Spring 2-3-2013

Proceeding of International Conference on Computer Science and Information Technology ICCSIT-2013

Prof. (Dr.) Srikanta Patnaik

Follow this and additional works at: https://www.interscience.in/conf_proc_volumes

Part of the Computational Engineering Commons, and the Digital Communications and Networking Commons
**Editorial**

The Computer Science and Information Technology deals with Data Compression, Cryptography, Error Detection and Correction. More recently it also can serve Network Coding. Traditionally, Computing studies occupy two partitions—Sciences and Engineering, separated by a line roughly at the Computer Architecture level. A more effective organization for Computer Science and Engineering requires an intrinsically interdisciplinary framework that combines academics and systems-oriented computing perspectives. Researchers have been developing such a framework, which re-aggregates Computer Science and Computer Engineering, then repartitions the resulting single field into analysis and synthesis components. The framework is based on the notion that Science is foremost about dissecting and understanding. Engineering is mostly about envisioning and building. For modeling, it is believed that Computer is the basic Infrastructure to centralize communication. Any communication between people about the same concept is a common revelatory experience about informational models of that concept. Each model is a conceptual structure of abstractions formulated initially in the mind of one, and while communicating if it is different from those in the mind of other, there is no common model and no communication. The introduction of Computers Science and Technology paved the way for the emergence of new Technologies. Computer Engineers over the next two decades will be called upon to develop Technologies that foster a cleaner, healthier, safer and sustainable Global Environment. "Engineers will be able to act as independent operators interacting with colleagues around the world," the report says. Engineers can design at home with advanced CAD systems or in collaboration with their Global colleagues in virtual Worlds. They will be able to use home-based Fabrication Technology to test many of their designs. As Mechanical Engineering looks to 2028, leaders will value people with diverse expertise and experience. They will bring these Global professions together to keep the promise of Technology serving people. They will inspire men and women everywhere to believe that grand challenges are a rally cry for a profession that is ready for the adventure of making the difficult doable. The main aim of the Conference is to communicate high quality original research work, reviews, short communications, in the fields of Computer Science and Information Technology.

The Computer had a great effect on Communication. The idea of modeling in a computer and with the aid of a Computer is the basic Infrastructure to centralize communication. Any communication between people about the same concept is a common revelatory experience about informational models of that concept. Researchers are working on applying their wireless and mobile research to Transportation, Health care, Educational collaboration and Environmental Sustainability. Projects already underway include safe and efficient Road Transportation, Autonomous Driving, Wireless Medical Implants, Mobile Video Delivery, Multiparty Wireless Videoconferencing and Energy Harvesting. The Conference sometimes is conducted in collaboration with other Institutions. IRNet encourage and invite proposals from Institutes within India and Abroad to join hands to promote research in various areas of disciplines. These Conferences have not only promoted the International exchange and cooperation, but have also won favorable comments from National and International participants, thus enabled IRNet to reach out to a Global Network within three years time. The Conference is first of its kind and gets granted with lot of blessings.

I sincerely thank all the authors for their invaluable contribution to this Conference. I am indebted towards the Reviewers and Board of Editors for their generous gifts of time, energy and effort. It’s my pleasure to welcome all the participants, delegates and organizer to this International Conference on behalf of IRNet family members.

I wish all success to the paper presenters. The papers qualifying the review process will be published in the forthcoming IOAJ Journal.
ENERGY BASED COMPUTATION IN WIRELESS SENSOR NETWORK: A REVIEW

SONALI RAUT1 & HEMLATA DAKHORE2

1,2Department Computer Science & Engineering, GHRIETW, RTMNU University, Nagpur

Abstract-In wireless sensor network, all nodes or devices are generally battery power-driven devices. These nodes have restricted amount of initial energy that are consumed at different rates, depending on the power level. The lifetime of the network is defined as the time until the first node fails. In this paper different type of energy efficient routing algorithms based on linear programming are discussed and approach of these algorithms is to maximize the minimum lifetime of wireless sensor network. Special attention has been devoted for algorithms formulate the routing problem, which uses the optimal flow path for data transmission and gives the optimum results. So comparative study of these algorithms are also discussed in this paper.

Keywords-Lifetime of wireless sensor network, Battery power, Minimum energy cost routing, Optimal flow, Linear programming problem.

1. INTRODUCTION

Today’s wireless communication has been an exponential growth caused by the need for connectivity. A similar trend in wireless networking has been followed due to the increasing exchange of data in services such as the Internet. The increasing need for data throughput, capabilities needed to deliver such services. Other applications in fields such as, agricultural, industrial, vehicular, medical sensors, residential and actuators have more relaxed throughput requirements. This type of network consists of a collection of nodes and each node has limited battery power. There may be many possible routes available between two nodes. For example, if node has sufficient battery power, each node generated some information

and this information needs to be delivered to set of destination nodes. If the node lies in its vicinity than it transmits the data to other node without any interference. A large battery power is required to transmit the data to a node which is situated far from source node. After few transmissions a node reaches to its threshold battery level and it may exclude from the data transmission and overall lifetime of network will decrease. Whereas network lifetime is defined as the time until the first node in the network dies.

For maximizing the lifetime of network, the data should be transferred such that energy consumption is balanced among the nodes in proportion to their energy reserved, instead of routing to minimize consumed power. The minimum energy routing is proposed as in [1] approach in this work is to minimize the energy consumption to reach the destination by sending the traffic to same path but if all the traffic follows the same path then all the nodes of that path will drain out there energy quickly. Instead of trying to minimize the consumed energy the main objective is to maximize the lifetime of the system. The maximum lifetime problem is a linear programming problem and solvable in polynomial time defined by Chang and Tassiulas [2]. In this work considers the single commodity version of the problem but [3] the problem is extended to multicommodity case, where each commodity has its own set of destinations.

[2],[3] proposed the system in which maximizing the lifetime of a network when message rate is known but Q. Li, J. Aslam and D. Ras proposed online, hierarchical and scalable algorithms [4] that do not rely on knowing the message rate and optimize the lifetime of network. Routing algorithms [2],[3],[4] consider energy consumption on sender side only, but as in [5] the maximum lifetime routing problem is extended to include the energy consumption at the receiver during reception. The relation of maximizing the minimum lifetime of the nodes to minimizing the cost per packet was defined as in [2],[3],[4],[5] but this relation take one step further to provide a delay guarantee in the time the packets reach their destination, while maximizing network lifetime [6].

In [7] described that if nodes in an ad hoc wireless network expend most of their power on communication-related applications, then power aware routing protocols like minimum battery cost and min-max battery cost schemes can prevent nodes from being unwisely overused. This extends the time until the first node powers down and increases the operation time before the network is partitioned. Investigations reveal that these two goals i.e. to use each node fairly and to extend their lifetimes are not compatible. A trade-
off between them is needed. C.K Toh proposed conditional max-min battery capacity routing scheme which chooses the shortest path if all nodes in all possible routes have sufficient battery capacity. When the battery capacity for some nodes goes below a predefined threshold (\(\gamma\)), routes going through these nodes will be avoided, and therefore the time until the first node power-down is extended. By adjusting the value of \(\gamma\), either the time when the first node powers down or the lifetime of most nodes in the network can be minimized.

Routing algorithm for network capacity maximization in energy constrained ad hoc network provided by K. Kar, M Codialam, T. V. Lakshman and L. Tassiulas [8]. C. Pandana and Ray Liu proposed the Keep Connect algorithm along with flow augmentation or with Minimum Total Energy algorithm and these combine algorithm provided maximum connectivity of the network as well as maximize the lifetime of network [9].

In [10] LEACH is given by W.R.Heinzelman, A.Chandrakasan and H. Balakrishnan. LEACH is a clustering-based protocol that utilizes randomized rotation of local cluster base stations (cluster-heads) to evenly distribute the energy load among the sensors in the network. LEACH uses localized coordination to enable scalability and robustness for dynamic networks, and incorporates data fusion into the routing protocol to reduce the amount of information that must be transmitted to the base station. H.Hsu and Q.Liang extend the work of LEACH [11]. They extend LEACH stochastic cluster-head selection algorithm by a deterministic component to reduce energy consumption.

In [12] Gil Zussman and Adrian Segal give an energy efficient routing in ad hoc disaster recovery networks. Their network model is based on the model for energy conserving routing in a wireless sensor network, presented by Chang and Tassiulas [2]. Formulated as an any cast routing problem, in which the objective is to maximize the time until the first battery drains-out. They present iterative algorithms for obtaining the optimal solution of the problem. Then, they derive an upper bound on the network lifetime for specific topologies and describe a polynomial algorithm for obtaining the optimal solution in such topologies. Finally, numerical results regarding the upper bound and the algorithm are presented.

Various energy conservation schemes in wireless sensor network discussed by G. Anastasi, M. Conti, M. D. Francesco and A. passarella. To reduce power consumption in wireless sensor network, they identified three main enabling techniques namely duty cycling, data-driven approach and mobility [13]. Distributed energy balanced routing is proposed in [14]. This routing algorithm uses the energy balance path for data transmission. It firstly calculates the total energy cost of all the paths from source node to base station and then select energy efficient path for data transmission.

II. ENERGY EFFICIENT ROUTING ALGORITHM

A key challenge in ad hoc wireless sensor network is achieving a long lifetime of nodes that carry limited amount of battery energy. It could be impossible or inconvenient to recharge the battery in the remote location therefore, the crucial requirement is to prolong the network life time. Need to Prolong the WSN Lifetime is to devices are generally battery powered, devices may be embedded inside structures, Failure of some devices may result in the failure of entire network, sensor nodes cooperatively perform a single task so they must be alive for same amount of time, sensor nodes are use for application where they monitor particular region so they must go down in a fashion that the overall task may be accomplish.

In wireless sensor networks the main objective is to maximize the minimum lifetime of each node. Lifetime is maximized by balancing the energy consumption of each node, using energy efficient routing. To maximize the objective function, it is appropriate for an emergency network in which every node is critical.

So based on the above algorithm, various energy efficient routing algorithms maximize the lifetime of network. These routing algorithms are

Energy Efficient Routing Algorithm for Single Destination

Flow redirection algorithm (FR)

Definition: Largest longest length path (Lp): largest longest length path is the path in which has largest capacity in terms of battery power and have less energy consumption per bit transmission than all other nodes in network. Smallest longest length path (Sp): smallest longest length path is the path in which has minimum capacity in terms of battery power and have higher energy consumption per bit transmission than all other nodes in network.

Flow redirection algorithm (FR) is the redirection based algorithm where some amount of flow is redirected from smallest longest length path (Sp) to largest longest length path (Lp). In this algorithm firstly Determine the two paths from node i to the destination which are to be involved in redirection second step calculate the amount of redirection (\(\$\)) and in the final step add \(\$\) amount of flow to the largest longest length path Lp. After adding \(\$\) amount of flow algorithm checks one or possibly more loops in the path and then loops are removed link by link along the path so in this step properly increment or decrement the flows of the two paths.
III. MAXIMUM RESIDUAL ENERGY PATH ROUTING (MREP)

In MREP algorithm let P be the set of all paths from node i to the destination node d. For a path p ∈ P, define the path length lp as a vector whose elements are the reciprocal of the residual energy for each link in the path after the route has been used by a unit flow. We assume that the routing path is calculated for each unit flow. The idea was to augment the flow on the path whose minimum residual energy after the flow augmentation will be longest.

Advantages

According to simulation result of [2] compare with Minimum Total Energy (MTE) on an average Flow Redirection (FR) and Maximum Residual Energy Path Routing (MREP) were both close to optimum, while Minimum Total Energy (MTE) [1] was not as good as the two. The ratio Rx of FR and MREP were over (Rx>0.9) of the true optimum in 84% and 89% of the case respectively, while that of MTE was true optimum in only 37% of the case and perform arbitrarily bad in worst case, simulation results were in favor of MREP and FR.

Limitations

As in [2] topology of the network is static. Hence the results are applicable to networks which are either static like sensor network or whose topology changes slowly enough, but not applicable for ad hoc network whose topology changes very frequently.

Energy Efficient Routing Algorithm for Multi-commodity case

Flow augmentation algorithm (FA)

The objective of this algorithm is to find the best link cost function which will lead to the maximization of the system lifetime so at each iteration, each origin nodes o ∈ Oc of commodity c calculates the shortest cost path to its destination nodes in Dc Then the flow is augmented by an amount of Qi c on the shortest cost path, where _ is the augmentation step size. After the flow augmentation, the shortest cost path is recalculated and the procedure are repeated until any node i ∈ N runs out of its initial total energy Ei. As a result of algorithm, we obtain the flow which will be used at each node to properly split incoming traffic.

Flow Redirection algorithm (FR)

This flow redirection algorithm is define for multi-commodity case where each node i belongs to set of all nodes N – Dc, for each commodity. For node i firstly select two paths and redirect the flow from smallest longest length path Sp to largest longest length path Lp. So need to assign more flows to Lp path as compared to any other path. After that, algorithm calculate amount of redirection Ei(c). In final step add amount of redirection in to the longest length path, after adding flow redirection amount algorithm checks one or possibly more loops from the path and then loop or loops are removed link by link along the path. So in final step algorithm can properly increment or decrement the flows of the two paths.

Advantages

FA algorithm iteratively identified traffic flow along the shortest cost path, which will lead to maximization of the system lifetime, as well as in flow redirection algorithm redirects some amount of flow from one path to another path and lifetime of later path is better than all other paths in the network belongs to same commodity so Flow Redirection algorithm is also lead for the maximization of the system lifetime.

Limitation

Proposed algorithms used fixed information generation rates and required a priori knowledge of future information generation, means FR and FA algorithms [3] provides optimal flow rates based on knowing the complete topology and packet generation rate at each node.

Online Power aware Routing Algorithms

Max-Min z-Pmin algorithm

This algorithm combines the benefit of selecting the path with the minimum power consumption and the path that maximize the minimal residual power of the node in the network. This algorithm first finds the minimum transmission energy path Pmin and then removes all edges whose residual energy fraction after use is smaller than or equal to the minimum residual energy fraction on the minimum transmission energy path. It then repeats the same procedure on the sub graph until just before the total transmission energy of the chosen path exceed z times Pmin, where z _ 1.

Zone Based Routing

Zone based routing organize the network structurally in geographical zones, and hierarchically to control routing across the zones. Each zone contains various nodes and treats the zones as an entity in the network and allows each zone to decide how to route a message across. In this algorithm each zone has a global controller node for message routing manages the zones and this node has highest power. Each zone uses the max-min z-Pmin algorithm to route a message within a zone.

Advantages

An online approximation power aware routing optimizes the life time of the network. The max-min z-Pmin algorithm combines the benefits of selecting the path with the minimum power consumption and the path that maximize the minimal residual power in the nodes of the networks. Whereas the zone based routing is to divide the network in to small number of zones and calculate the optimal path for each message across the zone as well as
computing the best path for the message within each zone.

Limitations

Proposed max-min z-Pmin algorithm requires information about the power level of each node in the network. In small network knowing this information accurately is not a problem but for large network it is difficult to aggregate and maintain this information. So it is very hard to implement max-min z-Pmin algorithm for large networks. In zone based routing each zone has many nodes and thus a lot of redundancy in routing a message through it. In the analysis of max-min z-Pmin algorithm, authors assume that the message are generated in constant rate means it is assumed message generated cyclically, or in each interval of time the set of message are same [4].

Algorithm for energy conservation in multi-commodity case with energy consumption at the receivers

Flow augmentation algorithm (FA)

The objective of this algorithm is to find the best link cost function which will lead to maximize the system lifetime and also consider the energy expenditure for unit data transmission at receiver end also. Each iteration every origin nodes o _ Oc of commodity c calculates the shortest cost path to its destination nodes in Dc Then the flow is augmented by an amount of _Qic on the shortest cost path, where _ is the augmentation step size. After the flow augmentation, the shortest cost paths are recalculated and the procedure are repeated until any node i _ N runs out of its initial total energy Ei. As a result of this algorithm, we obtain the flow which will be used at each node to properly split incoming traffic.

Advantages

In proposed flow augmentation (FA) algorithm which iteratively augments traffic flow along the shortest cost path which will lead to the maximization of the system lifetime and according to simulation result [4] in multi-commodity case the ratio Rx of commodity c calculates the shortest cost path to its destination nodes in Dc Then the flow is augmented by an amount of _Qic on the shortest cost path, where _ is the augmentation step size. After the flow augmentation, the shortest cost paths are recalculated and the procedure are repeated until any node i _ N runs out of its initial total energy Ei. As a result of this algorithm, we obtain the flow which will be used at each node to properly split incoming traffic.

Limitation

- In proposed work the performance of algorithm is depend on the value of _Qic_. If augmented step size _Qic_ becomes larger the performance deteriorated. So the larger _Qic_ means less frequent update on the routing information.
- Flow augmentation algorithm assumes that sensor have global information about the topology of networks to find the sufficient energy path, so this algorithm is not sufficient for large size networks and is not adaptive to dynamic network environment.

Energy Efficient Routing Algorithm with Delay Guarantee

Centralize energy efficient routing

The objective of this routing algorithm is to determine the optimal path from each sensor node to Access Point (AP) based on the topology of the network and the packet generation rates at the sensor node. In this algorithm, author propose a routing protocol for centralized implementation of the Linear Programming solution means in this protocol LP solution is decompose in to multiple routing tree. The algorithm decompose the network with optimal flow in to multiple routing trees then the nodes in the network will schedule using Time Division Multiple Access (TDMA) protocol such that all the packet reach the destination (AP) before deadline.

Distributed energy efficient routing

There are two distributed routing algorithms presented by the author namely least sum-cost path algorithm and least max-cost path algorithm, to be implemented for each period. The objective of these algorithms is to minimize the cost of routing paths from each sensor node to Access Point (AP). This algorithm calculate the shortest path from AP to all other nodes in the network where the cost of the path from one node to another is the sum and maximum of the costs of the links for least sum-cost algorithm and least max-cost algorithm respectively.

Level restricted energy efficient routing (LR-ENR) with delay guarantee

This routing algorithm aims to give delay guarantee on the arrival of packets at the Access Point (AP) while generating energy efficient path. This algorithm firstly execute Bellman-Ford algorithm to find minimum cost paths from each node to AP, then the routing algorithm is same as the distributed implementation with inclusion of a counter and tree construction packet, where counter is initialize to zero at the AP and increased by one at each transmission and this counter is contained by tree construction packet and this packet also contained nodes on routing path starting at the AP with node cost and cost of the transmission node in the routing tree.

Hop restricted energy efficient routing (HR-ENR) with delay guarantee

In this routing protocol, the time is divided into time frames. At the beginning of the frame, the AP floods a tree construction packet to in to the network which contains the counter, its routing path, node cost and the cost of the transmitting node in routing tree. Upon reception of a tree construction packet, the node checks whether the cost of the path is smaller than that of previously learned paths of the same
length. If, so it updates its minimum cost path for that length, and broadcast the packet. At the end of flooding each node knows about the minimum cost path of each length. They then send only the cost of the paths corresponding to each length, not the path themselves, to the Access Point (AP). The AP finds the optimal path length for each node based on the Integer Programming model, and sends it back to the nodes in the network. The nodes then use the routing path of the optimal length until the end of the frame.

Advantages
Energy Efficient Routing may choose path that are much longer than the shortest path to the Access Point while avoiding nodes with small residual energy and longer paths may prevent the system from meeting the delay guarantee but the proposed algorithm [6] Energy Efficient Routing with Delay Guarantee generate energy efficient paths and gives a delay guarantee on the arrival of packets at the Access Point.

Limitation
- According to simulation result [6] LR-ENR and HR-ENR required more memory and CPU cycles then centralize and distributed energy efficient routing algorithm. For algorithm implementation CPU requirement of LR-ENR and HR-ENR were \(O(\text{deg max } \_V, 2)\) and for centralize and distributed algorithm its \(O(\text{deg max } \_V, 3)\).

Algorithm for maximum connectivity and maximum lifetime of wireless sensor network

Keep connect algorithm (KC)
The keep connect algorithm finds the weight of node based on how many components are connected with this node. The weight of node can be thought as the importance of the node. Most important node is the node that results in large number of disconnected component as it dies.

Minimum Total Energy-keep connect (MTE-KC)
This algorithm firstly finds the minimum total energy path with edge cost. After that algorithm transmit the data on MTE path. If any node die because of low battery power, then algorithm recomputed the remaining nodes weight using keep connect algorithm and also recomputed the minimum total energy path using MTE.

Flow augmentation-keep connect (FA-KC)
Firstly in every update time find the minimum total energy path using flow augmentation algorithm. In second step if node dies, recomputed the remaining nodes weight using keep connect algorithm and also recomputed the energy efficient path using FA algorithm.

Advantages
The proposed KC algorithm along with flow augmentation or with Minimum Total Energy algorithm provide the best result such as these combine algorithm provide maximum connectivity of the network as well as maximize the lifetime of network.

Limitation
The main limitation of proposed algorithm is that when it work alone then maximum lifetime routing in network is not possible because KC alone can only calculate the weight of node based on how many connected component.

Distributed Energy Balance Routing (DEBR)
This algorithm balances the data traffic of networks in a decentralized manner and prolongs the life span of system. The routing algorithm uses a path that achieves energy balance for entire network. For energy balancing the minimum energy cost path has been select. For calculating energy cost it considers not only energy efficiency but also the amount of energy remaining in each node.

Advantages
- The maintenance of the energy is efficient for networks.
- The robustness to various event generation functions of this algorithm.

Limitations
Distributed energy balanced routing algorithm considers a general multi hop scenario where few nodes can communicate with base station. For large network this algorithm is not works properly, a more specific routing algorithm and problem definition is required for this kind of scenario.

IV. CONCLUSION

There are various challenges for deploying wireless sensor network like quality of service, type of service, scalability, programmability, maintainability, fault tolerance and lifetime of network. For increasing the lifetime of the network, the important thing is the efficient utilization of energy. So battery energy is the most important resource, and route the traffic through the minimum energy path to the destination is fatal for the network because all the nodes in that path will drain out their battery power quickly. So it’s not a feasible solution and instead of this solution route the traffic such that energy consumption is balanced among the nodes. In this paper, various energy efficient routing algorithm have reviewed to increase the lifetime of the network. The lifetime of wireless sensor network is defined as the time till the first node not fails. All the energy efficient routing algorithms discussed in this paper balance the energy consumption rates among the nodes in proportion to remaining energy. Finally paper concludes with approaching for research direction about various energy efficient routing protocols.
REFERENCES


PROPOSED A GLOBAL INSPECTOR IN WIRELESS COMMUNICATION TO SOLVE THE PROBLEM OF GLOBAL EAVESDROPPER

HARSHA C. KUNWAR¹ & RANJANA S. SHENDE²

¹,²Department of Computer science and Engg, G.H.R.I.E.T For Women’s, Nagpur, India

Abstract— Due to open nature of wireless sensor network it is relatively very easy for an attacker to eavesdrop and trace the packet movement in the network in order to capture the location of node physically. Such sensitive information can be trace by an adversary to derive the location of monitored object and data sink in the network. Existing scheme first formalizes location privacy issues and then proposes two techniques to provide location privacy to monitored object & two techniques to provide location privacy to data sink. After studying the adversary behavior, we present a counter measure to this problem. We propose a global inspector to preserve the privacy of packets. Global inspector will make use of Adhoc on-demand distance vector (AODV) Routing protocol to ensure security at the source as well as at sink node. This paper then performs traffic analysis to reduce the time and communication overhead based on throughput, jitter and delay. Through analysis and simulation, we demonstrate that the proposed technique are more efficient and effective for source and sink node in sensor network.

Keywords— Sensor network, Location privacy, Global Inspector.

