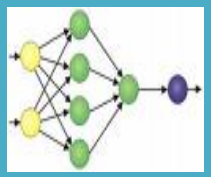
ISSN – 0974 - 1925

JOURNAL OF COMPUTER APPLICATIONS

Vol .II July - Sept 2009 Issue No.3







DEPARTMENT OF COMPUTER APPLICATIONS (MCA)

K.S.R. COLLEGE OF ENGINEERING

TIRUCHENGODE - 637 215

JOURNAL OF COMPUTER APPLICATIONS

JCA is a refereed research journal published quarterly by Department of Computer Applications, K.S.R College of Engineering, Tiruchengode. Responsibility for the contents rest upon the authors and not upon the **Journal of Computer Applications**.

JOURNAL OF COMPUTER APPLICATIONS

Vol II	July – Sept 2009	Issue No. 3

CONTENTS

NON-FUNCTIONAL REQUIREMENT PREFERENCES FOR 'ARDNAS' SYSTEM

P.Muthuchelvi
Sathyabama University,
Chennai, India
pranav4599@rediffmail.com

G.S.Anandha Mala,
Professor, Dept. of CSE,
St Joseph's College of Engineering
Chennai, India

V.Ramachandran
Professor, Dept. of CSE,
Anna University
Trichy, India

Abstract - As grid resources are geographically distributed, efficient resource discovery and management has become one of the important requirements. Besides, Grid users are independent identities and negotiation is necessary for reconciling their diverse characteristics. Therefore special mechanism is required to negotiate and discover the required resource or similar resource as an alternative when discovery fails. However, the quality of the service being provided in the grid environment depends on both functional as well as the non-functional requirements [NFR]. But conflicts between NFRs are not yet resolved effectively. Towards this end, a system of 'Non-Functional Requirement Preferences for 'ARDNAS' - (Agent Based Resource Discovery with Alternate Solution) 'NFR-ARDNAS 'System' is proposed to provide an expeditious and efficient resource and alternate resource when discovery fails with NFR preferences. In addition to the service provided to the grid user, the non-functional requirements preferences are also analyzed and the conflicts among them are resolved based on the trade-off analysis done with the help of fuzzy rule sets.

Keywords - Grid computing, Resource discovery, negotiation, alternate resource, agents, Non-Functional Requirements

INTRODUCTION

Grid computing is a hot research direction and drawing a lot of attentions from both academia and industry. A Grid is a set of resources distributed over a wide area networks that can support large scale distributed application. It will provide high-end computational and storage capabilities to a set of differentiated users. It has emerged to facilitate better utilization of under utilized heterogeneous and geographically distributed resources. The main motivating factor in grid computing is resource sharing. The resource management system, which is the central component of grid computing, provides highly available and adaptable computing capabilities to its user. This management provides efficient scheduling of applications and effective utilization of all resources available in the grid environment. Resource management acts not only as an interface between grid resource and grid application but also to provide reliable service to the user.

A grid service includes both functional and non-functional requirements. These properties can be obtained with the help of requirement Engineering. Functional requirements are associated with specific functions, tasks or behaviors the grid system must support, while non-functional requirements are determining the constraints on various attributes of these functions or tasks. Non-Functional requirements in requirement engineering presents a systematic and practical approach to "building quality into" grid services. Grid services must exhibit quality attributes, such as performance, fault tolerance, security and trust, platform independence, modifiability and etc.

There are many issues and challenges in the Grid environment. Amongst the challenges for Grid computing, to date, there is little work that addresses the issues of requirement engineering. To the best of the author's knowledge, at present, there are only a few (preliminary) efforts on considering NFRs to provide quality service. The quality of the service being provided depends on both functional as well as non-functional requirements (NFRs) like performance, fault-tolerance, security etc. These non-functional requirements are still not resolved effectively due to the conflicts among them. Hence, objective of the proposed system is to address this problem by getting the preferences from the grid user, analyze the conflicts and prioritize them. This approach makes use of the possible preferences of the grid users and non-functional requirement taxonomy to analyze the conflicts which are resolved based on trade-off analysis by prioritizing the preference. The prioritization depends on the dominating non-functional requirements from the inference engine.

In the current grid research, NFRP-GU[1] system was proposed to identify the non-functional requirements of the grid user request with their preferences and analyze the conflicts among them. This module is added into ARDNAS[2] system which was developed to provide an expeditious and efficient resource and alternate resource when discovery fails. In this paper, by integrating [1] and [2] 'NFR_ARDNAS - Non-Functional Requirement Preferences for ARDNAS system 'has been proposed in order to fulfill both functional and non-functional requirement of grid user. The contributions of this paper are listed as follows. Related works are discussed in section 2. The proposed system architecture and the functions of each component are explained in section 3. Section 4 discusses the implementation and results. Section 5 concludes this paper and discusses the future enhancements.

RELATED WORKS

A mobile agent toolkit designed for resource discovery [3]. Here mobile agents are used for resource discovery, load balancing and distribution. However this system does not have an economic model associated with it. An economic model is considered in [4]. Here mobile agent's traverse through the nodes in the network and when a suitable resource is found in any node, process is executed in that node, otherwise it moves to the next node. This is considered a time consuming process. Negotiation and re-negotiation between resource requester and resource provider agent with service level agreement between them is considered in [5].This paper however focused only on higher level functionalities of the system. A new approach based on agent teams to facilitate resource discovery with yellow page service was introduced in [6]. Here, the agent wants to select a team to execute its job. This also falls short of requirements in dynamic grid situation. Agent achieved higher success rate by slightly relaxing the bargaining terms, in intense pressure situation However, relaxation was decided on fuzzy decision controller [7].Elicitation of Non-Functional Requirement Preference for actors of use case from domain Model [8] was proposed to identify the non – functional requirements for a given use case description from the domain model such as unified modeling language class diagram and goal based questionnaires. However this system is not considering the grid user preferences. Considering the inadequacies of the efforts previously made and referred to above, the motivation behind the work is to sets out the system of NFR_ARDNAS for devising a speedy and very efficient method of resource discovery with the help of agents and analyzes the conflicts between NFR. The main objective of this work is to provide most relevant resource when necessary and to increase the success rate of agent with NFR user preferences. The quality of the service being provided in the grid environment depends on both functional as well as the non-functional requirements [NFR]. But conflicts between NFRs are not yet resolved effectively. This paper also presents an approach to identify the non-functional requirements of the grid user request with their preferences and analyze the conflicts among them.

PROPOSED SYSTEM ARCHITECTURE OF 'NFR-ARDNAS' SYSTEM

This system is proposed to provide an expeditious and efficient resource, alternate resource when discovery fails with NFR user preferences. An overview of the NFR_ARDNAS architecture is illustrated in Figure 1. The main components of this system include (i) Grid users (ii) ARDNAS system (iii) NFR extractor (iii) NFR Prioritizer and taxonomy (iv) Goal-based questionnaires. In this system, resource request is submitted by grid user in order to execute the application. After getting the request from the user, it searches for the match and provides the resource if discovery success otherwise provide alternate resource when discovery fails. However the service provided to the user is purely based on both functional and non-functional requirements. This paper presents an approach to identify the non-functional requirements of the grid user request with their preferences and analyze the conflicts among them and provide the service based on functional as well as non- functional requirements.

A. Grid Users

There are two categories of user in the grid computing environment namely 1.Resource requester and 2.Resource provider. Resource requester is a type of user who desiring to utilize the services from the Grid. They approach the grid for resources to execute the process. Conditions governing resource request generally are, name of the resource, speed, cost, time etc. The resource requester may seek the resource for application and trying to obtain the exact match of the resource with inherent constraints. Such users always strive to get the exact match of the resource at minimal cost. While they are getting service from the system, the quality of service is also depends on non-functional requirements. Grid user should specify their preferences while they are making a request for the resources. Second type of user is the resource providers, who can register their resources with specification and constraints to the agents. Resource owners strive to maximize their return-on investment. However resource providers and requester must be an authorized user for accessing the Grid.

B.'ARDNAS' system

ARDNAS system is designed specifically for intelligent and expeditious method of resource discovery. In such a system, the agents fall into several categories with different sets of behaviors to perform various operations. The various kinds of agents and their respective functions are

- Resource Requester Agents (RRA) – agents desirous of utilizing the services available in the grid on behalf of requester users.
- Resource Provider Agents (RPA) – agents able to contribute resources to the grid community on behalf of provider user..
- Negotiation and Alternate Solution Provider (NASP) agent – It maintains specialization wise classification list of agents and routes the request to the corresponding agents. It also negotiates and provides an alternate solution confirming to the constraints specified by the user when discovery fails.
- Cognitive agent – It is a customized agent who is collecting information on all the processes taking place in the system and replying to the queries by applying its intelligence.

C. Resource Requester Agents (RRA)

RRA reads the resource request with specification from user who desiring to utilize the services from the Grid. It approaches NASP for

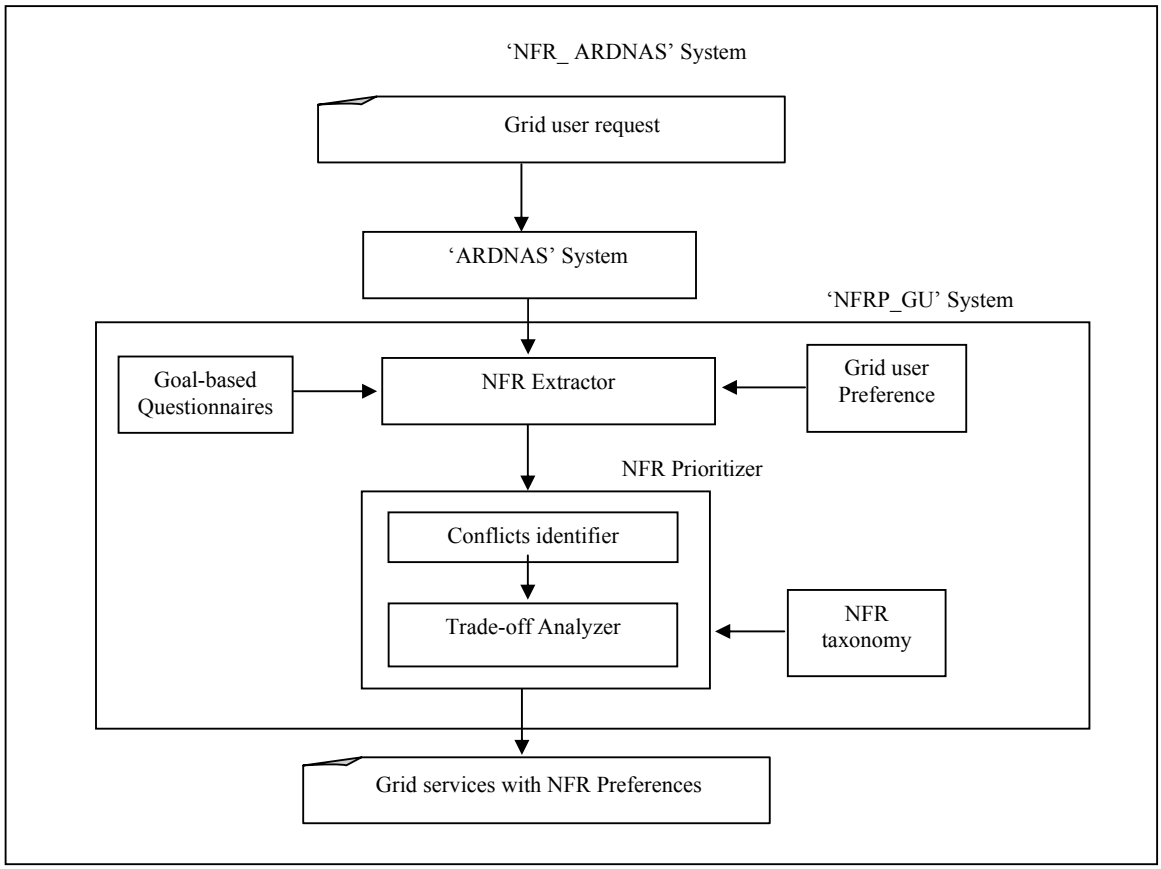


Fig.1 Architecture of 'NFR_ARDNAS System

resource with specification, using 'Call For Proposal (CFP) – Agent Communication Language (ACL)'. Given its pivotal position, NASP will route the request to the corresponding group of agents Thus the risk of referring the request to irrelevant agent can be avoided.

D. Resource Provider Agent (RPA)

This agent works on behalf of provider user who offers the resource to the grid community. When RPA approaches NASP for registering the resource with specifications, NASP includes it in the relevant group after due verification of service that can be provided. If the RPA dose not fit into any of the existing group, NASP creates a group for the new service and this RPA will become the first agent in that group. Thus the number of groups can be increased according to the service provided.

E. Cognitive Agent

The developments of the grid in recent times have significantly speeded-up its performance and yet the position has remained inadequate. One possible reason for this could be that RRA and RPA have not had the benefit of assistance of Cognitive agent before the process was set in motion. In the proposed system, Cognitive agent helps in this regard. This agent is equipped to play the crucial role. Acting as a back-end assistant for providing information when required, it acquires its knowledge from the processes happening in the system and updates itself while fulfilling its role. In the proposed system, Cognitive agent is thus customized agent and is responsible for the following services

- High/low demand resource
- Instantly available resource
- Transaction History
- Suitable resource identification
- Performance Evaluation

F. NASP Agent

This agent plays a critical role in the system and handles a large number of requests. It acts as a link between resource requester and resource provider agent. It performs many important functions like collecting the resource details from providers, match making, negotiating and providing alternate resources besides monitoring the quality of service in the system. RRA approaches NASP for specification governed resource for executing its task. When exact match is found, resource discovery succeeds.

Table - I
Goal Based Questionnaires

Event	Preference	NFR
Type grid user-id	Invalid user –id	Security
Type grid user-id	Provide Wizard or portal /Guide to enter the user-id without problem	Usability
Negotiation	Terminate the negotiation when two parties not come to an agreement	performance
Press login/logout button	Provide Wizard/Guide to press logout button	Usability
Press login/logout button	Enable user to use button effectively at reliable places	Reliability
Choose resource for application	Information for choosing resource is accurate and available	Correctness
Triggers alarm	Alarm resource repaired, changed and maintained	Maintainability
Triggers alarm	Constraints for trigger only when authorized users access the resource.	Security
Acquire Knowledge	Provide knowledge for the new user	performance

In case discovery fails, it tries to find alternate resource based on the type of processes (time /cost bound). To provide alternate resource, NASP relaxes attributes other than cost for cost bound process and time for time bound process. However for both the processes, relaxation is allowed only within the level of relaxation factor.

NASP agent performs the following functions

1. Receive resource specification from RPA.
2. Classify the RPA according to service and add the agents to that group and provide resource-id.
3. Receive the requests from RRA with resource specification.
4. Evaluate the proposals and short list the RPA based on the RRA specification.
5. Negotiate and decide the best resource for assigning to RRA.
6. If no such RPA exists, select some other RPA by relaxing some of the criteria based on the type of process for the suitable alternate resource.
7. Get Cognitive agents help whenever necessary.
8. Receive the utilization report from RRA after use of the allotted resource.

G. NFR Extractor

NFR preferences for the grid user requirements are extracted with the help of goal based questionnaires and grid user preferences. The goal based questionnaire includes all possible questions for the activities of both resource requester user and resource provider user. Users have to give their preferences by appropriately answering for the questions provided by the user friendly portal designed for this purpose. From this portal information 'NFR extractor' extracts the non-functional requirements preferences for the user and redirects them to the 'NFR prioritizer'. Sample goal based questionnaires are shown in the Table - I.

H. NFR Prioritizer

'NFR Prioritizer' consists of two components namely 1.Conflicts identifier 2.Trade-off analyzer. 'Conflicts identifier' analyzes the conflicts among the extracted NFRs with the help of NFR taxonomy. In 'NFR Taxonomy' all the NFRs are associated with other conflicting and dependable NFRs.

Table - II
NFR Taxonomy

Correctness#Reliability+#Efficiency+#Accuracy+#Conciseness+#Tolerance+#Precision+
Performance#Response+#Throughput+#Timeliness+#Availability-#Reliability-
Reliability#Efficiency+#Accuracy+#Latency-#Throughput-#Availability-
Security#Identification+#Authorization+#Immunity+#Nonrepudiation+#Privacy+#Performance-
Usability#Simplicity+#Accessibility+#Installability+#Operability+#Maintainability-
Maintainability#Flexibility+#Simplicity-#Operability-#Usability-#Portability+
Availability#Reliability-#Integrity-#Precision-#Throughput+#Tolerance-
Authorization#Security+#Performance-#Authentication+#Reliability+#Privacy+
Efficiency#Simplicity+#Maintainability+#Latency+#Performance-#Maintainability-
Identification# Security+# Performance-
Authentication# Security+# Performance-

Table - III
Sample Fuzzy Rules

1.If (Efficiency is low) and (Accuracy is low) then (Reliability is low) (1)
2. If (Efficiency is high) and (Accuracy is high) then (Reliability is high) (1)
3. If (Latency is low) and (Throughput is high) and (Availability is high) then (Reliability is high) (1)
4. If (Latency is high) and (Throughput is low) and (Availability is low) then (Reliability is low (1)

The entries in NFR taxonomy looks like,
Performance#Response+#Throughput+#Timeliness+
#Availability-#Reliability-

It states that 'Performance' is directly proportional to 'Response', 'Throughput', 'Timeliness', and 'Availability' but indirectly proportional with 'Reliability'. The sample NFR taxonomy is shown in the Table II. After identifying the conflicting NFRs, the NFRs are prioritized based on the trade-off analysis. Trade-off analysis explores the cost of relaxing one NFR in order to achieve an increase in another NFR. This is implemented using fuzzy rule sets. These rules are formulated for each NFR according to the conflicting and dependable NFR. Sample rules for the reliability are given in Table III. After the process of fuzzyfication and de-fuzzification, the NFRs are prioritized and the results are produced by the system.

IMPLEMENTATION AND RESULTS

The system has been implemented using JAVA and Java Agent Development framework-JADE. User friendly portals are created in ASP. The goal based questionnaires are stored in Ms-Access database. Trade-off analysis has been done in Mat lab with the help of fuzzy rule sets. The agent platform has been split on several hosts provided there is no firewall among them. Agents are created in distributed environments among five system. Agents are implemented as a java thread and ACL(Agent Communication Languages) messages are used for effective and lightweight communication between agents. The results produced by the ARDNAS system for 'with and without alternate solution' are compared. For the purpose of comparison, a set of hundred data has been analyzed. The success rate has been calculated as

Success rate (SR) = N success / N total

Where N success is the number of processes completed successfully and N total is the total number of processes submitted. It is observed that number of processes completed with alternate solution is consistently higher than number of processes completed without alternate solution in ARDNAS[2] system. In addition to the previous work, in this paper, conflicts between NFR are identified and prioritized using trade off analysis with the help of fuzzy rule sets. It was implemented by adjusting the weight values associated with each NFR. These weights assign the priority that each NFR has relative to the others. The system was executed several times with varying weight values to prioritize the NFR. The results produced by the system are shown in fig 2.

