ABSTRACT BOOK

M.TECH. DISSERTATION
(JAN-MAY 2015)

Department of Computer Engineering

YMCA UNIVERSITY OF SCIENCE AND TECHNOLOGY,
FARIDABAD-121006

2015

2015

INDEX

Sr. No.	Roll No.	Name of Student	Title	Supervisor
1	MCE-221-2K12	AKSHI YADAV	TEST CASE PRIORITIZATION BASED ON MULTIPLE FACTORS SOURCE CODE ANALYSIS	Mr Harish Kumar
2	MCE-222-2K13	AMIT KUMAR	A TEST CASE PRIORITISATION TECHNIQUE FOR OBJECT ORIENTED SOFTWARE BASED ON SOURCE CODE ANALYSIS	MrVedpal
3	MCE-223-2K13	ANIL KUMAR	TEXT SUMMARIZATION BY USING GA	Mr. Sushil Kumar
4	MCE-224- 2K13	ANSHUMAN MALHOTRA	ANALYZING AND ENHANCING TEST CASE MINIMIZATION & PRIORITIZATION TECHNIQUE	DR ParulTomar
5	MCE-225-2K13	HARISH CHAND	DESIGN OF A FRAMEWORK FOR SEMANTIC DOCUMENT SIMILARITY	DR. Manjeet
6	MCE-226-2K13	PRIYANKA	HYBRID ADAPTIVE REGRESSION TESTING STRATEGY	Mr Harish Kumar
7	MCE-227-2K13	RAHUL BANSAL	ONTOLOGY BASED RANKING IN SEARCH ENGINE	DR. Jyoti
8	MCE-229-2K13	REENU RANI	A TECHNIQUE FOR CONTEXT BASED RANKING FOR EFFICIENT INFORMATION RETRIEVAL	DR. C K Nagpal
9	MCE-230-2K13	RIMPY SINGLA	A FRAMEWORK FOR CONTEXT BASED INDEXING FOR EFFICIENT INFORMATION RETRIEVAL	DR. C K Nagpal
10	MCE-233-2K13	SHWETA	A HYBRID APPROACH FOR AVOIDING FRAUD CREDIT CARDS GENERATION WITH ENHANCED LLOYD'S METHOD AND THE VALIDATION METHOD	DR. Anuradha
11	MCE-234-2K13	SNEH DALAL	A NOVEL APPROACH FOR CONSTRUCTING ONTOLOGY BASED GROUNDED INDEX AND QUERY PROCESSOR	DR. Jyoti
12	MCE-235-2K13	SONIA GARG	POWER ALLOCATION BASED ROUTING IN COGNITIVE RADIO NETWORKS USING GAME THEORY	MrsPoonamGarg
13	MCE-237-2K13	VRISHA TICKOO	CONGESTION CONTROL IN TRANSPORT LAYER OF WBAN	DR. SapnaGambhir
14	MCE-238-2K13	YATISH MAHNA	SEMANTIC DOCUMENT SIMILARITY BASED ON ONTOLOGY AUGMENTED WITH RECENT TRENDS	DR. Manjeet

15	MIT-201-2K13	AMIT KUMAR	A COST AND CODE COVERAGE BASED TEST CASE PRIORITIZATION TECHNIQUE FOR OBJECT ORIENTED SYSTEM	Mr Vedpal
16	MIT-202-2K13	ANKIT MITTAL	KANBAN BOARD IN SCRUM OF SCRUM	Dr. NareshChauhan
17	MIT-204-2K13	AYYU AGGARWAL	TEXT SUMMARIZATION USING FREQUENCY BASED METHOD	Mr. Sushil Kumar
18	MIT-205-2K13	BHARAT CHANDRA JOSHI	SCALABILITY IN AGILE METHODOLOGY	Dr. NareshChauhan
19	MIT-207-2K13	GAURAV NAG	RANK-AWARE CRAWLING OF HIDDEN WEB	Mrs. Sonali
20	MIT-209-2K13	MANISH KUMAR	A TEST CASE METHOD USING GA	Dr. NareshChauhan
21	MIT-210-2K13	MANISHA YADAV	A NOVEL APPROACH FOR ANNOTATING WEB IMAGES USING SEMANTIC DISTANCE MEASURES BETWEEN WEB CONTENT BLOCKS	Dr. PayalGulati
22	MIT-211-2K13	MOHIT BHARDWAJ	A NEW APPROACH TO DETECT AND REMOVE DEADLOCK FROM DISTRIBUTED DATABASE SYSTEM	Dr. ParulTomar
23	MIT -212-2K13	NITIN ROHILLA	SCALING FRAMEWORK FOR HIDDEN WEB CRAWLING	Mrs. Sonali
24	MIT-213-2K13	POONAM BHATIA	INDEXING THE SEMANTIC WEB FOR EFFICIENT QUESTION ANSWERING SYSTEM	Dr. Komal Kumar Bhatia
25	MIT-214-2K13	PRACHI	EXCLASSIFIER: A NOVEL TECHNIQUE FOR DETECTING EXTREMIST VIDEOS ON SOCIAL MEDIA	Dr. Anuradha
26	MIT-215-2K13	SANGEETA	COLLABORATIVE FILTERING BASED RECOMMENDER SYSTEM USING ITEM GROUPS	Dr. Neelam Duhan
27	MIT-216-2K13	SEEMA RANI	NOVEL INDEXING TECHNIQUE FOR HIDDEN WEB PAGES	Mrs. Sonali
28	MIT-217-2K13	SURUCHI SINGH TEWATIA	AN IMPROVED METHOD FOR ASYNCHRONOUS CONCURRENCY CONTROL IN DISTRIBUTED DATABASE SYSTEM	Dr. ParulTomar
29	MIT-218-2K13	VINOD KUMAR	BIDIRECTIONAL TRUST CALCULATION IN SOCIAL NETWORK	Dr. sapnagambhir
30	MIT-05-2K12	JYOTI YADAV	DESIGN AND IMPLEMENTATION OF A NOVEL RANKING TECHNIQUE FOR HIDDEN WEB PAGES	Mrs. Manvi

31	MNW-241-2K13	AMRITA MALHOTRA	LEXICON BASED OPINION MINING SYSTEM USING HADOOP	Dr. NeelamDuhaan
32	MNW-242-2K13	ANKUR MITTAL	A NOVEL APPROACH TO PAGE RANKING MECHANISM BASED ON USER INTEREST	Mrs .ShilpaSethi
33	MNW-243-2K13	ANSHU	ONTOLOGY BASED WEB INDEXING	Ms. ManviSiwach
34	MNW-244-2K14	GARIMA FUPTA	TRUST BASED TECHNIQUE FOR DETECTION AND RESOLUTION OF BLACK HOLE ATTACK IN MANET	Dr. Atul Mishra
35	MNW-245-2K13	GAUTAM KUMAR	ENERGY EFFICIENT TASK SCHEDULING IN CLOUD COMPUTING	Dr. Atul Mishra
36	MNW-246-2K13	MAMTA YADAV	AGENT BASED SOLUTION TO ENERGY SINK HOLE PROBLEM IN FLAT WIRELESS SENSOR NETWORKS	Ms. PreetiSethi
37	MNW-247-2K13	MOHAMMED WASEEM	A NOVEL APPROACH FOR CHANGE DETECTION IN WEB PAGE	Ms. Ashlesha Gupta
38	MNW-248-2K13	MOHINI GOYAL	IMPLEMENTATION OF BOOK DOMAIN ONTOLOGY TO RETRIEVE HIDDEN WEB CONTENT	Ms. ManviSiwach
39	MNW-249-2K13	MUKTA RANI DHIMAN	AGENT BASED SOLUTION FOR PRECISION AGRICULTURE APPLICATION IN WSN	Ms. PreetiSethi
40	MNW-250-2K13	NAVNEET SHARMA	A NOVEL APPROACH IN DUPLICATE DOCUMENT DETECTION	Ms. Ashlesha Gupta
41	MNW-251-2K13	RAMANDEEP	EQUIVALENCE IN MUTATION TESTING	Ms. RashmiPopli
42	MNW-253-2K13	SACHIN CHHABRA	KEY DISTRIBUTION THROUGH FINGERPRINT BASED AUTHENTICATION USING MOBILE AGENT	Mr. Umesh
43	MNW-254-2K13	SAMRIDHI MANGLA	PROFILE MATCHING IN AD-HOC SOCIAL NETWORK BASED ON USER'S INTEREST	Dr. SapnaGambhir
44	MNW-255-2K13	SANJAY KUMAR	COST BASED SCHEDULING OF TASKS IN CLOUD COMPUTING ENVIRONMENT	Dr. Atul Mishra
45	MNW-256-2K13	SATISH KUMAR	IMPROVING THE PERFORMANCE OF FOCUSED WEB CRAWLER	Mrs. Shruti Sharma
46	MNW-257-2K13	SHIKHA ATWAL	IMPLEMENTATION AND COMPARISON OF ENCRYPTION ALGORITHMS IN WIRELESS NETWORK ON NS2 PLATFORM	Mr. Umesh Kumar
47	MNW-258-2K13	TANU	INTELLIGENT PCBB PROTOCOL FOR EMERGENCY MESSAGE DISSEMINATION IN VANET	Dr.Jyoti

2015

TEST CASE PRIORITIZATION BASED ON MULTIPLE FACTORS SOURCE CODE ANALYSIS

Testing is a common and necessary task carried out by software practioners to validate the quality of evolving software system. Unfortunately regression testing is often an expensive, time consuming process, particularly when applied to large software systems. Consequently, practitioner may wish to prioritize the test case in their regression test suit to execute more important test case earlier in the regression testing process. One common goal of this test case prioritization is to increase test suit's rate of fault detection, a measure of how quickly fault are detected by test suit.

