SEABIRDS

IEEE 2012 – 2013
SOFTWARE PROJECTS IN VARIOUS DOMAINS

| JAVA | J2ME | J2EE |
DOTNET |MATLAB |NS2 |

www.ieeeproject.in

SBGC

24/83, O Block, MMDA COLONY

ARUMBAKKAM

CHENNAI-600106

SBGC

4th FLOOR SURYA COMPLEX,

SINGARATHOPE BUS STOP,

OLD MADURAI ROAD. TRICHY- 620002

Web: www.ieeeproject.in

E-Mail: ieeeproject@hotmail.com

Trichy

Mobile:- 09003012150

Phone:- 0431-4012303

Chennai

Mobile:- 09944361169

SBGC Provides IEEE 2012-2013 projects for all Final Year Students. We do assist the students with Technical Guidance for two categories.

Category 1 : Students with new project ideas / New or Old IEEE Papers.

Category 2 : Students selecting from our project list.

When you register for a project we ensure that the project is implemented to your fullest satisfaction and you have a thorough understanding of every aspect of the project.

SBGC PROVIDES YOU THE LATEST IEEE 2012 PROJECTS / IEEE 2013 PROJECTS

FOR FOLLOWING DEPARTMENT STUDENTS

B.E, B.TECH, M.TECH, M.E, DIPLOMA, MS, BSC, MSC, BCA, MCA, MBA, BBA, PHD, B.E (ECE, EEE, E&I, ICE, MECH, PROD, CSE, IT, THERMAL, AUTOMOBILE, MECATRONICS, ROBOTICS) B.TECH(ECE, MECATRONICS, E&I, EEE, MECH , CSE, IT, ROBOTICS) M.TECH(EMBEDDED SYSTEMS, COMMUNICATION SYSTEMS, POWER ELECTRONICS, COMPUTER SCIENCE, SOFTWARE ENGINEERING, APPLIED ELECTRONICS, VLSI Design) M.E(EMBEDDED SYSTEMS, COMMUNICATION SYSTEMS, POWER ELECTRONICS, COMPUTER SCIENCE, SOFTWARE ENGINEERING, APPLIED ELECTRONICS, VLSI Design) DIPLOMA (CE, EEE, E&I, ICE, MECH,PROD, CSE, IT)

MBA(HR, FINANCE, MANAGEMENT, HOTEL MANAGEMENT, SYSTEM MANAGEMENT, PROJECT MANAGEMENT, HOSPITAL MANAGEMENT, SCHOOL MANAGEMENT, MARKETING MANAGEMENT, SAFETY MANAGEMENT)

We also have training and project, R & D division to serve the students and make them job oriented professionals

www.ieeeproject.in

PROJECT SUPPORTS AND DELIVERABLES

- Project Abstract
- IEEE PAPER
- IEEE Reference Papers, Materials & Books in CD
- PPT / Review Material
- Project Report (All Diagrams & Screen shots)
- Working Procedures
- Algorithm Explanations
- Project Installation in Laptops
- Project Certificate

www.ieeeproject.in

TECHNOLOGY	: JAVA
DOMAIN	: IEEE TRANSACTIONS ON DATA MINING

S.No.	IEEE TITLE	ABSTRACT	IEEE YEAR
1.	A Framework for Personal Mobile Commerce Pattern Mining and Prediction	Due to a wide range of potential applications, research on mobile commerce has received a lot of interests from both of the industry and academia. Among them, one of the active topic areas is the mining and prediction of users' mobile commerce behaviors such as their movements and purchase transactions. In this paper, we propose a novel framework, called Mobile Commerce Explorer (MCE), for mining and prediction of mobile users' movements and purchase transactions under the context of mobile commerce. The MCE framework consists of three major components: 1) Similarity Inference Model ðSIMÞ for measuring the similarities among stores and items, which are two basic mobile commerce entities considered in this paper; 2) Personal Mobile Commerce Pattern Mine (PMCP-Mine) algorithm for efficient discovery of mobile users' Personal Mobile Commerce Patterns (PMCPs); and 3) Mobile Commerce Behavior Predictor ðMCBPÞ for prediction of possible mobile user behaviors. To our best knowledge, this is the first work that facilitates mining and prediction of mobile users' commerce behaviors in order to recommend stores and items previously unknown to a user. We perform an extensive experimental evaluation by simulation and show that our proposals produce excellent results.	2012
2.	Efficient Extended Boolean Retrieval	Extended Boolean retrieval (EBR) models were proposed nearly three decades ago, but have had little practical impact, despite their significant advantages compared to either ranked keyword or pure Boolean retrieval. In particular, EBR models produce meaningful rankings; their query model allows the representation of complex concepts in an and-or format; and they are scrutable, in that the score assigned to a document depends solely on the content of that document, unaffected by any collection statistics or other external factors. These characteristics make EBR models attractive in domains typified by medical and legal searching, where the emphasis is on iterative development of reproducible complex queries of dozens or even hundreds of terms. However, EBR is much more	2012

www.ieeeproject.in

		computationally expensive than the alternatives. We consider the implementation of the p-norm approach to EBR, and demonstrate that ideas used in the max-score and wand exact optimization techniques for ranked keyword retrieval can be adapted to allow selective bypass of documents via a low-cost screening process for this and similar retrieval models. We also propose term independent bounds that are able to further reduce the number of score calculations for short, simple queries under the extended Boolean retrieval model. Together, these methods yield an overall saving from 50 to 80 percent of the evaluation cost on test queries drawn from biomedical search.	
3.	Improving Aggregate Recommendation Diversity Using Ranking-Based Techniques	Recommender systems are becoming increasingly important to individual users and businesses for providing personalized recommendations. However, while the majority of algorithms proposed in recommender systems literature have focused on improving recommendation accuracy (as exemplified by the recent Netflix Prize competition), other important aspects of recommendation quality, such as the diversity of recommendations, have often been overlooked. In this paper, we introduce and explore a number of item ranking techniques that can generate substantially more diverse recommendations across all users while maintaining comparable levels of recommendation accuracy. Comprehensive empirical evaluation consistently shows the diversity gains of the proposed techniques using several real-world rating data sets and different rating prediction algorithms.	2012
4.	Effective Pattern Discovery for Text Mining	Many data mining techniques have been proposed for mining useful patterns in text documents. However, how to effectively use and update discovered patterns is still an open research issue, especially in the domain of text mining. Since most existing text mining methods adopted term-based approaches, they all suffer from the problems of polysemy and synonymy. Over the years, people have often held the hypothesis that pattern (or phrase)-based approaches should perform better than the term-based ones, but many experiments do not support this hypothesis. This paper presents an innovative and effective pattern discovery technique which includes the	2012

