department of Neurology and Clinical NeuroscienceNorthwestern University
Feinberg School of Medicine

Dr. Pina C. Sanelli
Associate Professor of Public Health
Weill Cornell Medical College
Associate Attending Radiologist
NewYork-Presbyterian Hospital
MRI, MRA, CT, and CTA
Neuroradiology and Diagnostic Radiology
M.D., State University of New York at Buffalo,School of Medicine and Biomedical Sciences

Dr. Roberto Sanchez
Associate Professor
Department of Structural and Chemical Biology
Mount Sinai School of Medicine
Ph.D., The Rockefeller University

Dr. Wen-Yih Sun
Professor of Earth and Atmospheric SciencesPurdue University Director
National Center for Typhoon and Flooding Research, Taiwan
University Chair Professor
Department of Atmospheric Sciences, National Central University, Chung-Li, TaiwanUniversity Chair Professor
Institute of Environmental Engineering, National Chiao Tung University, Hsin-chu, Taiwan.Ph.D., MS The University of Chicago, Geophysical Sciences
BS National Taiwan University, Atmospheric Sciences
Associate Professor of Radiology

Dr. Michael R. Rudnick
M.D., FACP
Associate Professor of Medicine
Chief, Renal Electrolyte and Hypertension Division (PMC)
Penn Medicine, University of Pennsylvania
Presbyterian Medical Center, Philadelphia
Nephrology and Internal Medicine
Certified by the American Board of Internal Medicine

Dr. Bassey Benjamin Esu
B.Sc. Marketing; MBA Marketing; Ph.D Marketing
Lecturer, Department of Marketing, University of Calabar
Tourism Consultant, Cross River State Tourism Development Department
Co-ordinator , Sustainable Tourism Initiative, Calabar, Nigeria

Dr. Aziz M. Barbar, Ph.D.
IEEE Senior Member
Chairperson, Department of Computer Science
AUST - American University of Science & Technology
Alfred Naccash Avenue – Ashrafieh

Contents of the Volume

From the Chief Author's Desk

We see a drastic momentum everywhere in all fields now a day. Which in turns, say a lot to everyone to excel with all possible way. The need of the hour is to pick the right key at the right time with all extras. Citing the computer versions, any automobile models, infrastructures, etc. It is not the result of any preplanning but the implementations of planning.

With these, we are constantly seeking to establish more formal links with researchers, scientists, engineers, specialists, technical experts, etc., associations, or other entities, particularly those who are active in the field of research, articles, research paper, etc. by inviting them to become affiliated with the Global Journals.

This Global Journal is like a banyan tree whose branches are many and each branch acts like a strong root itself.

Intentions are very clear to do best in all possible way with all care.

Dr. R. K. Dixit
Chief Author
chiefauthor@globaljournals.org

Release and Deployment Management using ITIL

G.Rasa[1], S.Jagadheesh Kumar[2], Dr.R.S.D.Wahida Banu[3]

GJCST Classification D.2.9

Abstract **- Release management is the process of determining, acquiring, releasing and deploying changes into an Information Technology (IT) environment. The Service Delivery management strategy provides support to the Service Support process and ensures that the required services are delivered to the customer on time. Information Technology Infrastructure Library (ITIL) is a framework available in the form of documentation that defines the best-practices and approaches to manage IT services. This paper describes how to release a change into the IT environment using ITIL.**

Keywords- ITIL, Change Management, Configuration Management, Release Management, Deployment Management.

I. INTRODUCTION TO ITIL

ITIL is the most widely adopted approach for IT Service Management in the world. It provides a practical framework for identifying, planning, delivering and supporting IT services to the business This is a compendium of best practices from many companies in many industries. It represents the best thinking of thousands of people about how IT should be run, what the impact IT can have on the business it supports, and how to gain the most value from your IT investments. It provides guidance to organizations on how to use IT as a tool to facilitate business change, transformation and growth. One of the stated goals of the ITIL is to help decision makers make better decisions by ensuring the adequate IT information is available to support those decisions.

Adopting ITIL can offer users a huge range of benefits that include:

- improved IT services
- reduced costs
- improved customer satisfaction through a more professional approach to service delivery
- improved productivity
- improved use of skills and experience
- improved delivery of third party service.

ITIL is the product of the Office Government of Commerce (OGC), United State. ITIL originally emerged in the mid to late 1980s. The CCTA (Central Computer and Telecoms Agency) was a major UK government department, with an IT budget of around £8 billion. ITIL fulfill the ISO standard: ISO/IEC 20000. There are three versions of ITIL available which are 1.0, 2.0 and 3.0 . Among them, version 2 and 3 are the key versions. The difference between the two versions is a few changes in service life cycle structure [1]. Fundamentally, ITIL is exactly what its name implies "a collection of books" The common theme of the library is that all of the books provide guidelines that can help organizations implement the best practices that have been learned the hard way by the pioneering few. The library continue to grow as more successful techniques are documented and guidelines established for what can make others successful.

1) *ITIL version 2.0*

In 2000/2001, to make ITIL more accessible, ITIL v2 consolidated the publications into 8 logical "sets" that grouped related process-guidelines to match different aspects of IT management, applications, and services. However, the main focus was known as the Service Management sets (Service Support and Service Delivery) which were by far the most widely used, circulated, and understood of ITIL v2 publications.

- In April 2001 the CCTA was merged into the Office of Government Commerce (OGC), an office of the UK Treasury.[2]
- In 2006, the ITIL v2 glossary was published.
- In May 2007, this organization issued the version 3 of ITIL (also known as the ITIL Refresh Project) consisting of 26 processes and functions, now grouped under only 5 volumes, arranged around the concept of Service lifecycle structure.
- In 2009, the OGC officially announced that ITIL v2 would be withdrawn and launched a major consultation as per how to proceed.[3]

a. Service Support
b. Service Delivery
c. ICT Infrastructure Management
d. Planning to Implement Service Management
e. Application Management,
f. The Business Perspective Management
g. Security Management[4]

2) *ITIL Version 3.0*

The following are the five guides comprise the ITIL v3, published in May 2007:

a. ITIL Service Strategy
b. ITIL Service Design
c. ITIL Service Transition
d. ITIL Service Operation
e. ITIL Continual Service Improvement

About[1] - Research Scholar, Dravidian University, Kuppam, AP, India.
About[2]- Project Manager, IBM India Pvt. Ltd, Bangalore, India. Research Scholar, Dravidian University, Kuppam,KA,India.
About[3]- Professor & Head, Department of Electronic and Communication Engineering, Government College of Engineering, Salem,TN, India.

The five core guides map the entire ITIL Service Lifecycle, beginning with the identification of customer needs and drivers of IT requirements, through to the design and implementation of the service into operation and finally, on to the monitoring and improvement phase of the service.

3) *ITIL Service Strategy*

As the center and origin point of the ITIL Service Lifecycle, the ITIL Service Strategy volume[5] provides guidance on clarification and prioritization of service-provider investments in services. More generally, Service Strategy focuses on helping IT organizations improve and develop over the long term. In both cases, Service Strategy relies largely upon a market-driven approach. Key topics covered include service value definition, business-case development, service assets, market analysis, and service provider types.
List of covered processes:

a. Financial management
b. Demand Management
c. Service Portfolio Management(Available Version 3.0 only)

4) *ITIL Service Design*

The ITIL Service Design volume[6] provides good-practice guidance on the design of IT services, processes, and other aspects of the service management effort. Significantly, design within ITIL is understood to encompass all elements relevant to technology service delivery, rather than focusing solely on design of the technology itself. As such, Service Design addresses how a planned service solution interacts with the larger business and technical environments, service management systems required to support the service, processes which interact with the service, technology, and architecture required to support the service, and the supply chain required to support the planned service. Within ITIL v2, design work for an IT service is aggregated into a single Service Design Package (SDP). Service Design Packages, along with other information about services, are managed within the service catalogs.
List of covered processes:

a. Service Catalogue Management
b. Service Level Management
c. Risk Management
d. Capacity Management
e. Availability Management
f. IT Service Continuity Management
g. Information Security Management
h. Compliance Management
i. IT Architecture Management
j. Supplier Management

5) *ITIL Service Transition*

Service transition, as described by the ITIL Service Transition volume,[7] relates to the delivery of services required by a business into live/operational use, and often encompasses the "project" side of IT rather than "BAU" (Business as usual). This area also covers topics such as managing changes to the "BAU" environment.
List of processes:

a. Service Asset and Configuration Management
b. Service Validation and Testing
c. Evaluation
d. Release Management
e. Change Management
f. Knowledge Management ITIL Service Operation

6) *Service Operation*

Best practice for achieving the delivery of agreed levels of services both to end-users and the customers Service operation, as described in the ITIL Service Operation volume,[8] is the part of the lifecycle where the services and value is actually directly delivered. Also the monitoring of problems and balance between service reliability and cost etc. are considered. The functions include technical management, application management, operations management and Service Desk as well as, responsibilities for staff engaging in Service Operation.
List of processes:

a. Event Management
b. Incident Management
c. Problem Management
d. Request Fulfillment
e. Access Management

7) *ITIL Continual Service Improvement*

The continual service improvements have seven methods to improve the service such as

a. Define what you should measure?
b. Define what you can measure?
c. Gather the data
d. Process the data
e. Analyze the data
f. Presenting and using the data
g. Implement the corrective action

II. RELEASE AND DEPLOYMENT MANAGEMENT

1) *Terminology:*

Release: A release is a collection of authorized changes to an IT service. i.e., A collection of hardware, software, documentation, processes or other components required to implement one or more approved changes to IT services. The contents of each release are managed, tested and deployed as a single entity.
Release Unit: A ‚release unit' describes the portion of a service or IT infrastructure that is normally released together according to the organization's release policy.
Release Package: A release package may be a single release unit or a structured set of release units, including the associated user or support documentation that is required. Like the definition of release units, factors such as the

modularity of components, the amount of change occurring and resources required will be considered when formulating a complete Release Package[13]

Change Management: Change management processes are used to deliver a finalized and tested change into a pre-production environment along with a set of tools and/or procedures for migrating the change into the live production environment [9].

Build Management: The software, hardware and documentation that comprise a release unit should be assembled in a controlled manner to ensure a repeatable process. This should include automation where possible for its compilation and distribution, which for large organizations can significantly reduce the Total Cost of Ownership for the services involved.

Release Management: Release Management is an important key technology for distributing the project/product to the customer.[12]

Deployment: The activity responsible for movement of approved releases of hardware, software, documentation, process etc. to test and production environments.

2) *Release and Deployment Management*

The process responsible for planning, scheduling and controlling the movement of releases to test, pre-production and production environments. The primary objective is to ensure the integrity of the production environment.In conjunction with the use of Change Management, Release and Deployment will enhance an organization's capabilities to develop, compile, reuse, distribute and rollback releases in accordance with defined policies that improve efficiency and reduce business disruption.

3) *Goal of Release and Deployment:*

To deploy new releases into production, transition support to service operation, and enable its effective use in order to deliver maximum value to the clients.

4) *Objectives of Release and Deployment:*

a. To define and agree upon Release policies, and Release and Deployment plans with customers and stakeholders.
b. To ensure the integrity of constructed release packages and that they are recorded accurately in the Configuration Management System.
c. To ensure that all release packages can be tracked, installed, verified, uninstalled or backed out if necessary.
d. To ensure the required skills and knowledge is transferred to support team, customers, end users, suppliers and any other relevant stakeholders.
e. There is minimal unpredicted impact on the production services, customers and service operations.

5) *Scope of Release and Deployment:*

Release and Deployment works closely in conjunction with the other Release, Control and Validation (RCV) processes to enable the quality transition of services. The role played specifically by Release and Deployment is to build, package, validate and distribute authorized service changes to the target systems.

6) *Benefits of Release and Deployment:*

a. Delivering change, faster, at optimum cost and minimized risk
b. Assuring customers and users can use the new or changed service in a way that supports the business goals
c. Improving consistency in implementation approach across the business change, service teams, suppliers and customers
d. Contributing to meeting auditable requirements for traceability through Service Transition.

7) *Triggers and Interfaces*

The primary interfaces of release and deployment exist with Change Management and the surrounding Service Transition processes. Other inputs will also be provided from Service Strategy and Service Design to ensure that the requirements for value provision have been met.

The inputs to release and deployment include:

a. Authorized Request for Changes(RFCs)
b. Service Packages
c. Service Design Package
d. Service Acceptance Criteria
e. Service Management policies and standards
f. Build Models and plans
g. Exit and entry criteria for each stage of release and deployment

The outputs include:

a. Release and deployment plan
b. Updated RFCs for any required activities
c. Updated service catalogue reflecting any service changes
d. New or modified service
e. New or modified processes
f. Skilled and knowledgeable support staff
g. End users with capabilities to use the service
h. SLAs (Service Level Agreements), OLAs (Operational Level Agreements) , UCs(Underpinning Contract)
i. Deployment plans and packages
j. Service Transition Report

8) *Release Design Options*

When planning individual releases or defining the policies that should exist, consideration about the potential impact and need for resources will affect how releases will be deployed to the target locations. The common options for

deploying releases are described below:

a. Big Bang or Phased Approach
b. The Push or Pull Approach
c. Automated or Manual

Big Bang: Where the new or changed service is deployed to all user areas in one operation. This will often be used when introducing an application change and consistency of service across the organization is considered important.

Phased Approach: The service is deployed to a part of the user base initially, and then this operation is repeated for subsequent parts of the user base via a scheduled rollout plan. This will be the case in many scenarios such as in retail organizations for new services being introduced into the stores' environment in manageable phases.

The Push Approach: service component is deployed from the centre and pushed out to the target locations. In terms of service deployment, delivering updated service components to all users, either in big bang or phased forms is the use of a push approach, since the new or changed service is delivered into the users' environment at a time not of their choosing.

The Pull Approach: used for software releases. Software is made available in a central location but users are free to pull the software down to their own location at a time of their choosing or when a workstation restarts. The use of „Pull' updating a release over the internet has made this concept significantly more pervasive. A good example is virus signature updates, which are typically pulled down to update P.C's and servers when it best suits the customer; however at times of extreme virus risk this may be overridden by a release that is pushed to all know users.Pull approaches do not rest so heavily on accurate configuration data and they can trigger an update to user records. This may be through new users appearing and requesting downloads or expected users not doing so, triggering investigation into their continued existence.

Automation: Helps to ensure repeatability and consistency. The time required to provide a well-designed and efficient automated mechanism may not always be available or viable. Typical examples of activities that are capable of a degree of automation are:

a. Discovery tools aid release Planning Automate builds reduce time taken this in turn can resolve scheduling conflicts and delays
b. Automated configuration baselines procedures save time and reduce errors in identifying the status of CI's and releases during build, test and deployment etc

Manual: Important to monitor and measure the impact of many repeated manual activities as they are likely to be inefficient and error prone. This will ultimately slow down the release team and create resource and capacity issues that affect the agreed service levels

9) *Release Policy*

A Release Policy is the formal documentation of the overarching strategy for releases and was derived from the Service Design phase of the Service Lifecycle. It is the governing policy document for the process and must accommodate the majority of releases being implemented. Typical contents of a Release Policy include:

a. Level of infrastructure to be controlled by Releases
b. Preferred structure and schedules for Release Packages
c. Definition of major and minor releases, emergency fixes
d. Expected deliverables for each type of Release
e. Policy on the production and execution of back out plans
f. How and where Releases should be documented
g. Blackout windows for releases based on business or IT requirements
h. Roles and responsibilities defined for the Release and Deployment process, Supplier contacts and escalation points

10) *Release and Deployment Activities*

The Release Policy is the overarching strategy for Releases and was derived from the Service Design phase of the Service Lifecycle. The Release Plan is the operational implementation for each release. The Deployment Plan is the documented approach for distributing a single Release.

Steps for Release and Deployment Activities:

A. Release planning
B. Preparation for build, test and deployment
C. Build and test
D. Service test and pilots
E. Plan and prepare for deployment
F. Perform transfer. Deployment and retirement
G. Verify deployment
H. Early Life Support
I. Review and close the deployment
J. Review and close Service Transition

A. *Release Planning*

Any plans created for the release and deployment will need to be integrated with the overall Service transition plan, and conform to any policies that have been defined. For each release, plans should be authorized by Change Management and used to assist in the evaluation of the risk, impact and resource requirements for components of the change. Typically the release and deployment plans should document the:

a. Scope and content of the release
b. The risk assessment for the release
c. Affected stakeholders
d. Teams responsible for the release
e. Communication strategy to be used during the release and deployment process

Plans should take into account the acceptance criteria that exist for the release and when authorization points will

verify a pass or fail. The processes of Evaluation and Service Validation and Testing will be integrated here to assist in the determination whether to continue, pause or revert to previous stages of the release.

Build and test planning: The approach taken for building, testing and maintaining the controlled environments to production will need to be planned in order to enable optimum use of resources for the development of the release. The activities that occur here are:

a. Developing build plans based on the Service Design Package and defining any environment requirements.
b. Scheduling the resources and time required to setup the environments
c. Testing the build and compilation procedures
d. Scheduling the build and compilation activities
e. Assigning resources, roles and responsibilities for any key activities

Environments that may be utilized during this period include:

a. Build environments
b. Testing and integration environments
c. Deployment environments
d. Pilot environments

Utilizing Pilots: Pilots may be useful for testing the service with a group of participants that are representative of the broader end-user community. For this to be effective the scope needs to be carefully determined, as being either too large or too small will likely result in some negatively impact to the overall success and quality of the release and deployment process. Pilots should include mechanisms by which feedback can be gathered about various aspects of the release and associated processes. For complex releases or in large and diverse organizations it may be appropriate to use more than one pilot for address the different implementation and support issues that exist.

Deployment Planning: There are many factors that will be considered when choosing the most appropriate deployment strategy. Questions that must be answered include:

a. What needs to be deployed?
b. Who are the users?
c. Are there location dependences?
d. Where are the users?
e. Who else needs to be prepared well in advance?
f. When does the deployment need to be completed?
g. Why is the deployment happening?
h. What are the critical success factors and exit criteria?
i. What is the current capability of the service provider?

Financial/Commercial Planning: Where necessary, various financial and commercial factors will need to be assessed before the deployment activities, including:

a. Working capital
b. Contracts and licenses
c. Funding
d. Intellectual property requirements

B. *Preparation for build, test and deployment*

Before the actual building of the release occurs, the release design must be validated against the requirements defined for the new or changed service offering. This should be an independent evaluation that checks the release will deliver the predicted outcomes, and any issues documented in an interim evaluation report. Training of involved release and deployment staff. In many cases the introduction of a release may require additional training for the release, deployment, build and test teams. Such training could be related to the:

a. Service management processes to be used
b. Changes in security or health and safety procedures
c. Understanding of the Service Design documentation and plans
d. Technology being utilized for the release.

C. *Build and Test*

Wherever possible, repeatable practices and reusable components should be utilized to during the build and test of releases. This includes managing the:

1) Build, test and packaging environments
2) Compilation and packaging tools
3) Configuration of the releases themselves:
   a. Version control
   b. Documentation templates for testing and validation
   c. Access rights and security procedures

Release and build documentation: Documentation templates, procedures, knowledge bases and other guidance should be consistently available to support the release team in the activities performed. Typical documentation that will be used by the release teams include:

1) Contract agreements
2) Purchase requests
3) Health and Safety guidelines
4) Security policies
5) Licence agreements
6) Procedures for:
   a. Distributing software
   b. Delivering, moving and installing equipment
   c. Wiping sensitive data and media

d. Publishing, sharing and archiving knowledge, information and data.

Acquire and test required components: Release and Deployment should be interfaced with the organization's existing procurement processes to acquire any required components for the release. This will save time and effort in verifying assets, capturing and recording information, ensuring proof of licence and triggering updates to the Asset Management System. As part of the overall Service Validation and Testing, each of the individual components should be tested to verify that any quality criteria has been met, initiating action where quality criteria is not met.

Release Packaging: Build management procedures, tools and checklists should be utilized as part of the release packaging, to provide repeatable practices and expected outcomes. When a definitive package has been assembled, a baseline should be taken of the release package and the correct versioning and naming conventions applied.

Managing the build and test environments: The need for multiple environments in which to build and test will depend on the size, complexity, frequency and impact of the releases being managed. Test environments should be protected using a range of testing best practices, and appropriate access to these environments given based on the priorities defined. Automating the installation of systems and software reduces the workload of people, but also requires testing of the scripts and mechanisms that will be used.

D. *Service testing and pilots*

As part of a coordinated effort with Service Validation and Testing, testing and validation must be performed at multiple levels. With particular focus on the release itself, service rehearsals may be used, which simulates as much of the service as possible in an extensive and widely involved practice session. This would normally occur after other pilots have run, and is designed to be the last measure to detect any potential issues that will arrive during or after the deployment to the live environment.

Pilots: Previous planning should have already identified what pilots will be used as part of the overall release and deployment. Key actions to take during pilots are:

a. Training of any people involved
b. Documentation of any required procedures
c. Continual communication and engagement with customers and users
d. Determine the levels of support required for the actual release
e. Discover and fix issues wherever possible before the final deployment
f. Document improvements where appropriate and incorporate them into future plans

E. *Plan and prepare for deployment*

At this stage the focus is to prepare the organization and people for organizational change and to refine and deployment plans that have been documented. These plans should include guidance regarding:

a. Risk mitigation plans
b. Disposal and retirement plans
c. The logistics for delivery
d. Knowledge transfer
e. Mobilizing users to be ready to use the service
f. Mobilizing the support staff for service readiness

F. *Perform transfer, deployment and retirement*

During the actual implementation itself, the activities performed can be grouped under the following tasks:

a. Transfer financial assets
b. Transfer changes required to business/organization
c. Deploy processes and materials
d. Deploy Service Management Capability
e. Transfer service
f. Deploy service
g. Decommissioning and service retirement
h. Remove redundant assets

These activities will need to be modified to accommodate any items specified in the deployment plan as part of the acceptance criteria for go live'.

G. *Verify deployment*

Once the activities for the deployment of releases are finished, verification should occur that users are capable of operating the service. Verification should ensure that:

a. The service/release components are in place by means of a configuration audit
b. Documentation has been updated accordingly
c. Roles and responsibilities have been correctly assigned
d. The measurement and reporting systems are established to measure performance of the service

Any noted deviations from plans or other items raised should be documented in order to improve future implementations and processes used.

H. *Early Life Support*

The quality of transition to Service Operation is a crucial element to the success of the overall service change that is being implemented. Rather than simply hand off support post-deployment, the release and deployment teams should assist in managing any calls, incidents and problems that are detected in the early of the new or modified service. This

enables more stability in this vulnerable period, increased customer and user satisfaction, enhanced learning and better momentum for continual improvement. The resource allocation from these teams will then be gradually reduced while Service Operation takes on more responsibility for support.

The example shown in the figures above demonstrate how the number of incidents for deployment A was significantly reduced through the use of Early Life Support, including the training of users and staff, and the transfer of knowledge to service desk staff.

a. The acceptance criteria for finalizing Early Life Support should be agreed early in the process, which might include such conditions as: Users can use the service effectively and efficiently for their business activities
b. Service delivery is managed and controlled across any service provider interfaces
c. Service levels and performance standards are being consistently achieved
d. Stakeholder groups verify the transfer of required knowledge
e. All deliverables required for the release have been signed off.

I. Review and close the deployment

A formal review of deployments should be performed for all releases of a determined level of size, impact, cost or risk to the organization. The review seeks to ensure that all requirements for the release have been met and to identify and potential improvements that can be made. Items that should be reviewed include:

a. Any quality criteria deviations that exist. Any open actions or necessary fixes that have been identified
b. Review open changes
c. Review performance targets and achievements
d. Experiences and feedback from customers, users and staff involved with the deployment. All problems and known errors are documented and accepted by the business and/or suppliers
e. Check that any redundant assets have been removed

J. Review and close Service Transition

The final step required in finalizing the completion of Service Transition is a formal review appropriate to the relative scale of the change. This will utilize the process of Evaluation and is driven by Change Management, which will verify successful completion and that the handover to Service Operation is complete. The “lessons learnt” should be documented to provide any improvement opportunities that can be made and developed by Continual Service Improvement for future transitions.

III. Conclusion

This paper helps to understand the basics of ITIL framework and release methodology. Good implementation of ITIL in the business has many benefits such as a higher efficiency, a better overview and control of the business’ needs and many more. The study gives the high level understanding of how to release a change into the IT environment using ITIL efficiently for better benefits for the organizations.

IV. References

1) Rob Addy, Effective IT Service Management, Springer, ISBN 978-3-540-73197-9, 2007
2) L.Bass, M.Klein, and F.Bachmann, Quality Attribute Design Primitives and the attribute Driven design method, Proceeding of Fourth on Product Family Engineering, Bilbao, Spain, October 4, 2001
3) Office of Government Commerce (UK)CCTA and OGC. Retrieved May 5, 2005
4) www.ogc.uk
5) Majid Iqbal and Michael Nieves (2007). ITIL Service Strategy. The Stationery Office. ISBN 9780113310456
6) Vernon Lloyd and Colin Rudd, ITIL Service Design, The Stationery Office, ISBN 9780113310470, 2007
7) Shirley Lacy and Ivor Macfarlane . ITIL Service Transition. The Stationery Office, ISBN 9780113310487, 2007
8) David Cannon and David Wheeldon. ITIL Service Operation. The Stationery Office, ISBN 9780113310463, 2007.
9) Larry Klosterboer, Implementing ITIL Configuration Management, IBM Press, 2007
10) Hoek, A.Van Der, Hall, R. S., Heimbigner D., Wolf, A. L. ,Software release management, Proceedings of the 6th European conference held jointly with the 5th ACM SIGSOFT international symposium on Foundations of software engineering, p.159-175, September 22-25, Zurich, Switzerland, 1997.
11) Erenkrantz, J. R.Release Management within Open Source Projects. In: proceedings of the 3rd open source software development workshop. Portland, oregon, USA, may 2003, pp 51–55.
12) Annie Ibrahim Rana, Muhammad Waseem Arfi, "SOFTWARE RELEASE METHODOLOGY:A CASE STUDY", Student Conference on Engineering Sciences and Technology, 2005.
13) Tim Malone, Ivanka Menken and Gerard Blokdijk, "Release, Control and Validation",The Art of Service Pvt. Ltd.,2009.

A Dynamic Terrain-Spaced Maze Generation Algorithm

Metin Turan[1], Kenan Aydın[2]

GJCST Classification
I.2.1, I.2.2

***Abstract*-Maze algorithms are generally developed to create mazes in a game board, which consist of single cell of passages and nearly all cells are accessible. However, it would be useful if some group of cells randomly separated for terrain design and some passages was randomly irregular in width or contains rooms for arrangement of game objects.In this study a randomly irregular and terrain-spaced maze generation algorithm has been developed. The randomly produced rooms within the generated passages can be used for planning game strategy. On the other hand the cells which are not used for passages may be used for terrain design. This algorithm only needs boundary check to prevent getting out of the game board. Moreover maze complexity can be identified by a ratio which is defined as the ratio of passage cells number over total number of cells on the game board.**

I. Introduction

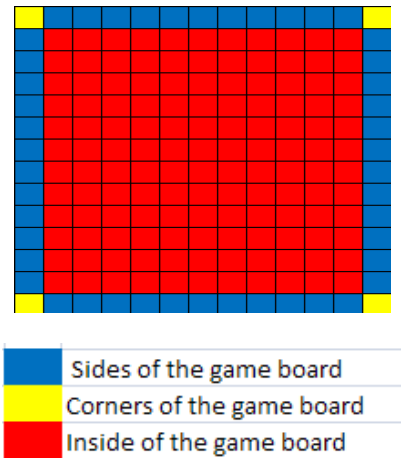
There are various kinds of maze algorithms in the literature; most of them use tree structure to generate a maze and others use sets to generate mazes. Maze generation algorithms generally focus on generating perfect mazes. A maze is perfect if every cell in the maze is accessible and there is exactly one path to another cell. Well known perfect algorithms using tree structure is Prim algorithm [1] where using sets is Kruksal algorithm [2]. Both use same approach to create a random maze which constructs cell walls using a random chosen side of the cell to create a wall.Maze algorithms (perfect, braid, unicursal, sparse, partial braid) [3,4,5,6,7] generate passages in a given two or n- dimensional game board which is composed of a number of cells. Passage width is generally regular for all passages through the entire game board. These algorithms aim is only to create a solvable maze. They haven't got ability to generate rooms within the passages. On the other hand, sparse algorithms [8] create passages where some cells are left uncreated, resulting in an irregular maze with wide passages and rooms. However the maze shapes not eligible to design rooms for game characters and to produce amusing maze games. Moreover it is still impossible saving empty spaces out of passages which would be used for designing terrains.However, it should be noted that although there are lots of tools available to make mazes more challenging when generating them on a computer, often the most interesting mazes are those designed and created by hand – "it is hard to have a random algorithm which will also generate a psychologically interesting maze"[9]. In other words, the passages of lots of two dimensional popular games (maze, labyrinth and racing game categories) have been designed statically. However, the proposed algorithm generates dynamic (random) passages for each new play which is more interesting for players.The proposed algorithm assumes game board is a wall initially. The aim is to construct randomly passages through the wall. This is a reverse approach if it is compared with wall adding techniques used in maze algorithms. The complexity of maze can be controlled by a parameter (filled ratio) given to proposed algorithm. It also creates loops and dead-ends in the maze. Loops are connected to at least one passage. On the other hand limited dead-ends help to create traps in the artificial games.Proposed algorithm starts with a random point on the board. Then it finds a random direction and a random passage length (number of cells) to construct passages in the maze. This step is repeated recursively until it reaches the given maze complexity target which is defined as the ratio of passage cells number over total number of cells on the game board.Unlike other known algorithms, proposed algorithm does not require extreme memory except an n-dimensional matrix to simulate the game board. A pixel or a predetermined area size (e.g. 20x20 pixels) could be represented with a cell in the matrix of size maxCOLUMN x maxROW, where for two dimensional matrix n shows x coordinate and m shows the y coordinate.The rest of the article is designed so to describe the work in Section II, statistical findings in Section III, algorithm analysis in section IV, game example(using proposed algorithm) in Section V, conclusion in Section VI and further works in section VII.

II. Work

The proposed algorithm (***generateMaze()***) depends on random values as the previous algorithms did. Due to the randomization generates same values over and over again a special randomize function is needed to construct reliable mazes. This general randomize function is constructed by adding CPU time on the current time which gives highly better results.First of all, ***generateMaze()*** finds a random starting point at one side of the board. It may be on horizontal (x is between 1 and maxCOLUMN) or vertical(y is 1 between maxROW) boundary which is determined by a parameter(boardside).Then generateMaze***()*** creates the passages. Before a new passage is created, it checks if the maze complexity target is obtained. The maze complexity is provided by a parameter (complexity). New passage direction is produced by general randomize function. The board is divided into three regions

About[1] - Kültür University, Istanbul, Turkey m.turan@iku.edu.tr
About[2]- Kültür University, Istanbul, Turkey heret1que@hotmail.com

where each cell has different possible directions to route. First region is the corner cells, where routing is lowest and there are two possible directions to move since they are on the corners of the game board. Second region is the side cells excluding the corner cells, where routing has three possible directions. Third region is the inner section of the board, where routing has four directions. This is outlined at Picture I.



Picture I: The regions of a game board

By the time the direction is selected, ***generateMaze()*** uses a random length (number of cells) to go forward in the selected direction. It checks the case of exceeding boundaries. A generation method which doesn't use a random length after direction decider will not produce a complex maze. However it will generate different type of terrain as in Picture II.

Picture II: A maze generated by proposed algorithm without "how many cells it will go" function.

generateMaze() called recursively until it reaches maze complexity.

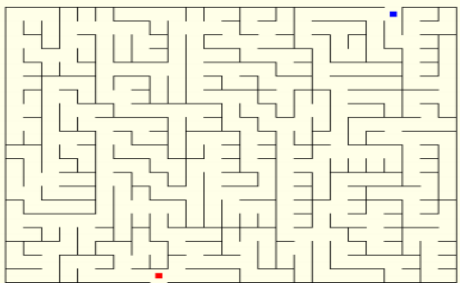
Checking the complexity procedure calculates the ratio of passage cells number over total number of cells on the game board

Unlike other known maze-generation algorithms, the ***generateMaze()*** doesn't create passages cell by cell, instead use number of cells. This speeds up the construction time of maze in the case of big boards. The ***generateMaze()*** pseudo-code is given below.

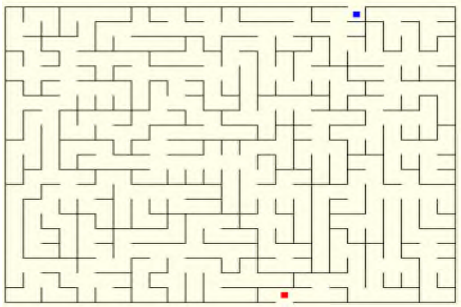
```
void generateMaze( boardside, complexity)
{
col,row=call startPoint(boardside);
call generatePassage(col,row,complexity);
}
col,row startPoint(boardside)
{
return random row and column value using boardside;
/* on the vertical or on the horizontal side of the board /*
}
void generatePassage(currentRow,currentColumn,
complexity)
{
direction=decideDirection(randomize());

passageLength=randomize()
if (checkBoundary())
currentRow,currentColumn=createPassage(passageLength);
else
currentRow,currentColumn=createPassage(Length to the
boundary);
if (checkComplexity(complexity))
generatePassage(currentRow,currentColumn,complexity);
}
```

Example outputs of Kruskal, Prim and ***generateMaze***() is given on Picture III, Picture IV and Picture V respectively.



Picture III: A maze generated by Kruskal Algorithm

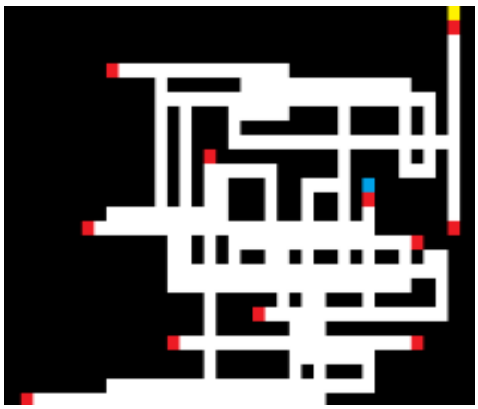


Picture IV: A maze generated by Prim Algorithm

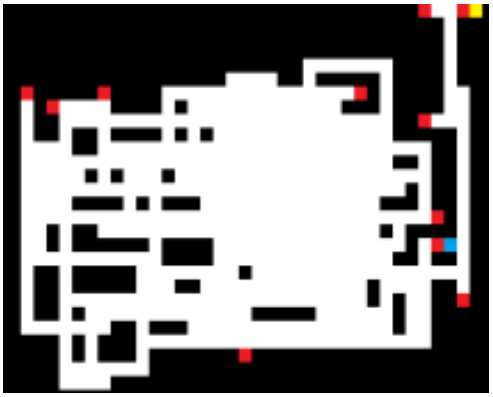
Picture V: A maze generated by Proposed Algorithm

Locating and storing the coordinates of the dead ends is also possible by findDeadEnds() algorithm.
After the generation of the maze is completed, findDeadEnds() will be called and simply looking for the cells which only have one neighbor cell to access these cells.

```
void findDeadEnds()
{
for each counting numbers between 0 and maxROW , called ROW
for each counting numbers between 0 and maxCOLUMN , called COL
If(NeighborCellsCount(ROW,COL)= =1)
DeadEnds[ROW][COL]=true;
}
```



Picture VI: A Maze generated by Proposed Algorithm by using 30% Filled Ratio

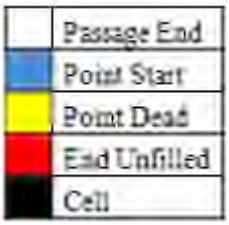


Picture VII: A Maze generated by Proposed Algorithm by using 45% Filled Ratio

Picture VIII: A Maze generated by Proposed Algorithm by using 60% Filled Ratio

Picture IX: A Maze generated by Proposed Algorithm by using75% Filled Ratio

Table I
Visual information of simulated mazes.



Above example maze outputs is given in pictures Picture VI, Picture VII, Picture VIII and Picture IV per different filled ratio values (0.3, 0.45, 0.6, 0.75). The meaning of the colors in the mazes is given in Table I.

III. STATISTICAL FINDINGS

Statistical work has been done for step, dead end, turn and collision properties of ***generateMaze() algorithm*** by running four different filled ratio values, each time it has been executed 200 times (Table II).

Table II
Average values of Steps, Dead Ends, Turns and Collisions per filled ratio

Filled Ratio	Step	Dead Ends	Turn	Collision
30%	1420	24	469	2546
45%	2851	23	935	6520
60%	5876	20	1914	16661
75%	16784	16	5474	57130

Step means that the number of passage construction movement (for each movement a number of cells are processed- reverse movement on the existing passages is also considered) produced by ***generateMaze().***

- Dead ends are the cells which only have one neighbor cell to access these cells.
- Turn is defined as changing previous direction.
- Filled ratio is the ratio of passage cells number over total number of cells on the game board.
- Collision is the number of how many times passages re-filled by reversing or overlapping.

In Figure I, it can be viewed that the average step values gets higher as the filled ratio gets higher. As a result average turn values gets higher. Nearly in every 3 steps ***generateMaze()*** makes a turn.

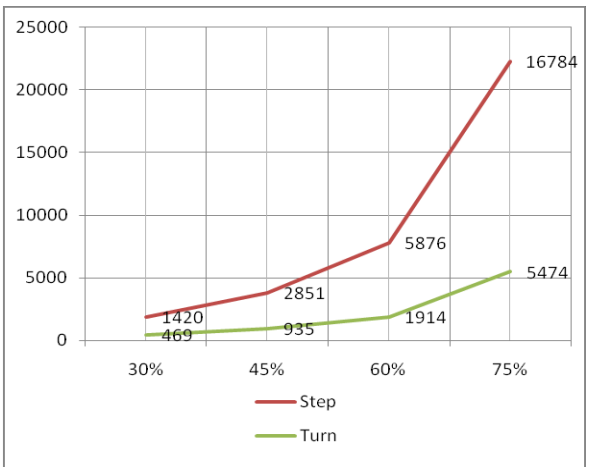


Figure I: Average Step and Turn Values per Filled Ratio

In Figure II, it can be seen that the dead ends are gets lower as the filled ratio gets higher. This means that there is a negative correlation between these two properties. It is an expected result of the ***generateMaze().*** It makes various kinds of turns which eliminate existing dead ends by connecting them with other new passages that has been intersected by the existing dead ends, since a dead end occurs at cells which connected with only one neighbor cell.

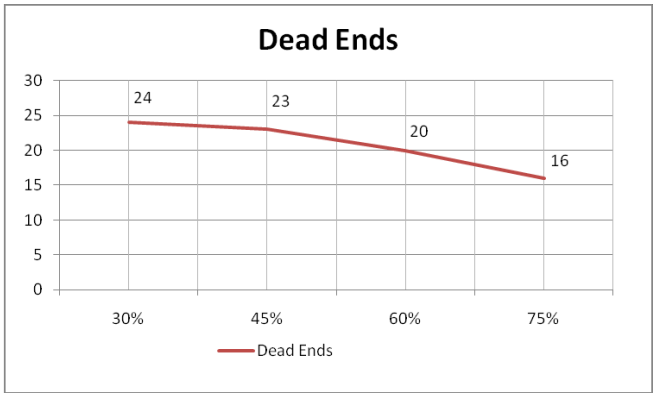


Figure II: Number of Dead Ends versus Filled Ratio

Table III
Correlation values between Turn – Collision and Dead End

Correlation Values	30%	45%	60%	75%
Turn – Collision	~1	~1	~1	~1
Dead End – Turn	~0	~0	~0	~0
Dead End - Collision	~0	~0	~0	~0

On the other hand, there is a positive correlation between turn and collision. This is an expected result due to the ***generateMaze()*** runs to create passages to reach to the filled ratio(when steps increases, turns increases also Figure I), while makes a lot of turns in random directions which causes more collisions. However, Dead End is independent property from turn and collision. It can be concluded from Table III.