While performing the white box testing there may be large number of test cases executed by the developer to ensure the correct functionality of their code. This process involves a lot of efforts. To reduce these efforts if somehow a developer is able to get the prioritized order of the test cases which he/she is going to execute to ensure the correct functionality during the process of white box testing is a challenging task. Keeping this idea in mind a test case prioritization technique for unit testing is proposed. The proposed technique is based on the analysis of the source code written by the developers. Various factors of source code are considered and different weight is assigned to each factors. Then the weights are multiplied with the value of factors and the result is arranged in an order. From this result prioritised order of test cases is obtained. To show the effectiveness of the proposed approach it is compared with non prioritized and random approach. The APFD value obtained by proposed approach is more, showing the efficiency of the proposed approach.

AKSHI YADAV

MCE-221-2K13

2015

A TEST CASE PRIORITISATION TECHNIQUE FOR OBJECT ORIENTED SOFTWARE BASED ON SOURCE CODE ANALYSIS

Scheduling test cases by using test case prioritization technique enhances their efficiency of attaining some performance criteria. The rate at which the errors are detected within the testing process is one such criterion. An enhanced rate of fault detection during testing can provide quicker feedback on the system under test thereby allowing software engineers to rectify errors before usual time. An application of prioritization techniques includes regression testing of software which has undergone alterations. To optimize regression testing software testers may assign test case preferences so that to some extent the more significant ones are run earlier in the regression testing process.

Software testing and retesting is an inherent part of software development as far as the errors are detected at the earliest that may not change the sequence hence gaining confidence. Regression testing has been proved to be crucial stage of software testing .Regression test prioritization techniques manipulates the execution of test case so that faults are detected at the earliest. To achieve performance requirements test cases with higher priority are executed than those with lower priority by test case prioritization techniques.

This proposed test case prioritization algorithm prioritizes the test cases using some factors. These factors are class covered, method covered, assignments, variable and test case dependency. In the presented approach firstly the considered factors covered by the test case are determined. After determined the covered factors test case prioritization value is calculated (TCPV). The test cases are prioritized on the basis of the value of TCPV. For experimental verification and validation the proposed approach has been applied on a software module implemented in Java .The Result shows the efficacy of the presented approach.

Amitkumar

MCE-222-2k13

2015

TEXT SUMMARIZATION BY USING GA

Now a day's size of document is very large, it is difficult to read or understand whole document so, summary of document is needed without changing the main content of document. Text Summarization provides large text data into a shorter version without changing its information content and meaning. It is very difficult for human beings to read whole document to understand

and manually summarize the documents. It can be easy by automatic text summarization. Text Summarization methods are divided into two types: extractive and abstractive summarization. An abstractive summarization is understanding of document, finding the new concepts and providing summary in few words (sentences different form texts sentences).it is very hard or impossible to design (now a days). In Extractive summarization method select important sentences, paragraphs etc from the original text document and concatenating them into shorter form. The importance of sentences is decided based on some features of sentences.

In this thesis, Automatic Text Summarization using Extractive techniques with the help of Genetic Algorithm has been presented. With the help of GA generates features of every sentences some features are old and some are newly included. Feature score will be between zero to one. Score of every sentence is generated and best score sentences are selected for summary.

Anil Kumar
MCE-223-2K13

2015

ANALYZING AND ENHANCING TEST CASE MINIMIZATION & PRIORITIZATION TECHNIQUE

Software testing is a critical element of software quality assurance and represents the ultimate process to ensure the correctness of the product. The quality product always enhances the customer confidence in using the product thereby increases the business economics. In other words, a good quality product means zero defects, which is derived from a better quality process in testing. Testing an application was considered to be one of the phases of Software Development life Cycle (SDLC) but these days testing is performed in each and every phase of the SDLC so as to reduce the maintenance cost

Software testing is major building block of the software development life cycle. Software testing is possible through number of test cases execution on the software. Software testing helps in producing the high quality software. There may be large number of test cases available for software to be tested. To execute all test cases available in test suite may take a lot of efforts to execute all test cases, a lot cost incurred and also more time taken to execute. Thus in our work, some efforts are contributed to reduce number of test cases available in the test suite.

Test case minimization and prioritization technique is adopted and enhanced to reduce number of test cases and also prioritized them. Code coverage technique is an existing technique which is taken as a base technique of our work and some proposed factors are introduced in collaboration of code coverage technique. In our work, an algorithm and a block diagram is designed through which whole proposed work is represented.

AnshumanMalhotra

MCE 224 2K13

2015

DESIGN OF A FRAMEWORK FOR SEMANTIC DOCUMENT SIMILARITY

The World Wide Web is a very large resource centre in which information is present in the form of web pages. These web pages are basically a collection of related concepts which try to impart certain information about some particular domain. It may be possible that two web pages containing different concepts are trying to communicate similar information. So, to enhance the search engines in order to handle this situation a novel approach has been proposed which takes into account the latest trends prevalent in daily conversation. In this approach a given web document is processed to obtain its concepts and then, these concepts are randomly replaced by the captured trends in order to induce semantics thereby making the given document similar to other related documents. A semantic score indicating the level of similarity is computed for documents with the help of an ontology augmented with the latest trends giving the relations between the concepts. In this way the limitations of comparing the documents lexicographically as well as with traditional ontology concepts are removed to a very large extent. Our proposed approach has outperformed the previous techniques by showing promising results which have been analyzed for both the cases, firstly, when only trends are considered and secondly, when both trends as well as traditional ontology concepts are considered.

Harish Chand

MCE-225-2K13

2015

HYBRID ADAPTIVE REGRESSION TESTING STRATEGY

Software testing is considered an important phase for developing and maintaining a software. It controls the quality and reliability of software. There is a great misconception that testing is a debugging process whose work is to just remove the errors. But actually testing is an activity which control the quality and maintain the reliability of software by finding more and more errors from the software. As the software systems evolve, different amounts and types of code modifications can be involved in different versions. These factors can affect the costs and benefits of regression testing techniques in many ways, and thus, there may be no single regression testing technique which is the most cost-effective technique to use on every version. Till date, many regression testing techniques have been proposed, but no research has been done on the problem of helping practitioners in selecting the appropriate techniques for new versions as systems evolve. To address this problem, a new strategy is proposed which is used to identify the regression testing techniques that will be the most cost-effective for each regression testing session considering organization's situations and testing environment.

In this work two step methods is used consist of AHP method and TOPSIS method. In first step AHP method is used for calculating the weights of the attributes or criteria. In second step these weights are considered and used in TOPSIS process. Then TOPSIS method is applied for the evaluation problem and the result shows the preference order of alternatives.

To analyze and validate the proposed method a set of 2 programs have been taken. Firstly AHP(Analytical Hierarchy Process) has been applied to calculate weights of each criterion then these weights has been used as a input in TOPSIS(Technique for Order Preference by Similarity to Ideal Solution) method. Now TOPSIS presented the rank of alternatives. Now to validate this result four test case prioritization techniques has been applied on set of two programs and calculated the APFD (Average percent of fault detection) for each prioritization technique. Finally conclusion is made that technique selected by proposed approach has the highest APFD value and declared it the best technique among others.

PRIYANKA (MCE-226-2K13)

2015

ONTOLOGY BASED RANKING IN SEARCH ENGINE

The World Wide Web is a huge repository of information. It contains collection of hyperlinked documents. The information on the web is not organized, thus it is very difficult to find relevant information from this huge repository of information. The huge amount of information on the web as well as the growth of new users creates newchallenges for information retrieval. The current existingKeyword Based Search Engines (such as Google, Bing and Yahoo) offer an efficientway to browse the web content. But they do not consider the context of the user query and return a large result set, out of which very few are relevant to the user. The ordering of the result set is also important. Ranking factors employed by these search engines to rank a web page do not take into account the context or domain of the web page in which it lies.In this thesis, an Ontology based Search Engine is considered which is a context based search engine. It first determines the context of the user query and then shows the results accordingly. It does this by using an Ontological Grounded Index which is an extension of inverted index. To rank a context sensitive web page, a ranking factor is developed which considers the underlying ontology of the particular domain. Here a weight is assigned to each web page according to the number of Data Properties present in it. This weight factor can also act as an offline factorlike the PageRank for Keyword Based Search Engine to rank the web pages.

Rahul Bansal

MCE-227-2k13

2015

A TECHNIQUE FOR CONTEXT BASED RANKING FOR EFFICIENT INFORMATION RETRIEVAL

With the massive growth of the web, we have been confronted with a flood of information and therefore search engine has become the most important and helpful tool to find desired information from the web. The keyword based method has been the most popular search method in the search engine since it provides simple and user friendly interface. In general, the keyword based search find the relevance of the documents according to the query words by simply checking the occurrence of that query keywords in the document collection. It doesn't consider the context of the query. And if even searching provides relevant results, what is the guarantee that user will be satisfied by this because without ranking of results, user can be frustrated in finding relevant information.

Now-a-days only just keyword matching is not enough, consideration of context of query is also required. So that up to a limit, some irrelevant results can be reduced. And in case of single word having multiple meanings, results belonging to all possible domains increase irrelevant results, which is also a problem for SE & user. So it has become necessary to precept user interested domain & to provide results only in that domain.

In this dissertation, we propose a technique that ranks results by considering context of query, page content information & user feedback information. In beginning, SE will not have any information about user interest except query. So feedback information is useful to SE to know about what the user wants. That's why feedback information is important factor for ranking. As we know description in meta_tag information provides proper information about what the document is, can be used to handle polysemous words. Hence proposed ranking technique helps to increase precision & recall of search results.

REENU RANI

MCE-229-2K13

2015

A FRAMEWORK FOR CONTEXT BASED INDEXING FOR EFFICIENT INFORMATION RETRIEVAL

With the massive growth of the web, we have been confronted with a flood of information and therefore search engine has become the most important and helpful tool to find desired information from the web. The keyword based method has been the most popular search method in the search engine since it provide simple and user friendly interface. In general, the keyword based search find the relevance of the documents according to the query words by simply checking the occurrence of that query keywords in the document collection. It doesn't consider the context of the query. So, it can't ensure that the results returned by the search engine are relevant to the user query or not.