www.ieeeproject.in

		processes of pattern deploying and pattern evolving, to improve the effectiveness of using and updating discovered patterns for finding relevant and interesting information. Substantial experiments on RCV1 data collection and TREC topics demonstrate that the proposed solution achieves encouraging performance.	
5.	Incremental Information Extraction Using Relational Databases	Information extraction systems are traditionally implemented as a pipeline of special-purpose processing modules targeting the extraction of a particular kind of information. A major drawback of such an approach is that whenever a new extraction goal emerges or a module is improved, extraction has to be reapplied from scratch to the entire text corpus even though only a small part of the corpus might be affected. In this paper, we describe a novel approach for information extraction in which extraction needs are expressed in the form of database queries, which are evaluated and optimized by database systems. Using database queries for information extraction enables generic extraction and minimizes reprocessing of data by performing incremental extraction to identify which part of the data is affected by the change of components or goals. Furthermore, our approach provides automated query generation components so that casual users do not have to learn the query language in order to perform extraction. To demonstrate the feasibility of our incremental extraction approach, we performed experiments to highlight two important aspects of an information extraction system: efficiency and quality of extraction results. Our experiments show that in the event of deployment of a new module, our incremental extraction approach reduces the processing time by 89.64 percent as compared to a traditional pipeline approach. By applying our methods to a corpus of 17 million biomedical abstracts, our experiments show that the query performance is efficient for real-time applications. Our experiments also revealed that our approach achieves high quality extraction results.	2012
6.	A Framework for Learning Comprehensible Theories in XML Document Classification	XML has become the universal data format for a wide variety of information systems. The large number of XML documents existing on the web and in other information storage systems makes classification an important task. As a typical type of semi structured data, XML documents have both structures and contents. Traditional text learning techniques are not very suitable for XML document classification as structures are not	2012

www.ieeeproject.in

		considered. This paper presents a novel complete framework for XML document classification. We first present a knowledge representation method for XML documents which is based on a typed higher order logic formalism. With this representation method, an XML document is represented as a higher order logic term where both its contents and structures are captured. We then present a decision-tree learning algorithm driven by precision/recall breakeven point (PRDT) for the XML classification problem which can produce comprehensible theories. Finally, a semi-supervised learning algorithm is given which is based on the PRDT algorithm and the cotraining framework. Experimental results demonstrate that our framework is able to achieve good performance in both supervised and semi-supervised learning with the bonus of producing comprehensible learning theories.	
7.	A Link-Based Cluster Ensemble Approach for Categorical Data Clustering	Although attempts have been made to solve the problem of clustering categorical data via cluster ensembles, with the results being competitive to conventional algorithms, it is observed that these techniques unfortunately generate a final data partition based on incomplete information. The underlying ensemble-information matrix presents only cluster-data point relations, with many entries being left unknown. The paper presents an analysis that suggests this problem degrades the quality of the clustering result, and it presents a new link-based approach, which improves the conventional matrix by discovering unknown entries through similarity between clusters in an ensemble. In particular, an efficient link-based algorithm is proposed for the underlying similarity assessment. Afterward, to obtain the final clustering result, a graph partitioning technique is applied to a weighted bipartite graph that is formulated from the refined matrix. Experimental results on multiple real data sets suggest that the proposed link-based method almost always outperforms both conventional clustering algorithms for categorical data and well-known cluster ensemble techniques.	2012
8.	Evaluating Path Queries over Frequently Updated Route Collections	The recent advances in the infrastructure of Geographic Information Systems (GIS), and the proliferation of GPS technology, have resulted in the abundance of geodata in the form of sequences of points of interest (POIs), waypoints, etc. We refer to sets of such sequences as route collections. In this work, we consider path queries	2012

www.ieeeproject.in

		on frequently updated route collections: given a route collection and two points ns and nt, a path query returns a path, i.e., a sequence of points, that connects ns to nt. We introduce two path query evaluation paradigms that enjoy the benefits of search algorithms (i.e., fast index maintenance) while utilizing transitivity information to terminate the search sooner. Efficient indexing schemes and appropriate updating procedures are introduced. An extensive experimental evaluation verifies the advantages of our methods compared to conventional graph-based search.	
9.	Optimizing Bloom Filter Settings in Peer-to-Peer Multi keyword Searching	Peer-to-Peer multi keyword searching requires distributed intersection/union operations across wide area networks, raising a large amount of traffic cost. Existing schemes commonly utilize Bloom Filters (BFs) encoding to effectively reduce the traffic cost during the intersection/union operations. In this paper, we address the problem of optimizing the settings of a BF. We show, through mathematical proof, that the optimal setting of BF in terms of traffic cost is determined by the statistical information of the involved inverted lists, not the minimized false positive rate as claimed by previous studies. Through numerical analysis, we demonstrate how to obtain optimal settings. To better evaluate the performance of this design, we conduct comprehensive simulations on TREC WT10G test collection and query logs of a major commercial web search engine. Results show that our design significantly reduces the search traffic and latency of the existing approaches.	2012
10.	Privacy Preserving Decision Tree Learning Using Unrealized Data Sets	Privacy preservation is important for machine learning and data mining, but measures designed to protect private information often result in a trade-off: reduced utility of the training samples. This paper introduces a privacy preserving approach that can be applied to decision tree learning, without concomitant loss of accuracy. It describes an approach to the preservation of the privacy of collected data samples in cases where information from the sample database has been partially lost. This approach converts the original sample data sets into a group of unreal data sets, from which the original samples cannot be reconstructed without the entire group of unreal data sets. Meanwhile, an accurate	2012

www.ieeeproject.in

		decision tree can be built directly from those unreal data sets. This novel approach can be applied directly to the data storage as soon as the first sample is collected. The approach is compatible with other privacy preserving approaches, such as cryptography, for extra protection.	