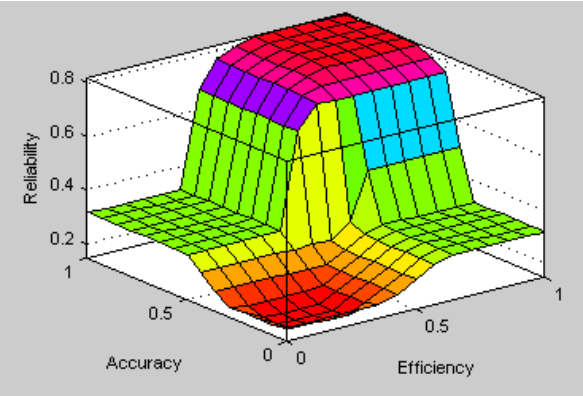


Fig.2 Trade - off analysis for reliability

CONCLUSIONS

The quality of the service being provided by the 'ARDNAS' system also depends on the non-functional requirements such as feasibility, reliability etc., But NFR' s are still not derived effectively due to the conflicts between them. The NFR_ARDNAS system is proposed to enhance the known methods of grid resource discovery. It plays a vital role in bridging the seemingly wide gap between requirement engineering and grid environment. A novel approach of deploying NASP, Cognitive agent and NFR Prioritizer used is suggested for gratifying the critical functions of linking two different domain.. This twin task greatly promotes the overall efficiency of the grid service. This system can be included as one of the services in the real Grid environment created with the help of GLOBUS tool kit.. Further the trade-off analysis can be automated with the help of knowledge - base.

REFERENCES

[1] Muthuchelvi P, Anandha Mala G.S. and Ramachandran V , "Non-Functional Requirement Preferences for Users of Grid Environment" , National Conference on Information and Software Engineering Organized by IEEE CS Student Branch Chapter and Department of Information Technology Vinayaka Missions University. 13th&14th February 09. Volume 1, pp02.

[2] Muthuchelvi P, Anandha Mala G.S. and Ramachandran V , " ARDNAS : Agent based Resource Discovery with Negotiated Alternate Solution", International conference on advances in Computing ,Communication and Control.- ICAC3'09. Organized with ACM - SIGART. January 23 & 24. 2009. PP.239-245.
[3] Rana.O.F, Di Martino.B , Grid performance and resource management using mobile agents, In: Performance analysis and Grid computing, 2004, 251-263.
[4] Manvi.S.S, M.N. Birje, B. Prasad, An Agent based Resource Allocation Model for computational Grids, Multi Agent and Grid Systems, 1(1), 2005, 17–27.
[5] Ouelhadj.D, Garibaldi.J, MacLaren..J, R. Sakellariou, K. Krishnakumar, A Multi-agent infrastructure and a service level agreement negotiation protocol for robust scheduling in Grid Computing' In: Peter M. A. Sloot et. al. (eds.), Advances in Grid Computing—EGC, Lecture Notes in Computer Science, 3470, Springer-Verlag, 2005, 651–660.
[6] Dominiak.M, Kuranowski.W, Gawinecki.M, M. Ganzha, M. Paprzycki , Utilizing agent teams in Grid resource management-preliminary considerations, In: Proceedings of the IEEE J. V. Atanasoff Conference, IEEE CS Press, Los Alamitos, CA, 2006, 46–51.
[7] Sim.K.M, Negotiation agents that make prudent compromises and are slightly flexible in reaching consensus, Computational Intelligence , Vol 20, No.4, 2004. pp- 643-662.
[8] .G.S.Anandha Mala, G.V.Uma, "Elicitation of Non-Functional Requirement Preference for Actors of Use case from Domain Model", Lecture Notes in Artificial Intelligence - 4303, Springer-verlag, 2006, pages 238 – 243.
[9] Muthuchelvi.P and Ramachandran.V, "ABRMAS: Agent Based Resource Management with Alternate Solution", Sixth International Conference on Grid and Cooperative Computing - GCC 2007, IEEE Computer Society Press, USA, pp. August 2007, 147 – 153
[10] Muthuchelvi P, Anandha Mala G.S. and Ramachandran V), "KBRMAS - Knowledge Based Grid Resource Management for Compromised Alternate Solution", " sai: International Transactions on System Science and Application ITSSA .2008.01.022, Vol 3, No 4, Jan 2008, pp. 338-345.
[11] .Rosa M. Badia, Jes´us Labarta, Ra¨ul Sirvent, Josep M. P´erez, Jos´e M. Cela, and Rogeli Grima. Programming Grid Applications with GRID Superscalar. Journal of Grid Computing, 2003, 1(2):151- 170.
[12] Arenas, A. Bilas, J. Luna, M. Marazakis and etl " Knowledge and Data Management in Grids: Notes on the State of the Art" CoreGRID White Paper Number WHP-0002, May 2, 2008, Institute on Knowledge and Data Management.
[13] Muthuchelvi P, Anandha Mala G.S. and Ramachandran V, "Agent Based Grid Resource Discovery with Negotiated Alternate Solution and Non-Functional Requirement Preferences" Journal of Computer Science, " 5 (3): pp: 191-198, 2009 ISSN 1549-3636, © 2009 Science Publications, USA.

BIOGRAPY

Prof. P.Muthuchelvi received B.E degree from University of Madras in Computer Science & Engineering in 1991, M.E degree from University of Madras in 2001. Currently she is working as Assistant Professor in St.Joseph's college of Engineering, Chennai, India.. She has 18 years of teaching experience on graduate level. Her area of interest includes Agent based Grid Computing. Currently she is pursuing Ph.D degree from Sathyabama University, India.

Dr.G.S.Anandha Mala received B.E degree from Bharathidhasan University in Computer Science & Engineering in 1992, M.E degree in University of Madras in 2001 and Ph.D degree from Anna University in 2007. Currently she is working as Professor in St.Joseph's college of Engineering, Chennai, India, and heading the department of Computer Science and Engineering. She has published 10 technical papers in various international journal / conferences. She has 15 years of teaching experience on graduate level. Her area of interest includes Software Engineering and Grid Computing.

Dr.V.Ramachandran served as a Professor in the Department of Computer Science and Engineering and Associate Professor in the Department of Information Technology in Anna University, Chennai, India. He obtained his doctorate degree from Anna University in 1991 and served as Visiting Professor in several national and international institutes. He has authored more than 70 research publications. He has more than 25 years of teaching experience. Currently he is occupying the coveted position as Vice Chancellor of Anna University, Trichy, India. His area of interest includes Grid computing, Distributed Computing, cloud computing and Webservices.

EXTENDED ENTERPRISE APPLICATION SOFTWARE - AN INDIAN PERSPECTIVE - "ZEAL TO ZENITH"

R.Seranmadevi
Lecturer, Dept. of MBA,
CMS College of Engineering, Namakkal,
seranmadevi@yahoo.co.in

Dr.M.LathaNatarajan
Professor / Head, Dept. of MBA,
Vivekanandha Engineering College for Women,
Tiruchengode. rosesbyangel@gmail.com

Abstract - ERP is essentially an integrated software system consisting of multi-module applications and a common database. ERP systems help an organization to manage crucial business processes such as planning, production, inventory, purchasing, sales, customer support, etc. The ultimate aim of ERP is to integrate the various departments within an organization based on the enterprise – wide data model. cERP is a cohesive framework to facilitate information exchange for doing business. It is applied across the extended enterprise using e-business solutions. ERP II enables businesses to compete by providing information online and by adding real value to businesses of all types and sizes. In India various enterprise are implementing this cross-functional system.

EAI is the set of technologies that allow the movement and exchange of information between various applications and business processes within and between the organizations. EAI products can be applied for integrating ERP with SCM and CRM.

All ERP implementations are not successful. Implementations succeed or fail due to a number of reasons. Depends upon the industry and company the success and failure will differ. The successive stories of many Indian Companies and the viewpoints of experts clearly codified that ERP implementation in India is already initiated, but the problem in it is sustainable development and uninterrupted involvement till to reach the peak or attain the zenith.

INTRODUCTION

ERP is essentially an integrated software system consisting of multi-module applications and a common database. Such software systems help coordinate business activities and facilitate the flow of information across an enterprise. ERP systems help an organization to manage crucial business processes such as planning, production, inventory, purchasing, sales, customer support, etc. ERP systems provide not only the core functionality that most large corporations depend on, but also a number of financial applications. The ultimate aim of ERP is to integrate the various departments within an organization based on the enterprise wide data model.[1]

BUSINESS INTEGRATION THROUGH ERP SYSTEM

ERP II is an application and development strategy that expands out of ERP functions to achieve the integration of an enterprise's key, domain-specific internal and external collaborative, operational, and financial processes. The extended components of ERP systems are composed of CRM, SCM, BI and e-Commerce applications.[14]

ERP II has led to the advent of Collaborative ERP (cERP). cERP is a cohesive framework to facilitate information exchange for doing business. It is applied across the extended enterprise using e-business solutions. The mission of cERP is to increase partnering between retailers and manufacturers, and manufacturers and their suppliers through co-managed processes and shared information.

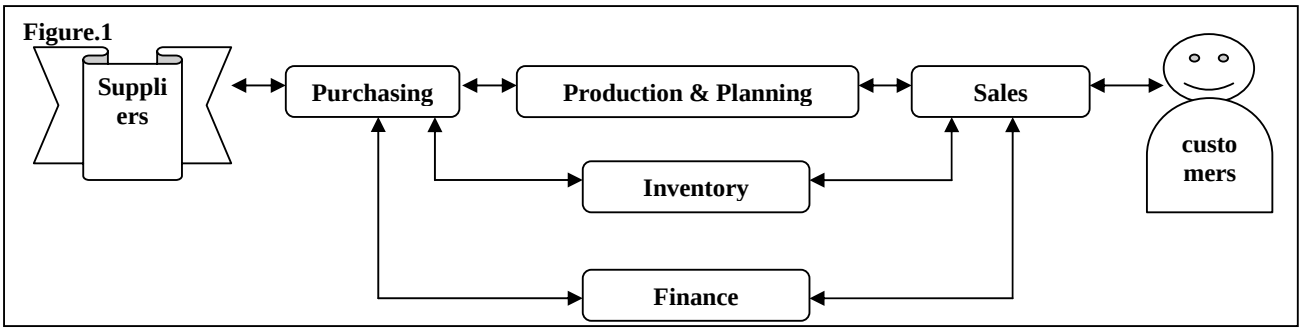
EXTENDED ERP

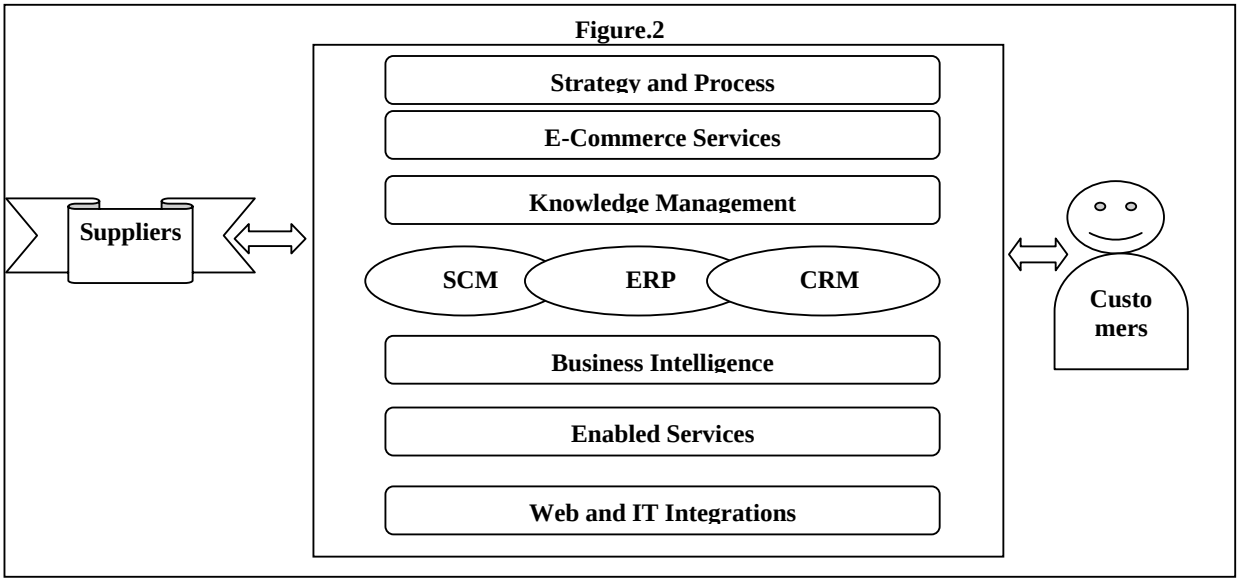
Lately ERP has evolved into Extended ERP or ERPII. It uses the internet to reach out to suppliers, customers and a wider range of employees. Applications such as Customer Relationships Management (CRM), Supply Chain Management (SCM), Business Intelligence (BI), and sell-side/buy-side e-Business provide the much needed functionality for ERP-II. Extended ERP applications have led to the advent of Collaborative Commerce (C-Commerce), C-Commerce is the electronic interaction of different businesses within the supply chain or other an industry.[15]

Over the past few years solutions such as CRM and SCM have leveraged the Internet to support these processes. ERP II incorporates all these processes (CRM, SCM, BI, e-Commerce and APS) in a single package. To be globally competent, an organization needs to open up and reach out to its collaborative partners.

ERP II enables businesses to compete by providing information online and by adding real value to businesses of all types and sizes. Various enterprise are implementing this cross - functional system. According to Gartner (2000), ERP II can be differentiated from ERP on the basis of six elements that touch business, application and technology strategy. These elements are: the role of ERP II, its business domain, the functions addressed within that domain, the processes required by those functions, the system architectures that can support those processes, and the way in which data is handled within those architectures.[2]

As the role of ERP is enterprise automation, ERP II automates and optimizes its entire value chain, comprising customers, suppliers and third party manufacturers.



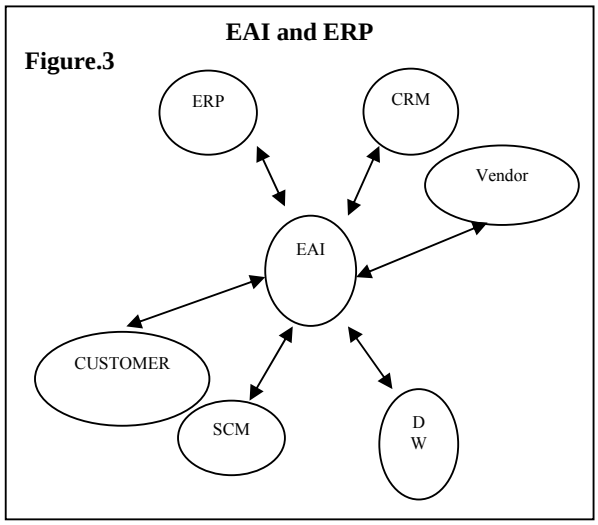


EAI AND ERP

EAI can be defined as application independent, business-oriented software that integrates the applications of an enterprise, or several enterprises. These are used to automate multi-enterprise business processes wherein the internet is the backbone of communication devices and departments. Thus EAI is the set of technologies that allow the movement and exchange of information between various applications and business processes within and between the organizations.

Messaging is the foundation of an EAI framework. This backbone transports messages between resources, reconciling network and protocol differences. It allows applications to share information with the outside world by sending and receiving messages. It adds quality –of-service options to message delivery, such as security and queuing.

EAI products can be applied for integrating ERP with SCM and CRM. EAI acts as glue between ERP, SCM, CRM and Data Warehouse and establishes a rapid integration framework both internally and externally. EAI sits between a wide range of applications in the enterprise brokers messages between them, and adds processing wherever required. [10]



An EAI solution is composed with services like Business Process Management, Application Connectivity, Translation and Formatting and Communication Middleware.

Generally, there are four types of EAI; data-level, application program interface-level, method-level, and user interface-level.

ERP AND BI

The rapid progress of BI applications, such as Data Warehouse and Data Mining, and IT in the last ten years has helped ERP systems to improve vastly and meet customer demands. Data Warehouse is a single, complete, and consistent store of data, obtained from a variety of different sources and made available to end users in a way that they can understand and use in a business context.