Like for the query '*good girl*' it returns document containing only good as well as document containing only girl which are irrelevant according to the query. Because of this reason, current search engines sometime miss highly relevant results and return some irrelevant results for user queries.

Most of the search engines do searching of documents by considering only generic synonyms of the query word to consider related words. They results documents containing every synonym of the target word. It is the wastage of efforts on the part of search engine & at the same time bombarding the user with lot of un-required information. Like when a user search for 'good girl', the search engine gives all documents containing all synonyms of the query word 'good' without considering the context of the query.

In this dissertation, we propose a contextual search framework to overcome such limitations of the traditional keyword based search and consideration of generic synonyms. Here we consider the whole context of the query and find the semantic similarity between words or their synonyms on the basis of the context of the query.

RIMPY SINGLA

MCE-230-2K13

2015

A NOVEL APPROACH FOR CONSTRUCTING ONTOLOGY BASED GROUNDED INDEX AND QUERY PROCESSOR

The World Wide Web is a large repository of information. It is a distributed information system that provides access to its various clients. The World Wide Web is growing at an exponential rate. This situation is likely to continue in the foreseeable future, as more and more information services move onto the Web. Because of the dynamic nature of the Web, it isdifficult for search engine to find the relevant documents for user queries. For this purpose, search engine maintainsthe index of downloaded documents stored in the localrepository. Whenever user fires query, search engine searchesthe index in order to find the relevant documents to bepresented to the user. The quality of the matched resultsdepend on the information stored in the index. The moreefficient the structure of index is,the more efficient will be theperformance of search engine. Keyword based search engines employ inverted index to find the associated documents according to the keywords present in the query. Generally the inverted index arebased solely on the occurrences of keywords in documents. It has the disadvantage of lower precision rate as it completely ignores the context of query. It returns those pages as well which contain keywords of the query but their context is entirely different from the fired query. In order to improve the efficiency of the searchengine, an improved indexing mechanism to index the webdocuments is being proposed that keeps the context relatedinformation in the index structure. In this proposed technique, Ontology Based Grounded Index is built that considers the context of documents and associate each term with an Ontological class by using underlying ontology of that domain. Similarly a Context Based Query Processor is proposed in which the context of the query will be analyzed to return the most relevant results to the user and thus increases the precision rate.

SnehDalal

(MCE-234-2K13)

2015

A HYBRID APPROACH FOR AVOIDING FRAUD CREDIT CARDS GENERATION WITH ENHANCED LLOYD'S METHOD AND THE VALIDATION METHOD

As the credit card has become the most prevailing mode of payment for both offline as well as online mode of transaction. However making safe and secure transaction is a real challenge. One step towards this safety is to generate credit cards with a special validation checking method. This may reduce the number of frauds to some extend as Fraudster many a time try to generate credit card numbers by just altering the digits. Luhn algorithm is one way. The Luhn check needs to be applied on each and every card separately making it infeasible for the large dataset. Thus for a large dataset, clustering can be done to effectively implement such systems. The Lloyd's algorithm is one of the most popular clustering algorithm used in data mining for real world applications, especially for the numerical dataset. K-mean performance and efficiency is greatly affected by the initial cluster centers, as different initial centers may leads to different cluster formations. In this thesis, Lloyd's is used in application with a modification which has an additional procedure for selecting better cluster centroids. If the initial centroids are properly selected then this results in better and efficient clustering. This is done by keeping intercluster similarity as low while intracluster similarity as high. This research work proposed a hybrid approach which combines both the techniques i.e. modified centroid problems in classical Lloyd's method along with the validation method.The proposed idea can be used in many other applications for clustering purpose which uses numeric values such as detecting whether the consecutives amount of transactions seems to be fraud. Analysis of large data in many field and many other applications over web.

SHWETA

MCE-233-2K13

2015

POWER ALLOCATION BASED ROUTING IN COGNITIVE RADIO NETWORKS USING GAME THEORY

Due to continuously increasing demand of wireless communication technology, the problem of spectrum shortage arises. Cognitive radio technology enables efficient utilization of the existing wireless spectrum resources. CRN utilizes spectrum band by allowing secondary users to use the spectrum opportunistically, by changing their transmission parameters. We need to analyze the performance of a wireless network. To analyze the performance of a wireless network, game theory is used due to its ability to model individual, independent decision makers. In this thesis we model how game theory can be used to allocate power levels to different CR nodes in a CRN at physical layer to ensure a certain QoS.

Cognitive Radio (CR) is a novel concept for improving the utilization of the radiospectrum. This promises the efficient use of scarce radio resources. In this thesis titled **"Power Allocation Based Routing In Cognitive Radio Networks Using Game Theory"**we aim to find out a routing algorithm which requires minimum transmission energy required to send packets in the network, while maximizing the overall throughput of the system. The thesis covers a brief introduction to cognitive radio networks, game theory, power allocation and how the game theory can be used to address many issues in cognitive radio networks. Then an algorithm for power efficient routing is designed. The algorithm attempts to maximize the total throughput of the Cognitive Radio system(secondary users) subject to the total power constraint of the CR system and tolerableinterference from and to the licensed band (primary users) with the help of Game Theory. In this thesis we designed a transmission power efficient algorithm using game theory and then compared its performance with various already existing routing algorithms. And in the last of thesis it is concluded that, the designed algorithm works better in a number of scenarios.

SONIA GARG

MCE-235-2K13

2015

CONGESTION CONTROL IN TRANSPORT LAYER OF WBAN

In 1996, T.G. Zimmerman came up with the idea of a new emerging field of research known as WBAN (**W**ireless **B**ody **A**rea **N**etwork). WBAN is a network in which tiny, intelligent patches known as sensors are placed on or inside the body of a living being. These sensors are connected to a base station known as sink which further communicates with a remote monitoring server having a database. It is standardized as IEEE 802.15.6. WBAN finds its applications in healthcare, sports, defence and entertainment. It has various issues like quality of service, reliable data transmission, flow control, congestion control, latency, scheduling, traffic classification, etc. One of the issues of transport layer of WBAN is congestion control. Congestion occurs when the quality of service deteriorates due to over burdened data load on node or link. As the problem of congestion and its consequences can cost life of a living being under observation, so it's very important to keep a check on the occurrence of congestion.

Since WBAN is in its puerility, no protocol for congestion control of transport layer has been designed as such. In this thesis, a protocol is proposed to detect, notify and avoid congestion in the network. The protocol considers the priority of each node which is provided at the time of network setup. Packets coming from the highest priority node is acknowledged and processed first by the sink than the packets arriving from lower priority nodes. In addition to that, the highest priority node sends most packets into the network than low priority node at a specific interval of time. Based on queue occupancy with packet loss, congestion is detected and a notification of the occurrence of congestion is send to all the sources. Four congestion levels (0, 1, 2, 3) are used in congestion avoidance. According to the each congestion level, sending rate of packets is increased or decreased thus, minifying the effects of congestion in the network. The algorithms of the proposed protocol are implemented in MATLAB. The simulated results show that the throughput of the proposed protocol is better than conventional TCP congestion control protocol. Moreover, the simulation results show a direct relationship between the number of nodes in the network and the throughput of the network.

VRISHA TICKOO

MCE-237-2K13

2015

SEMANTIC DOCUMENT SIMILARITY BASED ON ONTOLOGY AUGMENTED WITH RECENT TRENDS

The World Wide Web is a very large resource centre in which information is present in the form of web pages. These web pages are basically a collection of related concepts which try to impart certain information about some particular domain. It may be possible that two web pages containing different concepts are trying to communicate similar information. So, to enhance the search engines in order to handle this situation a novel approach has been proposed which takes into account the latest trends prevalent in daily conversation. In this approach a given web document is processed to obtain its concepts and then, these concepts are randomly replaced by the captured trends in order to induce semantics thereby making the given document similar to other related documents. A semantic score indicating the level of similarity is computed for documents with the help of an ontology augmented with the latest trends giving the relations between the concepts. In this way the limitations of comparing the documents lexicographically as well as with traditional ontology concepts are removed to a very large extent. Our proposed approach has outperformed the previous techniques by showing promising results which have been analyzed for both the cases, firstly, when only trends are considered and secondly, when both trends as well as traditional ontology concepts are considered.

YatishMahna

MCE-238-2K13

2015

A COST AND CODE COVERAGE BASED TEST CASE PRIORITIZATION TECHNIQUE FOR OBJECT ORIENTED SYSTEM

Test Cases are treated as one of the most important part of software testing activity. They are responsible for the validation of the software under inspection. Test suites are also used to test changes in the source code during regression testing. In number of cases, the test suites are so big that executing all tests for every source code change is absurd. Developers/testers need to prioritize the test suite so that most beneficial test cases are executed first. This can result in increasing the effectiveness of testing and saving a lot of time and cost.

Testing and maintenance of Object-Oriented (OO) software is expensive and difficult. Many researchers have found several approaches to schedule an order of test execution. Unfortunately, existing test prioritization techniques are failed to prioritize multiple test suites and test cases with same priority values. Consequently, those techniques are inefficient to prioritize tests and the cost is overrun during the prioritization process.

Test case prioritization by cost and coverage is a powerful technique for assessing relationships among software components .in a test suite , There are different factor which can affect the total cost and the coverage of a test case. The factor which can affect the cost are operator, variable, native, predicate statement, assignment. The coverage of the test case is depend on the no. of line statement is covered and in this the priority is also assigned to the factors which can affect the cost of the test case i.e. priority is given to the operators, variables, external system calls etc. Again there are different priority is assigned to the different operators .i.e. multiplication operator will be assigns a higher priority than the addition operator.