IV. ALGORITHM ANALYSIS

Both Prim and Kruksal algorithm create vertices and edges to construct their mazes. Their time complexity nearly same and can be defined as O(E log V), where E and V implies edges and vertices respectively[10]. However, generateMaze() doesn't create edges or vertices. It randomly produces passages. Passages are constructed on a board which is represented by a matrix (maxCOLUMN is ***x***, maxROW is ***y***). Moreover the ***filled ratio*** (sparse area percentage) and random ***length*** (number of cells to fill for each step) are also determines the time complexity. Using these parameters we calculate the following game board construction complexities for both the worst and the best conditions under assumption of no collision [11].For the worst condition, the random length is 1. This time ***generatePassage()*** is called **x*y** times.
filled ratio* (max(**x**,**y**)²)*(4+**length**) where 4 shows the number of steps executed in auxiliary functions and inside the ***generatePassage().*** If the matrix NxN size then the worst game board construction complexity is;
O(5*filled ratio*N²) For the best condition, the random length is max(**x**,**y**). This time ***generatePassage()*** is called only min(**x**,**y**) times. **filled ratio*** (min(**x**,**y**)*(4+max(x,y)))
And if the matrix NxN size then the shortest game board construction complexity is; O (filled ratio*(N²+4N)) This complexity values are acceptable and applicable if they compared with the Kruskal's (O(E log V)) and Prim's (O(E + V log V)) algorithms, where E and V shows edges and vertices respectively.

V. AN EXAMPLE GAME

A simple 2-Dimensional Windows based game developed by using the ***generateMaze()*** (Picture X)[12]. The goal of the game is to collect all the stars and proceed through levels (1 to 10) to complete the game. Player has to collect more stars for each next level. By the time, player shouldn't collect skulls. It decreases lives of the player (4). If player collect first-aid bags then it increases player lives. Player character is chased by an AI controlled character while collecting the stars. If AI character catches the player the live is decreased by one. ***generateMaze()*** allows creating dead end passages which make this game more amusing. Also an algorithm is developed so dispersing collectable items on dead-end passages to challenge the game.

Picture X: An example game created by proposed algorithm.

VI. CONCLUSION

The ***generateMaze()*** is a flexible(parametric) algorithm and if it is compared with the sparse algorithms in literature, it is truly random (not stable) as a result of the used technique. By the way, statistics from the hundreds of experiments have shown that created mazes are not related to each other and well-shaped. Moreover, computational complexity is less than any sparse algorithm. The worst case is to check only next matrix location O(1). On the other hand, memory requirements depends on the game board matrix sized N column and M rows. The worst case is not more than O(M x N). It decreases by the dimension of the passage width.The only problem using ***generateMaze()*** is the stack overflow error above the 0.75 filled ratio. When the filled ratio gets higher, collision increases and recursive call of ***generatePassage()*** reaches a value which causes stack overflow error. When filled ratio is greater than the 0.75 then generated maze is not well (Picture IX). So this is not a big problem for algorithm acceptance.

VII. FURTHER WORKS

Algorithm optimization is needed to control dead end generation and reduce the collision values which will speed up the process of maze generation.This algorithm may also be generalized to use in 3D terrain-spaced maze generation easily.

VIII. REFERENCES

1) [R. C. Prim: Shortest connection networks and some generalizations. In: Bell System Technical Journal, 36 (1957), pp. 1389–1401
2) Joseph. B. Kruskal, "On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem", Proceedings of the American Mathematical Society, vol. 7, no. 1 (Feb, 1956), pp. 48–50
3) Fisher A., "The Amazing Book of Mazes", Harry R. Abrams Inc.,2006

4) http://www.conceptispuzzles.com
5) Kern H., “Through the Labyrinth: designs and meanings over 5000 years”, Prestel , 2000
6) http://www.astrolog.org/labyrnth
7) T.H. Cormen, C. E. Leiserson, R.L. Rivest, and C. Stein, “Introduction to Algorithms, MIT Press and McGraw-Hill, 2001
8) Hans Pedersen and Karan Singh, “Organic labyrinths and mazes”, Proceedings of the 4th international symposium on non-photorealistic animation and rendering NPAR ’06, ACM Press, June 2006
9) Danny Kodicek, “ Mathematics and Physics for Programmers”, Charles River Media , 2005
10) Anany Levitin, “Introduction to The Design & Analysis of Algorithms”, Addison Wesley, International Edition, 2003.
11) Oliver Kullman, CS-232 algorithms and complexity notes, October 2005.
12) K. Aydın, M. Turan, “A Simple Windows Based Game”, Student Final Project, İstanbul Kültür University, Dept. Of Comp. Science, January 2010.

Truth of D-DoS Attacks in MANET

Gaurav Kumar Gupt[1], Mr.Jitendra Singh[2]

GJCST Classification
I.2.3, C.2.0

***Abstract*-Network security is a weak link in wired and wireless network systems. Malicious attacks have caused tremendous loss by impairing the functionalities of the computer networks. Denial of Service (DoS) and Distributed DoS (DDoS) attacks are two of the most harmful threats to the network functionality. Mobile Ad Hoc Networks (MANET) are even more vulnerable to such attacks.. Denial of Service (DoS) is the degradation or prevention of legitimate use of network resources. The wireless ad hoc network is particularly vulnerable to DoS attacks due to its features of open medium, dynamic changing topology,cooperative algorithms, decentralization of the protocols, and lack of a clear line of defense is a growing problem in networks today. Many of the defense techniques developed on a fixed wired network are not applicable to this new mobile environment. How to thwart the DoS attacks differently and effectively and keep the vital security-sensitive ad hoc networks available for its intended use is essential.**

Background -This research work concentrates on developing defense mechanism against certain types of DoS attacks in the Ad Hoc network environment

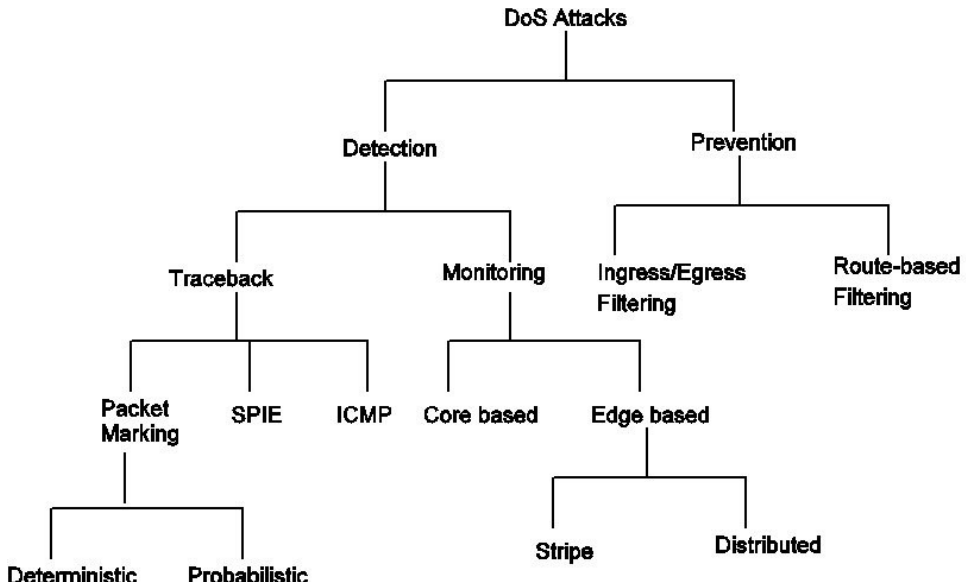
HOW THEY WORK

To give a very simple example, let us assume that there is already a small ad-hoc network in place. When a new node in this example it can be the PDA of Tom joins the ad-hoc network, there are a number of things to do: The device needs to set up contact to other nodes in range, telling them: I am here. By this, the new node learns who the neighbour nodes are, and vice versa. Another point is that the new node, in this example the PDA, needs a unique identifier to make it addressable an IP address in IP networks. For all this, the new node is on its own, as there is neither a central controlling entity nor a pre-existing fixed infrastructure in adhoc networks. When Tom wants to send a message from his PDA to that of Maria, other nodes serve as a relay station in a process called multi-hop routing, if the PDA of Maria is not in direct reach, using one of the routing protocols designed for ad-hoc networks. This small example shows a few imminent advantages of ad-hoc networks: They can extend the range of the wireless technology in use, e.g. WLAN or Bluetooth, they can reduce the nodes power consumption due to a lower transmission power required, and they increase the nodes mobility. To make this work, though, ad-hoc networks require a critical mass of well-behaving nodes, willing to forward others traffic.

About[1] -Department:Computer science Srm university (Mtech (cse) student) E-mail-gauravkumarblack@yahoo.coml.com
About[2]-Department:Computer science Srm university E-mail:jitendra.jit@gmail.com

I. INTRODUCTION

An ad hoc network is a collection of wireless mobile nodes that forms a temporary network without any centralized administration. In such an environment, it may be necessary for one mobile node to enlist other hosts in forwarding a packet to its destination due to the limited transmission range of wireless network interfaces. Each mobile node operates not only as a host but also as a router forwarding packets for other mobile nodes in the network that may not be within the direct transmission range of each other. Each node participates in an ad hoc routing protocol that allows it to discover multihop paths through the network to any other node. This idea of Mobile ad hoc network is also called infrastructureless networking, since the mobile nodes in the network dynamically establish routing among themselves to form their own network on the fly. So, Mobile Ad Hoc Network (MANET) is a collection of communication devices or nodes that wish to communicate without any fixed infrastructure and pre-determined organization of available links. The nodes in MANET themselves are responsible for dynamically discovering other nodes to communicate.Now-a-days, Mobile ad hoc network (MANET) is one of the recent active fields and has received marvelous attention because of their self-configuration and selfmaintenance capabilities. While early research effort assumed a friendly and cooperative environment and focused on problems such as wireless channel access and multihop routing, security has become a primary concern in order to provide protected communication between nodes in a potentially hostile environment. Recent wireless research indicates that the wireless MANET presents a larger security problem than conventional wired and wireless networks.



SECURITY GOALS:For analyzing the security of wireless mobile adhoc networks, we need certain parameters. The basic parameters for a secure system are:

- Availability
- Confidentiality
- Authentication
- Integrity
- Non-repudiation
- Scalability

II. SECURITY ATTACKS IN MANETS

The security attacks in MANETs can be categorized as active attacks and passive attacks. Active attack is an attack when misbehaving node has to bear some energy costs in order to perform the threat. On the other hand, passive attacks are mainly due to lack of cooperation with the purpose of saving energy selfishly. Nodes that perform active attacks with the aim of damaging other nodes by causing network outage are considered as malicious while nodes that make passive attacks with the aim of saving battery life for their own communications are considered to be selfish. Various types of attacks in MANETs are: Modification, Impersonation, Fabrication, Eavesdropping, Replay, Denial of Service, Malicious Software and Lack of Cooperation. Denial of Service attack is described below. Network Protocol Stack Based Attack Classification

Stack Layer	Attacks
Application	Backdoor, Virus, Data corruption or deletion, Repudiation
Transport	Desynchronization, Session hijacking, SYN flooding
Network	Blackhole, Byzantine, Flooding, Location disclosure, Misdirection, packet dropping, Resource consumption (Sleep deprivation), Rushing, Selfish, Spoofing, Wormhole
Link	Collision, Disruption MAC (802.11), Exhausting, Monitoring (Traffic analysis), Unfairness, WEP weakness
Physical	Eavesdropping, Interceptions, Jamming, Tampering
Multi-layer attacks	DoS, impersonation, replay, man-in-the-middle

Attacks could also be classified according to the target layer in the protocol stack

III. DENIAL OF SERVICE (DOS)

A denial of service (DoS) attack is characterized by an explicit attempt by an attacker to prevent legitimate users of a service from using the desired resources. Examples of denial of service attacks include:

- attempts to "flood" a network, thereby preventing legitimate network traffic
- attempts to disrupt connections between two machines, thereby preventing access to a service
- attempts to prevent a particular individual from accessing a service
- attempts to disrupt service to a specific system or person.

IV. DISTRIBUTED DENIAL OF SERVICE ATTACK(DDOS)

A DDoS (Distributed Denial-Of-Service) attack is a distributed, large-scale attempt by malicious users to flood the victim network with an enormous number of packets. This exhausts the victim network of resources such as bandwidth, computing power, etc. The victim is unable to provide services to its legitimate clients and network performance is greatly deteriorated. The distributed format adds the "many to one" dimension that makes these attacks more difficult to prevent. A distributed denial of service attack is composed of four elements, as shown in Figure 1.First, it involves a victim, i.e., the target host that has been chosen to receive the brunt of the attack. Second, it involves the presence of the attack daemon agents. These are agent programs that actually conduct the attack on the target victim. Attack daemons are usually deployed in host computers. These daemons affect both the target and the host computers.

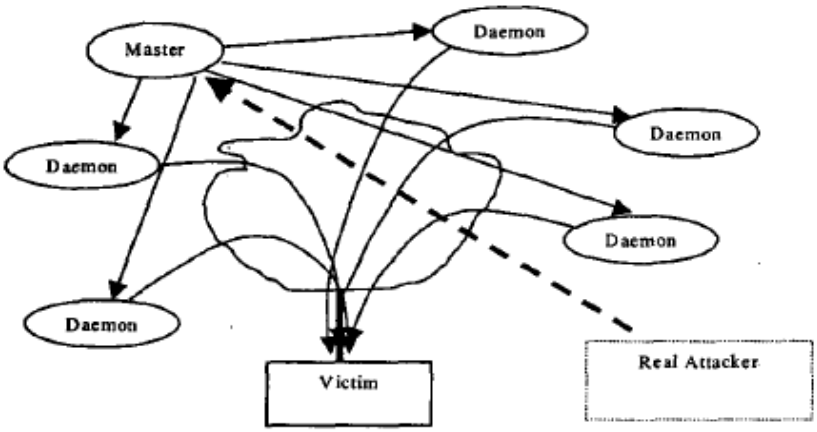


Figure 1: The four components of a distributed denial of service attack: a real attacker, a control master program, attack daemons and the victim

The task of deploying these attack daemons requires the attacker to gain access and infiltrate the host computers. The third component of a distributed denial of service attack is the control master program. Its task is to coordinate the attack. Finally, there is the real attacker, the mastermind behind the attack. By using a control master program, the real attacker can stay behind the scenes of the attack. The following steps take place during a distributed attack:

- The real attacker sends an "execute" message to the control master program.
- The control master program receives the "execute" message and propagates the command to the attack daemons under its control.
- Upon receiving the attack command, the attack daemons begin the attack on the victim.

V. DEFENSE MECHANISMS

We classify defense mechanisms to DDoS attacks into two broad categories: local and global. As the name suggests,local solutions can be implemented on the victim computer or its local network without an outsider's cooperation.Global solutions, by their very nature, require the cooperation of several Internet subnets, which typically cross company boundaries.

VI. LOCAL SOLUTIONS

Protection for individual computers falls into three areas.

1) *Local Filtering*

The timeworn short-term solution is to try to stop the infiltrating IP packets on the local router by installing a filter to detect them. The stumbling block to his solution is that if an attack jams the victim's local network with enough traffic, it also overwhelms the local router, overloading the filtering software and rendering it inoperable.

2) *Changing IPs*

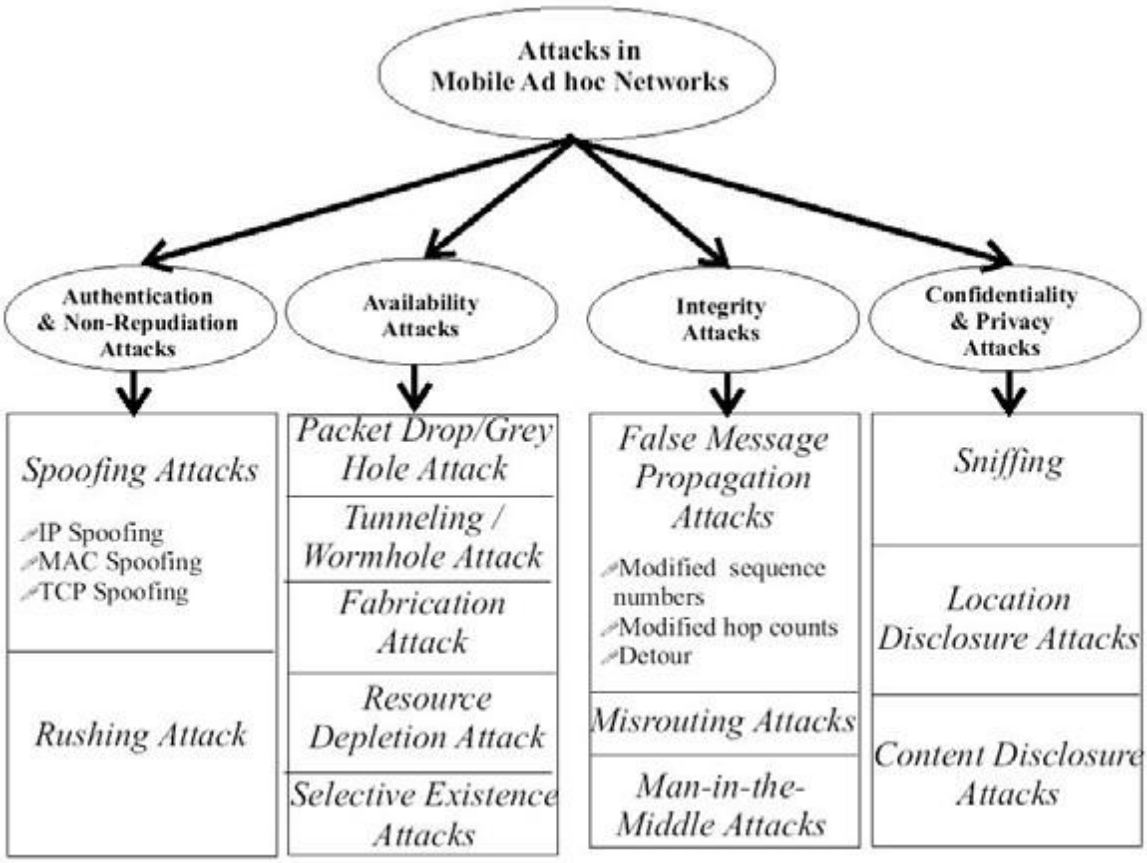
A Band-Aid solution to a DDoS attack is to change the victim computer's IP address, thereby invalidating the old address. This action still leaves the computer vulnerable because the attacker can launch the attack at the new IP address. This option is practical because the current type of DDoS attack is based on IP addresses. System administrators must make a series of changes to domain name service entries, routing table entries, and so on to lead traffic to the new IP address. Once the IP change which takes some time is completed, all Internet routers will have been informed, and edge routers will drop the attacking packets.

3) *Creating Client Bottlenecks*

The objective behind this approach is to create bottleneck processes on the zombie computers, limiting their attacking ability. Examples of this approach include

- RSA Security Corp. Client Puzzles: RSA's Client Puzzles algorithm (see http://www.rsasecurity.com/rsalabs/staff/ajuels/papers/c lientpuzzles.pdf) requires the attacking computer to correctly solve a small puzzle before establishing a connection. Solving the puzzle consumes some computational power, limiting the attacker in the number of connection requests it can make at the same time.
- Turing test: Software implementing this approach requires the attacking computer to answer a random question before establishing the connection. The question should be easy for humans to answer but not computers for example, "Which film won the Oscar for best picture in 2000?"

Attacks in MANET



VII. GLOBAL SOLUTIONS

Clearly, as DDoS attacks target the deficiencies of the Internet as a whole, local solutions to the problem become futile. Global solutions are better from a technological standpoint. The real question is whether there is a global incentive to implement them.

1. Improving the Security of the Entire Internet: Improving the security of all computers linked to the Internet would prevent attackers from finding enough vulnerable computers to break into and plant daemon programs that would turn them into zombies.

2. Using Globally Coordinated Filters: The strategy here is to prevent the accumulation of a critical mass of attacking packets in time. Once filters are installed throughout the Internet, a victim can send information that it has detected an attack, and the filters can stop attacking packets earlier along the attacking path, before they aggregate to lethal proportions. This method is effective even if the attacker has already seized enough zombie computers to pose a threat.

3. Tracing the Source IP Address: The goal of this approach is to trace the intruders' path back to the zombie computers and stop their attacks or, even better, to find the original attacker and take legal actions. If tracing is done promptly enough, it can help to abort the DDoS attack. Catching the attacker would deter repeat attacks.

However, two attacker techniques hinder tracing:

- IP spoofing that uses forged source IP addresses, and

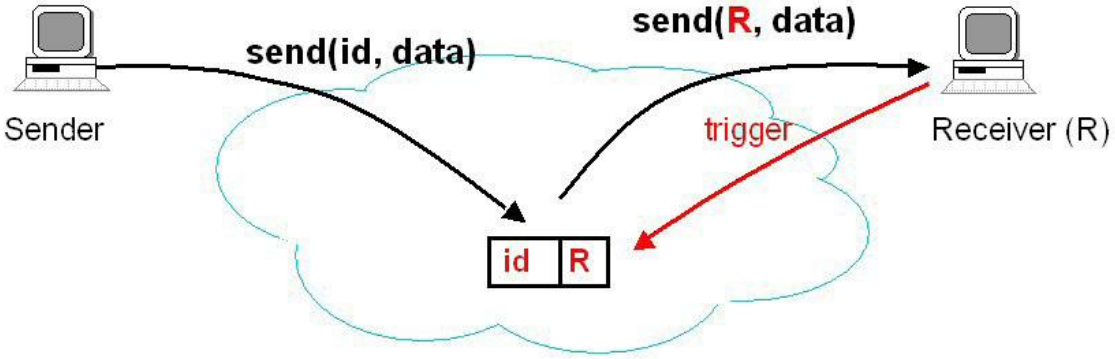
- The hierarchical attacking structure that detaches the control traffic from the attacking traffic, effectively hiding attackers even if the zombie computers are identified.

VIII. REVIEW

This section gives a brief overview of the. The contents of this are all about the various related to mobile ad hoc network, distributed denial of service attack and defense against this attack.

1) *Wireless Ad hoc Networks*

Lu Han describes [1] that the wireless ad hoc networks were first deployed in 1990's, Mobile Ad-hoc networks have been widely researched for many years. Mobile Ad-hoc Networks are collection of two or more devices equipped with wireless communications and networking capability. These devices can communicate with other nodes that immediately within their radio range or one that is outside their radio range. For the later, the nodes should deploy an intermediate node to be the router to route the packet from source toward destination. The Wireless Ad-hoc Networks do not have gateway, every node can act as the gateway. As per this paper, although, lots of research has been done on this particular field, it has often been questioned as to whether the architecture of Mobile Ad-hoc Networks is a fundamental flawed architecture. The main reason for the argument is that Mobile Ad-hoc Networks are not much used in practice, almost every wireless network nodes communicate to base station and access points instead of cooperating to forward packets hop-by-hop. As per the contents of this paper the key technologies to Wireless Adhoc Networks were not implemented as we expect. That is to say, many problems are inherently unsolvable.



2) *Security Threats In Mobile ad-hoc Networks*

Kamanshis Biswas and Md. Liakat Ali describes [2] that Mobile Ad Hoc Network (MANET) is a collection of communication devices or nodes that wish to communicate without any fixed infrastructure and pre-determined organization of available links. The nodes in MANET themselves are responsible for dynamically discovering other nodes to communicate. Although the ongoing trend is to adopt ad hoc networks for commercial uses due to their certain unique properties, the main challenge is the vulnerability to security attacks. A number of challenges like open peer-to-peer network architecture, stringent resource constraints, shared wireless medium, dynamic network topology etc. are posed in MANET. As MANET is quickly spreading for the property of its capability in forming temporary network without the aid of any established infrastructure or centralized administration, security challenges has become a primary concern to provide secure communication. This paper gives information about various security threats an ad-hoc network faces, the security services required to be achieved and the countermeasures for attacks in each layer. As per the contents of this paper, secure routing protocol is still a burning question. There is no general algorithm that suits well against the most commonly known attacks such as wormhole, rushing attack etc. In short, we can say that the complete security solution requires the prevention, detection and reaction mechanisms applied in MANET.

3) *Security In Ad-hoc Networks*

Vesa Kärpijoki describes [3] that in ad hoc networks the communicating nodes do not necessarily rely on a fixed infrastructure, which sets new challenges for the necessary security architecture they apply. In addition, as ad hoc networks are often designed for specific environments and may have to operate with full availability even in difficult conditions, security solutions applied in more traditionalnetworks may not directly be suitable for protecting them. A short literature study over papers on ad hoc networking shows that many of the new generation ad hoc networking proposals are not yet able to address the security problems and they face. Environment-specific implications on the required approaches in implementing security in such dynamically changing networks have not yet fully realized.

4) *Distributed Denial of Service:*

Taxonomies of Attacks, Tools and Countermeasures Stephen M. Specht and Ruby B. Lee describe [4] that Distributed Denial of Service (DDoS) attacks have become a large problem for users of computer systems connected to

the Internet. DDoS attackers hijack secondary victim systems using them to wage a coordinated large-scale attack against primary victim systems. As new countermeasures are developed to prevent or mitigate DDoS attacks, attackers are constantly developing new methods to circumvent these new countermeasures. This paper gives us information about DDoS attack models and proposed taxonomies to characterize the scope of DDoS attacks, the characteristics of the software attack tools used, and the countermeasures available. These taxonomies illustrate similarities and patterns in different DDoS attacks and tools, to assist in the development of more generalized solutions to countering DDoS attacks, including new derivative attacks. It is essential, that as the Internet and Internet usage expand, more comprehensive solutions and countermeasures to DDoS attacks be developed, verified, and implemented.

5) *Distributed Denial of Service Attacks*

Felix Lau, Stuart H. Rubin, Michael H. Smith and Ljiljana TrajkoviC describe [5] Distributed Denial of Service attacks in the Internet. They were motivated by the widely known February 2000 distributed attacks on Yahoo!, Amazon.com, CNN.com, and other major Web sites. A denial of service is characterized by an explicit attempt by an attacker to prevent legitimate users from using resources. An attacker may attempt to: "flood" a network and thus reduce a legitimate user's bandwidth, prevent access to a service, or disrupt service to a specific system or a user. This paper gives information about methods and techniques used in denial of service attacks, and list possible defenses. In this paper, distributed denial of service attack is simulated using ns-2 network simulator. This paper gives information about how various queuing algorithms implemented in a network router perform during an attack, and whether legitimate users can obtain desired bandwidth. In short, simulation results indicate that implementing queuing algorithms in network routers may provide the desired solution in protecting users in cases of distributed denial of service attacks.

6) *Denial of Service and Distributed Denial of Service Attacks*

Andrim Piskozub describes [6] main types of DoS attacks which flood victim's communication channel bandwidth, is carried out their analysis and are offered methods of protection from these attacks.

7) *A Survey of DDoS Defense Mechanisms*

Antonio Challita, Mona El Hassan, Sabine Maalouf and Adel Zouheiry describe [7] different types of DDoS attacks, present recent DDoS defense methods as published in technical papers, and propose a novel approach to counter DDoS. Based on common defense principles and taking into account the different types of DDoS attacks, this paper survey defense methods and classify them according to several criteria. This paper propose a simple-to-integrate DDoS victim based defense method, Packet Funneling, which aims at mitigating an attack's effect on the victim. In this approach, heavy traffic is "funneled" before being passed to its destination node, thus preventing congestion at the node's access link and keeping the node on-line. This method is simple to integrate, requires no collaboration between nodes, introduces no overhead, and adds slight delays only in case of heavy network loads. The proposed packet funneling approach promises to be a suitable means of coping with DDoS traffic, with easy integration at minimal cost

8) *Framework for Statistical Filtering Against*

DDoS Attacks in MANETs Hwee-Xian Tan and Winston K. G. Seah describes [8] that A DDoS (Distributed Denial-Of-Service) attack is a distributed, large-scale attempt by malicious users to flood the victim network with an enormous number of packets. This exhausts the victim network of resources such as bandwidth, computing power, etc. The victim is unable to provide services to its legitimate clients and network performance is greatly deteriorated. There are many proposed methods in the literature which aim to alleviate this problem; such as hop-count filtering, rate-limiting and statistical filtering. However, most of these solutions are meant for the wired Internet, and there is little research efforts on mechanisms against DDoS attacks in wireless networks such as MANETs. This paper gives information about the vulnerability of MANETs to DDoS attacks and provide an overview of statistical filtering, which is commonly used as a security mechanism against DDoS attacks in wired networks and then propose a framework for statistical filtering in MANETs to combat DDoS attacks. This paper also simulates some DDoS attacks in MANETs without any filtering mechanisms to explore and understand the effects of such attacks on the performance of the network.

9) *Defeating Distributed Denial of Service Attacks*

Xianjun Geng and Andrew B. Whinston describes [9] that the notorious, crippling attack on e-commerce's top companies in February 2000 and the recurring evidence of active network scanning a sign of attackers looking for network weaknesses all over the Internet are harbingers of future Distributed Denial of Service (DDoS) attacks.They signify the continued dissemination of the evil daemon programs that are likely to lead to repeated DDoS attacks in the foreseeable future. This paper gives information about network weaknesses that DDoS attacks exploit, the technological futility of addressing the problem solely at the local level, potential global solutions, and why global solutions require an economic incentive framework.

10) *On the Effectiveness of DDoS Attacks on Statistical Filtering*

Qiming Li, Ee-Chien Chang and Mun Choon Chan describes [10] that Distributed Denial of Service (DDoS) attacks pose a serious threat to service availability of the victim networkby severely degrading its performance. There has been significant interest in the use of statistical-based filtering to defend against and mitigate the effect of DDoS

attacks. Under this approach, packet statistics are monitored to classify normal and abnormal behaviour. Under attack, packets that are classified as abnormal are dropped by the filter that guards the victim network. This paper gives the effectiveness of DDoS attacks on such statistical-based filtering in a general context where the attackers are "smart". We first give an optimal policy for the filter when the statistical behaviours of both the attackers and the filter are static. Next, this paper considers cases where both the attacker and the filter can dynamically change their behavior, possibly depending on the perceived behavior of the other party. This paper observes that while an adaptive filter can effectively defend against a static attacker, the filter can perform much worse if the attacker is more dynamic than perceived.

IX. CHALLENGES IN MANETS

- MANETs face challenges in secure communication. For example the resource constraints on nodes in ad hoc networks limit the cryptographic measures that are used for secure messages. Thus it is susceptible to link attacks ranging from passive eavesdropping to active impersonation, message replay and message distortion.
- Mobile nodes without adequate protection are easy to compromise. An attacker can listen, modify and attempt to masquerade all the traffic on the wireless communication channel as one of the legitimate node in the network.
- Static configuration may not be adequate for the dynamically changing topology in terms of security solution. Various attacks like DoS (Denial of Service) can easily be launched and flood the network with spurious routing messages through a malicious node that gives incorrect updating information by pretending to be a legitimate change of routing information.
- Lack of cooperation and constrained capability is common in wireless MANET which makes anomalies hard to distinguish from normalcy.
- In general, the wireless MANET is particularly vulnerable due to its fundamental characteristics of open medium, dynamic topology, and absence of central authorities, distribution cooperation and constrained capability.

X. PROBLEMS DUE TO DOS/DDOS ATTACKS

- A denial-of-service attack (DoS attack) or distributed denial-of-service attack (DDoS attack) is an attempt to make a computer resource unavailable to its intended users.
- The bandwidth of a router between the Internet and a LAN may be consumed by DoS, compromising not only the intended computer, but also the entire network.
- Slow network performance (opening files or accessing web sites) due to Distributed Denial of Service attacks.
- Unavailability of a particular web site due to Distributed Denial of Service attacks.
- Inability to access any web site due to Distributed Denial of Service attacks.
- Dramatic increase in the number of spam emails received due to Distributed Denial of Service attacks.

Lot of work has to be done to solve these problems as given in literature survey. But there is no exact solution to these problems and I will try to detect and prevent DDoS attack.

XI. OBJECTIVES

To measure network performance which include parameters like first packet received at [s], last packet received at [s], average end-to-end delay [s], total number of bytes received, total number of packets received etc.

- Study the effect of Distributed Denial of Service (DDoS) attacks under different attack intensities, different number of attackers and different node mobilities.
- To measure impact of Distributed Denial of Service (DDoS) attacks on network performance.
- Detection of Distributed Denial of Service attacks in Mobile Ad-hoc Network.
- Prevention of Distributed Denial of Service attacks in Mobile Ad-hoc Network using defense techniques.
- Analysis of the effectiveness of the prevention techniques.

XII. CONCLUSION

Introductory of future plan,

- Used to improve Network Performance.
- Used to improve wireless network efficiency.
- Provide better security to the intended user.
- Solve the problems of bandwidth depletion and resource depletion.
- Helps in providing network resource such as a website, web service, or computer system to the legitimate users.

XIII. REFERENCES

1) Han L; Wireless Ad hoc Network; October 8, 2004.
2) Kamanshis Biswas and Md. Liakat Ali; Security Threats in Mobile Ad Hoc Network; Master Thesis; Thesis no: MCS-2007:07; March 22, 2007.
3) Vesa Kärpijoki; Security in Ad Hoc Networks; Helsinki University of Technology; HUT TML 2000.
4) Stephen M. Specht and Ruby B. Lee; Distributed Denial of Service: Taxonomies of Attacks, Tools, and Countermeasures; Proceedings of the 17th International Conference on Parallel and

Distributed Computing Systems, 2004 International Workshop on Security in Parallel and Distributed Systems, pp. 543-550; September 2004.

5) Felix Lau, Stuart H. Rubin, Michael H. Smith and Ljiljana TrajkoviC; Distributed Denial of Service Attacks; 2275-2280/2004 IEEE.
6) Andrim Piskozub; Denial of Service and Distributed Denial of Service Attacks; TCSET 2002; February 18-23, 2002; Lviv – Shavsko, Ukraine.
7) Antonio Challita, Mona El Hassan, Sabine Maalouf and Adel Zouheiry; A Survey of DDoS Defense Mechanisms; Department of Electrical and Computer Engineering American University of Beirut; {asc04,mhe03,sem05,atz00}@aub.edu.lb.
8) Hwee-Xian Tan and Winston K. G. Seah; Framework for Statistical Filtering Against DDOS Attacks in MANETs; Proceedings of the Second International Conference on Embedded Software and Systems; 2005 IEEE.
9) Xianjun Geng and Andrew B. Whinston; Defeating Distributed Denial of Service Attacks; February 2000.
10) Q. Li, E-C. Chang and M. C. Chan; On the Effectiveness of DDoS Attacks on Statistical Filtering; Proceedings of the 24th Annual Conference of the IEEE Communications Society (INFOCOM 2005), Miami; Mar 13-17, 2005.
11) Y. Kim, W. Lau, M. Chuah and J. Chao, PacketScore: Statistical-based Overload Control against Distributed Denial-of-Service Attacks, Proceedings of the 23rd Conference of the IEEE Communications Society (INFOCOM 2004), Hong Kong, Mar 7-11, 2004.
12) Bechler M, Hof H J, Kraft D, Pahlke F, Wolf L; A Cluster Based Security Architecture for Ad hoc Networks; 0-7803-8356-7/2004 IEEE.
13) P. Michiardi and R. Molva; Ad hoc Networks Security; IEEE Press Wiley; New York; 2003.
14) L. Zhou and Z.J. Haas, Cornell Univ.; Securing Ad hoc Networks; IEEE Network; Nov/Dec 1999; Volume: 13; pp. 24-30, ISSN: 0890-8044.
15) J. Mirkovic and P. Reiher; A Taxonomy of DDoS Attack and DDoS Defense Mechanisms; ACM Sigcomm Computer Communications Review; Vol. 34, No. 2, Apr. 2004.
16) Y. Kim, W. Lau, M. Chuah and J. Chao; PacketScore: Statistical-based Overload Control against Distributed Denial-of-Service Attacks; Proceedings of the 23rd Conference of the IEEE Communications Society (INFOCOM 2004); Hong Kong; Mar 7-11, 2004.
17) S. Bellovin; Distributed denial of service attacks; Feb. 2000; http://www.research.att.com/-smbhalks.
18) H. Wang, D. Zhang, and K. Shin; Detecting SYN flooding Attacks; IEEE INFOCOM 2002, Jun. 2002, pp. 1530-1539.
19) http://www.merl.com/projects/DenialServiceAttacks/

Semantic Web: Improving Web Search Using RDF Instead of XML

GJCST Classification
H.3.3, H.3.7

Fadi Al-Kalani[1],Mamoun .G. Awad[2],Nabeel Bani Hani[3]

***Abstract*- The aim of this research is to present the possible using of XML and RDF in developing the concept of Semantic Web. It is provide a brief introduction to the Semantic Web and defines related terms, in addition to discuss the optimize method to describe web contents using both XML and RDF. However, there have been different work dealing with the same issue, but our approach is more intended to deal with semantic information and produce different standard format. The approach focus on developing a simulate web search engine to describe and emphasize whether to deploy XML or RDF in web search.**

keywords- XML, RDF, OWL, W3C.

I. INTRODUCTION

The future and present generation of the Web is often characterized as the "Semantic Web" the concept emphasis that the information will no longer only be intended for human readers, but also for processing by machines, enabling intelligent information services, personalized Web-sites, and semantically empowered search-engines. The Semantic Web requires interoperability on the semantic level as well as semantic interoperability requires standards not only for the syntactic form of documents, but also for the semantic content. Proposals aiming at semantic interoperability are the results of recent W3C standardization efforts, notably Extensible Markup Language (XML)/XML Schema and Resource Description Framework (RDF)/RDF Schema (RDFS) [1].W3C standard document format for writing and exchanging information on the Web emphasis that XML is mostly concerned with syntax that does not make sense without semantics, and many recent activities aim at adding more semantic capabilities to XML(s.Decker ets 2000).However, RDF is mostly concerned about semantics which is not very useful in a computer system without syntax, and many recent activities aim at providing a syntactic grounding for RDF. In particular, the research introduce positive impact of using RDF instead of XML through the concept "Semantic Web" , with the knowledge that many of the researches demonstrated preference RDF for the many reasons which include simplicity, abstract syntax , and providing a data model.

About[1] -Delmon University,Bahrain Email:fadek2000a@yahoo.com
About[2]-University of Bahrain, Bahrain Email:mawad@asc.uob.bh
About[3]- University of Bahrain, Bahrain Email:nbanihani@asc.uob.bh

II. BACKGROUND

Theoretically, many of the researches demonstrated preference RDF rather than XML for following reasons (s.Decker ets 2000):

- The main advantage of RDF over the basic XML is its simplicity. Unlike the order of elements in XML, the order of RDF properties does not matter. RDF offers a very appealing and flexible solution to any web designer.
- RDF has an abstract syntax that reflects a simple graph-based data model, and formal semantics with a rigorously defined notion of entailment providing a basis for well founded deductions in RDF data.
- XML and RDF are the current standards for establishing semantic interoperability on the Web, but XML addresses only document structure. RDF better facilitates interoperation because it provides a data model that can be extended to address sophisticated ontology representation techniques (s.Decker ets 2001).

In Fact, many of these advantages are theoretically described and does not have a clear measure with the concept Semantic web , furthermore those advantages were build on with the assumption of the features of the both languages XML and RDF ,in addition, the impact of using RDF rather than XML is not measurable specially through the concept of semantic web. Actually, there have been different previous works dealing with similar problems. However, our proposal is more intended to deal with Semantic information and produce different standard formats. In particular, our main goal is to define the appropriate elements to develop a semantic web using both XML and RDF. Our approach achieve by developing a simulated web search engine to describe and emphasizes the positive role of using RDF rather than XML in web search.

III. RESOURCES AND METHODOLOGY OF THE RESEARCH

Currently most of the Web content is suitable for human use. Furthermore the typical uses of the Web today are information seeking, publishing, and using, searching for people and products, shopping, reviewing catalogues. But, today's Web search face many Limitations summarized as (Payam Barnaghi 2008):

- The Web search results are high recall, low precision.
- Results are highly sensitive to vocabulary.
- Results are single Web page.
- Most of the publishing contents are not structured to allow logical reasoning and query answering.

HTML is the language used to display graphics and text, but the contents describes cannot be processed by the machine. The semantic web is a tool to address this issue by introducing XML (extensible markup language), RDF (resource description framework), RDFS (RDF Schema) and OWL (web ontology language) to describe web contents that enable automated information access.

Initial work in this field of research is carried out by introducing the objectives of the research that will be achieve by developing application file system that offer specifically suited search and processing methods for given address file .In detail each file would contain information that describes common biography of apple. Deeply, results would have contextual analysis of the represented information and knowledge. However these aims should be addressed by the following steps:

- Develop two separate XML files contains common biography of apple. Typically, the property of the root apple consists of hotel, restaurant, hardware, and fruit. In addition, both files would contains information which represents summarize descriptions of the web contents .the structure of the first file shown in (fig. 1) where used to restore the data of apple with the categories elements defined as classification , name , and address. While the second file shown in (fig. 2) organized elements with name given as hotel, restaurant, hardware, and fruit. In fact, both files would present same information with different document structure.