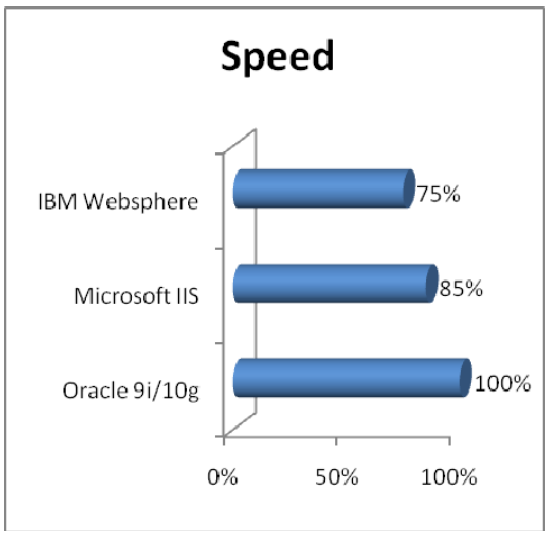
By integrating ERP with data warehousing, the company obtained the business benefits like ability to rapidly change and adopt new business structures, Increased profitability due to improved customer profiling and reporting, and Time and cost saving due to minimal training. So in this case was realized by updating ERP functions and integrating them into one of the big data warehousing vendors namely, Cognos.[11]

EAI VENDORS[12]

Some key EAI vendors, published in the EAI vendor guide with their respective market share, are; Table.1

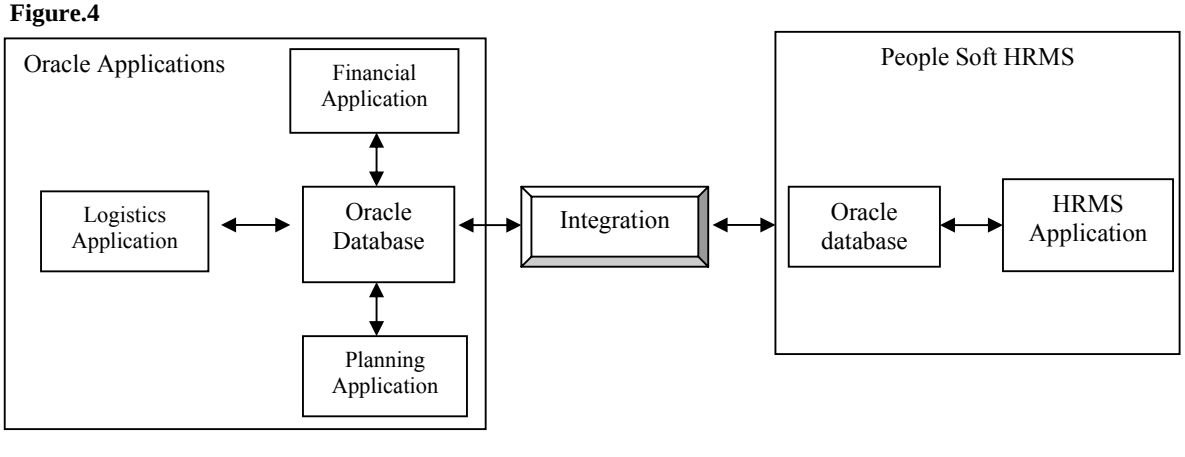
EAI Vendors	Market Share (%)
IBM	12.0
Neon	10.8
Mercater	10.0
TIBCO	9.1
Sun	6.4
BEA Systems	5.1
STC	4.4
Vitria	3.9
Active	3.7
Extricity	3.0

support, Compatibility with other systems, Ease of customization, Market position of the vendor and domain knowledge, Cross module integration and fit with organization structure, Reference of the vendor, Implementation time, Consultancy, Methodology of the software and System reliability. But these factors are depends upon the industry and company will differ, so keen interest and careful selection must be made.



INTEGRATED ERP APPLICATIONS

Figure.4



ERP SELECTION CRITERIA:[8]

All ERP implementations are not successful. Implementations succeed or fail due to a number of reasons. The factors should consider before selecting the appropriate ERP software for our organization are Functionality, Technical criteria, Cost, Service and

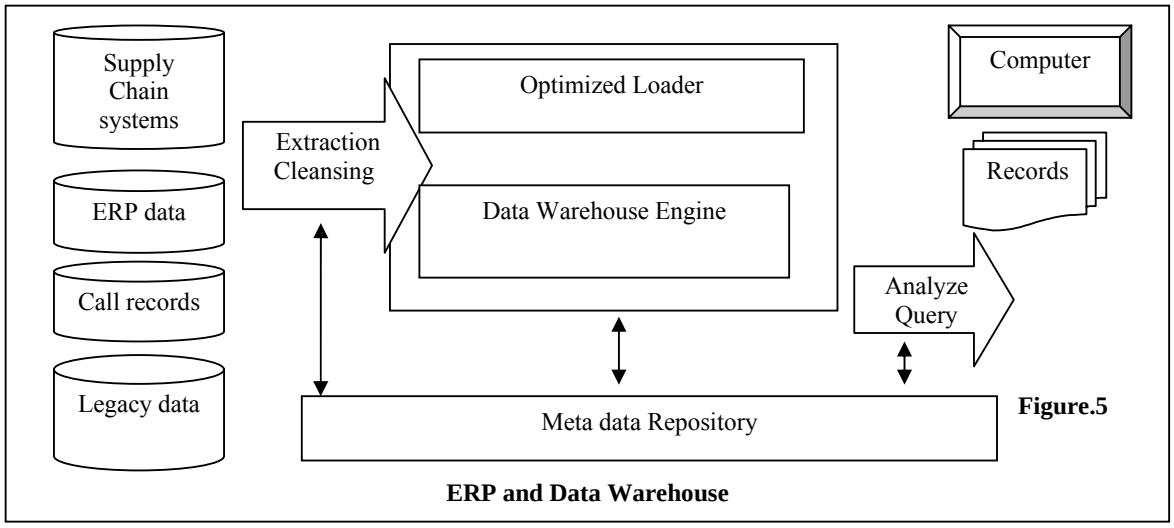


Figure.5

ERP and Data Warehouse

Figure.7[18]

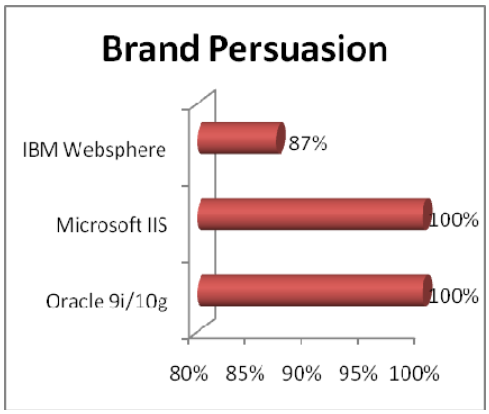


Figure.8 [18]

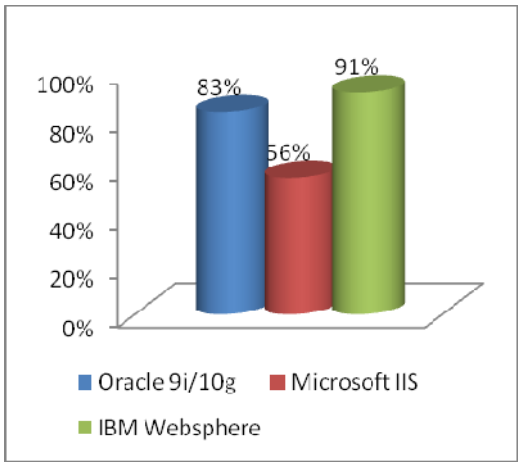
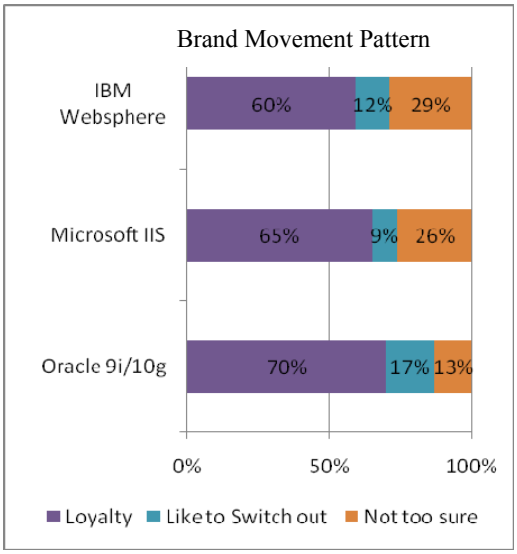


Figure .9 [18]



MOVEMENT OF APPLICATION SERVER VENDORS OF ERP SOFTWARE

The following are the notorious vendors are supplying the ERP Software viz., SAP AG, Oracle, People Soft, SSA Global, Sage Group Baan, Microsoft Dynamics. The later had released several versions like MS Dynamics CRM, MS Dynamics AX (Formerly, Axapta), MS Dynamics GP (Formerly, Great Plains), MS Dynamics NAV (Formerly, Navision), MS Dynamics SL (Formerly, Soloman), Microsoft Retail Management System (Formerly, Quicksell). Out of all the application servers out there, only three managed to get into the User's choice. Out of these, Oracle 9i/10g emerged as the most future ready application server, followed by Microsoft's IIS and IBM Websphere. [17]

Figure. 10 [17]

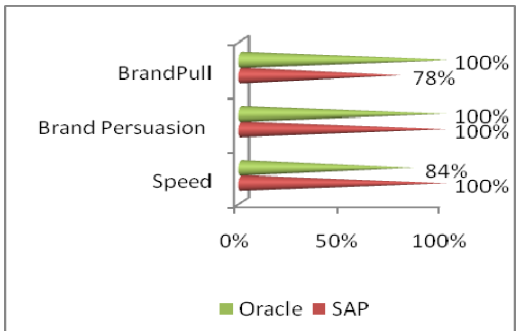
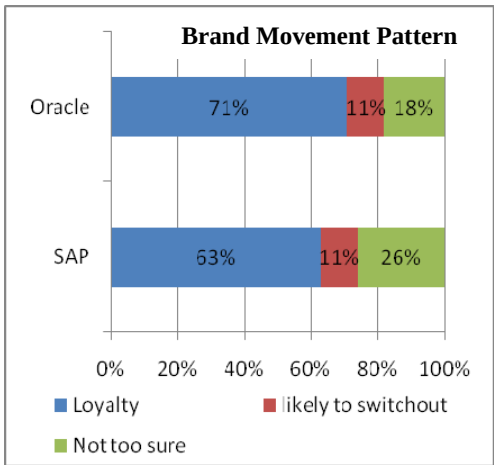


Figure. 11



ERP SOFTWARE VENDORS

Apart from application software ERP Software Sap leads a good career in Business Analytics, CRM, and Enterprise Messaging Solution. ERP software is yet another segment without any surprises, as it's largely dominated by two players – Sal and Oracle. Out of these, SAO continues its winning streak and remains the most future ready brand. Close on its heels is Oracle, with a relative speed of 84%. SAP's MySAP ERP continues to remain as the most future ready brand, with Oracle close on its heels with its multiple HRMA offerings from PeopleSoft, J.D.Edwards, and Oracle's own E-Business Suite. The situation hasn't changed much over last year, except for the entry of one more brand into the club-Automatic Data Processing Enterprise HRMS. MySAP enjoys the highest brand loyalty at 74%, with zero likely to switch outs. [16]

ERP IMPLEMENTATION IN INDIA

ERP systems are in place in many Indian Business houses. In fact, these systems are evolving day-by-day. There are several Indian enterprises that are running successful ERP systems.[10]

SAP IMPLEMENTATION AT BHEL[11]

BHEL (Bharat Heavy Electricals Limited) is the largest public-sector engineering and manufacturing enterprise in India in the energy related/infrastructure sector. BHEL manufactures over 180 products under 30 major product groups, and caters to core sectors of the Indian economy. It has so many core business which is located at various parts of India. In order to streamline information availability, BHEL decided on end-to-end implementation of the SAP solution. This meant that multiple legacy systems of different functions were to be integrated into a single system in SAP.

MAHINDRA & MAHINDRA MYSAP IMPLEMENTATION: [11]

Mahindra & Mahindra Limited (M&M) is a leading $800 Million automobile manufacturer, employing an estimated 12,000 people. It is the flagship organization of the Mahindra group. The company has been the market leader in farm equipment and machinery in the highly competitive e Indian Market since 1986. On one hand, M7M wants to compete for more global business, and on the other hand, international competitors want a piece of the Indian business as well. So M7M implemented mySAP Supply Chain Management to link all its plants and to streamline the manufacturing process. The solution enables a pull-based replenishment system that optimized logistics and manufacturing operations, which helped M&M improve margins and reduce costs, while enabling it to respond quickly to consumer needs.

DABUR PHARMACEUTICALS LIMITED: [12]

Dabur Pharma Limited is a leading Indian Pharma company engaged in cancer research and manufacturing related products. It is an associate company of Dabur India Limited and was incorporated in March 2003.it operates in Europe and in some other markets of the world through its fully-owned subsidiary-Dabur Oncology Plc. The Dabur group of companies has a legacy of trust and excellence in the healthcare business. The organization implemented SAP ERP to help create, consolidate and maintain a database in real time to help the organization to make faster decisions, achieve cost cutting and maintain regulatory compliance. The implementation was take place phase by phase. Then, according to the business blueprint requirements, the actual SAP system was configured. This was followed by baseline demonstrations of each module to SAP Users.

HINDUSTAN PRODUCTS LIMITED (HPL) [3]

HPL is an FMCG based leading organization, which has a wide range of products from personal and household-care products to foods, beverages and many more. The organization owns over 110 manufacturing units and third party manufacturers. It

also has a wide distribution network consisting of over 110 ware houses. The organization operates through its own supply chain network, which had serious inventory problem. The organization, in 1977, implemented the Baan ERP package at over 250 sites, including manufacturing, sales and warehouse departments to integrate its manufacturing, financial and distribution processes. HPL's e-commerce initiatives used the ERP infrastructure to extend the organization through the Web, to its partners, suppliers, vendors, and customers. This integration of ERP and related SCM tools resulted in significant reduction in inventory levels, reduced stock levels, and lower working capital requirements, while improving the response time and customer-service levels.

ADITYA BIRLA CARBON BLACK BUSINESS MANUFACTURING LIMITED [3]

Aditya Birla Group's Carbon Black business spans four companies in four countries namely Egypt, India, Thailand, and in China. With a total annual capacity of 7,90,000 mt, the group is the fourth largest producer of Carbon Black in the World. The group company was using different custom-developed applications at all its units. This meant lack of uniformity in business processes functions. This made consolidation of business level data difficult. Yet another problem was the high cost of maintaining and upgrading disparate systems. A single enterprise application across business units was required. Since the group company already using SAP across all its business, SAP R/3 was the preferred choice for its Carbon Black business.

DIGITAL SHOPPY-DIGITAL INTEGRATION [16]

One of India's fast growing consumer electronics, home appliances, computers and communication devices retailer, Digital Shoppy a part of the Pratyankara Electronics and has stores spread across 33 locations. Like any other retail chain, Digital Shoppy was also in an expansion ode but with the growth in the number of stores came the added pressure of managing store dependently without an integrated ERP.

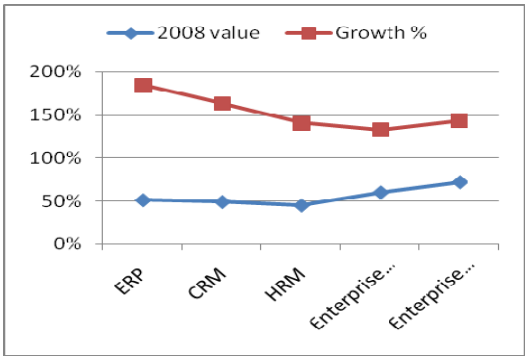
SURVEY RESULT OF PC QUEST ABOUT INDIA'S FAVORITE ENTERPRISE SOFTWARE BRANDS: [17]

The number of CIOs who intend to purchase ERP software has increased by 134% as compared to last year, indicating that ERP will be a high growth area.

The integration of SAP with their own methods will brings the company "The Best IT Implementation Awards 2009" under various categories. Table.2

Brand Category	2007 Value	2008 Value	Growth %
ERP	22%	51%	134%
CRM	23%	49%	115%
HRM	23%	45%	96%
Enterprise Messaging	34%	60%	73%
Enterprise Connectivity	42%	72%	72%

Figure.12



BSES POWER [6]

Best IT Implementation of the year 2009 Awards won by the company for largest scale in core application. For an electricity distribution organization, the operations & Maintenance department requires data exchange and communication with related support functions like stores & Meter management. OMS (Outage Management System) is the main O& M Module (Operations and Maintenance) used for monitoring the electricity supply. Further it has automated SMS escalations of faults and breakdowns. The OMS has been seamlessly integrated with SAP R/3 to make the operations efficient. It brings the benefits like 304 offices connected across 950 sq km, 1500 stakeholders have access to unified data at any moment and 90% customer complaints attended less than 2 hours, finally power outages decreased by 98%.

ARTEMIS HEALTH INSTITUTE [7]

Atremis is a new hospital in Gurgaon which became operational in 2007 won the Best Automation Award in 2009. The hospital went for the latest Hospital Information System (HIS) package. The HIS caters to the front office operations however back office operations could not be carried out through the HIS and documents related to these processes existed in independent silos. To make these operations system enables and to ensure seamless integration with HIS, there was a need to implement an ERP package that catered to these functions. Since the group company Apollo Tyres was already using SAP for all its functions, they decided to implement the ERP for back-office functions of the hospital as well. Implementing SAP is a huge task keeping in mind the adoption of best practices available in SAP.

INDIAN OIL CORPORATION - AUTOMATED B2B PROCESS INTEGRATION [7]

The project was implemented by IOC, which is India's largest Pubic sector petroleum company and is 116th on Fortune Global 500 listing 2008. Keeping track of every transaction has always been difficult for them as it is spread across the country. To solve this issue, they have integrated their SAP implementations which are considered among the largest in the country.

ERP IMPLEMENTATION AT THE VIEWPOINT OF INDIAN EXPERTS [6]

- Rajendra B.Vattikuti, Group President, CSC India stated that "CSC's global practice dedicated to supporting SAP systems combines a wealth of skills and physical resources including world-class, full service outsourcing, consulting and systems integration teams".
- Swapan Bardhan, Assistant Manager, IT, Calcutta Medical Research Institute sharing the thoughts regarding ERP implementation that a few years back CMRI had implemented ERP based on the Microsoft platform. Previously, entire processes to run on ne server.
- Navtej Matharu, CTO, InfoVision Group stresses the importance of innovation in the company's line of work specialized in CRM services. He insisted that "… changes and innovation are thought through to ensure successful implementation and a regression path, should the need arise".
- Dinesh Kumar, Executive director, IT,NTPC is a man who thrives on challenges and he has never faced the difficulty of "running out of things to do". He has been instrumental in keeping the company at the leading edge of technology adoption, whether it be usage of satellite based communication in 1982 or ERP in 2003.
- Ray D'Souza, Director of Systems & Technology, Lowe Lintas is an ex-navy personnel heads the IT at advertising agency. He also has the distinction of making the company the first advertising agency in India which implemented an ERP System.
- P.Shyam Sunder, Vice President, Quality and Head of IT, Britannia Industries took over this role after handling several other responsibilities. As a matter of fact, Britannia is the first FMCG company in India to comprehensively outsource infrastructure solutions, SAP application services, and consulting, among others. He created an IT strategy at Britannia, and also oversaw the implementation of all modules of SAP- big bang, multi location networking in a record time of under a year, with no cost and time overrun.
- GS Ravikumar, CIO, GATI recalled by Gati 2000 for making a failed project successful, Ravikumar had a lot of expectations riding on his shoulders. Responsible for designing, developing and implementing customized ERP across 300 locations including some remote places, was a challenge and he succeeded.
- Sanjay Rawal, CIO, Coca Cola, India is beginning his career as a Programmer analyst and moved up with other directions. As bottling units are a critical business segment, Rawal is now involved with standardizing the bottling franchisees for bringing them on an ERP platform across locations and putting a common set of processes in place.
- Amit Mukherjee, group CIO, RPG Enterprise he is the technology face of his group, he ensured that all retail formats of RPG's business, runs smoothly providing a high degree of customer satisfaction. Apart from retail, he is also the one who spearheaded Saregama, becoming India's first media-entertainment company to implement SAP solutions for its intellectual Property Management.
- TG Dhandapani, Corporate CIO, TVS Group has helped the way IT is managed at TVS group. He was closely involved in eight successful SAP implementations in his organization with the in house team. He facilitated the in-house developed Dealer Management System, an ERP for TVS Motor dealers, and so far 325 dealers have adopted this system that enables seamless interaction with ERP and BI.
- Bihag Lalaji, VP (IT), Ambuja Cement working for the last 12 years and implemented SAP in a record time of 14 months for 2,500 users. "The challenge was to bring on eight different platforms on a single system; user training; and to switch from old to new system", Lalaji says.
- Shobhana Ravi, CIO, TAFE. Under her leadership, TAFE has developed its own data warehousing applications using SAP and Oracle databases. She not only redefined the IT set up in the company but has infused modern technologies that have enables the company to accrue more profitability in the competitive automotive space.