I. INTRODUCTION

As the popularity and deployment of pervasive computing technologies grow, privacy of individuals is slowly steaming away. People are often grateful to exchange their privacy for small benefits and conveniences brought by the modern devices and neglect the consequences of potential privacy violations. So a responsible design of new technologies should take privacy risks into account. One of the new technologies posing a serious privacy risk is the wireless sensor network. Wireless sensor network A wireless sensor network (WSN) is a heterogeneous network composed of a large number of tiny low-cost devices, denoted as nodes, and one or few general purpose computing devices referred to as base stations (or sinks). A general purpose of the WSN is to monitor some physical phenomena (e.g., temperature, barometric pressure, light) inside an area of deployment. Nodes are equipped with a communication unit (e.g., radio transceiver), processing unit, battery and sensor(s). Nodes are constrained in processing power and energy, whereas the base stations have laptop capabilities and not severely energy resources [1]. The base stations usually act as gateways between the WSN and other networks (e.g., Internet). There is a wide variety of applications for WSNs [2], ranging from military applications (e.g., perimeter monitoring through environmental (e.g., animal habitat monitoring and health applications (e.g., patient health monitoring) to commercial applications (e.g., shopping habits monitoring, bridge structural health monitoring. WSNs can be classified according to several aspects with impact on the security protocol design. One such aspect is the mobility of nodes and the base station. The nodes can be mobile or placed on static positions. The same holds true for the base station. Another consideration is the way the nodes are placed. The nodes can be deployed manually on specific locations following some predefined network topology or randomly deployed in an area, e.g., by dropping from a plane. The number of nodes is also a very important factor – number of nodes in a network can range from tens to tens of thousands.

II. REVIEW LITERATURE

Location privacy need to be developed to prevent the adversary from determining the physical location of source sensors and sink. Due to limited energy lifetime of battery powered sensors-nodes, these method have to be energy efficient. Mehta et al proposed a technique source simulation, periodic collection at source node and sink simulation, backbone flooding at sink node to provide location privacy and also formalizes the location privacy issues under a global eavesdropper and estimate average communication overhead needed to achieve a given level of privacy by imposing Lightfoot et al proposed technique as the SinkToroidal Region (STaR) routing [6]. With this technique, the source node randomly selects an intermediate node within a designed Star area located around the SINK node. The Star area is large enough to make it unpractical for an adversary to monitor the entire region. This routing protocol ensures that the intermediate node is neither too close, nor too far from the SINK node in relations
to the entire network. STaR routing scheme can achieve excellent performance in energy consumption and delivery latency. Main limitation of this technique is message delivery ratio is slightly lower than the other schemes. Bamba et al described the Privacy Grid framework [5] that allows users the customization based on privacy requirements in terms of location hiding and QoS measures to control query processing overheads. Three dynamic grid-based spatial cloaking algorithms are developed for providing location k-anonymity and location l-diversity in a mobile environment. Experimental evaluation results reported and showed that compared to existing grid cloaking approaches, the dynamic grid cloaking algorithms provide much higher anonymization success rate and yet are highly efficient in terms of both time complexity and update cost. Kamat et al portrayed that Sensor networks can be deployed to monitor valuable assets. The author studied the ability of different routing protocols to obfuscate the location of a source sensor. To achieve improved location privacy, the author proposed a new family of routing techniques, called phantom routing, for both the flooding and single-path classes that enhance privacy protection. Phantom routing techniques are desirable since they only marginally increase communication overhead, while achieving significant privacy amplification. Ouyang et al proposed a new approach, Cyclic Entrapment, to lead adversaries into traffic loops in a sensor network. A comparison of CEM with existing methods shows that it can get a comparable source location protection while adding a comparatively low cost in terms of message latency and energy usage. As an advantage over existing techniques, it can protect a source’s location while allowing for an optimal routing time for messages from that source. However investigation of the impact of source mobility, multiple sources, and message rate from the source on this problem and our approach is not known. Deng et al addressed the issue securing a wireless sensor network against a variety of threats that can lead to the failure of the base station. First, multipath routing to multiple destination base stations is analyzed as a strategy to provide tolerance against individual base station attacks or sensor node compromises. Second, confusion of address and identification fields in packet headers via hashing functions is explored as a technique to help disguise the location of the base station from eavesdroppers. Third, relocation of the base station in the network topology is studied as a means of enhancing resiliency and mitigating the scope of damage. The author had extensively experimented with all three strategies both via simulation in ns2 and implementation on Berkeley MICA sensor motes. The results from these experiments show that a wireless sensor network can be secured quite well against attacks on base stations and compromises of sensor nodes.

III. SYSTEM ARCHITECTURE

The important part is to provide more security or authentication using Global Inspector algorithm.

A. Creation of Network

Create a network of more than ten nodes in a network. Each node having a capacity of sending a packet and receiving the same.

B. Imposing a GI in Network

- In a network select one node as a global inspector i.e GI which will authenticate that the packet is from trusted party.
- GI will make use of Adhoc on-demand distance vector routing i.e AODV technique to provide security at source as well as sink node.

C. Pass packet from trusted node

If packets travel from source to global inspector and from global inspector to destination node then destination node will come to know that the data is from party not from adversary.

D. Packets receive from untrusted source

- If any packet receive from untrusted party (i.e. not from GI) global eavesdropper.
When packet receives from untrusted node i.e. global eavesdropper that packet is dropped in a network.

IV. POSSIBLE CONTRIBUTION

So far to remove the drawback of existing technique we are implementing a GI i.e. global inspector and AODV routing technique to preserve the privacy at source as well as at the sink node. In GI i.e. global inspector algorithm if a packet is sent by the GI, then the nodes would accept the packet and use the same and if a packet is not sent by the GI, then the nodes would not accept the packet and drop it as it is not from a trusted source, this would ensure the privacy of the packets, and thus securing the source and sink node. The technique is shown by the following figure.

We are also performing traffic analysis based on various factor such throughput, jitter and delay to reduce the communication overhead at the source and destination node. The following graph shows the performance of technique.

V. DESIGN

Providing location privacy in wireless sensor network using global inspector i.e. GI is implemented in NS-2.34 environment installed on Fedora Operating System in VMware Workstation and is divided into various modules as fallows

A. Creation of wireless Environment and performing ping procedure module to perform verification of nodes.
B. Selection of global inspector in a network to define trusted node.
C. Verification of packets either is that from trusted source or not. If packet is from trusted source then process that packet.
D. Received packets are not from trusted source then drop packet instead of processing.

VI. CONCLUSION

Prior work on location privacy in sensor network assumed a global eavesdropper and provides two different techniques to protect source as well as two techniques to protect destination. We also presented techniques to provide the location privacy to object and sink against a global eavesdropper.

ACKNOWLEDGMENT

This work was supported by the US National Science Foundation under grant CNS-0916221. The author would like to thank the reviewers for their valuable comments.
REFERENCES


INTEGRATION OF SOUND SIGNATURE AND GRAPHICAL PASSWORD AUTHENTICATION SYSTEM

SUYOG S. NISCHAL¹, SACHIN GAIKWAD², KUNAL SINGH³ & A. DEVARE⁴

¹,²,³ Dnyanganga College of Engineering And Research, Pune
⁴ Computer Department, ZES’s DCOER, Pune

Abstract - A graphical password system with a supportive sound signature and video clip to enhance the security level in authentication system and it is cued click point based system. In this system password consist of graphical images in which user can select one cued click-point per image and video clip for authentication. Systems shared very good performance in term of speed, accuracy and enhance the security. User firstly need to specify the enter the ccp’s of image, in second step volume level and ternary allowed to play the video for secured login into account. In this system Cued-Click point to users receive immediate implicit feedback as to whether they are on the correct path of logging in. CCP offers both improved usability and security.

Keywords – Image Authentication, Sound signature, cued-click point, Video timing.

I. INTRODUCTION

Passwords are used for – (a) Authentication (Establishes that the user is who they say they are). (b) Authorization (The process used to decide if the authenticated person is allowed to access specific information or functions) and (c) Access Control (Restriction of access-includes authentication & authorization). Mostly user select password that is predictable. This happens with both graphical and text based passwords. Users tend to choose memorable password that are easier for hackers to guess the passwords. Number of graphical password systems has been developed; Study shows that a text-based password tends to lead inadequate security and usability problems. It is well known that the human brain is better at recognizing and recalling images than text, graphical passwords exploit this human characteristic that’s why in this system combination of ccp’s, volume level and video timing. A big necessity to have a strong authentication way is needed to secure all our login accounts as possible, so researches come out with advanced password called graphical password where they trying to improve the password techniques and avoid the weakness of normal password. Based on the two assumptions; first, humans can remember pictures better than alphanumeric characters and second, a picture worth a thousand passwords; some psychological studies and company software seem to agree with these assumptions. As known generally, the main drawbacks for the current graphical password schemes are the shoulder surfing problem and usability problem. Even though graphical passwords are easier to guess and break, but in this system combination of all three types i.e. ccp, volume level and video timing.

II. PREVIOUS WORK

Considerable work has been done in this area, the best known of these systems are Passfaces. Brostoff and Sasse (2000) carried out an empirical study of Passfaces, which illustrates well how a graphical password recognition system typically operates. Blonder-style passwords are based on cued recall. A user clicks on several previously chosen locations in a single image to log in. As implemented by Passlogix Corporation (Boroditsky, 2002), the user chooses several predefined regions in an image as his or her password. To log in the user has to click on the same regions. The problem with this scheme is that the number of predefined regions is small, perhaps a few dozens in a picture. The password may have to be up to 12 clicks for adequate security, again tedious for the user. Another problem of this system is the need for the predefined regions to be readily identifiable. In effect, this requires artificial, cartoon-like images rather than complex, real-world scenes. Cued Click Points (CCP) is a proposed alternative to Passpoints. In CCP, users click one point on each of 5 images rather than on five points on one image. It offers cued-recall and introduces visual cues that instantly alert valid users if they have made a mistake when entering their latest click-point (at which point they can cancel their attempt and retry from the beginning). It also makes attacks based on hotspot analysis more challenging. Each click results in showing a next-image, in effect leading users down a “path” as they click on their sequence of points. A wrong click leads down an incorrect path, with an explicit indication of authentication failure only after the final click. Users can choose their images only to the extent that their click-point dictates the next image. If they dislike the resulting images, they could create a new password involving different click-points to get different images.
III. PROPOSED WORK

In the proposed work we have integrated sound signature to help with the password. No system has been developed so far which uses sound signature and graphical password authentication. Study says that sound signature or tone can be used to add facts like images, text etc. Our idea is inspired by this novel human ability. Research says that human can remember images as well as sound tone easily; by applying this method we design our project so it will provide more security. Observed that all student who were registered entered their graphical password and video sound clip and it will be more secured from their point of view it is very good for Graphical and sound clip password authentication system.

IV. SYSTEM ARCHITECTURE

Incipient working: Firstly we need to enter the CCP of image. If entered CCP’s are correct then system will allows user for next level of logging. In next level user required to enter the volume level, if volume level is correct system will allows for next authentication level. In last stage of logging user need to enter correct video timing. If any of them (CCP’s, Volume level, Video timing) are incorrect then system will go in halt state for next 12 hours. After completion of 12 hours reboot again and user can try for uploading and downloading of data by entering correct password for all stages.

V. EXPERIMENTAL RESULTS

Data collected from 10 participants. Each participant was asked to register himself/herself and then each was invited to for login trail 5 times as legitimate user and 5 times as impostor randomly. Participants were final year engineering students of age group 20-24 Year. According to our survey we got instantaneous positive feedback and response.

VI. CONCLUSION

We have proposed a novel approach which uses sound signature and graphical password click points. Previously developed system never used this approach this system is helpful when user is logging after every single cycle. In future systems other patterns may be used for security purpose like touch of smells, video graphical click point, study shows that these patterns are very useful in secure login the associated objects like images, text and video clip.

REFERENCE


VII. AUTHORS

Suyog S. Nischal
Pursuing BE Computer (Final Year).

Kunal Singh
Pursuing BE Computer (Final Year).

Sachin B. Gaikwad
Pursuing BE Computer (Final Year).
Abstract—“Visual Java Editor” is complete user friendly program with we can create java program without using the Notepad and Dos BIOS. We just can create Java Program run and compile in a single window.

I. INTRODUCTION

Java visual editor is designed by java tools. We can create and save open java files by our software. We can compile and run java files by our project and view the result in our GUI window.

This application has limitations that in some aspects. It cannot be implemented in system supporting different operating systems than windows. It cannot process editor based compile and execute the java program and also make its text file is created by manual or other editor. Java file must compile by prompt.

This application can be added more functionality through create, open and save any java file make our editor tools implemented in effectively in windows based systems. Its improvement can be done regarding other operating systems like windows system etc. The applications can be made for system supporting more than one operating system. This application can also be made of process queries such as ‘java editor systems in first row”. Some more functionality such as compile and execute java file.

II. EASE OF USE

We are developing a Java-based component ware toolkit, with direct manipulation interface and repository component Builders, which supports the development of multiview applications. This consists of three components:

- **BuildByWire**, a direct manipulation editor for specifying the user interface components and basic layout behaviour for visual notations, together with editors for constructing and manipulating diagrams using the notations.

- **JViews**, a component ware toolkit for designing and implementing the data, behavior and intercomponent consistency management aspects of view and repository components. JViews view components are interfaced to BuildByWire components to form a multi-view, distributed editing system. JViews also provides support for data persistency, collaborative editing and version control.

- **JComposer** is a direct manipulation editor for specifying and generating JViews components (itself built with BuildByWire and JViews). JComposer allows BuildByWire components and third-party Java Beans to be combined with JViews components, resulting in multi-view, multi-user, customizable component ware solutions for a wide range of editor applications.

A. Project Scope

This project deals with implementation of Visual Java Editor. This project makes the user work in only one window to create any Java program and all the possible implementation that can happen in Java program.

B. Better performance:

The performance of each individual member will affect positively on the overall organization when the tasks requirements are clear and tasks are distributed in a balance manner. So employees will be able to produce a quality work and get the tasks done quickly in a correct manner from the first time since the tasks is clear and there are no room for misunderstanding or confusion.

Overall Performance:

If each and every team member knows exactly and understands prominently about the task to be accomplished, its relation to other tasks, and how it fits and contribute to the overall organizational performance, then defiantly his/her individual performance will improve which will also enhance the team performance that will contribute to the overall performance of the organization. This will have a positive effect on increasing the overall productivity and the quality of organization products or services, since fewer error and defaults will accrue.

a) Focus on core businesses and cut costs:

(Java Editor) will eliminate the need for excessive managerial effort in organizing and managing the work flow or in solving conflicts between team member, which will enable the organization to place more effort and focus on the core businesses, to save more dollars and employ them in a better way.

b) Customer satisfaction:

When an organization focuses on its core businesses and concentrate on its main competences,
it will increase the market share and become more competitive than other companies which will lead to gaining the most important factor that ensure long success of organization; customer satisfaction and off course that will be reflected in organization Profitability.

Assumptions and Dependencies
The following list presents the general constraints, assumptions, dependencies, or guidelines that are imposed upon the implementation of my project.

- All the system must have to connect with internet.
- The product must have a user-friendly interface that is simple enough to understand.
- Response time for compiling should take no longer time.
- A general knowledge of basic Java Language and Net beans is required to use the product.

1) Hardware Assumptions:
The server should have proper configuration hardware.

2) Software Assumptions:
J2SDK1.4, J2EE1.4 and Net Bean IDE are assumed to be installed on the system.

3) Operating System
WIN XP operating system is available.

III. TOOLS

A. Class Class
Instances of the class Class represent classes and interfaces in a running Java application. Every array also belongs to a class that is reflected as a Class object that is shared by all arrays with the same element type and number of dimensions. The primitive Java types (boolean, byte, char, short, int, long, float, and double), and the keyword void are also represented as Class objects.

Class has no public constructor. Instead Class objects are constructed automatically by the Java Virtual Machine as classes are loaded and by calls to the defineClass method in the class loader.

B. Class Process
The Runtime.exec methods create a native process and return an instance of a subclass of Process that can be used to control the process and obtain information about it. The class Process provides methods for performing input from the process, performing output to the process, waiting for the process to complete, checking the exit status of the process, and destroying (killing) the process.

The Runtime.exec methods may not work well for special processes on certain native platforms, such as native windowing processes, daemon processes, Win16/DOS processes on Microsoft Windows, or shell scripts. The created sub process does not have its own terminal or console. All its standard io (i.e. stdin, stdout, stderr) operations will be redirected to the parent process through three streams (Process.getOutputStream(), Process.getInputStream(), Process.getErrorStream()). The parent process uses these streams to feed input to and get output from the sub process. Because some native platforms only provide limited buffer size for standard input and output streams, failure to promptly write the input stream or read the output stream of the sub process may cause the sub process to block, and even deadlock.

The sub process is not killed when there are no more references to the Process object, but rather the sub process continues executing asynchronously. There is no requirement that a process represented by a Process object execute asynchronously or concurrently with respect to the Java process that owns the Process object.

C. Class Runtime
Every Java application has a single instance of class Runtime that allows the application to interface with the environment in which the application is running. The current runtime can be obtained from the getRuntime method.

An application cannot create its own instance of this class.

1) Source Editor
The Source Editor is a full-featured text editor that is integrated with the Form Editor, Projects window, compiler, and debugger. The Source Editor is automatically opened when a new and editable object is created from the available templates in the IDE. It can also be opened by double-clicking the node of an editable object in the Projects window.

2) Form Editor
The Form Editor is a visual tool for creating and editing GUI code written in the Java™ programming language. The Form Editor opens a design-time view of your GUI form in a Source Editor tab. The Form Editor also opens a Palette window of components you can add to your form, an Inspector window that shows the structure of your form, and a Properties window where you can modify the properties and behavior of each component.

3) Compiler
The Build menu contains commands for compiling applications. You can compile an entire project or compile single files. You can also access the most common compiling actions by using the commands in the Build toolbar or by using the keyboard equivalents.

4) Debugger
The IDE provides full integration of the JDPA debugger. When you start a debugging session, the IDE automatically opens a set of Debugger windows that display runtime information about your program. You can set breakpoints and watches, step through code, examine and modify variables, examine classes and the call stack, follow threads, and run multiple sessions.

5) Favorites Window
The Favorites window lets you access any location on your computer. This is handy for accessing files and directories that are outside of your project directories. The Favorites window does not know anything about project class path and membership, so none of the project-related commands like Compile File are available in the Favorites window.

IV. TECHNICAL SPECIFICATION

A. Advantages
- This project is easy to understand and user friendly for a fresher user.
- User can compile and interpret simultaneously in only one window.

B. Disadvantages
- This is whole dependent on java SDK. If it is not installed in the machine, it would not perform the operation.

C. Applications
- It can be used as multiple Java Programming System.
- Useful to various projects that can run on Java.

REFERENCES


Abstract—In the present system, users have to manually select the specific application or configuration utility for different settings like network, firewall, kernel modules, etc. However, in many Linux distributions, even this functionality is lacking a GUI front end. Hence, it becomes cumbersome for novice users to set up their settings through the terminal. Remembering and entering command-line options is a tedious task and many times errors can occur due to simple mistakes in commands or due to user’s misunderstanding of the extensive parameters available. Even for experienced users, while setting up a new machine or changing some configuration settings quickly and easily, there is a need for a better and more user-friendly GUI alternative to get the job done. Considering the above mentioned drawbacks of the present system, it is obvious that a centralized program with GUI will be helpful to all Linux users, novice and expert alike. A central place for all major settings that users perform will not only make configuration an easy task but will be helpful in saving time and other errors that may arise while doing it manually through command-line.

I. INTRODUCTION

The existing system does not support making profiles for different settings in the system. It requires certain settings to be done. Every time a setting is changed, the previous setting is lost. The user has to remember commands for changing each setting. There is no way of saving which loadable kernel modules are in use. If the user wants to load a new module, he has to do it every time the machine is started. We propose a set of adaptive strategies that help in taking better GUI based model. The GUI enabled application helps user to make his own settings easily. The GUI will enable the user to change many settings such as network configuration, firewall configuration and Loadable Kernel Modules (LKM). Some of the features can be enabled or disabled depending on the profile. These different features can be saved and can be loaded easily. There will be user preferences given which will help user to make changes in system easily. Also, these profiles can be saved to a file which can then be sent to anyone anywhere to get the same settings at the other PC.

II. RELATED WORK

In Linux GUI is developed using qt designer that is present. [1]Networking related files and commands used are referred from red hat [2]. Linux had grown to be as stable and as reliable as a very powerful operating system. Linux kernel is ‘copylefted’ software, patented under the GNU GPL, and thus, nobody actually owns it. But more significantly, Linux is sheltered by the Open Source community.[3]

III. PROPOSED SOLUTION

Considering the above mentioned drawbacks of the present system, it is obvious that a centralized program with GUI will be helpful to all Linux users, novice and expert alike. A central place for all major settings that users perform will not only make configuration an easy task but will be helpful in saving time and other errors that may arise while doing it manually through command-line.

Now a day, it has become increasingly important to make software to save users’ time and efforts doing monotonous and unnecessarily time consuming tasks. This helps users concentrate on the real work at hand instead of spending time and effort setting up the system. In addition to this, a portable one-click solution that we propose will make it very easy for all users to have a consistent system experience even if they are using different machines or networks. Hence, it helps users work just as efficiently as they would at their homes or workplaces using their own personal computers. There will no longer be a need to remember one’s personal settings, work settings, or the array of commands that accompany them. Linux has been avoided by many a novice users because of its lack of user friendly UI. Our software helps to tackle this problem in the configuration department, which is very commonly used by many users.

Figure1: This explains the system design of LSM.
IV. FUNCTIONAL SPECIFICATION

To create Linux Settings Manager, a GUI based software which will provide a central settings management location for Linux users. It will host the most commonly required configuration options for immediate or later use. Profiles will allow multiple users as well as offer portability to the users. The system is aimed at all Linux users as an alternative to command-line arguments required to achieve the same goal. It will provide the following functions and features:
1) GUI for network, firewall, kernel modules settings
2) Ability to create different profiles for multiple users
3) Ability to store settings of each user in personal settings file
4) Easy loading of settings from settings files.
5) Instant application of all settings selected
6) Option to reset settings if needed
7) Support for multiple network adapters

V. CONCLUSION

This paper presents a practical implementation of a Linux Setting Manager.

ACKNOWLEDGMENT

First and foremost, I would like to thank my guide, Prof. Mrs S. M. Murge, for her guidance and support. I will forever remain grateful for the constant support and guidance extended by guide, in making this report. Through our many discussions, she helped us to form and solidify ideas. The invaluable discussions I had with her, the penetrating questions she has put to me and the constant motivation, has all led to the development of this project.

REFERENCES

SYSTEM DESIGN TO ACHIEVE SCALABLE VIDEO CODING

MANIKANTA GUDIPATI1, SATWIK NARLANKA2 & SANOOP MALLISSERY3

1,2,3Department of Information and Communication Technology
Manipal Institute of Technology, Manipal University, Manipal-576104, India

Abstract—Introduction of video conferencing and video tele-phony led to a significant change in mode of communication. Video streaming has also gained popularity over the past few years and is currently used in many application scenarios such as surveillance and monitoring, computer based education and training which require transmission of video over the internet satisfying real time constraints. Given the heterogeneity of the internet, guaranteeing specific bandwidth for real time video transmission over a network is improbable. Scalable Video Coding (SVC), the Annex G extension of the H.264/MPEG-4 AVC video compression standard aims at providing high video quality to users in heterogeneous multimedia environments by transmitting a scalable bit stream with one or more subset bit streams [1]. This paper discusses about system components of SVC and a decision taking algorithm which aids in creation of adaptable bit stream. This algorithm was implemented in an Adaptation Decision Taking Engine (ADTE) and was tested in an environment simulated using Joint Scalable Video Model (JSVM) for SVC. In addition, description tools part of MPEG-21 Digital Item Adaptation such as, Usage Environment Description (UED), Universal Constraints Description (UCD) and Adaptation Quality of Service (Adaptation QoS) were also used.