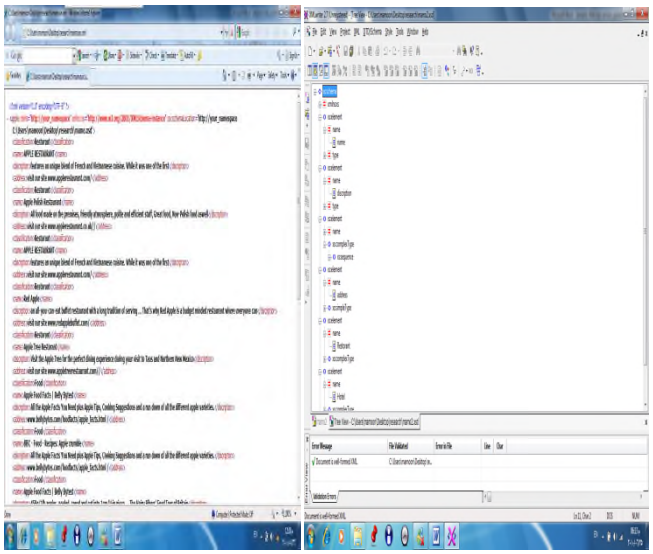


Fig. 1. XML source code and schema for the first file

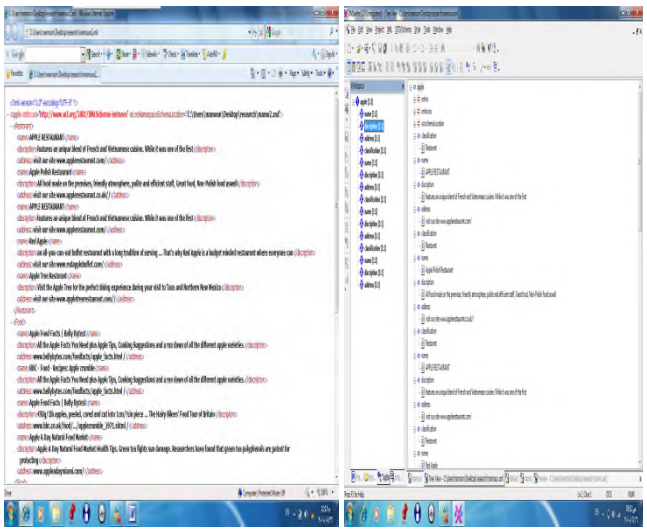


Fig. 2. XML source code and schema for the second file

- Develop java script files used to retrieve data from both XML and RDF file and display the results in web pages format.
- In order to complete the idea about our application's capabilities we present RDF file depicted in (fig. 3) to restore the same information described in XML files.

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
 xmlns:dc="http://APPLE_POLISH_RESTORANT"
xmlns:rd="http://APPLE_RED_RESTORANT">
  <rdf:Description rdf:about="APPLE_POLISH RESTORANT">
<dc:discribtion>features an unique blend of French and Vietnamese cuisine. While it
was one of the first</dc:discribtion>
   <dc:address>visit our site www.applerestaurant.com</dc:address>
  </rdf:Description>