WHY DOES ERP FAILURES IN INDIA? [14]

Implementation risks occur throughout the ERP system life cycle, which ranges from go-no-go decision on ERP to the time the system has gone live, including training issues.

- Poor leadership and lack of top management's commitment
- Unrealistic user expectation
- Inadequate control of system
- Lack of project Management structure and methodologies
- Insufficient user training
- Ineffective communication of project objectives
- Lack of user participation

- Lack of required skill-set and skill-mix
- Lack of software design standards
- Business process re-engineering
- System migration and control risks
- Module specific risks

HOW TO MAKE AN ERP IMPLEMENTATION A SUCCESSFUL ONE? [16]

- Requiring appropriate business and it legacy systems
- Creating a business plan and vision
- Preparing for business process re-engineering
- Developing change management culture and programmes
- Designing proper communication
- Fixing ERP teamwork and composition
- Monitoring and evaluating the performance
- Appointing a project champion
- Instituting project management
- Developing software, testing and trouble-shooting
- Expecting top management's support

CONCLUSION

The ERP Implementation project posed the challenges like Multi-location, Multi-language, Multiple currency implementation, and Network availability from a single instance and Adopting change management methodology during all phases. The other benefits can be brought down by the ERP System to the company are Significant cost savings, Global consolidation, Support for concurrent access demands of the company, Better analysis of operations and faster decision making, and Improved responsiveness to changing global market scenario. The successive stories of many Indian Companies and the viewpoints of experts clearly codified that ERP implementation in India is already initiated, but the problem in it is sustainable development and uninterrupted involvement till to reach the peak or attain the zenith.

REFERENCES

1. Asim Raj Singla, "Enterprise Resource Planning", 1st Edition 2008, Cengage Learning, New Delhi , pp.95.
2. Daniel E. O'Leary, ERP Implementation 2000. pp.115.
3. Data Quest, Volume XXVI No: 19, October 2008.
4. Data Quest, Volume XXVI No.22, November 2008 "Power CIOs", pp.56-62, 65-68, 74-78.
5. Data Quest, Volume XXVI No: 21, November 2008, pp. 41-43.
6. Data Quest, Volume XXVII, No.12, June 2009 "Infrastructure – IT industry's Top wish from the new Government", pp.26-35.
7. *Data Quest*, Volume XXXII No:11, 2009 Construction – the good, bad and Ugly -, pp. 60.
8. Dharminder Kumar & Sangeeth Gupta "Management Information Systems", 1st Edition, Excel Books, 2008,pp.109,155.
9. George Reynolds & Ralph Stairs "Principles of Information Systems", Cenage Learning, 1st Edition, 2008, pp.371-403.
10. Glynn C.William (2009), ERP Implementation – Implementing ERP Sales & Distributions, PHI, New Delhi.
11. Industrial Automation, Volume No:11, Issue No :V, December 2008, pp.14-19.
12. Information Technology, Vol 17, No: 01 Issue: 193, November 2007, pp.23.
13. IT Case Book 2009, Data Quest (Supplementary Issue), Vol XXVI, No: 23, December 2008.
14. James O' Brien & George M.Marakas "Management Information Systems", TMH, 7th Edition, 2007 Enterprise Resource Planning-The Business Backbone – pp.253-259.
15. Kenneth C.Laudon & Jane P.Laudon "Management Information Systems", Prentice Hall of India Publications Limited, 8th Edition, pp. 363.
16. PC Quest volume I, June 2009 "Best It Implementation of the Year 2009", pp.24-28, 91, 111-122.
17. PC Quest, published by Cyber Media, Volume II, July 2009 "Measuring the Impact", pp.36, 68, 71.
18. PC Quest, Volume IV,2009 "India's Favorite IT Brands", pp. 44, 51-56, 69-74, 119.
19. Waman S.Jawadekar "Management Information Systems", TMH, 3rd Edition, 2009, Enterprise Management Systems, pp. 473-503.
20. www.army.mil
21. www.sap.com

BIOGRAPHY

I am Ms. **R.Seranmadevi** qualified with B.com in UG level and enriched knowledge in versatile fields viz., MBA, M.Com, and MCA and undergone the research program in Management and awarded with M.Phil degree. At present working for CMS College Engineering, Namakkal affiliated to Anna University, Coimbatore. Earned experience in different fields including industry and educational area for about eight years. My profile is constructed with many paper presentations in both national and international level seminars / conferences, which depicts my thrust of interest.

Dr.M.Latha Natarajan qualified MBA., M.Phil., Ph.D., PGDPM. Totally she crossed 13 years in the field of academic. Profile is constructed with many paper presentations in various national and international level seminars and workshops and published 3 national level publications. The author has published two subject books to Periyar Institute of Distance Education.

IMPLEMENTATION OF MATRIX KEYBOARD IN DIGITAL X-RAY MACHINE BY USING EMBEDDED SYSTEM DESIGN

Vismita.Nagrale, R.N.Moghe, Sameer Kanse
MGM's Jawaharlal Nehru Engineering College
Department of E&TC, Aurangabad, India
Email : vismita.nagrale@gmail.com

Abstract - This paper describe the development and implementation of rotary switches converted into digital switch using embedded system design being used in x-ray machine for production facilities. Digital techniques are discrete techniques and they have been applied to automatic services, as this technology is economically viable.

The accuracy of voltage and current should be as required for x-ray of different parts of body. By using analog to digital converter, voltage and current can be increased or decreased in steps to ensure proper x-ray of the body part with high efficiency.

Keywords - Complementary metal–oxide–semi conductor (CMOS), metal–oxide–semiconductor field-effect transistor (MOSFET), Transistor–Transistor Logic (TTL), KiloVoltAmpere (KVA), American Standard Code for Information Interchange (ASCII).

INTRODUCTION

There are numerous changes in technology and electronics play an important role in innovations and discoveries. Industrial designs must satisfy time to market requirements. During the design phase the designer must be able to evaluate the overall system performance. The control panel of x-ray machine which was consisting of rotary switches converted fully to digital using embedded system design.

The development of the X-ray art reveals that during most of the period the principal advances were made in the field of medicine, with applications in two broad categories, namely, diagnosis and therapy. However, at the present time an important proportion of the X-ray apparatus in existence is devoted to industrial applications. In comparing industrial and medical applications of X rays, it becomes apparent that by far the greater proportion of industrial applications are diagnostic in nature; that is, the discovery of certain information about the internal structure of the material being irradiated is the object of the operation. In many cases, differential absorption, detected by film, fluorescent screen, or ionization device, is utilized to reveal the desired information just as in the case of medical diagnosis. In other cases, the diffraction of the rays into definite patterns, recorded by similar means, provides the information.

X-ray illumination, direction and with controllable, variable depth from the object surface, the main aim of X-ray machine is to find the defects, fractures or broken bones of human body. The X-ray machine's rays are passed through the concerned part of human body. These rays penetrate through the muscles and through the bones. Due to this property of X -ray an image is created and a full view of fractured bone is detected. These rays are harmful to human body; so basic parameters to be considered are

1. Power (KVA rating) of X -rays and
2. Time duration of X -rays existing on human body.

FACTORS AFFECTING X-RAY OUTPUT

The principal factors which affect the intensity of the X-ray beam issuing from an X-ray tube are:

(1) Target material;
(2) Anode voltage;
(3) Anode current;
(4) Absorption by intervening material; and
(5) Distance from focal spot to point of utilization.

TYPES OF SWITCHES

A. Rotary Switch

A rotary switch is a type of switch that is used on devices which have two or more different states or modes of operation, such as a three-speed fan or a CB radio with multiple frequencies of reception or channels.

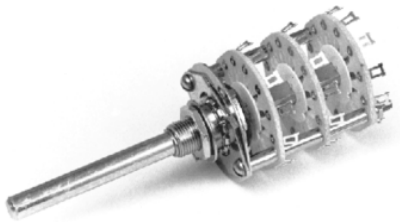


Fig.1. Rotary Switch

A rotary switch[1], consists of a spindle or rotor that has a contact arm or spoke which projects from its surface like a cam. It has an array of terminals, arranged in a circle around the rotor, each of which serves as a contact for the spoke through which any one of a number of different electrical circuits can be connected to the rotor. The switch is layered to allow the use of multiple poles, each layer is equivalent to one pole. Usually such a switch has a detent mechanism so it clicks from one active position to another rather than stalls in an intermediate position.

Rotary switches were used as channel selectors, range selectors on electrical metering equipment, as band selectors on multi-band radios, etc.

B. Analog Switch

The analogue (or analog) switch[2], also called the bilateral switch, is an electronic component that behaves in a similar way to a relay, but has no moving parts. The switching element is normally a MOSFET (Metal Oxide Semiconductor Field Effect Transistor), which is a type of transistor. The control input to the switch is a standard CMOS (Complementary Metal Oxide Semiconductor) or TTL (Transistor-Transistor Logic) logic input, which is shifted by internal circuitry to a suitable voltage for switching the MOSFET[3]. The result is that a logic 0 on the control input causes the MOSFET to have a high resistance, so that the switch is off, and a logic 1 on the input causes the MOSFET to have a low resistance, so that the switch is on. Analogue switches are usually manufactured as integrated circuits in packages containing multiple switches (typically two, four or eight).

Fig.2. shows a MOS (Metal Oxide Semiconductor) analog switch, M3 is a controlled transmission gate, M1 and current source form source follower. Input signal is connected to one side of M3, the other side of M3 is connected to the gate of M1, control signal Vc is connected to the gate cjf M3. When Vc is high, input signal passes M3 and reaches the gate of M1, output signal is achieved from the source of M1. Because of high input impedance of M1, the plug-in consumption is low, isolation between input signal and output signal is also realized.

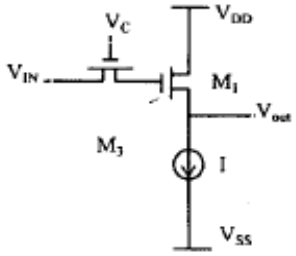


Fig.2. MOS analog switch

The resolution of switch depends on source follower. Unlike a relay, however, the analogue switch does not provide electrical isolation between the analogue signal and the control signal. This means that it should not be used in high-voltage circuits where such isolation is desired. Also, since there is only a low current path between the input and output, the maximum current allowed through the switch may be smaller than that in a typical relay. There are also some constraints on the polarity and range of voltages of the signal being switched.

C. Matrix Keyboard

To use different keyboards and keypads for different application needs certain data information or codes. In the proposed design of the keyboard matrix the codes are specially designed in ASCII and in return their HEX codes are given to the controller.

Fig.3. shows a conventional keyboard matrix scanning structure. The scanning lines of the keyboard which respectively connects to an X-port and a Y-port are typically arranged into a matrix, such as a matrix consisting of n rows and m columns of scanning lines. Each cross point corresponds to a position that one key is located and thus there are totally n x m keys. In the proposed designed it's a 5x7 matrix key board.

One scanning line of each key is electrically connected to the Y-port which serves as a scanning signal output port and has totally n lines. The other scanning line of each key is connected to X-port which serves as a scanning signal input port and has totally m lines. When a key is pressed, an electric conduction is constructed between two corresponding lines which cross at its corresponding position, and thus a scanning signal output from a corresponding line of Y-port will be sent to a line of X-port, thereby the position of the key being pressed is able to be detected. Depending upon the logic levels of both X-port and Y-port the codes have been calculated. Initially both are held at high logic level.

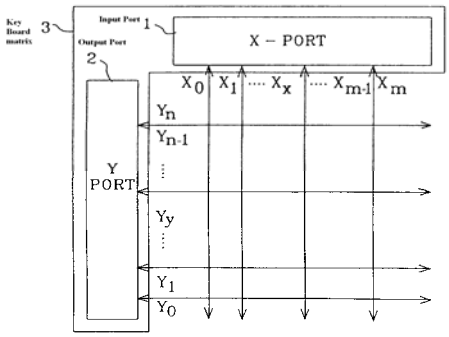


Fig.3.Matrix Keyboard structure

BASIC STRUCTURE

The basic structure of rotary switch control of X-ray Machine is shown in Fig. 4. The insufficient current selection of this control, a new enhanced digital system keyboard is shown in Fig. 5, where it automatically selects the current from the range of 25mA to 100mA where where it has a selection of 25, 50, 100mA for radiography for specified part.

It hunts down automatically to the necessary current and power rating and gives the clear image on the computer and then the X-ray is shooted. The image of X-ray is of enhanced quality

Fig.4.Rotary switch control panel of X-ray Machine

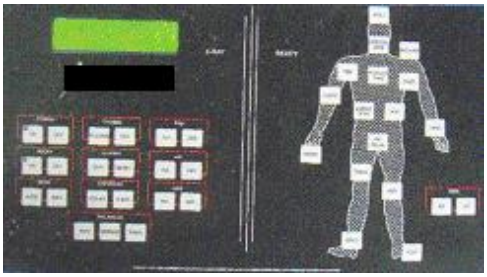


Fig.5.Digital control panel of X-ray Machine

The keyboard operates in two modes:

a. Manual

b. Automatic

a. Manual Mode Flow

The first few steps to be used as shown in automode. After selecting all parameters, KVP is set according to body part, mA is set and MAS is set i.e time for exposure.

For example, if mA is set to 25mA, KVP is set to 40 and time is for 2 sec., then so product of 25 mA X 2sec = 50 as this value is being used by MAS for exposure which is better for X-ray.

Then for firing, pressing the stand by key first then for 0.8 sec it is on stand by or more after that X-ray key is fired. Setting of KVP is actually setting values from Variac, the output of variac is fed to X-ray tube and to Port P1 and P2 which in return given to MAS, mA is given to the filament to F1 and F2.

b. Auto mode Flow

X-ray machine is put on. First auotmode is selected by default, if wanted to select manual mode then switch to manual key and press it to start the manual mode.

To select view, whether AP view or Lateral view, press respective switch, then select the body type like thin, medium or thick from the digital control of key board.

Next select which body part to be X-ray. All these switches are operated by operator and then automatically proper values of KVP are done and x-ray is shooted

The block diagram in Fig. 6 shows two micro-controller working specifically with keyboard through which input can be changed and to shoot the X-ray of required body part.

The design of Digital X-ray Machine can be divided in to three main parts:

1. Power control unit
2. Controller Card
3. Keyboard matrix

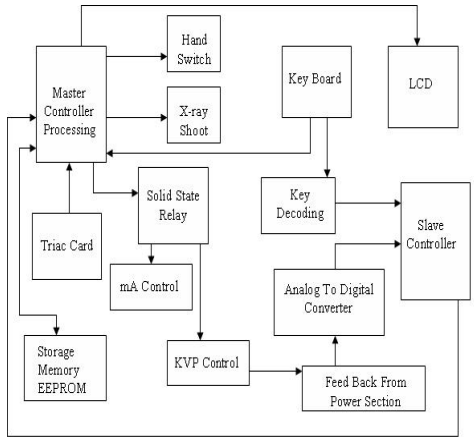


Fig.6.Block diagram of Digital X-ray Machine

Fig. 7 shows how the controls are automated as soon as power is on. There is no need to adjust voltage and current ratings manually as when the system is power on, the system scans the full body and comes to a centre point and by pressing the digital key of that particular part, voltage and current ratings are selected. The shooting of X-ray is completed and a optimized quality of X-ray is generated.

The voltage and current is inputted through keyboard and given to dedicated microcontroller, in return the microcontroller decodes the keys and translates the information to ASCII format. This information is fed to the master controller and it processes to the relay to hand shoot and X-ray is shooted. The machine is so compact and mobile that one not has to move the patient from one place to other to take X-ray in Radiography section.

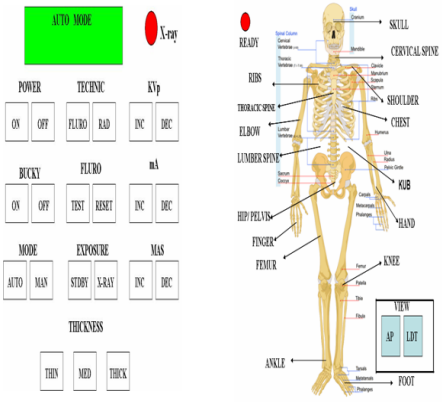


Fig.7.Body scan after power on. Display shows voltage and current ratings of digital control panel

HARDWARE DESIGN

Fig.8.300 volt span variac / 20 amp current rating, sweep 360 degree and motorosed. 230 volt servo motor

Fig.9.Power supply card

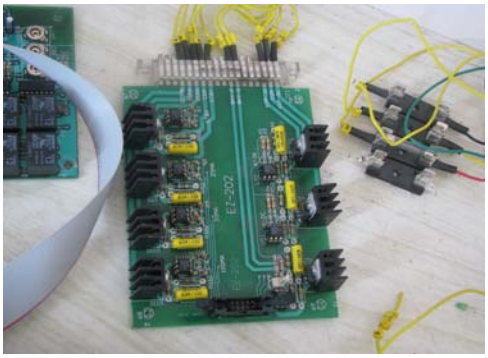


Fig.10.Triac card

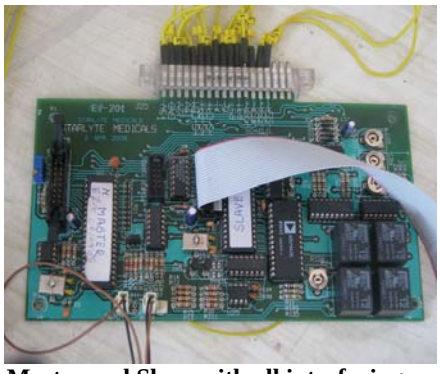


Fig.11.Master and Slave with all interfacing peripherals

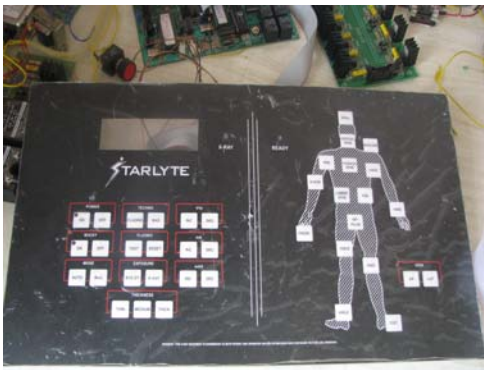


Fig.12.Digital matrix keyboard

EXPERIMENTAL RESULTS

The table no.1 shows the values of various body parts of human being (normal body) taken for quality X-ray.