I. INTRODUCTION

During the past few years, an exponential increase is observed in creation and consumption of multi-media content over the internet[2]. In general, the internet is becoming content centric network and content aware network which not only enriches QoS but also enables a whole range of services especially those which include real time streaming over the internet. Online classroom portals, Doctor Consultancy services through video conferencing are examples of such services. The emergence of broadband networks can be considered as the major factor for the observed increasing demand of information in the form of videos over the internet. This section of the paper mainly focuses on the reasons which imply that usage of an adaptable bit stream is advantageous over conventional methods especially in the case of real time video transfer. Given the rapid advances in network technologies and increased demand in multimedia services, there are numerous challenges that are to be considered especially when transmitting a video with real time constraints. These so called challenges can be briefly categorized in to network specific constraints which degrade the Quality of Service or result in disruption of transfer and user specific quality requirements. Considering the case of wireless networks, high BER is noticed due to small scale and large scale fades. The resulting bit errors cause severe degradation in the quality of the video that is to be presented [3]. Also, video packets arriving with additional delay can result in the play process to pause. Another network specific factor that can impact the QoS is bandwidth fluctuation. There are several reasons such as distance from the base station, co-channel interference that might result in fluctuation of the bandwidth. Consider a case of handoff where the available unused radio may not be enough to accommodate a newly joined mobile host or the case when a mobile terminal moves from a WLAN to a wireless wide area network. The bandwidth variation in this case can be from few Mbps to a few Kbps. Consequences of the above mentioned cases are problematic during a real time video transmission over a wireless network. For a better understanding to why bit stream adaptation is required, a comparative unicast and multicast scenario is considered in this paper. Fig. 1. depicts a unicast scenario where five different bit streams are transferred over to the receiver. And Fig. 2. depicts a multicast scenario wherein a single bit stream is traversed over the network. The multicast approach is bandwidth efficient but the visual quality requirements of the user and processing power of the end device are not considered. Whereas in the case of unicast transfer, bandwidth consumption is high since five different components are to be traversed over the Networks.

Fig. 1. A Unicast Scenario

The need for a standard which supports adaptive bit stream has been noticed as early as 1992 [5]. This need was acknowl- edged with the specification of bi-directionally coded pictures of MPEG visual. Though
the term "scalability" was not used, a feature of Coarse Granular scalability- discarding without any effect on integrity of video [4]. Given the heterogeneity of the current networks, network and user specific constraints, it is impossible to determine the exact bit rate for optimal video delivery. Hence the current objective of video coding for real time video transfer has changed to optimizing the video quality over a bit rate range instead of single bit rate. One way of implementing this is to use an encoding technique that encodes videos in to several layers. Scalable Video Coding (SVC), the Annex G extension of the H.264/MPEG-4 AVC video compression standard aims at providing high video quality to users in heterogeneous multimedia environments by transmitting a scalable bit stream with one or more subset bit streams.

II. OVERVIEW OF SCALABLE VIDEO CODING

Bit streams can generally be classified in to single layered bit streams and scalable bit streams. Scalable bit streams are those bit streams which form valid bit streams for the target decoder even if the parts of bit stream are removed or missing. The bit streams which do not have this property are called single layered bit streams. Scalable video coding involves encoding of the bit stream in to a base layer and several enhancement layers [5]. Base layer contains the minimum data required and the purpose of the other layers is to enhance the video quality. It is to be noted that base layer streams encoding with Scalable Video technique can be decoded by H.264/AVC products such as set top boxes. The usual modes of scalability are Quality, Temporal and Spatial Scalabilities.

A. Types of Scalabilities

1) Quality Scalability: SNR Scalability, also referred as Quality Scalability is achieved by encoding the bit stream in to desired number of layers (two or more) with same spatiotemporal resolution but different video quality. The base layer provides basic quality and addition of enhancement layers compliments the quality. Since in this case addition of enhancement layers also said to enhance the SNR, this technique is prominently known as SNR scalability. Quality scalability can be benefitting when used in video applications of telecommunications such as video on demand services.

2) Spatial Scalability: This type of scalability involves encoding ensuring that the coded base layer provides basic spatial resolution. The enhancement layers can be coded at different spatial resolutions. For example, the base layer can use QCIF resolution (176x144) where as the enhancement layer can use CIF resolution (356x288). Dependency Identifier D is used to refer to the corresponding resolution of a particular layer. The Dependency identifier of base layer is 0 and the Dependency Identifier of enhancement layer increases by a value of 1 from one layer to the next [6].

3) Temporal Scalability: In this type of scalability, base layer is coded to provide basic frame rate and the enhancement layers are coded with temporal prediction with respect to the base layer. The layers can have same or different resolutions but when encoded, provide full temporal resolution at the decoder. The temporal layers can be identified by a temporal identifier T. The temporal identifier T starts from 0 i.e. for the base layer, T=0 and increases by 1 from one temporal layer to the next.
System Design to Achieve Scalable Video Coding

III. SUPPLEMENTAL ENHANCEMENT INFORMATION MESSAGES

An SEI NAL unit contains one or more SEI messages. SEI messages do not play any role during decoding of output pictures, but can assist in related processes such as picture output timing, rendering, error detection, concealment of errors, and resource reservation. It is necessary that a common understanding regarding the properties of a layered bit stream needs to be established during the session setup, using signaling protocols [7]. One of the main reasons for the introduction of SEI messages as a concept has been to allow system specifications, such as those defined by 3 GPP or DVB, to interpret the supplemental information identically, and hence interoperate. Session announcement and session negotiation are equally important, as they apply to different use scenarios. In current times, protocols and protocol extensions are currently under development to fulfill the above requirements for IP networks. At this point, it is safe to assume that the signaling will reuse few of the syntax structures and mechanisms defined in the SVC specification. During session setup, however, these SEI messages are not conveyed as part of a video session as the transport for the session has not been established. They are instead conveyed through the signaling protocols. SEI messages specified in SVC also relate to the indication of decoding dependency and scalability characteristics, which are important for SVC specific transport and system level operations. SVC specifies a set of SEI messages to signal the scalability features and to allow system components to conveniently access that information. These SEI messages include:

Scalability Information SEI message (SSEI) Sub picture Scalable Layer SEI message Layers not Present SEI message (LNP-SEI) Layer Dependency Change SEI message (LDC-SEI) Quality Layer Information SEI message Scalable Nesting SEI message A. Scalability Information SEI message

SSEI message includes most relevant properties of a scalable bit stream such as: Bit rate, Frame size, Frame rate. Information regarding the no. of scalable layers present following by information regarding each layer. Layer identifier, decoding dependency information i.e., temporal-id, dependency-id, quality-id Layered Dependency: Based on layered dependency, non required layers can be identified to avoid transmission. Other information that is either explicitly included, or referenced from the description of another layer. The layer identifier is assigned as follows. A larger value of layer identifier indicates a higher layer. A value 0 indicates the lowest layer. Decoding and presentation of a layer is independent of any higher layer but may be dependent on a lower layer. Therefore, the lowest layer can be decoded and presented independently, decoding and presentation of layer 1 may be dependent on layer 0, decoding and presentation of layer 2 may be dependent on layers 0 and 1, and so on[8]. B. Sub-picture Layer SEI message A sub-picture based layer consists of[9]-sub-pictures where each sub picture is a subset of coded slice of a coded picture. Sub-picture scalable layer information SEI message can be used to map the corresponding NAL units. Sub-Picture based layer flag: Indicates if the current layer is sub-picture based layer.

C. Scalable Nesting SEI message

Scalable Nesting SEI message contains more than one SEI message This SEI message indicates the scope of layer representations to which SEI messages apply. Also enables the reuse of SEI message syntax defined for H.264/AVC SEI messages. For Scalable Nesting SEI message, applies-to-all-layers-in-access-unit field defines what values to be considered. If the applies-to-all-layers-in-access-unit flag is set, temporal/dependency/quality-id are not used.

IV. DESIGN AND IMPLEMENTATION OF DECISION ENGINE

A. Description of the Usage Environment Usage Environment Description(UED), Usage Constraints Description(UCD), Adaptation Quality of Service(Adaptation QoS) are description tools part of MPEG-21 Digital Item Adaptation(DIA) are used to characterize scalable bit streams and usage environments. UED describes a usage environment in terms of terminal and network characteristics, natural environment properties, and user preferences [10]. For instance, extracted UED from an end-user makes it possible to express that the device of that end-user is a net book, having support for SVC and connected to a network with an average bandwidth of 400 Kbps. Further, Additional constraints which help in describing the usage environment such as resolution of the end-device, supported frame rate, priorityID of the scalable bit-stream are specified in UCD. Finally, AdaptationQoS provides information in great detail which concerns the influence of scalability on QoS and the way a scalable bit-stream can be adapted.[11]

B. AdaptationQoS In Adaptation QoS, Bit rate, spatial scalability, temporal scalability, and priority id are described for each extraction point. Fig. 5 shows an example of a simplified Adaptation- QoS description. As for spatial and temporal scalability, the scalability values described in the actual Network Abstraction Layer (NAL) unit header can be obtained by subtracting the scalability values described in the AdaptationQoS description from the maximum scalability values used in the bit-stream. Therefore, contrary to the scalability values described in a NAL unit header, the scalability values described in AdaptationQoS represent a higher quality (higher resolution, higher frame rate) as they approach 0. Further, it is important to know that the perceptual
video quality increases as the value of priority id increases.

- <description xsi:type="AdaptationDescription">
  - <attribute xsi:type="AdaptationDescription">
    - <Constraint id="Normal bitrate">
      - <vector xsi:type="Vector">
        <value xsi:type="Integer">150 135 120 108 96 84 72 60</value>
      </vector>
    </Constraint>
    - <vector xsi:type="Vector">
      <value xsi:type="Integer">400 300 200 100 50</value>
    </vector>
  </description>
  - <adapterOperator xmlns="SpatialLayer">
    - <vector xsi:type="Vector">
      <value xsi:type="Integer">0 1 2 3</value>
    </vector>
  </adapterOperator>
  - <adapterOperator xmlns="TemporalLevel">
    - <vector xsi:type="Vector">
      <value xsi:type="Integer">0 1 2 3</value>
    </vector>
  </adapterOperator>
  - <adapterOperator xmlns="PriorityId">
    - <vector xsi:type="Vector">
      <value xsi:type="Integer">0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40</value>
    </vector>
  </adapterOperator>
</configurationDescription>

C. UCD
Fig. 6 shows an example of a simplified UCD. Spatial resolution, frame rate and the value of priority id are contents of UCD. In particular, the device of a particular end-user is restricted to CIF resolution, a frame rate of 30 Hz, and a maximum value for priority id of 40.

D. Adaptation decision-taking for SVC
A scenario depicting working of an ADTE is considered in [12] and is shown in Fig.7. Bit-streams compliant with SVC can be adapted in various ways in order to take into account different network conditions, device resolutions, and supported frame rates. Fig.7 demonstrates the use of an ADTE. An ADTE is capable of determining an optimal adaptation method, using information stored in the different MPEG-21 descriptions about the characteristics of a usage environment and the properties of a scalable bit-stream. The ADTE consists of an XML parser and an adaptation decision-taking capability by referring to UCD and UED, as illustrated in Fig.8. The XML parser extracts information about the usage environment and the scalable bit-stream from the MPEG-21 descriptors. Using this information, an adaptation decision is taken, which is then signaled to the extractor. The decision takes the form of scalability coordinates \([D, T, P]\), where D, T, and P denote spatial scalability, temporal scalability, and priority id, respectively. Finally, the extractor creates an adapted bit-stream, taking the scalability coordinates and the original bit-stream as an input. Since the ADTE uses information about the usage environment of each end-user, personalized consumption of multimedia content can be made possible.[13]
The system will consist of XML files with five variable parameters: Average Bandwidth, Preference, Resolution, Frame Rate, and SNR value. These values are received from the client in UCD and UED files.

### TABLE I
SUMMARY OF ENCODED SCALABLE BIT STREAM

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Frame Rate</th>
<th>Bit Rate</th>
<th>Min. BitRate</th>
<th>Y-PSNR</th>
<th>U-PSNR</th>
<th>Y-PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>176x144</td>
<td>1.875</td>
<td>61.5645</td>
<td>61.5645</td>
<td>35.3542</td>
<td>40.7845</td>
<td>41.477</td>
</tr>
<tr>
<td>176x144</td>
<td>3.750</td>
<td>85.9069</td>
<td>85.9069</td>
<td>52.977</td>
<td>40.569</td>
<td>41.756</td>
</tr>
<tr>
<td>176x144</td>
<td>7.500</td>
<td>105.1526</td>
<td>105.1526</td>
<td>32.8424</td>
<td>40.1248</td>
<td>41.680</td>
</tr>
<tr>
<td>176x144</td>
<td>15.000</td>
<td>124.3888</td>
<td>124.3888</td>
<td>32.3648</td>
<td>40.4054</td>
<td>41.535</td>
</tr>
<tr>
<td>352x288</td>
<td>1.875</td>
<td>283.3855</td>
<td>283.3855</td>
<td>57.2751</td>
<td>41.740</td>
<td>45.549</td>
</tr>
<tr>
<td>352x288</td>
<td>3.750</td>
<td>389.5216</td>
<td>389.5216</td>
<td>35.9198</td>
<td>41.312</td>
<td>45.242</td>
</tr>
<tr>
<td>352x288</td>
<td>7.500</td>
<td>492.5689</td>
<td>492.5689</td>
<td>34.8390</td>
<td>41.071</td>
<td>45.060</td>
</tr>
<tr>
<td>352x288</td>
<td>15.000</td>
<td>521.1558</td>
<td>521.1558</td>
<td>34.3389</td>
<td>40.945</td>
<td>45.194</td>
</tr>
<tr>
<td>352x288</td>
<td>30.000</td>
<td>754.4538</td>
<td>754.4538</td>
<td>35.5579</td>
<td>40.908</td>
<td>45.796</td>
</tr>
</tbody>
</table>

### V. RESULTS AND CONCLUSION

This paper informs the reader about salient features of Scalable Video Coding and demonstrates the advantages of SVC over conventional methods in the case of real time video transfer. This paper also focuses on structure of SEI NAL units and portrays the role played by these messages. This paper discusses about system components of SVC and ADTE which aids in creation of adaptable bit stream. The designed Adaptation Decision Taking Engine (ADTE) was tested in an environment simulated using Joint Scalable Video Model (JSVM) for SVC [14]. To verify the usability of the designed ADTE, a video BUSCIF30 was encoded using JSVM for SVC. Data regarding various outputs is shown in Table I. In addition, description tools part of MPEG-21 Digital Item Adaptation such as, Usage Environment Description (UED), Universal Constraints Description (UCD) and Adaptation Quality of Service (Adaptation QoS) were also used.

### REFERENCES


 PACKET BUFFERS DISTRIBUTION FOR HIGH-BANDWIDTH SWITCHES AND ROUTERS  
KALYANI TAMBE$^1$ & SHUBHANGI VAIKOLE$^2$  
$^{1,2}$Datta Meghe college of Engg., Airoli,Navi Mumbai 400708  

Abstract—High-speed routers rely on well-designed packet buffers that support multiple queues, provide large capacity and short response times. Some researchers suggested combined SRAM/DRAM hierarchical buffer architectures to meet these challenges. However, these architectures suffer from either large SRAM requirement or high time-complexity in the memory management. In this paper, we present scalable, efficient, and novel distributed packet buffer architecture. Two fundamental issues need to be addressed to make this architecture feasible: 1) how to minimize the overhead of an individual packet buffer; and 2) how to design scalable packet buffers using independent buffer subsystems. We address these issues by first designing an efficient compact buffer that reduces the SRAM size requirement by $(k_1 1)k$. Then, we introduce a feasible way of coordinating multiple subsystems with a load-balancing algorithm that maximizes the overall system performance. Both theoretical analysis and experimental results demonstrate that our load-balancing algorithm and the distributed packet buffer architecture can easily scale to meet the buffering needs of high bandwidth links and satisfy the requirements of scale and support for multiple queues.

I. INTRODUCTION  
The phenomenal growth of the Internet has been fueled by the rapid increase in the communication link bandwidth. Internet routers play a crucial role in sustaining this growth by being able to switch packets extremely fast to keep up with the growing bandwidth (line rate). This demands sophisticated packet switching and buffering techniques. Packet buffers need to be designed to support large capacity, multiple queues, and provide short response times. The router buffer sizing is still an open issue. The traditional rule of thumb for Internet routers states that the routers should be capable of buffering RTT$^R$ data, where RTT is a round-trip time for flows passing through the router, and R is the line rate. In the author claimed that the size of buffers in backbone routers can be made very small at the expense of a small loss in throughput. Focusing on the performance of individual TCP flows, the author claimed that the output/input capacity ratio at a network link largely determines the required buffer size. If the output/input capacity ratio is lower than one, the loss rate follows a power-law reduction with the buffer size and significant buffering is needed. Given everlasting controversy, nowadays, routers manufacturers still seem to favor the use of large buffers. For instance, the Cisco CRS-1 modular service card with a 40 Gbps line rate incorporates a 2 GB packet buffer memory per line card. In order to support fine-grained IP quality of service (QoS) requirements, nowadays, a packet buffer usually maintains thousands of queues. For example, the Juniper E-series router maintain as many as 64,000 queues. Given the increasing popularity of OpenFlow, a packet buffer that supports millions of queues is always desired. Furthermore, a packet buffer should be capable of sustaining continuous data streams for both ingress and egress. With the ever-increasing line rate, current available memory technologies, namely SRAM or DRAM alone cannot simultaneously satisfy these three requirements. This prompted researchers to suggest hybrid SRAM/DRAM (HSD) architecture with a single DRAM interleaved DRAMs or parallel DRAMs sandwiched between SRAMs. In this paper, we briefly review previous work on packet buffer architectures and present scalable and efficient hierarchical packet buffer architecture. This is our first attempt to combine the merits of two previously published packet buffer architectures. Consequently, the SRAM occupancy has been significantly reduced. By fully exploring the advantage of parallel DRAMs, we first propose a memory management algorithm (MMA) called Random RoundRobin (RRR). Thereafter, we devise a “traffic-aware” approach which aims to provide different services for different types of data streams. This approach further reduces the system overhead. Both mathematical analysis and simulation demonstrate that the proposed architecture together with its algorithm reduce the overall SRAM requirement significantly while providing guaranteed performance in terms of low time complexity, upper bounded drop rate, and uniform allocation of resources. In a simulation, the proposed architecture reduces the size of SRAM by more than 95 percent and the maximal delay is only us-level, when the traffic intensity is 76 percent. In Section we briefly review the related work from the literature. Whet show their architectural scalability limitations analytical we introduce the concept of traffic-aware approach in designing a packet buffer. We propose a new buffering architecture reducing the requirement of SRAM size in Section 5. Its performance is
studied. Further using it as a basic building block, we present distributed packet buffer architecture. The corresponding mathematical analysis is shown in, the simulations are carried out to verify these results.

II. BACKGROUND AND RELATED WORK

2.1 SRAM and DRAM Technology
Current SRAM and DRAM cannot individually meet the access time and capacity requirements of router buffers. While SRAM is fast enough with an access time of around 2.5 ns, its largest size is limited by current technologies to only a few MB. On the other hand, a DRAM can be built with large capacity, but the typical memory access time (i.e., TRC) is too large, around 40 ns. Over the last decade, the DRAM memory access time decreases by only 10 percent every 18 months. In contrast, as the line-rate increases by 100 percent every 18 months, DRAM will fall further behind in satisfying the requirements of high-speed buffers. Given a DRAM family, in order to keep the DRAM modules busy, we need to transfer a minimum size chunk it is also called as block in of data into it to effectively utilize the bandwidth provided by the DRAM module. Large memory access time of DRAM requires the system to read/write data from/to any memory address for at least TRC time units. According to our investigation, the current chunk size of DRAMs could range from 64 to 320 Bytes. However, given much higher price and smaller capacity of low latency DRAM products, nowadays, high latency DRAM products such as the DDR3 actually dominates the market making the typical chunk size become 320 Bytes.

2.2 Packet Buffer Architectures
Bridging the speed gap between the SRAM and the DRAM becomes a major challenge. This speed mismatch does not refer to the bandwidth but the access time and the concomitant access granularity. Due to the variable packet sizes that the IP protocol allows, it is common for packet processors to segment packets into fixed-size cells, to make them easier to manage and switch. A common choice for the cell size is 64 Bytes because it is the first power of two larger than the size of a minimum packet (i.e., 40 Bytes). Thus, a packet buffer should be able to access data at the granularity of a cell. This requirement however is not applicable to the DRAM. In a cell-based packet buffer where a chunk is much larger than a cell, the payload efficiency and the effective throughput of the entire system are dramatically reduced. This prompted researchers to suggest hybrid SRAM/DRAM (HSD) architecture. To conduct a quantitative analysis, a parameter called b was introduced to denote the ratio of access time between the DRAM and the SRAM. Accordingly, the access granularity of DRAM is b times that of the SRAM, i.e., b cells. This description is simple and straightforward. However, it becomes inadequate under some circumstances. First, the access time alone cannot determine the access granularity. The minimal access granularity has to be related to the other factors, such as the bandwidth, and the frequency. For example, given b equals to 10, the access time of the DRAM is 10 times that of the SRAM. According to the definition of b, the access granularity of the DRAM should be 10 cells. However, if the bandwidth of the SRAM is twice that of the DRAM, the access granularity of the DRAM becomes five cells. The odd situation also happens when a DRAM with shorter access time and higher bandwidth is introduced, where the access granularity of new DRAM depends on the product of both its bandwidth and the current access time. There is no guarantee that the speed mismatch is improved. Second, the definition of b encounters some troubles in modeling a complicated architecture that consists of multiple SRAM and DRAM devices. It becomes meaningless to compare the access time of individual memory devices. Third, the definition of b becomes inapplicable when the allowed minimal access granularity of SRAM is less than a cell. Given the same bandwidth, as long as the SRAM still adopts the cell-based access, the access granularity of DRAM has to be less than b cells, even if the access time of DRAM is indeed b times that of the SRAM. Being aware of these drawbacks, a new description of this problem was introduced in. Using the same parameter b, it directly refers to the chunk size of a single DRAM. Thus, the speed mismatch of SRAM and DRAM now changes into the size mismatch of cell and chunk. We use this definition in our latter statement and redescibe the previous work in this way for consistency. An additional challenge in designing packet buffers is that we need to maintain multiple queues, rather than just one single FIFO queue. Intuitively, dispatching and storing packets in multiple separate queues entail significant overhead for the memory management algorithms. In short, the fundamental problem in designing a packet buffer is to find an efficient way to bridge the gap of size between the cell and the chunk. It must introduce minimal overhead while satisfying the aforementioned requirements, viz., SRAM-level access time, DRAM-level storage capacity, and large-scale multiple queuing. Generally speaking, there are three ways of organizing a packet buffer, viz., the hybrid SRAM/DRAM architecture BW, b ¼ BW/TRC. Give fixed TRC and cell-size, b can be reduced by replacing a fast DRAM with multiple slower DRAMS, i.e., smaller value of BW for each. In this way, b could match the size of cell, thus the batch load is no longer needed. Each DRAM is now capable of accommodating cells independently. Accordingly, the MMA now needs to coordinate the data transfers between Q queues and kDRAMs, which can be formalized as a bipartite graph for maximum matching problem. Given the original b in a packet buffer with only one DRAM is b1, k should
be not less than \((b_1 \geq 64 \text{ Bytes})\), in order to provide an equivalent overall bandwidth. Shrimali and McKeown proposed a memory architecture with \((b_1 \geq 64 \text{ Bytes})\) interleaved DRAMs, based on a randomized algorithm, where each cell is written/read into/from the interleaved DRAM memories randomly, thus it seriously suffers from an out-of-sequence problem. Using \((b_1 \geq 64 \text{ Bytes})\) interleaved DRAMs as well, Garcia et al suggested that a per-queue Round-Robin dispatching scheme (i.e., if a cell is the \(i\)th cell in a queue, then it should be dispatched into DRAM \(j\), where \(j = i \mod k\)). This could guarantee a maximum matching when the system is equipped with a sufficient large tail buffer, typically twice the size of Nemo. Wang and Hamdi further suggested that the tail buffer can be implemented using \(k\) distributed SRAMs, which is easier to implement in practice. With per-queue Round-Robin, the out-of-sequence problem is solved. However, the high time complexity in finding the maximum matching remains. Given the fact that a single chunk could require \(O(b_1Q)\) iterations before finding a maximum matching, the author acknowledged that the time complexity in achieving the maximum matching could be \(O(b_1Q)\) in the worst case.