TECHNOLOGY : DOTNET

DOMAIN : IEEE TRANSACTIONS ON DATA MINING

S.No.	IEEE TITLE	ABSTRACT	IEEE YEAR
1.	A Probabilistic Scheme for Keyword-Based Incremental Query Construction	Databases enable users to precisely express their informational needs using structured queries. However, database query construction is a laborious and error-prone process, which cannot be performed well by most end users. Keyword search alleviates the usability problem at the price of query expressiveness. As keyword search algorithms do not differentiate between the possible informational needs represented by a keyword query, users may not receive adequate results. This paper presents IQP—a novel approach to bridge the gap between usability of keyword search and expressiveness of database queries. IQP enables a user to start with an arbitrary keyword query and incrementally refine it into a structured query through an interactive interface. The enabling techniques of IQP include: 1) a probabilistic framework for incremental query construction; 2) a probabilistic model to assess the possible informational needs represented by a keyword query; 3) an algorithm to obtain the optimal query construction process. This paper presents the detailed design of IQP, and demonstrates its effectiveness and scalability through experiments over real-world data and a user study.	2012
2.	Anomaly Detection for Discrete Sequences: A Survey	This survey attempts to provide a comprehensive and structured overview of the existing research for the problem of detecting anomalies in discrete/symbolic sequences. The objective is to provide a global understanding of the sequence anomaly detection problem and how existing techniques relate to each other. The key contribution of this survey is the classification of the existing research into three distinct categories, based on the problem formulation that they are trying to solve.	2012

www.ieeeproject.in

		These problem formulations are: 1) identifying anomalous sequences with respect to a database of normal sequences; 2) identifying an anomalous subsequence within a long sequence; and 3) identifying a pattern in a sequence whose frequency of occurrence is anomalous. We show how each of these problem formulations is characteristically distinct from each other and discuss their relevance in various application domains. We review techniques from many disparate and disconnected application domains that address each of these formulations. Within each problem formulation, we group techniques into categories based on the nature of the underlying algorithm. For each category, we provide a basic anomaly detection technique, and show how the existing techniques are variants of the basic technique. This approach shows how different techniques within a category are related or different from each other. Our categorization reveals new variants and combinations that have not been investigated before for anomaly detection. We also provide a discussion of relative strengths and weaknesses of different techniques. We show how techniques developed for one problem formulation can be adapted to solve a different formulation, thereby providing several novel adaptations to solve the different problem formulations. We also highlight the applicability of the techniques that handle discrete sequences to other related areas such as online anomaly detection and time series anomaly detection.	
3.	Combining Tag and Value Similarity for Data Extraction and Alignment	Web databases generate query result pages based on a user's query. Automatically extracting the data from these query result pages is very important for many applications, such as data integration, which need to cooperate with multiple web databases. We present a novel data extraction and alignment method called CTVS that combines both tag and value similarity. CTVS automatically extracts data from query result pages by first identifying and segmenting the query result records (QRRs) in the query result pages and then aligning the segmented QRRs into a table, in which the data values from the same attribute are put into the same column. Specifically, we propose new techniques to handle the case when the QRRs are not contiguous, which may be due to the presence of auxiliary information, such as a comment, recommendation or advertisement, and for handling any nested structure that may exist in the QRRs. We also design a new record alignment algorithm that	2012

www.ieeeproject.in

		aligns the attributes in a record, first pairwise and then holistically, by combining the tag and data value similarity information. Experimental results show that CTVS achieves high precision and outperforms existing state-of-the-art data extraction methods.	
4.	Creating Evolving User Behavior Profiles Automatically	Knowledge about computer users is very beneficial for assisting them, predicting their future actions or detecting masqueraders. In this paper, a new approach for creating and recognizing automatically the behavior profile of a computer user is presented. In this case, a computer user behavior is represented as the sequence of the commands she/he types during her/his work. This sequence is transformed into a distribution of relevant subsequences of commands in order to find out a profile that defines its behavior. Also, because a user profile is not necessarily fixed but rather it evolves/changes, we propose an evolving method to keep up to date the created profiles using an Evolving Systems approach. In this paper, we combine the evolving classifier with a trie-based user profiling to obtain a powerful self-learning online scheme. We also develop further the recursive formula of the potential of a data point to become a cluster center using cosine distance, which is provided in the Appendix. The novel approach proposed in this paper can be applicable to any problem of dynamic/evolving user behavior modeling where it can be represented as a sequence of actions or events. It has been evaluated on several real data streams.	2012
5.	Horizontal Aggregations in SQL to Prepare Data Sets for Data Mining Analysis	Preparing a data set for analysis is generally the most time consuming task in a data mining project, requiring many complex SQL queries, joining tables, and aggregating columns. Existing SQL aggregations have limitations to prepare data sets because they return one column per aggregated group. In general, a significant manual effort is required to build data sets, where a horizontal layout is required. We propose simple, yet powerful, methods to generate SQL code to return aggregated columns in a horizontal tabular layout, returning a set of numbers instead of one number per row. This new class of functions is called horizontal aggregations. Horizontal aggregations build data sets with a horizontal denormalized layout (e.g., point-dimension, observation variable, instance-feature), which is the standard layout required by most data mining algorithms. We propose three fundamental methods to evaluate horizontal aggregations: CASE: Exploiting the	2012

www.ieeeproject.in

		programming CASE construct; SPJ: Based on standard relational algebra operators (SPJ queries); PIVOT: Using the PIVOT operator, which is offered by some DBMSs. Experiments with large tables compare the proposed query evaluation methods. Our CASE method has similar speed to the PIVOT operator and it is much faster than the SPJ method. In general, the CASE and PIVOT methods exhibit linear scalability, whereas the SPJ method does not.	
6.	Slicing: A New Approach for Privacy Preserving Data Publishing	Several anonymization techniques, such as generalization and bucketization, have been designed for privacy preserving micro data publishing. Recent work has shown that generalization loses considerable amount of information, especially for high dimensional data. Bucketization, on the other hand, does not prevent membership disclosure and does not apply for data that do not have a clear separation between quasi-identifying attributes and sensitive attributes. In this paper, we present a novel technique called slicing, which partitions the data both horizontally and vertically. We show that slicing preserves better data utility than generalization and can be used for membership disclosure protection. Another important advantage of slicing is that it can handle high-dimensional data. We show how slicing can be used for attribute disclosure protection and develop an efficient algorithm for computing the sliced data that obey the ‘-diversity requirement. Our workload experiments confirm that slicing preserves better utility than generalization and is more effective than bucketization in workloads involving the sensitive attribute. Our experiments also demonstrate that slicing can be used to prevent membership disclosure.	2012
7.	Tree-Based Mining for Discovering Patterns of Human Interaction in Meetings	Discovering semantic knowledge is significant for understanding and interpreting how people interact in a meeting discussion. In this paper, we propose a mining method to extract frequent patterns of human interaction based on the captured content of face-to-face meetings. Human interactions, such as proposing an idea, giving comments, and expressing a positive opinion, indicate user intention toward a topic or role in a discussion. Human interaction flow in a discussion session is represented as a tree. Tree based interaction mining algorithms are designed to analyze the structures of the trees and to extract interaction flow patterns. The experimental results show that we can successfully extract several interesting patterns that are useful for the	2012

www.ieeeproject.in

		interpretation of human behavior in meeting discussions, such as determining frequent interactions, typical interaction flows, and relationships between different types of interactions.	