Table -1
Ratings of Different Body Parts for a Normal Body type

Body parts	Current in mA	Exposure time to shoot X-ray (MAS) (sec)	KVP
Skull	50	62.5	65
Cervical spine	50	20	60
Shoulder	50	20	60
Ribs	50	50	65
Thoracic	50	62.5	65
Chest	50	20	55
Elbow	50	10	48
Lumber	50	75	76
Kub	50	75	80
Hand	50	6.5	45
Finger	50	3	43
Hip	50	75	75
Femur	50	20	60
Knee	50	20	53
Ankle	50	15	50
Foot	50	5	45

The figure 13, 14, 15 shows the amount of voltage and current ratings passed to shoot X-ray of the specific part of the body.

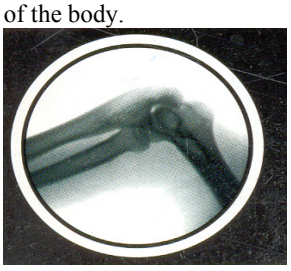


Fig.13.X-ray of leg part with tibia and fibula when the voltage and current are not according to the ratings required for proper intensity taken using rotary switch.

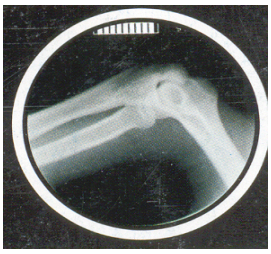


Fig.14.X-ray of leg part with tibia and fibula with all specified ratings taken using digital control

Fig.15.X-ray of hand wrist with carpals, metacarpals and phalanges taken using digital control

CONCLUSION

This paper has discussed some aspects of rotary switch replaced by matrix keyboard digital. Design techniques for digital are essential for high performance to shoot X-ray. A top down design methodology, based on digital to X-ray radiation analysis from early stages, improves the robustness and reduces the risk of failures in shooting X-ray of proper intensity of a specific body part. It is seen that for various types of bodies the mode of operation is different and the ratings are also different, to get a quality X-ray.

REFERENCES

[1] Joel. A. Jr., "Digital Switching--How It Has Developed," Communications, IEEE Transactions on [legacy, pre - 1988] Volume 27, Issue 7, Jul 1979 Page(s):948 – 959.

[2] Wu Ge, Shi Yin, "A one-way MOS analog switch," Solid-State and Integrated Circuit Technology, 1998. Proceedings. 1998 5th International Conference on, 21-23 Oct. 1998 Page(s):413 – 415.

[3] Ruegg, H.W., "An integrated FET analog switch," Proceedings of the IEEE Volume 52, Issue 12, Dec. 1964 Page(s):1572 – 1575.

[4] Encyclopedia, "X- ray," "X-ray Machine," "Rotary switch," "Analog switch," en.wikipedia.org.

BIOGRAPHY

Miss. Vismita Nagrale is a student of MGM's JNEC, Aurangabad and also a lecturer at AISSMS College of Engineering, Pune of Pune University. She received her B.Tech (App. Electronics) from University of Nagpur in 1994, her Masters Degree M.E. in electronics from Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra is going on. Her research interests include planning, implementation of design, simulation on various R&D projects in embedded systems and vlsi.

SPAM FILTERING USING K - NN

S.Ananthi
Senior Lecturer, Dept. of Applied Sciences,
Sethu Institute of Technology,
Viruthunagar (Dt)

Dr. S. Sathyabama
Professor, Dept. of MCA,
K.S. Rangasamy College of Technology
Tiruchengode-637 215

Abstract - This project performs a survey on the current spam filtering techniques, explores the use of k-Nearest-Neighbor algorithm as the basis for personalized spam filters. Several other classifiers such as Naive Bayesian classifier, Random Forest Tree, and Heuristic rules are combined to construct hybrid spam filtering systems to, if possible, improve the performance of classification. At the same time, heavy experiments are performed in preprocessing steps to compare their impacts on the same algorithm and the results are reported. Finally, this project participates spam filtering using k-NN algorithm.

INTRODUCTION

According to www.dictionary.com, spam is "unsolicited e-mail, often of a commercial nature, sent indiscriminately to multiple mailing lists, individuals, or newsgroups". Legitimate email, also called ham, on the opposite, is the email we want. It is reported that the first spam appeared in 1970s.Nowadays, spam is a problem for everyone with e-mail. Spam Cop, a web site to help users to report spam, estimated that between 60 and 80% of all emails is spam. There are hundreds of millions of spam per week and the number is still growing constantly. It is inconvenient and time-consuming to read and delete spam manually. Spammers bear little or no cost for distribution, yet normal receivers are forced to spend substantial time and effort deleting unwanted emails from their mailboxes. Fortunately, help has arrived. In the recent past, a large number of freely available software helps users to filter out annoying emails in a considerably good performance. Such software is called spam filter. The following section introduces the typical use of a spam filter, a categorization of spam filters, and a unified model of spam filtration.

SPAM FILTER

"A spam filter is a piece of software which takes an input of an email message. For its output, it might pass the message through unchanged for delivery to the user's mailbox, it might redirect the message for delivery elsewhere, or it might even throw the message away. Some spam filters are able to edit message during processing".

ALGORITHM APPLICATION

K - NN Algorithm

The "k"-nearest neighbors algorithm'' is amongst the simplest of all [machine learning] algorithms. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its "k" nearest neighbors. ''k" is a positive [integer], typically small. If "k" = 1, then the object is simply assigned to the class of its nearest neighbor. In binary (two class) classification problems, it is helpful to choose "k" to be an odd number as this avoids tied votes. The same method can be used for regression, by simply assigning the property value for the object to be the average of the values of its "k" nearest neighbors. It can be useful to weight the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones.

The neighbors are taken from a set of objects for which the correct classification (or, in the case of regression, the value of the property) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. In order to identify neighbors, the objects are represented by position vectors in a multidimensional feature space. It is usual to use the Euclidean distance, though other distance measures, such as the [[Manhattan distance]] could in principle be used instead. The "k"-nearest neighbor algorithm is sensitive to the local structure of the data. The test sample (green circle) should be classified either to the first class of blue squares or to the second class of red triangles. If "k = 3" it is classified to the second class because there are 2 triangles and only 1 square inside the inner circle. If "k = 5" it is classified to first class (3 squares vs. 2 triangles inside the outer circle).]

The training examples are vectors in a multidimensional feature space. The space is partitioned into regions by locations and labels of the training samples. A point in the space is assigned to the class "c" if it is the most frequent class label among the "k" nearest training samples. Usually Euclidean distance is used as the distance metric, however this will only work with numerical values. In cases such as text classification another metric, such as the '"overlap metric"' or Hamming distance can be used. The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples. In the actual classification phase, the test sample (whose class is not known) is represented as a vector in the feature space. Distances from the new vector to all stored vectors are computed and "k" closest samples are selected. There are a number of ways to classify the new vector to a particular class, one of the most used techniques is to predict the new vector to the most common class amongst the K nearest neighbors. A major drawback to using this technique to classify a new vector to a class is that the classes with the more

frequent examples tend to dominate the prediction of the new vector, as they tend to come up in the K nearest neighbors when the neighbors are computed due to their large number. One of the ways to overcome this problem is to take into account the distance of each K nearest neighbors with the new vector that is to be classified and predict the class of the new vector based on these distances.

The best choice of "k" depends upon the data; generally, larger values of "k" reduce the effect of noise on the classification, but make boundaries between classes less distinct. A good "k" can be selected by various [[heuristic (computer science)|heuristic]] techniques, for example, cross-validation. The special case where the class is predicted to be the class of the closest training sample (i.e. when "k" = 1) is called the nearest neighbor algorithm. The accuracy of the "k"-NN algorithm can be severely degraded by the presence of noisy or irrelevant features, or if the feature scales are not consistent with their importance. Much research effort has been put into feature selection selecting or scaling features to improve classification. A particularly popular approach is the use of evolutionary algorithms to optimize feature scaling. Another popular approach is to scale features by the mutual information of the training data with the training classes.

The naive version of the algorithm is easy to implement by computing the distances from the test sample to all stored vectors, but it is computationally intensive, especially when the size of the training set grows. Many nearest neighbor search algorithms have been proposed over the years; these generally seek to reduce the number of distance evaluations actually performed. Some optimizations involve partitioning the feature space, and only computing distances within specific nearby volumes. Several different types of nearest neighbor. The nearest neighbor algorithm has some strong consistency (statistics) consistency results. As the amount of data approaches infinity, the algorithm is guaranteed to yield an error rate no worse than twice the [Bayes error rate] (the minimum achievable error rate given the distribution of the data). "k"-nearest neighbor is guaranteed to approach the Bayes error rate, for some value of "k" (where "k" increases as a function of the number of data points).

The "k"-NN algorithm can also be adapted for use in estimating continuous variables. One such implementation uses an inverse distance weighted average of the "k"-nearest multivariate neighbors. This algorithm functions as follows:

Compute Euclidean from target plot to those that were sampled.

Order samples taking for account calculated distances.

Choose heuristically optimal "k" nearest neighbor based on RMSE done by cross validation technique.

Calculate an inverse distance weighted average with the "k"-nearest multivariate neighbors.

The k-nearest neighbors algorithm is amongst the simplest of all machine learning algorithms. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors. k is a positive integer, typically small. If k = 1, then the object is simply assigned to the class of its nearest neighbor. In binary (two class) classification problems, it is helpful to choose k to be an odd number as this avoids tied votes. The same method can be used for regression, by simply assigning the property value for the object to be the average of the values of its k nearest neighbors. It can be useful to weight the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones. the neighbors are taken from a set of objects for which the correct classification (or, in the case of regression, the value of the property) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. In order to identify neighbors, the objects are represented by position vectors in a multidimensional feature space. It is usual to use the Euclidean distance, though other distance measures, such as the Manhattan distance could in principle be used instead. The k-nearest neighbor algorithm is sensitive to the local structure of the data.

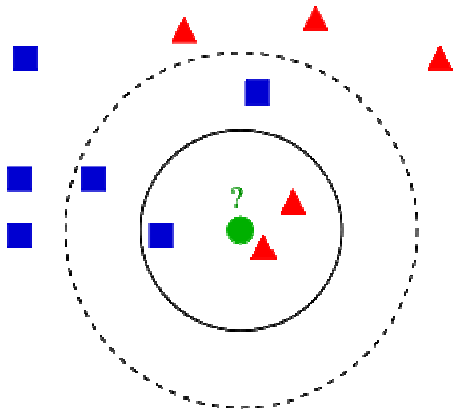


Fig.1. K-NN Classification

Example of *k*-NN classification. The test sample (green circle) should be classified either to the first class of blue squares or to the second class of red triangles. If $k = 3$ it is classified to the second class because there are 2 triangles and only 1 square inside the inner circle. If $k = 5$ it is classified to first class (3 squares vs. 2 triangles inside the outer circle).

The training examples are vectors in a multidimensional feature space. The space is partitioned into regions by locations and labels of the training samples. A point in the space is assigned to the class *c* if it is the most frequent class label among the *k* nearest training samples. Usually Euclidean

distance is used as the distance metric, however this will only work with numerical values. In cases such as text classification another metric, such as the overlap metric (or Hamming distance) can be used.

The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples. In the actual classification phase, the test sample (whose class is not known) is represented as a vector in the feature space. Distances from the new vector to all stored vectors are computed and *k* closest samples are selected. There are a number of ways to classify the new vector to a particular class, one of the most used techniques is to predict the new vector to the most common class amongst the K nearest neighbors. A major drawback to using this technique to classify a new vector to a class is that the classes with the more frequent examples tend to dominate the prediction of the new vector, as they tend to come up in the K nearest neighbors when the neighbors are computed due to their large number. One of the ways to overcome this problem is to take into account the distance of each K nearest neighbors with the new vector that is to be classified and predict the class of the new vector based on these distances.

PARAMETER SELECTION

The best choice of *k* depends upon the data; generally, larger values of *k* reduce the effect of noise on the classification, but make boundaries between classes less distinct. A good *k* can be selected by various heuristic techniques, for example, cross-validation. The special case where the class is predicted to be the class of the closest training sample (i.e. when $k = 1$) is called the nearest neighbor algorithm. The accuracy of the *k*-NN algorithm can be severely degraded by the presence of noisy or irrelevant features, or if the feature scales are not consistent with their importance. Much research effort has been put into selecting or scaling features to improve classification. A particularly popular approach is the use of evolutionary algorithms to optimize feature scaling. Another popular approach is to scale features by the mutual information of the training data with the training classes.

In Sorting Spam with K-Nearest-Neighbor and Hyperspace Classifiers Spam classification[1] continues to provide an interesting, if not vexing, field of research. This particular classifier problem is unique in the machine-learning field as most other problems in machine learning are not continuously made more difficult by intelligent and motivated minds. A number of different approaches have been taken for spam filtering beyond the single-ended machine learning filter and a reasonable survey really requires a full book; in this paper we will restrict ourselves to post-SMTP acceptance filtering and the sorting of email into two classes: good and spam.

Even within post-acceptance, single-ended filtering, there are a number of techniques available. One of the most common is a Naive Bayesian filter (usually using a limiting window of the most significant N words. Other common variations use chi-squared analysis, a Markov random field even compressibility of the unknown text given basis vectors representing the good and spam classes [Willets 2003]. Although K-NN filters [Fix and Hodges, 1951] [Cover and Hart, 1967] [have been considered for spam classification in the past (John Graham-Cumming's early POP file used a K-NN) they have fallen into disfavor among most filter authors. We reconsider the use of K-NNs for classification and attempt to quantize their qualities.

PURE VERSUS INCREMENTALLY TRAINED K-NN

One disadvantage of standard Cover and Hart style K-NNs is every known input is added to the stored data; this can cause very long compute times. To mitigate this, we have used selectively trained K-NNs; rather than adding every known text immediately to the stored data, we incrementally test each known text and only add the known text to the stored data if the known text was judged incorrectly. This speeds up filter classification tremendously. Those familiar with Cover and Hart's limit theorem should realize that this modification results in a K- NN that the Cover and Hart limit theorem does not necessarily apply to. We will consider extension of the Cover and Hart theorem to cover incremental trained K-NNs in future work.

In Non-Parametric Spam Filtering Based On K-NN and LSA[2]. The paper proposes a non-parametric approach to filtering of unsolicited commercial e-mail messages, also known as spam. The email messages text is represented as an LSA vector, which is then fed into a k - NN classifier. The method shows a high accuracy on a collection of recent personal email messages. Tests on the standard LINGSPAM collection achieve an accuracy of over 99.65%, which is an improvement on the best-published results to date.

The amount of unsolicited commercial e-mails (also known as spam) has grown tremendously during the last few years and today it already represents the majority of the Internet traffic. Although the spam is normally easy to recognize and delete, doing so on an everyday basis is inconvenient. Once an e-mail address has entered in the widespread spam distribution lists it could become almost unusable unless some automated measures are taken. Nowadays, it is largely recognized that the constantly changing spam form and contents requires filtering using machine learning (ML) techniques, allowing an automated training on up-to-date representative collections. The potential of learning has been first demonstrated by Sahami et al who used a Naïve Bayesian classifier. Several other researchers tried this thereafter and some specialized collections have been created to train ML algorithms. The most famous one LINGSPAM (a set of e-mail messages

from the *Linguist List*) is accepted as a standard set for evaluating the potential of different approaches.

TEXT CATEGORIZATION WITH K-NN

We used the k-nearest-neighbor classifier, which has been proved to be among the best performing text categorization algorithms in many cases [2]. K-NN calculates a similarity score between the document to be classified and each of the labeled documents in the training set. When k = 1 the class of the most similar document is selected. Otherwise, the classes of the k closest documents are used, taking into account their scores. We combined k - NN with latent semantic analysis (LSA). This is a popular technique for indexing, retrieval and analysis of textual data, and assumes a set of mutual latent dependencies between the terms and the contexts they are used in. This permits LSA to deal successfully with synonymy and partially with polysemy, which are the major problems with the word-based text processing techniques (due to the freedom and variability of expression). LSA is a two-stage process including learning and analysis. During the learning phase it is given a text collection and it produces a real-valued vector for each term and for each document. The second phase is the analysis when the proximity between a pair of documents or terms is calculated as the dot product between their normalized LSA vectors [2]. In our experiments, we built an LSA matrix (TF.IDF weighted) from the messages in the training set. The e-mail message to be classified is projected in the LSA space and then compared to each one from the training set. Then a K-NN classifier for a particular value of k predicts its class.

In Spam Mail Filtering Based On Network Processor[3], As the rapid development of the Internet, the occurrence of more and more spam mails becomes harmful to users. Content-based spam filtering technologies become the mainstream anti-spam mail methods so far. Support vector machine (SVM), Bayes, windows and K-N are excellent ones of these technologies and they have advantages and disadvantages respectively. The common shortage of content-based methods is that they can't filter spam mails as far as white-list-based, black-list-based or rule-based methods. This paper proposes a spam mail filtering system based on SVM[4] and Bayes which is implemented on Network Processor (NP). To content-based methods, this system preserves the filtering accuracy and takes advantage of the parallel processing abilities of NP to improve the filtering speed.

As the popularization of Internet, the efficient, convenient and cheep email has become the substitute of the conventional papery mail. But at the same time, spam mails cause lots of problems. At present, the anti-spam technologies can be divided as black-list-based, white-list-based, ruled-based and content-based. The black-list-based method has to maintain a real-time black list, which increases the management cost. The white-list-based method is able to filter 100% spam mails, but it may also filter some legitimate mails during the first communication. The rule-based method could find out 80% spam mails through simple rules, but it is difficult to improve the percentage and requires the user to be professional to build the rule-set. Compare with black-list-based, white-list-based and ruled-based methods, the content-based method has advantage of filtering accuracy, but it is slower. SVM and Bayes are two excellent methods to filter spam mails based on content. SVM is a statistical learning method which was developed during the 90s of twenty century. It classifies documents by constructing the best linear classifying plane and is recognized one of the best technologies for document classifying. .