The theoretical model and the measurement set serve as a foundation for exercising complexity analysis on various problems that are related to the interaction among components. The purpose work describes how we can prioritize. In the proposed work higher quality software.

AMIT KUMAR

MIT-201-2K13

2015

KANBAN BOARD IN SCRUM OF SCRUM

In the last few years Agile Methodologies appeared as a reaction to traditional software development methodologies. Agile has become most prominent type of software development method. The main reason behind this is its fast result and its acceptable nature, as it accepts changes in the project at any level of work.

Kanban is another framework used to implement agile. Back in the 1940s, Toyota optimized its engineering process by modelling it after how supermarkets stock shelves. Supermarkets stock just enough product to meet consumer demand, a practice that optimizes the flow between the supermarket and the consumer. Because inventory levels match with consumption patterns, the supermarket gains significant efficiency in inventory management and optimizing for the customer. When Toyota brought that idea to it's factory floors, teams would deliver a card, or "Kanban", to each other (say, to the team that assembles the doors) to signal that they have excess capacity and are ready to pull more materials.

This research work starts with dividing the project in subprojects on some basis. The criteria for dividing the project are client preferences, interdependency of modules, project domain and ease of coding, technical ability. Then Scrum of Scrum is to be applied on the project. As the project is divided into various parts, i.e. subprojects, scrum process will be applied to each subproject. The Scrum framework consists of scrum teams and their associated roles, events, artifacts, and rules. After that Scrumban technology is applied in which various Kanbans are to be designed for the project and they are managed by the teams. They are called self-managed Kanbans. Visualize the flow of work into and out of the sprint. Once the visualizing of workflow is achieved, it helps team including product owners to know the bottleneck areas within the workflow. Further, visualizing helps in knowing who is working on what tasks, and what is progress state of different stories at a particular point of time. After that in proposed work, an architecture of a big Kanban for the whole project is introduced. . The idea is to make the remaining product backlogs in pending stage, and the selected userstories from the product backlog in development phase or deployment phase according to their state in the reference of small kanban at that time. This work will allow the companies to better screen job candidates and assess their internal talent for skills development.

ANKIT MITTAL

MIT-202-2K13

2015

TEXT SUMMARIZATION USING FREQUENCY BASED METHOD

With the rapid growth of the web, users get easily lost in the rich hyper structure and providing relevant information to the users to cater to their needs is the primary goal of website owners. Therefore, finding the content of the web and retrieving the user's interests and needs from their behavior have become increasingly important.

As the result of rapid advancements in Information Technology, Information Retrieval on Internet (Internet-Searching) is gaining importance, day-by-day. Search Engines are admittedly essential tools for this purpose. But, like a two side of the same coin, search engines performance degrades due to some critical issues like redundancy of data and irrelevant data.

Our basic focus is on Text Summarization which converts a large text data into a shorter version retaining the important content and its meaning using extractive methods. In Extractive method, the key sentences are extracted from the text and put together to form a summary. Many important applications have used text summarization in order to reduce redundant and non-relevant information presented to users of the document retrieval systems.

This thesis proposes a new approach for text summarization i.e. frequency method along with the Point-wise Mutual Information (PMI) to calculate the semantic similarity between the words on the document level. Frequency and PMI value is calculated for every sentence. Then, the sentences with the total value above the threshold value are selected to be in the summary. Further, an algorithm has been proposed.

AYYU AGGARWAL

MIT-204-2K13

2015

SCALABILITY IN AGILE METHODOLOGY

Over the last decades more and more software development companies transfer at least a part of their development process to the so-called offshore countries. To increase the productivity of these projects “the new agile process for remote location projects” is given in this research. The research question of this work is how “projects are made while sitting in a remote location by following the agile methodology” and some more questions that are considered as the limitation of agile will be resolve by this methodology. This research uses a case study methodology to identify waste and gaps in the process. Studies of other processes, conducted mainly by literature review, are used to fulfil the identified gaps. As a result of the research several improvements were proposed and questions for the future research were identified.

Agile software development in its current meaning appeared in mid-1990s as an alternative to the standard of the time heavily regulated and strictly document based software development processes. That time in the developer’s community emerged the theory that traditional processes require huge amount of paperwork, these papers and artifacts slow down the development and are sometimes even useless. Another very important issue for that time was inability of the customer to define the final set of the requirements and features for the software from the start of the project which caused changes during the development process. A traditional heavily regulated and regimented development model for example the waterfall model was not able to handle this issue properly. Due to that, the focus in the development process was switched from fulfilling clearly stated project requirements to continuous delivery up to date value to the customer.

BHARAT CHANDRA JOSHI

MIT-205-2K13

2015

RANK-AWARE CRAWLING OF HIDDEN WEB

The Hidden web alludes to the collection of Web information which is high quality is hidden behind the search forms. As the number of Internet users and the number of accessible web pages grows, it is becoming difficult for users to find the documents that are relevant to their particular needs. Traditional crawler finds information from keyword while there is Search form provided on the hidden web. These search forms have various fields. To get the relevant information these search forms fills automatically.

In this Dissertation, a novel technique is proposed for crawling a Hidden Web site by taking the ranking of the results into account. So there is a technique to calculate that how to approximate the site's ranking function so that we can compute the top results based on the data collected so far. The technique used spearman coefficient for calculating the ranking score for the archive of URL.

GAURAV NAG

(MIT-207-2K13)

A TEST CASE METHOD USING GA

Web application has been accepted in a wide range of software application field, and many of our daily activities depend on the services provided by them. For this, web applications quality has a great influence on our daily lives. Therefore effective testing of web application is important to provide trusty services for the fast growing demand. A web application is often collection of relatively small software components, created with so many technologies. So web applications are complex and changeable, that why effective testing of web application is very important. For the development of internet technology, the web applications become very complex with strict requirements of quality needs.

The main area of discussion in this thesis is to generate automatic test case for the web application using the concept of genetic algorithm. This thesis introduces the concept of grey box testing which combine the white box testing and black box testing. First user session is collected from the log and then a graph is constructed which is called Request Dependence Graph (RDG). This thesis introducing a concept of covering important nodes in the test case generation which leads to reduction of test cases and give high efficiency in the testing of the web application. Importance of the node is calculated on the basis of weight of the nodes in the form of In_degree ,Out_degree and entropy. Once weight of each node is calculated then weight matrix and Adjacency Matrix is constructed. On the basis of the matrix important node is find out and that node is covered in the test cases. To generate test case automatically, genetic algorithm approach is applied. Theoretical discussion and experimental results are shown. Feasibility of actual existence of the path which is covered in the test cases is finding out.

MANISH KUMAR

MIT-209-2K13

2015

A NOVEL APPROACH FOR ANNOTATING WEB IMAGES USING SEMANTIC DISTANCE MEASURES BETWEEN WEB CONTENT BLOCKS

In recent years, the digital images are increasing rapidly due to advancement of storage technology, facility of sharing of images on the web through the news websites and social networking sites like facebook, instagram, google+, etc and as well as availability of economical digital cameras. Therefore, images over www are exponentially increasing, leading to difficulty in searching relevant web images. To retrieve relevant images, image retrieval system came into account. As, it is difficult to manually tag or annotate huge amount of images present on web, creates the need of automatic image annotation. It provides several keywords automatically for an image to describe its content which is useful in image retrieval. Various researchers are working on text based and content based image annotations [1, 2].

Text based annotation approach uses the contextual information such as Page title of web page, image name, alternative text, associate text etc. to annotate them. Associate text which is semantically close to the image provides better description of the image. But in existing approaches, semantic correlation among web content blocks is also not considered during the extraction of contextual information.

This research work provides a novel approach for annotating web images using web page segmentation and semantic distance measures among web content blocks. In the proposed work, Web page segmentation segments the web page into web content blocks from which contextual information relevant with the image is extracted. Semantic distance measures are used to check semantic closeness among the image blocks and text blocks present on the web page. Apart from the text extracted from semantically close text blocks, frequent keywords are also extracted from contextual information such as Page Title, metadata, alternative text and ImgSrc to annotate the image.

MANISHA YADAV
MIT-210-2K13

2015

A NEW APPROACH TO DETECT AND REMOVE DEADLOCK FROM DISTRIBUTED DATABASE SYSTEM

Distributed database system provides resource sharing environment for optimal performance of various database activities, especially when data is spread over a large number of sites. Distributed nature of transactions occurring at different sites and requiring resources from diverse sites pose various operational problems, such as deadlocks, concurrency and data recovery. A deadlock is a state where some processes request for some resources but those resources are held by some other processes. Occurrence of deadlock in a system will lead to resource wastage and breakdown of the system. A deadlock may occur when a transaction enters into wait state which request resource from other blocked transactions. The deadlocks are handled in three phases namely deadlock detection, deadlock avoidance and deadlock detection.

Detailed discussion about deadlock and distributed database is provided in this thesis. Advantages and disadvantages of distributed database along with the distributed transaction model is explained in the first chapter of this thesis work.

Many algorithms have been provided by many scientists and many model for deadlock detection and removal have been proposed. A detailed review of some of these algorithms has been discussed in the thesis. The distributed nature of database demand full proof control structure for its proper and effective functioning. Therefore the allocation of the resources should be properly controlled otherwise it may lead to several anomalies such as concurrency of transaction, synchronizing of events and deadlocks. This thesis provides a new approach for detection and removal of deadlock in distributed databases along with the review of various algorithms for deadlock detection and removal in distributed database systems.

MOHIT BHARDWAJ
MIT-211-2K13

2015

SCALING FRAMEWORK FOR HIDDEN WEB CRAWLING

The Web, the largest database of the world has greatly improved access to the documents. As the number of Internet users and the number of accessible web pages grows, it is becoming difficult for users to find documents that are relevant to their particular needs. Users must either browse through a hierarchy of concepts to find the information they need or submit a query to a Search Engine and continually searching through hundreds of results most of them irrelevant .Web Crawlers are one of the most crucial components used by the Search Engines to collect pages from the Web. Traditional crawler finds information from keyword while hidden web crawler automatically fills forms, search and finds relevant information.