TECHNOLOGY : JAVA

DOMAIN : IEEE TRANSACTIONS ON NETWORKING

S.No.	IEEE TITLE	ABSTRACT	IEEE YEAR
1.	Adaptive Opportunistic Routing for Wireless Ad Hoc Networks	A distributed adaptive opportunistic routing scheme for multi-hop wireless ad hoc networks is proposed. The proposed scheme utilizes a reinforcement learning framework to opportunistically route the packets even in the absence of reliable knowledge about channel statistics and network model. This scheme is shown to be optimal with respect to an expected average per-packet reward criterion. The proposed routing scheme jointly addresses the issues of learning and routing in an opportunistic context, where the network structure is characterized by the transmission success probabilities. In particular, this learning framework leads to a stochastic routing scheme that optimally "explores" and "exploits" the opportunities in the network.	2012
2.	Efficient Error Estimating Coding: Feasibility and Applications	Motivated by recent emerging systems that can leverage partially correct packets in wireless networks; this paper proposes the novel concept of error estimating coding (EEC). Without correcting the errors in the packet, EEC enables the receiver of the packet to estimate the packet's bit error rate, which is perhaps the most important meta-information of a partially correct packet. Our EEC design provides provable estimation quality with rather low redundancy and computational overhead. To demonstrate the utility of EEC, we exploit and implement EEC in two wireless network applications, Wi-Fi rate adaptation and real-time video streaming. Our real-world experiments show that these applications can significantly benefit from EEC.	2012
3.	Exploiting Excess Capacity to Improve Robustness	Excess capacity (EC) is the unused capacity in a network. We propose EC management techniques to improve network performance. Our techniques exploit the EC in two ways. First, a connection pre provisioning algorithm is used to reduce the connection setup time. Second,	2012

www.ieeeproject.in

	of WDM Mesh Networks	whenever possible, we use protection schemes that have higher availability and shorter protection switching time. Specifically, depending on the amount of EC available in the network, our proposed EC management techniques dynamically migrate connections between high-availability, high-backup-capacity protection schemes and low-availability, low-backup-capacity protection schemes. Thus, multiple protection schemes can coexist in the network. The four EC management techniques studied in this paper differ in two respects: when the connections are migrated from one protection scheme to another, and which connections are migrated. Specifically, Lazy techniques migrate connections only when necessary, whereas Proactive techniques migrate connections to free up capacity in advance. Partial Backup Reprovisioning (PBR) techniques try to migrate a minimal set of connections, whereas Global Backup Reprovisioning (GBR) techniques migrate all connections. We develop integer linear program (ILP) formulations and heuristic algorithms for the EC management techniques. We then present numerical examples to illustrate how the EC management techniques improve network performance by exploiting the EC in wavelength-division-multiplexing (WDM) mesh networks.	
4.	Improving Energy Saving and Reliability in Wireless Sensor Networks Using a Simple CRT-Based Packet-Forwarding Solution	This paper deals with a novel forwarding scheme for wireless sensor networks aimed at combining low computational complexity and high performance in terms of energy efficiency and reliability. The proposed approach relies on a packet-splitting algorithm based on the Chinese Remainder Theorem (CRT) and is characterized by a simple modular division between integers. An analytical model for estimating the energy efficiency of the scheme is presented, and several practical issues such as the effect of unreliable channels, topology changes, and MACoverhead are discussed. The results obtained show that the proposed algorithm outperforms traditional approaches in terms of power saving, simplicity, and fair distribution of energy consumption among all nodes in the network.	2012
5.	Independent Directed Acyclic Graphs for Resilient	In order to achieve resilient multipath routing, we introduce the concept of independent directed acyclic graphs (IDAGs) in this paper. Link-independent (node-independent) DAGs satisfy the property that any path from a source to the root on one DAG is link-disjoint	2012

www.ieeeproject.in

	Multipath Routing	(node-disjoint) with any path from the source to the root on the other DAG. Given a network, we develop polynomial- time algorithms to compute link-independent and node-independent DAGs. The algorithm developed in this paper: 1) provides multipath routing; 2) utilizes all possible edges; 3) guarantees recovery from single link failure; and 4) achieves all these with at most one bit per packet as overhead when routing is based on destination address and incoming edge. We show the effectiveness of the proposed IDAGs approach by comparing key performance indices to that of the independent trees and multiple pairs of independent trees techniques through extensive simulations.	
6.	Latency Equalization as a New Network Service Primitive	Multiparty interactive network applications such as teleconferencing, network gaming, and online trading are gaining popularity. In addition to end-to-end latency bounds, these applications require that the delay difference among multiple clients of the service is minimized for a good interactive experience. We propose a Latency EQualization (LEQ) service, which equalizes the perceived latency for all clients participating in an interactive network application. To effectively implement the proposed LEQ service, network support is essential. The LEQ architecture uses a few routers in the network as hubs to redirect packets of interactive applications along paths with similar end-to-end delay. We first formulate the hub selection problem, prove its NP-hardness, and provide a greedy algorithm to solve it. Through extensive simulations, we show that our LEQ architecture significantly reduces delay difference under different optimization criteria that allow or do not allow compromising the per-user end-to-end delay. Our LEQ service is incrementally deployable in today's networks, requiring just software modifications to edge routers.	2012
7.	Opportunistic Flow-Level Latency Estimation Using Consistent NetFlow	The inherent measurement support in routers (SNMP counters or NetFlow) is not sufficient to diagnose performance problems in IP networks, especially for flow-specific problems where the aggregate behavior within a router appears normal. Tomographic approaches to detect the location of such problems are not feasible in such cases as active probes can only catch aggregate characteristics. To address this problem, in this paper, we propose a Consistent NetFlow (CNF) architecture for measuring per-flow delay measurements within routers. CNF utilizes the existing NetFlow architecture that already reports the first	2012

www.ieeeproject.in

		and last timestamps per flow, and it proposes hash-based sampling to ensure that two adjacent routers record the same flows. We devise a novel Multiflow estimator that approximates the intermediate delay samples from other background flows to significantly improve the per-flow latency estimates compared to the naïve estimator that only uses actual flow samples. In our experiments using real backbone traces and realistic delay models, we show that the Multiflow estimator is accurate with a median relative error of less than 20% for flows of size greater than 100 packets. We also show that Multiflow estimator performs two to three times better than a prior approach based on trajectory sampling at an equivalent packet sampling rate.	