CONCLUSION

Considering the disadvantages of discriminating the spam by mail body classification, is paper brings up a spam discriminating model based on SVM and D-S Identity Theory, the analysis of mail body and mail header features. Theory, there are a lot of ways to analyze Mail headers and mail bodies such as SVM with DSA.

REFERENCE

[1] William Yerazunis, Fidelis Assis, Christian Siefkes, Shalendra Chhabra Sorting Spam With K-Nearest-Neighbor And Hyperspace Classifiers

[2] Preslav Ivanov Nakov Panayot Markov Dobrikov Non-Parametric Spam Filtering based on k-NN and LSA

[3] Lian Lin, Zhongwen Li, Liang Shi 2008 IFIP International Conference on Network Spam Mail Filtering Based on Network Processor and Parallel Computing

[4] Chih-Chung Chang and Chih-Jen Lin_Last updated: February 27, 2009 LIBSVM: a Library for Support Vector Machines

FINDING ANOMALIES IN DATABASES

Mrs.K.Poongothai
Lecturer,
Dept. of Applied Science,
Selvam College of Technology

Mrs.M.Parimala
Lecturer,
Dept. of MCA,
M.Kumarasamy College of Engineering

Dr.S.Sathiyabama
Professor, Dept. of MCA,
K.S. Rangasamy College of Technology
Tiruchengode-637 215

Abstract - Association rules have become an important paradigm in knowledge discovery. Nevertheless, the huge number of rules which are usually obtained from standard datasets limits their applicability. In order to solve this problem, several solutions have been proposed, as the definition of subjective measures of interest for the rules or the use of more restrictive accuracy measures. Other approaches try to obtain different kinds of knowledge, referred to as peculiarities, infrequent rules, or exceptions. In general, the latter approaches are able to reduce the number of rules derived from the input dataset. This paper is focused on this topic. We introduce a new kind of rules, namely, anomalous rules, which can be viewed as association rules hidden by a dominant rule. We also develop an efficient algorithm to find all the anomalous rules existing in a database.

INTRODUCTION

Association rules have proved to be a practical tool in order to find tendencies in databases, and they have been extensively applied in areas such as market basket analysis and CRM (Customer Relationship Management). These practical applications have been made possible by the development of efficient algorithms to discover all the association rules in a database [11, 12, 4], as well as specialized parallel algorithms [1]. Related research on sequential patterns [2], associations varying over time [15], and associative classification models [5] have fostered the adoption of association rules in a wide range of data mining tasks.

Despite their proven applicability, association rules have serious drawbacks limiting their effective use. The main disadvantage stems from the large number of rules obtained even from small-sized databases, which may result in a second- order data mining problem. The existence of a large number of association rules makes them unmanageable for any human user, since she is overwhelmed with such a huge set of potentially useful relations. This disadvantage is a direct consequence of the type of knowledge the association rules try to extract, i.e., frequent and confident rules. Although it may be of interest in some application domains, where the expert tries to find unobserved frequent patters, it is not when we would like to extract hidden patterns. It has been noted that, in fact, the occurrence of a frequent event carries less information than the occurrence of a rare or hidden event. Therefore, it is often more interesting to find surprising non-frequent events than frequent ones [7, 12, 15]. In some sense, as mentioned in [7], the main cause behind the popularity of classical association rules is the possibility of building efficient algorithms to find all the rules which are present in a given database.

The crucial problem, then, is to determine which kind of events we are interested in, so that we can appropriately characterize them. Before we delve into the details, it should be stressed that the kinds of events we could be interested in are application-dependent. In other words, it depends on the type of knowledge we are looking for. For instance, we could be interested in finding infrequent rules for intrusion detection in computer systems, exceptions to classical associations for the detection of conflicting medicine therapies, or unusual short sequences of nucleotides in genome sequencing.

Our objective in this paper is to introduce a new kind of rule describing a type of knowledge we might me interested in, what we will call anomalous association rules henceforth. Anomalous association rules are confident rules representing homogeneous deviations from common behavior. This common behavior can be modeled by standard association rules and, therefore, it can be said that anomalous association rules are hidden by a dominant association rule. In the following section, we review some related work. We shall justify the need to define the concept of anomalous rule as something complementary to the study of exception rules. Section 3 contains the formal definition of anomalous association rules. Section 4 presents an efficient algorithm to detect this kind of rules. Finally, Section 5 discusses some experimental results.

MOTIVATION AND RELATED WORK

Several proposals have appeared in the data mining literature that try to reduce the number of associations obtained in a mining process, just to make them manageable by an expert. According to the terminology used in [6], we can distinguish between user-driven and data-driven approaches (also referred to as subjective and objective interestingness measures, respectively [15], although we prefer the first terminology).Let us remark that, once we have obtained the set of good rules (considered as such by any interestingness measure), we can apply filtering techniques such as eliminating redundant tuples [15] or evaluating the rules according to other interestingness measures in order to check (at least, in some extent) their degree of surprising ness, i.e., if the rules convey new and useful information which could be viewed as unexpected [8, 9, 11, 6]. Some proposals [13, 15] even intro- duce alternative interestingness measures which are strongly related to the kind of knowledge

they try to extract. In user-driven approaches, an expert must intervene in some way: by stat- ing some restriction about the potential attributes which may appear in a reallocation [12], by imposing a hierarchical taxonomy [10], by indicating potential useful rules according to some prior knowledge [15], or just by eliminating non-interesting rules in a first step so that other rules can automatically removed in subsequent steps [18].On the other hand, data-driven approaches do not require the intervention of a human expert. They try to autonomously obtain more restrictive rules. This is mainly accomplished by two approaches:

a) Using interestingness measures differing from the usual support-confidence pair [14, 12].
b) Looking for other kinds of knowledge which are not even considered by classical association rule mining algorithms.

The latter approach pursues the objective of finding surprising rules in the sense that an informative rule has not necessary to be a frequent one. The work we present here is in line with this second data-driven approach. We shall introduce a new kind of association rules that we will call anomalous rules. Before we briefly review existing proposals in order to put our approach in context, we will describe the notation we will use henceforth. From now on, X,Y, Z, and A shall denote arbitrary item sets. The support and confidence of an association rule X) Y are defined as usual and they will be represented by supp(X) Y) and conf(X) Y), respectively. The usual minimum support and confident thresholds are denoted by M inSupp and M inc onf, respectively. A frequent rule is a rule with high support (greater than or equal to the sup- port threshold M inSupp), while a confident rule is a rule with high confidence (greater than or equal to the confidence threshold M inC onf). A strong rule is a classical association rule, i.e., a frequent and confident one [7] try to find non-frequent but highly correlated item sets, whereas [12] aims to obtain peculiarities defined as non-frequent but highly confident rules according to a nearness measure defined over each attribute, i.e., a peculiarity must be significantly far away from the rest of individuals. [9] finds unusual sequences, in the sense that items with low probability of occurrence are not expected to be together in several sequences. If so, a surprising sequence has been found. Another interesting approach [13, 10, 3] consists of looking for exceptions, in the sense that the presence of an attribute interacting with another may change the consequent in a strong association rule. The general form of an exception rule is introduced in [13, 15] as follows:

X) Y X Z): Y X 6) Z

Here, X) Y is a common sense rule (a strong rule). X Z): Y is the exception, where: Y could be a concrete value E (the Exception [12]). Finally, X 6) Z is a reference rule. It should be noted that we have simplified the definition of exceptions since the authors use five [13] or more [15] parameters which have to be settled beforehand, which could be viewed as a shortcoming of their discovery techniques. In general terms, the kind of knowledge these exceptions try to capture can be interpreted as follows:

X strongly implies Y (and not Z).
But, in conjunction with Z, X does not imply Y
(Maybe it implies another E)

For example [14], if X represents antibiotics, Y recovery, Z staphylococci, and E death, then the following rule might be discovered: with the help of antibiotics, the patient usually tends to recover, unless staphylococci appear; in such a case, antibiotics combined with staphylococci may lead to death. This is a very interesting kind of knowledge which cannot be detected by traditional association rules because the exceptions are hidden by a dominant rule. However, there are other exceptional associations which cannot be detected by applying the approach described above. For instance, in scientific experimental, it is usual to have two groups of individuals: one of them is given a placebo and the other one is treated with some real medicine. The scientist wants to discover if there are significant differences in both populations, perhaps with re spect to a variable Y. In those cases, where the change is significant, an ANOVA or contingency analysis is enough. Unfortunately, this is not always the case. What the scientist obtains is that both populations exhibit a similar behavior except in some rare cases. These infrequent events are the interesting ones for the scientist because they indicate that something happened to those individuals and the study must continue in order to determine the possible causes of this unusual change of behavior. In the ideal case, the scientist has recorded the values of a set of variables Z for both populations and, by performing an exception rule analysis, he could conclude that the interaction between two item sets X and Z (where Z is the item set corresponding to the values of Z) change the common behavior when X is present (and Z is not). However, the scientist does not always keep records of all the relevant variables for the experiment. He might not even be aware of which variables are really relevant. Therefore, in general, we cannot not derive any conclusion about the potential changes the medicine causes. In this case, the use of an alternative discovery mechanism is necessary. In the next section, we present such an

alternative which might help our scientist to discover behavioral changes caused by the medicine he is testing

DEFINING ANOMALOUS ASSOCIATION RULES

An anomalous association rule is an association rule that comes to the surface when we eliminate the dominant effect produced by a strong rule. In other words, it is an association rule that is verified when a common rule fails. In this paper, we will assume that rules are derived from item sets containing discrete values. Formally, we can give the following definition to anomalous association rules:

Definition 1. Let X, Y, and A be arbitrary item sets. We say that X A is an anomalous rule with respect to X) Y , where A denotes the Anomaly, if the following conditions hold:

a) X) Y is a strong rule (frequent and confident)
b) X :Y) A is a confident rule c) X Y): A is a confident rule

It should be noted that, implicitly in the definition, we have used the common minimum support (M inSupp) and confidence (M inC onf) thresholds, since they tell us which rules are frequent and confident, respectively. For the sake of simplicity, we have not explicitly mentioned them in the definition. A minimum support threshold is relevant to condition a), while the same minimum confidence threshold is used in conditions a), b), and c). The semantics this kind of rules tries to capture is the following:

X strongly implies Y, but in those cases where we do not obtain Y, then X confidently implies A

In other words: When X, then we have either Y (usually) or A (unusually) Therefore, anomalous association rules represent homogeneous deviations from the usual behavior. For instance, we could be interested in situations where a common rule holds:

If symptoms-X then disease-Y

Where the rule does not hold, we might discover an interesting anomaly:

if symptoms-X then disease-A

When not disease-Y

If we compare our definition with Hussain and Suzuki's [13, 12], we can see that they correspond to different semantics. Attending to our formal definition, our approximation does not require the existence of the conflictive itemset (what we called Z when describing Hussain and Suzuki's approach in the previous section). Furthermore, we impose that the majority of exceptions must correspond to the same consequent A in order to be considered an anomaly. In order to illustrate these differences, let us consider the relation shown in Figure 1, where we have selected those records containing X. From this dataset, we obtain conf(X)Y)=0:6, conf(X Z):Y)=conf(XZ)A)=1, and conf(X)Z)=0:2. If we suppose that the item set XY satisfies the support threshold and we use 0:6 as confidence threshold, then \X Z)A is an exception to X)Y , with reference rule X):Z". This exception is not highlighted as an anomaly using our approach because A is not always present when X: Y. In fact, conf(X:Y)A) is only 0:5, which is below the minimum confidence threshold 0:6. On the other hand, let us consider the relation in Figure 2, which shows two examples where an anomaly is not an exception. In the second example, we find that conf(X)Y) = 0:8, conf(X Y) :A) = 0:75, and conf(X :Y) A) = 1. No Z -value exists to originate an exception, but X

X Y A4 Z3 · · ·
X Y A1 Z1 · · ·
X Y A2 Z2 · · ·
X Y A1 Z3 · · ·
X Y A2 Z1 · · ·
X Y A3 Z2 · · ·
X Y1 A4 Z3 · · ·
X Y2 A4 Z1 · · ·
X Y3 A Z · · ·
X Y4 A Z · · ·
· · ·

Fig.1. A is an exception to X Y when Z, but that anomaly is not confident enough to be considered an anomalous rule

The table in Figure 1 also shows that when the number of variables (at-tributes in a relational database) is high, then the chance of finding spurious Z item sets correlated with :Y notably increases. As a consequence, the number of rules obtained can be really high (see [15, 13] for empirical results). The semantics we have attributed to our anomalies is more restrictive than exceptions and, thus, when the expert is interested in this kind of knowledge, then he will obtain a more manageable number of rules to explore. Moreover, we do not require the existence of a Z explaining the exception. In particular, we have observed that users are usually interested in anomalies involving one item in their consequent. A more rational explanation of this fact might have psychological roots: As humans, we tend to find more problems when reasoning about negated facts. Since the anomaly introduces a negation in the

X Y Z1 · · ·
X Y Z2 · · ·
X Y Z · · ·
X Y Z · · ·
X Y Z · · ·
X Y Z · · ·

X A Z · · ·
X A Z · · ·
X A Z · · ·
X A Z · · ·
· · ·
X Y A1 Z1 · · ·
X Y A1 Z2 · · ·
X Y A2 Z3 · · ·
X Y A2 Z1 · · ·
X Y A3 Z2 · · ·
X Y A3 Z3 · · ·
X Y A Z · · ·
X Y A Z · · ·
X Y3 A Z · · ·
X Y4 A Z · · ·
· · ·

Fig.2. X→ A is detected as an anomalous rule, even when no exception can be found through the Z-values

In the Figure 2 A is detected as an anomalous rule, even when no exception can be found through the Z -values. rule antecedent, experts tend to look for `simple' understandable anomalies in order to detect unexpected facts. For instance, an expert physician might directly look for the anomalies related to common symptoms when these symptoms are not caused by the most probable cause (that is, the usual disease she would diagnose). The following section explores the implementation details associated to the discovery of such kind of anomalous association rules. Remark. It should be noted that, the more confident the rules X:Y) A and X Y):A are, the stronger the anomaly is. This fact could be useful in order to define a degree of strength associated to the anomaly.

DISCOVERING ANOMALOUS ASSOCIATION RULES

Given a database, mining conventional association rules consists of generating all the association rules whose support and confidence are greater than some user-specified minimum thresholds. We will use the traditional decomposition of the association rule mining process to obtain all the anomalous association rules existing in the database:

{Finding all the relevant item sets.
{Generating the association rules derived from the previously-obtained item- sets.

For instance, Apriori-based algorithms are iterative [8]. Each iteration consists of two phases. The first phase, candidate generation, generates potentially frequent k-item sets (Ck) from the previously obtained frequent (k-1)-item set (Lk−1). The second phase, support counting, scans the database to find the actual frequent k-item sets (Lk). Apriori-based algorithms are based on the fact that that all subsets of a frequent item set is also frequent. This allows for the generation of a reduced set of candidate itemsets. Nevertheless, it should be noted that the there is no actual need to build a candidate set of potentially frequent item sets [11]. In the case of anomalous association rules, when we say that X A is an anomalous rule with respect to X) Y , that means that the item set X[:Y[A appears often when the rule X) Y does not hold. Since it represents an anomaly, by definition, we cannot establish any minimum support threshold for X[:Y[A. In fact, an anomaly is not usually frequent in the whole database. Therefore, standard association rule mining algorithms cannot be used to detect anomalies without modification. Given an anomalous association rule X A with respect to X) Y, let us denote by R the subset of the database that, containing X, does not verify the association rule X)Y. In other words, R will be the part of the database that does not verify the rule and might host an anomaly. When we write suppR (X), it actually represents supp(X[:Y) in the complete database. Although this value is not usually computed when obtaining the item sets, it can be easily computed as supp(X) supp(X[Y). Both values in this expression are always available after the conventional association rule mining process, since both X and X [Y are frequent item sets. Applying the same reasoning, the following expression can be derived to represent the confidence of the anomaly X A with respect to X) Y :

ConfR (X A) = supp(X [A) − supp(X [Y [A)
Supp(X) − supp(X [Y)

Fortunately, when somebody is looking for anomalies, he is usually interested in anomalies involving individual items. We can exploit this fact by taking into account that, even when X [A and X [Y [A might not be frequent, they are extensions of the frequent item sets X and X [Y, respectively. Since A will represent individual items, our problem reduces to being able to compute the support of L [i, for each frequent item set L and item i potentially involved in an anomaly. Therefore, we can modify existing iterative association rule mining algorithms to efficiently obtain all the anomalies in the database by modifying the support counting phase to compute the support for frequent item set extensions:

{Candidate generation : As in any Apriori-based algorithm, we generate potentially frequent k-item sets from the frequent item sets of size k-1.
{Database scan: The database is read to collect the information needed to compute the rule confidence for potential anomalies. This phase involves two parallel tasks Candidate support counting: The frequency of each candidate k-item set is obtained by scanning the database in order to obtain the actual frequent k-item sets. Extension support counting: At the same time that candidate

support is computed, the frequency of each frequent k-1-itemset extension can also be obtained. Once we obtain the last set of frequent item sets, an additional database scan can be used to compute the support for the extensions of the larger frequent item sets. Using a variation of a standard association rule mining algorithm as TBAR [4], nicknamed ATBAR (Anomaly TBAR), we can efficiently compute the support for each frequent item set as well as the support for its extensions. In order to discover existing anomalies, a tree data structure is built to store all the support values needed to check potential anomalies. This tree is an extended version of the typical item set tree used by algorithms like TBAR [3]. The extended item set tree stores the support for frequent item set extensions as well as for all the frequent item sets themselves. Once we have these values, all anomalous association rules can be obtained by the proper traversal of this tree-shaped data structure. In interactive applications, the human user can also use the aforementioned extended item set tree as an index to explore a database in the quest for anomalies.

CONCLUSION AND FUTURE WORK

In this paper, we have studied situations where standard association rules do not provide the information the user seeks. Anomalous association rules have proved helpful in order to represent the kind of knowledge the user might be looking for when analyzing deviations from normal behavior. The normal behavior is modeled by conventional association rules, and the anomalous association rules are association rules which hold when the conventional rules fail. We have also developed an efficient algorithm to mine anomalies from databases. Our algorithm, ATBAR, is suitable for the discovery of anomalies in large databases. We intend to apply our technique to real problems involving datasets from the biomedical domain. Our approach could also prove useful in tasks such as fraud identification, intrusion detection systems and, in general, any application where the user is not really interested in the most common patterns, but in those patterns which differ from the norm.