2.2.3 Parallel DRAM Architecture

Wang proposed a parallel hybrid SRAM/DRAM architecture with \((k \geq k_0 \geq 64 \text{ Bytes})\) DRAMs named PHSD. Compared with the interleaved architectures, the PHSD reduces the time complexity of MMA to \(O(kbP)\) by introducing \(k\) arbiters working in parallel. By further setting a limitation on the burst size of incoming traffic, \(O(kb1lnQ)\) size requirement of SRAM was derived. However, the size requirement of SRAM still accounts for \(O(kb1Q)\) in the worst case.

2.2.4 Other Approaches

Designing a general-purpose packet buffer is always a difficult task. In contrast, for specific applications where buffer behavior is predictable, the task can be greatly simplified. Kabra et al. introduced a parallel DRAM approach that is optimized for the fixed departure time. The architecture proposed in which differs from the reservation-based design in that the solution is based on aggregating packets into blocks so that the amount of book-keeping information in SRAM is minimized. Lin and Hamdi proposed an approximation algorithm, which serves the application of fairness queuing. The solution proposed in is a virtually pipelined memory architecture that aims to mimic the behavior of a single SRAM bank using multiple DRAM and SRAM banks. Such general memory architecture is applicable to both the packet buffering problem as well as the various network flow state implementation problems where same data potentially be read for several times. However, this architecture assumes perfect randomization in the arrival requests to the memory banks, the buffer will overflow and start dropping requests, which greatly degrades the system performance. To make the system robust to adversarial memory access patterns, additional content addressable memory devices are introduced, making the extended memory architecture too cumbersome for pure packet buffering applications.

III. THE SCALING LIMITATION OF PREVIOUS ALGORITHMS

Single DRAM packet buffer architectures like Nemo cannot meet the requirements of ever increasing bandwidth and storage capacity. A straightforward idea to improve its performance is to replace the single DRAM in the middle with multiple parallel DRAMs. These DRAMs share the same address bus, thus can be still regarded as a single DRAM from outside. On the other hand, a packet buffer architecture with multiple DRAM requires the use of slow DRAMs compared to the Nemo architecture, in order to maintain the same bandwidth. Memory vendors always seek to provide DRAM modules with higher bandwidth and larger storage capacity. In order to avoid the size mismatch between a chunk and a cell, it becomes less cost-effective to choose slow DRAMs or degrade with use fast DRAMs deliberately. To conduct a fair comparison, we must assume that all the architectures use the same raw materials in building a packet buffer. In other words, we assume that a packet buffer always consists of \(k\) identical DRAMs and the physical features of these DRAMs remain the same no matter whichever kind of packet buffer architecture is adopted. The chunk size of a single DRAM is always \(b\). We restrict the definition of \(b\), \(b\) always equals to \(b_1\) in the following description: We mainly focus on the scalability issue of previous algorithms; i.e., whether the overall bandwidth accounts for \(k_{BW}\), or not, the required size of SRAM, the time-complexity of MMA and the other system overheads.

3.1 Buffer Behaviors

When we carefully examine the hierarchical packet buffer architectures by using the aforementioned methodology, whether the HSD architecture, interleaved DRAMs, or parallel DRAM, they all rely on three parameters, \(k\), \(b\), and \(Q\). The required size of SRAM is always \(O(kbQ)\) to understand this phenomenon for its further study, we first examine the buffer behavior of previous hybrid SRAM/DRAM architectures and algorithms, especially in the Nemo and the PHSD architectures. As shown in Fig. 2, the DRAM structure in this extend version of Nemo is implemented as a composition of \(k\) DRAMs that simply provides a data bus of width \(k\) times that of a single DRAM data bus. Given a fixed chunk size of \(b\) for a single DRAM, Nemo increases the scale of
batch load by k times, which requires each of the Q queues to maintain kb-size of data. Whenever kb-size of data is accumulated in a queue, it will be written into k DRAMs through a mutual data bus. In this way, the size gap between cell and chunk is compensated. One major drawback of Nemo is that the first (kb - 1) size of data cannot depart from the queue until the last bit arrives. This increases the buffering requirement. Like the other independent address bus architectures, the PHSD adopts a distributed implementation that uniformly dispatches the traffic of each queue to the k DRAMs, respectively; i.e., if a cell is the ith cell in a queue, then it will be dispatched to DRAM j, where j = i mod k. By using so-called per-queue Round-Robin scheme, the previous single kb-size queues are replaced by k queues of size b each. As arestult, the system maintains 2kQ queues for both ingress and egress. Fig. 3 illustrates the organization of a single queue in PHSD. After carefully analyzing its buffer behavior, we make the following observations:

1. Given that only the first k b = 64 size of data arrives to a queue, the data could stay in the SRAMs infinitely. So the maximum SRAM size in PHSD is roughly equal to that of Nemo, when b is much larger than 64 Bytes.

2. The departure time of each b-sized chunk in PHSD relies on the corresponding traffic pattern. Given that only the first k b = 64 size of data increases, the data could stay in the SRAMs infinitely, rendering that the typical values of b for currently available DRAMs are around 64 to 320 Bytes size of data arrives in each of the multiple queues, it creates an unbalanced traffic allocation among the DRAMs, i.e., only the first DRAM receives b size of data.

3. Although the long-term output is balanced by the per-queue Round Robin scheme, short-term biased output still exists, e.g., the heads of currently active queues in the output may all be allocated to a particular DRAM. Accordingly, this DRAM becomes the bottleneck of entire system in egress.

4. Per-queue Round Robin increases the scale of queueing and maximum matching problem by k times making the corresponding MMA less efficient.

5. The same Round Robin sequence adopted by multiple queues may create synchronization effect that overwhelms the first DRAM in the initial stage.

6. Each DRAM maintains exactly Q queues and the entire system maintains kQ queues in total. In practice, maintaining so many dynamic queues on such a big scale is a real challenge. For queues bearing modest traffic, this incurs significant overhead, even though there is no practical reason to maintain queues among multiple DRAMs to support it. According to the aforementioned analysis, it seems that the per-queue Round-Robin does more harm than good, and both Nemo and PHSD fail to exploit the advantage of having parallel DRAMs. The requirement of SRAM size in the worst case is always 0k bQ.

3.2 Exponentially Increasing SRAM Size

Due to the sluggish advance in the access delay of DRAM, b will remain more or less constant, making k and Round robin algorithms where each queue follows the same Round Robin sequence, the RRR scheme randomly selects the starting DRAM of each queue and keeps updating it for every k rounds. It is not a truly random algorithm as well; because every kb-sized chunk is still processed in sequence internally. As demonstrated in Fig. 4, i, j, m, and n are random numbers that determine the dispatching sequence of a queue. Accordingly, log2k bits information has to be attached to the head of every kb-sized chunk for the sake of reordering it in egress. In this way, RRR avoids the synchronization effect while still guaranteeing the in-order processing and uniform allocation of kbsize data within each queue. Fig. 5 shows the architecture of a packet buffer where k independent address bus based DRAMs serve as the main storage sandwiched between two SRAMs. In the ingress, a bQ-size SRAM is maintained as tail cache. The Random Round Robin algorithm manages to uniformly dispatch b-sized chunks among the multiple DRAMs. Given the possibility that multiple chunks could be allocated to the same DRAM in a short period of time, intermediate FIFOs (i.e., inner-front-buffer) are introduced to hold these chunks temporarily so as to prevent unnecessary drops. At the egress we maintain a bQ-size SRAM serving as the head cache which is 1-k that of in Nemo.3 Whenever the scheduler requests data from a queue, its corresponding b-sized data (i.e., the first chunk of this queue in the order of FIFO) are fetched from the corresponding DRAM arbitrary. When multiple chunks need to be fetched out from the same DRAM (i.e., the heads of current active queues locate at the same DRAM chip), conflict happens. An arbiter is introduced to solve this problem. The internal structure of the arbiter is shown in Fig. 6. For each DRAM, the arbiter maintains a separated request list and k request lists in total. Whenever a new request arrives at the arbiter, it will first be identified which specific queue it belongs to. Then, its current RRR sequence will be derived according to the log2k bits information attached to the head of every kb-sized chunk. Referring to its round robin counter, the location of its head chunk and the corresponding request list can be determined. For every round, only the oldest request of each request list can be issued to the DRAMs. Since the queuing length of the request lists could be different, the delay that a request is satisfied varies. In the next section, we will prove that such kind of request queuing delay and the queuing delay in the inner-front-buffers can be upper bounded, given satisfying certain conditions. To prevent any blocking of data during the initial stage, as we start to write into the packet buffer, packets are written to the head cache first; so they are available immediately if
a queue is read. To accomplish this, the architecture in Fig. 5 has a direct-write path for packets from the writer, to be written directly into the head cache.

IV. PERFORMANCE ANALYSIS AND SIMULATION STUDY FOR COMPACT PACKET BUFFER

4.1 Mathematically Guaranteed Performance

We present a detailed analysis showing that the proposed compact buffer design provides a guaranteed performance in terms of an upper bounded delay and drop rate, when a small speedup factor is provided. Since the signal transferring delay between SRAMs and DRAMs depends on the design of specific circuit board, we omit it in this paper. This assumption applies for both mathematical analysis and simulation unless otherwise stated. Therefore, the overall delay of the proposed system in real-life has to be slightly higher than that of derived in this paper, e.g., less than 10TRC. As a tradeoff between storage and uniform allocation of incoming traffic, the RRR adopts a fast batch load scheme that may lead to a short-term unbalanced traffic allocation among DRAMs. However, for a specific single queue, the RRR guarantees the traffic allocation gap between any pair of DRAMs to be no more than b, as kb-size data of each queue is always uniformly distributed among k DRAMs. Given a finite depth of the inner-front buffers, it provides the system with a fixed time window that allows an unbalanced traffic to be temporarily buffered. Based on the worst case, we assume that the buffer time window is so small that none of the queues has the chance to spread their traffic among multiple DRAMs. In other words, if the queue does contribute some traffic in a small time window, it always contributes exactly one b-sized chunk to a single DRAM that results in the most unbalanced traffic allocation. Further considering the random sequence generated by the RRR, the proposed architecture can be modeled as k independently and identically distributed finite capacity M=D=1=n queuing systems, definition of the exponential series, we finally arrive at. Please refer to for detailed derivation process. In Table 1, we list the values of n needed to guarantee an steady state buffer (i.e., 10_9 loss probability). For example, with 90 percent traffic intensity, we choose n = 101b as a typical value of inner-front-buffer size to achieve a loss probability less than 10_9. Consequently, the corresponding response time is upper bounded by 1

4.2 Influence of Extra Delay

In a compact buffer, any request must be postponed for additional 202TRC to ensure in-order processing. Since such delay lasts for only a few microseconds, it will not decline the quality of service significantly. However, it may change the time complexity of MMA and affect the scalability of the entire compact buffer. For both ingress and egress, the dispatching of cells and requests requires nothing other than an O(1)-complexity mapping. However, in order to ensure the tail cache never underruns, a so-called lookahead request window is required by the ECQF-MMA, in order to predict the most critical queue and reduce the required SRAM size. According to the conclusions presented, the relationship between the sizes of head cache and the lookahead window, Given enough requests are accumulated; a feasible scheduling can be found which reduces the required size of head cache. Consequently, the time complexity of ECQF-MMA is strictly related to the size of lookahead window. Due to the additional delay, our compact buffer requires to increase the lookahead window by 202b. Apparently, when Q is much larger than 202, it affects the time complexity and the delay of the overall system very little. The overall time complexity of the ECQF-MMA still ranges from O(1) to O(Q) depending on the chosen size of lookahead window. Although many previous algorithms accept such Q_1b _ 1 lookahead windows by default, we must acknowledge that the ECQF-MMA itself may not scale when the number of queues is very large. Therefore, the actual size of head cache should be much larger than bQ when Q is quite large, e.g., (bQ/nQ). The RRR and its corresponding architecture we proposed in this paper do not reduce the time complexity of MMA, but reduce the size of SRAM to 1=k that of in Nemo.

4.3 Simulation Study and Results

We simulate a compact packet buffer with 10 DRAMs based on the system architecture shown in Fig. 5. The simulation is based on 64 Byte cells. Given the symmetric structure, only the simulation results of the ingress are illustrated for the sake of simplicity. We test the system performance with four different chunk sizes, viz., 64, 128, 320, and 640 Bytes, each of which consists from 1 to 10 cells, respectively. We define a time slot as the minimum working span where each DRAM is capable of processing exactly one cell. Since there are 10 DRAMs, at most 10 cells are generated in each time slot depending on the traffic intensity. The arrival process of individual cell submits to Poisson distribution in this paper unless other mentioned. Since we consider fixed size cells in the simulation, we generate only two kinds of traffic patterns: uniform traffic and unbalanced (hotspot) traffic. For the uniform traffic, incoming cells are uniformly distributed across all the Q queues. Analysis of real-life traces indicated that the top 10 percent of flows account for over 90 percent of the packets and the bytes transmitted. Therefore, for the unbalanced (hotspot) traffic, 90 percent of the cells are destined to only 10 percent of the queues. For both traffic patterns, there are 104 active queues. At the egress, we simulate two kinds of output behaviors: random
scheduling and round robin scheduling. Due to the similarity of the results, only the result under random
scheduling is revealed. Figs. 7a and 7b show the number of backlogged cells in inner-front-buffers
under different traffic patterns. We clearly observe that the average queue length is always less than 50,
while the single inner-front-buffer still attains its peak size of around 16. The results are almost the same
except that the system with uniform traffic suffers from more dramatic buildup during the initial period
(i.e., first 104 time slots). The simple reason is that the uniform traffic distorts the traffic intensity. In the
beginning of the simulation when the tail cache is empty, uniform traffic evenly allocates cells among
all queues creating less “dispatchable” chunks than the unbalanced traffic. Thus, it actually leads to a
lighter workload in the first 3 * 10^3 time slots. However, when the backlogged cells reside in the tail
cache are finally released, it creates a burst which causes a sudden buildup in the inner-front-buffers.
Because of space limitations, we only show the simulation results with the uniform traffic as it is the
most rigorous traffic pattern for the proposed compact buffer architecture. By further prolonging the
simulation to 109 time slots, we test the system performance under different configurations. With b = 1
1 cell, the tail cache is abolished. In Fig. 8a, we observe that the practical result is always well bounded
by the theoretical value. By doubling the size of b, the distortion effect appears. With low
traffic intensity injected to the system, its maximum delay could go beyond the theoretical estimation.
For example, with 30 percent traffic intensity and b = 2
2 cells, a maximal delay of 26 time slots is recorded
which is 18.2 percent higher than the theoretical value of 22 time slots. The reason is that the tail
buffer distorts the traffic pattern dramatically creating a lot of new bursts. Detailed traffic traces are shown
in Fig. 8b, that there are more red points than the black ones in the region of high traffic intensity, e.g.,
>40%. With the increasing traffic intensity and the growing size of b, the distortion effect has been
greatly weakened. In contrast, the system benefits from the increasing size of b that provides positive
reshaping of traffic. From both Figs. 8a and 8c, we clearly find that the practical system performance
under higher traffic intensity is always better than the theoretical value. With 70 percent traffic intensity and
b = 5
5 cells, a maximal delay of 132 time slots is recorded, while the theoretical value is 160 time slots.
Meanwhile, with 90 percent traffic intensity and b = 10
10 cells, a maximum delay of 839 time slots is recorded. In contrast to the theoretical value of 1,010
time slots, the practical performance gains 16.9
percent improvement.

V. DISTRIBUTED MEMORY HIERARCHY

Compared with the previous approaches, the fast
batchload scheme and the RRR algorithm reduce the
SRAM sizerequirement to l = k when a small amount
of speedup is provided. Moreover, it does not rely on
any traffic pattern, thus is capable of serving
anyapplication. However, it is still a traffic-agnostic
approach which requires O(bQ) size of SRAM.
The scaling limitation remains. To overcome this
limitation, we first consider a simplearchitecture.
packet buffer system consisting of three compact
packet buffers. Queues are mapped to the compact
packet buffers and this information is tracked in the
queue table. A dispatching module located between
the queue table and packet buffers delivers cells
according to the tags. A FIFO queue (i.e., outer-front-
buffer) in front of each compact buffer deals with
short-term bursty traffic, and further forms a
subsystem together with this compact buffer. At the
outlet, we must clarify a few important issues. First, a
compact packet buffer is injected with a maximum of
90 percent traffic intensity, in order to fulfill the
speedup requirement. In other words, the output
speed of an outerfront- buffer never exceeds 90
percent of the capacity of a single compact buffer.
Second, a compact buffer is capable of forwarding
data in cells instead of chunks, because of its internal
SRAM/DRAM hybrid architecture and the
corresponding MMA. It provides a short access time
in the ingress, however suffers from relatively large
delay in avoiding the out-of-sequence
problem. As we have proved that such delay can be
upper bounded, a compact buffer can be simply
regarded as an array of multiple FIFO queues which
is able to process individual cells with a constant
delay. Accordingly, a physical-queue is referred to a
FIFO queue of a compact buffer and there are Q
physical queues in a single compact buffer. Third,
regarding a packet buffer as a black box, a
logicalqueue must be established whenever a stream
of data (e.g., one flow or an aggregation of multiple
flows) needs to be processed in a manner of FIFO.
Given multiple compact buffers, we need to figure
out a feasible scheme to map a logical-queue to
single/multiple physical-queue(s). When a single
physical-queue cannot satisfy the buffering needs of a
logical-queue (e.g., the compact buffer where this
physical-queue located has no resource to spare), then
the packet distributor will map this logical-queue to
multiple physical-queues that belong to different
compact buffers. Similarly, mapping multiple active
logical-queues to the physical queues located at the
same compact buffer may overwhelm this compact
buffer. The packet distributor implements a suitable
load-balancing scheme that keeps track of the
information of each compact buffer, including the
utilization of storage and bandwidth, and the number
of active physical queues. Using this information, we
need to devise a load-balancing algorithm that can
figure out the best configuration of the queue table
and queue mapping which uses the smallest number
of physical queues to support the largest number of
logical queues.

International Conference on Computer Science and Information Technology (ICCSIT), 3rd Feb-2013, Nagpur, ISBN: 978-93-82208-60-0
28
In order to achieve the above, we first build a basic framework that allows logical queues to dynamically switch from one physical queue to another without any blocking. Unlike the simple linked-list-based scheme in, in our system, any logical queue can be mapped to:

1. A single physical-queue: we refer to such logical queues as "small" logical queue.
2. K physical queues that allocate in K subsystems, respectively, (K denotes the total number of subsystems in a distributed packet buffer): we refer to such logical queues as "large" logical queue.

This distinction for a logical queue can be applied both at the ingress (distributor) and egress (aggregator), thus leading to combinations of states. For example, a logical queue is in the state of "large-small" if it is served by all subsystems in ingress and served by only a single subsystem in egress. Fig. 10 shows the state machine we have defined. Although there could be thousands of combinations, we only reserve six critical states. They are "unallocated," "large," "small," and three intermediate states "largesmall," "small-large," and "large-small-large." Any logical queue can switch its state between "small" and "large" smoothly with certain constrains. Meanwhile, it is strictly controlled that any logical queue can only possess no more than three serving states at any time, i.e., at most two turning points. This helps the system minimize the overhead of state maintenance. Based on the state machine above, we devise a load-balancing algorithm. The algorithm is naturally separated into three tasks that are implemented at the distributor, compact packet buffer subsystem, and the aggregator, respectively. The tasks communicate with each other through the centralized queue table. Here, are some typical behaviors of the load-balancing algorithm. Whenever the first cell of a new logical queue arrives, the distributor maps it to a subsystem that is currently the lightest loaded. For this new logical queue, the destination subsystem reserves an empty physical queue and updates the queue table and changes the state of this logical queue from "unallocated" to "small." For a single logical queue, the state of "small" could last for quite a long time. When the logical queue becomes empty, the state of logical queue is changed back to "unallocated." If a subsystem is temporarily overloaded, (i.e., the backlogged cells residing at the outer-front-buffer is beyond a threshold) the subsystem can divert the newly arriving cells to other subsystems. The diverting can be achieved by randomly changing the state of any new arrival cells to "large-small" if it is originally served as "small." To be more precise, the subsystem will still accept any new cells. But if the new arrival cells belongs to a logical queue which is originally served by this subsystem only (i.e., its state is "small"), the subsystem will mark the cell, and update the queue table by changing its corresponding state from "small" to "large-small." At the ingress, if the serving state of a logical queue is changed to "large," the cells of this logical queue will be dispatched to all subsystems in a per-flow round-robin manner. In this fashion, given a distributed system consisted with K subsystems, \(\delta K \_ 1 \_ 1 \_ 0 \_ K\) of the traffic of this "small" logical queue can immediately be diverted to other subsystems which helps to relieve the burden of the overloaded subsystem. As the output continues, the "large-small" logical queue will be updated to "large" logical queue when the previously marked cell is fetched. Now this large logical queue can be absorbed by lightly loaded subsystems. Generally speaking, it is a reverse process of the diversion described above where an intermediate state called "small-large" is introduced.

To be more comprehensive, there is one additional state called "large-small-large" in our state machine. The purpose of introducing this state is to prevent any subsystem from absorbing logical queues that bear considerable amount of traffic that cannot be easily handled by a single subsystem. With the state of "large-small-large," it provides us with an alternative choice that any logical queue can be diverted immediately at any time so as to prevent overwhelming of a single subsystem only.

VI. MATHEMATICAL ANALYSIS OF DISTRIBUTED PACKET BUFFER

In this section, we present a detailed explanation about the parameters of our load-balancing algorithm, such as MaxDivertTimes, THRESHOLD and DelayFactor. In particular, how these parameters affect system performance. It is difficult to estimate the traffic of a logical queue in real life. Our load-balancing algorithm introduces a probabilistic method which distinguishes the type of logical queue dynamically. Assume a logical queue bears \(P_0 < P_1 \_ 1 \_ 0\) unit of traffic and subsystem is capable of serving \(1 \_ 1 \_ 0\) unit of traffic at most (after considering the speedup), this logical queue has a probability of \(P\) to be diverted when it is served as "small." As the queue state changes dynamically, any long-lasting active logical queue is finally served as a large logical queue. Because the logical queue which has been diverted for no less than MaxDivertTimes cannot be absorbed any longer according to the load-balancing algorithm. Our key observation is that a logical queue bearing more traffic has higher probability to be diverted. Hence, the expected time that a logical queue achieves its final state could vary greatly, regarded as 81 percent traffic intensity, when the speedup inside a compact buffer is taken into consideration. Meanwhile, in order to observe the dynamic behavior of the entire system, the simulations are always separated into three phases. Assume the simulation lasts for \(X\) time slots. For the first \(0 \_ 2 \_ 0\) time slots, there is only input without output where cells are backlogged. In this...
way, we can create an initial backlog and also simulate the situation when the congestion happens. After 0:2; X \( \geq 1 \) time slots, a full-speed output (i.e., 90 percent when the internal speedup of a compact buffer is considered) based on page 5 random scheduling (i.e., every logical queue has the same probability to be fetched out) begins while input maintains, which represents the normal conditions. With backlogged cells in the first phase, we can monitor the system performance in detail, especially how the load-balancing algorithm behaves. After 0:5; X time slots, the input stops while only the output maintains fetching any backlogged cells. In this way, we can simulate the situation when the system is lightly loaded. Moreover, we choose the parallel system (i.e., PHSD) as the basic reference standard of our distributed system. Because it is the best parallel architecture we know so far which represents the previous “traffic-agnostic” approaches. To make it a fair comparison, the PHSD architecture also consists of 40 DRAMs. Accordingly, a logical queue is always mapped to 40 physical queues which require 2b size of SRAM for each. Thus, a logical queue inside the PHSD requires 80b size of SRAM in total. In contrast, a physical queue inside the distributed buffer requires 2b size of SRAM. It can be mapped to at most four physical queues simultaneously. In the first part of our simulations, we inject the distributed packet buffer with only 100 active data streams to have a clear observation on the load-balancing algorithm. The input traffic pattern is unbalanced (hotspot) traffic that top 10 percent active data streams account for 90 percent of the cells. In other words, a large data stream contributes 81 times as many cells as that of a small data stream. We arbitrarily selected two data streams (one large, one small) to have a close-up view of their states. As illustrated in Fig. 13, the queue states of these two data streams change frequently. We observe that the large data stream is mapped to multiple physical queues most of the time (i.e., the queue state finally remains as “large” after MaxDivertTimes changes of queue state), while the small data stream is served by a single physical queue mostly. The probability method works. Further increasing the number of active data streams to 4,105, we simulate the system performance under both uniform traffic and unbalanced (hotspot) traffic (i.e., top 10 percent active data streams account for 90 percent of the overall traffic). Fig. 14 presents the results of total number of active queues under different situations. The simulation lasts for 107 time slots. As shown in the figures, the parallel system (i.e., PHSD in [101]) always dispatches the logical queues in per-queue round-robin introducing a lot of overheads. The total number of active physical queues always achieves 1:6; 106 during the initial periods regardless of the intensity of injected traffic. In contrast, our distributed system which much less active physical queues. Taking an unbalanced traffic pattern and 78 percent traffic intensity as an example, in the first phase, the distributed scheme only maintain around 4:8; 104 active physical queues which is about 3 percent that of PHSD. As the second phase starts, the number of active physical queues for both architectures drops greatly. However, the distributed one still outperforms the PHSD where 1:2; 104 and 1:7; 105 physical-queues are maintained, respectively. As the further decreasing of traffic intensities, our distributed system shows more obvious advantages. Fig. 15 presents the comparison of SRAM occupancy between both architectures. Given a modest traffic intensity (i.e., \(<76\%\)), the distributed architecture reduces the SRAM occupancy by more than 95 percent no matter what the traffic pattern is. Even with a maximal 81 percent of the traffic intensity, still, the distributed architecture requires \(<9:5\%\) of the SRAM than the PHSD during its peak time. With 63 percent traffic intensity, the distributed architecture requires only 2.6 percent the overall size of SRAM.