In this Dissertation, a novel technique is proposed which scale the web crawler by parallelising it. This web crawler works for extracting the hidden data from various domains websites running parallely at multiple machines. It will balance the load distribution, enhance the performance and provide reliability.

NITIN ROHILLA
MIT -212-2K13

2015

INDEXING THE SEMANTIC WEB FOR EFFICIENT QUESTION ANSWERING SYSTEM

Search engine are the program that search document for specific keywords and return a list of documents where the keywords were found. Search engine search the query in the form of links of pages but the user wants the exact answer instead of links containing the answer. Search engines cannot differentiate between the variable documents and spams.Some search engine crawler retrieve only document title not the entire text in the document. Question answering system aims to develop techniques that can go beyond the retrieval of relevant documents in order to return exact answers to natural language questions. Answering natural language questions requires more complex processing of text than employed by current information retrieval systems. A number of question answering systems have been developed which are capable of carrying out the processing required to achieve high levels of accuracy. However, there is little progress on techniques for quickly finding exact answers. Existing Question Answering system are unable to handle variety of questions and reasoning based question. Sometimes the data sources for answering a question is not present then also the QA system fails to answer the query. This thesis investigates a novel technique for "indexing the Semantic web for efficient question answering system". Investigated techniques include: manual constructed question classifier based on <Subject, Predicate, Object>, document retrieval specifically for question answering, semantic type answer extraction, answer extraction via manually constructed index for every category of Question.

POONAM BHATIA

MIT-213-2k13

EXCLASSIFIER: A NOVEL TECHNIQUE FOR DETECTING EXTREMIST VIDEOS ON SOCIAL MEDIA

With the growing popularity of social media, social network (you tube) remains the largest as well as the most popular video sharing site. However, terrorists groups have made YouTube as a focal point for targeting innocent and vulnerable people. They propagate their ideologies to mainstream audience who otherwise would not visit their website. Hence, there is a need to detect such videos to prevent online radicalization among the users. The thesis proposes a metadata and audio based classification method for detecting such videos which promote hate and violence. It does this by extracting the user generated metadata such as title and description which the uploader of the video adds. It also finds the patterns to classify an audio into violence class such as gunshots and screams. The proposed methodology promises a recall rate of 84%.

PRACHI

MIT-214-2K13

2015

COLLABORATIVE FILTERING BASED RECOMMENDER SYSTEM USING ITEM GROUPS

Everything in this world keeps on changing because change is the rule of the world. In recent years of exponential growth of information present on the internet, it has become a universal repository of human knowledge and culture, thus enabling an exchange of ideas and information on a global level. The incredible success of internet is based on its usage, simplicity, efficiency and enormous market potential. To find relevant information, search engines came into account. But search engines can not help if a user does not know what exactly he/she wants to search. There comes recommender systems. Recommender systems are tools for interacting with large and complex information spaces. They provide a personalized view of such spaces, prioritizing items likely to be of interest to the user. The field has grown enormously in a the variety of problems addressed and techniques employed, as well as in its practical applications. Recommender systems research has incorporated a wide variety of artificial intelligence techniques including machine learning, data mining, user modeling etc. Personalized recommendations are an important part of many on-line e-commerce applications such as Amazon.com, Netflix, YouTube and Facebook. Most widely used techniques in recommender systems are: collaborative filtering and content-based filtering. These two techniques possesses their own limitations such as cold start problem, sparsity, scalability and accuracy which can be solved by hybrid approach for recommender systems.

This research work will provide a system which give personal recommendation to the target user. Movie domain is considered to explain the proposed technique. In this work, movie groups are formed based on their genre information which are represented in the form of group rating matrix to solves the sparsity problem. Groups of movies also help to overcome the cold start problem. Mean absolute error metric is used to evaluate the accuracy of the system. It has been observes by the experimental results that the predictions provided by the system are accurate.

SANGEETA

MIT-215-2K13

2015

NOVEL INDEXING TECHNIQUE FOR HIDDEN WEB PAGES

Hidden Web is a collection of web pages which cannot be accessed by a generic crawler by simply following the link structure. The information stored in hidden web is very important. This high quality content of hidden web is hidden behind the search forms. Search forms are provided on the hidden web. These search forms have various fields. To get the relevant information these search forms need to be filled. To retrieve this information, a hidden web crawler fills these search forms automatically with the help of its local repository. After submission of the search forms, information is retrieved and stored. The web pages fetched by hidden web crawler need to be indexed for fast searching process. Searching process is solely based on the Indexing technique. Efficient indexing technique reduces searching time for a query. Efficient indexing reduces the query processing time up to large extent. Many indexing techniques exist in literature to index the content retrieved by a crawler from surface web. But these indexing techniques are not good for hidden web. So, the hidden web contents still need to be indexed efficiently. In this thesis, we are proposing an indexing technique which indexes the hidden web pages more efficiently using semantic analysis. This technique uses attributes value pair and keywords to index the information retrieved by hidden web crawler. The ultimate goal of this indexing technique is to reduce query processing time and give more specific result pages to the user's query.

SEEMA RANI

MIT-216-2K13

2015

AN IMPROVED METHOD FOR ASYNCHRONOUS CONCURRENCY CONTROL IN DISTRIBUTED DATABASE SYSTEM

A distributed database is a set of several parts that correlate with each other logically over a network of interconnected computers. Collections of data in a database can be distributed across multiple physical locations connected via local area network. Since the database is distributed, different users can access it without interfering with each other. However, the DBMS must periodically synchronize the scattered databases to make sure that they all have consistent data.

Concurrency control is the activity of coordinating concurrent accesses to a database in a multiuser database management system (DBMS). Concurrency control permits users to access a database in a multiprogrammed fashion while preserving the illusion that each user is executing alone on a dedicated system. The main technical difficulty in attaining this goal is to prevent database updates performed by one user from interfering with database retrievals and updates performed by another. The concurrency control problem is exacerbated in a distributed DBMS (DDBMS) because users may access data stored in many different computers in a distributed system, and a concurrency control mechanism at one computer cannot instantaneously know about interactions at other computers.

The proposed algorithm in this thesis solves the problems of consistency and dirty read on the database system. To make a database consistent at any point of time, central node concept is introduced in the majority voting protocol. The central node is the node where all the data items have the most current value in them. All the transactions are forwarded to the central node which executes the request and forwards the update request to all other nodes. All the nodes then start their voting process. Also the database remains consistent as multiple voting is used. Also all the READ data request are forwarded to the central node, which holds the most current value of the data item. This solves the problem of dirty read in the database system.

SURUCHI SINGH TEWATIA

MIT-217-2K13

2015

BIDIRECTIONAL TRUST CALCULATION IN SOCIAL NETWORK

With the evolution of Internet, web 2.0 provides us many services, social network is one of them. It allows the user to connect with each other and share the various information they have. Popular social networks are: Facebook, Google+, LinkedIn, Instagram, Reddit, Pinterest, Vine, Tumblr, and Twitter are widely used worldwide. With the advancements in social network, it also surge with many issues like identity theft, privacy, trolling, access to information, data mining and so many others. Trust is a major issue in the social network. Trust has high impact in the social network. In real life, trust is a factor on the basis of which a person determines whether someone is close to them or not. Similarly, to distinguish between their social network friends, trust is calculated. It is nothing but a security measure to control some of the issues of OSN. Trust can be calculated on the basis of actions performed in OSN. These actions can be messages, posts, comments, shares, etc., that can be possible in the network.

Calculation of trust has been an integral part of the social networking scene for a long time now, and lot of work has been done in this area. The existing methods resolved many of the problem related to the trust calculation in OSN but also they lack in some area. They calculate trust on the user end but not consider at the friend's end as trust in real life is always bidirectional not unidirectional. There might be possibility that this trust factor is less from the other side but for the case of zero trust on either side, the relation is also at non-existence state. Also the methods don't provide us any specific value for those actions performed in OSN. So in order to resolve these problems a trust calculation system is proposed which calculates trust with different weightage to those actions also calculates trust bidirectional i.e. consider the trust value level at the friend's end of the user. The proposed method categories friends on OSN into three different list named as *unprivileged* list, *privileged* list and the *most_trusted* list as per the level of trust of friend and the security window set by the user in the form of two *upper threshold* and *lower threshold*. These categories decide what can friend access or what not i.e. giving access of user's data as per their trust formed with the user. The proposed method has successfully implemented and working has shown that the proposed method calculates trust well in an OSN, also improving the deficiency that already existing in some.

VINOD KUMAR

MIT-218-2K13

2015

DESIGN AND IMPLEMENTATION OF A NOVEL RANKING TECHNIQUE FOR HIDDEN WEB PAGES

Web is a wide term which mainly consists of surface web and hidden web. The hidden or deep web refers to content that is hidden behind HTML forms. This contains a large collection of data which is unreachable by link-based search engines. A study conducted at University of California, Berkeley estimated that the deep web consists of around 91,000 terabytes of data, whereas the surface web is only about 167 terabytes. Several researchers have explored various methods for crawling deep web content. To access this content one must submit valid input values to the HTML forms. There are many methods for extraction of data from hidden web. Various crawlers to access deep web content has been already explored by many researchers. These hidden web crawlers return huge result set for the user query. But users commonly look at top ten or twenty results that can be seen without scrolling. Users rarely look at results coming after first response page so ranking of the results is needed. So to maintain quality of results returned by the crawler, efficient ranking technique is required. Here efficiency means in terms of relevant, correct results and minimum data retrieval.

Till now ranking of hidden web pages is a big challenge, enough work has not been done in this area. In this thesis, a novel technique for the ranking of hidden web pages is designed and implemented. Various algorithms used in process of implementation are also developed. Results of the proposed technique are compared with the general techniques such as techniques used by Google etc.