TECHNOLOGY : JAVA

DOMAIN : IEEE TRANSACTIONS ON MOBILE COMPUTING

S.No.	IEEE TITLE	ABSTRACT	IEEE YEAR
1.	Acknowledgment-Based Broadcast Protocol for Reliable and Efficient Data Dissemination in Vehicular Ad Hoc Networks	We propose a broadcast algorithm suitable for a wide range of vehicular scenarios, which only employs local information acquired via periodic beacon messages, containing acknowledgments of the circulated broadcast messages. Each vehicle decides whether it belongs to a connected dominating set (CDS). Vehicles in the CDS use a shorter waiting period before possible retransmission. At time-out expiration, a vehicle retransmits if it is aware of at least one neighbor in need of the message. To address intermittent connectivity and appearance of new neighbors, the evaluation timer can be restarted. Our algorithm resolves propagation at road intersections without any need to even recognize intersections. It is inherently adaptable to different mobility regimes, without the need to classify network or vehicle speeds. In a thorough simulation-based performance evaluation, our algorithm is shown to provide higher reliability and message efficiency than existing approaches for non safety applications.	2012
2.	FESCIM: Fair, Efficient, and Secure Cooperation Incentive	In multihop cellular networks, the mobile nodes usually relay others' packets for enhancing the network performance and deployment. However, selfish nodes usually do not cooperate but make use of the cooperative nodes to relay their packets, which has	2012

www.ieeeproject.in

	Mechanism for Multihop Cellular Networks	a negative effect on the network fairness and performance. In this paper, we propose a fair and efficient incentive mechanism to stimulate the node cooperation. Our mechanism applies a fair charging policy by charging the source and destination nodes when both of them benefit from the communication. To implement this charging policy efficiently, hashing operations are used in the ACK packets to reduce the number of public-key-cryptography operations. Moreover, reducing the overhead of the payment checks is essential for the efficient implementation of the incentive mechanism due to the large number of payment transactions. Instead of generating a check per message, a small-size check can be generated per route, and a check submission scheme is proposed to reduce the number of submitted checks and protect against collusion attacks. Extensive analysis and simulations demonstrate that our mechanism can secure the payment and significantly reduce the checks' overhead, and the fair charging policy can be implemented almost computationally free by using hashing operations.	
3.	Characterizing the Security Implications of Third-Party Emergency Alert Systems over Cellular Text Messaging Services	Cellular text messaging services are increasingly being relied upon to disseminate critical information during emergencies. Accordingly, a wide range of organizations including colleges and universities now partner with third-party providers that promise to improve physical security by rapidly delivering such messages. Unfortunately, these products do not work as advertised due to limitations of cellular infrastructure and therefore provide a false sense of security to their users. In this paper, we perform the first extensive investigation and characterization of the limitations of an Emergency Alert System (EAS) using text messages as a security incident response mechanism. We show emergency alert systems built on text messaging not only can meet the 10 minute delivery requirement mandated by the WARN Act, but also potentially cause other voice and SMS traffic to be blocked at rates upward of 80 percent. We then show that our results are representative of reality by comparing them to a number of documented but not previously understood failures. Finally, we analyze a targeted messaging mechanism as a means of efficiently using currently deployed infrastructure and third-party EAS. In so doing, we demonstrate that this	2012

www.ieeeproject.in

		increasingly deployed security infrastructure does not achieve its stated requirements for large populations.	
4.	Handling Selfishness in Replica Allocation over a Mobile Ad Hoc Network	In a mobile ad hoc network, the mobility and resource constraints of mobile nodes may lead to network partitioning or performance degradation. Several data replication techniques have been proposed to minimize performance degradation. Most of them assume that all mobile nodes collaborate fully in terms of sharing their memory space. In reality, however, some nodes may selfishly decide only to cooperate partially, or not at all, with other nodes. These selfish nodes could then reduce the overall data accessibility in the network. In this paper, we examine the impact of selfish nodes in a mobile ad hoc network from the perspective of replica allocation. We term this selfish replica allocation. In particular, we develop a selfish node detection algorithm that considers partial selfishness and novel replica allocation techniques to properly cope with selfish replica allocation. The conducted simulations demonstrate the proposed approach outperforms traditional cooperative replica allocation techniques in terms of data accessibility, communication cost, and average query delay.	2012
5.	Local Broadcast Algorithms in Wireless Ad Hoc Networks: Reducing the Number of Transmissions	There are two main approaches, static and dynamic, to broadcast algorithms in wireless ad hoc networks. In the static approach, local algorithms determine the status (forwarding/nonforwarding) of each node proactively based on local topology information and a globally known priority function. In this paper, we first show that local broadcast algorithms based on the static approach cannot achieve a good approximation factor to the optimum solution (an NP-hard problem). However, we show that a constant approximation factor is achievable if (relative) position information is available. In the dynamic approach, local algorithms determine the status of each node “on-the-fly” based on local topology information and broadcast state information. Using the dynamic approach, it was recently shown that local broadcast algorithms can achieve a constant approximation factor to the optimum solution when (approximate) position information is available. However, using position information can simplify the problem. Also, in some applications it may not be practical to have position information. Therefore, we wish to know whether local broadcast algorithms based on the dynamic	2012

www.ieeeproject.in

		approach can achieve a constant approximation factor without using position information. We answer this question in the positive—we design a local broadcast algorithm in which the status of each node is decided "on-the-fly" and prove that the algorithm can achieve both full delivery and a constant approximation to the optimum solution.	