Since our algorithm does not refer to the physical queues assignment status in balancing, we are curious about the actual distribution of active queues. Fig. 16 shows the actual distribution of active physical queues among four subsystems where 4,104 data streams are generated. We observe that the unbalanced distribution is always trivial for both traffic patterns. We also monitor the length of outer-front-buffers. The distributed system performs almost the same as the PHSD, in terms of the FIFO lengths, which are much less than 101 for both architectures.

VII. IMPLEMENTATION ISSUES

In this section, we discuss some implementation concerns and their possible solutions. First, high-speed dispatchers and aggregators are very difficult to extend to large scale. The distributed architecture enables us to build them in a hierarchical way that each level only handles limited number of physical ports. Second, a centralized queue table has been widely used in a packet buffer to track the status of individual logical queues, including the memory address of current head/tail of a physical queue, the per-queue round robin counters for both ingress and egress, and, etc. With the increasing number of logical queues and the overall throughput, a centralized queue table could become the bottleneck of an entire system. The proposed distributed packet buffer architecture adopts a hierarchical structure which reduces the number of physical queues significantly leading to a much smaller queue table. Taking the top level queue table as an example, it records the mapping information of only four compact buffers instead of 40 DRAMs. To further optimize the system, we propose a pipelined query scheme which allows the queue table to be
implemented by using off-chip memories. Fig. 17 is a simple demonstration. Instead of keeping only one FIFO, the system maintains two FIFOs which are used to store the data and control information of cells, respectively. Whenever a cell arrives, it will be directly dispatched to the data FIFO while its corresponding query is issued to the queue table immediately. As soon as the queue table returns the result, the control information will be stored in the control FIFO. Finally, a cell tagged with the destination information can be derived by fetching data from both FIFOs, before being forwarding to a dispatcher. On the other hand, being aware of the data consistency problem introduced by the pipelined query, the system must provide a small fully index table to store the most recent updates. Fig. 18 demonstrates how the value of counters can be increased without disturbing the pipelined query. The depth of these two FIFOs and the entries number of a little and can be implemented inside a chip. Besides, we also notice that majority of the queue table updates rarely. For example, the mapping information of a single logical queue can be updated for at most MaxDivertTimes times in its entire lifetime. Accordingly, multiple read-only copies of the queue table can be maintained simultaneously to maximize the throughput. Meanwhile, a queue table can be decoupled into multiple subtables. For example, the RRin and RRout can be maintained by ingress and egress, respectively. These schemes could further increase the throughput of a queue table.

VIII. CONCLUSIONS

Building packet buffers based on a hybrid SRAM/DRAM architecture while introducing minimum overhead is the major issue discussed in this paper. To distinctly increase the throughput and storage capacity of a packet buffer, a parallel mechanism using multiple DRAM chips should be deployed. Our analysis shows that previous algorithms make very little effects in exploring the advantage of parallel DRAMs leading to the requirement of large size SRAM and high time complexity in memory management. In this paper, we present a novel packet buffer architecture by using both fast batch load scheme and a hierarchal distributed structure. It reduces the requirement of SRAM size greatly. Both mathematical analysis and simulation results indicate that the proposed architecture provides guaranteed performance in terms of the low time complexity, short access delay, and upper bounded drop rate, when small speedup is provided. fully index table rely on the round trip time of a query, typically less than 30. Therefore, they cost just

REFERENCES

LOAD OPTIMIZATION ON VIDEO ON DEMAND SERVER

GANESH B. INGLE

Tolani Maritime Institute, Pune, Maharashtra, India

Abstract: Theme of this paper is to find the optimum solution for load assignment to the Video on demand Source. The algorithm checks and investigates the different channel allocation for video transmission. The channel allocation and higher resource utilization can be achieved by Genetic algorithm (GA efficiently).

Index Terms : Video On Demand Server (VOD Server), Dynamic bandwidth allocation, Genetic algorithm (GA), Quality of service (QOS), makespan, Average Utilization.

I. INTRODUCTION

Video on Demand (VOD) or Audio and Video on Demand (AVOD) are systems which allow users to select and watch/listen to video or audio content on demand. IPTV technology is often used to bring video on demand to televisions and personal computers[1]. The job of the video server is to deliver isochronous (one second of content per second of real time) compressed digital video to the correct port (the one connected to the downstream modulator which can reach the customer who has made the request), and to address that video data stream so that the customer's STB can recognize and retrieve it. It also communicates with the Conditional Access system to set up encryption for that stream, so that only that customer can receive and decode it [2]. Different types of video on demand request are handled by video server like EVOD (Exclusive Video on Demand), IVOD (Impulse video on demand), QVOD (Quasi Video on Demand), TVOD (Transactional Video on Demand), FVOD (Free Video on Demand) [3]. During peak hours the load increased on the server can cause problems like low quality of service, data loss. The load optimization techniques has crucial importance for the increase in performance of the server.

II. LITERATURE REVIEW

There are many algorithms and techniques for the dynamic channel allocation. Gary Chan and Pengye in their work discussed optimal band width allocation using simulated annealing by proving the given problem is NP hard, but they have not provided the bandwidth optimization for Server (IEEE VOL 13 No 2 April 2011). Kevin C. Almeroth and Mostafa H. Ammar in their work discussed use of multicast delivery to provide a scalable and interactive video on demand service. In their work they have not mentioned transfer of video on heterogeneous network (IEEE VOL 14 No 6 August 2011). Cyrus C.Y. and Mounir hamid in their work discussed a scalable video on demand systems using multiple batch buffering techniques, but they have not mention the the delay caused in the the data loss. To optimize a load distribution to the server Genetic algorithm is the best solution.

A **genetic algorithm (GA)** is a method for solving both constrained and unconstrained optimization problems based on a natural selection process that mimics biological evolution. The algorithm repeatedly modifies a population of individual solutions. At each step, the genetic algorithm randomly selects individuals from the current population and uses them as parents to produce the children for the next generation. It involves following steps. After an initial population is randomly generated the GA based on following steps

- **Selection** which equates to survival of the fittest;
- **Crossover** which represents mating between individuals;
- **Mutation** which introduces random modifications.

**a. Selection**
- key idea: give preference to better individuals, allowing them to pass on their genes to the next generation.
- The goodness of each individual depends on its fitness.
- Fitness may be determined by an objective function or by a subjective judgment.

**b. Crossover**
- Prime distinguished factor of GA from other optimization techniques
- Two individuals are chosen from the population using the selection operator
- A crossover site along the bit strings is randomly chosen
- The values of the two strings are exchanged up to this point
- If S1=000000 and s2=111111 and the crossover point is 2 then S1'=110000 and s2'=001111
- The two new offspring created from this mating are put into the next generation of the population
• By recombining portions of good individuals, this process is likely to create even better individuals
  c. Mutation
  • With some low probability, a portion of the new individuals will have some of their bits flipped.
  • Its purpose is to maintain diversity within the population and inhibit premature convergence.
  • Mutation alone induces a random walk through the search space
  • Mutation and selection (without crossover) create a parallel, noise-tolerant, hill-climbing algorithms
  
We can apply the genetic algorithm to solve problems that are not well suited for standard optimization algorithms, including problems in which the objective function is discontinuous, no differentiable, stochastic, or highly nonlinear.

The basic genetic algorithm works as follows

Formulate Initial population
Randomly initialize population
repeated
evaluate objective function
find fitness function
apply genetic operators

✓ Reproduction

The following fig illustrates the GA

![Genetic Algorithm Illustration](Image)

Fig 1. Genetic Algorithm Illustration

III. LOAD ALLOCATION BY VIDEO SERVER DYNAMICALLY

Unlike the static channel, dynamic channel allocation is efficient and is used in areas where the traffic is no uniform and heavy. Five main assumptions must be considered in dynamic channel allocation.

1. Station Model:
   Stations are also called terminals. The number of independent stations are N, with independent constant arrival rates lambda, and probability of a frame being generated in a time interval of (delta t) is (delta t x lambda). Once a frame has been generated the station does nothing until the frame has successfully been transmitted.

2. Single Channel Assumption:
   From hardware point of view, all stations are equal. A single channel is available for communication on which all stations can transmit on it and all can receive from it.

3. Collision Assumption:
   If two frames are transmitted simultaneously, they will collide resulting in a false signal. Each station has the ability to detect collision and it must be kept in mind that collided frame must be retransmitted later.

DCA takes advantages different wireless networks. During peak hours numbers of users are more where as in low hours numbers of users are less.

1. During peak hours 97% user are connected to the networks.
2. During Low hours 30% users are connected to the networks.
3. Even in the peak hours not all the users are downloading or uploading the videos.
4. There are gaps in the packet to be sent on the network. That gap can be filled up with other data.

The Load allocation depends on the priority of transmission and type of data. In a dynamic assignment method, in contrast, all channels are potentially available to all cells and are assigned to cells dynamically as calls arrive. If this is done right, it can take advantage of temporary changes in the spatial and temporal distribution of channel in order to serve more users. For example, when packets are concentrated in a few cells, these cells can be assigned more channels without increasing the blocking rate in the lightly used cells. If a bandwidth is more then the client can take more bandwidth to achieve highest quality of service. Every multimedia data is stream further partitioned into layers of different priorities. Layers of lowest priorities are discarded till minimum quality of service is achieved. All the clients will get minimum bandwidth insuring that minimum bandwidth is not exceeding the total bandwidth limit. It means that most of the bandwidth goes unused.

![Bandwidth assignment over the wireless network](Image)

There are different layers of band provided such as audio and, video band which further classified into low priority and high priority band. Low priority band provides low quality video streaming. High priority band provides high quality video streaming.
IV. PROBLEM FORMULATION USING GENETIC ALGORITHM

Assuming distributed network, client requested for the video to the server. Initially if sum of all the request are not exceeding the total band with limit the channel will be allocated on first come first serve basis. As the number of demand for the channel allocation increase and surpasses the total bandwidth then the bandwidth allocation depends on the priority of the demand from the client.

Let there be some video servers and the task A consists of (A1,A2,..Ai) tasks. Now the challenge is to map n tasks to m possible servers. So challenge is to find best server selection strategy to improve the performance metrics. Many facts that affects the QOS are server load, response time, Make span (is the largest completion time of all the tasks in the server).

\[
\text{Average Utilization (AU)} = \frac{1}{\text{Makespan}} \text{ number of processors} \quad (1)
\]

\[
\text{Fitness value} = \frac{\text{Cumulative quality}}{\text{number of processors}} \quad (2)
\]

To get the optimized result perform crossover operation until fitness value > 0.3

V. PROPOSED GENETIC ALGORITHM

The proposed system is used to apply the to obtain optimum bandwidth for MDC. There for population Po. and comparison set S is created along with the comparison set to calibrate individual reproduced through crossover and mutation give optimum solution.

a. Steps Involved for proposed algorithm
1. Identify the video servers and sources
2. Enlist number of clients and tasks (A1,A2...Ai)
3. Derive parameter of initial request
4. Create initial population with list requests
5. Calculate makespan
6. Calculate Average utilization
7. Derive the fitness value for quality parameter
8. If sum of number of user bandwidth < total allocated Load perform step 9 otherwise step 10.
9. Assign Load on first come first serve basis.
10. Generate comparison set S from the population
11. While fitness difference <Precision (0.3)
    Repeat step 12 to 16 else goto 17
12. Accept more request
13. Compare new request value with those of S.
14. Verify bandwidth can be reallocated to client so that all the request will get satisfied. If yes go to step 10.
15. Select two random string from the population.
16. Perform cross over and mutation operation perform step 12.
17. Stopping criteria.

VI. DERIVATION FORMULATION

Suppose N number of video sources each jth source is encoded into Li layer of video stream. The layers are base layers and L2,L3,L4,...,Li are enhanced layers. Let Pik and bik, be quality and bit rate of kth layers of jth source. K=1,2,...Li

Cumulative quality, Qij and bit rate, Sij are represented by equation.

\[
q_{ij} = \sum_{k=1}^{k=i} P_{ik} \quad j = 1,2,3,...,Li \quad (1)
\]

\[
s_{ij} = \sum_{k=1}^{k=i} P_{ik} \quad j = 1,2,3,...,Li \quad (2)
\]

The quality of the jth source with L layers is Vj={q1,j,s1,j,q2,j,s2,j,...,(q3,j,s3,j)}
Where i=1,2,3,...L
The string is before cross over
Prev[0][0]=13 16 10 14 10 10
Prev[0][1]=14 4 22 4 7 3
Prev[0][2]=9 12 16 12 7 19
Prev[0][3]=17 14 1 23 125 23
Prev[1][0]=16 19 3 12 5 14
Prev[1][1]=14 14 1 19 19 9
Prev[1][2]=8 9 12 9 11 19
Prev[1][3]=12 14 15 12 0 2

VIII. GENERATING POPULATION

First step is to generate population of feasible solution of size P. For each chromosome K, which is video source to produce random set of bit of each sub string to select one layer of every source V. The fitness values are calculated. Each chromosome is checked if it is identical to the old or existing chromosome population. Identical binary strings are discarded so that the chromosome population should have unique solutions. Once unique chromosome set is generated we select two parent for generating offspring. Which represent video source. Now we will partition the initial population S into sub population T1 and T2. Fittest individual is selected from each population. Random portioning is better than always selecting the two parents from the given population.

IX. CROSSOVER

Instead of randomly exchanging genetic information between two parents even odd mask segmentation is used. Every segment represents video source. Two basic modes of operation are odd and even. In odd mode offspring inherits the properties from odd segment from one parent p1 and even segment for other parent p2 and vice versa. For N=5

That is 5 is a video source. m1 is applied to p1 by performing and operation to extract the odd segment and m2 is applied to derive even segment. Offspring’s are reproduced be recombining the resultant chromosome. Alternating these modes ensures the algorithm is not going to stuck up.

After cross over the strings are
after[0][0]=13 16 10 14 10 10
after[0][1]=14 4 22 4 7 3
after[0][2]=9 12 16 12 7 19
after[0][3]=12 14 15 12 0 2
after[1][0]=16 19 3 12 5 14
after[1][1]=147 14 1 19 19 9
after[1][2]=8 9 12 9 11 19
after[1][3]=17 14 1 23 15 23

X. MUTATION

Mutation is applied to number of bits in binary string of video source. The probability of inverting each bit depends on mutation rate. Mutation rate is inversely proportional to selective pressure that is if mutation rate is high the selective pressure is low and vice versa.

The strings after mutation are
after[0][0]=13 16 10 14 10 10
after[0][1]=14 17 22 4 7 3
after[0][2]=9 12 16 12 7 19
after[0][3]=12 14 15 12 0 2
after[1][0]=16 19 3 12 5 14
after[1][1]=16 19 3 12 5 14

XI. COMPARING FCFS AND GA ON THE BASIS OF LOAD DISTRIBUTION ON SERVER

The following result illustrates the comparative result of GA vs FCFS.

<table>
<thead>
<tr>
<th></th>
<th>Average CPU Utilization of video server</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA</td>
<td></td>
</tr>
<tr>
<td>FCFS</td>
<td></td>
</tr>
</tbody>
</table>

![Fig. 4 GA vs FCFS average CPU utilization](image)

![Fig. 4 GA vs FCFS on makespan](image)

XI. CONCLUSION

The purpose of paper is to improve the load distribution for the video on demand server. The load distribution can be assigned properly using genetic algorithm. The future scope is to highest optimize the load distribution in to the heterogeneous networks and distributed networks as well.
REFERENCES

[1] Pengye Xia and S.H. Gary Chan ,“Optimal Bandwidth assignment for Multiple-Description-coded Video “. IEEE, VOL.13, No2 April 2011


[5] Y.F. Chen Huang R. Jana H. Jiang ,M. Rabinovich “When is p2p technology is beneficial for IPTV services”


BIOMETRIC AUTHENTICATION BASED ON KEYSTROKE DYNAMICS USING NEURAL NETWORK.

SUSHANT DAGARE¹ & VINAYAK JADHAV²

¹²Department of Computer Science and Engineering, SSPM College of Engineering, Sindhudurg, Maharashtra, India

Abstract— As methods of authentication changes day by day computing and validating authentication parameters becomes critical issue. Prior keystroke dynamics validates user identity by their keyword latency. In this paper we proposed system which uses keyword latency and neural network algorithm for validating authorized use and decrease false rate.

Keywords— Keystrokes, backpropagation, latency

I. INTRODUCTION (KEYSTROKE DYNAMIC AND NEURAL NETWORK APPROACH)

Authentication is the process of determining whether a user should be allowed access to a particular system or resource. It is a critical area of security research and practice. Alphanumeric passwords are used widely for authentication, but other methods are also available today, including biometrics and smart cards. However, there are problems of these alternative technologies [1]. Biometrics raise main concerns such as acceptability and lack of flexibility and smart cards usually need a Personal Identification Number (PIN) because cards can fall in the hand of imposters. As a result, passwords are still dominant and are expected to continue to remain so for some time [1]. Traditional measures such as passwords and PINs need more advanced safeguards against unauthorized access to information and computer resources [19]. One such safeguard is keystroke dynamics. This method analyzes the way a user types at a terminal by monitoring the keyboard inputs and aims to identify users based on certain habitual typing rhythm patterns [19]. Keystroke dynamics is a biometric based on the assumption that different people type in uniquely characteristic manners. Observation of telegraph operators in the 19th century revealed personally distinctive patterns when keying messages over telegraph lines, and telegraph operators could recognize each other based on only their keying dynamics [6]. Keystroke dynamics is known with a few different names: keyboard dynamics, keystroke analysis, typing biometrics and typing rhythms [6]. Currently users begin information to computer systems via physical keyboards or keyboards on touch screens. The main advantage of using keystroke dynamics is that it can be used without any additional hardware. Thus it is inexpensive. The user acceptance of a keystroke dynamics biometric system is very high, since it is not intrusive and users do not necessarily even notice that such a system is used [6, 16]. The reset of this paper is organized as follows: The first part of paper shows how key dynamics works for user authentication and second part states how we can reduce errors or false rate by back propagation algorithm , a short review of the previous studies is then presented; this is followed by theoretical concepts. The proposed algorithm is presented. This is followed by experiments and discussion of results. Finally, conclusions and suggestions for further research are given.

II. LITERATURE REVIEW:

Compared to other biometrics, keystroke biometrics has additional desirable properties due to its user-friendliness Of late, keystroke dynamics has become an active and non-intrusiveness. Keystroke dynamics data can be research area due to the increasing importance of cyber collected without a user’s cooperation or even awareness. Security and computer or network access control. Most of the existing approaches focus on static verification, where and evaluation criteria across the studies. To address this a user types specific pre-enrolled string, e.g., a password issue, keystroke dynamics databases including benchmark during a login process, and then their keystroke features results of popular keystroke biometrics algorithms have are analyzed for authentication purposes [24]. Only a few have been published [17][19 ] to provide a standard research studies address the more challenging g problem of experimental platform for progress assessment. Kill our hy keystroke biometrics using “free text”, where the users can and Maxi on collected and published a keystroke dynamics type arbitrary text as input [22][26]. benchmark dataset containing 51 subjects with 400 keystroke dynamics collected for each subject [17]. Keystroke dynamics features are usually extracted using Furthermore, they evaluated fourteen available keystroke the timing information of the key down/hold/up events ,dynamics algorithms on this dataset, including Neural. The hold time or dwell time of individual keys, and the Networks [4], K-means [12], Fuzzy Logic [11], KNNs, latency between two keys, i.e., the time interval between Outlier Elimination [11], SVMs
[19], etc. Various distance the release of a key and the pressing of the next key are metrics, including Euclidean distance [3], Manhattan typically exploited. Digraphs, which are the time latencies distance [1][16] and Mahalanobis distance [3][4] were between two successive keystrokes, are commonly used. This keystroke dataset along with the evaluation Trigraphs, which are the time latencies between every methodology and state of the art performance provides a three consecutive keys, and similarly, n-graphs, have been benchmark to objectively gauge the progresses of new investigated as well.

III. KEYSTROKE DYNAMICS CLASSIFIER’S:

We frame keystroke dynamics based authentication as a one-class classification problem which learns a model for a user, rejects anomalies to the learned model as imposters, and accept inliers as the genuine user. Although the use of negative examples in training could significantly improve the accuracy of the classifier, it is unrealistic to assume prior knowledge about the keystroke features from imposters, let alone the availability of their training data.

We used the Nearest Neighbor classifier with the new distance metric defined in to either as certain keystroke dynamics feature as originating from the genuine user when the distance to its nearest neighbor in the training data is below a threshold value, or reject it as an imposter, otherwise. The covariance matrix is computed using all the training keystroke feature vectors from the intended user. The adoption of the new distance metric helps suppress the adverse effects of outliers during the classification stage. However, outliers could still corrupt the training data and deteriorate the authentication performance.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Equal-error rate</th>
<th>Algorithm</th>
<th>Zero-miss false-alarm rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nearest Neighbor</td>
<td>0.084 (0.056)</td>
<td>Nearest Neighbor</td>
<td>0.405 (0.268)</td>
</tr>
<tr>
<td>(new distance metric) + outlier removal</td>
<td></td>
<td>(new distance metric) + outlier removal</td>
<td></td>
</tr>
<tr>
<td>Manhattan (scaled)</td>
<td>0.096 (0.069)</td>
<td>Manhattan (scaled)</td>
<td>0.482 (0.273)</td>
</tr>
</tbody>
</table>

IV. CLASSIFICATION AND IDENTIFICATION OF DATA:

The problem of recognizing a given pattern belonging to a particular person either after exhaustive search through a large database, or by simply comparing the pattern with a single authentication template can be formulated within the framework of statistical decision theory. By this approach one can convert the problem of pattern recognition into a much more expedient task, which involves the execution of tests of statistical independence. The approaches described in the following paragraphs adhere to this model. The classification technique employed by Joyce and Gupta [14] represents the mean reference signature for a given user as MD {Musername; Mpassword; Mfirstname; Mlastname}. Verification is performed by comparing the test signature T (acquired at login time) with M and determining the magnitude of difference between the two profiles. Given MD (.m1; m2;:::; mn) and TD.(t1:t2;:::; tn) where n is the total number of latencies in the signature, the verifier computes the magnitude of difference using an L1 norm. Positive identification is declared when this difference is within a threshold variability of the reference signature. The mean and standard deviation of the norms jj M−Si jj, where Si s one of the eight training signatures, are used to decide the threshold for an acceptable difference vector between a given T and M. Although these absolute verification rates are encouraging, Joyce and Gupta[14] tested using a replacement methodology, which means that the distribution of the training set is necessarily representative of the learning set. The use of separate data sets, recorded at different times, would be more reliable. Therefore, we investigated the performance of classifiers based on studies where users were allowed to participate in experiments conducted at varied times under no supervision. The reference profiles collected were represented as N-dimensional feature vectors and processed in a manner similar to that of [14]. The data was split into learning and testing sets.