  <rdf:Description rdf:about="APPLE_RED RESTORANT">
<rd:discribtion>an all-you-can-eat buffet restaurant with a long tradition of serving ...
That's why Red Apple is a budget minded restaurant where everyone
can</rd:discribtion>
   <rd:address>visit our site www.redapplebuffet.com</rd:address>
  </rdf:Description>
</rdf:RDF>
```

Fig. 3. RDF source code

As mentioned, our application process is based on the concept of Information Retrieval , where a structured XML file and RDF are used as input, and where the result is a presentation document in HTML format with using of the transformation language java script.

IV. Results of Research

The following part of our research examines and describes the output that has shown after the process of search using the input which representing XML and RDF files.

- *Why not XML*

The wise of develop two different structures XML files is attempt to answer the question: Is the web search limitation is a result of using XML in describing web content? Or if we think more in XML we can gain better results? Although, we used the two different structure XML files, mostly we found UN satisfactions results. The UN satisfactions results appear through a single result only or the Web search results are high recall

a) *Single result*

When uncompleted or the information is missed as a process of any search the impression of UN certainly and UN satisfaction of the process of the search would express.
On the one hand the information is available; on the other hand complete information is not displayed. Using of our developed simulated web search engine, we attempt to reach to specific information about APPLE RESTAURANT meaning that we need information about restaurant related to apple. According to that the equation assumed to be (word_search= "APPLE RESTAURANT"). Surprisingly, the result shown in (fig. 4) is totally uncompleted since there is more information that contains the word search.

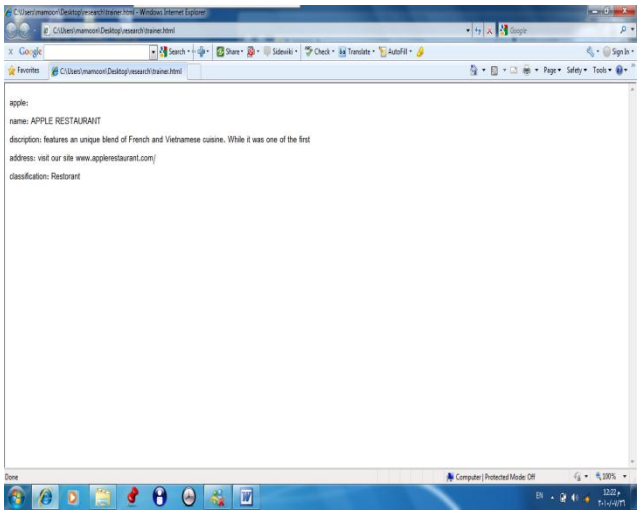


Fig.4. Single result displayed

Despite the fact that there are more results described APPLE RESTAURANT, but the result displayed show that this is the only information that is available. Essentially, we refer such of these results to the structure of XML which face Some of the problems with using attributes, XML Attributes are difficult to read and maintain, in addition it cannot contain multiple values as well as it cannot contain tree structures .those facts allow us to use elements for data and use attributes for information that is not relevant to the data. In other words XML is not support the idea that the attribute used to describe the element which allows us to give more information about each element and then specifically reach to a target element. Even using the two XML files described; the result is closely, Thus because XML structure enforce the element to be isolate from it description. In fact mostly we found this drawback result in many of library and bookshop web sites. For example unless typing the same book name, we wouldn't get a target result.Basically, we may refer the limitation search to the extremely using of XML in developing of library and bookshop web sites.

b) *High recall results*

The drawback characteristic of using attributes in XML described may cause another web search limitation represented in displaying many of the results that are not related to the target search. In our developed simulated web search engine we need to reach to the specific information about "Apple Hotel In the world ". Originally, the word search follow classification "Hotel", user expect to only get a fully information about Hotels in the world. In fact the results appear in (fig. 5) show information about Foods, Hardware, Restaurants and Hotels that may dispersal the user.

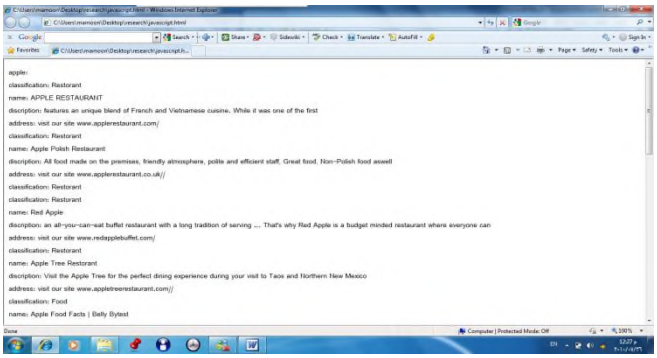


Fig.5. Many results displayed

Because we don't specify to each element it description so we couldn't specify a target word search. In other words the limitation appeared is a result of XML couldn't assign attribute to the element?

- *Why RDF*

As recognizable, XML itself is not concerned with meaning; however the standard of XML doesn't indicate how to derive a fact from a document. In fact XML documents are not useful for understandable but display documents.
RDF provides some basic level of meaning particularly, RDF is designed to represent knowledge in a distributed world, means RDF is particularly concerned with meaning.The second key aspect of RDF is that it works well for distributed information by establishing some relationships between documents that allow RDF applications files to put together .It does this in two ways, first by linking documents together by the common vocabularies they use, and second by allowing any document to use any vocabulary (http://www.rdfabout.com/intro/). Insider structure, RDF has XML syntax, consist of triples or sentences it consists of a resources, a property, and a value. In other sequence Object -> Attribute-> Value triples. Every resource has a URI

which can be a URL (a web address) or a some other kind of identifier; the Properties which are special kind of resources; used to describe relations between resources; and lastly Statements that is an object-attribute-value triple [9].Back to our developed simulated web search engine, according to (fig. 6) we convert part of XML file into RDF with the same given information.

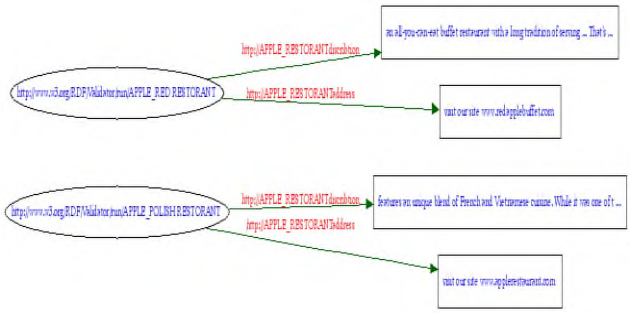


Fig.6. Graphical presentation of RDF file

Notable, we can easy reach to a particular and fully information depending on target word. for example if we looking to word search “APPLE RESTAURANT” then the result appear will contain only all the information describe "APPLE RESTAURANT "

V. Conclusions

XML is mostly concerned with syntax that does not make sense without semantics and can specify the structure of documents, not the meaning of the document contents .While Resource Description Framework (RDF) provides a standard for describing resources on the Web which gets us into metadata (data about data) and that is where things start getting particularly Semantic and exponentially more exciting.

VI. References

1) Stefan Decker, Frank van Harmelen, Jeen Broekstra,Michael Erdmann, Dieter Fensel, Ian Horrocks , Michel Klein & Sergey Melnik. (2001). The Semantic Web - on the respective roles of XML and RDF, 13, 44, 53231.
2) Stefan Decker, Sergey Melnik, Frank Van Harmelen, Dieter Fensel, Michel Klein, Jeen Broekstra, Michael Erdmann & Ian Horrocks. (2000, Sep./Oct.). The Semantic Web: The Roles of XML and RDF, IEEE Internet Computing, vol. 4(5), pp. 63-74.
3) Martha M. Yee, (2009, June). Can Bibliographic Data be Put Directly onto the Semantic Web? Information Technology and Libraries, University of California.
4) Enrique Chavarriaga & José A. Maciac. (2009). A model-driven approach to building modern Semantic Web-Based User Interfaces, Elsevier Ltd.
5) Nazaraf Shah, Chen-Fang Tsai1, Milko Marinov2, Joshua Cooper3, Pavel Vitliemov4 & Kuo-Ming Chao.(2009) Ontological On-line Analytical Processing for Integrating Energy Sensor Data”, IETE Technical Review, Vol 26(5).
6) Jorge Martinez-Gil, Enrique Alba & Jos_e F. Aldana-Montes. Optimizing Ontology Alignments by Using Genetic Algorithms , Boulevard Louis Pasteur s/n 29071 M_alaga (Spain) .
7) Payam .M. Barnaghi. (2007, April) Automatic Hypermedia Presentation Generation System, University of Malaya.
8) Payam Barnaghi. (2008). Introduction to the Semantic Web the University of Nottingham, Malaysia Campus.
9) W3C Semantic Web. Retrieved from: http://www.w3.org/2001/sw/ http://www.w3.org/TR/rdf-concepts/

Metrics and Heuristics in Software Engineering

Rakesh Kumar[1], Deepali Gupta[2]

GJCST Classification
I.2.8, D.2.8, D.4.8

Abstract- **Heuristics plays an important role in software development and are widely used to provide a link between design principles and software measurement. They offer insightful information based upon experience that is known to work in practice. Heuristics are not meant to be exact; in fact, they derive their benefits from this imprecision by providing an informal guide to good and bad practices. They provide a means by which knowledge and experience can be delivered from the expert to the novice. The paper is set out to bring techniques for building maintainable object oriented software closer to the developer in the form of design heuristics. Heuristics document common design problems that developers encounter during software development. Some heuristics in software engineering can be expressed in high-level abstract terms while others are more specific. The heuristic catalogue provides a comprehensive reference point for both novice and expert developers to apply well-documented techniques for building maintainable software.**

Keywords- Heuristics, OO Design, Metrics, Software Engineering and software metrics.

I. INTRODUCTION

Software engineering is the systematic collection of decades of programming experience together with the innovations made by researchers towards developing high quality software in a cost effective manner. In other words, software engineering is a systematic & cost effective approach to develop software. The basic objective of software engineering is to develop methods for developing software that can scale up & can be used to consistently develop high quality software at low cost. The improvements to software engineering over the last four decades have indeed been remarkable. Notable changes have occurred in software from error correction to error prevention. Now there are several development activities apart from coding like design, testing & maintenance. A lot of effort is now paid to requirements specification. Periodic reviews, testing, documentation, software project management are carried out during all stages of software development process. Software engineering has traditionally been an expensive and time-intensive process. Object-oriented analysis and design is the principal industry-proven methodology that answers the call for a more cost-effective, faster way to develop software and systems [6]. The basic design behind OOD is fundamentally different from the paradigm of function oriented design. In Object Oriented Design, data and operations are considered together, where as in case of function oriented approach the two are kept separate.Hence, OO design is one where the final system is represented in terms of object classes, relationships between objects and relationships between classes. An object is an instance of a class & has state, behaviour and identity. Key features of OOD are inheritance, information hiding, encapsulation, polymorphism, low coupling, high cohesion and modularity.A key element in software process is measurement. Software engineering, by is nature, is a quantitative discipline. Within the software engineering context a measure provides a quantitative indication of the extent, amount, dimension, capacity or size of some attribute of a product or process. Measurement is the act of determining a measure. The IEEE standard glossary [2] defines metric as "a quantitative measure of the degree to which a system, component, or process processes a given attribute". Software metrics provide a quantitative way to assess the quality of internal product attributes, there by enabling a software engineer to assess quality before the product is built. Metrics provide the insight necessary to create effective analysis and design models, solid code, and thorough tests.

There are set of attributes [1] defined that should be encompassed by effective software metrics. These are: -

1. **Simple and computable: -** It should be relatively easy to learn how to derive the metric and its computation should not demand inordinate effort and time.
2. **Empirically and intuitively persuasive: -** The metric should satisfy engineer's intuitive notions about the product attribute under consideration.
3. **Consistent and objective: -** The metric should always yield results that are unambiguous.
4. **Consistent in the use of units and dimensions: -** The mathematical computation used should be consistent in the use of units and dimensions.
5. **Programming language independent: -** Metric should be based on analysis model, design model, or the structure of the program itself.
6. **An effective mechanism for high quality feedback: -** Metric should lead to higher quality end product.

II. OBJECTIVES

The aim of Object Oriented (OO) Metrics is to predict the quality of the object oriented software products. Various attributes, which determine the quality of the software, include maintainability, defect density, fault proneness, normalized rework, understandability, reusability etc.Object Oriented (OO) Metrics are required because in OO code, complexity lies in interaction between objects, a large

About[1] - Reader, Department of Computer Science and Applications, Kurukshetra University, Kurukshetra, India. (Telephone: 09896336145 email: rsagwal@rediffmail.com)
About[2]- Dean Academics, Geeta Institute of Management and Technology, Kanipla, Kurukshetra, India. (Telephone: 09215710291 email: deepali_gupta2000@yahoo.com)

portion of code is declarative, OO models real life objects: classes, objects, inheritance, encapsulation, message passing. Software metrics helps to improve software process and its product. The use of object oriented software development techniques introduces new elements to software complexity both in software development process and in the final product [8].Our research investigates ways to help designers with the task of understanding, evaluating and improving their software products. While we view the art of design and the judgments of how to apply heuristics as beyond the reach of current technology, we argue that tools can provide valuable information to assist the designer with these judgments. OOD heuristics encapsulate software problems and their solutions in supporting an informal approach to design evaluation. Software design and development involves a range of practices with varying levels of formality: examples include formal methods, coding styles, design patterns and test-driven development. The common goal is the production of high quality software.However, quality is a concept that can not be measured directly. In order to measure and understand quality, it is necessary to relate it to measurable quantities. The field of software metrics deals with the identification of meaningful quantitative measures of specific properties of software.Heuristics enable a softer model to be constructed in order to obtain a more holistic and subjective, view of quality. This potentially places a greater burden on the developers who must interpret this view since it consists of potentially conflicting indicators with varying degrees of precision and relevance. Heuristics may occur as individual pieces of developers' or may be presented as a suite covering multiple aspects of software development.

III. Heuristics in Software Engineering

Everyday in our life, we do make the use of heuristic to solve the problem and software engineering is not an exception. In the past also we are seeing the use of some metric based heuristics in design and development of the software. For example, if the number of parameters in a function is more than five gives impression that module may not be having function cohesion. The heuristics are not written as hard and fast rules; they are meant to serve as warning mechanisms which allow the flexibility of ignoring the heuristic as necessary. Use of heuristics in modern OO software engineering has also been observed. Design is a difficult task because it involves finding compromises between conflicting pressures, cost and reliability. Designers must find ways to provide specific capabilities required by stakeholders, while attaining sufficient quality in emergent properties such as usability, efficiency, and flexibility. Software designers aim to satisfy the expectations of stakeholders by meeting functional and non-functional requirements.But in order to make this possible, they must first address the needs of the software developers themselves. Keeping the complexity of the design in check is foremost among these. Object-orientation (OO) allows software to be structured in a way that helps to manage complexity and change. However, as software reuse practitioners have discovered, realizing the benefits of OO is not straightforward. Competence with the mechanisms of classes and objects, attributes and methods, inheritance and polymorphism is far from sufficient to ensure successful designs.Metric-based heuristic framework is used to detect and locate object-oriented design flaws from the source code [7]. It is accomplished by evaluating design quality of an object-oriented system through quantifying deviations from good design heuristics and principles.Classes and Objects are the Building Blocks of the Object-Oriented Paradigm. Some of the heuristics proposed by Riel [3] are listed in Table 1 as follows:

Sno.	Heuristics
1	All data should be hidden within its class.
2	Users of a class must be dependent on its public interface, but a class should not be dependent on its users.
3	Minimize the number of messages in the protocol of a class.
4	Implement a minimal public interface that all classes understand.
5	Do not put implementation details such as common-code private functions into the public interface of a class.
6	Do not clutter the public interface of a class with things that users of that class are not able to use or are not interested in using.
7	A class should only use operations in the public interface of another class or has nothing to do with that class.
8	A class should capture one and only one key abstraction.
9	Keep related data and behavior in one place.
10	Spin off non related information into another class (i.e., non communicating behavior).
11	Be sure the abstractions that you model are classes and not simply the roles objects play.

Table 1. List of heuristics proposed by Riel.

Riel [3] documents 61 "golden rules" for OO design, while Fowler and Beck describe 22 code smells [4]. Smells evokes a subjective, subtle process of perceiving something about a design. Beck and Fowler noted that code smells do not lend themselves to automatic quantification [4]. The designer must form an impression of the net product of many factors at work in the design. This requires judgment and insight beyond the capabilities of simple automata. A notable characteristic of design patterns is that they often break rules. For example, the Composite pattern advocates the use of methods that are overridden to do nothing, contrary to a common maxim, expressed by Riel's heuristic as "It should be illegal for a derived class to override a base class method with a NOP method, i.e. a method which does nothing." However, the Gang of Four chose to break this rule

deliberately, in their words preferring transparency over safety. Many similar examples of conflicting forces can be found. Some conflicts are so pervasive that they apply to nearly all design situations. Separation of concerns, for example, encourages decoupling portions of a design, while another heuristics, "Keep related data and behavior in one place" often suggests the opposite. Even within an organized set of heuristics, conflicts occur. One heuristic says "Theoretically, inheritance hierarchies should be deep, i.e. the deeper the better", while another adds the qualification that "In practice, inheritance hierarchies should be no deeper than an average person can keep in his or her short-term memory. A popular value for this depth is six". Heuristics are a valuable tool for identifying design forces (whether conflicting or not) and evaluating design quality, but their application is not straightforward for many reasons [5], such as:

Lack of consensus on which heuristics should be adopted: Some conflicting heuristics usefully illuminate matters of concern to the designer. Other conflicts, however, reflect differing design philosophies, and a particular designer is likely to be interested only in one side of the debate. Many of the tenets arising from software reuse culture, for example, are in opposition to more recent refactoring and agile methods approaches. The open/closed principle, for example, encourages anticipation of future needs by making the design open for extension (reusable), but without requiring modification of existing code; refactoring culture discourages anticipation of future needs and prefers modifying existing code when necessary. This cultural difference might show up in unexpected ways, such as a stronger preference for small methods in the reuse culture, so that methods constitute small overridable units.

Nebulous definitions: One heuristic, for example, says "A class should capture one and only one key abstraction", but rigorously specifying the meaning of "key abstraction" is problematic. Similarly, heuristic "Model the real world whenever possible", is only as firm as our grip on reality.

Subjectivity and calibration: Code smells require the designer to judge when some intangible threshold has been crossed. The "large class smell", "lazy class smell" and "long method smell" are obvious examples where different standards might apply. The relative importance of conflicting heuristics is also dependent on the value system of the designer. If breaking up a large class produces a lazy class, is the result better?

Interpretation in different contexts: Many heuristics are expressed abstractly, in order to apply to any OO design. It may be necessary, however, to adapt a heuristic to local conditions. For example, when deciding if an inheritance hierarchy is too deep, should the root class be counted in programming languages that enforces a single root? Or, in an organisation that has adopted a refactoring approach to software development, how much emphasis should be placed on a heuristic motivated by software reuse, such as heuristic "All base classes should be abstract classes"?

Diverse levels of abstraction. Some heuristics can be interpreted at different levels. For example, "All data should be hidden within its class", might be viewed as a syntactic restriction make attributes private or as a semantic one, which might also discourage the use of getters. A "long method smell" could be detected at a lexical level by counting lines of code, at a syntactic level by counting statements and expressions, at a language semantic level by counting method invocations, collaborators, etc, or at a problem-domain semantic level by gauging the conceptual size of the method.

Information overload: Heuristics are intended to help software engineers manage the complexity of software, but injudicious application of heuristics could compound the problem.

Acquiring relevant data and relating it to heuristics: Many heuristics require substantial data gathering. Heuristic "Minimize fan-out in a class" and another "Most of the methods defined on a class should be using most of the data members most of the time" are examples. Additionally, the correspondence between available information and heuristics is not always clear.These issues, and the inherent fuzziness of heuristics, make automated support of heuristics difficult. In consequence, designers usually must gauge the quality of their products without assistance from tools. The designer builds a mental model of the software, and evaluates, according to a subjective, and perhaps even subconscious, process that is likely to be informed by heuristics, but may explicitly apply few.

IV. CONCLUSION

One of the mature engineering disciplines is the ability of its practitioners to quantify the quality of a product that is the ability to establish metrics. The use of metrics can be a valuable aid in understanding the effect of actions that are implemented for improving the software development process. The metrics provide visibility and control for the complex software development process, and therefore they are valuable for providing guidance on improving the software development process, and for meeting organizational goals to improve software quality and productivity. This causes new requirements for software metrics. While some of the traditional metrics can be used, new metrics must be introduced. The introduction of object-oriented (OO) methods to software development has changed the process of building and managing software in a profound way.Quality of software is increasingly important and testing related issues are becoming crucial for software. Although there is diversity in the definition of software quality, it is widely accepted that a project with many defects lacks quality. Methodologies and techniques for predicting the testing effort, monitoring process costs, and measuring results can help in increasing efficiency of software testing. Prediction of fault-prone modules supports software quality engineering through improved scheduling and project control. It is a key step towards steering the software testing and improving the effectiveness of the whole process In order to measure and understand quality, it is necessary to relate it to measurable quantities. Heuristics provide a link between sets of abstract design principles and

quantitative software metrics. They are an important part of software design and are becoming more widely used. Effective visualization of heuristics includes quantitative, qualitative and ambient aspects. Visualisation of heuristics provides many challenges. Heuristics are likely to be studied both individually and in comparison with others.The researchers are not primarily concerned with the relevance or validity of individual heuristics: the main focus is on their evaluation and interpretation. Our work is intended to provide the basis for an exploratory framework in which heuristics may be postulated, explored and managed.

V. REFERENCES

1) Lem O. *Ejiogu*, *Software engineering with formal metrics*, QED Information Sciences, Inc., Wellesley, MA, 1991.
2) *IEEE Standards Collection: Software Engineering,* IEEE Standard 610.12-1990, IEEE, 1993.
3) Riel A., *Object-Oriented Design Heuristics,* Addison-Wesley, 1996.
4) Fowler, M., *Refactoring: Improving the Design of Existing Code*, Addison - Wesley, 1999.
5) Churcher, N., et al., "Supporting OO Design Heuristics", *Proceedings of the 2007 Australian Software Engineering Conference,* IEEE Computer Society*,* 2007, pp. 101-110.
6) Berard, E.V, "Essays on Object-Oriented Software Engineering"*,* vol. 1, Addison Wesley, 1993.
7) Salehie, M., Li, S., Tahvildari, L., "A Metric-Based Heuristic Framework to Detect Object-Oriented Design Flaws"*,Proceedings of 14th IEEE International Conference on Program Comprehension,* IEEE Computer Society, 2006, pp.159-168.
8) Brooks, I., "Object-Oriented Metrics Collection and Evaluation with a Software Process," Proc. OOPSLA '93 Workshop Processes and Metrics for Object-Oriented Software Development, Washington, D.C., 1993.

The Role of Boolean Function in Fractal Formation and it's Application to CDMA Wireless Communication

GJCST Classification
C.2.1, I.2.3, G.1.2

Somnath Mukherjee[1], Pabitra Kumar Ghosh[2]

Abstract- **In this paper, a new transformation is generated from a three variable Boolean function 3, which is used to produce a self-similar fractal pattern of dimension 1.58. This very fractal pattern is used to reconstruct the whole structural position of resources in wireless CDMA network. This reconstruction minimizes the number of resources in the network and so naturally network consumption costs are getting reduced. Now -a -days resource controlling and cost minimization are still a severe problem in wireless CDMA network. To overcome this problem fractal pattern produced in our research provides a complete solution of structural position of resources in this Wireless CDMA Network.**

Keywords- Boolean functions, Level of Boolean function, Fractal Pattern, Wireless CDMA Networks, BTS, Wireless Network Port**.**

I. Introduction

A three variable Boolean function is used to generate the fractal pattern for the implementation of Wireless CDMA Network. More precisely, one transformation is generated from the three variable Boolean function named as 3 (named according to the Wolfram naming conventions [2]) which can produce the self-similar fractal pattern in a significant manner of increasing the level of Boolean function. In papers [1], [2] we have explored an application in the formation of self- similar and chaotic fractal formations, using one transformation named as „Carry Value Transformation' (CVT). In this paper, we have explored the algebraic beauties of the newly generated function and their application towards wireless communication problem basically on Wireless Code-Division-Multiple-Access (CDMA) network. It is to be noted that we have generated the fractals using a computational program written in C language and the compiler version is Borland C -3.0 by the newly defined transformation from the Boolean function 3. At first, we have generated one square matrix using the defined transformation as obtained in [1], and then we got the fractal using the matrix entities. The Wireless CDMA Network is designed based on the above generated fractal pattern. The recourses of the wireless CDMA Network if placed in the fractal pattern then the efficiency and effectiveness would be enhanced on the basis of design and maintenance cost. Interestingly, the services into this designed network are uniformly distributed.

About[1]-Dr.B.C.Roy Engineering College Email: somnath.7.mukherjee@gmail.comDurgapur, West Bengal-713206, India
About[2]-Dr.B.C.Roy Engineering College Email: pabitra06114@gmail.com. Durgapur, West Bengal-713206, India

II. Review of Earlier Works and Fundamental Concepts

In paper [3], resource allocation for the purpose of energy efficiency has been explored, but in this paper with the help of fractal geometry we implement the designing of the positional structure of resources in wireless CDMA network. Let us first warm up ourselves with some fundamentals, which are related to the current paper.

1) *Boolean Function*

A Boolean function f $(x_1,x_2,,x_3,x_4\ldots\ldots\ldots x_n)$variables is defined as a mapping from $\{0,1\}^n$ into $\{0,1\}$. It is also interpreted as the output column of its truth table f which is a binary string of length 2^n . For n -variables the number of Boolean functions is 2^{2n} and each Boolean function is denoted as f_R^n known as the function number R (also interpreted as rule number R), of n -variable. Here R is the decimal equivalent of the binary sequence (starting from bottom to top, with top is the LSB) of the function in the Truth Table, and numbering scheme is proposed by Wolfram and popularly known as Wolframs naming convention.

2) *Fractal Pattern*

The pattern, which reserves a fractional real number, as its fractal dimension is known as Fractal pattern. Here we use the similarity dimension as a fractal dimension. The similarity dimension is defined as follows
For a self-similar pattern, there is a relation between the scaling factor „S' and the number of pieces 'N' into which the pattern can be divided and that relation is
$N = 1/S^D$ this relation can be equivalently written
$D = \log N/\log (1/S)$.
This „D' is called fractional dimension or fractal dimension (self-similarity dimension).

3) *BTS (Base Transceiver Station)*

A base transceiver station or cell site (BTS) is a piece of equipment that facilitates wireless communication between user equipment (UE) and a network. UEs are devices like mobile phones (handsets), WLL phones, computers with wireless internet connectivity, WiFi and WiMAX gadgets etc. The network can be that of any of the wireless communication technologies like GSM, CDMA, WLL, WAN, WiFi, WiMAX etc. BTS is also referred to as the

radio base station (RBS), node B (in 3G Networks) or, simply, the base station (BS).

4) *Wireless Network Port*

Wireless Network ports are the points that contains the device to emit web to provide users uninterrupted network, it may be compared with the BTS(Base Transceiver Station) of the mobile communication

5) *Why. Wireless CDMA Network*

There are several wireless networks like TDMA(Time Division Multiple Access), FDMA(Frequency Division Multiple Access),CDMA(Code Division Multiple Access).In FDMA the available bandwidth is divided into frequency bands, that means each station is allocated to send its data. In TDMA each station share the bandwidth of the channel in time , each station is allocated a time slot during which it can send data. In both (FDMA & TDMA) case a switching technique is required to provide the network service to the station. CDMA(Code Division multiple Access) Network is a third-generation (3G) wireless communications and it is a form of multiplexing, which allows numerous signals to occupy a single transmission channel, optimizing the use of available bandwidth. The technology is used in ultra-high-frequency (UHF) cellular telephone systems in the range from 800-MHz to 1.9-GHz. In CDMA one channel carries all transmissions simultaneously where channel means a common path between sender and receiver [4]. CDMA simply means communications with different codes. Let us assume we have three stations namely station-1, station-2 and station-3 are connected with the same channel The data from station-1 is d1, from station-2 is d2 and so on are allocated for each station. Each station have to be assigned also a particular code e.g. c1,c2 and so on. The data carried out by the channel is the sum of the terms(d1.c1,d2.c2,d3.c3…), that means the data in the channel at any instant time is the sum of the values of d1.c1,d2.c2,d3.c3 and so. Any station wants to receive the data that is sent from the other station ,have to multiply the data on the channel by the code of the sender. That means here is no time delay for transmission of data like any others wireless communication. In case of CDMA there is no such switching technique is required to provide the network service. So the implementation of CDMA wireless network by the model of fractal pattern is possible in practically to provide the network services among the stations.

6) *Carry Value transformation (CVT)*

In [1] we have defined a new transformation named as CVT and shown its use in the formation of fractals.If and are two n-bit strings then is an (n+1) bit string, belonging to the set of non-negative integers, and can be computed bit wise by logical AND operation followed by a 0.Conceptually, CVT in binary number system is same as performing the bit wise XOR operation of the operands (ignoring the carry-in of each stage from the previous stage) and simultaneously the bit wise ANDing of the operands to get a string of carry-bits, the latter string is padded with a '0' on the right to signify that there is no carry-in to the LSB (the overflow bit of this ANDing being always „0' is simply ignored).

Example:Consider the CVT of the numbers (13)10 ≡ (1101)2 and (14)10 ≡ (1110)2. Both are 4-bit numbers. The carry value is computed as follows:

Carry: 1 1 0 0 0
Augend: 1 1 0 1
Addend: 1 1 1 0
XOR: 0 0 1 1

Carry genereted in ith column saved in (i-1)th column

In the above example, bit wise XOR gives (0011)2 ≡ (3)10 and bit wise ANDing followed by zero-padding gives (11000)2 ≡ (24)10. Thus and equivalently in decimal notation one can write . In the next section, a new notion of CVT named as Level Sensitive Carry Value Transformation is discussed.

III. Level Sensitive Carry Value Transformation

The Level Sensitive Carry Value Transformation is defined on the domain (ZxZxN) and it maps to Z. e. In other words, LSCVT is a mapping from ZxZxN->Z where Z is set of non-negative integers and N is the set of all natural number . Here we have considered Boolean function 3 and firstly the the decimal number „3' is converted to it's binary form then the corresponding binary values are assigned according to the functional values of the truth table of a three variable Boolean function. For Boolean function 3 the functional values of the truth table are assigned to the corresponding binary values of 3.

Function	----------	Value
f(0,0,0)	------	1.
f(0,0,1)	------	1.
f (0,1,0)	------	0.
f(0,1,1)	------	0.
f (1,0,0)	------	0.
f(1,0,1)	------	0.
f(1,1,0)	------	0.
f(1,1,1)	------	0.

Now the binary values of the corresponding any number in a matrix are

Z1 ----- the binary form of Z1
Z2 ----- the binary form of Z2
N ----- the binary form of N

The corresponding functional value from the above truth table.

Example:-

Z1=4, Z2=5, N=4

Then the corresponding binary values are

4 ----- 1 0 0
5 ----- 1 0 1
4 ----- 1 0 0

---

0 1 0

So the decimal value of the corresponding binary value is 2. Thus LSCVT(100,101,100) = 010 and equivalently in decimal notation it can be written LSCVT(4,5,4) = 2.

IV. Generation of Self-Similar Fractal Pattern Using LSCVT

A matrix is constructed that contains only the carry values (or even terms) defined above between all possible integers a's , b's and c's are arranged in an ascending order of x, y and z-axis respectively. We observe some interesting patterns in the matrix. We would like to make it clear how the matrix is constructed

Step 1: Arrange all integers 0,1,2,3,4 …(as long as we want) in ascending order and place it in three axis(x,y,z) in the matrix.

Step 2: Compute LSCVT (a,b,c) for c =1,2,3…using Boolean function 3 of three variable and store it in decimal form at (a,b,c) position.

Step 3: Then we notice on the pattern of value „0' and we have made it a specific color (green) and the other values except „0' also made a specific color (red) in the matrix. The pattern made by '0' in the matrix shown as a fractal (describe in figure -1below).

Z= 4 to 7

Y \ X	0	1	2	3	4	5	6	7
0	7	6	5	4	3	2	1	**0**
1	6	6	4	4	2	2	**0**	**0**
2	5	4	5	4	1	**0**	1	**0**
3	4	4	4	4	**0**	**0**	**0**	**0**
4	3	2	1	**0**	3	2	1	**0**
5	2	2	**0**	**0**	2	2	**0**	**0**
6	1	0	1	**0**	1	0	1	**0**
7	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

Figure-1Fractal formation in a matrix

Here Y represents the row of the matrix, X represents the Column and Z represents the level of the matrix. For an example, the positional decimal value of the matrix in the position (6,4 ,4) is 1. For Boolean function 3 the functional values of the truth table are assigned to the corresponding binary values of 3.

Function	----------	Value
f(0,0,0)	------	1.
f(0,0,1)	------	1.
f (0,1,0)	------	0.
f(0,1,1)	------	0.
f (1,0,0)	------	0.
f(1,0,1)	------	0.
f(1,1,0)	------	0.
f(1,1,1)	------	0.

Now the binary values of the corresponding positional number in the matrix are

6	-----	1 1 0
4	-----	1 0 0
4	-----	1 0 0
		----------
		0 0 1

So the decimal value of 001 is 1. In this way the total posional values are calculated in the matrix

1) *Observation*

We have observed the matrix and also found some interesting fractal pattern in specific level wise(level 0,1,2…,). We also consider the level of the Boolean function upto 255 from 0.and consider the level in a signifant of manner ,e.g. from level 0 to 1, from level 2 to 3,from level 4 to 7,from level 8 to 15, from level 16 to 31 , from level 32 to 63, from level 64 to 127, from level 128 to 255.1. From level (Z) 0 to 1 :- In this two level if we consider a 2x2 square matrix the following result is obtained:-At level 0 a square matrix is found which is formed by only 0 entities and at level 1 a square matrix is also found formed by 0 and 1 entities.The basic structure of the pattern of the square matrix at level 0 is below:-

Z=0

Y \ X	0	1
0	**0**	**0**
1	**0**	**0**

Figure-2 Formation of Euclidean geometry by the element 0

In this square matrix all combination is formed by 0 The basic structure of the pattern of the square matrix in level (Z) 1 as bellow-

Z=1

Y \ X	0	1
0	1	**0**
1	**0**	**0**

Figure-3 Formation of Fractal by the element 0

This is also a square matrix formed by 1 and 0. The pattern which is formed by 0 is symmetric to a right angle triangle In this level(z=1) if we calculate a 4x4 matrix ,then we can not get fractal pattern formed by matrix entities .of 0. The pattern is given below in the matrix:-

Z =2 to 3

Y \ X	0	1	2	3
0	3	2	1	**0**
1	2	2	**0**	**0**
2	1	**0**	1	**0**
3	**0**	**0**	**0**	**0**

Figure- 4 Formation of Fractal by the element 0

In this level if we also calculate 8x8 matrix, then we can not get the fractal pattern formed by 0 of the matrix . The pattern is given below:-

Z= 2 to 3

	0	1	2	3	4	5	6	7	X
0	3	2	1	**0**	3	2	1	**0**	
2	2	2	**0**	**0**	2	2	**0**	**0**	
2	1	**0**	1	0	1	**0**	1	**0**	
3	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	
4	3	2	1	**0**	3	2	1	**0**	
5	2	2	**0**	**0**	2	2	**0**	**0**	
6	1	**0**	1	**0**	1	**0**	1	**0**	
7	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	

Y

Figure-5 Formation of Euclidean geometry by the element 0

3. From level 4 to 7 :- In this four levels if we consider a 8x8 square matrix ,then at all levels (e.g. 4,5,6,7)the same basic fractal pattern formed by the 0 entities of the square matrix of level 2 to 3 is repeated and also regenerated ,The pattern is given below:-

Z=4 to 7

	0	1	2	3	4	5	6	7	X
0	7	6	5	4	3	2	1	**0**	
1	6	6	4	4	2	2	**0**	**0**	
2	5	4	5	4	1	**0**	1	**0**	
3	4	4	4	4	**0**	**0**	**0**	**0**	
4	3	2	1	**0**	3	2	1	**0**	
5	2	2	**0**	**0**	2	2	**0**	**0**	
6	1	**0**	1	**0**	1	**0**	1	**0**	
7	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	

Y

Figure -6 Formation of Fractal by the element 0

4.From level 8 to 15 :- In this levels if we consider 16x16 square matrix, at all levels (e.g. 8,9,..15 levels) the same basic pattern formed by 0 entities of the square matrix of level 4 to 7 is repeated and regenerated upto level 15

Z=8 to 15

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	X
0	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	**0**	
1	14	14	12	12	10	10	8	8	6	6	4	4	2	2	**0**	**0**	
2	13	12	13	12	9	8	9	8	5	4	5	4	1	**0**	1	**0**	
3	12	12	12	12	8	8	8	8	4	4	4	4	**0**	**0**	**0**	**0**	
4	11	10	9	8	11	10	9	8	3	2	1	**0**	3	2	1	**0**	
5	10	10	8	8	10	10	8	8	2	2	**0**	**0**	2	2	**0**	**0**	
6	9	8	9	8	9	8	9	8	1	**0**	1	**0**	1	**0**	1	**0**	
7	8	8	8	8	8	8	8	8	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	
8	7	6	5	4	3	2	1	**0**	7	6	5	4	3	2	1	**0**	
9	6	6	4	4	2	2	**0**	**0**	6	6	4	4	2	2	**0**	**0**	
10	5	4	5	4	1	**0**	1	**0**	5	4	5	4	1	**0**	1	**0**	
11	4	4	4	4	**0**	**0**	**0**	**0**	4	4	4	4	**0**	**0**	**0**	**0**	
12	3	2	1	**0**	3	2	1	**0**	3	2	1	**0**	3	2	1	**0**	
13	2	2	**0**	**0**	2	2	**0**	**0**	2	2	**0**	**0**	2	2	**0**	**0**	
14	1	**0**	1	**0**	1	**0**	1	**0**	1	**0**	1	**0**	1	**0**	1	**0**	
15	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	

Y

Figure-7 Formation of Fractal by the element 0

5 From level 16 to 31 :- In this levels if we consider 32x32 square matrix, at all levels (e.g. 16,17,18…….31 levels) the same basic pattern formed by 0 entities of the square matrix of level 8 to 15 is repeated and also regenerated upto level 31.

The same case is repeated and regenerated in also from level 32 to 63 if we consider 64x64 square matrix, and from 64 to 127 consider 128x128 square matrix ,and from 128 to 255 level consider 255x255 square matrix .

2) *Dimension of the pattern*

There are some level(Z) where we have got an Euclidean geometric figure(Figure-1,Figure-3,Figure-5),so we can not calculate the dimension of that particular pattern.

From level 1 to 255 if we calculate a particular square matrix, then the dimension of the pattern formed only by the entities 0 is fixed and it is D= log3/log2,

That is 1.5849 which is a Fractal.

3) *Analysis of the matrix*

We have analyzed the matrix and constructed a rule to get the corresponding level in which the fractal pattern belongs ..

- From level 1 if we consider 2x2 matrix then we get the expected pattern of Fractals.
- From level 2 to 3 if we consider 4x4 matrix then we get the expected pattern of Fractals.
- Level 4 to 7 if we consider 8x8 matrix then also get the expected pattern of Fractals.
- Level 8 to 15 and 16 to 31 and so on upto level 255 we get the expected pattern of fractals if we consider the order of matrix nXn,

Where n= the highest no. of level +1.

Illustratati
Level 1
if we consider the 1+1 =2 order matrix ,that is 2x2 matrix then we get the expected self symmetric pattern of Fractals.
Level 2 to 3
If we consider the 2+2=4 order matrix ,that is 4x4 matrix then we get the same result.
In the same iteration process all levels satisfy the condition as below:-

Level 2^0 to $(2^{(0+1)} - 1)$ ----calculate $2^{(0+1)}X2^{(0+1)}$ order matrix.
Level 2^1 to $(2^{(1+1)} - 1)$ ----- calculate $2^{(1+1)}X2^{(1+1)}$ order matrix.
Level 2^2 to $(2^{(2+1)} - 1)$ ----- calculate $2^{(2+1)}X2^{(2+1)}$ order matrix
Level 2^3 to $(2^{(3+1)} - 1)$ ------calculate $2^{(3+1)}X2^{(3+1)}$ order matrix.
Level 2^4 to $(2^{(4+1)}-1)$ ------calculate $2^{(4+1)}X2^{(4+1)}$ order matrix.
……
……..
……
……...
……
……
……
……
By method of Induction We can get
Level (Z variable in the square matrix) 2^n to $(2^{(n+1)} - 1)$ ----calculate $2^{(n+1)}X2^{(n+1)}$ order matrix to get the expected original pattern of fractals.
All levels L Є (2^n to $(2^{(n+1)} - 1)$.
Where n Є Z+ (set of all positive integer.)

V. Application

The application of these fractal patterns at different matrix dimension is towards the efficient as well as cost effective design of Wireless CDMA Networks. Our aim of the implementation is to share small resources among the vast wireless CDMA network. Let us take a small example; our basic pattern comes like :-

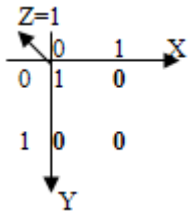
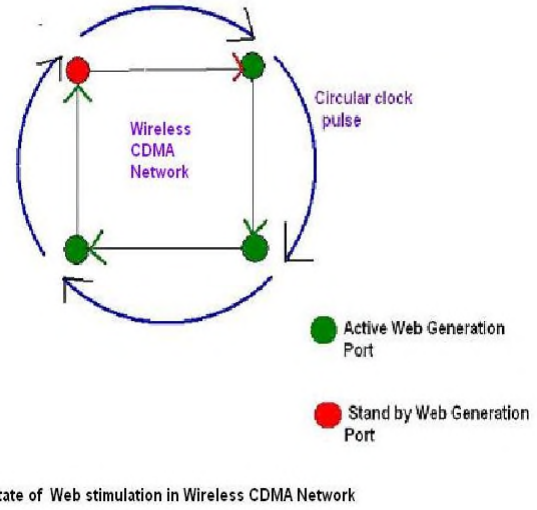


Figure-8 Basic Fractal pattern by the element 0 in 2x2 square matrix.

Suppose in this 2 x 2 matrix 4 blocks represent 4 resources and they access the wireless network. Now to provide the web service we need to have 4 active wireless web ports, now in our technique 4 ports are needed but we have to make active only 3 ports at an instant time and there is a certain switching devices which virtually rotates the apparent position of active wireless port of a certain clock speed in a specific direction. So, according to these concepts the signal that is sent to each user is same in respect of signal measurement and the clock speed is so high that it is seemed that the signal is continuous to the users. Now the quantities measure in respect of resource saving comes out in a very significant way. This process is like that whether we need to active 4 ports all times in a normal network structure, here just we need to active only 3 ports at all times and one port should stand by in its position. By some mathematical calculation we can say that our plan is {(1/4)*100} % = 25% efficient to save the resource for this basic pattern of Fractal.A typical diagram shows below to represent the Fractals in Wireless CDMA Network :-



The State of Web stimulation in Wireless CDMA Network

Figure -9

If we consider our general formula that is Level 2^n to $(2^{(n+1)} - 1)$ -------calculate $2^{(n+1)}X\ 2^{(n+1)}$ order matrix to get the expected original pattern of fractals.

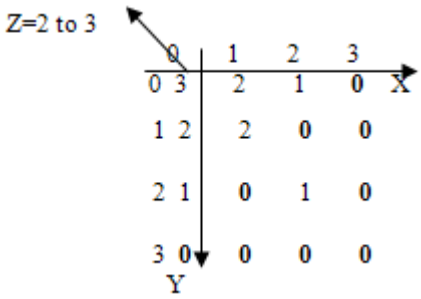


Figure-10 Formation of Fractal by the entities 0 in 4x4 Square matrix.

As the rule if we consider 4x4 matrix for level 3 the same pattern of the fractals is repeated and regenerated.
In this case we also apply the previous concept, if the total pattern rotates in a certain clock speed x in a certain direction;
then the elementary pattern that is depicted below

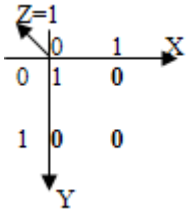


Figure -11 Elementary Fractal pattern by the element 0 in 2x2 Square matrix.

have to rotate in 4x speed at the same direction. Here the prefix 4 is to be added because after changing the initial position again returning to the initial position one pattern has to change position 4 times.By the help of simple arithmetic as the previous one we are able to calculate the % of efficient shave of the recourse;

{(7/16)*100}%=43.75%

It seems that our efficiency has grown up.Then if we calculate as our proposed formula 8x8 matrix at level 7 then the pattern is repeated as well as regenerated .

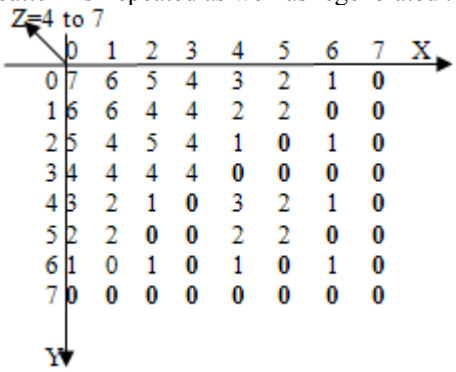


Figure-12 Fractal pattern by the element 0 in 8x8 Square matrix

Apply and approach same as the previous one. Now in this case if the total pattern rotates in certain clock speed x in certain direction then the corresponding elementary pattern level 1

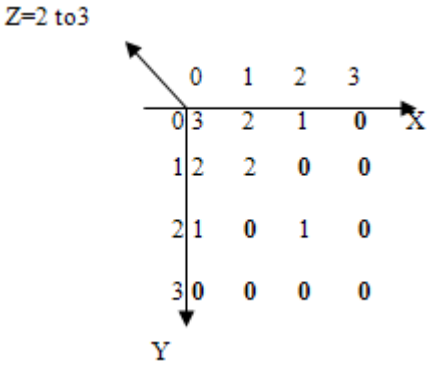


Figure-13 Fractal pattern by the element 0

have to switch in 4x speed in same direction and the corresponding elementary pattern level 2

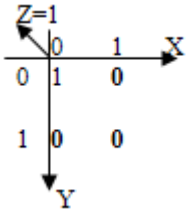


Figure- 14 Elementary Fractal pattern by the element 0

have to rotate in 4^2x speed. Same simple clock logic as the previous one. From the calculation as the previous one we are able to calculate the % of efficiency to optimizes the resources.

{(37/64)*100}% = 57.8125%

It shows that our efficiency has grown up gradually. Then the same occurrence repeated of the matrix arrangement at level 7 as the pattern is generated in case of 16 x 16 matrix and also the

% of efficiency for saving the resources increases;

{(172/256)*100}%=67.18%

Same pattern is regenerated for consideration of 32 x 32 matrix at level 16, the % of efficiency here is also save of the recourse and becomes;

{(781/1024)*100}%=76.26%

As the process goes on we can see that the efficiency of our implementation increases; So we can say that our proposed concept is more efficient for the large coverage area of a wireless CDMA network. In all above case we have to increase the speed elementary pattern level value ,(the level value may vary from 1 ,2,3,4,,5, ,n; where n is any natural positive number.) at the order of 4 level value . So according to this if the total pattern rotates at certain clock speed x ; then the speed of the then the elementary pattern at different level rotates according to the multiple of their corresponding level value . So, from the above discussion we can conclude that the speed of the particular elementary pattern at the particular level depends on the level value and that can be expressed as [$4^{\text{level value}} * x$].

VI. Comparison of our CDMA system with existing CDMA system

1) *Advantages*

There are mainly three advantages of our proposed CDMA system over the existing CDMA system. First of all number of resources are getting reduced at any instant time without affecting the network service to the station . Secondly as the number of resources getting reduced, so the cost of network service is also getting reduced. Lastly at any instant time the number of resources will remain standby in our CDMA system, so the energy will also save.