JYOTI YADAV

MIT-05-2K12

2015

LEXICON BASED OPINION MINING SYSTEM USING HADOOP

When purchasing a product for the first time one usually needs to choose among several products with similar characteristics. Companies use to promote their brands and products pointing out good characteristics avoiding to mention the poor ones. The best way to choose the most suitable product is to rely upon the opinions of others.

Sentiment analysis or opinion mining is the computational study of people's opinions, appraisals, attitudes, and emotions toward entities, individuals, issues, events, topics and their attributes [1]. This is an Information Extraction task which is technically very challenging but also practically very useful. With the advent of web 2.0, huge volumes of opinionated text are available on web. To extract sentiment about an object from this huge web, automated opinion mining systems are thus needed. By devising an accurate method to identify the sentiments behind any text; one can predict the mood of the people regarding a particular product or service.

However, with so much social media available on the web sentiment analysis is now considered as a big data task. Hence the conventional sentiment analysis approaches fails to efficiently handle the vast amount of sentiment data available now a days.

This thesis, successfully addresses the significant challenges and has resolved many of them.

The main focus of the thesis was to find such a technique for analysing sentiments posted on web that can efficiently perform sentiment analysis on big data sets. A technique that can categorize the text as positive, negative and neutral in a fast and accurate manner.

In this thesis, sentiment analysis is performed on a large data set of tweets using Hadoop and the performance of the technique is measured in form of the speed and accuracy. The experimental results shows that this technique exhibits very good efficiency in handling big sentiment data sets.

AMRITA MALHOTRA

MNW-241-2K13

2015

A NOVEL APPROACH TO PAGE RANKING MECHANISM BASED ON USER INTEREST

WWW is a vast resource of hyperlinked and heterogeneous information including text, audio, video and metadata. It is estimated that www is doubling in size every six to ten month. Due to the rapid growth of information resources on WWW, it is difficult to manage the information on the web. A search tool on the web helps user to search the document in this large repository. Search engine database contains the bulk of web pages so they must be able to distinguish high quality pages form low quality pages. Hence ranking of the web page is essential for fulfil the user needs. Present page ranking algorithm considers either only the link structure or keywords enter in the query for ranking purpose. Ranking algorithm that considers link structure will return popular but less relevant pages. Algorithm considering content of web page only consider the keyword entered in the query for rank purpose they do not consider user browsing history and user interest, which will return the results more relevance to user needs. In this dissertation work a new ranking approach considering user interest is being proposed.

ANKUR MITTAL

MNW-242-2K13

2015

ONTOLOGY BASED WEB INDEXING

The main aim of search engines is to provide most relevant documents to the users in minimum possible time. So granting efficient and fast access to the index is a major issue for performance of Web Search Engines. Indexing is performed on the web pages after they have been gathered into a repository by the crawler. The existing architecture of search engine shows that the index is built on the basis of the terms of the document and consists of an array of the posting lists where each posting list is associated with a term and contains the term as well as the identifiers of the documents containing the term. The current information retrieval systems use terms to describe documents and search engines consider term frequency as a factor to determine its importance in a page and does not take into account the context in which term appears in document and how much a term represent a page, which can be done by analysing its relation with other terms by using concept of ontology. Firstly, a specific domain is chosen that we want to index. It can be any website containing information about automobiles, ticket reservation, books etc. The pages of website we want to index are downloaded by web crawler. Parser read each document line by line and tokenize the document. Tokens (Keywords) are passed to Label and frequency based ranker. Both frequency and label based ranker produce ranked lists of keywords. These lists are given to Average Rank calculator which assigns ultimate rank to keywords. Stemming and Lemmatization is performed on ranked list. Context of every Keyword present in ranked list is calculated using WordNet. On the other side, Ontology of specific domain that we have considered for indexing is downloaded from ontology database. Class Extractor extracts all the classes from downloaded ontology. Context Extractor extracts context of class. Context of class and Context of Keyword are passed to context matcher which assigns mapping score to each pair (keyword, Ontology class). The keyword is mapped to ontology class for which it is having highest mapping score.

As keyword is mapped to Ontology class so it can be easily determined what is its relationship with other keywords. Based on number and type of relationship a word have with other words it can be determined how much a keyword represent a web page. So, Ontology based index is created which is having four fields that is Keyword, Context, Ontology class, Doc_id. B+ tree data structure is used for storing the index. Its time complexity is $log_b n$ where b and n are maximum number of child each node can have and number of terms in index respectively. So, Ontology based Indexing system creates an index that attaches context and ontology class with keyword and hence produces relevant results with increased speed

ANSHU

MNW-243-2K13

2015

TRUST BASED TECHNIQUE FOR DETECTION AND RESOLUTION OF BLACK HOLE ATTACK IN MANET

An Ad hoc Network is a pool of wireless mobile nodes energetically forming a network without the use of any pre-accessible network infrastructure or centralized administrator. These ad hoc networks are infrastructure less, decentralized and are used where very fast deployment is needed. These nodes communicate with each other by sending data and other control packets between the source and destination, even though the sender and receiver are not in range of each other. Thus the communication here enabled is hop-to-hop communication or multi hop communication. Having dynamic topology, the mobile ad-hoc networks (MANETs) allows nodes to get attached and leave the network at any second of time.

Thus MANET can be used in a variety of fields. Current MANETs are designed primary for military utility, emergency services and other rescue services. This generic characteristic of MANET has rendered its vulnerability to various security attacks. Due to which unprotected attacks of the malicious nodes can occur at any time.

Recently, researchers started to study malicious nodes and their effects on mobile ad-hoc networks. That resulted in creating a new thread of research in the MANET field. A number of research papers discussing different cooperation enforcement schemes for detecting and defending against malicious nodes and their disturbance to mobile ad-hoc networks are published. Still none of these proposed cooperation enforcement schemes provide better performance against malicious nodes attacks and all of them make certain premise which makes them more suitable for some scenarios and less suitable for others.

This thesis thus, here, focuses on one such attack known as “Black hole attack” which is a form of Denial of Service attack (DoS). This is an active MANET attack which is one of the crucial types of attack being recognized in military environment. This thesis describes the presence and consequences of black hole attack in AODV one of the most commonly used routing protocol in MANET. The thesis also presents a brief overview of various pre-existing solutions of cooperative black hole detection and proposes a novel approach to detect and remove the black hole node from the mobile ad hoc network by using trust based method.

GARIMA
GUPTA

MNW-244-2K13

ENERGY EFFICIENT TASK SCHEDULING IN CLOUD COMPUTING

Cloud Computing offers service to end-users rather than a product, by sharing resources, software and other information under a usage based payment model. It enables hosting of various kinds of applications such as business, scientific, social network, etc. as it has key characteristics like multi-tenancy, scalability, performance, security, etc. Economic benefits are the main driver for the Cloud, since it promises the reduction of capital expenditure (CapEx) and operational expenditure (OpEx).

Cloud Computing is facing many challenges like Data Security, Energy Consumption, Server Consolidation, Virtual Machine Migration, etc. Existing approaches of VMs management have investigated the consolidation of VMs using the similar type of workload in cloud environment to reduce the power. This research work focuses on the study of management of virtual machines in a cloud environment. Management of virtual machines helps in reducing the energy consumption of data centers, achieving the efficient resource usage and reduction in operational costs which benefits the end-users from decreased prices for resource usage.

Proposing the new solution that can be much more reliable, feasible and energy efficient than the others existing ones and to lessen the energy consumption. In this approach the main focus is on utilization of the cpu frequencies in the form of MIPS and try to utilize the cpu frequency efficiently upto a certain limit. In this the proposed work is try to reduce the active host due to which the supply of the energy is lessen and we applied a frequency and voltage control techniques that provides them dynamically i.e DVFS. DVFS plays a major role in controlling the power supply at an instance of time. Finally, develop a fully utilizationapproach of host. The real contribution of the thesis is discussed as follows:

1) **Increase utilization**: The increasing the utilization of the cpu frequencies of host as much as possible. It may leads to reduction in power up of another host.
2) **Minimum the power**: There is a need to reduce the power at the time of running of all the hosts; it is well managed by the DVFS.
3) **Minimum theenergy**: The power consumption at a certain time should be lessening to reduce the carbon footprints.

In this thesis, various techniques to reduce the energy have been compared. A Virtual machine management Cloud environment has been designed, developed and presented for the implementation of VMs management technique for green cloud. This technique allows the

dynamic request of cpu frequencies and resources to the VM allocation at run-time according to the current utilization of resources and thus minimizing energy consumption. A simulated environment, CloudSim Toolkit has been used to validate the experiment. The experiment demonstrates that the proposed approach can effectively handle power consumption, in the heterogeneous infrastructure and heterogeneous VMs.

GAUTAM
KUMAR

MNW-245-2K13

2015

AGENT BASED SOLUTION TO ENERGY SINK HOLE PROBLEM IN FLAT WIRELESS SENSOR NETWORKS

Wireless sensor networks consist of many small nodes with sensing, computation and communication capabilities. All the sensors are low power devices and once deployed they remain unattended. One of the major problems observed in flat static wireless sensor networks is the faster depletion of the energy of the sensor nodes nearer to the sink. This problem occurs because the data transmission to the sink is in the form of multiple hops and the nodes nearer to the sink carry more load .This problem is known as ***energy sink holeproblem***. The resulting disadvantage due to sink hole is early failure of network even when there is surplus amount of residual energy left in it. Though the research community has provided various solutions to deal with this problem, the use of mobile agents to meet the above stated problem is still in its infancy. As the name suggests mobile agents are a type of software agents with mobility. They autonomously travel from computer to computer performing the desired task. In thesis, a mobile agent based solution has been proposed for solving energy sink hole problem. The proposed solution extends the network life by reducing redundant data passed to the nodes near to the sink thus reducing the load and saving battery life. The algorithm is easy to implement and the analytical results show that this algorithm can significantly improve the network lifetime. The proposed solution is realized using a tool “Aglets. This tool helps in developing of mobile agents and provides a single uniform paradigm for distributed object computing which includes providing synchrony and asynchrony, message-passing and object-passing, and stationary objects and mobile objects all together.