**Part I: Creating Training Pattern for Neural Network:**

In this part we will first calculate training sets for neural network by calculating latency of characters in data (i.e. password).
Biometric authentication based on keystroke Dynamics using Neural Network.

The features extracted for formation of the pattern form the Features Vector possessing keystroke duration and keystroke latency. Where, Itp is the Keystroke Duration of the key (I), that is, time that the user leads for press and liberate the key (I). VIt1 is the Keystroke Latency between of the keys (V) and (I), that is, interval of time that the user leads for liberate the key (I) and press the key (V).

Features in Typing Patterns
Keystroke dynamics is not what you type, but how you type. Features commonly used to describe a user’s typing pattern are:
• Latencies between successive keystrokes (the elapsed time between the release of the first key and the depression of the second)
• Duration of each keystroke (How long is the key held down)
• Finger placement.
• Pressure applied on the keys
• Overall typing speed
• For known regularly-typed strings (e.g., username and password), such features are quite consistent
• However, features are a function of the user and the environment

Tri-graph Features:
• Three consecutively typed keys is called a tri-graph
• Duration of the tri-graph – time between the pressing of the first key and third key
• Consider that the user types the text: MYWORLD
• A sequence of tri-graphs and durations (msec) is

\[
\text{M(\text{mean of data in msec})} = \frac{\text{d(s1,1-s2,1)+ (s1,2-s2,2)+ (s1,3-s2,3)+ (s1,4-s2,4)+ (s1,5-s2,5)}}{6}
\]

Password Hardening:
• Text password + Typing pattern = Hardened password
• Encode the typing pattern along with the normal text password to create a longer password
• Advantage: More security
• Limitation: False rejects- if the user’s typing pattern varies significantly (e.g., due to change of keyboard), the user will not be able to login
• Challenge: attacker should not be able to find out what specific typing features (e.g., latency between the first and second keystroke) are used; otherwise the text password could be easily revealed

Part 2: Reducing false rate by Back propagation algorithm:

Above network have three layers (input layer, hidden layer and output layer).

1. In first pass we provide value of M calculated by tri-graph method.
2. Network will pass that value to output layer as training set.
3. In next pass test data which i.e. M_test (latency of user password) supplied to input layer.
4. Output layer compare difference between latency of training data and test data.
5. If test latency is nearer to training latency then again same test data with some weight W_j and bias j adjustment feed to network.
6. Process stop when test data exactly equal to training data value.

In this way we can classify input patterns with respect to user authentication.

Method [12]:

1) Initialize all weights and biases in network;
2) While term initiating condition is not satisfied {
3) For each training tuple X in D {
4) // propagate the inputs forward:
5) For each input layer unit j /
6) \( O_i = I_j; // \text{out of an input unit is its actual input value} \)
7) For each hidden or output layer unit j /
8) \( I_j = J^I W_{ij} O_i + 9_j; // \text{compute the net input of unit} j \text{ with respect to the previous layer} j. \)
9) \( O_i = \); // compute the output of each unit j
11) For each unit j in the output layer
12) \( E_{\text{Errj}} = O_j (1 - O_j) (T_j - O_j); \) // compute the error

13) For each unit \( j \) in the hidden layers, from the last to the first hidden layer

14) \( E_{\text{Errj}} = O_j (1 - O_j) \cdot \text{Lk}\cdot \text{Err} \cdot w_{jk}; \) // compute the error with respect to the next higher layer, \( k \)

15) For each weight \( w_{ij} \) in network

16) \( A_{w_{ij}} = (l) E_{\text{Errj}} O_i; \) // weight increment

17) \( W_{ij} = w_{ij} + A_{w_{ij}}; \) // weight update

18) For each bias \( B_j \) in network

19) \( A B_j = (l) E_{\text{Errj}} B_j; \) // bias increment

20) \( B_j = B_j + A B_j; \) // bias update)}

VI. APPLICATIONS :-

Keystroke dynamics has many applications in the computer security arena. One area where the use of an adaptive approach to keystroke dynamics may be particularly appealing is in restricting root level access to the master server hosting a Kerberos [21] key database. Any user accessing the server is prompted to type a few words of a pass phrase in conjunction with his/her username and password. Access is granted if his/her typing pattern matches within a reasonable threshold of the claimed identity. This safeguard is effective as there is usually no remote access allowed to the server, and the only entry point is via console login. Alternatively, dynamic or continuous monitoring of the interaction of users while accessing highly restricted documents or executing tasks in environments where the user must be “alert” at all times (for example air traffic control), is an ideal scenario for the application of a keystroke authentication system. Keystroke dynamics may be used to detect uncharacteristic typing rhythm (brought on by drowsiness, fatigue etc.) in the user and notify third parties.

VII. FUTURE WORK

This project uses statistical analysis approach. Other methods that could be used are neural network and fuzzy logic to reduce false rate. Also, larger user trail could have been used for testing purposes. Other areas that could use as future work parameter are keyboard type, time of day and user behavior. By using JavaScript or ASP for better data capture method can be implemented.

REFERENCES


Biometric authentication based on keystroke Dynamics using Neural Network.


“ONLINE INTRUSION DETECTION AND INTRUSION ALERT AGGREGATION WITH GENERATIVE MODELING”

MANDA.B.KALKUMBE & SUJATA G TUPPAD


Abstract—Intrusion Detection System provides protective measures for the public and private networks. It guarantees information security. At present scenario, most IDS are quite reliable in detecting suspicious intrusion attacks, but when the intrusion actions caused by a single attack instance of particular type, often results in hundreds or even thousands of alerts instead of single alert. This makes ambiguity to network security engineer. The primary goal of this paper is to identify and to cluster different alerts belonging to a specific attack instance with the concept of alert aggregation. Alert aggregation is an important subtask of intrusion detection. The goal is to identify and to cluster different alerts—produced by low-level intrusion detection systems, firewalls, etc.—belonging to a specific attack instance which has been initiated by an attacker at a certain point in time. In this method, different types of alerts will be clustered into different groups called meta-alerts. Thus, meta-alerts can be generated for the clusters that contain all the relevant information whereas the amount of the alerts can be reduced substantially. Meta-alerts may then be the basis for reporting to security experts. Meta-alerts are generated with a delay of typically only a few seconds after observing the first alert belonging to a new attack instance.

Index Terms—Intrusion detection, alert aggregation, generative modeling, data stream algorithm.

I. INTRODUCTION

Intrusion detection system can be classified into two categories
- Host based IDS
- Network Based IDS

Host based IDS product protect an end system or network application by auditing system & event logs. Network based IDS product can be deployed on the network monitoring network traffic for attack. Intrusion detection is an important security tool. It has the possibility to provide valuable information about the current status of security. Intrusion detection system (IDS) monitors network traffic and monitors for suspicious activity & alert the system or network administrator. They help to defend against the various threats to which networks and hosts are exposed to by detecting the actions of attackers or attack tools in a network or host-based manner with misuse or anomaly detection techniques [1]. At present, most IDS are quite reliable in detecting suspicious actions by evaluating TCP/IP connections or log files, for instance. Once an IDS finds a suspicious action, it immediately creates an alert which contains information about the source, target, and estimated type of the attack (e.g., SQL injection, buffer overflow, or denial of service). As the intrusive actions caused by a single attack instance— which is the occurrence of an attack of a particular type that has been launched by a specific attacker at a certain point in time—are often spread over many network connections or log file entries, a single attack instance often results in hundreds or even thousands of alerts. IDS usually focus on detecting attack types, but not on distinguishing between different attack instances. In addition, even low rates of false alerts could easily result in a high total number of false alerts if thousands of network packets or log file entries are inspected. As a consequence, the IDS creates many alerts at a low level of abstraction. It is extremely difficult for a human security expert to inspect this flood of alerts, and decisions that follow from single alerts might be wrong with a relatively high probability. In our opinion, a “perfect” IDS should be situation-aware [2] in the sense that at any point in time it should “know” what is going on in its environment regarding attack instances (of various types) and attackers. In this paper, we make an important step toward this goal by introducing and evaluating a new technique for alert aggregation. Alerts may originate from low-level IDS such as those mentioned above, from firewalls (FW), etc. Alerts that belong to one attack instance must be clustered together and meta-alerts must be generated for these clusters. The main goal is to reduce the amount of alerts substantially without losing any important information which is necessary to identify ongoing attack instances. We want to have no missing meta alerts, but in turn we accept false or redundant meta-alerts to a certain degree. This problem is not new, but current solutions are typically based on a quite simple sorting of alerts, e.g., according to their source, destination, and attack type. Under real conditions such as the presence of classification errors of the low-level IDS (e.g., false alerts), uncertainty with respect to the source of the attack due to spoofed IP addresses, or wrongly adjusted time windows, for instance, such an approach fails quite often.

Our approach has the following distinct properties:
It is a generative modeling approach [3] using probabilistic methods. Assuming that attack instances can be regarded as random processes “producing” alerts, we aim at modeling these processes using approximative maximum likelihood parameter estimation techniques. Thus, the beginning as well as the completion of attack instances can be detected. It is a data stream approach, i.e., each observed alert is processed only a few times [4]. Thus, it can be applied online and under harsh timing constraints. The remainder of this paper is organized as follows: In Section 2 some related work is presented. Section 3 describes the proposed alert aggregation approach, and Section 4 provides experimental results for the alert aggregation using various data sets. Finally, Section 5 summarizes the major findings.

II. RELATED WORK

Most existing IDS are optimized to detect attacks with high accuracy. However, they still have various disadvantages that have been outlined in a number of publications and a lot of work has been done to analyze IDS in order to direct future research (cf. [3], for instance). Besides others, one drawback is the large amount of alerts produced. Recent research focuses on the correlation of alerts from (possibly multiple) IDS. If not stated otherwise, all approaches outlined in the following present either online algorithms or as we see it can easily be extended to an online version. Probably, the most comprehensive approach to alert correlation is introduced in [4]. One step in the presented correlation approach is attack thread reconstruction, which can be seen as a kind of attack instance recognition. No clustering algorithm is used, but a strict sorting of alerts within a temporal window of fixed length according to the source, destination, and attack classification (attack type). In [5], a similar approach is used to eliminate duplicates, i.e., alerts that share the same quadruple of source and destination address as well as source and destination port.

In addition, alerts are aggregated (online) into predefined clusters (so-called situations) in order to provide a more condensed view of the current attack situation. The definition of such situations is also used in to cluster alerts. In [6], alert clustering is used to group alerts that belong to the same attack occurrence. Even though called clustering, there is no clustering algorithm in a classic sense. The alerts from one (or possibly several) IDS are stored in a relational database and a similarity relation which is based on expert rules is used to group similar alerts together. Two alerts are defined to be similar, for instance, if both occur within a fixed time window and their source and target match exactly. As already mentioned, these approaches are likely to fail under real-life conditions with imperfect classifiers (i.e., low-level IDS) with false alerts or wrongly adjusted time windows.

Another approach to alert correlation is presented in [7]. A weighted, attribute-wise similarity operator is used to decide whether to fuse two alerts or not. However, as already stated in [8] and [9], this approach suffers from the high number of parameters that need to be set. The similarity operator presented in [10] has the same disadvantage there are lots of parameters that must be set by the user and there is no or only little guidance in order to find good values. In [11], another clustering algorithm that is based on attribute-wise similarity measures with user defined parameters is presented. However, a closer look at the parameter setting reveals that the similarity measure, in fact, degenerates to a strict sorting according to the source and destination IP addresses and ports of the alerts. The drawbacks that arise thereof are the same as those mentioned above. An attack instance in our sense can also be seen as a kind of root cause, but in [8] root causes are regarded as “generally persistent” that does not hold for attack instances that occur only within a limited time window. Furthermore, only root causes that are responsible for a majority of alerts are of interest and the attribute-oriented induction algorithm is forced “to find large clusters” as the alert load can thus be reduced at most. Attack instances that result in a small number of alerts (such as PHF or FFB) are likely to be ignored completely. The main difference to our approach is that the algorithm can only be used in an offline setting and is intended to analyze historical alert logs. In contrast, we use an online approach to model the current attack situation. The alert clustering approach described in [9] is based on [8] but aims at reducing the false positive rate. The created cluster structure is used as a filter to reduce the amount of created alerts. Those alerts that are similar to already known false positives are kept back, whereas alerts that are considered to be legitimate (i.e., dissimilar to all known false positives) are reported and not further aggregated. The same idea but based on a different offline clustering algorithm is presented in [15]. There, the reconstruction error of an autoassociator neural network (AA-NN) is used to distinguish different types of alerts. Alerts that yield the same (or a similar) reconstruction error are put into the same cluster. The approach can be applied online, but an offline training phase and training data are needed to train the AA-NN and also to manually adjust intervals for the reconstruction error that determine which alerts are clustered together. In addition, it turned out that due to the dimensionality reduction by the AA-NN, alerts of different types can have the same reconstruction error which leads to erroneous clustering. In our prior work, we applied the well-known e-means clustering algorithm in order to identify attack instances. However, this algorithm also works in a purely offline manner.
III. ONLINE ALERT AGGREGATION TECHNIQUES

In this section, we describe our new alert aggregation approach which is at each point in time based on a probabilistic model of the current situation. To outline the preconditions and objectives of alert aggregation, we start with a short sketch of our intrusion framework. Then, we briefly describe the generation of alerts and the alert format. We continue with a new clustering algorithm for offline alert aggregation which is basically a parameter estimation technique for the probabilistic model. After that, we extend this offline method to an algorithm for data stream clustering which can be applied to online alert aggregation. Finally, we make some remarks on the generation of meta-alerts.

3.1 COLLABORATING INTRUSION DETECTION AGENTS

As can be seen in fig. 1, there are many layers namely sensor layer, detection layer, alert processing layer and reaction layer. The sensor layer is responsible to generate TCP or UDP traffic that is given to the next layer for detection. The detection layer is responsible to detect intrusion based on misuse and anomaly detection. The generated flood of alerts is given to the next layer known as alert processing layer. The alert processing layer makes use of the proposed probabilistic technique in order to aggregate alerts. The aggregated alerts are given to reaction layer which provides meaningful reports to security personnel besides taking prevention measures. [24] in more detail. In our layered ID agent architecture, each layer assesses, filters, and/or aggregates information produced by a lower layer. Thus, relevant information gets more and more condensed and certain, and, therefore, also more valuable. We aim at realizing each layer in a way such that the recall of the applied techniques is very high, possibly at the cost of a slightly poorer precision. In other words, with the alert aggregation module on which we focus in this paper we want to have a minimal number of missing meta-alerts (false
negatives) and we accept some false meta alerts (false positives) and redundant meta-alerts in turn.

3.2 ALERT GENERATION AND FORMAT
At the sensor layer, sensors determine the values of attributes that are used as input for the detectors as well as for the alert clustering module. Attributes in an event that are independent of a particular attack instance can be used for classification at the detection layer. Attributes that are (or might be) dependent on the attack instance can be used in an alert aggregation process to distinguish different attack instances. A perfect partition into dependent and independent attributes, however, cannot be made. Some are clearly dependent (such as the source IP address which can identify the attacker), some are clearly independent such as the destination port which usually is 80 in case of web based attacks), and lots are both (such as the destination port which can be a hint to the attacker’s actual target service as well as an attack tool specifically designed to target a particular service only). In addition to the attributes produced by the sensors, alert aggregation is based on additional attributes generated by the detectors. Examples are the estimated type of the attack instance that led to the generation of the alert (e.g., SQL injection, buffer overflow, or denial of service), and the degree of uncertainty associated with that estimate.

3.3 DATA STREAM ALERT AGGREGATION
In this section, offline approach is extended to an online approach working for dynamic attack situations. Assume that in the environment observed by an ID agent attackers initiate new attack instances that cause alerts for a certain time interval until this attack instance is completed. Thus, at any point in time the ID agent—which is assumed to have a model of the current situation, the correct assignment of alerts to attack instances is known (indicated by different symbols). Actual observations: The alerts produced by a real detection layer. The task of the alert aggregation is to reconstruct the attack situation by means of these observations only (including false alerts). Reconstruction: The result of the aggregation (correspondence of alerts and clusters/meta-alerts) together with four different types of problems that may arise.

1. Component adaption: Alerts associated with already recognized attack instances must be identified as such and assigned to already existing clusters while adapting the respective component parameters.
2. Component creation (novelty detection): The occurrence of new attack instances must be stated. New components must be parameterized accordingly.
3. Component deletion (obsoleteness detection): The completion of attack instances must be detected and the respective components must be deleted from the model.

That is, the ID agent must be situation-aware and try to keep his model of the current attack situation permanently up to date. Clearly, there is a trade-off between runtime (or reaction time) and accuracy. For example, it is hardly possible to decide upon the existence of a new attack instance when only one observation is made. From the viewpoint of our objectives (cf. Section 3.1), the tasks 1 and 2 are more time critical than task 3.

From a probabilistic viewpoint we can state that our overall random process is nonstationary in a certain sense which can be regarded as being equivalent to changing the mixing coefficients at certain points in time. A mixing coefficient is either zero or the reciprocal of the number of “active” components (for the time interval of the respective attack instance). With appropriate novelty and obsoleteness detection mechanisms, we aim at detecting these points in time with both sufficient certainty and timeliness.

3.4 META-ALERT GENERATION AND FORMAT
With the creation of a new component, an appropriate meta alert that represents the information about the component in an abstract way is created. Every time a new alert is added
1. Sequences of meta-alerts may be investigated further in order to detect more complex attack scenarios (e.g., by means of hidden Markov models).
2. Meta-alerts may be exchanged with other ID agents in order to detect distributed attacks such as one-to-many attacks.
3. Based on the information stored in the meta-alerts, reports may be generated to inform a human security expert about the ongoing attack situation.

Meta-alerts could be used at various points in time from the initial creation until the deletion of the corresponding component (or even later). For instance, reports could be created immediately after the creation of the component or which could be more preferable in some cases a sequence of updated reports could be created in regular time intervals. Another example is the exchange of meta alerts between ID agents: Due to high communication costs, meta-alerts could be exchanged based on the evaluation of their interestingness [18].

According to the task for which meta-alerts are used, they may contain different attributes. Examples for those attributes are aggregated alert attributes (e.g., lists or intervals of source addresses or targeted service ports, or a time have been created by the alert aggregation module. These components are the basis for meta-alert generation.

(b) Assignment problem: New observations must either be assigned to an existing component which is then adapted or a new component must be created. Also, outdated components must be deleted. (c) Adapted model: The new situation after a few steps. One component has been created, one component has been deleted, and the other components have been adapted.
Online intrusion detection and intrusion alert aggregation with generative modeling

In this paper, the existing approach of intrusion detection systems alert aggregation have been analyzed. From this analysis, the existing method produces hundreds or even thousands of alerts instead of single alert and chances of missing alert. In our approach, an online intrusion alert aggregation with data stream and generative modeling will be developed. This approach is effectively reduces the number of alerts without missing a single alert, which result efficient clustering with no false clustering, true positive rate is improved and well suited for online applications (real time) will be implemented. In future, same concept of alert aggregation will be upgrade for heterogeneous networks.

IV. CONCLUSION

In this paper, the existing approach of intrusion detection systems alert aggregation have been analyzed. From this analysis, the existing method produces hundreds or even thousands of alerts instead of single alert and chances of missing alert. In our approach, an online intrusion alert aggregation with data stream and generative modeling will be developed. This approach is effectively reduces the number of alerts without missing a single alert, which result efficient clustering with no false clustering, true positive rate is improved and well suited for online applications (real time) will be implemented. In future, same concept of alert aggregation will be upgrade for heterogeneous networks.

REFERENCE


AUTOMATED FAILURE DETECTION AND ANALYSIS

SWAPNIL B. KADAM¹ & SHASHANK P. NAIK²

¹²S.S.P.M College of Engg. Computer Engineering dept,Kankavli

Abstract—When a system fails to function properly, health-related data are collected for troubleshooting. However, it is challenging to effectively identify anomalies from the voluminous amount of noisy, high-dimensional data. The traditional manual approach is time-consuming, error-prone, and even worse, not scalable. We present an automated mechanism for node-level anomaly identification in large-scale systems. A set of techniques is presented to automatically analyze collected data. Data transformation to constructu a form data format for data analysis, feature extraction to reduce data size, and unsupervised learning to detect the node acting differently from others. Moreover, we compare two techniques, principal component analysis (PCA) and independent component analysis (ICA), for feature extraction.

I. INTRODUCTION

While working in large scale network, number of problems arise, the total time required to deal with such type of problems depends on the severity of problems. As system takes more time to recover from failures, maintenance cost goes on increasing, it also causes loss of processing cycles. To deal with such type of loss, the information at various nodes in network is collected and verification of failure reason is performed. In traditional system, this process of dealing with failures was handled by humans, but such manual process was leading to various problems such as consumption of time, scalability of network and many more. As scalability of network goes on increasing, it is caused by handling a large amount of nodes.

II. RELATED WORK

Lots of work is going on anomaly identification and failure diagnosis [1]. Basically, this is without applying any extraction technique or by using feature extraction technique based on principal component analysis (PCA) [1][2] and independent component analysis (ICA) [1][3] used for feature extraction in anomaly identification scheme. These two techniques are two well-known pattern recognition techniques. Both of these techniques are supervised and unsupervised, it will never mind the abnormal or normal. The accuracy of the extraction technique is measured by its sensitivity and specificity. The sensitivity (i.e., the proportion of correctly recognized faults) is the measure of the number of actual faults that are detected. The specificity (i.e., the proportion of correctly recognized 1) Principle component analysis: PCA is a well-known dimensionality reduction technique. Basically, it performs linear transformation on the input data to find a new set of features which are uncorrelated. The principal component analysis is performed by finding the direction of maximum variance in the data set.

2) Independent component analysis: ICA is also a so-called blind source separation technique, it is responsible for finding a new set of basis vectors for vectorspace transformation. However, it finds the directions of maximum independence with the help of higher order statistics. Outlier detection is a technique used to perform identification of abnormal nodes.

III. TECHNICAL CHALLENGES

The traditional anomaly identification schemes are either low consuming or error-prone. Basically, the accuracy of automated anomaly identification using PCA and ICA based extraction techniques [1] is comparable in case of single fault but ICA outperforms PCA in case of multiple fault occurrences. 1) Data diversity: Depending on the monitoring tools used, the data collected often have different formats and semantics, thereby making it hard to process them in a uniform way.

2) Datavolume: Due to the size of modern systems, data collected for analysis is characterized by its huge volume, e.g., an order of gigabytes or more. Finding anomalies from such huge data is like finding needles in a haystack.

3) Datadependency: Most measured data are mixtures of independent signals, and they often contain noise. A naive method that directly compares the measured data for anomalies is generally inaccurate by producing substantial amount of false alarms.