2) *Limitations*

There is a limitation to construct this CDMA network on the basis of fractal pattern, To get the fractal pattern the number of resources should be in our proposed manner, that means if we imagine the resources of the network as in the position of a matrix entities, then all the resources will be

reconstructed according to the position of the entities of a square matrix. If the above conditions will satisfy then, it is possible to construct such network model on the basis of fractal pattern.

VII. Conclusion and Further Research

In this paper we have used three variable Boolean function 3 in formation of fractal pattern and also completing the analysis we can say at level 2^n to $(2^{(n+1)}-1)$ of a level of a square matrix if we consider the $2^{(n+1)}x2^{(n+1)}$ square matrix the fractal pattern of the dimension same as the dimension of sierpinski triangle that means the dimension will be 1.5849 is produced. From this pattern model we can designing a new area of Wireless CDMA Network that can be used for a efficient and effective resources saving Wireless CDMA Communication .In further case if we try to implement the a new Wireless LAN technology ,it is possible to construct from our research .

VIII. References

1) P. P. Choudhury, S. Sahoo, B. K Nayak, and Sk. S. Hassan, Carry Value Transformation: It s Application in Fractal Formation 2009 IEEE International Advanced Computing Conference (IACC 2009), Patiala, India, 6-7 March, pp 2613-2618, 2009.
2) P. P. Choudhury, S. Sahoo, B. K Nayak, and Sk. S. Hassan, Act of CVT and EVT in Formation of Number-Theoretic Fractals (Communicated to World Scientific journal FRACTALS. Wolfram, Cellular Automata and Complexity Collected Papers, Addison-Wesley Publishing
3) Stefano Buzzi, Senior Member, IEEE, and H. Vincent Poor, Fellow, IEEE on Joint Receiver and Transmitter Optimization for Energy-Efficient CDMA Communications.
4) ISBN 0-201-62664-0, 1994. Data Communications and Networking by BEHROUZ A FOROUZAN forth edition published from Tata McGraw-Hill Publishing Company Limited, New Delhi. Special Indian Edition 2006, ISBN 0-07-063414-9

Multicategory Classification Using Support Vector Machine for Microarray Gene Expression Cancer Diagnosis

GJCST Classification
J.3, H.2.8

Dr.S.Santhosh Baboo[1], Mrs.S.Sasikala[2]

***Abstract*-This paper deals with the advanced and developed methodology know for cancer multi classification using Support Vector Machine (SVM) for microarray gene expression cancer diagnosis, this is used for directing multicategory classification problems in the cancer diagnosis area. SVMs are an appropriate new technique for binary classification tasks, which is related to and contain elements of non-parametric applied statistics, neural networks and machine learning. SVMs can generate accurate and robust classification results on a sound theoretical basis, even when input data are non-monotone and non-linearly separable. The performance of SVM is evaluated for the multicategory classification on benchmark microarray data sets for cancer diagnosis, namely, the SRBCT Data set. The results indicate that SVM produces comparable or better classification accuracies when the data given as input are preprocessed. SVM delivers high performance with reduced training time and implementation complexity is less when compared to artificial neural networks methods like conventional back-propagation ANN and Linder's SANN.**

***Keyword*-** SVM, ANOVA, Dataset,Cancer Classification and Gene Expression

I. Introduction

Cancer is one of the atrocious diseases found in most of the living organism, which is one of the challenging studies for scientist towards 20th century. There were lot of proposal from various pioneers and detailed picture study was still going on. Basically Cancer is characterized by an abnormal, uncontrolled growth that may destroy and invade adjacent healthy body tissues or elsewhere in the body. Living organisms such as animals and plants are made of cells. The simplest organisms consist of just a single cell. The human body compromises of billions of cells; most of the cells have a limited life-span and need to be replaced in cyclic manner. Each cell is capable of duplicating themselves. Millions of cell divisions and replications take place daily in the body and it is astounding that the process occurs so perfectly most of the time every cell division requires replication of the 40 volumes of genetic coding. On rare circumstances there is some defect in a division and a rogue, potentially malignant cell arises. The immune system seems to recognize such occurrences and is generally capable of removing the abnormal cells before they have an opportunity to proliferate. Rarely, there is a failure of the mechanism and a potentially malignant cell survives, replicates and cancer is the result.High-density DNA microarray computes the activities of several thousand genes simultaneously and the gene expression profiles have been used for the cancer classification recently. This new approach promises to give improved therapeutic measurements to cancer patients by diagnosing cancer types with improved accuracy. SVM is successfully applied to the cancer identification problems. However, it is most favorable extension to more than two classes was not obvious, which might impose limitations in its application to multiple tumor types. The binary SVM is applied to multiclass cancer diagnosis problems. SVM is a learning machine used as a tool for data classification, function approximation, etc, due to its generalization ability and has found success in many applications. Feature of SVM is that it minimizes the upper bound of generalization error through maximizing the margin between separating hyper plane and dataset. SVM has an added advantage of automatic model selection in the sense that both the optimal number and locations of the basis functions are automatically obtained during training. The performance of SVM largely depends on the kernel [23], [24].This paper presents a new technique for Multicategory Classification for Microarray Gene Expression Cancer Diagnosis Using Support Vector Machine for predicting cancer cells in living organism by the technique of ANOVA (Analysis Of Variance). This proposed system put fourth an accuracy of 100% by second level combinational technique. The multi-category cancer classification performance of SVM is evaluated on SRBCT Dataset. The evaluation results indicate that SVM produces better classification accuracy with reduced training time and implementation complexity compared to earlier implemented models. The remainder section of this paper is organized as follows. Section 2 discusses cancer classification systems with various classifying approach that were earlier proposed in literature. Section 3 explains the proposed work of developing a cancer classification system using a Support Vector Machine. Section 4 illustrates the results for experiments conducted on sample dataset in evaluating the performance of the proposed system. Section 5 concludes the paper with fewer discussions.

About[1]- Reader, PG and Research department of Computer Science, Dwaraka Doss Goverdhan Doss Vaishnav College Chennai, Tamil Nadu,INDIA.santhos2001@sify.com
About[2]- ResearchScholar,Dravidian University,AndraPradesh, Head Department of Computer Science, Sree Saraswathi Thyagaraja College, Pollachi, Coimbatore, Tamil Nadu, INDIA. Sasivenkatesh04@gmail.com

II. Related Work

Sridhar ramaswamy et al. [16] describes about multiclass cancer diagnosis using tumor gene expression signatures, which deliberately says about, the complex combination of clinical and histopathological data for optimal treatment of patients with cancer depends on establishing accurate diagnoses; it seems to be difficult because of atypical clinical presentation or histopathology. To determine whether the identification of multiple common adult malignancies could be achieved purely by molecular classification, for example the author, subjected 218 tumor samples, spanning 14 common btumor types, and 90 normal tissue samples to oligonucleotide microarray gene expression analysis. Here by using SVM the accuracy of multi class is predicted by expressing 16,063 genes and sequence tags. So this had an output of 95%, much greater than the accuracy of random classification that is about 9%. In recent times, [6] [7] DNA microarray-based tumor gene expression profiles have been used for cancer diagnosis. Anyhow, studies have been limited to few cancer types and have spanned multiple technology platforms complicating comparison among different datasets. The possibility of cancer diagnosis across all of the common malignancies based on a single reference database has not been explored. For a sample 314 tumors and 98 normal tissues were considered, in that 218 tumor and 90 normal tissue samples passed quality control criteria and were used for subsequent data analysis. The remaining 104 samples of the data will either fail the quality control measures or the quality of RNA, as assessed by spectrophotometric measurement of OD and agarose gel electrophoresis, or yielded poor-quality scans. Scans are discarded if the mean chip intensity exceeded 2 SDs from the average mean intensity for the whole scan set, if the proportion of present calls was less than 10%, or if microarray artifacts were visible. The problem of biological and measurement noise, contaminating nonmalignant tumor components, and inclusion of genetically heterogeneous samples within clinically defined tumor classes may all effectively decrease predictive power in the multiclass setting. Increased gene number likely allows for accurate prediction despite these factors. A greater variety and large number of tumors with detailed clinic pathological characterization will be required to fully explore the true limitations of gene expression-based multiclass classification. Lipo wang et al proposed the accurate cancer classification using expression of very few genes, the author aim at finding the smallest set of genes that can ensure highly accurate classification of cancers from microarray data by using supervised machine learning algorithms. The importance of finding the minimum gene subsets is three-fold: 1) It greatly reduces the computational burden and "noise" arising from irrelevant genes. From the examples stated in this paper, finding the minimum gene subsets even allows for extraction of simple diagnostic rules which lead to accurate diagnosis without the need for any classifiers. 2) The gene expression tests are simplified to include only a very small number of genes rather than thousands of genes, which can bring down the cost for cancer testing significantly.3) It calls for additional investigation into the possible biological relationship between these small numbers of genes and cancer development and treatment. Our simple yet very effective method involves two steps. In the first step, the author chooses some important genes using a feature importance ranking scheme. In the second step, the author tests the classification capability of all simple combinations of those important genes by using a good classifier. For three "simple" and "small" data sets with two, three, and four cancer (sub) types, our approach obtained very high accuracy with only two or three genes. For a "large" and "complex" data set with 14 cancer types, the author divided the whole problem into a group of binary classification problems and applied the 2-step approach to each of these binary classification problems. Through this "divide-and-conquer" approach, the author obtained accuracy comparable to previously reported results but with only 28 genes rather than 16,063 genes. In general, this method can significantly reduce the number of genes required for highly reliable diagnosis by the technique of SVM-T test analysis. The author analyzed finally and gave the accuracy rate of 100% by three combinational iteration techniques. Ahmad M. Sarhan suggests ,The cancer classification based on microarray gene expression data using DCT and ANN'. The author mainly deals about, a stomach cancer detection system based on Artificial Neural Network (ANN), and the Discrete Cosine Transform (DCT), is developed. The developed system extracts classification features from stomach microarrays using the DCT. The extracted features from the DCT coefficients are then applied to an ANN for classification (tumor or non tumor). The microarray images used in this study were obtained from the Stanford Medical Database (SMD). Simulation results showed that the developed system produces a very high success rate. DNA Microarrays are glass microscope slides onto which genes are attached at fixed and ordered locations. Each gene sequence is identified by a location of a spot in the array. Using a Microarray printer, the DNA is spotted directly onto the slide. With microarrays, it is possible to examine a gene expression within a single sample or to compare gene expressions within two tissue samples, such as in tumor and non tumor tissues. In this paper, a robust system for stomach cancer detection using microarrays is presented. The system consists of a feature extraction stage followed by an ANN classification stage. The feature extraction stage uses the 2 D DCT to compress the input microarray. Low frequency components of the DCT array constitute most of the energy/information of the input microarray. These components were, thus, used as distinctive features and were extracted using a windowing technique. The paper also investigates through simulations, optimal parameters such as the optimal number of DCT coefficients/features and the optimal ANN structure for the recognition of stomach cancer. The proposed method produces a success rate of 99.7%. The sensitivity, specificity, and accuracy of the system were found to be equal to 99.2%, 100%, and 99.66%

respectively. Experimental tests on the SMD Database achieved 99.7% of recognition accuracy using only100 DCT coefficients, with a simple 2-layer ANN structure and low computational cost. Runxuan Zhang en al. in [6] proposed a fast and efficient classification method called ELM algorithm. In ELM one may choose at random and fix all the hidden node parameters and then analytically determine the output weights. Studies have shown [2] that ELM has good generalization performance and can be implemented easily. Many nonlinear activation functions are used in ELM, like sigmoid, sine, hard limit [5], radial basis functions [3] [4], and complex activation functions [1]. In order to evaluate the performance of ELM algorithm for micro category cancer diagnosis, three benchmark micro array data sets, namely, the GCM, the lung and the lymphoma data sets are used. For gene selection recursive feature elimination method is used. ELM can perform multicategory classification directly with out any modification. This algorithm achieves higher classification accuracy than the other algorithms such as ANN, SANN and SVM with less training time and a smaller network structure.

III. Methodology

This proposed system mainly deals with cancer prediction by using SVM classification technique. SVM technique uses ANOVA test for grouping up the sample amount of sequential data. SVM technique overcomes the previous classification methodology by means of time consumption and by giving best accuracy rate. This projected method is comprised of two steps. In Step 1, all genes in the training data set are ranked using a scoring scheme. Then, the genes with high scores are retained. In Step 2, the classification capability of all simple combinations is tested among the genes selected in Step 1 using a good classifier. This paper proposes a new method of ranking with ANOVA and classifying with SVM. The mechanisms for Step 1 and Step 2 are described as follows.

Step 1: Gene Importance Ranking

In Step 1, the importance ranking of each gene is computed using a feature ranking measure, two of which are described below.

Only the most important genes are retained for Step 2.

1) *ANOVA (ANalysis Of VAriance)*

ANOVA is a technique, which is often used in analysis of data, and to draw interesting information based on P-values. The ANOVA is known to be robust and assumes that all the sample populations are normally distributed with equal variance and all observations (samples) are mutually independent. The approach chosen in this paper is the one-way ANOVA which performs an analysis on comparing two or more groups (samples) which in turn returns a single p-value that is significant for groups that are different from others. The most significant varying information has the smallest p-values. Within groups estimate of

$$\sigma_y^2 = \frac{\sum_{ij}(y_{ij}-\bar{y}_j)^2}{\sum_j(n_j-1)} = \frac{SS_{WG}}{df_{WG}} = MS_{WG} \quad (1)$$

Between –group estimate of

$$\sigma_y^2 = \frac{\sum_j n_j(y_j-\bar{y}_{ij})^2}{(K-1)} = \frac{SS_{BG}}{df_{BG}} = MS_{BG} \quad (2)$$

$$F\,(df_{BG}, df_{WG}) = b\frac{Between\ \ Group\ estimate\ of\ \sigma_y^2}{Within\ Groupestimate\ of\ \ \sigma_y^2} = \frac{MS_{WG}}{MS_{BG}} \quad (3)$$

Of all the information existing in the ANOVA table, if the p value for the F- ratio is less than the critical value (α), then the effect is said to be significant. In this paper the α value is set at 0.05, any value less than this will result in important effects, while any value greater than this value will result in non-significant effects. The very small p-value indicates that differences between the column means (group means) are highly significant. The probability of the F-value arising from two similar distributions gives us a measure of the significance of the between-sample variation as compared to the within-sample variation. Small p-values indicate a low probability of the between-group variation being due to sampling of the within-group distribution and small p-values indicate interesting features. This study uses the p-values to rank the important features with small values and the sorted numbers of features are used for further processing.

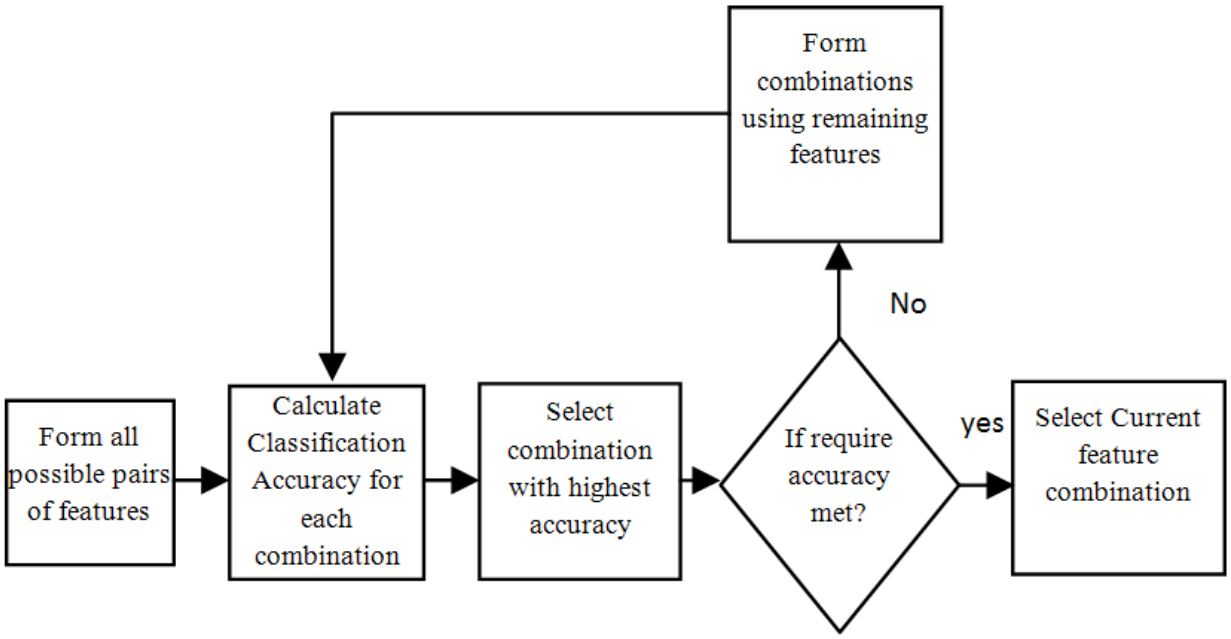


Fig.1: Proposed Feature Selection Method

Initially, all the features are ranked using a feature ranking measure and the most important features alone are retained for next the step. After selecting some top features from the importance ranking list, the data set is attempted to classify with only one feature. In this paper, the Support Vector Machine (SVM) classifier is used to test n-feature combinations.

a) *Class Separability*

Another frequently used method for gene importance ranking is the class separability (CS) [8]. The CS of gene I is defined as

$$CS_i = SB_i/SW_i \quad (4)$$

$$SB_i = \sum_{k=1}^{K}(\bar{x}_{ik} - \bar{x}_i)^2 \quad (5)$$

$$SW_i = \sum_{k=1}^{K}\sum_{j\in C_k}(x_{ij} - \bar{x}_{ik})^2 \quad (6)$$

For gene i, SB_i (the distances between samples of different classes) is the sum of squares of the interclass distances. SW_i (the distances of samples within the same class) is the sum of squares of the intraclass distances. A larger CS denotes a greater ratio of the interclass distance to the intraclass distance and, therefore, can be used to measure the capability of genes to separate different classes. In fact, the CS used here is similar to the F-statistic that is also widely used for ranking genes in literature (see, e.g., [12], [13]). The difference between the CS and the F-statistic F is:

$$CS = F.(K-1)/(\sum_{k=1}^{K} n_k\text{-}1) \quad (7)$$

Because the term

$$F.(K-1)/(\sum_{k=1}^{K} n_k\text{-}1) \quad (8)$$

CS equation is a constant for a specific dataset; the CS can be regarded as a simplification of F-statistic. The two methods will guide to the same ranking results for the same data set.

Step 2: Finding the Minimum Gene Subset

After selecting some top genes from the importance ranking list, the data set is attempted to classify with only one gene. Each selected feature is given as input into our classifier. If no good accuracy is obtained, continued classifying the data set with all the possible 2-feature combinations within the selected feature. If still no good accuracy is obtained, this procedure with 2-features combination is repeated and so on, until a good accuracy is obtained.

2) *Support Vector Machine (SVM)*

SVM is usually used for classification tasks introduced by Cortes [25]. For binary classification SVM is used to find an optimal separating hyper plane (OSH) which generates a maximum margin between two categories of data. To construct an OSH, SVM maps data into a higher dimensional feature space. SVM performs this nonlinear mapping by using a kernel function. Then, SVM constructs a linear OSH between two categories of data in the higher feature space. Data vectors which are nearest to the OSH in the higher feature space are called support vectors (SVs) and contain all information required for classification. In brief, the theory of SVM is as follows [27].

Consider training set $D = \{(x_j, y_i)\}_{i=1}^{L}$ with each input n i x $\in R^n$ and an associated output $y_i \in \{-1, +1\}$. Each input x is firstly mapped into a higher dimension feature space F, by

$z=\varphi(x)$ via a nonlinear mapping $\varphi: R^n \rightarrow F$. When data are linearly non-separable in F, there exists a vector $w \in F$ and a scalar b which define the separating hyper plane as:

$$y_i(w'.z_i + b) \geq 1 - \xi_i, \forall i \quad (9)$$

where $\xi(\geq 0)$ are called slack variable. The hyper plane that optimally separates the data in *F* is one that

$$minimise \frac{1}{2}.w'.w + C. \quad (10)$$
$$subject\ to\ y_i(w'.z_i + b) \geq 1 - \xi_i, \xi_i \geq 0, \forall i$$

where *C* is called regularization parameter that determines the tradeoff between maximum margin and minimum classification error. By constructing a Lagrangian, the optimal hyper plane according to previous equation, may be shown as the solution of

$$maximize\ W(\alpha) = \sum_{i=1}^{L} \alpha_i - \frac{1}{2}\sum_{i=1}^{L}\sum_{j=1}^{L} \alpha_i\alpha_j y_i y_j K(x_i, x_j) \quad (11)$$
$$subject\ to\ \sum_{i=1}^{L} y_i\,\alpha_i = 0, 0 \leq \alpha_i \leq C, \forall i$$

where $\alpha_1,\ldots..,\alpha_L$ are the nonnegative Lagrangian multipliers. The data points *i x* that correspond to $\alpha_i>0$ are SVs. The weight vector *w* is then given by

$$w = \sum_{ieSVs} \alpha_i y_i z_i \quad (12)$$

For any test vector $x \in Rn$, the classification output is then given by

$$y = sign(w.z + b) = sign(\sum_{ieSVs} \alpha_i y_i K(x_i, x) + b) \quad (13)$$

To build an SVM classifier, a kernel function and its parameters need to be chosen. So far, no analytical or empirical studies have established the superiority of one kernel over another conclusively. In this study, the following three kernel functions have been applied to build SVM classifiers:

1) Linear kernel function, $K(x,z) = \langle x,z \rangle$;

2) Polynomial kernel function $K(x, z) = (\langle x, z \rangle + 1)^d$ is the degree of polynomial;

3) Radial basis function $K(x,z) = \exp\left\{-\frac{||x-z||^2}{2\sigma^2}\right\}$, σ is the width of the function.SVM kernel functions The classification ability of feature combinations in cancer applications is obtained with first attempt work of SVM kernel function. The three main kernel functions are used for our study here. Partial kernel function, influence to data near test points. The above mentioned kernel functions are briefly explained in this chapter. The most used kernel function for SVM is Radial Basis Function (RBF). Radial Basis Function Kernel: The B-Spline kernel is defined on the interval [−1, 1]. It is given by the recursive formula:

$$k(x,y) = B_{2p+1}(x - y) \quad (14)$$
$$where\ p \epsilon N\ with\ B_{i+1} := B_i \otimes B_0$$

In the work by Bart Hamers it is given by:

$$k(x,y) = \prod_{p=1}^{d} B_{2n+1}(x_p - y_p) \quad (15)$$

Alternatively, B_n can be computed using the explicit expression (Fomel, 2000):

$$B_n(x) = \frac{1}{n!}\sum_{k=0}^{n+1} \binom{n+1}{k} (-1)^k (x + \frac{n+1}{2} - k)^n \quad (16)$$

Where x+ is defined as the truncated power function:

$$x_+^d = \begin{cases} x^d, & if\ x > 0 \\ 0, & otherwise \end{cases} \quad (17)$$

1. Linear Kernel: The Linear kernel is the simplest kernel function. It is given by the inner product <x,y> in addition with an optional constant c. Kernel algorithms using a linear kernel are often equivalent to their non-kernel counterparts.

$$k(x,y) = x^T y + c \quad (18)$$

2. Polynomial Kernel: The Polynomial kernel is a non-stationary kernel. Polynomial kernels are apt for problems where all the training data is normalized.

$$k(x,y) = (\propto x^T y + c)^d \quad (19)$$

Modifiable parameters are the slope **alpha**, the constant term **c** and the polynomial degree **d**.

IV. Experimental Results

In order to evaluate the performance of the SVM algorithm for multicategory cancer diagnosis SRBCT Dataset used in this paper. The SRBCT data set [28] contains the expression data of 2,308 genes. There are a total of 63 training samples and 25 testing samples already provided in [2]; five of the testing samples are not SRBCTs. The 63 training samples hold 23 Ewing family of tumors (EWS), 20 rhabdomyosarcoma (RMS), 12 neuroblastoma (NB), and eight Burkitt lymphomas (BL). And, the 20 SRBCT testing samples hold six EWS, five RMS, six NB, and three BL. As introduced in [6], for a microarray data with n genes, each ANOVA classifier produces a hyperplane w, which is a vector of n elements, each corresponding to the expression of a particular gene. The absolute magnitude of each element in w can be considered as a measure of the importance of each corresponding gene. Each ANOVA-SVM classifier is first trained with all of the genes, then the gene corresponding to the bottom 10 percent, w_{ij}, are removed. Each classifier is then again trained after the

removal of genes. This process is repeated with iterations and a rank of all of the genes based on the statistical significance of each class can be obtained. In the first step, the entire set of 2,308 genes is ranked according to their TSs in the training data set. Then, 50 genes with the highest TSs are picked out as showninTable1 (http://www.ntu.edu.sg/home5/pg02317674). SVM is used as the classifier for the SRBCT microarray data set. The expression data of the gene ranked 1 in referenceTable1 (http://www.ntu.edu.sg/home5/pg02317674) is used to train and then test the SVM. This process is repeated with the top two genes in Table 1 (http://www.ntu.edu.sg/ home5/pg02317674), then the top three genes, and so on. The testing error reduced to 0 when the top 16 genes were input into the SVM. To further decrease the number of genes required for accurate classification of the four cancer subtypes, all possible combinations of one gene and two genes within the 50 selected genes are tested. None of them can lead to 100 % CV accuracy for training data.Then, all possible combinations of three genes within the 50 selected genes were tested. Among all 19,600 such 3-gene combinations, the 5-fold CV accuracy for the training data found to be reached 100 percent with the combination of Gene 5, Gene 12, and Gene 30. The testing accuracy of this combination is 95 % (one error for the 20 testing samples). Because this combination contains only three genes, it is still possible for us to visualize it. Two different views of the 3D plot of this 3-gene combination is shown in fig.2.

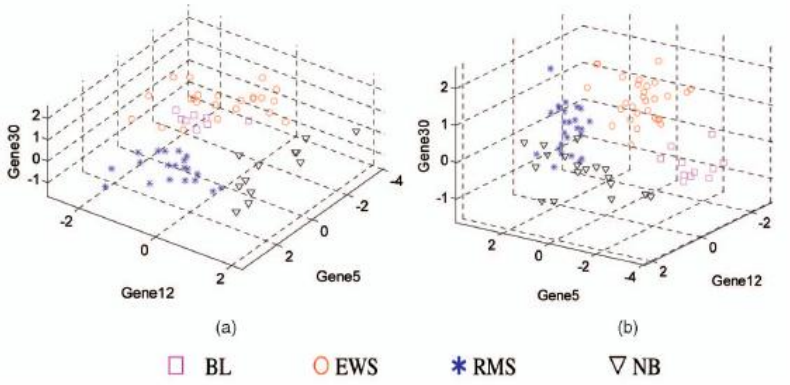


Fig.2: Three-dimensional views of gene expression

levels of 3-gene combination (5, 12, 30) that separates the four SRBCT subtypes, i.e., BL, EWS,RMS, NB: (a) A view in which RMS and NB can be seen to be separated from other subtypes. (b) A view in which BL can be seen to be separated from EWS. From the two views, the four subtypes are clearly separable. Here, all of the genes are labeled according to their TS ranks. ANOVA found only one such 3-gene combination that achieved 100 percent 5-fold CV accuracy.In the view of Fig. 2a, RMS and NB are well separated from other types. In the view of Fig. 2b, BL and EWS are well separated with a clear boundary. It is obvious that the four SRBCT subtypes are separated from one another, when observed from these two views. Except for (5, 12, and 30), no other gene combinations obtained 100 % CV accuracy for the training data. Even some combinations achieved 100% testing accuracy.

Table 1: Comparison of Accuracy of SVM with PCA and T-Test

Algorithm	Accuracy (%)
SVM with Raw data	95.51
SVM with preprocessed data	97.28

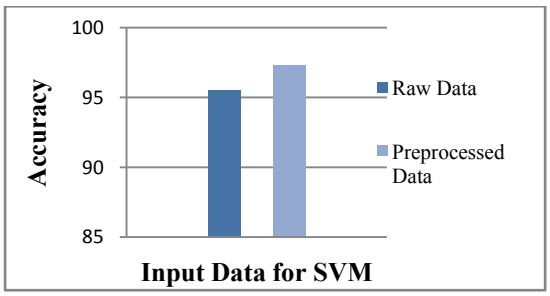


Fig. 3: Comparison of accuracy for the Raw and preprocessed dataset on SVM

Fig.3 displays the comparison of achieved by SVM for two input datasets which are preprocessed and raw dataset. The preprocessed dataset input delivers high accuracy.

V. Conclusion

In this paper, a fast and efficient classification method called the SVM algorithm for a multicategory cancer diagnosis problem based on microarray data is presented. Its performance has been compared for the raw data and ANOVA preprocessed data. It is found that SVM performs better with high accuracy when the data is preprocessed and given as input. The previous methods inevitably involve more classifiers, greater system complexities and computational burden, and a longer training time. SVM can

carry out the multicategory classification directly, without any modification. Study results are consistent with our hypothesis that, even when the number of categories for the classification task is large, the SVM algorithm achieves a higher classification accuracy than the other algorithms with less training time and a smaller network structure. It can also be seen that SVM achieves more balanced and better classification for individual categories as well. Theoretical investigation on these is currently under way.

VI. References

1) M. Ringner, C. Peterson, and J. Khan, "Analyzing Array Data Using Supervised Methods," Pharmacogenomics, vol. 3, no. 3, pp. 403-415, 2002.
2) G.-B. Huang and C.-K. Siew, "Extreme Learning Machine: RBF Network Case," Proc. Eighth Int'l Conf. Control, Automation, Robotics, and Vision (ICARCV '04), Dec. 2004
3) D. Serre, Matrices: Theory and Applications. Springer-Verlag, 2002.
4) G.-B. Huang, L. Chen, and C.-K. Siew, "Universal Approximation Using Incremental Constructive Feedforward Networks with Random Hidden Nodes," IEEE Trans. Neural Networks, vol. 17, no. 4, pp. 879-892, 2006.
5) S. Dudoit, J. Fridlyand, and T.P. Speed, "Comparison of Discrimination Methods for Classification of Tumors Using Gene Expression Data," J. Am. Statistical Assoc., vol. 97, no. 457, pp. 77-87, 2002.
6) Runxuan Zhang, Guang-Bin Huang, Narasimhan Sundararajan, and P. Saratchandran, " Multicategory Classification Using an Extreme Learning Machine for Microarray Gene Expression Cancer Diagnosis, " vol 4,no 3,july-september 2007.
7) M. Schena, D. Shalon, R.W. Davis, and P.O. Brown, "Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray," Science, vol. 270, pp. 467-470, 1995.
8) S. Dudoit, J. Fridlyand, and T.P. Speed, "Comparison of Discrimination Methods for the Classification of Tumors Using Gene Expression Data," J. Am. Statistical Assoc., vol. 97, pp. 77-87, 2002.
9) R. Linder, D. Dew, H. Sudhoff, D. Theegarten, K. Remberger, S.J. Poppl, and M. Wagner, "The 'Subsequent Artificial Neural
10) Network' (SANN) Approach Might Bring More Classificatory Power to ANN-Based DNA Microarray Analyses," Bioinformatics, vol. 20, no. 18, pp. 3544-3552, 2004
11) G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks," Proc. Int'l Joint Conf. Neural Networks (IJCNN '04), July 2004.
12) G.-B. Huang and C.-K. Siew, "Extreme Learning Machine: RBF Network Case," Proc. Eighth Int'l Conf. Control, Automation, Robotics, and Vision (ICARCV '04), Dec. 2004.
13) G.-B. Huang and C.-K. Siew, "Extreme Learning Machine with Randomly Assigned RBF Kernels," Int'l J. Information Technology, vol. 11, no. 1, 2005.
14) G.-B. Huang, Q.-Y. Zhu, K.Z. Mao, C.-K. Siew, P. Saratchandran, and N. Sundararajan, "Can Threshold Networks Be Trained Directly?" IEEE Trans. Circuits and Systems II, vol. 53, no. 3, pp. 187-191, 2006.
15) M.-B. Li, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "Fully Complex Extreme Learning Machine," Neurocomputing, vol. 68, pp. 306-314, 2005.
16) S. Ramaswamy, P. Tamayo, R. Rifkin, S. Mukherjee, C.-H. Yeang, M. Angelo, C. Ladd, M. Reich, E. Latulippe, J.P. Mesirov, T. Poggio, W. Gerald, M. Loda, E.S. Lander, and T.R. Golub, "Multiclass Cancer Diagnosis Using Tumor Gene Expression Signatures," Proc. Nat'l Academy Sciences, USA, vol. 98, no. 26, pp. 15149-15154, 2002.
17) R. Linder, D. Dew, H. Sudhoff, D. Theegarten, K. Remberger, S.J. Poppl, and M. Wagner, "The 'Subsequent Artificial Neural Network' (SANN) Approach Might Bring More Classificatory Power to ANN-Based DNA Microarray Analyses," Bioinformatics, vol. 20, no. 18, pp. 3544-3552, 2004.
18) O. Troyanskaya et al., "Missing Value Estimation Methods for DNA Microarrays," Bioinformatics, vol. 17, pp. 520-525, 2001.
19) M. West, C. Blanchette, H. Dressman, E. Huang, S. Ishida, R. Spang, H. Zuzan, J.A. Olson Jr., J.R. Marks, and J.R. Nevins, "Predicting the Clinical Status of Human Breast Cancer by Using Gene Expression Profiles," Proc. Nat'l Academy of Sciences USA, vol. 98, pp. 11 462-11 467, 2001.
20) E. Freyhult, P. Prusis, M. Lapinsh, J.E. Wikberg, V. Moulton, and M.G. Gustafsson, "Unbiased Descriptor and Parameter Selection Confirms the Potential of Proteochemometric Modelling," BMC Bioinformatics, vol. 6, no. 50, 2005.
21) S. Dudoit, M.J.V.D. Laan, S. Keles, A.M. Molinaro, S.E. Sinisi, and S.L. Teng, "Loss-Based Estimation

Novel Search and Retrieval Based on Domain Ontology

Hasnain Ali Salman, Xu De Zhi

GJCST Classification H.3.3

Abstract- **This paper proposes a framework of domain Ontology - based scientific and information retrieval (IR), and makes in - depth study on information organization and semantic retrieval. A prototype information retrieval information system is also implemented via a series of retrieval effects tests. The domain Ontology - based retrieval information system and the retrieval process of the information systematic prototypes are also depicted. In the end, the model is validated through a platform of trial information system. The result shows that when contrasted with traditional retrieval information system; this information system has a strong function in extending the connotation and denotation of the search words. So, it has the superiority in enhancing greatly the precision and recall.**

Keywords- Information retrieval (IR); Domain Ontology; Information organization; Semantic annotation; Semantic retrieval

I. Introduction

With the development of information technology, network-based environment of professional information resources increasing, and gradually evolved into a distributed, loosely network information environment. Information on the diversity of complex and heterogeneous information systems and so the resources of the professional literature, information organization and retrieval and use of proposed New challenges. Traditional information retrieval (IR) with the information system is mainly the professional literature to string matching and meta-data as a basis, but that expression match the same string Issues, rather than the meaning word form matching, vocabulary, and other isolated defects [1 - 4]; meta-data program is also the applicability of their capabilities and the evolution of different knowledge information systems and "Granularity"[2] of the resources there exist certain limitations described in, which have led to the current professional literature on information retrieval (IR) information system is less desirable[3]. As a new concept and method of information organization, ontology for the semantic retrieval theory provides an important way to solve the ideas and implementation. Since the ontology is cited in to the field of information retrieval since researchers in ontology construction [6], semantic annotation, retrieval, matching, similarity algorithms and search results in such areas as research, proposed Some distinctive ontology (IR) models and information systems [5 -7]. These studies indicate that contribute to the semantic level through ontology information retrieval, and to a certain extent, improve search results. However, current ontology-based information retrieval research focuses on enterprise information systems integration, Web Page retrieval and knowledge in the field, research priorities, but focused on search technology, reasoning logic and algorithm level, on the body of information retrieval in the professional literature Role, the information system architecture and implementation of less involved. As an important part of, (IR) professional literature has its own characteristics and research value [8]. Professional literature is not the same as unstructured text, thematic data Library contains a large number of staff by the use of indexing knowledge organization tools in understanding the original Based on the literature began to extract the contents of the semantic metadata[9-10-11]. In the information group Organization and information retrieval should take full advantage of the semantics of these meta-data value. And literature from the areas of information organization and heterogeneous concept, the concept query expansion and Semantic retrieval in various aspects of the implementation of the overall study. The paper is organized as follows: section 2 introduces related work on experimentation, while in section 3 we briefly introduce and the methods. The main contribution of this work is presented in section 4, which describes the experiments performed and the result analysis, and in section 5 that draws conclusions and outlines future work.

About- This paper is supported by the national natural science foundation of china (60970096)

II. Information Retrieval (IR) Model Based of Ontology

To solve the above problem, this paper specifically based on domain ontology Industry Information Retrieval model, its basic ideas are:-

1) With the current more mature information organizing tool to reflect Concepts and knowledge areas of the domain ontology.
2) In accordance with the domain ontology on the concept of the professional literature in the areas of progress Line automated semantic annotation, build the concept of the document vector.
3) When users search using the domain ontology of the query request extraction and semantic concept expansion of the concept of building a query vector.
4) Calculate the query concept vector and documents related to the concept of vector degree retrieve the relevant documents to meet client requirements.

The model framework Figure 1

Figure 1 Search model of professional literature based on domain ontology

Compared with existing research, retrieval model proposed in this paper mainly the following characteristics:-

1) In reference classification, topic and other traditional information organization method And meta-data methods based on the concept of ontology by constructing field, coding information System ontology and the formal description of the professional literature to the field of ontology concepts and heterogeneous integration and organization of the professional literature information.
2) The use of thesauri and classification to achieve the concept of ontology from the field Based on the dynamic transformation of semantic distance between concepts, semantic coincidence degree, Degree of relationship between the concepts of correlation is calculated to achieve a quantitative check Expansion.
3) The proposed relevance weights with semantic annotation method automatic processing without manual indexing of the professional literature, and expertise that have been indexing Industry documents, you can use its semantic metadata, semantic annotation to enhance effectiveness.
4) In the search stage, the proposed algorithm has certain adaptability, can be avoided because of incomplete ontology domain concepts influencing the cable results.

The following are the purposes of this organization in (IR) model, query processing, check Implementation in such areas as cable further elaboration.

1) *Professional Literature Information Organizations*

To achieve the professional literature of information organization, we takes the field was constructed concept ontology, ontology coding information system and the professional literature three inter-related body, domain and ontology. Among them, the concept of ontology to relevant areas of professional leader domain concepts and their hierarchical relationships, the use of natural words, etc For formal description coding information system ontology description of the professional literature Some external features of the above property values of the standardized, centralized management and maintenance of protection-related code, to ensure the encoding flexibility, scalability and reusability. Ontology defines the professional literature, professional literature class and is Resistance to the class instance to represent the professional literature and an example of real property are the external features of the literature, semantic information, and examples of inter-related documents formal description.

a) *Ontology*

The concept of ontology construction field is the basis of the whole information system and key High-quality domain ontology can bring better search results, but also means relatively high costs of ontology construction [1]. From the applicability, development is difficult degree of maturity and technical aspects, this article uses a thesaurus and classification to achieve the automatic transformation of ontology domain concepts [12-13-14], and also on the semantic distance between concepts, semantic coincidence degree, the concept of relevance, etc. Quantitative terms, the concept of similarity were constructed and the correlation matrix, a numerical Act sees Ontology coding information system is relatively simple, relevant standards and regulations [15] on the professional literature of the "type", "format", "Language", "time and space Range "and other attributes of the code are more clearly defined, by coding body Department of ontology on the definition of the formal description. Ontology construction in the professional literature, first established “Professional literature" Class as a top-level class, then according to relevant standards and norms [16-17] for the top layer type set up "Name", "Creator" "Association" And "Semantic Annotation" and other 16

properties, a number of attributes and then set the child property. For example, "Semantic Annotation". These attributes have "field concept" and "relevance weights". Two sub-attributes, "Professional Literature" was established under the top level category "Books" ,"Journal" ,"Conference paper", "Standard" , sub-class of 11 to represent a variety of documents types, these sub-classes inherit "Professional Literature" Class of all properties. Taking into account the special nature of various types of literature, but also the addition of these sub-categories Special attributes, such as "Conference paper" ,Class set", "Conference" Is a knowledge organization and knowledge management of, and contains The meeting name "TECHNOLOGY" ,"Meeting Date" ,,"Meeting place" and "Conference Name" and other sub-attributes; "Standard", Class from ,"Professional Literature" class inheritance to the "Other responsible person" Attributes increased "Approved body" Sub-attributes Ontology in the professional literature, each one corresponding to the professional literature, literature class model with an instance of the attribute value is the corresponding recording in the literature and Indexing items. "Language", "type", "format" and need to be standardized body from the coding information system of values to obtain, and the "field concept" property value is obtained from the field of ontology concepts. Use of "associated" attribute can also be real are examples of connections between different documents, such as a "journal paper" Real Cases can be "part" of a "journal" example; different versions of the book also can be "inherited version of" attribute to associate. Professional literature ontology can In order to better achieve different types of professional literature of information integration, and is very easy to expand the type and attributes.

b) *Semantic annotation of the professional literature*

Ontology building is completed; one must also mark out the field of ontology concepts each instance in the areas of literature and the relevance of the concept of weight, and save to the professional literature examples of "semantic annotation" property in the process shown in Figure 2 as follows:

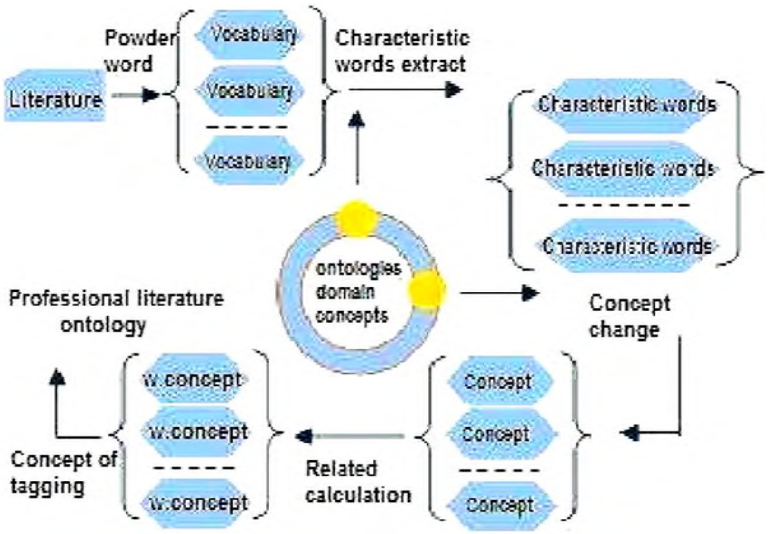


Figure 2 Semantic annotation process professional

Correlation weights in the calculation, considering the concept involved Words appear in the literature of different positions, in which documents and fields of long degree of several aspects to the concept of computational units, on the TF × IDF Operators Method [7, 18, 19] is improved, the concept of d c of the documents related to the right degree Value of w is defined as:

w = TF (c in d) · IDF (c) · field Boost (c. field in d) ·length Norm (C. field in d) (1)

Table 1 Correlation effects indicate that weight

Factors	Formula	Explain
Concept c in the literature d in the frequency of occurrence Rate	$tf(c \ \ in \ \ d) = \sqrt{frequency}$	Formula for the concept of frequency c in the literature in d frequency. Concept in the literature the frequency of a more High, then the concept of the higher correlation
Concept c in the whole Literature collection Reverse Frequency	$idf(c) = 10 + \ln \frac{numDocs}{docFreq + 1}$	num Docs the literature for the document collection Total doc Freq package for the literature collection Literature with the concept of the number of c. In a document collection, if the concept contains a The fewer the number of documents, Explain the concept of these The more relevant documents
Concept c appears in Literature in different fields The weighted value	Fields Boost(c field in d)= $\begin{cases} b1, C \text{ in Key word} \\ [b2, C \text{ in Title Keyword}] \\ b3, C \text{ in Abstracts} \\ b4, C \text{ in Other locations} \end{cases}$	The literature of a fixed set of terms b1, b2, b3, b4 are constants,respectively concept appears in the literature on different fields Correlation of factors. In general, Concept than appears in the title appears in the text Summary of important, out now surpasses Digest Important in the body of the
Concept c in which the word The length	Length Num(c field in d)= $\frac{1}{\sqrt{numfields}}$	Num Fields that c lies in the field The concept included in the total number of

Equation (1) the meaning of the various parts of the formula as shown in table i.

2) *Query processing*

Express the information needs of users, that search query is usually the form of the word. To realize the concept of matching based on semantic retrieval, on must deal with the conceptualization of the query request. Meanwhile, to increase the recall rate, also need to implement query expansion [20]. As the domain ontology contains a large number of concepts and their relations information, we use ontology to enter line a synonym expansion, semantic implication, the associated extension and semantic extension like [21]. In order to implement query expansion, first of all need to establish that the concept of inter- quantitative indicators of the degree associated the concept of similarity and relevance Matrix.

a) *The concept of similarity and the correlation matrix*

The concept of the concept of similarity that can replace compliance process level and semantic degree reflects the concept of their mutual aggregation characteristics. Conceptual similarity Is a very strong concept of subjectivity [22], from the perspective of different applications and start, the researchers made a variety of similarity calculation method [21], the considering the concept of culture between the structural level network diagram of a variety of factors, will The concept of ontology in the field of any two concepts X and Y define the similarity To:

$$Sim(x,y) = \frac{Size[P(x) \cap P(y)]}{Max\{Size(P(X)), Size(P(Y))\}} \bullet \frac{\alpha}{Dis(X,Y) + \alpha} \qquad 2$$

EQ (2) the right of the left part of the calculation of the equal sign is the language X and Y Justice degree, elements for the X and Y the number of concepts shared by the host, the denominator for the X or Y of the upper maximum number of concepts. Where, P(X) that takes study the concept of node X and all upper set, P(X) ∩ P(Y) table said the concept of the concept of X and Y by the concept of shared set of upper, Size (A) a number of elements within that set. Eq (2) the right of the right half of the equal sign calculate the semantic similarity of distance, where Dist (X, Y) for the X and Y of the semantic distance; α as regulatory factors, can approximate geographic Solution does not consider semantics in the case of coincidence degree, when the similarity is (0. 5) the semantic distance.

The correlation refers to the concept of linkages between the concepts together appears extent; reflect the characteristics of the combination between concepts. Taking into account the concept of and documentation related to the concept of degree of weight and is now in the literature of total frequency, almost Read X and Y can be defined as the correlation.

$$Col(X,Y) = \frac{\sum_{di \in D(X \cap Y)} W_{(X,di)} + W_{(Y,di)}}{\sum_{di \in D(X)} W_{(X,di)} + \sum_{di \in D(Y)} W_{(Y,di)}} \quad 3$$

Eq (3) WX,di said the concept of X in the document the correlation di Weight, WY, di , said the concept of Y in the document di correlation in weight. D(X) that all include the concept of X's document collection, D (X∩Y) that contains both X and Y are two concepts of the document collection, Eq(3) the concept of the molecular part of the said X and Y and all contain both of these two concepts the weight of the document and, to part denominator X and Y and for all with X or Y value of the document and the right. Definition of similarity between two concepts and the relevance algorithm, you can on the concept of ontology in the field of all the concepts two by two Calculate the similarity and phase Related degrees, and thus construct the concept of relevance and similarity matrix.

b) *Conversion and expansion of the semantic query*