MAMTA YADAV

MNW-246-2K13

2015

A NOVEL APPROACH FOR CHANGE DETECTION IN WEB PAGE

Search engines are useful tool to retrieve information from Internet. Although, Internetusers speak different languages most of resources are written and published in the English. All Internet search engines provide a lingual search, although some of them enable the searcher to select the language of the results, Internet users need a multilingual search engine, which can give them results in English language and in their own language.

Web crawlers are the heart of search engines. Web crawlers continuously keep on crawling the web and find any new web pages that have been added to the web, pages that have been removed from the web. A novel architecture for incremental parallel web crawler has been designed that helps to reduce the overlap and network bandwidth problem among crawlers while working in parallel. The hallmark of the proposed mechanisms is that changes occurring at micro level i.e. paragraph levels are also identified with a capability to identify three major types of changes namely: content level changes, structural changes and presentation changes.

In this thesis, we have proposed an approach to detect changes in web pages for future updating or replacement of web pages. Our approach overcomes the drawback of previous proposed methods. It is more efficient in terms of complexity. Our proposed work will provide a wide scope for future work.

MOHAMMED WASEEM

MNW 247 2K13

IMPLEMENTATION OF BOOK DOMAIN ONTOLOGY TO RETRIEVE HIDDEN WEB CONTENT

A lot of research is going on to devise new methods for indexing and accessing hidden web available in the form of databases. To extract this hidden web data, user has to fill various search forms available on WWW with appropriate values .To find these appropriate values ontology may be one of the method. With help of ontology, a semantic database can be constructed which will provide correct values to be filled in various fields of search form interfaces. In order to create such a database, an ontology based data extraction approach is required dealing with large amount of info available on the world wide web in form of HTML, OWL and RDF/XML files. Till now work has been done in direction of filling forms but by using relational database which fills the search form on basis of attribute label-column matching scheme which matches labels of form interface with columns of database to find search results. These are limited to simple keyword searches for retrieving data and making it difficult to obtain good results with high precision. Therefore ,for finding appropriate value which required to be filled in form interface, a

semantic database is required to be created from such an information extraction component that can extract the document structure and other relevant meta data from semantic web available in form of RDF/XML,OWL files and store these files in form of Ontology graph as RDF triples. To access deep web several issues need to be addressed .First, a method is needed to Create semantic ontological database. Second, interaction of query interface with server database and how to formulate valid queries for database .Third, a method is needed to automatically generate queries for database to maximize response with small no of queries. Fourth, a method is needed to automatically extract the search result records embedded in response pages in response of queries.

Search interfaces lies in surface web which allows the user to enter the query either by typing or selecting options and search the items of his domain of interest.Atwo step approach is needed to find it :

• First, it needs a crawler to crawl web pages of surface web and download source code for that pages.

• Second it need to check for each downloaded page whether it contains a <form>….</form> tag. used to create HTML forms which requires human interaction to accept input for form fields.

2.) Query interface analysis: It includes extraction of web database connection information. A major work in query interface analysis is to understand the meaning of each input field so that appropriate type of query values can be used for the input field. Final is to determine valid queries.A query is called a valid query for a search interface if it is acceptable to the query interface regardless of whether any results are retrieved by the query.

3.)Automatic query generation: It is not practical to form all possible queries for query interface because lot of combinations are possible.This effort can be reduced by relationship link between two attribute values with help of ontology.

In this work focus is on creation of ontology database which provide us relevant information about book domain avoiding irrelevant results as was generated previously from traditional database.

MOHINI GOYAL
MNW-248-2K13

AGENT BASED SOLUTION FOR PRECISION AGRICULTURE APPLICATION IN WSN

Wireless sensor network is an adhoc collection of large number of spatially distributed, integrated, low power, low cost nodes with sensing capability. In wireless sensor network deployment of nodes is random and on large scale. Although use of WSNs is widespread, there are many challenging unsolved problems in WSN research. Most existing WSNs use a traditional server-client based model. However, the model exhibits unsatisfying performance in data-gathering when a network is composed of many sensor nodes. In WSNs, sensor nodes work as clients and each of them transfers data to a server working as a processing center.This gives birth to massive sensory data which is redundant in nature. If the transferred data-size is quite large, there will be a lot of network traffic in the WSN. This causes some sensor nodes to delay data transmission so that total performance of the WSN falters. Routing of such kind of unnecessary data not only saturates network resources, but also consumes immense nodes energy.

To solve the above-mentioned problems in WSNs, an alternative idea is to introduce Mobile Agent (MA) instead of using the server/client based model in WSNs. Mobile agents are used to reduce the communication cost, especially over low bandwidth links, by moving the processing function to the data rather than bringing the data to a central processor (sink).

The main performance metrics include execution time and energy consumption, and the performance is better if both of these can be minimized. To satisfy both metrics, we need to carefully consider not only algorithms but also well-balanced parameters such as the number of MAs, size of MAs, the number of sensor nodes.

Precision farming is the ability to handle variations in productivity within a field and maximize financial return, reduce waste and minimize impact of the environment using automated data collection, documentation and utilization of such information for strategic farm management decisions through sensing and communication technology.

Using WSN in PF applications will revolutionize the data collection in agricultural field, support the sought highlyautomated agriculture system which requires intensive sensing of environmental conditions at the ground level and rapid communication of the raw data to a local or remote server where the availability of computational and storage power, the

identification of pests in the crops, drought or increased moisture, the decision making, and the control of farm equipment is done in real time (automated actuation devices like sprinklers, foggers, valve-controlled irrigation system, etc, can be used to control irrigation, fertilization and pest control in order to offset the adverse conditions forming.

A few of the researchers have worked on the WSN in precision agriculture and the work done is commendable but still there are some challenges left unresolved. So in this thesis various challenges involved in WSN in precision agriculture are addressed and the possible resolution to these problems is suggested. Implementation of data aggregation is also performed by using mobile agent.

MUKTA RANI DHIMAN
MNW-249-2K13

2015

A NOVEL APPROACH IN DUPLICATE DOCUMENT DETECTION

The development of the World Wide Web in the recent times has made the concept of Web Crawling receive remarkable significance. Users of World Wide Web utilize search engines for information retrieval in web as search engines play a vital role in finding information on the web. But the voluminous amount of web documents has weakened the performance and reliability of web search engines. As, the subsistence of near-duplicate data is an issue that accompanies the growing need to incorporate heterogeneous data. These pages either increase the index storage space or increase the serving costs thereby irritating the users. Near-duplicate detection has been recognized as an important one in the field of plagiarism detection and spam detection. The detection of duplicate and near duplicate web documents has long been recognized in web crawling research community. It is an important requirement for search engines to provide users with the relevant results for their queries in the first page without duplicate and redundant results. Such near-duplicates can be detected and eliminated using the novel approach for duplicate document detection. Detection of duplicate web documents carried out by crawling the web documents from the repositionary. At first, keywords are extracted from the crawled documents. Then obtaining the clustering and fingerprint of the documents. The documents having duplicity value greater than the threshold value are considered as duplicate documents depend on the author name and format of the documents. The detection of the duplicate documents has resulted in reduced memory for repositionary and improved search engine quality.

NAVNEET SHARMA
MNW-250-2K13

2015

EQUIVALENCE IN MUTATION TESTING

Mutation testing has been used as an essential method to evaluate the quality of test suites. In mutation testing artificial defects are seeded into the program. If a test suite fails to detect a mutation, it may also fail to detect real defects and so it is needed to be improved. However, in some cases, there are mutations which keep the program semantics unchanged and thus cannot be detected by any test suite. Such mutants are called as equivalent mutants. Equivalents mutants can never be killed, as their behavior is identical to original program. Equivalent mutants add no value to the process as the behavior of these mutants is similar to the original implementation and therefore cannot be killed by the test cases. The equivalent mutant is one of the crucial problems in mutation testing, as test case fails to detect them. In this thesis work, an algorithm is proposed to resolve the problem of equivalent mutants. JUnit framework is used in creating these mutants. JUnit is a unit testing framework for the Java programming language. JUnit has been important in the development of test-driven development, and is one of the family of unit testing frameworks. JUnit is linked as a JAR at compile-time; the framework resides under package junit.framework. According to proposed algorithm, the generation of mutant depends on three factors: input, output values and operators. If same output generates after applying different input on same operator mutants, then these mutants will be avoided. Similarly if after using same input values on different operators, generates the same output, in that case the mutant is again avoided.

RAMANDEEP

MNW-251-2K13

2015

KEY DISTRIBUTION THROUGH FINGERPRINT BASED AUTHENTICATION USING MOBILE AGENT

In a communication network, security of information is a major prerequisite which can be achieved with cryptography. Key management is of vital concern in any cryptographic method. Symmetric key algorithms use the same cryptographic keys for both encryption of plaintext and decryption of cipher text. The key can be distributed between the two parties by face-to-face or by any previous secured key or with the help of any third party. The first one is vulnerable and not safe while the second one depends on the security of previous key exchanged hence it is advocated to use a third party known as KDC. The KDC operates mostly on the symmetric key algorithm andintended to reduce the risks inherent in exchanging keys. In the KDC, the secret key is provided in encrypted form by using a unique key to both the sender and receiver.

The literature reviews key management mechanisms proposed so far and brings up their strengths and shortcomings. This thesis additionally displays an approach uses fingerprint biometric of communication parties as unique key to distribute the session key fingerprint based key of user. This thesis presents a systematic and organized study of key exchange using mobile agents. The strengths and weaknesses of each technique is also a matter of focus in this thesis. The literature survey shows that most of the algorithms fail to secure the key management without being compromised by passive or active attack.