REFERENCES

[1] R.Agrawal and J.Shafer. Parallel mining of association rules. IEEE Transactions on Knowledge and Data Engineering, 8(6):962{969, 1996.

[2] Rakesh Agrawal and Ramakrishnan Srikant. Mining sequential patterns. In Philip S. Yu and Arbee S.P.Chen, editors, Eleventh International Conference on Data Engineering, pages 3{14, Taipei, Taiwan, 1995. IEEE Computer Society Press.

[3] Yonatan Aumann and Yehuda Lindellt. A statistical theory for quantitative association rules. Journal of Intelligent Information Systems, 20(3):255{283, 2003.

[4] F. Berzal, J.C. Cubero, J.M. Marn, and J.M. Serrano. An efficient method for association rule mining in relational databases. Data and Knowledge Engineering, 37:47{84, 2001.

[5] F. Berzal, J.C. Cubero, Daniel Snchez, and J.M. Serrano. Art: A hybrid classifi- cation model. Machine Learning, 54(1):67{92, 2004.

[6] DR Carvalho, AA Freitas, and NFF Ebecken. A critical review of rule surprising-ness measures. In NFF Ebecken, CA Brebbia, and A Zanasi, editors, Proc. Data Mining IV Int. Conf. On Data Mining, pages 545{556. WIT Press, December 2003.

[7] Edith Cohen, Mayur Datar, Shinji Fujiwara, Aristides Gionis, Piotr Indyk, Rajeev Motwani, Jeffrey D. Ullman, and Cheng Yang. Finding interesting associations without support pruning. IEEE Transactions on Knowledge and Data Engineering, 13(1):64{78, 2001.

[8] AA Freitas. On Rule Interestingness Measures. Knowledge-Based Systems, 12(5-6):309{315, October 1999.

[9] Alex Alves Freitas. On objective measures of rule surprisingness. In Principles of Data Mining and Knowledge Discovery, pages 1{9, 1998.

[10] J.Hanand Y.Fu.Discovery of multiple-level association rules from large databases. In Proceedings of the VLDB Conference, pages 420{431, 1995.

[11] Jiawei Han, J. Pei, and Y.Yin. Mining frequent patterns without candidate generation. In Proceedings of the 1998 ACM SIGMOD international conference on Management of Data, pages 1{12, 2000.

[12] C. Hidber. Online association rule mining. In Proceedings of the 1999 ACM SIG-MOD international conference on Management of Data, pages 145{156, 1999.

[13] Farhad Hussain, Huan Liu, Einoshin Suzuki, and Hongjun Lu. Exception rule mining with a relative interestingness measure. In Pacific-Asia Conference on Knowledge Discovery and Data Mining, pages 86{97, 2000.

[14] Yves Kodratoff. Comparing machine learning and knowledge discovery in DataBases : An application to knowledge discovery in texts. In Machine Learn- ing and its Applications, volume 2049, pages 1{21. Lecture Notes in Computer Science, 2001.

[15] Bing Liu, Wynne Hsu, Shu Chen, and Yiming Ma. Analyzing the subjective interestingness of association rules. IEEE Intelligent Systems, pages 47{55, 2000.

AN EMERGING ANT COLONY OPTIMIZATION ROUTING ALGORITHM (ACORA) FOR MANETS

M. Subha,
Lecturer,
Dept. of MCA,
Vellalar College For Women, Erode-12,
Email: subha_sangeeth@rediffmail.com

Dr. R. Anitha,
Director,
Dept. of MCA,
K.S.Rangasamy College of Technology,
Tiruchengode,
Email:aniraniraj@rediffmail.com

Abstract - Ad hoc networks are characterized by multi-hop wireless connectivity, frequently changing network topology and the need for efficient dynamic routing protocols. A mobile ad hoc network (Manet) is a collection of mobile nodes which communicate over radio. These kinds of networks are very flexible, thus they do not require any existing infrastructure or central administration. Therefore, mobile ad hoc networks are suitable for temporary communication links. The biggest challenge in this kind of networks is to find a path between the communication end points, what is aggravated through the node mobility.

In this paper we present a new ad-hoc routing algorithm ACORA, which is based on swarm intelligence. We refer to the protocol as the Ant Colony Optimization Routing Algorithm (ACORA). Ant colony algorithms are a subset of swarm intelligence and consider the ability of simple ants to solve complex problems by cooperation. Several algorithms which are based on ant colony problems were introduced in recent years to solve different problems, e.g. optimization problems. We aim to show that the approach has the potential to become an appropriate algorithm for mobile multi-hop ad-hoc networks, which are based on simulations made with the implementation in ns-2.

Keywords - Manet, Swarm Intelligence, ACO, ACORA

INTRODUCTION

An ad-hoc network consists of a set of nodes that communicate using a wireless medium over single or multiple hops and do not need any preexisting infrastructure such as access points or base stations. Therefore, mobile ad-hoc networks are suitable for temporary communication links. The biggest challenge in this kind of networks is to find a path between the communication end points, what is aggravated through the node mobility.

The routing scheme in a MANET can be classified into two major categories-Proactive and Reactive. The proactive or table driven routing protocols (DSDV) maintain routes between all node pairs all the time. It uses periodic broadcast advertisements to keep routing table up-to-date. This approach suffers from problems like increased overhead, reduced scalability and lack of flexibility to respond to dynamic changes. The reactive or on-demand (DSR, AODV) approach is event driven and the routing information is exchanged only when the demand arises. The route discovery is initiated by the source. Hybrid approaches combines the features of both the approaches [2].

In this paper we present a new routing algorithm ACORA for mobile, multi-hop ad-hoc networks to improve the performance of the existing protocol of mobile ad hoc network. The protocol is based on swarm intelligence and especially on the ant colony based metaheuristic. The proposed algorithm is implemented in ns-2 [10, 11 12 and 13].

Basics of Swarm Intelligence Systems

The emergent behavior of self-organization in a group of social insects is known as swarm intelligence. There are two popular swarm-inspired methods in computational intelligence areas: Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO) which is out off focus in this paper. ACO was inspired by the behavior of ants [1, 3, 4, and 5].

The basic idea of the ant colony optimization metaheuristic is taken from the food searching behavior of real ants. When ants are on they way to search for food, they start from their nest and walk toward the food. When an ant reaches an intersection, it has to decide which branch to take next. While walking, ants deposit *pheromone*1, which marks the route taken. The concentration of pheromone on a certain path is an indication of its usage. With time the concentration of pheromone decreases due to diffusion effects. This property is important because it is integrating dynamic into the path searching process.

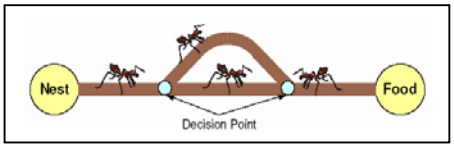


Fig.1 All ants take the shortest path after an initial searching time

Figure 1 shows a scenario with two routes from the nest to the food place. At the intersection, the first ants randomly select the next branch. Since the below route is shorter than the upper one, the ants which take this path will reach the food place first. On their way back to the nest, the ants again have to select a path. After a short time the pheromone concentration on the shorter path will be higher than on the longer

path, because the ants using the shorter path will increase the pheromone concentration faster. The shortest path will thus be identified and eventually all ants will only use this one.

This behavior of the ants can be used to find the shortest path in networks. Especially, the dynamic component of this method allows a high adaptation to changes in mobile ad-hoc network topology, since in these networks the existence of links are not guaranteed and link changes occur very often. Also ACO mainly suits for ad hoc networks due to link quality, local work and support for multipath [4 and 5].

Why ACO suits to ad hoc networks ?

We discuss various reasons by considering important properties of mobile ad hoc networks.

Dynamic topology is responsible for the bad performance of several routing algorithms in mobile multi-hop ad hoc networks. The ant Colony optimization meta-heuristic is based on agent systems and works with individual ants. This allows a high adaptation to the current topology of the network.

Local work - In contrast to other routing approaches, ant Colony optimization meta-heuristic is based only on local information i.e., no routing tables or other information blocks have to be transmitted to neighbors or to all nodes of the network.

Link quality is possible to integrate the connection/link quality into the computation of the pheromone concentration, especially into the evaporation process. This will improve the decision process with respect to the link quality. It is here important to notice, that the approach has to be modified so that nodes can also manipulate the pheromone concentration independent of the ants, i.e. data packets, for this a node has to monitor the link quality.

Support for multi-path - Each node has a routing table with entries for all its neighbors, which contains also the pheromone concentration. The decision rule, to select the next node, is based on the pheromone concentration on the current node, which is provided for each possible link. Thus, the approach supports multi-path routing.

ANT COLONY OPTIMIZATION ROUTING ALGORITHM (ACORA)

The ant colony optimization algorithm (ACORA), is a probabilistic technique for solving computational problems which can be reduced to finding good paths through graphs. This algorithm is a member of ant colony algorithms family, in swarm intelligence methods, and it constitutes some meta heuristic optimizations.

ACORA has two phases. They are: Route Discovery phase and Route Maintenance phase. Both the phases use these FAnt (figure 2a), BAnt (Figure 3a) and merging FAnt and BAnt (Figure 4a). The FAnt is for the collection of information and BAnt is for feedback to the forwarding mode.

Route Discovery

In the route discovery phase new routes are created. The creation of new routes requires the use of a *Forward Ant* (FAnt) and a *Backward Ant* (BAnt). A FAnt is an agent which establishes the pheromone track to the source node. In contrast, a BAnt establishes the pheromone track to the destination node. The FAnt is a small packet with a unique sequence number. Nodes are able to distinguish duplicate packets on the basis of the sequence number and the source address of the FAnt.

A Forward Ant is broadcasted by the sender and will be relayed by the neighbors of the sender (figure 2b). A node receiving a FAnt for the first time creates a record in its routing table. A record in the routing table is a triple and consists of (destination address, next hop, pheromone value). The node interprets the source address of the Forward Ant as destination address, the address of the previous node as the next hop, and computes the pheromone value depending on the number of hops the FAnt needed to reach the node.

Then the node relays the FAnt to the neighbors. Duplicate FAnt are identified through the unique sequence number and destroyed by the nodes. When the FAnt reaches the destination node, it is processed in a special way.

The destination node extracts the information of the FAnt and destroys it. Subsequently, it creates a BAnt and sends to the source node (Figure 3b). The BAnt has the same task as the FAnt, i.e. establishing a track to this node. When the sender receives the BAnt from the destination node, the path is established and data packets can be sent.

```
FOR EACH periods DO
Get timer (); Set prone ACORA ();
IF (pheromone ACORA ant return) THEN
Get timer (); bandwidthi = Get bandwidth (pathi);
delayi =Get delay (pathi); packet-lossi = Get packet-
loss (pathi);
Throughput =Get Destination (packeti)-Get Source-
send (packet i)*100;
END IF
```

Fig.2a. FAnt algorithm for ACORA.

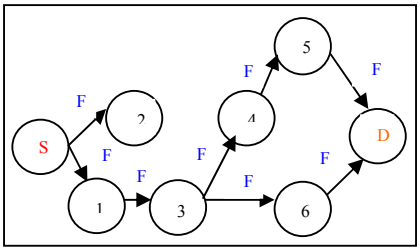


Fig.2b. A FAnt (F) is send from the sender (S) towards the destination node (D). The FAnt is relayed by other nodes, which initialize their routing table and the pheromone values

IF ACORA = TRUE THEN
DAi = Get Destination addr (packeti); SAi = Get Source addr (packeti);
IF transport layer packet = TRUE THEN
DPAi = Get Destination port (packeti); SPAi = Get Sourceport (packeti);
next-hop address lookup (DAi, SAi, DPAi, SPAi);
ELSE next-hop address lookup (DAi, SAi);
END IF
ELSE next-hop address lookup (DAi)

DA → Destination Address
SA→ Source Address
SPA→ Source Port Address
i→ is the number of data transmitted in path
DPA→ Destination Port Address

Fig.3a. BAnt algorithm for ACORA

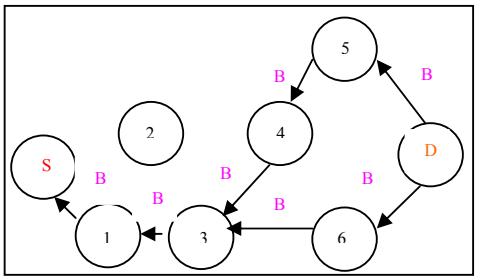


Figure 3b. The BAnt (B) has the same task as the FAnt. It is send by the destination node toward the source node

WHILE NOT (all packets ACORA grouped)
FIND DAi = Get Destination addr (packeti); SAi = Get Source addr (packeti) pair have the largest value in the pheromone
IF (DAi or SAi) already grouped means change the index value
GROUP (DAi or SAi)
COMPUTE (distance) VALUE of DAi and SAi
DELETE DAi, SAi (repeated key value)
END

Fig.4a. Merging FAnt and BAnt algorithm for ACORA

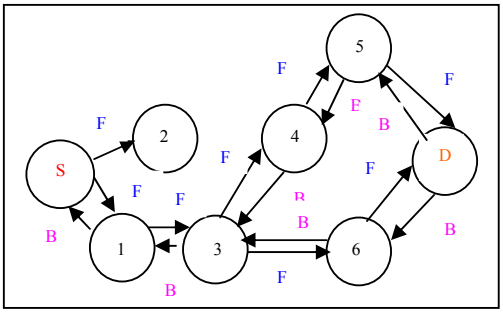


Fig.4b. Route Establishment between source node (S) and destination node (D) to send data packets by FAnt and BAnt.

Route Maintenance

The second phase of the routing algorithm is called route maintenance, which is responsible for the improvement of the routes during the communication. ACORA does not need any special packets for route maintenance. Once the FAnt and BAnt have established the pheromone tracks for the source and destination nodes, subsequent data packets are used to maintain the path. Similar to the nature, established paths do not keep their initial pheromone values forever. When a node (relay node) relays a data packet toward the destination (destination address*)* to a neighbor node (next hop), it increases the pheromone value of the entry (destination address, next hop, pheromone value*)* by pheromone function, i.e., the path to the destination is strengthened by the data packets. In contrast, the next hop (next hop) increases the pheromone value of the entry (source address, relay node, pheromone value) by pheromone function, i.e. the path to the source node is also strengthened. The evaporation process of the real pheromone is simulated by regular decreasing of the pheromone values.

The above method for route maintenance could lead to undesired loops. ACORA prevents loops by a very simple method, which is also used during the route discovery phase. Nodes can recognize duplicate receptions of data packets, based on the source address and the sequence number. If a node receives a duplicate packet, it sets the DUPLICATE ERROR flag and sends the packet back to the previous node. The previous node deactivates the link to this node, so that data packets cannot be sending to this direction any more.

ACORA handles routing failures, which are caused especially through node mobility and thus very common in mobile ad-hoc networks. ACORA recognizes a route failure through a missing acknowledgement. If a node gets a ROUTE ERROR message for a certain link, it first deactivates this link by setting the pheromone value to 0. Then the node searches for an alternative link in its routing table. If there is a second link it sends the packet via this path. Otherwise the node informs its neighbors, hoping that they can relay the packet. Either the packet can be transported to the destination node or the backtracking continues to the source node. If the packet does not reach the destination, the source has to initiate a new route discovery phase.

SIMULATION RESULTS

The ACORA is simulated under Linux Fedora-8, using the network simulator NS2 version ns-allinone-2.33. The network surface used is 1000m*1000m. The mobility scenarios are generated by the automatic generator *setdest* provided by NS2. The maximal speed of members is defined at 5km/h. The pause time is 20 seconds. The simulation duration is 300 seconds. Physical/Mac layer used is IEEE 802.11. The Mobility model used is random waypoint model with pause time equal to 20 sec and with maximum nodes movement speed equal to 3 m/s for the ACORA protocol.

The ACORA is examined the performance metrics of packet delivery ratio in nodes are in static as well as in mobility.

Packet Delivery Ratio

The number of packets originated by the MAC layer to the number of packets received by the destination is packet delivery ratio. Figure 5a shows the number of packet delivered, delay and the energy consumption of the ACORA.

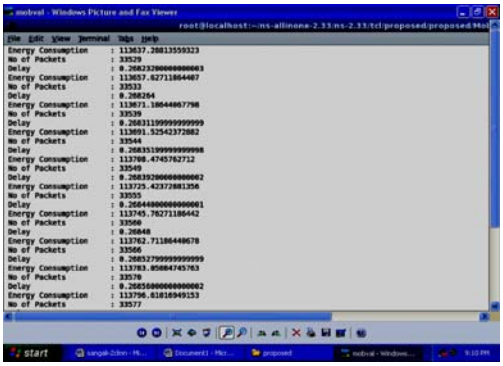


Fig.5a. Packet delivery ratio in ACORA during mobility

Packet Delivery Ratio of ACORA in Static and Mobility Nodes

Figure 6a and 6b shows the performance of the ACORA in static i.e. the packets delivered when the nodes are in static. In Figure 6a only three hops are taken, in static. In Figure 6b five hops are taken in static way for packet delivery.

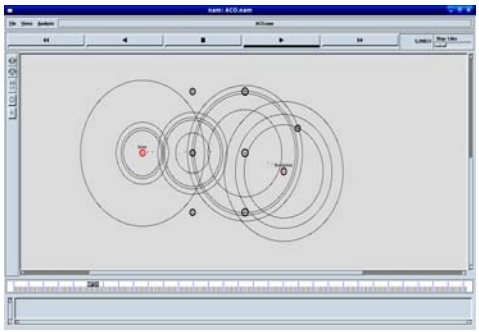


Fig.6a. ACORA for packet delivery in static (Three hops)

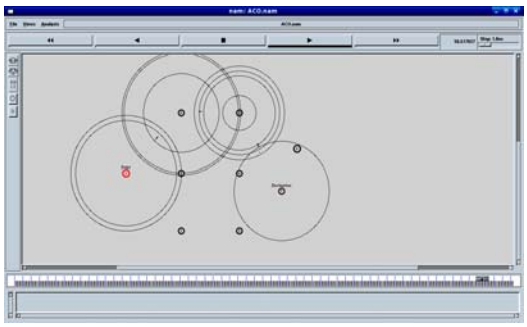


Fig.6b. ACORA for packet delivery in static (Five hops)

Also ACORA is applied for the mobility nodes. In Figure 6c the nodes are in mobility.