TECHNOLOGY : JAVA

DOMAIN : IEEE TRANSACTIONS ON IMAGE PROCESSING

S.No.	IEEE TITLE	ABSTRACT	IEEE YEAR
1.	A Primal–Dual Method for Total-Variation-Based Wavelet Domain Inpainting	Loss of information in a wavelet domain can occur during storage or transmission when the images are formatted and stored in terms of wavelet coefficients. This calls for image inpainting in wavelet domains. In this paper, a variational approach is used to formulate the reconstruction problem. We propose a simple but very efficient iterative scheme to calculate an optimal solution and prove its convergence. Numerical results are presented to show the performance of the proposed algorithm.	2012
2.	A Secret-Sharing-Based Method for Authentication of Grayscale Document Images via the Use of the PNG Image With a Data Repair Capability	A new blind authentication method based on the secret sharing technique with a data repair capability for grayscale document images via the use of the Portable Network Graphics (PNG) image is proposed. An authentication signal is generated for each block of a grayscale document image, which, together with the binarized block content, is transformed into several shares using the Shamir secret sharing scheme. The involved parameters are carefully chosen so that as many shares as possible are generated and embedded into an alpha channel plane. The alpha channel plane is then combined with the original grayscale image to form a PNG image. During the embedding process, the computed share values are mapped into a range of alpha channel values near their maximum value of 255 to yield a transparent stego-image with a disguise effect. In the process of image authentication, an image block is marked as tampered if the authentication signal computed from the current block content does not match that extracted from the shares embedded in the alpha	2012

www.ieeeproject.in

		channel plane. Data repairing is then applied to each tampered block by a reverse Shamir scheme after collecting two shares from unmarked blocks. Measures for protecting the security of the data hidden in the alpha channel are also proposed. Good experimental results prove the effectiveness of the proposed method for real applications.	
3.	Image Reduction Using Means on Discrete Product Lattices	We investigate the problem of averaging values on lattices and, in particular, on discrete product lattices. This problem arises in image processing when several color values given in RGB, HSL, or another coding scheme need to be combined. We show how the arithmetic mean and the median can be constructed by minimizing appropriate penalties, and we discuss which of them coincide with the Cartesian product of the standard mean and the median. We apply these functions in image processing. We present three algorithms for color image reduction based on minimizing penalty functions on discrete product lattices.	2012
4.	Vehicle Detection in Aerial Surveillance Using Dynamic Bayesian Networks	We present an automatic vehicle detection system for aerial surveillance in this paper. In this system, we escape from the stereotype and existing frameworks of vehicle detection in aerial surveillance, which are either region based or sliding window based. We design a pixel wise classification method for vehicle detection. The novelty lies in the fact that, in spite of performing pixel wise classification, relations among neighboring pixels in a region are preserved in the feature extraction process. We consider features including vehicle colors and local features. For vehicle color extraction, we utilize a color transform to separate vehicle colors and non-vehicle colors effectively. For edge detection, we apply moment preserving to adjust the thresholds of the Canny edge detector automatically, which increases the adaptability and the accuracy for detection in various aerial images. Afterward, a dynamic Bayesian network (DBN) is constructed for the classification purpose. We convert regional local features into quantitative observations that can be referenced when applying pixel wise classification via DBN. Experiments were conducted on a wide variety of aerial videos. The results demonstrate flexibility and good generalization abilities of the proposed method on a challenging data set with aerial surveillance images taken at different heights and under different camera angles.	2012
5.	Abrupt	The robust tracking of abrupt motion is a challenging	2012

www.ieeeproject.in

	Motion Tracking Via Intensively Adaptive Markov-Chain Monte Carlo Sampling	task in computer vision due to its large motion uncertainty. While various particle filters and conventional Markov-chain Monte Carlo (MCMC) methods have been proposed for visual tracking, these methods often suffer from the well-known local-trap problem or from poor convergence rate. In this paper, we propose a novel sampling-based tracking scheme for the abrupt motion problem in the Bayesian filtering framework. To effectively handle the local-trap problem, we first introduce the stochastic approximation Monte Carlo (SAMC) sampling method into the Bayesian filter tracking framework, in which the filtering distribution is adaptively estimated as the sampling proceeds, and thus, a good approximation to the target distribution is achieved. In addition, we propose a new MCMC sampler with intensive adaptation to further improve the sampling efficiency, which combines a density-grid-based predictive model with the SAMC sampling, to give a proposal adaptation scheme. The proposed method is effective and computationally efficient in addressing the abrupt motion problem. We compare our approach with several alternative tracking algorithms, and extensive experimental results are presented to demonstrate the effectiveness and the efficiency of the proposed method in dealing with various types of abrupt motions.	

TECHNOLOGY : JAVA

DOMAIN : IEEE TRANSACTIONS ON SOFTWARE ENGINEERING

S.No.	IEEE TITLE	ABSTRACT	IEEE YEAR
1.	Automatic Detection of Unsafe Dynamic Component Loadings	Dynamic loading of software components (e.g., libraries or modules) is a widely used mechanism for an improved system modularity and flexibility. Correct component resolution is critical for reliable and secure software execution. However, programming mistakes may lead to unintended or even malicious components being resolved and loaded. In particular, dynamic loading can be hijacked by placing an arbitrary file with the specified name in a directory searched before	2012

www.ieeeproject.in

		resolving the target component. Although this issue has been known for quite some time, it was not considered serious because exploiting it requires access to the local file system on the vulnerable host. Recently, such vulnerabilities have started to receive considerable attention as their remote exploitation became realistic. It is now important to detect and fix these vulnerabilities. In this paper, we present the first automated technique to detect vulnerable and unsafe dynamic component loadings. Our analysis has two phases: 1) apply dynamic binary instrumentation to collect runtime information on component loading (online phase), and 2) analyze the collected information to detect vulnerable component loadings (offline phase). For evaluation, we implemented our technique to detect vulnerable and unsafe component loadings in popular software on Microsoft Windows and Linux. Our evaluation results show that unsafe component loading is prevalent in software on both OS platforms, and it is more severe on Microsoft Windows. In particular, our tool detected more than 4,000 unsafe component loadings in our evaluation, and some can lead to remote code execution on Microsoft Windows.	
2.	Fault Localization for Dynamic Web Applications	In recent years, there has been significant interest in fault-localization techniques that are based on statistical analysis of program constructs executed by passing and failing executions. This paper shows how the Tarantula, Ochiai, and Jaccard fault-localization algorithms can be enhanced to localize faults effectively in web applications written in PHP by using an extended domain for conditional and function-call statements and by using a source mapping. We also propose several novel test-generation strategies that are geared toward producing test suites that have maximal fault-localization effectiveness. We implemented various fault-localization techniques and test-generation strategies in Apollo, and evaluated them on several open-source PHP applications. Our results indicate that a variant of the Ochiai algorithm that includes all our enhancements localizes 87.8 percent of all faults to within 1 percent of all executed statements, compared to only 37.4 percent for the unenhanced Ochiai algorithm. We also found that all the test-generation strategies that we considered are capable of generating test suites with maximal fault-localization effectiveness when given an infinite time budget for test generation. However, on	2012