4) Anomaly characteristics: In large-scale systems, anomaly types are many and complex. Moreover, node behaviors are dynamically changing during operation. To meet these challenges, basically two main feature extraction techniques were used: PCA and ICA.
IV. PROPOSED WORK

In proposed work node level information such as net work traffic, memory utilization, I/O read/written, data are collected across the system and used for purpose of automated data analysis. The automated mechanisms are applied periodically with a predefined frequency by a system monitoring tool in case of unusual events. The anomalies are then provided for manual validation to the system administrator. In the diagram shown below, the shaded boxes indicate areas of interest. By applying fast processing capability of computers with human expertise, the proposed mechanism can quickly identify the anomalies with very high accuracy (i.e., with high sensitivity and high specificity). The group analysis is performed on the basis of two scenarios. First, based on the nodes which are performing activities similar to the majority of nodes. Second, fault occurs rarely, so the rest of the system nodes behave normally. The group analysis is divided into three modules. It can be applied to find the nodes that exhibit different behaviors from the majority as follows:

1) Data transformation: It is a process of collecting related information across the system and assembling them into a uniform format. Here, a feature of a node is defined as any individually measurable characteristic of the node being observed, such as CPU utilization, memory size, I/O, network traffic, etc.

2) Feature extraction: A feature extraction based on ICA [1] is applied on the feature matrix to obtain a matrix which has much lower dimensionality while keeping the most relevant information in the data. This not only gives acceleration to data analysis but also reduces the dimensionality. For better and faster convergence, the FastICA algorithm is proposed.

3) Outlier detection: It is used to determine the nodes that are behaving differently from the majority of nodes. An abnormal behavior can be identified by analyzing the matrix generated by feature extraction. An outlier detection algorithm can quickly identify the outliers.

In manual validation once outliers are identified, these outliers are focused to troubleshoot the problems at a particular node. This increases the chances of improving the performance of the network.

Thistotalprocesshelpsinachievingthefollowing:

1) Sensitivity: It is the proportion of correct faulty classification to the number of faulty nodes.
2) Specificity: It is the proportion of correct non-faulty classification to the number of non-faulty nodes.

Sensitivity of 1.0 means that the mechanism recognizes all the faulty nodes and specificity of 1.0 means that the mechanism identifies all the non-faulty nodes.

V. CONCLUSION

System proposed in the paper should identify anomalies with highest probability and identifying nodes under failure. Making fault-tolerant systems, fault-tolerance sensitivity (very close to 1.0) and high specificity (above 0.9).

REFERENCES

[1] Toward Automated Anomaly Identification in Large-Scale Systems Zhiling Lan, Member, IEEE Computer Society, Ziming Zheng, Student Member, IEEE, and Yawei Li, Member, IEEE, Transaction on Parallel and Distributed Systems, Vol. 21, No. 2, February 2010
SMART FRAMES DECRYPTING ALGORITHM FOR HYDRO ACOUSTIC COMMUNICATION

SUMEDH S GAWANDE¹ & NIDHI MALVIYA²

¹,²Trinity Institute of Technology & Research, Bhopal, Madhya Pradesh

Abstract- The electromagnetic wave propagation is very poor in water. Therefore it cannot be used for under water applications. Hence the ultimate solution is acoustic communication. In the acoustic communication we are using the acoustic waves for communication. The hydro acoustic modulator will be used at physical layer. But to increase the security we are proposing the smart deciphering algorithm at the receiver end to avoid the unwanted listeners. By using this smart deciphering the only the required node will be able to decipher the frames and unwanted listeners will be avoided. The proposed algorithm will be designed using MATLAB.

Keywords- Booth multiplier, Low power, modified booth multiplier, VHDL, partial product generation (PPG).

I. INTRODUCTION

As electromagnetic waves propagate poorly in sea water, acoustics provides the most obvious medium to enable underwater communications. High-speed communication in the underwater acoustic channel is challenging due to extended multi-path, refractive properties of the medium, severe fading, rapid time-variation and large Doppler shifts. Communication techniques originally developed for terrestrial wired and wireless channels need significant modifications to suit underwater channels.

The development of digital communications for undersea applications dates back to simple ping-based use of sonar that operates in the audible band. The use of "one ping only" in the fictional Hunt for Red October that was used to communicate from submarine to submarine is an example of digital communications, which while primitive, certainly is sufficient when only one bit is necessary.

High-speed communication in the underwater acoustic channel has been challenging due to a number of reasons. Extended multi-path, rapid time-variation and severe fading are also common in many underwater channels. The relatively slow signal propagation speed in water implies large Doppler shifts at moderate speeds of communicating platforms.

The multipath and time-varying are the two key characters of underwater acoustic channel, and the major obstacle of high-speed and reliable transmission of underwater acoustic channel is narrow band and serious signal fading. In order to transmit data high-speedily and more reliably in such harsh and complex underwater acoustic channel, a good error correction code requires be adopted to improve the reliability of transmission. And Turbo code is one of the best error correction channel code in the world nowadays, therefore, it is applied as channel coding of coherent communication on "JiaoLong" manned submersible.

At the MAC layer Long-baseline (LBL) acoustic positioning systems, widely covered in various scientific publications, are based on specific protocols and were designed to solve one particular task, namely the positioning of remote targets. The proposed algorithm is solutions which extend the application range of such systems by combining them with underwater acoustic communication systems. Both systems can share the same electro-acoustic circuit for transmitting and receiving signals and the same signal processing module. The same acoustic signal can be used for both transmission of digital information and estimation of its source position. Hence, the algorithm is proposed that will decipher the data frames efficiently or at minimum error rate at receiving node end. Therefore in this manner the network reliability will be enhanced.

II. LITERATURE REVIEW

The wireless communications field started with the discovery of the long range propagation characteristics of radio waves. Based on sonar development, underwater acoustic communications was then developed [1].Acoustic signal is the only physical feasible tool that works in underwater environment. Compared with it, electromagnetic wave can only travel in water with short distance due to the high attenuation and absorption effect in underwater environment. It is found that the absorption of electromagnetic energy in sea water is about \(45 \times f \text{ dB per kilometer, where } f \text{ is frequency in Hertz; In contrast, the absorption of acoustic signal over most frequencies of interest is about three orders of magnitude lower [2]. As electromagnetic waves propagate poorly in sea water, acoustics provides the most obvious medium to enable underwater communications. High-speed communication in the underwater acoustic channel is challenging due to limited bandwidth, extended multi-path, refractive properties of the medium, severe fading, rapid time-
variation and large Doppler shifts. The capability to communicate efficiently underwater has important applications including oceanographic studies, offshore oil prospection and extraction, and defense operations. As electromagnetic waves cannot propagate over long distances underwater, acoustic communication assumes an important role for such applications [3].

Multicarrier modulation in the form of orthogonal frequency division multiplexing (OFDM) has been shown feasible for underwater acoustic communications via effective algorithms to handle the channel time-variability. New method to construct non-binary regular and irregular LDPC codes that achieve excellent performance, match well with the underlying modulation, and can be encoded in linear time and in a parallel fashion. Based on the fact that the generator matrix of LDPC codes has high density, we further show how to reduce the PAPR considerably with minimal overhead [4].

The propagation medium of underwater acoustic channel exhibits distinct characteristics when contrasted with other common propagation media such as copper, fiber, and radio. In particular there is the extremely slow propagation speed of sound in water, high signal attenuation due to absorption, significant delay spreads and inter symbol interference, and range-dependent transmission bandwidth. These features make the delay reliability tradeoff for underwater acoustic channels fundamentally different from other channels [5].

Long baseline (LBL) acoustic positioning systems are widely covered in scientific publications and are available as commercial solutions. The task of acoustic positioning basically comes down to precise measurement of the signals arrival time and its propagation velocity. Acoustic positioning is thus affected by in homogeneity of the propagating signal’s environment, variability of environmental parameters, multipath propagation of the signal, movements of positioning targets and the antenna’s baseline nodes along with other challenges, common for acoustic communication [6].

For divergence of the adaptive DFE leading to loss of data frames and affecting the performance of Turbo decoding in harsh UAC (Underwater Acoustic Communication) channel, inter-frame interleaved Bi-SOVA decoding algorithm was proposed. Inter-frame interleaved Bi-SOVA algorithm encodes and decodes the whole data packet (28 frames) to correct the lost frames, since it can distribute error bits randomly [7].

III. METHODOLOGY

The underwater network using acoustic communication is versatile for the military & research applications. Hence, these networks like other networks must be secured & efficient enough to locate the destination & retrieve the data packets without error. This method can improve & solve the problem of accessing the wireless network in hydro acoustic communication. The acoustic communication best suited to the underwater implication for network. But the issues are same for this communication system to like the other networks using electromagnetic communication. The retrieval of the data packets in network is one of big issue to avoid data loss. The solution required for efficient communication & retrieval of packets without loss should be accomplished. This requirement is fulfilled by advanced MAC layer algorithm for locating & deciphering data packets for implication in the underwater acoustic networks.

The proposed algorithm will decipher the data packets in the network at destination node. The proposed MAC layer algorithm combines the merits of LBL (LONG BASELINE) positioning with Bi-SOVA algorithm to locate & retrieve the data packets. Long baseline (LBL) acoustic positioning systems are widely covered in scientific publications and are available as commercial solutions.

In this paper we are explaining the dual encryption transmission of spread spectrum. And then decrypt that encrypted data without any prior knowledge of encrypting PN sequence. First data sequence is directly encrypted by PN sequence in direct sequence spread spectrum (DSSS). Then this encrypted data is modulated & its modulation is again encrypted by PN sequence in frequency hopping spread spectrum (FHSS). As data is encrypted two times the data is more secure from the unwanted user. To describe in details the operations of spread spectrum detection & PN sequence retrieval, the simplified block diagram of a transmitter & receiver for spread spectrum multiple access system is illustrated in Fig. 1 & Fig. 2.

![Transmitter block diagram](image-url)
Here, in this block diagram of transmitter first we the data sequence b(n) is directly multiplied with the PN sequence c(t) in product modulator or modulator. The output of the modulator is wide spectrum signal m(N). The spectrum of this signal is quite high compared to that of narrow band data signal b(n).

Then, this signal m(N) is applied to the M-ary FSK modulator. The modulator output is particular frequency depending upon the input symbol. The output of the modulator is applied to mixer. The other input of the mixer is the particular frequency from frequency synthesizer. The output of the frequency synthesizer at particular instant is frequency slot or ‘hop’. The output of the mixer is DS/FH/M-ary FSK signal & is transmitted over the wideband channel.

The frequency hops given to mixer are generated by the frequency synthesizer. The input of the frequency synthesizer is controlled by the pseudo-noise (PN) sequence generator. The ‘t’ successive bits of PN sequence generator control the frequency hops generated by synthesizer. Since the bits of the PN sequence generator change randomly, the frequency hops generated also change randomly. Since ‘t’ bits PN sequence controls frequency hops, there will be distinct ‘2^t’ frequency hops generated. The total bandwidth of the output signal is very high.

The received DS/FH/ MFSK signal is applied to hopping pattern recognition (HPR) block. This block consists of the functional stages initial frequency detection (IFD), remaining frequency detection (RFD), compute hopping pattern (CHP). The goal of the detection device is to detect the HP (hopping pattern) without any prior knowledge about the HP used by the sender at the receiver. The HP of wireless network will be known using this block without any prior knowledge, about the HP used by the sender at the receiver.

Hopping pattern (HP) may consist of two to m random frequencies. Each frequency may exist only once. Assume H.P is F with frequencies F₀, F₁, F₂,......, Fₘ. To detect that a frequency being transmitted clear channel assessment is required. It requires 50 µs to indicate busy. To recognize Hopping Pattern formula is expressed as:

\[
\text{Hopping Pattern} = \sum_{n=2}^{m} f_{x} \cdot d_{x}
\]

Here, dₙ is the dwell time of frequency Fₓ. Only one frequency will be detected by detection device at a time. This detection occurs by calling the CCA procedure. To detect the specific frequency transmitted currently by clear channel assessment transmission should be of complete test period. When on transmission stops on the detected frequency before test period then, the correct frequency may not be available. This failure when the frequency hop is less than 50 µs can overcome in stage six. The output of the third stage is then given to the decision device, then to the fourth stage M-ary FSK detector.

The detector detects the particular symbol transmitted. In the FH/M-ary FSK the individual frequency of smallest duration is called ‘chip’. Now, at the output of the M-ary FSK detector we get the ‘N’ bit symbol in parallel. From these ‘N’ bits‘t’ bit LSB’s are then transferred to the PN sequence register r (t). These are the PN sequence for that ‘N’ bit symbol. Then the ‘N’ bit symbol is converted to the serial bit stream by parallel to serial converter. Then this ‘N’ serial bits are applied to the mixer with the ‘t’ bit LSB’s. The output of the mixer is the data signal b(n). This ‘t’ bit LSB’s are then used to generate the frequency hops by frequency synthesizer. This is fed to the mixer with the input signal & compared with the output of HPR algorithm in decision block. The output of the both HPR algorithm & the mixer should be same. If the frequency hop is lesser than the 50 µs than HPR algorithm fails. Then the output of the mixer is passed to the M-ary FSK detector.

IV. CONCLUSION

This method can improve & solve the problem of accessing the wireless network in hydro acoustic communication. As the error rate decreases the network will be more reliable. The data frames will be retrieved with very less error & secured way. The destination nodes will be accessed reliably. The proposed algorithm will allow implementing custom media access schemes that require precise synchronization of network nodes. The hydro acoustic network will be more reliable.

REFERENCES


TIME EFFECTIVE INTRUSION DETECTION MODEL BASED ON K-MEANS CLUSTERING AND PROBABILISTIC CLASSIFICATION USING FUZZY GNP HYBRID RULE MINING

NIVEDITA P. CHAUDHARI¹ & PRAVIN J. NIKUMBH²

¹Computer Engineering, R.A.I.T, Nerul, Mumbai, India
²4/20, Pujari Apt., Devnar, Mumbai 88., India

Abstract— Intrusion Detection System (IDS) plays an effective way to achieve higher security in detecting malicious activities. Building effective IDS is an enormous knowledge engineering task. Data mining has emerged as an important research area and association rule mining become one of the most attractive tasks of data mining. Network intrusion detection model based on evolutionary optimization technique called Genetic Network Programming (GNP) with sub attribute utilization mechanism and probabilistic classification is proposed with combination of K means clustering algorithm. K-means algorithm is used for filtering purpose. From data of K number of clusters, irrelevant data is filter out, so that remaining data can passed to next processing of generation of fuzzy class association rules. The main aim of using K-means algorithm is to generate less but strong rules. Hybrid rule mining uses fuzzy class association rule mining algorithm to extract rules with different classes. Then, using different class rules and the classification of data is done probabilistically.

Index Terms- Data Mining, Intrusion Detection System, K-means Clustering Algorithm, Genetic Network Programming, Probabilistic Classification.

I. INTRODUCTION

Nowadays, intrusion and crimes related to computer system have been increasing as computer and communication technologies are developing. Therefore, for provision of stable services network security become an important matter. Intrusion detection system emerges to detect malicious behavior that compromises security. Intrusion detection system is used to differentiate between normal activities and abnormal activities of system. Misuse and anomaly detection are two main IDS approaches. Misuse detection approach detects intrusion based on pattern of known attacks while anomaly detection approach detects new attacks. KDD99Cup [last] dataset and DARPA dataset are widely used as training and testing dataset for evolution of IDS. Among these, in this paper KDD99Cup dataset is used as a standard. Data mining can be defined as process of extracting useful rules from dense database. Due to rapid development of data mining, wide varieties of algorithms are developing. These algorithms can help out for network intrusion detection problems. Intrusion detection aids to classify data into normal and intrusion kind.

K-means is well known unsupervised learning algorithm meant for clustering purpose. K-means algorithm is used to finding out natural grouping of data based on similarities among pattern [7]. In proposed algorithm, K-means clustering is used as filter that filters out data from such a group which contains less number of records. Filtered group can be thought as composition of dissimilar types of records from dataset. Advantage of filtration is to smooth the process of building strong and efficient association rules using fuzzy Genetic Network Programming. Genetic Network Programming (GNP) is evolutionary optimization technique applied to dynamic problems. GNP is extended evolutionary algorithm of genetic algorithm (GA) and genetic programming (GP) gives solution based on graph structures has been proposed [2], [3]. The distinct inherent features of graph structure such as reusability, compact structure making GNP mine sufficient number of important rules instead of all rules meeting criteria [6]. Hybrid rule mining uses combination of GNP with probabilistic classification. Fuzzy logic theory is used with GNP to extract fuzzy rules. Matching degree of data with extracted rule is calculated using probability matching function. Based on calculated matching degree, new connection data can be classified into normal data, known and unknown intrusion.

II. K-MEANS ALGORITHM

Clustering is a method of grouping data into specified number of clusters, such as data from same clusters are quite similar where as data from different clusters are quite different from each other. K-means is well known unsupervised learning algorithm used for clustering [10]. The procedure is simple and easy to group a given data through certain number of clusters, let say k number of clusters. The mean value of numerical data contained within each cluster is called as centroid. In proposed algorithm, k must be greater than number of attacks. The algorithm is composed of following steps:
I. Place K number of records, represented by data that are being clustered, as initial centroids.
2. Assign each record to the cluster that has closest centroid.
3. When all records have been assigned, recalculate the position of K centroids.
4. Repeat step 2 and 3 until centroids no longer move. This results into separation of records into respective groups.

The main goal of K-means clustering algorithm is to group data into specified number of clusters. The approach is generally used in intrusion detection system to classify data into normal and attack instances. In proposed algorithm, the approach of K-means clustering as classification has been changed to filtration. The referred dataset contains both discrete and continuous attribute. These both attributes before applying to K-means algorithm are converted into numerical format to deal with clustering. The data records then grouped into k clusters. The clusters containing minimum number of records are filtered out so as to dataset contain only similar type of records. Because of filtration, from remaining dataset strong and efficient number of rules can be generated with less time required as compared with existing system.

III. OVERVIEW OF GNP AND CLASS ASSOCIATION RULE MINING

In this section, a class association rule mining using GNP is explained.

A. Architecture of GNP

A large number of studies been conducted on evolutionary optimization techniques. Genetic Algorithm and Genetic Programming are typical evolutionary algorithms. GA evolves strings and mainly applied to optimization problem. GP was devised later which uses tree structure for solution. GNP is one of the evolutionary algorithm which uses directed graph structure instead of trees and strings. Directed graph structure of GNP is represented by gene structure which consists of judgment nodes and processing nodes [2]. Function of judgment node is to examine attribute of tuple and return a judgment result either “Yes” or “No”. Then, corresponding branch of judgment node is connected to next node. J1, J2...Jn are judgment functions by judgment nodes showing attributes to examine where N is the total number of judgment functions. On the other hand processing nodes are set of P1, P2...PM where M is total number of processing functions which works as action/processing functions. Execution starts from start node, then next node to be executed is determined according to the connection between nodes and a judgment result of current activated nodes. Fig. 1 shows the gene of a node in GNP individual. Fig. 2 is the genotype expression of GNP in GNP individual. 0 is start node, 1 is processing node, 2 is judgment node. ID, works as identification number. C1, C2...Cn means the nodes connected from node i, firstly, secondly and so on.

Selection, crossover and mutation are three kinds of genetic operators.

B. GNP Based Class Association Rule Mining

Association rule mining find correlation between features or attributes used to describe a data set. Let I = {A1, A2,...An} be set of attributes and each tuple T has a set of attributes satisfying T ⊆ I. Association rule can be represented as X ⇒ Y. It means, tuple satisfying X are likely to satisfy Y, where X is antecedent and Y is conclusion of the rule. Support(X) = X and Support(Y) = Y. If event X and Y are independent, then support (X ∩ Y) = xy. Then, \( \chi^2 \) of rule X ⇒ Y can be calculated as [4]

\[
\chi^2 = \frac{N(\text{xy})^2}{\text{xy}(1-x)(1-y)}
\]

(1)

Important association rule mining rule should satisfy following conditions:
Hybrid Rule Mining based on Fuzzy GNP

The conventional representation of class association rule based on GNP is shown in Fig. 3. Hybrid rule mining is combination of fuzzy GNP and conventional class association rule mining. Hybrid rule mining have the advantage of utilizing both discrete and continuous attributes in one single rule. Fig.4 describes an example of hybrid rule mining representation. Rule extraction starts from processing node $P_r$. The first judgment node examines the fuzzy membership value of continuous attribute $A_1$, the second judgment node examines the value of discrete attribute $A_2$, and the third judgment node examines the value of symbol attribute $A_3$. $N$ is the number of total tuples. $P_r$ is the probability moving to Yes-side. When random value selected in $(0, 1)$ is smaller than the certain probability $P_r$, GNP selects Yes-side and goes to the next judgment node. Probability $P_r$ is the fuzzy value of corresponding attribute. Otherwise, transition starts from next processing node to find new rule.

B. Fuzzy Membership Function for Continuous Attributes

In this paper, we utilize advantage of fuzzy theory to have every continuous attribute value in $[0, 1]$. Fuzzy theory allows complex system to have linguistic description [5]. In the paper, each continuous attribute in the database is transformed into 5 linguistic terms (Very low, Low, Middle, High, Very high). By using 5 linguistic terms for single continuous attribute, we will get more accurate membership value for corresponding continuous attribute.

The linguistic terms are defined by the combination of trapezoidal and triangular membership functions. The parameters $x_1$, $x_2$, $x_3$, $x_4$ and $x_5$ are also evolved along with the evolution of GNP. Each continuous attribute have its own membership value. The parameters for each continuous attribute are initialized by analyzing the distribution of data. During evolution process, the parameter value of fuzzy membership function should be adjusted generation by generation. Fuzzy membership values are used to determine the transition in GNP individuals while searching for association rules.

C. Fitness Function

In this paper, fitness function of GNP individual is as

$$ F = \sum_{r=1}^{R} \left( \chi^2(r) + 10(n(r) - 1 + \alpha_{max}(r)) \right) $$

IV. HYBRID RULE MINING

In this paper, hybrid rule mining is used for extraction of rules. Hybrid rule mining utilizes both discrete and continuous attributes in one single rule.

A. Hybrid Rule Mining based on Fuzzy GNP

The conventional representation of class association rule based on GNP is shown in Fig. 3. Hybrid rule mining is combination of fuzzy GNP and conventional class association rule mining. Hybrid rule mining have the
where, $R$ is set of important extracted rules; $n(r)$ is the number of attributes in antecedent part of rule $r$; $\alpha_{new}(r)$ is defined as follows:

$$
\alpha_{new}(r) = \begin{cases} 
\alpha_{new}, & \text{if } r \text{ is new} \\
0, & \text{otherwise}
\end{cases}
$$

(7)

Rule $r$ is new when $\alpha_{new}$, otherwise it is 0.

V. CLASSIFIER BASED ON PROBABILITY

Classifier is constructed to classify new connection data into normal, misuse and anomaly intrusion correctly. Classification is done after extraction of important class association rule including normal and intrusion.

A. Probability Distribution Function

Probability distribution function is constructed according to the matching probability of all data with rules of normal and intrusion. The matching probability of data with a rule is defined as follows.

$$
Match_{Pr}(d) = \frac{N_{d}}{N_{r}}
$$

(8)

where, $N_{d}(d)$ is the number of matched attributes of data $d$ with antecedent part of rule $r$ in class $k$. $N_{r}$ is the number of attributes in the antecedent part of rule $r$.

Then, the average matching probability of the data with all the rules in class $k$ is calculated as follows.

$$
Avg_{Match}(d) = \frac{1}{|R_k|} \sum_{r \in R_k} Match_{Pr}(d)
$$

(9)

where, $R_k$ shows set of rules in class $k$.

Finally, average probability distribution function can be created by distribution of average matching probability of all the training data $d \in D_{train}$ with the rules $r \in R_k$, where $D_{train}$ is set of training data.

B. Classification using Probability Distribution Function

The probability that $P_k(d)$ that new connection data i.e. testing data belongs to class $k$ can be calculated using average probability distribution function. The probabilities of belonging normal class $k = 1$, known intrusion class $k = 2$ and unknown intrusion class $k = 0$ are calculated by Eq. (10), (11) and (12) as follows.

$$
P_{k=1} = (1 - Avg_{Match}(d)) \times (Avg_{Match}(d))
$$

(10)

$$
P_{k=2} = (Avg_{Match_2}(d)) \times (1 - Avg_{Match_1}(d))
$$

(11)

$$
P_{k=0} = 1 - \sum_{k \in C} P_k(d)
$$

(12)

where, $C = \{1, 2, \ldots, K\}$ is set of classes. $K=2$ is used in this paper. Based on calculation of these probabilities, $d$ is assigned to class with highest probability. Positive false rate (PFR) increases when normal data labeled as intrusion and negative false rate (NFR) increases when intruded data labeled as normal. In order to balance PFR and NFR, medication of probability $P_k(d)$ is introduced.