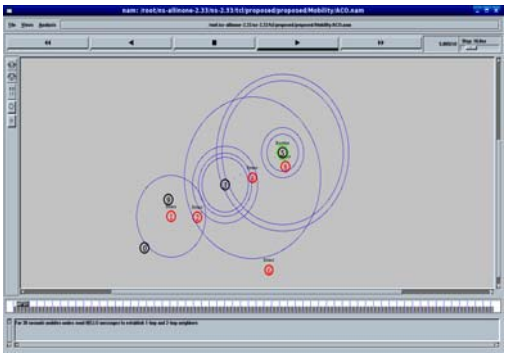


Fig.6c. ACORA for packet delivery when nodes are in mobility

The graph of the ACORA in static and mobility for the packet delivery ratio is given. The figure 6d shows the graph of the packet delivered in destination by ACORA in static.

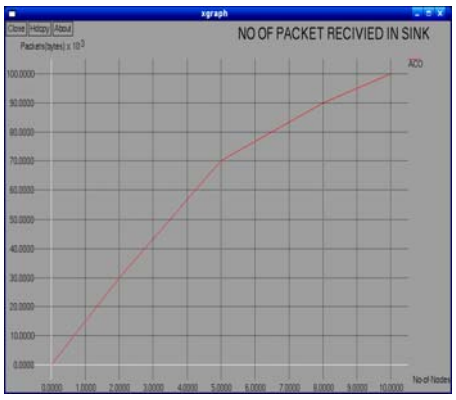


Fig.6d. ACORA for packet delivery in static

The figure 6e shows the graph of the packet delivered in destination by ACORA when nodes are in mobility. The paths vary due to mobility of the nodes even though the proposed protocol maintains the path and delivered the packets.

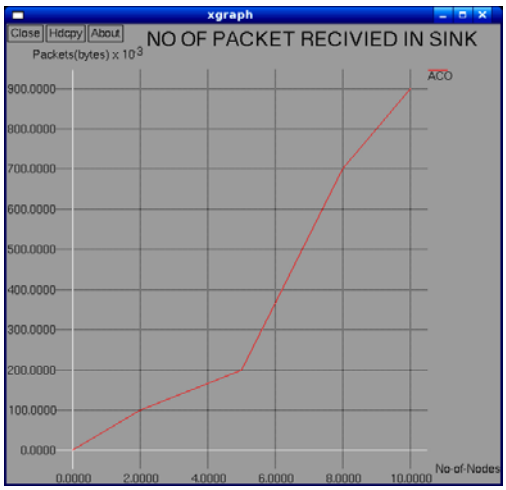


Fig.6d.ACORA for packet delivery in mobility

CONCLUSIONS AND FUTURE WORK

In this paper we have implemented the proposed ACORA protocol in C++ and integrated the module in the ns Simulator. The performance of the proposed protocol was measured with respect to metrics like Packet delivery ratio and also the packet delivery ration is examined under the static nodes and the nodes are in mobility.

The results of the simulation indicate that performance of ACORA remains stable, in both the cases, static and as well as in the mobility.

In future, the performance comparison can be made between the proposed protocol and the existing protocol for performance metrics such as end-to-end delay, routing overhead, etc. of ad hoc routing protocols with different simulation parameters.

REFERENCES

[1] Marco Dorigo and Thomas Stutzle, Ant Colony Optimization, 3rd ed. Harlow, England : Addison-Wesley, 1999.

[2] C.-K. Toh. Ad hoc mobile wireless networks: protocols and systems. Prentice Hall, 2002. ISBN: 0-13-007817-4.

[3] Frederick Ducatelle, Gianni Di Caro and Luca Maria Gambardella, "Using ANT Agents to combine reactive and proactive strategies for routing in mobile ad-hoc networks"

[4] Mesut Gunes, Udo Sorges, Imed Bouazizi, "ARA – The Ant-colony based routing algorithm for MANETs", International workshop on Ad-hoc Networking (IWAHN 2002).

[5] S. Prasad1, Y.P.Singh2, and C.S.Rai2, "Swarm Based Intelligent Routing for MANETs", International Journal of Recent Trends in Engineering, Vol 1, No. 1, May 2009.

[6] C. E. Perkins and P. Bhagvat. Highly dynamic destination sequenced distance-vector routing (DSDV) for mobile computers. Computer Communications Rev., pages 234 - 244, October 1994.

[7] A. Boukerche, Performance Evaluation of Routing Protocols for Ad Hoc Wireless Networks, Mobile Networks and Applications 9, Netherlands, 2004, pp. 333-342.

[8] http //en.wikipedia.org/ wiki/ Destination Sequenced_ Distance_Vector_routing

[9] A.E. Mahmoud, R. Khalaf & A, Kayssi, Performance Comparison of the AODV and DSDV Routing Protocols in Mobile Ad-Hoc Networks, Lebanon, 2007.

[10] The VINT Project, The ns Manual, http://www.isi.edu/nsnam/ns/nsdocumentation.html

[11] Marc Gries, Tutorial for Network Simulator-ns, http://www.isi.edu/nsnam/ns/tutorial

[12] Eitan Altman and Tania Jimenez, "NS Simulator for beginners"

[13] K. Fall and K. Varadhan. The ns Manual, Nov 2000.

ECOMMERCE IN CLIENT / SERVER TECHNOLOGY USING SNMP PROTOCOL

V.Vallinayagi
Lecturer,
Dept. of MCA,
Sri Sarada College for Women,
Tirunelveli.

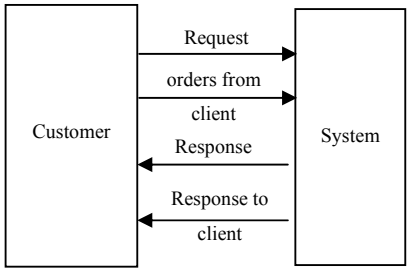
K.Sujatha
Lecturer,
Dept. of MCA,
St. Xavier's College (Autonomous) ,
Tirunelveli

Dr. S.P. Victor
Reader and Head,
Dept. of C.S,
St. Xavier's College (Autonomous) ,
Tirunelveli

Abstract - Ecommerce is perhaps the biggest change that business has feed after the industrial revolution. The revolution is of much greater magnitude and more global in nature and is beckoning business worldwide. Ecommerce business on net and it is challenging the very basics of the way inn which business is done. There are many Challenges to overcome when the business is made online. Data sharing is possible across various applications such as ERP, legacy application supplier system and ecommerce order taking system and collecting payments online. When the individuals are to conduct electronic transaction on the internet. It must be easy for the user to retrieve emails and net through mobiles and in systems using simple network management protocol which is used to store and manage the information in order to transmit and save sensing information. Electronic commerce enabled by internet era technologies for many large organization they connect ISP with the company, university and where ever it maybe. Here we use client/server technology, so the server may run number of applications. When the user want to access for ecommerce he need different applications to run both on client/server technology. It is easy for the consumer to browse an deactivate data from EDI. SNMP is standard protocol only one application is needed. Internet performance are determined by protocol processing. The basic elements of ecommerce are an e-shop on a server, a user with a web browser and internet connection between two. The large ecommerce sites cost millions to set up and sums to maintain and update. The benefits of the SNMP running on client / server technology application includes the abilities to prevent data easy to use and allow administration to control their need to maintain its network. Its popularity is to support by every enterprise network equipment into the word it can also manage any type of network devices. It is based on network management involves initializing, monitoring and modifying the operation of networks and elements connected to network using SNMP as management protocol. The data transfer are polling and trap to send messages from one device to another.

Keywords - SNMP, Ecommerce, Emarkreting, Client/ Server, Web Browser, Email.

INTRODUCTION



Ecommerce in Internet

The basics of ecommerce should be analogous to market stall or a pitch at a car boot sale. The large ecommerce sites cost millions to setup and similar sums to maintain and update.

- Visibility getting a site notices and on the line customer arrive at the site.
- Ease of user once the customer arrive at the site they were able to find what they want with the minimum of hassle.
- Order processing – online orders have to be processed logically electronic order are linked to computererised bank office systems
- After sales :- Queries and faults need to be processed online. The customer can longer simplify pop back to shop

Any form of business transaction in which the parties interact electronically rather than by physical exchange of document or direct meetings among officials. In real time ecommerce is defined as the marketing process where in one consumer can get his desired product at his door step by purchasing online.

A consumer can actually buy desired products by logging into the concerned sites. There are several advantages they are as like delivery at their door step and online purchase. Which stop you from carrying money.

There are challenges that business face while doing business online. If business organization and individuals are to conduct electronic transaction through the internet then the Intelligent Agents are there to conduct routine tasks search and retrieve information and act as domain experts.

ECOMMERCE CLIENT/SERVER TECHNOLOGY

It is a model of distributed technology where one program communicate exchanging information .client/server architecture concerns how processing

activity is distributed over the network several clients can access a single server as is the case of small lan(or) client can access data from database located on several servers.[1]

A typical client/server interaction consists of following sequences of steps

- The user runs client software to create a query
- the client connects to the server
- The client send the query to server
- The server analyze computes the result of query
- the server sends the result to the client
- The client presents the result to the user.

SNMP WITH CLIENT/SERVER TECHNOLOGY [2]

SNMP is the standard protocol only one application is needed. This application then uses SNMP to communicate with the desired device and fetch needed data Instead of having many applications running on the client and communicating with the database. One has only one applications communicating with many clients and database. The SNMP is the network management solution used by most of industry It is simple and web based to use all sorts of technologies.

Features : It should be standard for the agent to be able to represents the MIB that it is implementing

- may be an SNMP manageable device should be able to send messages representing MIB
- Affects convert from MIB to the required format.
- The queries are made as simple as possible.

Web based Client/Server with SNMP

It is used to locate each other so they can send request and response. The time needed to retrieve management data with in the system is higher than the time needed to encode message .data retrieved time increased linearly with number of retrieved objects. So we take a measure of cputime, memory and round trip delay routine time measurements For the retrieval of interface specific data from within the system we wanted to use the same code in our SNMP web service prototype. The time needed for the retrieval of data from within the system is higher than time needed to encode messages. Net SNMP does not catch previously fetched data so data retrieval time increases linearly with the number of retrieval objects. we focus our measurements on single request response interactions .In the web processing we measure the first TCP segment and last TCP segment. The time was measured in table.[3]

	Objects 1	22	66
Snmp-1	0.4	1.6	5.6
Snmp-2	0.9	1.1	4.2
Snmp-3	0.5	1.1	4.2
Snmp-4	0.5	1.7	4.8
Snmp-5	0.5	1.8	4.8
Snmp-6	0.7	1.6	5.7

(1, 22, 66 objects)

We focus our measurement on single request-response interaction. Delay time depends on number of objects retrieved and several SNMP agents would benefit from some form of caching and after long testing SNMP agent improve the performance of encoding in the web services.

We compare the web services with the SNMP agents and concluded that the web objects have time delay than SNMP agents through calculating cpu usage, roundtrip delay and memory usage.

IMPLEMENTATION

The system consists of 6 different modules interacting with each other. The modules are the following.

- client application for processing goods
- certifying authority
- merchant server
- acquiring bank server
- message provided to the user
- delivery of goods to the user

Description on Modules

Client Server

The order is placed with merchant server

- To get an token number
- To receive a catalog and order
- Place an order
- Receive confirmation

Certifying Authority

- receiving username and password
- authenticate users based on username and password
- send random token to the authenticated user
- maintain a list of users

Merchant Server

- sent catalogs and order forms to order list
- receive order from clients
- get payment
- send order confirmation

Acquiring Bank Server

- receive payment from issue bank
- send confirmation of receipt to merchant and issue bank
- bank server send queries to the consumer about the product
- payment details
- payment cheque and d no

Packet Filtering

- security is made
- set protocol is used to check the sign of the candidate. These are the steps for the ecommerce to activate the performance of their work .They communicate with the server and client and access the information from the browsers in web. In case particular product is wanted or bought by various dealers then this client/server technology is used to retrieve information at same time from the entire client to the server. In ecommerce they use the

forbidding auction for the consumer products. Immediate queries is sent to the browser by the request of the client. In many situation ecommerce is used an immediate reply and want to retrieve bulk amount of data from the web browser. So we use SNMP to retrieve objects as wells replying at same time to all the clients.

IMPLEMENTATION AND OBSERVATION

Server

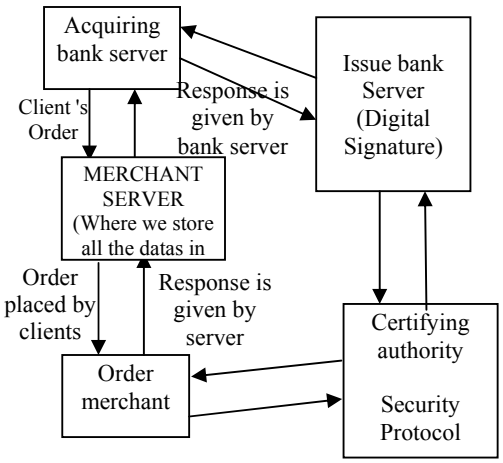
This is done by client/server technology Server side-Bank servers are accessed by the merchant server by an agent. SNMP server is simple and single process that repeatedly waits for an incoming message. It parses each incoming message and translates. To optimize the performance the SNMP server translates into ASN1.1 representation into an internal field format.

Client

By using MIB's it stores in database and access the database by the clients in all applications

The user working in this Network selects the data and corrects it to message type and converts into SNMP message. This message is sent down through UDP/IP stack to access the network .The message is sent to the agent to translate in local systems. the agent read the data and stores it for response through get response message.

IMPLEMENTATION OF CLIENT/SERVER IN ECOMMERCE



CONCLUSION

This way we activate the ecommerce in a better way with SNMP protocol in client/server technology. In future all will be accessing the SNMP network protocol in deciduous manner. All the clients are advised to activate in one application.

REFERENCES

[1] E-Commerce by David Frankly.

[2] Implementation of Internet Protocols by ManiSubramanian.

[3] Comparative study of web objects and SNMP Kugsan Jeoy.

[4] "W3C: XML Path Language (XPath) Version 1.0," <url:http://www.w3.org/TR/xpath/>. "HP - Web Services Management Framework," <url:http://devresource.hp.com/ drc/specifications/ wsmf/>.

[5] "OASIS Web Services Distributed Management TC,"<url:http://www.oasis-open.org/committees /wsdm/>.

[6] R. Neisse, R. L. Vianna, L. Z. Granville, M. J. B. Almeida, and L.M. R. Tarouco, "Implementation and Bandwidth Consumption Evaluation of SNMP to Web Services Gateways," IEEE / IFIP Network Operations & Management Symposium, Apr. 2004.

[7] "IETF: NETCONF Working Group," <url:http://www.ops. ietf.org/netconf/>.

[8] T. Goddard, "NETCONF over SOAP," Internet-Draft, Feb. 2004, <url:http://www.ietf.org /internet-drafts/draft-ietf-netconf-soap-01.txt>.

[9] Kugsang Jeong, Deokjai Choi, Soohyung Kim and Gueesang Lee, "An SNMP agent for context delivery in ubiquitous computing environments," ICCSA2005, proceeding part V, pp.203, May 2005.

JOURNAL OF COMPUTER APPLICATIONS

Information for Authors

1. All papers should be addressed to HOD - MCA*, Journal of Computer Applications,* K.S.R College Of Engineering, Tiruchengode 637 215, Namakkal District, Tamil Nadu, India. E-Mail:ksrjca@gmail.com, Mobile: 9865931266.
2. Two copies of manuscript along with soft copy in CD are to be sent.
3. A CD-ROM containing the text, figures and tables should separately be sent along with the hard copies.
4. Manuscript containing original research work, which has not been published earlier nor is under consideration for publications elsewhere may only be submitted.
5. Manuscript will be reviewed by experts in the corresponding research area, and their recommendations will be communicated to the authors.

Guidelines for submission

Manuscript Format

The manuscript should be about 8 pages in length, typed in double space with Times New Roman font, size 12,double column on A4 size paper with one inch margin on all sides and should include 100-200 words abstract,3-5 relevant key words, and a short (50-100 words) biography statement. The pages should be consecutively numbered, starting with the title page and through the text, references, tables, figure and legends. The title should be brief, specific and amenable to indexing. The article should include an abstract, introduction, body of the paper containing headings, sub-headings, illustration and conclusions.

References

A numbered list of references must be provided at the end of the paper. The list should be arranged in the order of citation in text, not in alphabetical order. Each reference number should be enclosed by square brackets.

In text, citations of references may be given simply as "[1]", Similarily, it is not necessary to mention the authors of a reference unless the mention is relevant to the text.

Example

[1] M.Demic,"Optimization of Characteristics of the Elasto-Damping Elements of Cars from the Aspect of Comfort and Handling", International Journal of Vehicle Design,13(1),1992,pp.29-46.

[2] S.A.Austin,"The Vibration Damping Effect of an Electro-Rheological Fluid", ASME Journal of Vibration and Acoustics, 115(1), 1993, pp.136-140.

SUBSCRIPTION

The annual subscription for **JOURNAL OF COMPUTER APPLICATIONS** is **Rs.1000/-** which includes postal charges. The Subscribe for **JOURNAL OF COMPUTER APPLICATIONS** a Demand Draft may be sent in favour of **The Principal**, **K.S.R College of Engineering, payable at Tiruchengode**.

SUBSCRIPTION FORM

Name : _____.

Mailing Address : _____.

_____.

City :_____ State :_____ Pin :_____.

Phone : _____ _____ E-Mail :_____.

Subscription For: _____ Years. Please find enclosed a sum of Rs._____ by DD / MO No._____.

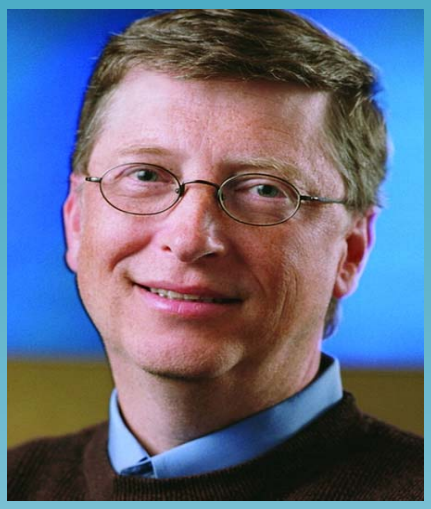
Bank _____ _____ Date _____.

In favor of "**The Principal"**, **K.S.R College of Engineering, payable at Tiruchengode**.

Please use the photocopies of the subscription form

CHARLES BABBAGE

Published by
Department of Computer Applications
K.S.R College of Engineering
Tiruchengode – 637 215, Tamilnadu
Mobile: 98659 31266
www.ksrce.ac.in, e-mail: ksrjca@gmail.com



BILL GATES