www.ieeeproject.in

		average, a directed strategy based on path-constraint similarity achieves this maximal effectiveness after generating only 6.5 tests, compared to 46.8 tests for an undirected test-generation strategy.	
3.	Input Domain Reduction through Irrelevant Variable Removal and Its Effect on Local, Global, and Hybrid Search-Based	Search-Based Test Data Generation reformulates testing goals as fitness functions so that test input generation can be automated by some chosen search-based optimization algorithm. The optimization algorithm searches the space of potential inputs, seeking those that are "fit for purpose," guided by the fitness function. The search space of potential inputs can be very large, even for very small systems under test. Its size is, of course, a key determining factor affecting the performance of any search-based approach. However, despite the large volume of work on Search-Based Software Testing, the literature contains little that concerns the performance impact of search space reduction. This paper proposes a static dependence analysis derived from program slicing that can be used to support search space reduction. The paper presents both a theoretical and empirical analysis of the application of this approach to open source and industrial production code. The results provide evidence to support the claim that input domain reduction has a significant effect on the performance of local, global, and hybrid search, while a purely random search is unaffected.	2012
4.	PerLa:A Language and Middleware Architecture for Data Management and Integration	A declarative SQL-like language and a middleware infrastructure are presented for collecting data from different nodes of a pervasive system. Data management is performed by hiding the complexity due to the large underlying heterogeneity of devices, which can span from passive RFID(s) to ad hoc sensor boards to portable computers. An important feature of the presented middleware is to make the integration of new device types in the system easy through the use of device self-description. Two case studies are described for PerLa usage, and a survey is made for comparing our approach with other projects in the area.	2012
5.	Comparing Semi-Automated Clustering Methods for Persona Development	Current and future information systems require a better understanding of the interactions between users and systems in order to improve system use and, ultimately, success. The use of personas as design tools is becoming more widespread as researchers and practitioners discover its benefits. This paper presents an empirical study comparing the performance of existing qualitative and quantitative clustering techniques for the task of	2012

www.ieeeproject.in

		identifying personas and grouping system users into those personas. A method based on Factor (Principal Components) Analysis performs better than two other methods which use Latent Semantic Analysis and Cluster Analysis as measured by similarity to expert manually defined clusters	
6.	StakeRare: Using Social Networks and Collaborative Filtering for Large-Scale Requirements Elicitation	Requirements elicitation is the software engineering activity in which stakeholder needs are understood. It involves identifying and prioritizing requirements-a process difficult to scale to large software projects with many stakeholders. This paper proposes StakeRare, a novel method that uses social networks and collaborative filtering to identify and prioritize requirements in large software projects. StakeRare identifies stakeholders and asks them to recommend other stakeholders and stakeholder roles, builds a social network with stakeholders as nodes and their recommendations as links, and prioritizes stakeholders using a variety of social network measures to determine their project influence. It then asks the stakeholders to rate an initial list of requirements, recommends other relevant requirements to them using collaborative filtering, and prioritizes their requirements using their ratings weighted by their project influence. StakeRare was evaluated by applying it to a software project for a 30,000-user system, and a substantial empirical study of requirements elicitation was conducted. Using the data collected from surveying and interviewing 87 stakeholders, the study demonstrated that StakeRare predicts stakeholder needs accurately and arrives at a more complete and accurately prioritized list of requirements compared to the existing method used in the project, taking only a fraction of the time	
7.	QoS Assurance for Dynamic Reconfiguration of Component-Based Software	A major challenge of dynamic reconfiguration is Quality of Service (QoS) assurance, which is meant to reduce application disruption to the minimum for the system's transformation. However, this problem has not been well studied. This paper investigates the problem for component-based software systems from three points of view. First, the whole spectrum of QoS characteristics is defined. Second, the logical and physical requirements for QoS characteristics are analyzed and solutions to achieve them are proposed. Third, prior work is classified by QoS characteristics and then realized by abstract reconfiguration strategies. On this basis, quantitative evaluation of the QoS assurance abilities of	2012

www.ieeeproject.in

		existing work and our own approach is conducted through three steps. First, a proof-of-concept prototype called the reconfigurable component model is implemented to support the representation and testing of the reconfiguration strategies. Second, a reconfiguration benchmark is proposed to expose the whole spectrum of QoS problems. Third, each reconfiguration strategy is tested against the benchmark and the testing results are evaluated. The most important conclusion from our investigation is that the classified QoS characteristics can be fully achieved under some acceptable constraints.	

TECHNOLOGY : JAVA

DOMAIN : IEEE TRANSACTIONS ON SECURE COMPUTING

S.No.	IEEE TITLE	ABSTRACT	IEEE YEAR
1.	Revisiting Defenses against Large-Scale Online Password Guessing Attacks	Brute force and dictionary attacks on password-only remote login services are now widespread and ever increasing.Enabling convenient login for legitimate users while preventing such attacks is a difficult problem. Automated Turing Tests (ATTs) continue to be an effective, easy-to-deploy approach to identify automated malicious login attempts with reasonable cost of inconvenience to users. In this paper, we discuss the inadequacy of existing and proposed login protocols designed to address large-scale online dictionary attacks (e.g., from a botnet of hundreds of thousands of nodes). We propose a new Password Guessing Resistant Protocol (PGRP), derived upon revisiting prior proposals designed to restrict such attacks. While PGRP limits the total number of login attempts from unknown remote hosts to as low as a single attempt per username, legitimate users in most cases (e.g., when attempts are made from known, frequently-used machines) can make several failed login attempts before being challenged with an ATT. We analyze the performance of PGRP with two real-world data sets and find it more promising than existing proposals	2012