Semantic query transformation and expansion of the query to retrieve the word form the concept of the request into the form of semantic information, and in accordance with the rules expand the concept. The process shown in Figure 3, the following steps, Method using word queries cut into several phrases, and with leading Extraction of domain ontology concepts related to the characteristics of the field of vocabulary.

According to the concept of ontology in the field of natural words the concept of identity, will feature words Meeting of the concept extracted, similarity threshold were set Rs (Rs∈[0, 1]) and correlation threshold Rc (Rc ∈[0, 1]), the control concept of similarity / correlation matrix, the domain concepts Ontology query request with the concept of similarity or correlation is greater than Rs greater than the concept of Rc extract, added to the original query, in order to achieve the requested query language Meaning extension. The query request cannot be converted into the concept of characteristic words are still words retain the form of exchange, to participate in subsequent retrieval.

III. Search Implementation

The implementation query request and quantification of the document collection and calculation. For the document to be retrieved set D in any text file dj ∈D, it can be expressed as s + t Dimensional vector of the form

$$d_j = (w_{1j}, w_{2j}, \cdots, w_{sj}, w'_{1j}, w'_{2j}, \cdots, w'_{tj}) \quad 4$$

Document vector by the concepts and vocabulary of two parts, Eq (4), s the number of concepts for the information system, t in the information system without a counterpart in the number of index terms. Wij for the concept ci of the document dj's relevance weights in on the text Offer examples for calculating the semantic annotation wij 'is the i-word ti on the document dj's relevance weights, using conventional TF × IDF method Obtained. For the expansion and transformation of the query request, it could be further form Show for the query vector.

$$q = (w_{1q}, w_{2q}, \cdots, w_{sq}, w'_{1q}, w'_{2q}, \cdots, w'_{tq}) \quad 5$$

Where s the number of concepts for the information system, t in the information system without the corresponding, the index number of the concept of the word. Vector component wiq i-a concept that ci In the query q In weight, and there wiq ≥ 0. wiq 'said the first i- Glossary ti in the query q in weight, and there wiq '≥ 0 Queries in the document and that the basis to quantify, document and The degree of similarity between the query can be transformed into two-dimensional to s + t Amount of similarity, and according to cosine law [19] to calculate the specifi Value.

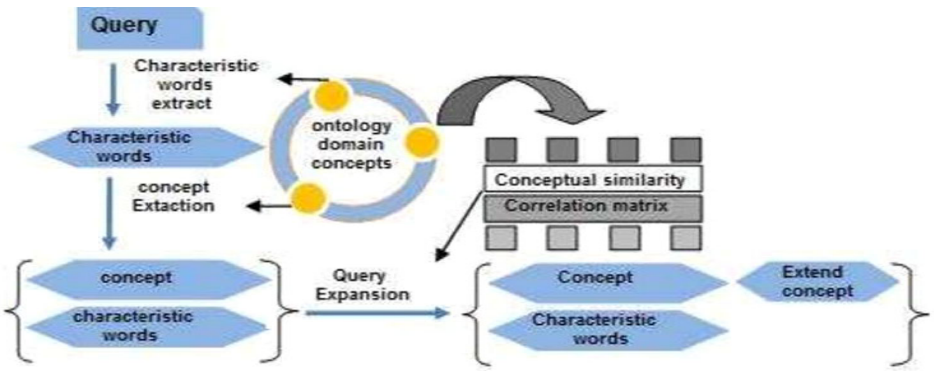


Figure 3 Semantic query transformation and growth process

IV. Evaluating IR systems

The methods are needed so as to be able to compare their abilities. In this paper you have two very different aspects of an IR system can be measured: efficiency, and effectiveness. Efficiency can be measured in terms of the resources required by the system, including the storage space required to store the document collection, and the computing resources needed to perform operations on the collection, such as the addition and removal of documents, and performing queries. Effectiveness attempts to measure, as the name implies, the effectiveness of an IR system at satisfying a set of queries. Given a sufficiently general document and query collection, the effectiveness should provide a domain neutral measure of the ability of the system. The measure of effectiveness is further complicated by the fact that it is dependent on the type of task being evaluated. Interactive systems must be evaluated in a different way to systems in which user feedback plays no role, Some of the issues have been discussed in the first international workshop on adaptive information retrieval (Joho et al.,2008) organized by the guest editors. They are Many IR researchers [Rijsbergen 2006] believe that a satisfactory approach to the evaluation of an information retrieval system is yet to be found. Since this is still a rich and ongoing area of research, we will only examine the most common evaluation methods. Furthermore, it is assumed that the system under evaluation operates with minimal user interaction. The most widespread method of evaluating an IR system involves providing precision-recall values for a set of queries posed on a specific document collection. Usually, a precision-recall diagram is plotted so that comparisons between IR systems can be made visually Use the document collection and a query, define R the number of relevant documents in the set for this query, and A be the number of documents retrieved by the IR system. Finally, let I be the number relevant documents within the documents retrieved by the IR system. Recall and precision can then be defined as

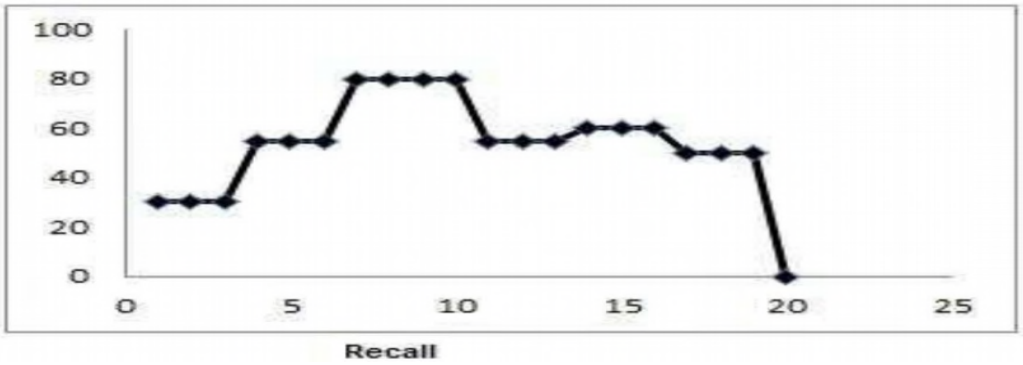
$$RECAL = \frac{I}{R}$$

$$PRECISION = \frac{I}{P}$$

The basic precision and recall measures assume that the IR system returns an unsorted list of results, which is then evaluated in full. If this is not the case, recall and precision values change as more documents within the result list are examined (the assumption is made that the list is ordered from the highest to least believed relevance). This is done by computing precision over the seen documents whenever a relevant document is found in the retrieved document list. For example, assume the following documents are retrieved in response to a query, with the documents marked by an asterisk indicating relevant documents:-

Rank	Document number	Relevant	Rank	Document number	Relevant
1	*dnu*19		11	*dnu* 77	*
2	*dnu* 40	*	12	*dnu* 53	
3	*dnu* 28		13	*dnu* 69	*
4	*dnu* 35		14	*dnu* 94	
5	*dnu* 3		15	*dnu* 100	*
6	*dnu* 9	*	16	*dnu* 65	
7	*dnu* 67		17	*dnu* 89	*
8	*dnu* 71	*	18	*dnu* 75	
9	*dnu* 43		19	*dnu* 13	*
10	*dnu* 36	*	20	*dnu* 5	

This paper is you assume that this query has 20 relevant documents, of which nine were retrieved as above. At a recall level of 6% (i.e.1 out of the 20 relevant documents have been seen), precision is 50%, since one out of the three seen documents are relevant. At the 60% recall level, precision increases to 50%. At 60% and 80% recall, precision values are 50% and 40%. Finally, at 90% recall, precision drops to 0%, as all relevant documents were not retrieved Precision-recall curves are normally drawn by computing precision at 20 standard recall values, namely, 0%, 5%, 10%... and 25%. If, as in the above example, insufficient relevant documents exist to compute recall at all these points, the values at the standard points are set as the maximum known precision at any known recall points between the current and next standard points. The example presented above would therefore yield the precision recall curve illustrated in Figure 4. Usually, precision-recall curves are computed by averaging the precisions obtained at the standard recall values over all queries posed to the system.



When averaged over a number of queries, precision-recall curves tend to follow an exponential decay curve. Intuitively this result from the fact that an algorithm would generally rank at least a few relevant documents quite highly, therefore yielding high precision for low recall values. As the number of relevant documents returned by the system increases however, more and more irrelevant documents are returned within the results. Obviously, a perfect algorithm would have a float precision-recall curve at the 100% level.

V. Conclusion

This paper analyzes the traditional information retrieval methods inadequate for achieving the semantic information system are described in four key modules, given A new ontology construction method based on this proposed ontology-based information system framework for information retrieval information systems, and describes the information system prototype Design and retrieval processes, effective solution to the traditional information retrieval recall and precision for low rate of problems; In addition, the paper also set up a A 100 or so computers in the field HTML format paper documents as a basis of knowledge of their search, by comparing the experimental results show that Ontology-based information retrieval compared to traditional information retrieval, it expanded the connotation and extension of search terms, making retrieval recall rate significantly Strengthened. Of course, the proposed ontology-based information retrieval information system prototype box leave many shortcomings, such as ontology and knowledge library of data increase, making information retrieval response time when the start search will be some effect, how to reduce the response time will be my they next step is to focus on the object of study.

VI. Acknowlegements

This work is supported by National Natural Science Foundation of China under grant No.60970096 and 90818004, the Postdoctoral Science Foundation of China under Grant No.20080440988 and the Natural Science Foundation of Hunan Province, China Under Grant No.09JJ403

VII. References

1) GÀbor NagypÀl. Imp roving Information Retrieval Effectiveness by Using Domain Knowledge Stored in Ontologies[C]. In: Proceedings of OTM Confederated International Conferences, LNCS3762. Ber2 lin: Sp ringer - Verlag, 2005: 780 - 789.
2) Zhang, L., Yu, Y., & Yang, Y. (2005). An enhanced Model for Searching in Semantic Portals. WWW 2005, May 10-14, Chiba, Japan
3) .A.Varzi, Basic problems of mereotopology, in: N. Guarino (Ed.), Formal Ontology in Information Systems, IOS Press,Amsterdam, 1998
4) Paralic J, Kostial I. Ontology - based Information Retrieval[C] . In: Proceedings of the 14 th International Conference on Inform ation and Intelligent information System s, Croatia. 2003: 23 - 28.
5) Tomassen SL. Searchingwith Document Space Adap ted Ontologies [C]. In: Proceedings of Em erging Technologies and Inform ation information System s for the Know ledge Society, Athens, Greece. Berlin: Sp ringer - Verlag, 2008: 513 - 522.
6) Vallet D, FernándezM, Castells P. An Ontology - based Information RetrievalModel[C]. In: Proceedings of the Sem anticW eb: Re2 search and Applications. 2005: 455 - 470.
7) Motiee S, Nematzadeh A, ShamsfardM. A Hybrid OntologyBased App roach for Ranking Documents [C]. In: Proceedings of World Academ y of Science, Engineering and Technology, Prague. Czech Republic: World Acad Sci, Eng & Tech - waset, 2006: 55 - 60.
8) Li, Y., & Bontcheva, K. (2007). Hierarchical, perceptron-like learning for ontology-based information extraction. In Proceedings of the 16th international conference on World Wide Web (pp. 777-786). Banff, Alberta, Canada.
9) H. Ding, I. T. Solvberg,and Y. Lin, “A Vision on Semantic Retrieval in P2P Network”, in Proceedings of the 18th International Conference on Advanced Information

Networking and Application, 2004,pp.177-182
10) Kim, S.-B., Seo, H.-C., & Rim, H.-C. (2004). Information retrieval using word senses: Root sense tagging approach. Proceedings of the 27th annual international ACM SIGIR conference on research and development in information retrieval (pp. 258-265). Sheffield, UK
11) Hahn U. Turning Informal Thesauri into Formal Ontologies: A Feasibility Study on Biomedical Knowledge Re - use[J]. Com parative and Functional Genom ics, 2003, 4 (1) : 94 - 97.
12) Sean Bechhofer , Carole Goble, Thesaurus construction through knowledge representation, Data & Knowledge Engineering, v.37 n.1, p.25-45,April2001 [doi>10.1016/S0169-023X (00)00052-5]
13) Hepp M, De Bruijn J. GenTax: A Generic Methodology for Deriving OWL and RDF – S Ontologies from Hierarchical Classifications, Thesauri, and Inconsistent Taxonomies[C]. In: Proceedings of 1st International Conference onW eb Reasoning and Rule information System s,Innsbruck, Austria. 2007: 103 - 117.
14) Dublin CoreMetadata Element Set, Version 1. 1 [EB /OL]. [2008- 05 - 15]. http: / / dublincore. org/documents/dces/.
15) Li Peng, Tao Lan,"Measuring semantic similarity in ontology", Computer Engineering and Applications,vol.28, no.1,Jan.2007,pp227-229.
16) Lucene 2. 3. 1 Core AP I[EB /OL]. [2008 – 05 - 06]. http: / / lucene. apache. org/ java /2_3_1 / ap i/ core / index. html.
17) Bhogal J, Macfarlane A, Smith P. A Review of Ontology Based Query Expansion [J]. Inform ation Processing and M anagem ent,2007, 43 (4) : 866 - 886.
18) L. Dey, S. Singh, R. Rai and S. Gupta, Ontology aided query expansion for retrieving relevant texts, Proceedings of Advances in Web Intelligence, Third International Atlantic Web Intelligence Conference, Lodz, Poland (2005), pp. 126-132.
19) Couto, F. M., Silva, M. J., and Coutinho, P. M. (2007). Measuring semantic similarity between gene ontology terms. Data Knowl. Eng., 61(1), 137-152.
20) Formica A. Ontology - based Concept Similarity in Formal Concept Analysis [J]. Inform ation Sciences, 2006, 176 (18) : 2624 -2641.1987.
21) Zobel J. How Reliable are the Results of Large - scale Information Retrieval Experiments[C]. In: Proceedings of the 21st Annual International ACM S IGIR Conference on Research and Development in Information Retrieval. Australia: Melbourne. 1998: 307 - 314.
22) Zobel J. How Reliable are the Results of Large - scale Information Retrieval Experiments. In: Proceedings of the 21st Annual International ACM S IGIR Conference on Research and Development in Information Retrieval.. 1998: 307- 314.

Software Risk Management- An Integrated Approach

GJCST Classification
B.7, K.6.1

Mr. N. Sathish Kumar[1], Mr. A. Vinay Sagar[2], Mr. Y. Sudheer[3]

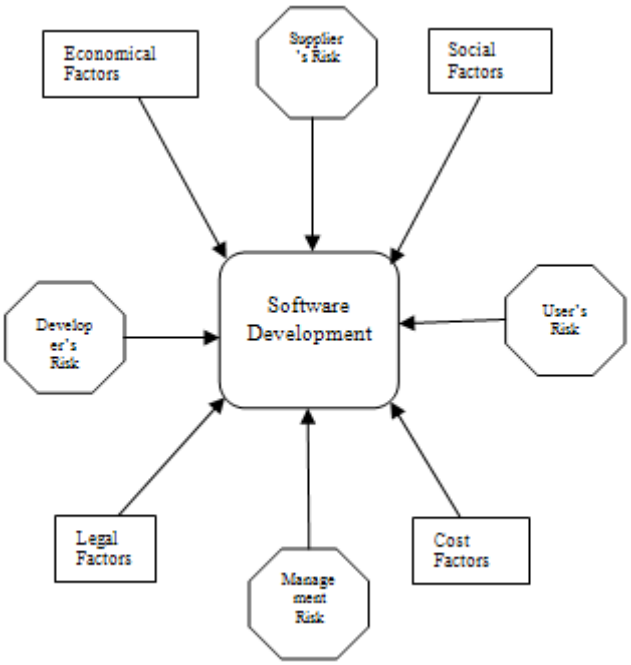
***Abstract*-The risk management is an integrated approach, it is essential for every project i.e. from small size to large size projects. In this paper, we present a new approach for managing the risk in easy and effective manner. This approach is for minimizing the confusion and loss associated with a project. It is performed continually over the life of a program and it can be achieved in several phases which include Risk Identification, Risk Priority, Risk Management Plan, Risk Monitoring and Internal Audit. The risk management plan includes product size, business impact, customer, process, technology, development environment, staffing (size and experience), schedule, and cost.**

I. INTRODUCTION

Software Risk Management is an integrated approach for minimizing the uncertainty and potential loss associated with a project by providing insights to support informed decision making. It is performed continually over the life of a program, from initiation to retirement. Some categories of risk include product size, business impact, customer-related, process, technology, development environment, staffing (size and experience), schedule, and cost. A risk is an event or condition that, if it occurs, has a positive or negative effect on a project's objectives. The three common characteristics of risk are (1) it represents a future event, (2) it has a probability of occurring of greater than 0%, but less than 100%, and (3) the consequence of the risk must be unexpected or unplanned for. Future events can be categorized as positive risk if their consequences are favorable or as negative risk if their consequences are unfavorable. Risk Management concentrates on performing bottom-up, detailed, continuous assessment of risk and opportunity. It focuses on addressing the day-to-day operational risks that a program faces. Risk Management follows a two-stage, repeatable process of assessment i.e., the identification, estimation and evaluation of the risks confronting a program and management i.e., the planning for, monitoring of, and controlling of the means to eliminate or reduce the likelihood or consequences of the risks discovered. It is performed continually over the life of a program.There are a variety of risks that confront the global software industry.

About[1] - Associate Professor, Department of CSE, SVS Institute of Technology,Bheemaram,Hanamkonda, Wranga-50600l, A.P., INDIA [2] satish4info@gmail.com
About[2] - Assistant Professor, Department of CSE, SVS Institute of Technology Bheemaram,Hanamkonda, Wranga-50600l, A.P., INDIA anchurivinay@gmail.com
About[3]- Assistant Professor, Department of CSE, SVS Institute of Technology Bheemaram,Hanamkonda, Warangal-50600l, A.P., INDIA sudheer.yethipathi@gmail.com

The characteristics of the legal, social, economic and competitive environments impose constraints and opportunities that help to define the nature of the risks for suppliers, buyers, and other stakeholders in the software acquisition and development process. Enterprise Risk Management (Corporate boards, CEOs, CFOs and other members) and Operational risk management (Focus on integrated governance, risk and compliance, regulatory measures) are the existing systems,
in which there are many problems are occurred.



II. THE MOST COMMON & SERIOUS SOFTWARE RISKS

There are various reasons as to why earlier risk management is difficult to implement effectively. We observed that few projects have more than 20 active risk factors at any one time. Another reason for the relatively low implementation of earlier risk management methods in practice that the fact that risk is a fuzzy concept for which users lack the necessary tools to more accurately define risk for a deeper analysis. Users may not have the ability to provide accurate estimates for probability and loss projections required for a reliable risk analysis. Existing risk management methods may not provide support for dealing with these differences.

Risks may also affect a project in more than one way and most risk management approaches focus on cost, schedule or quality risks, but there may be combinations of risks or other characteristics such as future required maintenance, company reputation, or potential liability that should be considered important in decision-making process. Finally, many current risk management techniques may be perceived as too costly or too complex to use. Simple, straightforward risk management techniques that require an acceptable amount of time to produce results might be the answer.A key software risk area, and one that has an immediate and potentially catastrophic impact on software projects, is related to the development of software. Software security policies do not fully address the risk of using foreign suppliers to develop weapon system software. Current guidance allows program managers discretion in managing foreign involvement in software development, without requiring them to identify such risks.

III. RISK MANAGEMENT PROCESS

An endless list of all possible risks that need to be considered in a software development process. There are any numbers of internal/external scenarios that may influence the possible risks that our software on is susceptible. The key to successful risk management lies in the ability to incorporate a new risk management process that addresses the complementary needs of the business and its customers.Risks can be identified at the very earliest stages of the software life-cycle. The ability to identify risks earlier translates into earlier risk removal, at less cost, which promotes higher project success probability.Software project risk defines the operational, organizational and contractual software development parameters. Software process risk includes both management and technical work procedures. In management procedures, process risk may be found in activities such as planning, staffing, tracking, quality assurance and configuration management. In technical processes, it may be found in activities such as requirements analysis, design, code and test. Planning is the management process risk that is most often reported. Software product risk contains intermediate and final work product characteristics. Primarily a technical responsibility, product risk may be found in requirements stability, design performance, code complexity and test specifications. Product risk is difficult to manage because software requirements are often perceived as flexible. Requirements are the most significant product risks in risk assessments. A software risk can be classified in the following figure. In general, a risk management process should consist of the following activities, acquisition, supply, development, operation and maintenance of software products and services

- Early planning for risk management
- Implement the Risk Management
- Managing the project risk profile
- Advanced risk analysis technique
- Risk monitoring
- Risk treatment
- Effective Evaluation

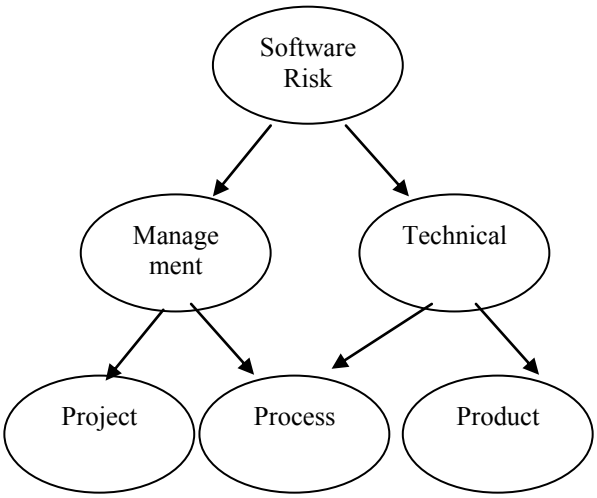


Figure2 : Software Risk classifications

IV. RISK MANAGEMENT TOOLS

1) *Risk Identification*

The Universal Risk Project identifies two basic types of risk statements that can be used to identify whether a set of circumstances represents a risk to the project:"IF technology is not available, Then we will not meet the requirement" and "IF we cannot hire sufficient qualified software engineers, Then We cannot meet the planned development scheduleA tool that can be used to aid in the identification of risks on a software project are strategically defined software measures. In this we proposes the use of a software measurement process as a integrated approach for identifying risks before they become problems.

- Software Component Status
- Staff Experience
- Reviews Completed
- Problem Report Status
- Resource Utilization

2) *Risk Analysis*

After risks are identified, they should be partitioned into categories such as technical, cost, schedule, management, etc. Note that some risks may fall into multiple categories.Why do risks need to be partitioned? First of all, some risks are more important than others. Also, different stakeholders may be concerned about different risks, or different personnel may bear responsibility for tracking/monitoring different risks. Finally, different risk types may require different mitigation strategies.

The initial activity in risk analysis is to identify contributing factors, then establish a hierarchy of those contributing factors. It illustrates a hierarchy of how a project might fail, given the contributing factors of Staffing, Funding, Performance Failures, and so on. The Staffing factor is further broken down to show, first, how staffing may

become a contributing factor to project failure and second, what the contributing factors might be that result in insufficient staffing (subsequently leading to project failure). All contributing factors defined within the hierarchy would be broken down to a correspondingly meaningful level of detail.Similarly, for positive risk, a hierarchy of contributing factors could also be created, this time highlighting those elements for which risk is being undertaken in order to leverage a perceived opportunity for the project, such as "Schedule Completion Will Be Early".There are any number of ways that risk can be partitioned, analyzed and quantified. The approach taken and method(s) used should always be tailored to meet the needs of the business, the customer and the project. If you reference the Tools and Methods section of this document, you can review several different MS Excel-based tools developed by the DACS from information available in the literature, each having its own partitioning scheme and offering a different approach to quantifying project risk.Once you have successfully partitioned and quantified software risk, prioritization of risk becomes the next logical activity.

3) *Risk Priority*

Risk priority is a critical characteristic of the formal risk management process, as it provides the opportunity to apply what are typically limited project resources to those risks having the largest potential impact on the project. For many risk priority approaches, risks are ranked and priority based on some combination of probability. This can be done qualitatively in a risk priority matrix using some type of composite probability-impact score. Note that the qualitative matrix has some basis in quantitative values for frequency and impact, although these can be subjectively defined and tailored to suit specific sets of circumstances. The highest priority risks would be those falling in the red region. The quantitative example of risk priority highlights the combined probability-impact scores for specific elements of the project. Of course, these are very high-level risks to try to deal with effectively. The highest priority risk is schedule failure, followed by technical failure, cost failure, etc. The MS Excel-based tools provided in the Tools and Methods section show how high-level risk areas can be further broken down to quantify and prioritize very specific elements of risk.The combined risk exposure number associated with performing regression testing indicates that it should be considered a much lower priority than if regression testing is not going to be performed at all based on the quantitative probability and loss associated with each potential undesired outcome.

4) *Risk Management Plan*

The risk management plan includes product size, business impact, customer, process, technology, development environment, staffing (size and experience), schedule, and cost. Risk management planning provides the basis for the identification of the monitoring procedures that should be put in place for each risk, including how to tell if a risk is going to manifest as a real problem, and how frequently each identified risk should be monitored. Risk planning also takes into account risk aversion planning i.e., what actions will be taken to mitigate risk before it occurs and contingency planning i.e., how to react if a risk actually manifests.The document "Risk Management Survey of Department of the Navy Programs" contains samples of Risk Management SOW Contract Clauses, a Risk Status Report format, a product/process combination risk management process, a technical risk assessment form, a technical risk status form, a risk management process and analysis methodology, an event-driven risk mitigation report, and risk management survey references.

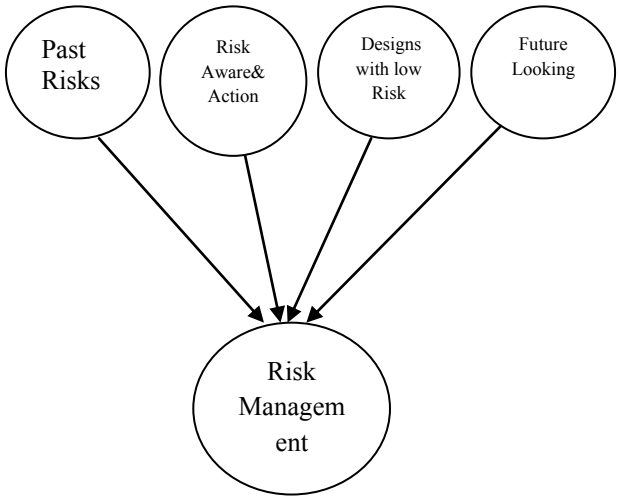
5) *Risk Monitoring*

There are a number of measurements/metrics and tools that can be used to monitor and report status of project risks. Following are the some of the basic measurements and metrics that can be considered to meet these needs.

- Number of Identified Risks
- Number of Active Risks
- Number of Risks Assessed "High", "Medium" and "Low"
- Expected Value (Probability x Cost)
- Overall Project Criticality Index

The number of identified and active risks provides a picture of how many risks have been mitigated, and how many remain as open. Prioritized risks can be categorized as the number considered "high", "medium" and "low". Combining risk probability and impact into a total composite score is another way to identify, track and monitor high priority risks. Identifying an "average" probability impact (PI) score subdivides the risk population into "greater than" and "less than" average risks, providing a very general guideline on which risks to address first. Another way to look at impact is risk probability times cost impact, rather than consequence or severity impact, in order to calculate the expected value of each potential risk, A useful metric for assessing project risk is the percent likelihood of meeting the target schedule and/or budget. Any significant negative deviation from this metric would not well for the success of the project.There are also a number of risk management audit measurements/metrics that are appropriate to monitor and track progress on controlling project risk. These include factors such as (1) the number and ratio of scheduled and actual risk management audits, (2) the effort required to support risk management audits, (3) the total number of risk-related problem reports (and the number of open and closed reports) that are measured in each risk management audit, and (4) the number and ratio of actual and deferred corrections that are reported in each risk management audit. This number of corrections can be further broken down into categories of major and minor corrections and, as with the number of actual corrections, can be assessed against the amount of effort required to implement them.One of the better tools to use in the risk management process is Risk Status Reports.

They represent a best practice in that they allow the communication of appropriate risk metrics to all stakeholders on the project. Risk Status Reports can include such things as a listing of the top ten risk items, the number of risks resolved as of the report date, the number of new risk items introduced since the last report, the total number of current unresolved risk items, the presence of unresolved risk items that lie on the project critical path, and assessment of the probable cost of unresolved risk in comparison to the amount of risk reserve. A second useful tool for monitoring risk and reporting risk status is the Event-Driven Management Report. This tool can be represented by a generic waterfall chart, where risks are identified and documented on an Integrated Product Team watch list. One general caution or observation regarding the Event-Driven Management Report, Be aware that risks may not always follow the waterfall, i.e., high- or medium-level risks could conceivable pop up anywhere in the life cycle, depending on how effective the overall risk management process is. Event-driven risks will not necessarily conveniently decrease in risk level as the life cycle evolves.



6) *Internal Audit*

A software solution that includes an audit component is a paperless audit management system proven to reduce the time and cost associated with internal audit processes. It is better for mid-market and large enterprise organizations to improve audit efficiency and productivity of the entire audit process including risk assessment, planning, scheduling, preparation, review, report generation, global issue tracking and administration.

- Analyst Reports
- Events Calendar
- Early Updates
- Success Stories
- Surveys

Effective implementation of risk management is based on the following set of benefits resulting from the process:

- The risks that could impact project success are identified
- A priority order in which risks should be addressed is established
- Mitigation alternatives appropriate for each potential problem are carefully considered based on project circumstances
- Optimized mitigation techniques for all risks above their thresholds are selected
- Contingency plans in case the risk mitigates are developed proactively, rather than as a result of fire-fighting
- Information to improve risk management policies is captured, analyzed and acted upon
- Risk management processes / procedures are systematically and periodically reviewed and improved to further reduce risk

7) *Safe Tracking And Implementation*

Results of the risk reduction implementation must be tracked. The tracking step involves gathering data, compiling that data into information and then reporting and analyzing that information. This includes measuring known risks and monitoring triggers as well as measuring the impacts of risk reduction activities. This includes

- Identify the new risks and add them to existing risk list.
- Information needed for additional requirements planning
- Implementing the plan

The implementation starts with the requirement specification and continues till release of software project. The major objective of this phase is to maximize the utilization and effectiveness of resources in time and within project budget in order to carry out the various activities during development.

V. CONCLUSION

Software risk management is the integrated approach in which risk factors are systematically identified, assessed, and implemented. The determination of the risk in a software either due to external or internal causes in a major part of software management. The process of risk analysis is continuous and applies to many different levels, at once identifying system level, assigning probability impact. For a software manager it is responsibility to manage the software risk in effective manner. The risk management plan is implemented as a part of initial project planning and utilized throughout all of the phases of the software development process. Risk management requires fear free environment where risks can be identified and discussed by conducting meetings and periodical reviews. We can reduce confusion,

time, development and management cost by using an integrated approach.

VI. REFERENCES

1) *Software Risk Mangament* by Barry W. boehm, Les Alamitos, CA, IEEE Computer Society 1989,
2) *Risk management for Software Projects*, Alex Down, Michael Coleman, Peter Absolon, Tata McGraw-Hill Book Company, 1994
3) *Practical Risk Assessment for Project Management*, by Stephen Grey, John Wiley&Sons, 1995.
4) Software Engineering Risk Management: Finding Your Path Through the Jungle, Version 1.0, Dale Karolak, IEEE Computer Society.
5) *Risk Management: Concepts and Guidance, 2nd edition*, by Carl Pritchard, ESI International, 2001
6) *Project and Program Risk Management: A Guide to Managing Project Risks and Opportunities*, Max Wideman, 1998.
7) *Making Hard Decisions with Decision Tools*, by Robert T. Clemen and Terence Reilly, Duxbury Press, 2001.
8) *Project Risk Management: Processes, Techniques and Insights*, by Chris Chapman and Stephen Ward, John Wiley & Sons, 1997.
9) *Stop IT project failures through risk management*, by Dan Remenyi, Butterworth Heinemann, 1999.
10) *Managing Risk: Critical issue for survival and success into the 21st century*, Alan Waring and A. Ian Glendon, Thompson Learning, 1998.
11) *Managing Risk: Methods for Software Systems Development*, Elaine Hall, Addison Wesley.
12) 12.*http://acc.dau.mil/simplify/ev.php?ID=1203_201&ID2=DO_TOPIC*

Integration of ICT in Teacher Education "Computer Assisted Instruction & E-Learning"

Mr. Vimal Kumar

GJCST Classification K.3.1

I. INTRODUCTION

Education is the life long process of acquiring new knowledge and skill through formal exposure of information, ideas and experiences. These can be done in the schools by way of systematic planning of instruction. In turn it needs proper method or technology to adopt in teaching the concepts of the subjects in the school. Now a day they call it as a educational technology which implies a behavioral science approach to teaching and learning, in trait, it makes use and pertinent, scientific and technological sociology, communicating linguistics and their related fields.Educational technology has grown as a resut of technological devices in the use of practies with the explored psychological of teaching, learning and behavioral modification. There are many means by which effective instruction can be imparted in the classroom. The use of Computer Assisted Instruction (CAI) in schools gaining momentum. Now days, more and more schools are having computers, so use of technology enhances effectiveness of a learning experience.

II. NEED AND SCOPE OF THE STUDY

The investigator was interested in studying the effectiveness of computer-assisted instruction in teaching General Science at secondary level. The subjects on science can develop ability and achievement among the XII students. Hence it is necessary to study the achievement of the students of the students in relation to computer-assisted instruction adopted in classroom teaching.

1) *Objectives of the study*

- To find out the effectiveness of teaching **General Science** for XII through conventional method.
- To find out the effectiveness of teaching **General Science** for XII through computer assisted instruction.
- To find out the effectiveness of teaching **General Science** for XII through computer assisted instruction over conventional method.
- To prepare a computer assisted instructional package on **"universe".**

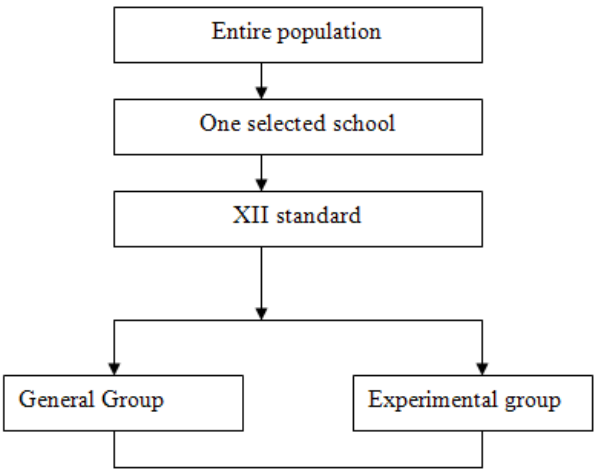
About-Assistant Registrar, Baba Farid Group of Institutions, Bathinda University Address: Research Scholar, Jodhpur National University, Jodhpur. E-mail: vimalkhurana2002@yahoo.com

2) *Hypotheses*

- There is no significant difference between mean gain scores of experimental and control group on pre – test.
- There is no significant difference between the mean gain scores of pre and post – test of control group.
- There is no significant difference between the mean gain score of pre and post – test of experimental group.
- There is no significant difference between the mean gain scores of post – test in general and experimental group.

III. SAMPLING

The present experimental study involved a parallel or equated group experimentation, which is more complete and accurate than the one group experimentation.The independent variable in the present study was CAI and the dependent variable was student achievement in terms of gain scores



In this study 32 students studying in XII of D.A.V Sen.Sec.School, Malout were taken as the sample. Experimental group consisted of 16 students who were taught "The Universe" by the CAI. The control group of another 16 students who were taught the same content by conventional method.

1) *Procedure in framing equated groups*

The sample of 32 students was divided into two equated groups of 16 students each. Two groups were equated as

nearly as possible in terms of their achievement scores of sciences in second semester examination. Students having similar range of marks in the second semester examination were allotted equally and randomly for experimental and general groups. To find out that there was no significance difference between the two groups, 't' test was applied to the scores of second semester examination.

2) *Construction of tools (Ready-made)*

Standardized Tool were use in this study prepared by JAGANATH.K. DANGE and SHAIK ABDUL WAHAB. (Research scholar), Lecturer in P.G. Studies in Education, Kuvempu University, Shankarghatta (Karnataka)

3) *Development of CAI*

The content is divided into different tasks and the tasks are presented in the form of Slides in Micro-Soft Power Point. After completion of every task a question will be there to test students learning. In the presentation needed pictures, background music and animations are added as special effects in the CAI.

IV. Data Analysis

The data collected by the investigator are analyzed and interpreted as given below.

TABLE 1
Test of significance of the scores obtained by experimental and controls group in pre – test.

Groups	N	M	S.D	't' value	Significant level at 0.05
Experimental group	16	13.56	528	071	NOT SIGNIFICANT
Control group	16	13.31	524		

According to the above table the obtained 't' value 071 is statistically no significant, because it is lesser than the critical value 2.13 for 30 df. At 0.05 level of significant from this, it can be concluded that there is no significant difference between two groups and a null hypothesis is accepted.

TABLE 2
Test of significance of the scores obtained by control group student in their pre and post – test.

Test	N	M	S.D	't' value	Significant level at 0.05
Pre-test	16	13.31	5.24	1	NOT SIGNIFICANT
Post-test	16	13.56	5.28		

The above table shows that the obtained‚t' valued 1 statistically not significant because it is lesser than the critical value 2.13 for 30 df. at 0.05 level of significant. From this it can be concluded that there is no significance difference between the mean gain scores of pre and post – test of control group.

TABLE 3
Test of significance of the scores obtained by experimental group students in their pre and post – test.

Test	N	M	S.D	't' value	Significant level at 0.05
Pre-test	16	13.56	5.28	8.47	SIGNIFICANT
Post-test	16	13.43	5.48		

According to the above table the obtained „t' value 8.47 is statistically significant, because it is greater than the critical 't' value 2.13 for 30 df. at 0.05 level of significant. From this it can be concluded that there is significance difference between the mean gain scores of pre and post – test of experimental group.

TABLE 4
Test of significance of the scores obtained by experimental group students in there post – test.

Groups	N	M	S.D	't' value	Significant level at 0.05
Experimental group	16	15.43	5.48	5.77	SIGNIFICANT
Control group	16	13.25	5.53		

The above table shows that the obtained„t' value 5.77 is statistically significant, because it is greater than the critical't' value 2.13 for 30 df. At 0.05 level of significant. From this, it can be concluded that the students taught by computer assisted instruction method performed well, than

the students taught by conventional method in learning the concepts of universe.

Findings

- There is no significant differences between mean gain scores of experimental and general group of pre –test.
- There is no significance difference between mean gain scores of pre test and post – test of general group.
- There is significance difference between mean gain scores of pre test and post – test of experimental group.
- There is significance difference between mean gain scores of post – test of general group and experimental group.

V. Education implications

CAI may be introduced in all the standards in the schools, as the study found favorable result and the students found to be interested to learn through CAI. Hence all the teachers may be given orientation about CAI and its effect on the achievement of among the students. The teacher must be given sufficient training and encouragement in preparing the CAI module. Govt. May distribute CAI packages of all subjects to all the schools that hey can use it in their daily teaching learning process.

VI. Reference

1) Aggarwal. J.C.,(1995), ”Essentials of educational technology: Teaching learning innovations in education” , Vikas Publiching House New Delhi.
2) Buch, M.B., (1998), “Fourth survey of research in education”, volume –II, NCERT, New Delhi.
3) Gupta, S.K., (1969), “Technology of science education, Vikas Publishing House Pvt. Ltd., New Delhi.
4) Sharma, S.R., (1998), “Educational technology”, Mohit Publication, New Delhi.
5) www.ICTEDU.Org
6) www.Indianedu.Com

An Analysis of Different Resampling Methods in Coimbatore, District

Dr.S. Santhosh Baboo[1], M.Renuka Devi[2]

GJCST Classification
I.4.6, I.3

***Abstract*-Image pre-processing of satellite Imagery. In order to actually geometrically correct the original distorted image, a procedure called resampling is used to determine the digital values to place in the new pixel locations of the corrected output image. The resampling Process calculate the new pixel values from the original digital pixel values in the uncorrected image. When remote sensing has been used to create an image, it needs to undergo some form of validation procedure using observational and/or sampling techniques. Failure to do so will reduce the confidence in the final product.This study describes the three methods applied in Coimbatore district. This study is used to evaluate the three resampling methods and how they are vary from their pixel calculation and accuracy.**

*Keywords:*Resampling,Satteliite imagery,pixel location

I. INTRODUCTION

When the first satellite, Sputnik, was launched in 1957 no one could have foreseen how its diverse its use would become. Today, we have Direct TV, On-Star, XM Radio and live up-to-the-second television coverage from every corner of the world. Today, satellite information is being relayed back to earth every second of every day. Before Sputnik had completed it first orbit it had relayed the first data back to earth. And it was not the "oldies" station on XM Radio. It was environmental data. More than forty years later, the use of satellite imaging continues as the most popular provider of environmental monitoring. With recent demands for new levels of data we are presented with the problem of how to manipulate our new raw satellite images so that these images can be integrated with pre-existing environmental observations and methods.In order to retrieve, manipulate and process raw satellite images we make use of commercial computer software, in particular ERDAS Imagine for Visualizing Images. ERDAS Imagine is used for data visualization and analysis of satellite images. With a full understanding the use of key components of the ERDAS, we are able to customize, compose and modify algorithms. This allows us to prompt and direct ERDAS to meet our specific needs and tailor, to our needs, the processing of the satellite data.Once the raw remote sensing digital data has been acquired, it is then processed into usable information. Analog film photographs are chemically processed in a darkroom whereas digital images are processed within a computer. Processing digital data involves changing the data to correct for certain types of distortions. Whenever data is changed to correct for one type of distortion, the possibility of the creating another type of distortion exists. The changes made to remote sensing data involve two major operations: *preprocessing* and *postprocessing*.The preprocessing steps of a remotely sensed image generally are performed before the postprocessing enhancement, extraction and analysis of information from the image. Typically, it will be the data provider who will preprocess the image data before delivery of the data to the customer or user. Preprocessing of image data often will include *radiometric correction* and *geometric correction*. lGeometric corrections are made to correct the inaccuracy between the location coordinates of the picture elements in the image data, and the actual location coordinates on the ground. Several types of geometric corrections include system, precision, and terrain corrections. Geometric correction contains two methods that is Parametric and Non-Parametric. Non-parametric method establishes mathematical relationships (mapping polynomials) between the coordinates of pixels in an image and the corresponding coordinates of those points on the ground (via a map). Two steps are involved in non-parametric corrections 1. Rectification 2. Resampling. Step one used to Calculate new output pixel locations (X, Y) and Relate image location to map location using a "mapping polynomial" function.Step two involves in the process of extrapolating data values to a new grid. Resampling is the step in rectifying an image that calculates pixel values from the original data grid. This also involves in determination of DN values to fill in the output matrix of the rectified or registered image.There are three methods in resampling:

- Nearest Neighbour
- Bilinear Interpolation
- Cubic convolution

II. NEAREST NEIGHBOUR

Nearest neighbor is a resampling method used in remote sensing. This resampling uses the digital value from the pixel in the original image, which is nearest to the new pixel location in the corrected image. This is the simplest method and does not alter the original values, but may result in some pixel values being duplicated while others are lost. This method also tends to result in a disjointed or blocky image appearance.The approach assigns a value to each "corrected" pixel from the nearest "uncorrected" pixel. The advantages of nearest neighbor include simplicity and the ability to preserve original values in the unaltered scene. The disadvantages include noticeable position errors, especially along linear features where the realignment of pixels is obvious.

About[1]- Lt.Dr.S.Santhosh Baboo, Reader, Postgraduate and Research department of Computer Science, Dwaraka Doss Goverdhan Doss Vaishnav College, Chennai E-mail id:santhosh2001@sify.com
About[2]- M.Renuka Devi,Asst.Prof, MCA Department, Sree Saraswathi Thayagaraja College, Pollachi, E-mail id:renuka_srk@yahoo.co.in

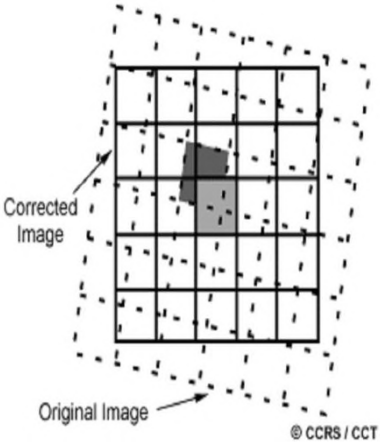


Figure 2.Nearest Neighbor pixel Calculation

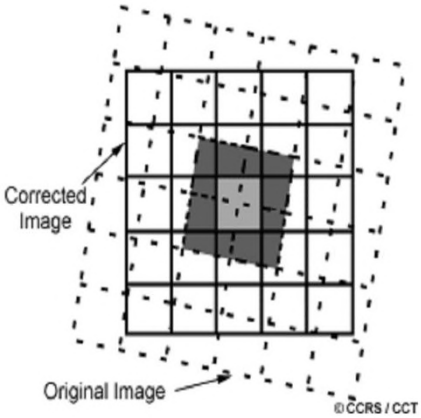


Figure3.Bilenear Interpolation

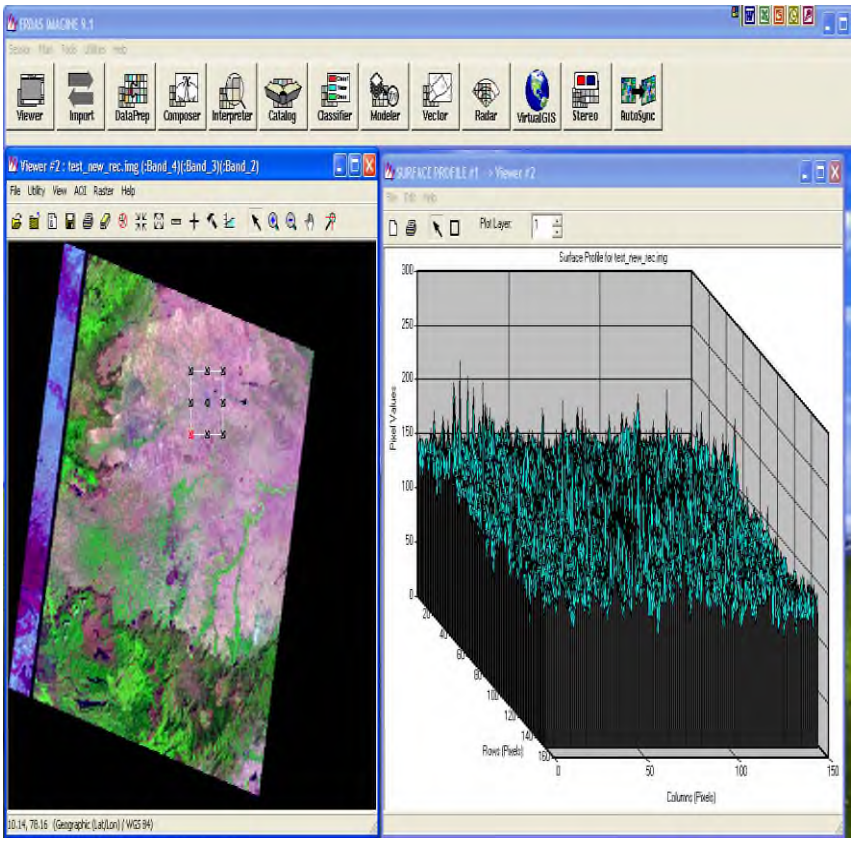


Figure.2.1 Surface profile of Nearest Neighbor resampling method

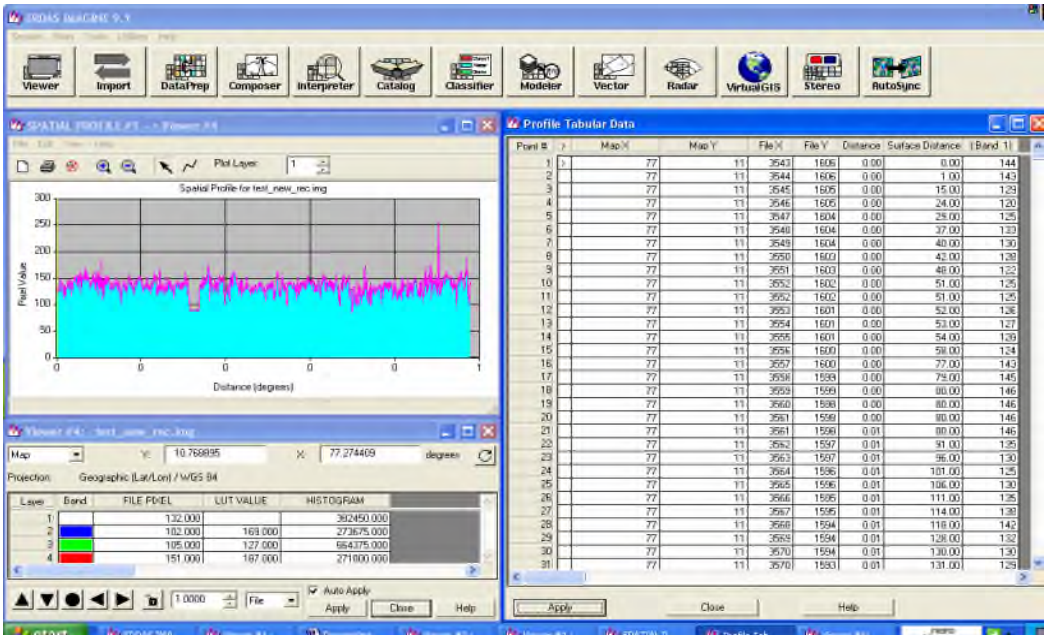


Figure. 2.2 pixel, histogram and spatial profile details of nearest neighbor

III. BILINEAR INTERPOLATION

Bilinear can refer to bilinear filtering or bilinear interpolation. Bilinear filtering is a method used to smooth out when they are displayed larger or smaller than they actually are. Bilinear filtering uses points to perform bilinear interpolation. This is done by interpolating between the four pixels nearest to the point that best represents that pixel (usually in the middle or upper left of the pixel). Bilinear interpolation resampling takes a weighted average of 4 pixels in the original image nearest to the new pixel location

The averaging process alters the original pixel values and creates entirely new digital values in the output image. This may be undesirable if further processing and analysis, such as classification based on spectral response is to be done. If this is the case,resampling may best be done after the classification process.It is shown figure 2.This resampling method assigns the average digital number (DN) of the four pixels closest to the input pixel (in a 2x2 window) to the corresponding output pixel. The mathematical function is bilinear.When apply this method in the coimbatore imagery we can get the following result in the figure 2.

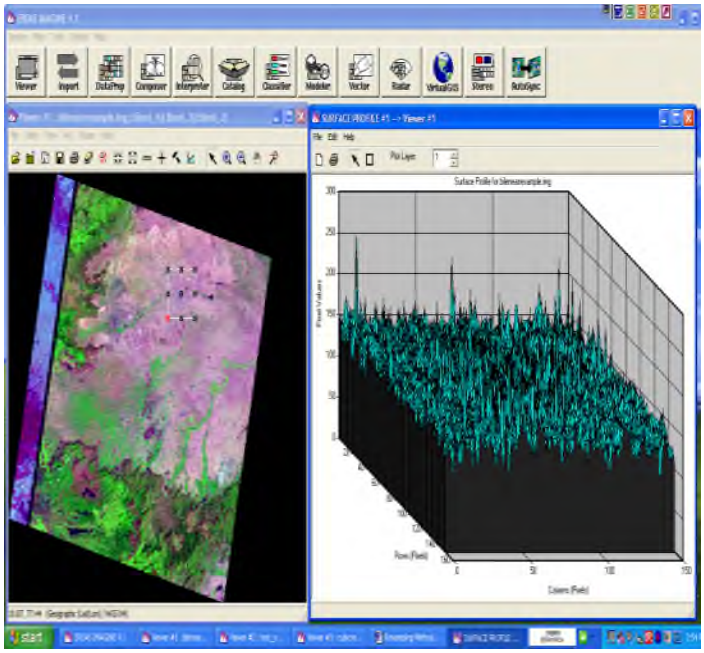


Figure 2.1Surface profile of bilinear resampling method

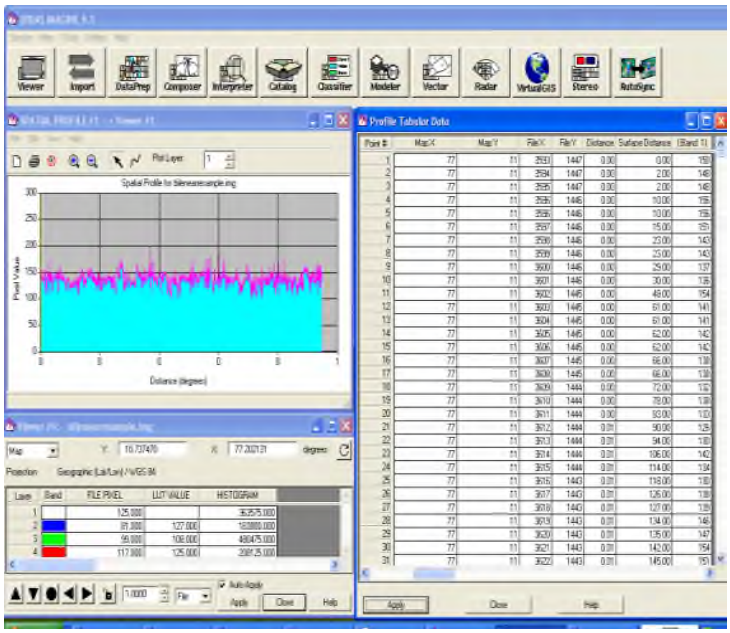


Figure 2.2.pixel,histogram and spatial profile details of bilinear resampling

IV. CUBIC CONVOLUTION

Cubic convolution is a method used to determine the gray levels in an image. This is determined by the weighted average of the 16 closest pixels to the input coordinates. Then that value is assigned to the output coordinates. This method is slightly better than bilinear interpolation, and it does not have the disjointed appearance of nearest neighbor interpolation. Cubic convolution requires about 10 times the computation time required by the nearest neighbor method. This resampling calculates a distance weighted average of a block of sixteen pixel from the original image which surround the new output pixel location

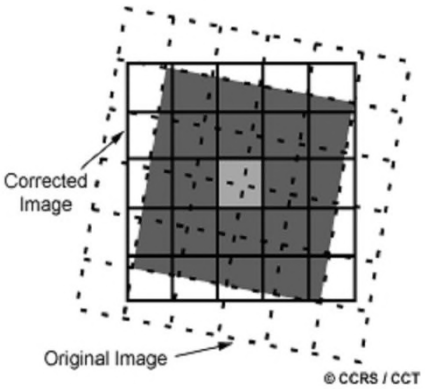


Figure 4 Cubic Convolution

This resampling method assigns the average DN of the sixteen pixels closest to the input pixel (in a 4x4 window) to the corresponding output pixel. The mathematical function is cubic.

V. RESULT AND DISCUSSION