The proposed framework uses mobile agent approach for key exchange because the agent can distribute the work for key generation and distribution between the nodes required of communication hence increasing its efficiency as in requirement of large network in which several nodes communicating at a same time. Additionally the approach protects the privacy of biometric data even when KDC is compromised. The implementation and results shows the proposed work is efficient and improves the overall security of a communication network.

SACHIN CHHABRA

MNW-253-2K13

PROFILE MATCHING IN AD-HOC SOCIAL NETWORK BASED ON USER'S INTEREST

An Ad hoc Network is a pool of wireless mobile nodes energetically forming a network without the use of any pre-accessible network infrastructure or centralized administrator. These ad hoc networks are infrastructure less, decentralized and are used where very fast deployment is needed. These nodes communicate with each other by sending data and other control packets between the source and destination, even though the sender and receiver are not in range of each other. Thus the communication here enabled is hop-to-hop communication or multi hop communication. Therefore, each node doesn't only plays the role of an end system, but also acts as a router, that sends packets to desired nodes. Having dynamic topology, the mobile ad-hoc networks (MANETs) allows nodes to get attached and leave the network at any second of time.

Thus MANET can be used in a variety of fields. Current MANETs are designed primary for military utility, emergency services and other rescue services. Since MANETs can be used for communication between two nodes, a social network can be built.

Recently, researchers started to study about the formation of social network with Ad-hoc communication modes ex. Bluetooth, WI-FI, or cellular radio. These modes of ad-hoc communication can be utilized to form a social network, where users having similar interests can connect with each other. A number of research papers discussing, ASN are published for profile matching between two users, or forming a mobile application with this idea. Still none of the research provides better profile matching considering similar interest of user.

This dissertation proposes a new approach for profile matching in ad-hoc social network wherein users can connect with each other if they have similar interest and are in proximity with each other. Here two things are considered, viz, similar interest of user that is found by considering semantic matching of interests between users and Proximity of user that is found by fetching their GPS locations. Keeping this idea, an algorithm is proposed.

User maintains two profile i.e. global profile and Local profile. Global profile of user is similar to online social networking and local profile of user is based on present interest and

location of user. Use of Local profile, do not have any effect on social media, rather it completes them. Meeting known and un-known people is an important motivation for such network.

The parameter on which the proposed algorithm i.e. Semantic based profile matching algorithm in ad-hoc social network is compared with other research work is number of friend suggestions received. This number varies for global and local profile of user. With Local profile it is more, resulting in a good and meaningful social network because the requests are only related to other users' interest and Location. For matching two profiles a threshold value is taken, which is 0.75. If any two users profile match more or equal to this threshold value, then they are considered as users with similar profiles, and friend request can be send. We concluded that using simple global profile, user cannot match their profile with other user having similar interest, but via local profiles this can be achieved.

SAMRIDHI MANGLA

MNW-254-2K13

COST BASED SCHEDULING OF TASKS IN CLOUD COMPUTING ENVIRONMENT

In modern era of technology, there are tremendous increases in demand of online hosted services and application led to development of the Cloud computing. Cloud computing technology involves the high cost infrastructure and highly scalable pool of computational resources which offer service to end-users by sharing those resources, software and other information under a usage based payment model. The provisioning of these resources to the end user should be done in most efficient way so that cloud is utilized more effectively and efficiently.

Cloud computing has many critical issues and challenges like resources provisioning, data security, sever consolidation; service provisioning, energy management and virtual machine migration etc. Resource provisioning consist both resource allocation and task scheduling. Task scheduling tasks are scheduled on to the allocated resources according to their requirement and characteristics. The scheduler should adapt its scheduling strategy according efficient execution of appropriate scheduling criteria tasks. This research work focuses on the study of task scheduling for resource provisioning in a cloud environment. Task scheduling helps in manage the various QoS (quality of service) parameters to achieving the efficient resource usage and reduction in operational costs which benefits the end-users from decreased prices for resource usage.

Now a days, cloud computing has emerged as a promising platform for executing scientific workflows. Under this new paradigm, the computing and storage resources are now delivered as a kind of utility charged in a pay-as-you-go manner to the general public. Composition of Scientific workflows in cloud computing environment can include tasks with highly variable computational granularity depending upon varying user's requirements. Scientific workflows may include fine tasks with less varying computational overheads to large tasks with highly variable computational overheads or combination of both. The actual execution time of these individual tasks may not be too long but the overheads involved prior to their actual scheduling on Virtual Machines (VMs) may be of too long duration as compared to their runtime duration. Clustering of tasks with maximum dependency into a single unit is a technique to efficiently utilize the cloud resources such as bandwidth.

In this thesis work a novel technique is being proposed for managing the cost based task scheduling. Task clustering merges the dependent tasks and allocates same host to the

tasks belonging to same clusters, thereby reducing the data transfer delay between dependent tasks over the network. Prioritization of the clusters on the basis of both cost of resource and computation performance reduces the computational cost of the cluster. The prioritization of the cluster helps in to allocate the beneficial host for the cloud provider. Existing clustering techniques provide strategies that rely on intercommunication between tasks. Proposed technique lay emphasis on locality of reference based dependencies and examine ways to efficiently provision cloud resources based on dependencies and cost between tasks and availability of resources. This resource provisioning techniques in cloud computing environment minimize the bandwidth usage among dependent tasks in workflows and also minimize the cost of resource allocation. To reduce the makspan of the task the conventional Max-Min algorithm is used The simulation of this technique on cloud based tool CloudSim shows that clustering technique reduces the bandwidth usage and cost; max-min show that efficiently utilizes all cloud resources as compared to technique that independently schedules the tasks on cloud resources.

SANJAY KUMAR

MNW-255-2K13

2015

IMPROVING THE PERFORMANCE OF FOCUSED WEB CRAWLER

In this thesis, we introduced an implementation of focused crawler based on multi agents and the works of the agent is to provide the freshness of data and these agents provide the fresh data time to time and by these agents it cover the problem of freshness of the data and new arrival of data at any web page. It overcomes that problem which I mention earlier in this thesis. The goal of focused crawler is fetching as more as possible web pages that the relevant to a user specified topics. The topics can be represented by a set of keywords (we call them seed keywords) or example urls. The key for designing a focus crawler is how to judge whether a web page is relevant to the topic or not.

To achieve such a crawler, we use seed keywords to find out seed urls from search engine (e.g. Google), and use seed urls and examples urls to fetch seed pages. Then we use TF-IDF algorithm to find out some more but limited number of keywords (we call them topic keywords) from seed web pages to represent the topic. We compute the vector similarity between web pages and topic keywords to see whether the page is relevant to topic or not.

SATISH KUMAR

MNW-256-2K13

2015

IMPLEMENTATION AND COMPARISON OF ENCRYPTION ALGORITHMS IN WIRELESS NETWORK ON NS2 PLATFORM

In this thesis, we design, and analyze the security of several encryption schemes. Our main goal is to provide schemes for which we can provide theoretical proofs of security, but which are also efficient and practical.

We begin by studying different encryption algorithms such as DES, AES, RSA, 3DES, DSS, etc. Then we have to choose some encryption algorithms so that we can get the best out of them and use them further in wireless protocols for improving security. So, finally we choose three encryption algorithms which are best at their level which are AES (Advanced Encryption Standard) [4], RSA (Rivest-Shamir-Adleman) [3] and DSS (Digital Signature Standard) [5, 6]. The main goal of this thesis is not only to choose best out of these three encryption algorithms, but also to give theoretical proof for the above said statement. AES has 10 rounds if key size is 128 bits, 12 rounds if key size is 192 bits and 14 rounds if key size is 256 bits. DSS uses digital signature for data transfer while encrypting and decrypting data. Our proofs are in the graphs; no random assumption is required. Now-a-days, we are using WPA2 (Wi-Fi Protected Access Version 2) [1, 2] for providing security in wireless network. WPA2 is equivalent to IEEE 802.11i standard [1].

Next, we study the graphs of these encryption schemes which have parameters viz. delay, throughput, pdr (packet delay ratio), and plr (packet loss ratio) and received packet count. These are schemes in which have their different ways to encrypt messages. In this thesis, we individually plot graphs of the encryption algorithms chosen and see the results which show significant changes.

We also study them together and compare them. Then, we plot graphs of all encryption algorithms based on one parameter. We get fifteen graphs if we plot graphs individually and five graphs if we take all encryption algorithms at the same time with one parameter at a time.

SHIKHA
ATWAL

MNW-257-2K13

2015

INTELLIGENT PCBB PROTOCOL FOR EMERGENCY MESSAGE DISSEMINATION IN VANET

Vehicular ad hoc network (VANET) is a new type of Mobile Ad hoc Network (MANET).It supports safety systems designed to avoid road accidents. It supports safety systems designed to avoid road accidents in two ways: (1) periodic transmissions (beacon) from all vehicles that inform neighbors about their current status and (2) dissemination of emergency messages to warn other vehicles to avoid the danger.

The successful dissemination of warning messages beyond the transmission range of vehicle faces four major issues: (1) The broadcast storm (2) The hidden nodes (3) The reliability problem (4) the severe interferencewith the existing periodic single-hop safety messages.

Broadcast transmission is used in VANET to alert all vehicles within a geographical area about an emergency situation. Typically, selection of next relaying hop is the major problem in VANET broadcasting. To get the smallest preoperational delay, the number of relaying hop must be minimize. Meanwhile, transmission reliability must also be preserved. Both of these two constrains must be taken into consideration. However, these two aspect often collide to each other since increasing one of them always result in decreasing of other In this thesis work, for fast and effective dissemination of emergency messages within a geographical area an intelligent Particle Swarm Optimization Contention Based Broadcast (PCBB) is proposed that significantly reduces the number of required transmissions and provide timely and successful delivery of warning messages.

TANU

MNW-258-2K13