www.ieeeproject.in

2.	Data-Provenance Verification For Secure Hosts	Malicious software typically resides stealthily on a user's computer and interacts with the user's computing resources. Our goal in this work is to improve the trustworthiness of a host and its system data. Specifically, we provide a new mechanism that ensures the correct origin or provenance of critical system information and prevents adversaries from utilizing host resources. We define data-provenance integrity as the security property stating that the source where a piece of data is generated cannot be spoofed or tampered with. We describe a cryptographic provenance verification approach for ensuring system properties and system-data integrity at kernel-level. Its two concrete applications are demonstrated in the keystroke integrity verification and malicious traffic detection. Specifically, we first design and implement an efficient cryptographic protocol that enforces keystroke integrity by utilizing on-chip Trusted Computing Platform (TPM). The protocol prevents the forgery of fake key events by malware under reasonable assumptions. Then, we demonstrate our provenance verification approach by realizing a lightweight framework for restricting outbound malware traffic. This traffic-monitoring framework helps identify network activities of stealthy malware, and lends itself to a powerful personal firewall for examining all outbound traffic of a host that cannot be bypassed	2012
3.	Design and Implementation of TARF:A Trust-Aware Routing Framework for WSNs	The multihop routing in wireless sensor networks (WSNs) offers little protection against identity deception through replaying routing information. An adversary can exploit this defect to launch various harmful or even devastating attacks against the routing protocols, including sinkhole attacks, wormhole attacks, and Sybil attacks. The situation is further aggravated by mobile and harsh network conditions. Traditional cryptographic techniques or efforts at developing trust-aware routing protocols do not effectively address this severe problem. To secure the WSNs against adversaries misdirecting the multihop routing, we have designed and implemented TARF, a robust trust-aware routing framework for dynamic WSNs. Without tight time synchronization or known geographic information, TARF provides trustworthy and energy-efficient route. Most importantly, TARF proves effective against those harmful attacks developed out of identity deception; the resilience of TARF is verified through extensive evaluation with both simulation and empirical experiments on large-scale	2012

www.ieeeproject.in

		WSNs under various scenarios including mobile and RF-shielding network conditions. Further, we have implemented a low-overhead TARF module in TinyOS; as demonstrated, this implementation can be incorporated into existing routing protocols with the least effort. Based on TARF, we also demonstrated a proof-of-concept mobile target detection application that functions well against an antidetection mechanism.	
4.	On the Security and Efficiency of Content Distribution via Network Coding	Content distribution via network coding has received a lot of attention lately. However, direct application of network coding may be insecure. In particular, attackers can inject "bogus" data to corrupt the content distribution process so as to hinder the information dispersal or even deplete the network resource. Therefore, content verification is an important and practical issue when network coding is employed. When random linear network coding is used, it is infeasible for the source of the content to sign all the data, and hence, the traditional "hash-and-sign" methods are no longer applicable. Recently, a new on-the-fly verification technique has been proposed by Krohn et al. (IEEE S&P '04), which employs a classical homomorphic hash function. However, this technique is difficult to be applied to network coding because of high computational and communication overhead. We explore this issue further by carefully analyzing different types of overhead, and propose methods to help reducing both the computational and communication cost, and provide provable security at the same time	2012
5.	Detecting Anomalous Insiders in Collaborative Information Systems	Collaborative information systems (CISs) are deployed within a diverse array of environments that manage sensitive information. Current security mechanisms detect insider threats, but they are ill-suited to monitor systems in which users function in dynamic teams. In this paper, we introduce the community anomaly detection system (CADS), an unsupervised learning framework to detect insider threats based on the access logs of collaborative environments. The framework is based on the observation that typical CIS users tend to form community structures based on the subjects accessed (e.g., patients' records viewed by healthcare providers). CADS consists of two components: 1) relational pattern extraction, which derives community structures and 2) anomaly prediction, which leverages a statistical model to determine when users have sufficiently deviated from communities. We further extend CADS into MetaCADS	2012

www.ieeeproject.in

		to account for the semantics of subjects (e.g., patients' diagnoses). To empirically evaluate the framework, we perform an assessment with three months of access logs from a real electronic health record (EHR) system in a large medical center. The results illustrate our models exhibit significant performance gains over state-of-the-art competitors. When the number of illicit users is low, MetaCADS is the best model, but as the number grows, commonly accessed semantics lead to hiding in a crowd, such that CADS is more prudent.	
6.	ES-MPICH2: A Message Passing Interface with Enhanced Security	An increasing number of commodity clusters are connected to each other by public networks, which have become a potential threat to security sensitive parallel applications running on the clusters. To address this security issue, we developed a Message Passing Interface (MPI) implementation to preserve confidentiality of messages communicated among nodes of clusters in an unsecured network. We focus on M PI rather than other protocols, because M PI is one of the most popular communication protocols for parallel computing on clusters. Our MPI implementation-called ES-MPICH2-was built based on MPICH2 developed by the Argonne National Laboratory. Like MPICH2, ES-MPICH2 aims at supporting a large variety of computation and communication platforms like commodity clusters and high-speed networks. We integrated encryption and decryption algorithms into the MPICH2 library with the standard MPI interface and; thus, data confidentiality of MPI applications can be readily preserved without a need to change the source codes of the MPI applications. MPI-application programmers can fully configure any confidentiality services in MPICHI2, because a secured configuration file in ES-MPICH2 offers the programmers flexibility in choosing any cryptographic schemes and keys seamlessly incorporated in ES-MPICH2. We used the Sandia Micro Benchmark and Intel MPI Benchmark suites to evaluate and compare the performance of ES-MPICH2 with the original MPICH2 version. Our experiments show that overhead incurred by the confidentiality services in ES-MPICH2 is marginal for small messages. The security overhead in ES-MPICH2 becomes more pronounced with larger messages. Our results also show that security overhead can be significantly reduced in ES-MPICH2 by high-performance clustersRequirements elicitation is the software engineering activity in which	

www.ieeeproject.in

7.	On the Security of a Ticket-Based Anonymity System with Traceability Property in Wireless Mesh Networks	In 2011, Sun et al. [CHECK END OF SENTENCE] proposed a security architecture to ensure unconditional anonymity for honest users and traceability of misbehaving users for network authorities in wireless mesh networks (WMNs). It strives to resolve the conflicts between the anonymity and traceability objectives. In this paper, we attacked Sun et al. scheme's traceability. Our analysis showed that trusted authority (TA) cannot trace the misbehavior client (CL) even if it double-time deposits the same ticket.	2012

www.ieeeproject.in