Every one method has some advantage and disadvantage for some type of data,when compare result of surface and spatial profile of nearest neighbor resampling method the Cubic Convolution is gives the better result in coimbatore imagery.we can see the difference between these method in the following figures.Figure 5 shows the result of nearest neighbor and bilenear interpolation methods.In figure 5.1 we see the difference between Bilinear Interpolation and Cubic Convolution. The advantage of Nearest Neighboris extremes subtletie sare not lost and fast computation is possible. But the distadvantage of this method is "stairstepped".That is resampling to smaller gridsize effect around diagonallines and curves.This problem is solved by Bilenear Interplation.Its Result are smoother,accurate,without “stairstepped” effect. But it has some disadvantage that is edges are smoothed and some extremes of the data file valuesare lost. The most accurate resampling method is cubic convolution. It gives the effect of cubiccurve weighting can sharpen the image and smoothout noise. But the disadvantage is most computational is needed and it does n’t provide desired result.

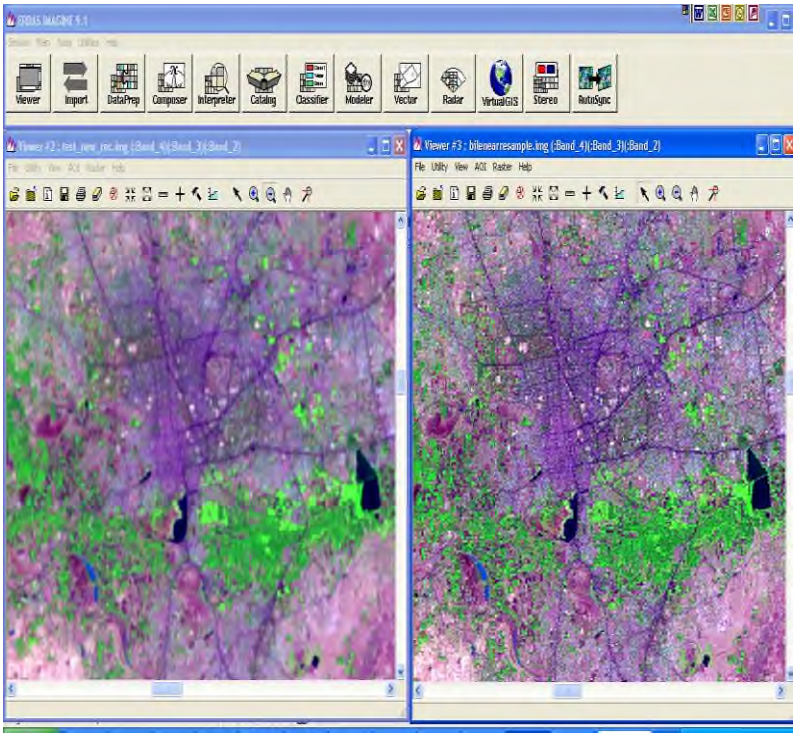


Figure 5.Result of nearest neighbor and Bilinear Interplation

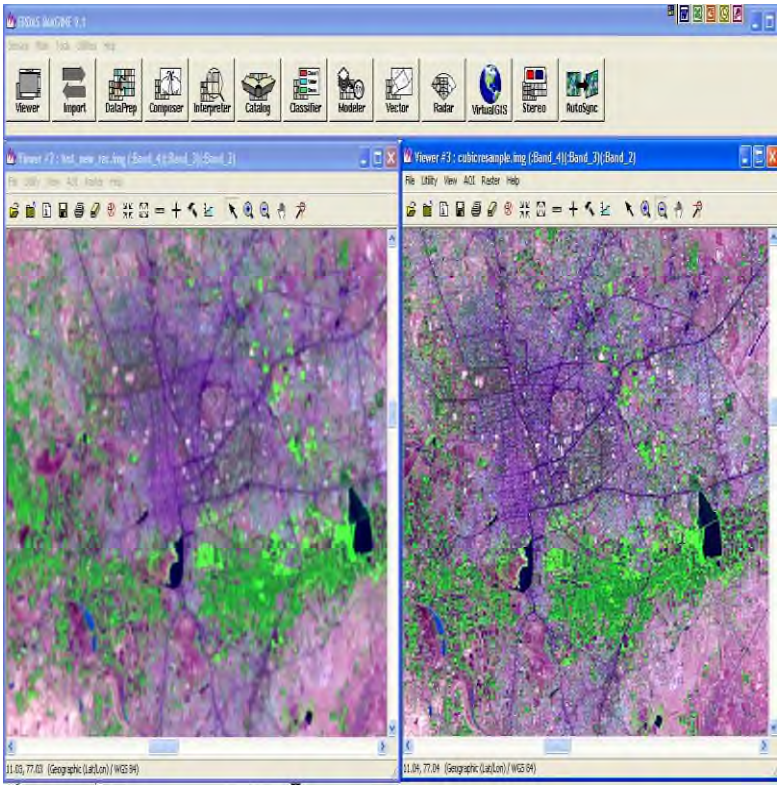


Figure 5.1.Result Bilinear Interpolation and Cubic Convolution

VI. CONCLUSION

The three resampling methods; Nearest Neighbor, Bilinear Interpolation and Cubic Convolution, determine how the cell values of an output raster are determined after a geometric operation is done. The method used depends upon the input data and its use after the operation is performed. Nearest Neighbor is best used for categorical data like land-use classification or slope classification. The values that go into the grid stay exactly the same, a 2 comes out as a 2 and 99 comes out as 99. The value of of the output cell is determined by the nearest cell center on the input grid. Nearest Neighbor can be used on Continuous data but the results can be blocky.Bilinear Interpolation uses a weighted average of the four nearest cell centers. The closer an input cell center is to the output cell center, the higher the influence of its value is on the output cell value. This means that the output value could be different than the nearest input, but is always within the same range of values as the input. Since the values can change, Bilinear is not recommended for categorical data. Instead, it should be used for continuous data like elevation and raw slope values. Cubic Convolution looks at the 16 nearest cell centers to the output and fits a smooth curve through the points to find the value. Not only does this change the values of the input but it could also cause the output value to be outside of the range of input values (imagine a sink or a peak occurring on a surface). This method is also not recommended for categorical data, but does an excellent job of smoothing continuous data.

VII. REFERENCES

1) Christensen, Albert H.J. (1996). The practice of piecewise fits with particular refer- ence to property mapping. Surveying and Land Information Systems, 56, 165-183.
2) Dowman,I., and Tao C.V.,2002.An update on the use of rational functions for Photogrammetric restitution,Highlights ISPRS,vol.7,No.3,pp.22-29.
3) Fraser,C.S.,Baltsavias, E.,Gruen, A.,2002.Processing of Ikonos imagery for submetre 3D positioning and building extraction,ISPRS Journal of photogrammetry & Remote Sensing 56 (2002) 177-194.
4) Fraser ,C., Hanley, H.,Yamakawa, T.2001.sub-Metre Geopositioning with Ikonos Geo imagery , ISPRS Joint Workshop on High Resolution Mapping from Space , Hanover,Germany ,2001.
5) Fritz,L.,1995.Recent Developments for Optical Earth Observation in the United states, Photogrammetric Week,pp75-84,Stuttgart ,1995.
6) Hanley,H.B. and Fraser,C.S.,2001.Geopositioning accuracy of Ikonos imagery : indications from 2D transformations,photogrammetric Record ,17(98) : 317-329.
7) Hattori,S.,one ,T., Fraser, C.S. and Hasegawa,H.,2000. Orientation of high-resolution satellite images based on affine Projection,International Archires of photogrammetry & Remote Sensing, Amsterdam,33(B3):59-366(on CD ROM).
8) Li, R., Zhou, G., Yang, S., Tuell, G., Schmidt, N. J. and Flower, C , 2000, Astudy of the potential attainable geometric accuracy of IKONOS satellite images, IAPRS, 33(B4), 587-595.
9) Jensen, John R. (1996). Introductory Digital Image Processing: A Remote Sensing Perspective (2nd ed.). Chapter 6, Image Preprocessing: Radiometric and Geometric Correction. Upper Saddle River, NJ: Prentice-Hall. pp. 107- 137.
10) ovak, Kurt (1992). Rectification of digital imagery. Photogrammetric Engineering & Remote Sensing, 58, 339-344.
11) Wolf, Paul. R. and Ghilani, Charles. D. (1997). Adjustment Computations: Statistics and Least Squares in Surveying and GIS. New York: John Wiley & Sons, 564 p.

A negative auction

GJCST Classification F.4

Lorenzo Cioni

***Abstract*-In this paper we describe a type of auction mechanism where the auctioneer A wants to auction an item ζ among a certain number of bidders $b_i \in B$ (i = 1, . . . , n) that submit bids in the auction with the aim of not getting ζ.Owing to this feature we call this mechanism a negative auction. The main motivation of this mechanism is that both the bidders and the auctioneer give a negative value to the auctioned item (and so they see it as a bad rather than a good).The mechanism is presented in its basic simple version and with some possible extensions that account for the payment of a fee for not attending the auction, the interactions among the bidders and the presence of other supporting actors.**

I INTRODUCTION

In this paper we describe a type of auction mechanism[1] where the auctioneer A wants to auction an item ζ among a certain number of bidders[2] $b_i \in B$ (i = 1, . . . , n) that submit bids in the auction with the aim of not getting ζ.Owing to this feature we call this mechanism a negative auction ([4]). The main motivation of this mechanism is twofold ([7] and [8]):

- *both the bidders and the auctioneer give a negative value to the auc tioned item (and so they see it as a bad rather than a good),*
- *the auctioneer has an imperfect knowledge of the bidders and so cannot contact any of them directly.*

The mechanism[3], at least in its basic version, is simple and will be described in detail in section 5 whereas the needed details will be presented in the sections 3 and 4.

Algorithm 1.1 The basic mechanism is based on the following steps.

- *A selects the bidders b_i according to some private criteria that depend on the nature of ζ;*
- *the b_i submit their bids in a sealed bid auction;*
- *once the bids have been submitted they are revealed so that:*
- *the bidder who made the lowest bid is the losing bidder and gets*[4] *ζ;*
- *the other bidders are termed winning bidders and get the benefit of having avoided the allocation of ζ;*
- *the losing bidder*[5] *b_1 gets ζ and, as a compensation, a sum equal to his bid x_1;- each winning bidder b_i pays to the losing bidder a properly defined frac tion of x_1.*

This simple mechanism will be described in some detail in the following sections together with the possible strategies of the bidders and some possible extensions. The extensions include a pre auction phase, where some of the bidders pay a fee for not attending the auction, and a post auction phase that can assume three forms and that aims at the reallocation of ζ depending on criteria that are different from those that drove the auction phase itself.

II PRE AUCTION AND POST AUCTION PHASES

In the pre auction phase the bidders are allowed to pay to A a fee f (that A fixed and made common knowledge among the bidders) for not attending the auction. In this case, depending on the amount of the fee, we can have that:

- *m bidders prefer to pay the fee in order to not attend the auction;*
- *k = n − m bidders prefer to attend the auction.*

In this case, at the end of the auction phase, A has collected an extra compensation equal to $e_C = mf$ that is awarded to the losing bidder.The value e_C (see also section 8) may be either a public knowledge among the bidders that therefore know m but not necessarily k (since the value n is not necessarily a common knowledge among the bidders) before the auction phase or it may be a private knowledge of A to be revealed only after the execution of the auction phase.

About-lcioni@di.unipi.it Computer Science Department, University of Pisa largo Pontecorvo 3, 56127, Pisa, Italy tel.: (+39) 050 2212741 fax: (+39) 050 2212726

[1]In this paper we are going to use the term mechanism in a rather informal sense as a set of rules, strategies and procedures. For a more formal use of the term we refer, for instance, to [7] and [9].

[2]In what follows we identify a bidder $b_j \in B$ also by the index $j \in N = \{1, \ldots, n\}$.

[3]The proposed mechanism is loosely inspired by the Contract Net Protocol ([6, 14]).

[4]Possible ties among two or more losing bidders are resolved through a properly designed random device.

[5]We assume that after the bids have been revealed we renumber the bidders so that the losing bidder is the bidder b_1 whereas all the other bidder b_i (with i 6= 1) are the winning bidders.

As to the last point we note how this feature may be guaranteed or at least enforced through the design of the structure of the pre auction phase that can be designed so to make the communication among the bidders either too difficult or too costly. The easiest solution is to have the bidders, at least in this phase, to be unaware one of the others so to make any inter bidders communication impossible. In the present paper we consider only the private knowledge case so that the value e_c has no influence on the behavior of the k attending bidders that do not have such information when they submit their bids (see section 8).We note indeed how even the m bidders who paid the fee can attend the possible post auction phase. This requires that in that phase the full set of bidders is revealed and becomes a common knowledge. In the post auction phase we introduce some mechanisms that try to correct a simplifying assumption that we have made in the basic mechanism. The basic mechanism is, indeed, based on the assumption that the various b_i are independent one from the others (in the sense that the allocation of ζ to one of the bidders has effect only on that bidder) and, similarly, do not influence any other actor[6].The mechanisms of the post auction phase aim, indeed, at accounting for the following facts:

(pa1) the bidders b_i are interdependent and so they may influence each other so that, for any pair of bidders (b_i, b_j), we can define as $d_{i,j}$ the damage caused to b_i from the allocation of ζ to b_j;

(pa2) the bidders b_i may influence the actors of the set S (see footnote 6) so that, for any actor $s_i \in S$, we can define as $D_{i,j}$ the damage caused to s_i from the allocation of ζ to b_j.

We may assume in general that $d_{i,j} \neq d_{j,i}$ so the cross damages between pairs of bidders are not symmetrically distributed.In the (pa1) case we assume that the bidders are interdependent but $S = \emptyset$. In this case the bidders can try to negotiate an allocation to another bidder that is more preferred by all the bidders depending on the values $d_{i,j}$ (for $i \neq j$) and not on the values $m_i = d_{i,i}$ that drive the auction phase.In the (pa_2) case, we assume that the bidders are independent but $S \neq \emptyset$. In this case the members of S may try to obtain a reallocation depending on the values $D_{i,j}$. Last but not least the two cases (pa1) and (pa2) can be merged in a single case where we have both interdependent bidders and $S \neq \emptyset$.In all the post auction cases the starting point is the allocation of ζ to one of the bidders on the basis of the outcome of the auction phase where each bidder is guided only by his self damage $m_i = d_{i,i}$.At the end of the auction phase we can have two cases:

- the resulting allocation is satisfactory[7];
- the resulting allocation is unsatisfactory.

In the former case no reallocation is required whereas in the latter case either the bidders of the set B or the supporters of the set S may try to renegotiate it, within the proposed mechanisms in order to identify a new bidder as the more preferred allocation. We underline how such reallocation may require the raising of a further compensation for the new bidder in order to have him accept the allocation of ζ.

III THE DEFINING PARAMETERS

Both the auctioneer A and the bidders of the set B are characterized by some parameters that depend heavily on the nature of the item ζ but also on their individual characteristics.

Definition 3.1 For what concerns **A** we have only one parameter: the value m_A that A assigns to ζ as a measure of his utility since the only benefit that

A receives from the auction is the allocation of ζ. With m_A we denote:

- *the damage or the negative utility that **A** will receive from ζ if the auction is void so the allocation fails;*
- *the benefit or the positive utility that **A** receives from the allocation of ζ to one of the $b_i \in B$.*

Observation 3.1 *In the former case m_A has a negative value whereas in the latter it has a positive value.*

Definition 3.2 *Every $b_i \in B$ is characterized by the following parameters (see also [7, 8]):*

- *a value m_i that he assigns to ζ;*
- *the amount x_i he is willing to bid;*
- *the random variables X_j that describe the bids of the other bidders;*
- *the interval of the values $[0, M_i]$ to which the m_i belong;*
- *the intervals of the values $[0, M_j]$ to which the X_j belong;*
- *the differentiable cumulative distributions F_j of the values X_j;*
- *the corresponding density functions $f_j = F_j'$ of such values.*

Observation 3.2 *We note that:*

(1) the parameter m_i has a dual meaning in the sense that:

- *it represents the damage that b_i receives from the allocation of ζ;*
- *it represents the benefit that b_i gets from the allocation of ζ to some other bidder;*

[6]With the term actor we denote a figure that is distinct from both A and the Bs but that wants to attend the auction since he thinks to be damaged from the allocation of ζ to one of the bidders. Such actors are termed supporters and form the set S.

[7]The concept of satisfaction will be defined for each post auction phase. For the moment we say that an allocation is satisfactory if there are no incentives for its modification either from the members of B or from the members of S or from both.

(2) the parameter x_i has a dual meaning in the sense that:

- it represents the sum that b_i asks as a compensation for the allo cation of ζ;

- it defines the fraction c_i of the compensation that b_i has to pay to the losing bidder.

We can also define the following probabilities:

- the probability p_i for b_i of losing the auction;

- the dual probability $q_i = 1 - p_i$ for b_i of winning the auction.

We recall that the losing bidder is the bidder who gets ζ and a compensation from the other bidders, the winning bidders.

IV THE BASIC ASSUMPTIONS

In this section we introduce the basic assumptions that we make on the parameters that characterize both the auctioneer and the bidders and that will be maintained through the rest of the paper.

Assumptions 4.1 The only assumption we can make on **A** is that his value m_A is a private information of the auctioneer so that it is not known to the bidders.If we relax this assumption so that m_A becomes a common knowledge among the bidders we may assume that such a knowledge may influence the evaluations of the bidders since they may derive form this knowledge hints on the real nature of the auctioned item.

Assumptions 4.2 The basic assumptions that involve the characteristic parameters of the bidders may be summarized as follows[8]:

-the bidders are assumed to be risk neutral so that their utility is linearly separable ([7]) and can be expressed as the difference between a benefit and a damage and so as $x_i - m_i$ if the bidder b_i loses the auction or as $m_i - c_i$ if he wins it;

-the random variables X_j are assumed to belong to a common interval [0, M] for a suitable $M > 0$;

-the random variables X_j are assumed to be independent random variables;

-the valuations m_i are assumed to be private values of the single bidders;

-the bidders b_j are assumed to be symmetric so they are characterized by the same F and by the same corresponding f;

- the random variables X_j are assumed to be uniformly distributed on the interval [0, M] so that we have, for $x \in [0, M]$:

and, correspondingly:

$$P(X_j \leq x) = F(x) = \frac{x}{M} \qquad (1)$$

$$f(x) = \frac{1}{M} \qquad (2)$$

From the foregoing assumptions we derive that the probability p_i for each bidder b_i of losing the auction is the same for all the bidders so we can denote it as p and use $q = 1-p$ to denote the dual probability of winning the auction.

Observation 4.1 Possible relaxations of the foregoing assumptions involve:

- the possibility that the bidders are risk adverse[9] so that his utility is no more linearly separable but it is a convex function of x_i;

-the possibility that the evaluations are either common or interdependent among the bidders;

-the possibility that the bidders are asymmetric so that we can have different intervals $[0, M_j]$ and different functions F_j and f_j for each bidder b_j as well as the possibility to have different distributions (such as a Gaussian or a triangular distribution) also under the symmetry assumption.

Such relaxations can be introduced either one at a time or in combinations. Their treatment, that makes the analysis more complex, is out of the scope of the present paper and is the subject of further research efforts (see section 8 for further details).

[8] See [7, 8]

[9] We recall that, in classical terms, a player is risk neutral ([5]) if he is indifferent between attending a lottery and receiving a sum equal to its expected monetary value whereas he is risk averse if he prefers the expected value to attending the lottery. We can also say that a player is risk neutral if his utility function is linearly separable in gain and loss whereas, if he is risk averse, it can be seen as a concave function. In our context we have to consider the opposite perspective and so we consider the utility function of risk averse bidders as a convex function of its meaningful parameters.

V THE BASIC MECHANISM AND ITS STRATEGIES

The basic mechanism has only the auction phase among independent bidders with $S = \emptyset$.

Algorithm 5.1 We can describe the basic mechanism with the following algorithm[10].

(ph1) A auctions ζ;

(ph2) the b_i make their bids x_i in a sealed bid one shot auction;

(ph3) the bids are revealed;

(ph4) the lowest bidding bidder[11] b_1 gets ζ and x_1 as a compensation for this allocation;

(ph5) each of the other bidders b_i pays to b_1 $\sum$ a fraction c_i of x_1 such that:

$$\sum_{i \neq 1} c_i = x_1 \qquad (3)$$

Observation 5.1 For what concerns the values c_i we assume a proportional repartition among the bidders so we have:

$$c_i = x_1 \frac{x_i}{X} \qquad (4)$$

where $X = \sum_{j \neq 1} x_j$. In this way we account for the fact that the bidders who receive a bigger advantage from the allocation of ζ to b_1 pay the higher fractions of the compensation. At this point we state and prove the following proposition.

Proposition 5.1 (Weakly dominant strategy) From the assumptions we made in section 4 we derive that it is a weakly dominant strategy for each bidder to submit a bid x_i equal to his evaluation m_i of the auctioned item ζ.

Proof

From what we have stated in sections 3 and 4 we derive easily that the expected utility from the auction for every bidder b_i when he faces the phase (ph_2) can be expressed as:

$$E(b_i) = p(x_i - m_i) + (1 - p)(m_i - x_1 \frac{x_i}{X}) \qquad (5)$$

as the sum of the utility if he loses the auction multiplied with the probability of losing it and the utility if he wins it multiplied with the probability of winning it.Relation (5) can be rewritten as:

$$E(b_i) = (1 - \frac{x_i}{M})^{n-1}(x_i - m_i) + (1 - (1 - \frac{x_i}{M})^{n-1})(m_i - x_1 \frac{x_i}{X}) \qquad (6)$$

by using the following equalities:

$$p = (1 - \frac{x_i}{M})^{n-1} \qquad (7)$$

$$q = 1 - p = 1 - (1 - \frac{x_i}{M})^{n-1} \qquad (8)$$

that have been derived by using the hypotheses of independence and identical and uniform distribution of the X_j and by imposing that the x_i is lower than any of the X_j for $j \neq i$. Since in relations (5) and (6) we want to impose that in any case each bidder b_i has a non negative utility we get the following constraints[12]:

- $y_1 = x_i - m_i \geq 0$
- $y_2 = m_i - x_1 \frac{x_i}{X} = m_i - x_1 \frac{x_i}{x_i + X'} \geq 0$

where y_1 is the utility for b_i if he loses and y_2 is his utility if he wins. From the former constraint we derive:

$$x_i \geq m_i \qquad (9)$$

For what concerns the latter constraint, from the definition of y_2 and by performing the derivations with respect to x_i, we easily derive that:

- $y_2' < 0$
- $y_2'' > 0$

[10] Also in this section we assume that, when the phase (ph_3) is over we can renumber the bidders so that b_1 is the losing bidder whereas the b_i (with $i \neq 1$) are the winning bidders.

[11] Possible ties are resolved with the random selection of one of the tied bidders.

[12] We note how we can write $X = x_i + X$ where X accounts for the bids of the bidders distinct from b_1 and b_i.

so y_2 is concave decreasing with:

-a maximum value equal to $y_2(0) = m_i$ for $x_i = 0$,

-a minimum value equal to:

$$y_2(M) = m_i - x_1 \frac{M}{M + X'} \qquad (10)$$

for $x_i = M$

It is easy to verify that we have $y_2(m_i) > 0$ whereas we cannot exclude that $y_2(M)$ may assume negative values though this is a rather unlikely event. From relations (7) and (8) we can easily see how:

-p has a maximum value of 1 for $x_i = 0$, decreases for x_i increasing and attains a null value for $x_i = M$;

- q has dual behavior since it has a minimum value of 0 for $x_i = 0$, increases for x_i increasing and attains the maximum value of 1 for $x_i = M$;

At this point we want to find the value $\bar{x}_i$ where we have

$$p = q \qquad (11)$$

so that for $x_i < \bar{x}_i$ we have that p dominates q whereas we have the opposite for $x_i > \bar{x}_i$. From relation (11) and relations (7) and (8) we get:

$$(1 - \frac{x_i}{M})^{n-1} = 1 - (1 - \frac{x_i}{M})^{n-1} \qquad (12)$$

From relation (12), with some easy algebra, we derive:

$$\bar{x}_i = (1 - (\frac{1}{2})^{\frac{1}{n-1}})M \qquad (13)$$

We note that $\bar{x}_i \to 0$ *as* $n \to \infty$ so that q tends to dominate p for any x_i. According to all this we have that b_i should maximize y_2 so to bid no less than m_i and so (given the constraint we have imposed on y_1) he should bid a sum equal to m_i.

Observation 5.2 We note that we have:

$$\frac{p}{p'} \to 0 \; as \; n \to \infty \qquad (14)$$

where p′ is the derivative of p as a function of x_i whereas:

$$\frac{q}{q'} \to \infty \; as \; n \to \infty \qquad (15)$$

where q' is the derivative of q as a function of x_i.

Observation 5.3 It is obvious that at phase (ph_3) each b_i knows if he is the loser or one of the winners. In the former case he has a utility:

$$x_1 - m_1 \qquad (16)$$

whereas in the latter he has a utility:

$$m_i - x_1 \frac{x_i}{X} \qquad (17)$$

Observation 5.4 We have in this way verified how the truthful bidding is a weakly dominant strategy for each bidder in the basic mechanism of the negative auction.

Observation 5.5 The proposed mechanism has a strong analogy with a First Price Sealed Bid auction ([7]). In the auctions of this type the winning bidder is the highest offering bidder who pays his bid. Under hypotheses similar to the ones we made in sections 3 and 4 we have that in a First Price Sealed Bid auction the best strategy for each bidder is to bid a little less than one's own evaluation or to bid $x_i = m_i - \delta$ *with* $\delta \to 0$ *for* $n \to \infty$.

If we suppose to use negative prices our mechanism is analogous to a First Price Sealed Bid auction so, in our case, the best strategy for each bidder is to bid a little more than one's own evaluation or to bid $x_i = m_i + \delta$ *with* $\delta \to 0$ *for* $n \to \infty$.

VI THE USE OF THE FEE

In this section we present the pre auction phase where:

- m bidders pay the fee f in order to not attend the auction;

- $k = n - m$ bidders prefer to attend the auction.

We make the hypothesis that the sum $e_c = mf$ is a private information of **A** so it is unknown to the other k bidders that neither know n. For the k attending bidders we can repeat what we have said in section 5.In this case the losing bidder, at the end of the auction phase, gets the following final compensation:

$$f_c = x_1 + e_c = x_1 + mf \quad (18)$$

If the mechanism has a post auction phase then all the initial n bidders can attend to it, as we will show in the following sections.At this point we define the following profiles:

(ne1) all the n bidders pay the fee f,

(ne2) none of the n bidders pays the fee f.

We want to see if such profiles are Nash Equilibria[13] *(NE)* or not.

In the case $(ne1)$ we have that if the bidders collude among themselves and decide that they all pay the fee f they collect the sum $e_C = nf$. In this case, every bidder would have a utility equal to[14] $m_i - f$. If only one bidder b_j individually violates the collusive agreement he gets a utility equal to:

$$(n-1)f - m_j \quad (19)$$

since no further compensation from the auction phase is possible. The individual deviation is profitable (so that $(ne1)$ is not a NE) if we have:

$$(n-1)f - m_j > m_j - f \quad (20)$$

or if:

$$m_j < f\frac{n}{2} \quad (21)$$

So if the fee *f* is such that the constraint (21) is satisfied for at least one b_j the collusive agreement is not a *NE* and the auction cannot be void since **A** is able to find a bidder to which to allocate ζ with a compensation paid by the other bidders. We note that if **A** fixes *f* as:

$$f > \frac{2M}{n} \quad (22)$$

we have:

$$\frac{n}{2}f > M \geq m_i \; \forall b_i \quad (23)$$

and so relation (21) is surely verified.

In the case $(ne2)$ the individual deviation depends on the possible policies of the single bidders since we have that $e_C = 0$ so from this condition we cannot derive any incentive for the bidders to deviate. In order to understand under which conditions the case $(ne2)$ can occur we therefore examine a more general case and so under which conditions a bidder is better off if he pays the fee than if he attends the auction.

A bidder b_i has indeed the following possibilities[15]:

(1) he pays the fee f and has an utility[16] $u_i^p = m_i - f$;

(2) he does not pay and attend the auction and so:

(2a) he has an utility $u_i^l = x_i - m_i$ if he loses the auction,

(2b) he has an utility $u_i^w = m_i - x_1\frac{xi}{xi+X'}$ if he wins the auction.

From the case (1) we derive the first constraint since we have that if $u_i^p < 0$ then b_i does not pay the fee and attends the auction. This requires that:

$$u_i^p = m_i - f \geq 0 \quad (24)$$

or:

$$f \leq m_i \quad (25)$$

If condition (25) is violated for every b_i so that we have:

$$f > m_i \quad (26)$$

for every bi we have that no bidder pays the fee. In this way we have that if $f > max\{mi\}$ or if f is very high no bidder pays the fee and so they all attend the auction phase. If *f* is assigned a lower value some of the bidders prefer to pay it whereas others prefer to attend the auction. Lastly, if *f* gets a very low value we have that all the bidders may prefer to pay it so that the auction phase is void, without any discordance with what we have seen with regard to *(ne1)*.Once we have established that relation (24) is satisfied we want to make a comparison with the cases (2*a*) and (2*b*) so to understand if a bidder is better o by paying the fee or by attending the auction.

[13]A Nash Equilibrium is a profile of strategies for the bidders where none of them has a gain from an individual deviation ([1, 2, 9, 10]).

[14]This requires f < mi for every bi. We comment on this assumption shortly.

[15]We use the decorations p, l and w as exponents to denote, in the order, a payment, a loss and a win.

[16]In this case we evaluate the utility under the hypothesis of risk neutrality and so as the di erence between the benefit, as represented by the missed allocation of , and the payment as represented by the fee f.

We can make the following comparisons:

$$m_i - f \geq x_i - m_i \quad (27)$$

and:

$$m_i - f \geq m_i - x_1 \frac{x_i}{x_i + X'} \quad (28)$$

If such relations are satisfied then b_i is better off by paying the fee and so by not attending the auction. From relation (27) we derive:

$$f \leq 2m_i - x_i \leq m_i \quad (29)$$

(since we have assumed *xi ≥ mi)* and so not really a new constraint since it coincides with relation (25). On the other hand from relation (28) we get:

$$f \leq x_1 \frac{x_i}{x_i + X'} \leq x_1 \frac{x_i}{(n-1)x_1} \leq \frac{M}{n-1} \quad (30)$$

since, by the definition of x1 and xi, we get $X = xi + X' \geq (n-1)x1$ and $x1 \leq xi \leq M$ for every *bi.* From relation (30) we derive that if the fee *f* is small enough then the bidders have incentive to pay it otherwise they have incentives to attend the auction. From this we may derive that if **A** fixes f high enough (for instance *f = M/2)* he can be sure to have a non void auction even if some bidders may prefer to pay the fee *f.*

VII THE POST AUCTION PHASE

1) *Introductory remarks*

In the simplest case, when the auction phase is over, the allocation is performed by the bidders on the basis of the values *mi = di,i* only. This way of proceeding is based on the assumption that the bidders are independent and so that the allocation damages only each individual bidder and neither other bidders nor any other of the actors of the set S (the supporters). n section 7.2 we see how we can account for the interdependence of the bidders and so for the damages among the bidders. We therefore present an algorithm based on a succession of push operations by which a bidder can push ζ towards another more preferred bidder (according to the valuesattributed to the cross damages di,j).In this case we have no supporters so that $S = \emptyset$. In section 7.3 we assume that the bidders are independent but $S = \emptyset$.and we examine if the supporters can push ζ towards another more preferred bidder (according to the values attributed to the cross damages Di,j by the $s_i \in S$) Last but not least in section 7.4 we present an attempt to merge the two approaches since we assume to have both interdependent bidders and $S = \emptyset$.

2) *The interaction among the bidders*

Definition 7.1 (The added parameters) In addition to the parameters we have seen in section 3 and the assumptions we have made in section 4 we introduce the following parameters for every bidder bi:
-di,j ≥ 0 is the damage that bi receives if is allocated to bj;
-ci,j ≥ 0 is the contribution that bi is willing to pay to bj to have him accept the allocation of .
Observation 7.1 It is obvious that *mi = di,i* and *ci,i = 0.*
Before going on we recall that the auction phase ends with the allocation of to b1 who receives a compensation equal to x_1. We can define the due payment that b1 receives from every bidder $b_i \neq b_1$ as:

$$\sigma_{i,1} = x_1 \frac{x_i}{X} \quad (31)$$

(with $X = \sum_{j \neq 1} x_j$) so that we have:

$$\Sigma_1 = \sum_{i \neq 1} \sigma_{i,1} = x_1 \quad (32)$$

We can also define:

$$\Sigma_j = \Sigma_1 - \sigma_{j,1} \quad (33)$$

to be used shortly.
Mechanism 7.1 In this case the mechanism has the following structure:
-possible pre auction phase,
-auction phase,
-allocation and compensation phase, -reallocation phase.
From the allocation and compensation phase b_1 would get, from the members of $N_{-1} = N \setminus \{1\}$ the commitments of payment $\sigma_{i,1}$ that form the compensatory sum $\sum_1$whereas the reallocation phase depends on the values *di,j.*When the allocation phase is over, b1 orders the $d_{1,j}$ $\forall j \neq 1$ with regard to $d_{1,1} = m_1$. We can have two cases:
- $d_{1,1} < d_{1,j}$ $\forall j \neq 1$ so b_1 is satisfied and no reallocation is required;
- $\exists J_1 \subseteq N_{-1}$ such that $\forall j \in J_1$ $d_{1,j} < d_{1,1}$.

In the former case the mechanism ends and b_1 receives the commitments at payment as effective compensations from the other bidders. In the latter case b_1 may negotiate a reallocation with the members of J1 that he orders in increasing order of the damages $d_{i,j}$. We note that for any b_j with $j \in J_1$ we define as $\bar{c}_{1,j} = d_{1,1} - d_{1,j}$ the maximum contribution that b_1 is willing

to have him accept ζ whereas with $c_{1,j} < \bar{c}_{1,j}$ we denote the current value of this contribution.

Algorithm 7.1 The attempt of reallocation may proceed along the following steps:
(1) b_1 defines J_1;
(2) we have two cases:
(2a) $J_1 = \emptyset$ so go to (5);
(2b) $J_1 \neq \emptyset$ so go to (3);
(3) b_1 contacts (in the order) a b_j with $j \in J_1$ and offers him a further compensation $c_{1,j} < \bar{c}_{1,j}$ so that b_j would get $\Sigma = \Sigma_j + c_{1,j}$;
(4) at this point we have two cases:
(4a) b_j accepts and so becomes the new b_1 with $\Sigma 1 = \Sigma j + c_{1,j}$; go to (1);
(4b) b_j refuses so we have two cases:
(4b$_1$) there is one more b_j that can be contacted so go to (3);
(4b$_2$)there is no b_j to contact so the procedure ends with a failure; go to (5);
(5) end;
The operation at step (3) is a push operation through which the current b_1 tries to allocate ζ to some other bidder b_j having a benefit equal to $d_{1,1} - d_{1,j} - c_{1,j}$. Such procedure may either succeed or fail. For it to succeed the current b_j must accept the proposal of b_1. It is easy to see that b_j accepts if the following conditions are verified:
(ac$_1$) $\Sigma \geq m_j$

(ac$_2$) $d_{j,1} \geq d_{j,j}$

If condition (ac$_1$) is violated b_j surely refuses the push proposal whereas if the condition (ac$_2$) is violated b_j can accept ζ , with a risky decision, if he is sure he can push it to some other bidder b_h such that $d_{j,h} < d_{j,1} < d_{j,j}$. The procedure has the following termination conditions:

- *when no bider accepts a push proposal from the current* b_1;
- *when for a bidder* b_1 *we have* $J_1 = \emptyset$ *so the currently losing bidder is satisfied with the allocation;*
- *when there would be a cycle.*

The last case deserves some more comments. If we have, avoiding to rename the successive losing bidders, the following succession of exchanges:

$$b_1 \rightarrow b_j \rightarrow b_h \rightarrow \cdots \rightarrow b_k \rightarrow b_1 \qquad (34)$$

we have a cycle that could even give rise to a money pump for the initial b_1. To prevent this from occurring we impose a cut on the cycle so that the final accepting bidder must be b_k. This fact requires the recording of the various passages so to detect any cycle and to apply the correcting action.

3) *The presence of the supporters*

In this case we make the following assumptions:
- *the bidders are independent so we have* $d_{i,j} = 0 \ \forall i \neq j$;
- *we have s supporters* $s_i \in S$ *so that for each* s_i *we have the damages* $D_{i,j}$ *that he receives from the allocation of* ζ *to each bidder* b_j. *Mechanism 7.2 Also in*

this case (see section 7.2) the mechanism has the following structure:
- *possible pre auction phase,*
- *auction phase,*
- *allocation and compensation phase,*
- *reallocation phase.*

The reallocation is driven, in this case, by the members of S with their values $D_{i,j}$. We can consider S as partitioned[17]:

$$S = A \cup D \qquad (35)$$

where:
- *A is the set of the* s_i *that agree with the allocation of* ζ *to* b_1 *so that* $s_i \in A$ *if and only if* $D_{i,1} < D_{i,j}$ *for every* $b_j \neq b_1$;
- *D is the set of the* s_i *that disagree with the allocation of* ζ *to* b_1 *so that* $si \in D$ *if and only if*[18] *exists at least a bidder* $j_i \neq 1$ *such that* $D_{i,j_i} < D_{i,1}$.

[17]In a classic way we have $S = A \cup D$ and $A \cap D = \emptyset$. [18]We note that every $s_i \in D$ may have his own j_i.

We can have the following cases:
(1) $A = S$ and $D = \emptyset$ so no reallocation is required;
(2) $A = \emptyset$ and $D = S$ so every s_i has at least a more preferred allocation;
(3) $A = \emptyset$ and $D \neq \emptyset$.
In the case (1) the procedure is obviously over.
In the case (2) for every $s_i \in D$ we can partition N as $N = L_i \cup \{b_1\} \cup U_i$ where:

- L_i identifies the bidders that cause to s_i a lower damage than b_1 or the more preferred bidders;
- U_i identifies the bidders that cause to s_i a greater damage than b_1 or the less preferred bidders.

We can have two cases:

- $\cap_{s_i} L_i = \emptyset$,
- $\cap_{s_i} L_i \neq \emptyset$

In the former case no compromise is possible among the members of D so the allocation of ζ at the current b_1 is unchanged. In the latter case we can have two sub cases. In the former sub case we have $\cap_{s_i} L_i = b_j$ so the members of D offer to b_j both Σ_j(see section 7.2) and $\gamma_j = x_j - \Sigma_j$ to be shared proportionally among the members of D as:

$$\gamma_{i,j} = \gamma_j \frac{D_{i,1} - D_{i,j}}{\sum_{s_i}(D_{i,1} - D_{i,j})} \qquad (36)$$

We note that a proposal to bj is feasible only if, for each supporter si, the following feasibility condition holds:

$$\gamma_{i,j} \leq D_{i,1} - D_{i,j} \qquad (37)$$

If condition (37) is violated for at least one supporter then no proposal can be made so the Ss must consider another of the available bidders, if they have one, otherwise the procedure ends with a failure. If b_j accepts we have a new allocation otherwise the procedure ends with a failure and the allocation is unchanged. For the conditions of acceptance for b_j we refer to section 7.2. In this case b_j accepts if the offered total compensation is enough to cover the damage m_j from the allocation of ζ since the bidders are assumed to be independent. In the latter sub case we have $L = \cap_{s_i} L_i \subset N$ so we identify a set of $l = |L|$ elements. In this case the members of D can use the Borda method[19] ([12, 13]) on such elements so to define the Borda winner (be it bj) and apply to it what we have seen for the single outcome sub case. In the case of a tie on the Borda winners one of such winners can be selected at random since they can be seen as equivalent alternatives. If the new allocation is feasible and the Borda winner accepts the procedure is over otherwise the members of D discard him and repeat the procedure on the reduced set $L \setminus \{b_j\}$ until one of the bidders accepts (so the procedure ends with success) or there is no more Borda winners to be contacted so that the procedure ends with a failure.In the case (3) we have:

- $\forall si \in A$ b_1 *is the best choice;*
- $\forall s_i \in D$ *there are preferred choices to* b_1.

If, for each $si \in D$, we define the set $Li = \{j \in N \mid Di,j < Di,1\}$ *we can define the* set $L = \cap si \in D\, Li$ so that we have three cases:

(a) $|L| = 0$,
(b) $|L| = 1$,
(c) $|L| > 1$.

In the case (a) no reallocation is possible since there is no possible compromise among the members of D that are not able to agree on a feasible alternative to b_1. In the case (b) we have a b_j (with $j \in N$) that is better than b_1 for the members of D. The members of D can proceed as follows:

- *each* $s_i \in D$ *evaluates his individual gain* $D_{i,1} - D_{i,j}$;
- *they evaluate the collective gain* $\Gamma i = \sum_{s_i \in D}(D_{i,1} - D_{i,j})$;

- *they ask to the member of* A *how much they (as a whole) want to be paid to switch from* b_1 to b_j, *be it* $\rho_{1,j}$. If the total of $\rho_{1,j}$ and the sum that the D have to pay to b_j (that accounts also of the payments of the other bidders but $b1$) to have him to accept ζ is lower than $\Gamma\ i$ the reallocation is feasible and the procedure may end with success otherwise it surely ends with a failure.We note that:
- *the reallocation actually succeeds if* b_j *accepts so if the proposed total compensation cannot be lower than* m_j;

[19]Given n alternatives the method is based on the fact that each voter assigns n − 1 points to the top ranked alternative, n − 2 to the second top ranked alternative up to 0 point to the lowest ranked alternative. The points are added together and the alternatives ordered in a weakly descending order (ties are indeed possible) so that the alternative that receives the highest number of points, in absence of ties, is the Borda winner. If we have ties on the top ranked alternatives we can choose one of them at random as the Borda winner.

-the sum $\rho_{1,j}$ is defined by the members of the set A through a negotiation and is proportionally shared among the members of A so that each can compensate the major damages deriving from the new allocation.

In the case (c) we have $L \subset N$ such that b_j is a better choice than b_1 for any $j \in L$. In this case themembers of D can use the Borda method to select the best choice from the set $L$and use it as in the case (b) . If they succeed the procedure is over otherwise they discard that bidder from the set L, choose another bidder from the reduced L (if there is at least one bidder available) and repeat the procedure. If all the attempts fail the procedure of reallocation ends with a failure.

4) *Interaction and support*

In this section we sketch a possible algorithm that can be used in the case where:
- *the bidders are interdependent so that we have, in general, $d_{i,j} \geq 0$ for any $i \neq j \in N$;*
- *$S \neq \emptyset$ so that we have, in general, $D_{i,j} \neq 0$ for any $s_i \in S$ and $j \in N$.*

Mechanism 7.3 Also in this case (see section 7.2) the mechanism has the following structure:
- *possible pre auction phase,*
- *auction phase,*
- *allocation and compensation phase,*
- *reallocation phase.*

The reallocation depends on both the values $d_{i,j}$ (where i and j identify the bidders) and the values $D_{i,j}$ (where i identify the supporters and j identify the bidders). In the current version of the proposed algorithm we assume that the sets B and S can act independently one from the other.

Algorithm 7.2 In this case we can adopt a procedure based on the following steps:
(1) the Bs define the set J_B of suitable bidders as we have seen in section 7.2;
(2) the Ss define the set J_S of suitable bidders as we have seen in section 7.3;
(3) they evaluate the set $J = J_B \cap J_S$;
(4) if $J = \emptyset$ go to 9;
(5) if $J \neq \emptyset$ order J;
(6) select the best b_j from J, $J = J \setminus \{b_j\}$;
(7) b_j is contacted and he is offered a compensation;
(8) b_j can:
(8a) accept so he gets ζ and the compensation; go to 9;
(8b) refuse so that if $J \neq \emptyset$ go to 6 else go to 9;
(9) end;

Observation 7.2 The steps (1) and (2) are simultaneous moves in the sense of Game Theory ([9, 10, 11]). The steps (4) and (8b) define the termination conditions with failure.At the step (8b) the contacted bidder has refused so that, if $J \neq \emptyset$, the mem bers of B and S have another bidder to contact otherwise the procedure must end with a failure. On the other hand, at step (4), if $J = \emptyset$ the procedure neither eff ectively starts since the two sets B and S have no common bidder to whom propose the allocation.At the step (5) the bidders of the set J are ordered[20] from the best to the worst by applying the Borda method to the following preference profiles:
- *the one produced by the members of B over the set J that derives from the ordering on the set J_B;*
- *the one produced by the members of S over the set J that derives from the ordering on the set J_S.*

The use of the Borda method avoids the carrying out of direct comparisons between the evaluations of the bidders through the use of scores that account for the position of each bidder in the corresponding ordering. If the resulting profile contains tied alternatives they can be contacted in any order since they are seen as equivalent from both the members of B and the members of S.

Observation 7.3 At the step (7) it is necessary to collect a sum equal to Σ so that the members of B must collect a sum c_B and the members of S must collect a sum c_S such that:
- *the offer Σ to b_j is enough to compensate him for the allocation of ζ and so together with what the bidders already committed to pay to b_1 is not lower than x_j or $\Sigma \geq x_j - \Sigma_j$;*
- *the sum is proportionally subdivided between the two sets B and S as, respectively:*

$$c_B = \frac{|B|}{|B| + |S|}\Sigma \qquad (38)$$

[20] If |J| = 1 the proposed ordering operation proves obviously useless since there is only one bidder to be contacted.

and:

$$c_S = \frac{|S|}{|B| + |S|}\Sigma \qquad (39)$$

- *the sum* c_B *is to be shared among the members of B proportionally according to ratios:*

$$\frac{d_{i,1} - d_{i,j}}{\sum_{i \neq j}(d_{i,1} - d_{i,j})} \qquad (40)$$

- *the sum cS is to be shared among the members of S proportionally according to ratios:*

$$\frac{D_{i,1} - D_{i,j}}{\sum_{i \neq j}(D_{i,1} - D_{i,j})} \qquad (41)$$

VIII Concluding Remarks and Future Plans

In this paper we presented the structure of a negative auction mechanism under the form of a basic mechanism together with some possible extensions. The extensions include both a pre auction phase and a post auction phase: the first aims at reinforcing the requirement of individual rationality[21] whereas the latter aims at describing possible interactions among the bidders and the supporters. The proposed extensions are still under development so that the full formal characterization is under way. One of the refinement we are planning to introduce, in the case of the interactions among the bidders without supporters (see section 7.2), is the use of pull operations (in addition to the push operations) through which a set of bidders distinct from the current losing bidder can try to pull the allocation of ζ towards other more preferred bidders by sharing among themselves the cost of this switching between bidders.A push operation can, indeed, be executed only by the currently losing bidder so that, if he is satisfied with the allocation, no reallocation is possible though some other bidders may wish to pay him to have the item to be pulled to another and more preferred bidder.Other future plans include the relaxations we have listed in section 4 so that we plan to examine what happens if we assume that:

- *the bidders are risk adverse so that they prefer either to pay the fee or to pay a fixed amount for not getting* ζ *for sure than attending the auction with the risk of getting* ζ *though together with a compensatory sum;*
- *the evaluations are either common or interdependent among the bidders and in any way may vary either after the pre auction phase (if the associated values are common knowledge) or after the auction phase itself if a post auction phase is present;*
- *the bidders are asymmetric so we can have different intervals* $[0, M_i]$ *and different functions* F_i *and* f_i *for each bidder* b_i.

Last but not least we are planning to see what changes we may have in the auction phase if the sum $e_c$is a common knowledge among the bidders before they attend the auction phase.As a first approximation we can expect that if the k attending bidders know the value of m (and so the number of the bidders who paid the fee) they may be willing to bid less than m_i since each of them may consider to have a fixed compensation equal to mf, in case of loss, and so he may wish to increase the probability of losing the auction and such an increase may be obtained by simply bidding less than m_i.

IX References

1) Ken Binmore. Playing fair, Game Theory and the Social Contract. Mas-sachusetts Institute of Technology, 1998.
2) Ken Binmore. Just playing, Game Theory and the Social Contract. Massachusetts Institute of Technology, 1998.
3) Ken Binmore. Rational Decisions. Princeton University Press, 2009.
4) Lorenzo Cioni. An inverse or negative auction. Technical Report TR 10-17, Computer Science Department, November 2010.
5) Simon French. Decision Theory. An introduction to the mathematics of rationality. John Wiley and Sons, 1986.
6) Michael N. Huhns and Larry M. Stephens. Distributed methods for decision support. Multiagent Systems and Societies of Agents, pages 78–120, 2001. editor Gerhard Weiss.
7) Vijay Krishna. Auction Theory. Academic Press, 2002.
8) Paul Milgrom. Putting Auction Theory to Work. Cambridge University Press, 2004.
9) Roger B. Myerson. Game Theory. Analysis of conflict. Harvard Univer sity Press, 1991.
10) Martin J. Osborne and Ariel Rubinstein. A Course in Game Theory. MIT Press, 1994.
11) Fioravate Patrone. Decisori (razionali) interagenti. Una introduzione alla teoria dei giochi. Edizioni plus, 2006.
12) Donald G. Saari. Decisions and Elections, Explaining the Unexpected. Cambridge University Press, 2001.
13) Donald G. Saari. Disposing Dictators, Demistifying Voting Paradoxes. Social Choice Analysis. Cambridge University Press, 2009.
14) Michael Wooldridge. An Introduction to MultiAgent Systems. John Wiley and Sons, 2002.

[21] A mechanism satisfies the property of individual rationality ([3], [7], [9]) if the involved players do not have a negative utility from attending to it and so have some incentives from attending the mechanism.

Global Journals Inc. (US) Guidelines Handbook 2011

www.GlobalJournals.org

FELLOW OF INTERNATIONAL CONGRESS OF COMPUTER SCIENCE AND TECHNOLOGY (FICCT)

- FICCT' title will be awarded to the person after approval of Editor-in-Chief and Editorial Board. The title 'FICCT" can be added to name in the following manner e.g. **Dr. Andrew Knoll, Ph.D., FICCT, Er. Pettor Jone, M.E., FICCT**
- FICCT can submit two papers every year for publication without any charges. The paper will be sent to two peer reviewers. The paper will be published after the acceptance of peer reviewers and Editorial Board.
- Free unlimited Web-space will be allotted to 'FICCT 'along with subDomain to contribute and partake in our activities.
- A professional email address will be allotted free with unlimited email space.
- FICCT will be authorized to receive e-Journals - GJCST for the Lifetime.
- FICCT will be exempted from the registration fees of Seminar/Symposium/Conference/Workshop conducted internationally of GJCST (FREE of Charge).
- FICCT will be an Honorable Guest of any gathering hold.

ASSOCIATE OF INTERNATIONAL CONGRESS OF COMPUTER SCIENCE AND TECHNOLOGY (AICCT)

- AICCT title will be awarded to the person/institution after approval of Editor-in-Chef and Editorial Board. The title 'AICCTcan be added to name in the following manner:
  eg. **Dr. Thomas Herry, Ph.D., AICCT**
- AICCT can submit one paper every year for publication without any charges. The paper will be sent to two peer reviewers. The paper will be published after the acceptance of peer reviewers and Editorial Board.
- Free 2GB Web-space will be allotted to 'FICCT' along with subDomain to contribute and participate in our activities.
- A professional email address will be allotted with free 1GB email space.
- AICCT will be authorized to receive e-Journal GJCST for lifetime.
- A professional email address will be allotted with free 1GB email space.
- AICHSS will be authorized to receive e-Journal GJHSS for lifetime.

ANNUAL MEMBER

- Annual Member will be authorized to receive e-Journal GJCST for one year (subscription for one year).
- The member will be allotted free 1 GB Web-space along with subDomain to contribute and participate in our activities.
- A professional email address will be allotted free 500 MB email space.

PAPER PUBLICATION

- The members can publish paper once. The paper will be sent to two-peer reviewer. The paper will be published after the acceptance of peer reviewers and Editorial Board.

Process of submission of Research Paper

The Area or field of specialization may or may not be of any category as mentioned in 'Scope of Journal' menu of the GlobalJournals.org website. There are 37 Research Journal categorized with Six parental Journals GJCST, GJMR, GJRE, GJMBR, GJSFR, GJHSS. For Authors should prefer the mentioned categories. There are three widely used systems UDC, DDC and LCC. The details are available as 'Knowledge Abstract' at Home page. The major advantage of this coding is that, the research work will be exposed to and shared with all over the world as we are being abstracted and indexed worldwide.

The paper should be in proper format. The format can be downloaded from first page of 'Author Guideline' Menu. The Author is expected to follow the general rules as mentioned in this menu. The paper should be written in MS-Word Format (*.DOC,*.DOCX).

The Author can submit the paper either online or offline. The authors should prefer online submission.

Online Submission: There are three ways to submit your paper:

(A) (I) Register yourself using top right corner of Home page then Login from same place twice. If you are already registered, then login using your username and password.

(II) Choose corresponding Journal from "Research Journals" Menu.

(III) Click 'Submit Manuscript'. Fill required information and Upload the paper.

(B) If you are using Internet Explorer (Although Mozilla Firefox is preferred), then Direct Submission through Homepage is also available.

(C) If these two are not convenient, and then email the paper directly to dean@globaljournals.org as an attachment.

Offline Submission: Author can send the typed form of paper by Post. However, online submission should be preferred.

Preferred Author Guidelines

MANUSCRIPT STYLE INSTRUCTION (Must be strictly followed)

Page Size: 8.27" X 11'"

- Left Margin: 0.65
- Right Margin: 0.65
- Top Margin: 0.75
- Bottom Margin: 0.75
- Font type of all text should be Times New Roman.
- Paper Title should be of Font Size 24 with one Column section.
- Author Name in Font Size of 11 with one column as of Title.
- Abstract Font size of 9 Bold, "Abstract" word in Italic Bold.
- Main Text: Font size 10 with justified two columns section
- Two Column with Equal Column with of 3.38 and Gaping of .2
- First Character must be two lines Drop capped.
- Paragraph before Spacing of 1 pt and After of 0 pt.
- Line Spacing of 1 pt
- Large Images must be in One Column
- Numbering of First Main Headings (Heading 1) must be in Roman Letters, Capital Letter, and Font Size of 10.
- Numbering of Second Main Headings (Heading 2) must be in Alphabets, Italic, and Font Size of 10.

You can use your own standard format also.

Author Guidelines:

1. General,

2. Ethical Guidelines,

3. Submission of Manuscripts,

4. Manuscript's Category,

5. Structure and Format of Manuscript,

6. After Acceptance.

1. GENERAL

Before submitting your research paper, one is advised to go through the details as mentioned in following heads. It will be beneficial, while peer reviewer justify your paper for publication.

Scope

The Global Journals Inc. (US) welcome the submission of original paper, review paper, survey article relevant to the all the streams of Philosophy and knowledge. The Global Journals Inc. (US) is parental platform for Global Journal of Computer Science and Technology, Researches in Engineering, Medical Research, Science Frontier Research, Human Social Science, Management, and Business organization. The choice of specific field can be done otherwise as following in Abstracting and Indexing Page on this Website. As the all Global

Journals Inc. (US) are being abstracted and indexed (in process) by most of the reputed organizations. Topics of only narrow interest will not be accepted unless they have wider potential or consequences.

2. ETHICAL GUIDELINES

Authors should follow the ethical guidelines as mentioned below for publication of research paper and research activities.

Papers are accepted on strict understanding that the material in whole or in part has not been, nor is being, considered for publication elsewhere. If the paper once accepted by Global Journals Inc. (US) and Editorial Board, will become the copyright of the Global Journals Inc. (US).

Authorship: The authors and coauthors should have active contribution to conception design, analysis and interpretation of findings. They should critically review the contents and drafting of the paper. All should approve the final version of the paper before submission

The Global Journals Inc. (US) follows the definition of authorship set up by the Global Academy of Research and Development. According to the Global Academy of R&D authorship, criteria must be based on:

1) Substantial contributions to conception and acquisition of data, analysis and interpretation of the findings.

2) Drafting the paper and revising it critically regarding important academic content.

3) Final approval of the version of the paper to be published.

All authors should have been credited according to their appropriate contribution in research activity and preparing paper. Contributors who do not match the criteria as authors may be mentioned under Acknowledgement.

Acknowledgements: Contributors to the research other than authors credited should be mentioned under acknowledgement. The specifications of the source of funding for the research if appropriate can be included. Suppliers of resources may be mentioned along with address.

Appeal of Decision: The Editorial Board's decision on publication of the paper is final and cannot be appealed elsewhere.

Permissions: It is the author's responsibility to have prior permission if all or parts of earlier published illustrations are used in this paper.

Please mention proper reference and appropriate acknowledgements wherever expected.

If all or parts of previously published illustrations are used, permission must be taken from the copyright holder concerned. It is the author's responsibility to take these in writing.

Approval for reproduction/modification of any information (including figures and tables) published elsewhere must be obtained by the authors/copyright holders before submission of the manuscript. Contributors (Authors) are responsible for any copyright fee involved.

3. SUBMISSION OF MANUSCRIPTS

Manuscripts should be uploaded via this online submission page. The online submission is most efficient method for submission of papers, as it enables rapid distribution of manuscripts and consequently speeds up the review procedure. It also enables authors to know the status of their own manuscripts by emailing us. Complete instructions for submitting a paper is available below.

Manuscript submission is a systematic procedure and little preparation is required beyond having all parts of your manuscript in a given format and a computer with an Internet connection and a Web browser. Full help and instructions are provided on-screen. As an author, you will be prompted for login and manuscript details as Field of Paper and then to upload your manuscript file(s) according to the instructions.

To avoid postal delays, all transaction is preferred by e-mail. A finished manuscript submission is confirmed by e-mail immediately and your paper enters the editorial process with no postal delays. When a conclusion is made about the publication of your paper by our Editorial Board, revisions can be submitted online with the same procedure, with an occasion to view and respond to all comments.

Complete support for both authors and co-author is provided.

4. MANUSCRIPT'S CATEGORY

Based on potential and nature, the manuscript can be categorized under the following heads: Original research paper: Such papers are reports of high-level significant original research work.

Review papers: These are concise, significant but helpful and decisive topics for young researchers.

Research articles: These are handled with small investigation and applications

Research letters: The letters are small and concise comments on previously published matters.

5. STRUCTURE AND FORMAT OF MANUSCRIPT

The recommended size of original research paper is less than seven thousand words, review papers fewer than seven thousands words also. Preparation of research paper or how to write research paper, are major hurdle, while writing manuscript. The research articles and research letters should be fewer than three thousand words, the structure original research paper; sometime review paper should be as follows:

Papers: These are reports of significant research (typically less than 7000 words equivalent, including tables, figures, references), and comprise:

(a)Title should be relevant and commensurate with the theme of the paper.

(b) A brief Summary, "Abstract" (less than 150 words) containing the major results and conclusions.

(c) Up to ten keywords, that precisely identifies the paper's subject, purpose, and focus.

(d) An Introduction, giving necessary background excluding subheadings; objectives must be clearly declared.

(e) Resources and techniques with sufficient complete experimental details (wherever possible by reference) to permit repetition; sources of information must be given and numerical methods must be specified by reference, unless non-standard.

(f) Results should be presented concisely, by well-designed tables and/or figures; the same data may not be used in both; suitable statistical data should be given. All data must be obtained with attention to numerical detail in the planning stage. As reproduced design has been recognized to be important to experiments for a considerable time, the Editor has decided that any paper that appears not to have adequate numerical treatments of the data will be returned un-refereed;

(g) Discussion should cover the implications and consequences, not just recapitulating the results; conclusions should be summarizing.

(h) Brief Acknowledgements.

(i) References in the proper form.

Authors should very cautiously consider the preparation of papers to ensure that they communicate efficiently. Papers are much more likely to be accepted, if they are cautiously designed and laid out, contain few or no errors, are summarizing, and be conventional to the approach and instructions. They will in addition, be published with much less delays than those that require much technical and editorial correction.

The Editorial Board reserves the right to make literary corrections and to make suggestions to improve briefness.

It is vital, that authors take care in submitting a manuscript that is written in simple language and adheres to published guidelines.

Format

Language: The language of publication is UK English. Authors, for whom English is a second language, must have their manuscript efficiently edited by an English-speaking person before submission to make sure that, the English is of high excellence. It is preferable, that manuscripts should be professionally edited.

Standard Usage, Abbreviations, and Units: Spelling and hyphenation should be conventional to The Concise Oxford English Dictionary. Statistics and measurements should at all times be given in figures, e.g. 16 min, except for when the number begins a sentence. When the number does not refer to a unit of measurement it should be spelt in full unless, it is 160 or greater.

Abbreviations supposed to be used carefully. The abbreviated name or expression is supposed to be cited in full at first usage, followed by the conventional abbreviation in parentheses.

Metric SI units are supposed to generally be used excluding where they conflict with current practice or are confusing. For illustration, 1.4 l rather than 1.4 × 10-3 m3, or 4 mm somewhat than 4 × 10-3 m. Chemical formula and solutions must identify the form used, e.g. anhydrous or hydrated, and the concentration must be in clearly defined units. Common species names should be followed by underlines at the first mention. For following use the generic name should be constricted to a single letter, if it is clear.

Structure

All manuscripts submitted to Global Journals Inc. (US), ought to include:
Title: The title page must carry an instructive title that reflects the content, a running title (less than 45 characters together with spaces), names of the authors and co-authors, and the place(s) wherever the work was carried out. The full postal address in addition with the e-mail address of related author must be given. Up to eleven keywords or very brief phrases have to be given to help data retrieval, mining and indexing.
Abstract, used in Original Papers and Reviews:
Optimizing Abstract for Search Engines
Many researchers searching for information online will use search engines such as Google, Yahoo or similar. By optimizing your paper for search engines, you will amplify the chance of someone finding it. This in turn will make it more likely to be viewed and/or cited in a further work. Global Journals Inc. (US) have compiled these guidelines to facilitate you to maximize the web-friendliness of the most public part of your paper.

Key Words
A major linchpin in research work for the writing research paper is the keyword search, which one will employ to find both library and Internet resources.
One must be persistent and creative in using keywords. An effective keyword search requires a strategy and planning a list of possible keywords and phrases to try.

Search engines for most searches, use Boolean searching, which is somewhat different from Internet searches. The Boolean search uses "operators," words (and, or, not, and near) that enable you to expand or narrow your affords. Tips for research paper while preparing research paper are very helpful guideline of research paper.

Choice of key words is first tool of tips to write research paper. Research paper writing is an art.A few tips for deciding as strategically as possible about keyword search:

- One should start brainstorming lists of possible keywords before even begin searching. Think about the most important concepts related to research work. Ask, "What words would a source have to include to be truly valuable in research paper?" Then consider synonyms for the important words.
- It may take the discovery of only one relevant paper to let steer in the right keyword direction because in most databases, the keywords under which a research paper is abstracted are listed with the paper.
- One should avoid outdated words.

Keywords are the key that opens a door to research work sources. Keyword searching is an art in which researcher's skills are bound to improve with experience and time.

Numerical Methods: Numerical methods used should be clear and, where appropriate, supported by references.

Acknowledgements: Please make these as concise as possible.

References

References follow the Harvard scheme of referencing. References in the text should cite the authors' names followed by the time of their publication, unless there are three or more authors when simply the first author's name is quoted followed by et al. unpublished work has to only be cited where necessary, and only in the text. Copies of references in press in other journals have to be supplied with submitted typescripts. It is necessary that all citations and references be carefully checked before submission, as mistakes or omissions will cause delays.

References to information on the World Wide Web can be given, but only if the information is available without charge to readers on an official site. Wikipedia and Similar websites are not allowed where anyone can change the information. Authors will be asked to make available electronic copies of the cited information for inclusion on the Global Journals Inc. (US) homepage at the judgment of the Editorial Board.

The Editorial Board and Global Journals Inc. (US) recommend that, citation of online-published papers and other material should be done via a DOI (digital object identifier). If an author cites anything, which does not have a DOI, they run the risk of the cited material not being noticeable.

The Editorial Board and Global Journals Inc. (US) recommend the use of a tool such as Reference Manager for reference management and formatting.

Tables, Figures and Figure Legends

Tables: Tables should be few in number, cautiously designed, uncrowned, and include only essential data. Each must have an Arabic number, e.g. Table 4, a self-explanatory caption and be on a separate sheet. Vertical lines should not be used.

Figures: Figures are supposed to be submitted as separate files. Always take in a citation in the text for each figure using Arabic numbers, e.g. Fig. 4. Artwork must be submitted online in electronic form by e-mailing them.

Preparation of Electronic Figures for Publication

Even though low quality images are sufficient for review purposes, print publication requires high quality images to prevent the final product being blurred or fuzzy. Submit (or e-mail) EPS (line art) or TIFF (halftone/photographs) files only. MS PowerPoint and Word Graphics are unsuitable for printed pictures. Do not use pixel-oriented software. Scans (TIFF only) should have a resolution of at least 350 dpi (halftone) or 700 to 1100 dpi (line drawings) in relation to the imitation size. Please give the data for figures in black and white or submit a Color Work Agreement Form. EPS files must be saved with fonts embedded (and with a TIFF preview, if possible).

For scanned images, the scanning resolution (at final image size) ought to be as follows to ensure good reproduction: line art: >650 dpi; halftones (including gel photographs) : >350 dpi; figures containing both halftone and line images: >650 dpi.

Color Charges: It is the rule of the Global Journals Inc. (US) for authors to pay the full cost for the reproduction of their color artwork. Hence, please note that, if there is color artwork in your manuscript when it is accepted for publication, we would require you to complete and return a color work agreement form before your paper can be published.

Figure Legends: Self-explanatory legends of all figures should be incorporated separately under the heading 'Legends to Figures'. In the full-text online edition of the journal, figure legends may possibly be truncated in abbreviated links to the full screen version. Therefore, the first 100 characters of any legend should notify the reader, about the key aspects of the figure.

6. AFTER ACCEPTANCE

Upon approval of a paper for publication, the manuscript will be forwarded to the dean, who is responsible for the publication of the Global Journals Inc. (US).

6.1 Proof Corrections

The corresponding author will receive an e-mail alert containing a link to a website or will be attached. A working e-mail address must therefore be provided for the related author.

Acrobat Reader will be required in order to read this file. This software can be downloaded

(Free of charge) from the following website:

www.adobe.com/products/acrobat/readstep2.html. This will facilitate the file to be opened, read on screen, and printed out in order for any corrections to be added. Further instructions will be sent with the proof.

Proofs must be returned to the dean at dean@globaljournals.org within three days of receipt.

As changes to proofs are costly, we inquire that you only correct typesetting errors. All illustrations are retained by the publisher. Please note that the authors are responsible for all statements made in their work, including changes made by the copy editor.

6.2 Early View of Global Journals Inc. (US) (Publication Prior to Print)

The Global Journals Inc. (US) are enclosed by our publishing's Early View service. Early View articles are complete full-text articles sent in advance of their publication. Early View articles are absolute and final. They have been completely reviewed, revised and edited for publication, and the authors' final corrections have been incorporated. Because they are in final form, no changes can be made after sending them. The nature of Early View articles means that they do not yet have volume, issue or page numbers, so Early View articles cannot be cited in the conventional way.

6.3 Author Services

Online production tracking is available for your article through Author Services. Author Services enables authors to track their article - once it has been accepted - through the production process to publication online and in print. Authors can check the status of their articles online and choose to receive automated e-mails at key stages of production. The authors will receive an e-mail with a unique link that enables them to register and have their article automatically added to the system. Please ensure that a complete e-mail address is provided when submitting the manuscript.

6.4 Author Material Archive Policy

Please note that if not specifically requested, publisher will dispose off hardcopy & electronic information submitted, after the two months of publication. If you require the return of any information submitted, please inform the Editorial Board or dean as soon as possible.

6.5 Offprint and Extra Copies

A PDF offprint of the online-published article will be provided free of charge to the related author, and may be distributed according to the Publisher's terms and conditions. Additional paper offprint may be ordered by emailing us at: editor@globaljournals.org .

Informal Tips for writing a Computer Science Research Paper to increase readability and citation

Before start writing a good quality Computer Science Research Paper, let us first understand what is Computer Science Research Paper? So, Computer Science Research Paper is the paper which is written by professionals or scientists who are associated to Computer Science and Information Technology, or doing research study in these areas. If you are novel to this field then you can consult about this field from your supervisor or guide.

Techniques for writing a good quality Computer Science Research Paper:

1. Choosing the topic- In most cases, the topic is searched by the interest of author but it can be also suggested by the guides. You can have several topics and then you can judge that in which topic or subject you are finding yourself most comfortable. This can be done by asking several questions to yourself, like Will I be able to carry our search in this area? Will I find all necessary recourses to accomplish

the search? Will I be able to find all information in this field area? If the answer of these types of questions will be "Yes" then you can choose that topic. In most of the cases, you may have to conduct the surveys and have to visit several places because this field is related to Computer Science and Information Technology. Also, you may have to do a lot of work to find all rise and falls regarding the various data of that subject. Sometimes, detailed information plays a vital role, instead of short information.

2. Evaluators are human: First thing to remember that evaluators are also human being. They are not only meant for rejecting a paper. They are here to evaluate your paper. So, present your Best.

3. Think Like Evaluators: If you are in a confusion or getting demotivated that your paper will be accepted by evaluators or not, then think and try to evaluate your paper like an Evaluator. Try to understand that what an evaluator wants in your research paper and automatically you will have your answer.

4. Make blueprints of paper: The outline is the plan or framework that will help you to arrange your thoughts. It will make your paper logical. But remember that all points of your outline must be related to the topic you have chosen.

5. Ask your Guides: If you are having any difficulty in your research, then do not hesitate to share your difficulty to your guide (if you have any). They will surely help you out and resolve your doubts. If you can't clarify what exactly you require for your work then ask the supervisor to help you with the alternative. He might also provide you the list of essential readings.

6. Use of computer is recommended: As you are doing research in the field of Computer Science, then this point is quite obvious.

7. Use right software: Always use good quality software packages. If you are not capable to judge good software then you can lose quality of your paper unknowingly. There are various software programs available to help you, which you can get through Internet.

8. Use the Internet for help: An excellent start for your paper can be by using the Google. It is an excellent search engine, where you can have your doubts resolved. You may also read some answers for the frequent question how to write my research paper or find model research paper. From the internet library you can download books. If you have all required books make important reading selecting and analyzing the specified information. Then put together research paper sketch out.

9. Use and get big pictures: Always use encyclopedias, Wikipedia to get pictures so that you can go into the depth.

10. Bookmarks are useful: When you read any book or magazine, you generally use bookmarks, right! It is a good habit, which helps to not to lose your continuity. You should always use bookmarks while searching on Internet also, which will make your search easier.

11. Revise what you wrote: When you write anything, always read it, summarize it and then finalize it.

12. Make all efforts: Make all efforts to mention what you are going to write in your paper. That means always have a good start. Try to mention everything in introduction, that what is the need of a particular research paper. Polish your work by good skill of writing and always give an evaluator, what he wants.

13. Have backups: When you are going to do any important thing like making research paper, you should always have backup copies of it either in your computer or in paper. This will help you to not to lose any of your important.

14. Produce good diagrams of your own: Always try to include good charts or diagrams in your paper to improve quality. Using several and unnecessary diagrams will degrade the quality of your paper by creating "hotchpotch." So always, try to make and include those diagrams, which are made by your own to improve readability and understandability of your paper.

15. Use of direct quotes: When you do research relevant to literature, history or current affairs then use of quotes become essential but if study is relevant to science then use of quotes is not preferable.

16. Use proper verb tense: Use proper verb tenses in your paper. Use past tense, to present those events that happened. Use present tense to indicate events that are going on. Use future tense to indicate future happening events. Use of improper and wrong tenses will confuse the evaluator. Avoid the sentences that are incomplete.

17. Never use online paper: If you are getting any paper on Internet, then never use it as your research paper because it might be possible that evaluator has already seen it or maybe it is outdated version.

18. **Pick a good study spot:** To do your research studies always try to pick a spot, which is quiet. Every spot is not for studies. Spot that suits you choose it and proceed further.

19. Know what you know: Always try to know, what you know by making objectives. Else, you will be confused and cannot achieve your target.

20. Use good quality grammar: Always use a good quality grammar and use words that will throw positive impact on evaluator. Use of good quality grammar does not mean to use tough words, that for each word the evaluator has to go through dictionary. Do not start sentence with a conjunction. Do not fragment sentences. Eliminate one-word sentences. Ignore passive voice. Do not ever use a big word when a diminutive one would suffice. Verbs have to be in agreement with their subjects. Prepositions are not expressions to finish sentences with. It is incorrect to ever divide an infinitive. Avoid clichés like the disease. Also, always shun irritating alliteration. Use language that is simple and straight forward. put together a neat summary.

21. Arrangement of information: Each section of the main body should start with an opening sentence and there should be a changeover at the end of the section. Give only valid and powerful arguments to your topic. You may also maintain your arguments with records.

22. Never start in last minute: Always start at right time and give enough time to research work. Leaving everything to the last minute will degrade your paper and spoil your work.

23. Multitasking in research is not good: Doing several things at the same time proves bad habit in case of research activity. Research is an area, where everything has a particular time slot. Divide your research work in parts and do particular part in particular time slot.

24. Never copy others' work: Never copy others' work and give it your name because if evaluator has seen it anywhere you will be in trouble.

25. Take proper rest and food: No matter how many hours you spend for your research activity, if you are not taking care of your health then all your efforts will be in vain. For a quality research, study is must, and this can be done by taking proper rest and food.

26. Go for seminars: Attend seminars if the topic is relevant to your research area. Utilize all your resources.

27. Refresh your mind after intervals: Try to give rest to your mind by listening to soft music or by sleeping in intervals. This will also improve your memory.

28. Make colleagues: Always try to make colleagues. No matter how sharper or intelligent you are, if you make colleagues you can have several ideas, which will be helpful for your research.

29. **Think technically:** Always think technically. If anything happens, then search its reasons, its benefits, and demerits.

30. Think and then print: When you will go to print your paper, notice that tables are not be split, headings are not detached from their descriptions, and page sequence is maintained.

31. Adding unnecessary information: Do not add unnecessary information, like, I have used MS Excel to draw graph. Do not add irrelevant and inappropriate material. These all will create superfluous. Foreign terminology and phrases are not apropos. One should NEVER take a broad view. Analogy in script is like feathers on a snake. Not at all use a large word when a very small one would be

sufficient. Use words properly, regardless of how others use them. Remove quotations. Puns are for kids, not grunt readers. Amplification is a billion times of inferior quality than sarcasm.

32. Never oversimplify everything: To add material in your research paper, never go for oversimplification. This will definitely irritate the evaluator. Be more or less specific. Also too, by no means, ever use rhythmic redundancies. Contractions aren't essential and shouldn't be there used. Comparisons are as terrible as clichés. Give up ampersands and abbreviations, and so on. Remove commas, that are, not necessary. Parenthetical words however should be together with this in commas. Understatement is all the time the complete best way to put onward earth-shaking thoughts. Give a detailed literary review.

33. Report concluded results: Use concluded results. From raw data, filter the results and then conclude your studies based on measurements and observations taken. Significant figures and appropriate number of decimal places should be used. Parenthetical remarks are prohibitive. Proofread carefully at final stage. In the end give outline to your arguments. Spot out perspectives of further study of this subject. Justify your conclusion by at the bottom of them with sufficient justifications and examples.

34. After conclusion: Once you have concluded your research, the next most important step is to present your findings. Presentation is extremely important as it is the definite medium though which your research is going to be in print to the rest of the crowd. Care should be taken to categorize your thoughts well and present them in a logical and neat manner. A good quality research paper format is essential because it serves to highlight your research paper and bring to light all necessary aspects in your research.

Informal Guidelines of Research Paper Writing

Key points to remember:

- Submit all work in its final form.
- Write your paper in the form, which is presented in the guidelines using the template.
- Please note the criterion for grading the final paper by peer-reviewers.

Final Points:

A purpose of organizing a research paper is to let people to interpret your effort selectively. The journal requires the following sections, submitted in the order listed, each section to start on a new page.

The introduction will be compiled from reference matter and will reflect the design processes or outline of basis that direct you to make study. As you will carry out the process of study, the method and process section will be constructed as like that. The result segment will show related statistics in nearly sequential order and will direct the reviewers next to the similar intellectual paths throughout the data that you took to carry out your study. The discussion section will provide understanding of the data and projections as to the implication of the results. The use of good quality references all through the paper will give the effort trustworthiness by representing an alertness of prior workings.

Writing a research paper is not an easy job no matter how trouble-free the actual research or concept. Practice, excellent preparation, and controlled record keeping are the only means to make straightforward the progression.

General style:

Specific editorial column necessities for compliance of a manuscript will always take over from directions in these general guidelines.

To make a paper clear

· Adhere to recommended page limits

Mistakes to evade

- Insertion a title at the foot of a page with the subsequent text on the next page

- Separating a table/chart or figure - impound each figure/table to a single page
- Submitting a manuscript with pages out of sequence

In every sections of your document

· Use standard writing style including articles ("a", "the," etc.)

· Keep on paying attention on the research topic of the paper

· Use paragraphs to split each significant point (excluding for the abstract)

· Align the primary line of each section

· Present your points in sound order

· Use present tense to report well accepted

· Use past tense to describe specific results

· Shun familiar wording, don't address the reviewer directly, and don't use slang, slang language, or superlatives

· Shun use of extra pictures - include only those figures essential to presenting results

Title Page:

Choose a revealing title. It should be short. It should not have non-standard acronyms or abbreviations. It should not exceed two printed lines. It should include the name(s) and address (es) of all authors.

Abstract:

The summary should be two hundred words or less. It should briefly and clearly explain the key findings reported in the manuscript--must have precise statistics. It should not have abnormal acronyms or abbreviations. It should be logical in itself. Shun citing references at this point.

An abstract is a brief distinct paragraph summary of finished work or work in development. In a minute or less a reviewer can be taught the foundation behind the study, common approach to the problem, relevant results, and significant conclusions or new questions.

Write your summary when your paper is completed because how can you write the summary of anything which is not yet written? Wealth of terminology is very essential in abstract. Yet, use comprehensive sentences and do not let go readability for briefness. You can maintain it succinct by phrasing sentences so that they provide more than lone rationale. The author can at this moment go straight to

shortening the outcome. Sum up the study, with the subsequent elements in any summary. Try to maintain the initial two items to no more than one ruling each.

- Reason of the study - theory, overall issue, purpose
- Fundamental goal
- To the point depiction of the research
- Consequences, including definite statistics - if the consequences are quantitative in nature, account quantitative data; results of any numerical analysis should be reported
- Significant conclusions or questions that track from the research(es)

Approach:

- Single section, and succinct
- As a outline of job done, it is always written in past tense
- A conceptual should situate on its own, and not submit to any other part of the paper such as a form or table
- Center on shortening results - bound background information to a verdict or two, if completely necessary
- What you account in an conceptual must be regular with what you reported in the manuscript
- Exact spelling, clearness of sentences and phrases, and appropriate reporting of quantities (proper units, important statistics) are just as significant in an abstract as they are anywhere else

Introduction:

The **Introduction** should "introduce" the manuscript. The reviewer should be presented with sufficient background information to be capable to comprehend and calculate the purpose of your study without having to submit to other works. The basis for the study should be offered. Give most important references but shun difficult to make a comprehensive appraisal of the topic. In the introduction, describe the problem visibly. If the problem is not acknowledged in a logical, reasonable way, the reviewer will have no attention in your result. Speak in common terms about techniques used to explain the problem, if needed, but do not present any particulars about the protocols here. Following approach can create a valuable beginning:

- Explain the value (significance) of the study
- Shield the model - why did you employ this particular system or method? What is its compensation? You strength remark on its appropriateness from a abstract point of vision as well as point out sensible reasons for using it.
- Present a justification. Status your particular theory (es) or aim(s), and describe the logic that led you to choose them.
- Very for a short time explain the tentative propose and how it skilled the declared objectives.

Approach:

- Use past tense except for when referring to recognized facts. After all, the manuscript will be submitted after the entire job is done.
- Sort out your thoughts; manufacture one key point with every section. If you make the four points listed above, you will need a least of four paragraphs.
- Present surroundings information only as desirable in order hold up a situation. The reviewer does not desire to read the whole thing you know about a topic.
- Shape the theory/purpose specifically - do not take a broad view.
- As always, give awareness to spelling, simplicity and correctness of sentences and phrases.

Procedures (Methods and Materials):

This part is supposed to be the easiest to carve if you have good skills. A sound written Procedures segment allows a capable scientist to replacement your results. Present precise information about your supplies. The suppliers and clarity of reagents can be helpful bits of information. Present methods in sequential order but linked methodologies can be grouped as a segment. Be concise when relating the protocols. Attempt for the least amount of information that would permit another capable scientist to spare your outcome but be cautious that vital information is integrated. The use of subheadings is suggested and ought to be synchronized with the results section. When a technique is used that has been well described in another object, mention the specific item describing a way but draw the basic

principle while stating the situation. The purpose is to text all particular resources and broad procedures, so that another person may use some or all of the methods in one more study or referee the scientific value of your work. It is not to be a step by step report of the whole thing you did, nor is a methods section a set of orders.

Materials:

- Explain materials individually only if the study is so complex that it saves liberty this way.
- Embrace particular materials, and any tools or provisions that are not frequently found in laboratories.
- Do not take in frequently found.
- If use of a definite type of tools.
- Materials may be reported in a part section or else they may be recognized along with your measures.

Methods:

- Report the method (not particulars of each process that engaged the same methodology)
- Describe the method entirely
- To be succinct, present methods under headings dedicated to specific dealings or groups of measures
- Simplify - details how procedures were completed not how they were exclusively performed on a particular day.
- If well known procedures were used, account the procedure by name, possibly with reference, and that's all.

Approach:

- It is embarrassed or not possible to use vigorous voice when documenting methods with no using first person, which would focus the reviewer's interest on the researcher rather than the job. As a result when script up the methods most authors use third person passive voice.
- Use standard style in this and in every other part of the paper - avoid familiar lists, and use full sentences.

What to keep away from

- Resources and methods are not a set of information.
- Skip all descriptive information and surroundings - save it for the argument.
- Leave out information that is immaterial to a third party.

Results:

The principle of a results segment is to present and demonstrate your conclusion. Create this part a entirely objective details of the outcome, and save all understanding for the discussion.

The page length of this segment is set by the sum and types of data to be reported. Carry on to be to the point, by means of statistics and tables, if suitable, to present consequences most efficiently.You must obviously differentiate material that would usually be incorporated in a study editorial from any unprocessed data or additional appendix matter that would not be available. In fact, such matter should not be submitted at all except requested by the instructor.

Content

- Sum up your conclusion in text and demonstrate them, if suitable, with figures and tables.
- In manuscript, explain each of your consequences, point the reader to remarks that are most appropriate.
- Present a background, such as by describing the question that was addressed by creation an exacting study.
- Explain results of control experiments and comprise remarks that are not accessible in a prescribed figure or table, if appropriate.
- Examine your data, then prepare the analyzed (transformed) data in the form of a figure (graph), table, or in manuscript form.

What to stay away from

- Do not discuss or infer your outcome, report surroundings information, or try to explain anything.
- Not at all, take in raw data or intermediate calculations in a research manuscript.

- Do not present the similar data more than once.
- Manuscript should complement any figures or tables, not duplicate the identical information.
- Never confuse figures with tables - there is a difference.

Approach

- As forever, use past tense when you submit to your results, and put the whole thing in a reasonable order.
- Put figures and tables, appropriately numbered, in order at the end of the report
- If you desire, you may place your figures and tables properly within the text of your results part.

Figures and tables

- If you put figures and tables at the end of the details, make certain that they are visibly distinguished from any attach appendix materials, such as raw facts
- Despite of position, each figure must be numbered one after the other and complete with subtitle
- In spite of position, each table must be titled, numbered one after the other and complete with heading
- All figure and table must be adequately complete that it could situate on its own, divide from text

Discussion:

The Discussion is expected the trickiest segment to write and describe. A lot of papers submitted for journal are discarded based on problems with the Discussion. There is no head of state for how long a argument should be. Position your understanding of the outcome visibly to lead the reviewer through your conclusions, and then finish the paper with a summing up of the implication of the study. The purpose here is to offer an understanding of your results and hold up for all of your conclusions, using facts from your research and generally accepted information, if suitable. The implication of result should be visibly described. Infer your data in the conversation in suitable depth. This means that when you clarify an observable fact you must explain mechanisms that may account for the observation. If your results vary from your prospect, make clear why that may have happened. If your results agree, then explain the theory that the proof supported. It is never suitable to just state that the data approved with prospect, and let it drop at that.

- Make a decision if each premise is supported, discarded, or if you cannot make a conclusion with assurance. Do not just dismiss a study or part of a study as "uncertain."
- Research papers are not acknowledged if the work is imperfect. Draw what conclusions you can based upon the results that you have, and take care of the study as a finished work
- You may propose future guidelines, such as how the experiment might be personalized to accomplish a new idea.
- Give details all of your remarks as much as possible, focus on mechanisms.
- Make a decision if the tentative design sufficiently addressed the theory, and whether or not it was correctly restricted.
- Try to present substitute explanations if sensible alternatives be present.
- One research will not counter an overall question, so maintain the large picture in mind, where do you go next? The best studies unlock new avenues of study. What questions remain?
- Recommendations for detailed papers will offer supplementary suggestions.

Approach:

- When you refer to information, differentiate data generated by your own studies from available information
- Submit to work done by specific persons (including you) in past tense.
- Submit to generally acknowledged facts and main beliefs in present tense.

Administration Rules Listed Before Submitting Your Research Paper to Global Journals Inc. (US)

Please carefully note down following rules and regulation before submitting your Research Paper to Global Journals Inc. (US):

Segment Draft and Final Research Paper: You have to strictly follow the template of research paper. If it is not done your paper may get rejected.

- The **major constraint** is that you must independently make all content, tables, graphs, and facts that are offered in the paper. You must write each part of the paper wholly on your own. The Peer-reviewers need to identify your own perceptive of the concepts in your own terms. NEVER extract straight from any foundation, and never rephrase someone else's analysis.

- Do not give permission to anyone else to "PROOFREAD" your manuscript.

- Methods to avoid Plagiarism is applied by us on every paper, if found guilty, you will be blacklisted by all of our collaborated research groups, your institution will be informed for this and strict legal actions will be taken immediately.)
- To guard yourself and others from possible illegal use please do not permit anyone right to use to your paper and files.

Criterion for Grading a Research Paper (Compilation) by Global Journals Inc. (US)

Please note that following table is only a Grading of "Paper Compilation" and not on "Performed/Stated Research" whose grading solely depends on Individual Assigned Peer Reviewer and Editorial Board Member. These can be available only on request and after decision of Paper. This report will be the property of Global Journals Inc. (US).

Topics	Grades		
	A-B	**C-D**	**E-F**
Abstract	Clear and concise with appropriate content, Correct format. 200 words or below	Unclear summary and no specific data, Incorrect form Above 200 words	No specific data with ambiguous information Above 250 words
Introduction	Containing all background details with clear goal and appropriate details, flow specification, no grammar and spelling mistake, well organized sentence and paragraph, reference cited	Unclear and confusing data, appropriate format, grammar and spelling errors with unorganized matter	Out of place depth and content, hazy format
Methods and Procedures	Clear and to the point with well arranged paragraph, precision and accuracy of facts and figures, well organized subheads	Difficult to comprehend with embarrassed text, too much explanation but completed	Incorrect and unorganized structure with hazy meaning
Result	Well organized, Clear and specific, Correct units with precision, correct data, well structuring of paragraph, no grammar and spelling mistake	Complete and embarrassed text, difficult to comprehend	Irregular format with wrong facts and figures
Discussion	Well organized, meaningful specification, sound conclusion, logical and concise explanation, highly structured paragraph reference cited	Wordy, unclear conclusion, spurious	Conclusion is not cited, unorganized, difficult to comprehend
References	Complete and correct format, well organized	Beside the point, Incomplete	Wrong format and structuring

Index

save our planet

Global Journal of Computer Science and Technology

Visit us on the Web at www.GlobalJournals.org | www.ComputerResearch.org
or email us at helpdesk@globaljournals.org

ISSN 9754350