Fig. 6 show flowchart for proposed method. Proposed method overcomes crisp boundary problem and deals with multi-data type database. Time required with comparison of conventional class association rule mining method is less as we are using K-means clustering algorithm for filtration purpose.
VI. CONCLUSION

With the increasing rate of internet, security becomes a serious issue. Therefore, intrusion detection system, which detects intrusion accesses, has to fight against security vulnerability. The proposed algorithm is combination of K-means clustering algorithm and probabilistic classification using fuzzy GNP hybrid rule mining. K-means clustering algorithm is used as filter which filters out data which is not similar among training data. Filtration of data is on the basis of relevancy of data contained in each cluster and numbers of clusters are greater than number of intrusion type. After filtration data is passed for further process i.e. hybrid rule extraction and classification of testing data. Hybrid rule mining algorithm using GNP includes Fuzzy-based class association rule mining and probabilistic classification. The aim is to extract strong and robust class association rules from database in time efficient manner.

REFERENCES


A REVIEW ON CONCURRENCY FAULT DETECTION TECHNIQUES

C.REVATHI1 & M.MYTHILY2

1,2 School of Computer Science and Engineering,, Karunya University, Coimbatore - 641114, Coimbatore, India.

Abstract—— The UML is the most common language that is used for system modeling. But, this language has been designed as a general purpose modeling language that might need modeling constructs for the specific real time embedded (RTE) domain. To fill this lack, OMG has standardized a UML addition, called MARTE. Domain specification UML provides a special way called a profile. The UML profile for MARTE is a concentration of UML that provides idea devoted to real time modeling for design and analysis of real time application and platforms. The existing work, can detect only the concurrency problem such as deadlock and starvation. But, the UML / MARTE profile is specifically designed with the genetic algorithms to detect other concurrency issues like data races. This paper reviews the performance of the search space size and concurrency detection techniques.

Keywords- Genetic Algorithm, Hill Climbing, MARTE Profile, Random Search component.

I. INTRODUCTION

Concurrency problem should be identified in the design phase of software engineering process. It is made progressively difficult in larger and more complex systems. The finding of concurrency issues is based on the design models articulated in UML. Once the UML representation is not sufficient to completely model a system for a particular purpose, the representation is extended by profiles. The adjustment of the MARTE (Modeling and Analysis of Real Time and Embedded Systems) profile [1] addresses domain specific parts of real time concurrent system modeling. The objective of this paper is to detect several types of concurrency errors (such as deadlocks, starvation, and data races, data flow problems) and that can be simply combined into a Model Driven Architecture (MDD) approach, an OMG standard by the UML based MDD [2]. A genetic algorithm (GA) is tailored to identify different types of concurrency issues.

The existing work uses genetic algorithm to detect deadlocks [3] and starvation [4] into a composed form. GA method can also detect the data races. The fitness function exactly designed to detect deadlock and starvation, correspondingly. Now, fitness functions needed towards data race detection and to improve the performance comparison.

The next section presents a comparative study to measure the performance and also compares with a hill climbing search and random search.

II. COMPARATIVE STUDY

This section compares the three techniques: Random Search (RS), Hill Climbing (HC) and Genetic Algorithm (GA). These three techniques can be used to detect the concurrency issues like deadlocks and starvation. The three techniques have been compared based on performance, execution time and search space.

A. Random Search (RS)

The search space, a point is randomly selected and checked for a concurrency fault. Deadlock detection of the random search techniques is capable of discovering a fault, but with very less probabilities. RS can be applied to the cruise design model [4]. The cruise model is used at very small search space. The random search is likely to be as useful as the GA. Random search also detects the starvation. Here, the random search techniques can be useful for the two design models ModPhil (Modified Dining Philosophers problem) and ModCruise. The starvation detection rate is very high, because the ModPhil have the highest search space and ModCruise has a smaller search space size. So, the random search cannot detect the concurrency problem in small search space.

Random generation is random and checked for data races in a point in the search space. Complexity is not an issue in random search as information about the landscape of the search space is not used through the search. Still, random search performs poorly in MEOS [5]. Running a pre-determined space involves running a random search, usually the search space has large number of points. The main disadvantage of RS is that poorly detects the deadlock and also starvation detection rate is very low compared to other two techniques. So, RS can be used only in small search space. Advantages of the RS are good response time and starvation can be detected in any search space.

B. Hill Climbing (HC)

Hill climbing is an optimization technique that works based on randomly chosen candidates. It identifies a set of neighbourhoods depending upon the problem that it represents. A closer move to the neighbor improves the fitness value of the problem solution.

In existing work, HC techniques can be used by the stochastic hill climbing [6]. One random point
and mutating the current point generated via neighboring point. It exchanges it and a new adjacent point is generated. The stopping criteria point is nothing better than the current point, then the execution can be stopped. This continuous fills process a maximum number of classifications needed. In [7], HC worked by randomly choosing a candidate solution and separate iteration the element of a set of near neighboring to the existing solution are considered. These techniques have a crucial error with the HC testing is that the Hill pointed out by the Algorithm may be local maxima. This can be for inferior then global maxima in the search space.

HC does not perform well in the some of the particular design models. It is powerless to detect a deadlock and variance comparing to which genetic algorithm is mathematically significant. Since its execution depends on the first randomly generated chromosome. The GA was analyzed and generated by the number of sequences, and the HC uses the same number of sequences always be generated. If, the optimal chromosome is very different from the initial hill climbing chromosome. Therefore, HC might tired to allocate the number of sequences, HC total detection rate does not differ much in Phl and tends to be poor overall.

With respect to starvation, all three techniques are also capable of detecting this concurrency problem. Here, the detection rate of the starvation is the better than the two design model i.e. ModPhil and ModCruise [5][4]. But, the detection rate of date race is very worst performance by HC techniques. Both hill climbing and the genetic algorithm is significantly better in cost reduction [8]. The main advantage of these techniques are execution time is very low. And the disadvantages are the detection rate of the deadlock and data races is very low in the large search space.

Table 1. Comparison of different techniques

<table>
<thead>
<tr>
<th>Techniques</th>
<th>Fault Types</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deadlock</td>
</tr>
<tr>
<td>RS</td>
<td>**</td>
</tr>
<tr>
<td>GA</td>
<td>****</td>
</tr>
<tr>
<td>HC</td>
<td>*</td>
</tr>
</tbody>
</table>

Table 2. Comparison of Merits and Demerits

<table>
<thead>
<tr>
<th>Techniques</th>
<th>Merits</th>
<th>Demerits</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS</td>
<td>Execution time is very less. The detection rate of the starvation is modualt.</td>
<td>Deadlock and starvation detection rate is very poor eventhought it use, of MARTE profile.</td>
</tr>
<tr>
<td>GA</td>
<td>Deadlock and starvation detection rate is more efficient. GA also detects the other concurrency issues like data races. The result of the data races detection also more efficient.</td>
<td>Total execution time is high. The cost of reduction also high.</td>
</tr>
<tr>
<td>HC</td>
<td>Total execution time is very less. HC can be detecting the issues in large and small search space.</td>
<td>Deadlock and data races detection rate is very poor performance in the search space.</td>
</tr>
</tbody>
</table>

III. SUPPORTING PROFILES

A. UML/SPT profile

The UML Profile for Schedulability, Performance and Time (SPT) that have been established to be curable. In [9] describe problems and possible solutions related to the usage of the profile in the representation of Schedulability analysis models for real-time distributed systems.

The work in [10] determines the new Request for Proposals issued through OMG for a new
UML Profile named “Modeling and Analysis of Real-Time and Embedded systems”. In this[10] explains first a few domain concepts for annotating Nonfunctional Properties (NFPs) and focuses on supporting temporal verification of UML-based models. The importance is given to Schedulability and performance analysis for real-time systems.

The early design process can only identify concurrency problems such as deadlock [3]. But MARTE method based on the analysis of particular models expressed in the UML (Unified Modeling language) that uses an exclusive designed GA to detect deadlocks. All related concurrency information is excerpted from the system UML models that fulfill with the UML Schedulability; Performance and Time profile, a standardized specialization of UML for real-time concurrency system. CFD tool is supporting the methodology. The CFD is decayed into three components: scheduler, GA and RAG evaluator. The SPT sub profiles are RT concurrency modeling, SA profile and RT resources modeling. Deadlock detection is performed using RAG evaluator. A chromosome results into starvation if at least two threads waiting on locks. If a cycle is present, CFD tool will delicts the deadlock. If the cycle is not present the deadlock cannot do detect.

B. UML/MARTE profile

In this profile the UML provides a common extension system for customizing UML models for particular domains and platforms. Extension system allows improving standard semantics in exact additive fashion, so that they can't contradict the standard semantics. Profiles are defined using stereotypes, tag definition and constraints. In Table.3, comparison of stereotypes and tag definitions that are applied to specific model elements such as class, Attributes, operation and activities are shown.

The MARTE profile is a replacement of the Schedulability; Performance and Time (SPT) profile [9]. MARTE is tackled for both the real time and embedded system domain. The profile is roughly divided into two subdivisions: the MA RTE design model and the MARTE analysis model.

In existing work, first UML/MARTE based methodology for executable RTE systems modeling and another one, a framework and its underlying model transformations required to execute UML models conforming to the MARTE standard [11]. In this paper proposed by Accord UML methodology to play the role of that structuring framework. An important aspect of models exploitation is to enable their executability. The Accord framework provides the infrastructure for MARTE model execution. The main disadvantages of Accord execution platform could be parameterized with a subset of MARTE concepts properties but do not support all parameter values, hence this introduces some restrictions on usage of MARTE.

In, [4] [3] UML/MARTE profile detect the concurrency problem such as deadlock and starvation. Here, CFD tool can detect the fault. As illustrated in Fig 1, the profile can be divided to the two subdivision: the MARTE Design model and MARTE analysis model. The design model can be define by models various features of real time and embedded systems. The analysis model is used for the purpose of system analysis. These two subdivision are based on the MARTE foundation, the MARTE foundation defines the time concepts and use of concurrent resources allocation. The concepts are first obtained from the foundation of MARTE, namely, the Generic Resource Molding (GRM) packages. The GRM introduces two stereotypes <<Acquire>>, <<Release>>. The Software Resource Modeling (SRM) is the sub profile of GRM. This presents for designing multitasking application. SRM divided into four packages: SW_Resource Core (resource concepts), SW_Courrrency (concurrent concepts), SW-interaction (communication and synchronization) and SW-Brokering (resource management). Another sub profile is Generic Quantitative Analysis Modeling (GQAM) sub profile. This sub profile defines stereotype <<SaStep>>. Its tag includes priority, interOccTime and execution time. Maximum and Minimum time ranges can be used specified to the execution time.

Table 3. Comparison of SPT and MARTE profile

<table>
<thead>
<tr>
<th>Parameters</th>
<th>SPT</th>
<th>MARTE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>[T, L, a]</td>
<td>[T, L, a] and [T, a]</td>
</tr>
<tr>
<td></td>
<td>(T is a thread , L-is a Lock, a is a specific time unit )</td>
<td></td>
</tr>
<tr>
<td>Tool</td>
<td>CFD</td>
<td>CFD (concurrency Fault Detector)</td>
</tr>
</tbody>
</table>
A Review on Concurrency Fault Detection Techniques

The UML/MARTE profile [3] proposed by cannot detect the starvation and data races. Here, the CFD (Concurrency Fault Detector) tool can be detects deadlock based on the analysis of design description in UML and SPT.

The UML/MARTE profile can be detecting the other concurrency problems such as deadlock and starvation, data races, data flow. In [5] [4] [3] detects the deadlock, starvation and also data races. The UML/MARTE profile can be used to CFD tool. This tool can be detecting the fault accurately. Here, GA is the better performance to other than two techniques (i.e. HC, RS). Because; the concurrency problem (deadlock, starvation and data races) detection rate is high performance in the search space.

In existing work, [5] the data race can be detected by the UML/MARTE sequence diagram of the shared resources in Therac. Three optimization techniques are capable of detecting data races in both MEOS and DECF (resources), but with extremely different probabilities. CFD is an automated system that identifies concurrency problems than any concurrent function modeled with the UML/MARTE schema. Here, GA is well known to produce much better results than the other two techniques. The main advantages are the detection rate of the deadlock, starvation and data races. GA is better than all other techniques because the total execution time is less a cost comparison to other two techniques.

It , describe a methodology and tool for detecting deadlocks and starvation based on the analysis of design representation in UML and MARTE profile. A UML/MARTE model analysis method can be detecting the other concurrency problems like data races. Here, also support the CFD tool. CFD is decomposed into three portions as a scheduler, GA and RAG evaluator.

A. CFD scheduler

Concurrency Fault Detector CFD is the tool used as an automated system that identifies concurrency faults in any concurrent modeled application. The CFD is categorized into 3 parts namely Scheduler, Genetic Algorithm, RAG Evaluation. The work of scheduler is when a system is designed; assumptions are made about the architecture it will be run on. And the deployment assumptions are incorporated in CFD in the form of the scheduler. Scheduler currently emulates single processor execution and is POSIX compliant.

B. RAG evaluation (Deadlock Detection)

If a chromosome result in waiting on locks Deadlock and starvation detection is performed using a RAG. If a cycle is found, CFD outputs the details of the chromosome causing it (executing threads and waiting threads for each lock as well as lock access times), the corresponding RAG, and the fitness value. If no deadlock is found, CFD terminates, showing both the fitness value and output details of the highest fitness chromosome found.

IV. A UML/MARTE MODEL ANALYSIS METHOD FOR CONCURRENT SYSTEM

![Fig 2. Deadlock](image)

For example Fig 2, if the cycle is present the deadlock can be occur. Here, T1 and T2 represent by the resources and P1 and P2(process). P1 request T1(indicated by the solid arrow) and T1 held by P1(indicated by the dotted arrow). Similarly P2 request by T2, but T2 currently held by P1. This situation represents the occurrence of deadlock.

C. Genetic Algorithm

Genetic algorithms operate on a set of possible solutions. Because of the random nature of genetic algorithms, solutions found by an algorithm can be good, poor, or infeasible [defective, erroneous], so there should be a way to specify how good that solution is. This is done by assigning a fitness value [or just fitness] to the solution. Chromosomes represent solutions within the genetic algorithm. The two basic components of chromosomes are the coded solution and its fitness value. To use a genetic algorithm, you must represent a solution to your problem as a genome (or chromosome). The genetic algorithm then creates a population of solutions and applies genetic operators such as mutation and crossover to evolve the solutions in order to find the best one(s).

D. Chromosome representation

A chromosome is collected of genes and perfects a solution to the optimization problem. A gene can be described as a three tuples(T.L,a), where T-thread, L-lock, a-specific time unit when T accesses L, the optimization is done on access times of threads two locks. Under the testing environment time interval should start with start time of system to detect the
starvation. But the verification cost more. To avoid this user defined heuristics developed for verification followed by reduced the time interval.

E. Crossover Operator

They are three constructive that have to be met for the formation of valid chromosome after performing the crossover operation. The three constraints are: 1) All genes within the chromosome are ordered according to increasing thread identifiers, then lock identifiers, then increasing access time. 2) Lock access times must fall within the specified time interval or are set to -1. 3) Consecutive genes for the same thread and lock identifiers must have access time differences equal to at least the minimum and at most the maximum lock access range of the associated thread and lock, if start and end times are defined as ranges.

After performing the crossover if constraint 3 is not met for any of the constraint genes, then the second genes access time is replaced with a randomly chosen access time from a set of possible access times. This is repeated until all the genes meet constraint 3.

F. Mutation

The mutation operator operates by altering the access time of the genes. The access times are moved within a specified time intervals with the aim of finding at the optimal lock access time so that it will not cause starvation. If the chosen value lies outside the time interval the access times is set -1.

V. CONCLUSION

Concurrency abounds in much software system, where systems usually include threads that access shared resources and difficult thread communication. If not handled accurately, such access can deadlock and starvation situation, which might delay system execution. In existing work, HC and RS can be detecting the concurrency problem such as deadlock and starvation. These two algorithm performance is very less efficient. So, in this paper, describe a methodology and tool for detecting deadlocks and starvation based on the analysis of design representation in UML and MARTE profile. A UML/MARTE model analysis method can be detecting the other concurrency problems like data races. MARTE provides a support for modeling the real time embed systems. MARTE focuses on performance and scheduling. GA is better than the other two techniques. Because, the detection rate of the GA performance is more efficient. So, we can conclude that the genetic algorithm is efficiently performed as well or better than the two other alternatives and the difference are driven by the search space size and complexity of the search space.

REFERENCES

A PIXEL-EXPANSION-FREE VCSS APPROACH USING LAGRANGE INTERPOLATING POLYNOMIAL METHOD

RAMAJAYAM.N  NATHELDA MARY NAVINA

Karunya university

Abstract-In the field of visual Cryptography to decrypt the image with quality and no pixel expansion has become a challenge. To face these kinds of challenge in this paper has been proposed a new method called Lagrange Interpolating Polynomial Method. Encryption process the input image is separate into $n$ number of shares using polynomial function techniques. When share has been created the splitted shares are revealed. In the decryption process the reverse process has been taking place to retrieve the original image. Using this proposed method there is no pixel expansion and original quality of image is reconstructed and proved in the experimental result.

Index Terms: Lagrange Interpolating Polynomial, Simulated annealing method, binary image

I. INTRODUCTION

C.blundo says Visual cryptography is a cryptographic paradigm introduced by Naor and Shamir. Some predefined set of participants can decode a secret message without any knowledge of cryptography and without performing any cryptographic computation. In this method we have analyze visual cryptography schemes for grey level images whose pixels have $g$ grey levels ranging from 0 (representing a white pixel) to $g$ -1 (representing a black pixel). A visual cryptography idea for a set $r$ of $n$ participant is a scheme to predetermine a top secret image $SI$ into more number of shadow images called share. Where each member in $r$ receive one share and Certain eligible qualified set of participants can “visually” recover the top secret image. The advantage of the visual secret sharing scheme is decryption process where not including any combination addition encrypted data is decrypted using Human Visual method. [2] But the encryption plan requirements cryptographic calculation to divide the image into a number of parts.[2]

B.w.leung, et al says Visual cryptography is a type of secret sharing techniques for images [3]. The idea of VCS is to divide an image into a group of random share which individually disclose no information about the unique top secret image [3]. The image is relaxing of black and white pixels, and can be recovered by superimpose a threshold number of share with no any operational out mixed up [3]. Within this method also splits a top secret image into [3] more number of shares, the black facade and the other three shares [3]. It was claim that with no significant the black mask, no information about the secret image can be obtained even if all the other three shares are known.[3]

Rezvan dastanian and hadi shahriar shahhoseini says The Information, image and media encryption is a method for preventing misuse of adversaries. Visual cryptography is a method in which decryption is performed with human visual system along with “Revealing” operation operation. In this method one secret image is divided between two shares so that by stacking the two shares secret image appears.[4]. With stacking two shares, secret image I appear and with stacking one of the shares with 90 degrees rotation in clockwise on other share appears the secret image II[4]

Abhishek parakh and subhash kak says a recursive hitting of secret, the user encodes further information about smaller secrets in the shares of a larger secret without an expansion in the size of the latter, thereby increasing the efficiency of secret sharing[5]. The proposed protocol is an application for images as well as text[5]. Thomas month and et al says Visual Cryptography Scheme (VCS) for a set $P$ of $n$ participants is a method to encode a Secret Image (SI) into $n$ shadow images called shares, where each participant in $P$ receives one share Certain qualified subsets of participants can visually recover the SI, but other, forbidden sets of participants have no information on SI.[6].S. Kirkpatrick et al says The method of simulated annealing is a technique that has attracted significant attention as suitable for optimization problems of large scale, especially ones where a desired global extreme is hidden among many, poorer, local extreme. For practical purposes, simulated annealing has effectively “solved” the famous traveling salesman problem of finding the shortest cyclical itinerary for a traveling salesman who must visit each of $N$ cities in turn. The method has also been used successfully for designing complex integrated circuits.[7]

Dimitries bertimas et al says simulated annealing
algorithm is a probabilistic method proposed in Kirkpatrick and cerny for finding the global minimum of a cost function that may possess several local minima. It works by emulating the physical process whereby a solid is slowly cooled so that when eventually its structure is “frozen,” this happens at a minimum energy configuration.[8] Lin TL et al. says The main concept of the original visual secret sharing (VSS) scheme is to encrypt a secret image into n meaningless share images. It cannot leak any information of the shared secret by any combination of the n share images except for all of images. The shared secret image can be revealed by printing the share images on transparencies and stacking the transparencies directly, so that the human visual system can recognize the shared secret image without using any devices. [9] Hiroki koga says The visual secret sharing scheme (VSSS) is a new paradigm of the secret sharing proposed by [1] author. Letting $\mathcal{S} = \{1, 2, \ldots, n\}$ be a set of par-tic pants, in the VSSS a black-white secret image is encrypted to n black-white images called shares.[11] Eric r. verheul, henk c.a says Secret sharing techniques belong to the larger area of information hiding that includes watermarking. In secret sharing, random looking shares when brought together recreate the secret. In recursive secret sharing, the shares themselves have components defined at a lower recursive level[12] Sandeep katta also says Secret sharing techniques belong to the larger area of information hiding that includes watermarking. In secret sharing, random looking shares when brought together recreate the secret. In recursive secret sharing, the shares themselves have components defined at a lower recursive level. The injection of the random bits in the shares may be done conveniently using d-sequences or other random sequences.[13]

Chih-ching thien and ja-chen Lin says a user-friendly image-sharing method for easier management of the shadow images. The sharing of images among several branches using the proposed method has several characteristics 1.fast transmission among branches 2.fault tolerance 3.a secure storage system 4.reduced chance of pirating of high-quality images and 5.most importantly, the provision to each branch manager an easy-to-manage environment.[14] R.w.eglese says simulated annealing algorithm and the physical analogy on which it is based. Some significant theoretical results are presented before describing how the algorithm may be implemented and some of the choices facing the user of this method.[15] The rest of the paper is organized as follows Proposed method, Encryption process, Decryption process, Experimental result and conclusion.

II. PROPOSED METHOD

Visual Cryptography is encryption technique to encrypt an image in such a way. In previous method the encrypted image outcome size is large when compare to input size of the image. Encrypted image is again can be decrypted by stacked together “Revealing” operation operation. In visual cryptography the decryption process finished the reconstructed image has been affect two main problem. The two problems are Image quality of the reconstructed image, and pixel expansion. In the above mentioned problem had been solve my proposed method.

First the original secret image is encrypted into n number of shares by Visual cryptography based on polynomial function techniques with no pixel expansion. Then the shares are revealed .Then the shares again decompressed. Based on this process we can improve the display quality of the recovered image as well as the security of the recovered image.

III. ENCRYPTION PROCESS

In visual cryptography two major processes have been involved the Encryption process and decryption process. The encryption only most steps have been involved to encrypt an image. But the decryption process little steps involve to decrypt an image. Below show a proposed diagram for entire process for this method. In visual cryptography original image is given input to the encryption process, then the image in encrypted by using polynomial function techniques.

Input images are gray level images that pixels have grey levels ranging from 0(represent a white pixel) to g-1(represent a black pixel).In visual cryptography based polynomial function techniques split secret images into n number of the shares. In the encryption process image is encrypted based on the black pixel and white pixel, the image is encrypted two shares with black and white pixel (shadow images).

STUDENT

Original image

Share.1

Share.2
In the field of visual cryptography, we have to decrypt the image with quality and no pixel expansion has become a challenge. To face these kinds of challenge in this paper, we have proposed a new visual cryptography scheme. Encryption process the input image is separate into number of shares using polynomial function techniques. Once share has been created the splitted shares are revealed. In the decryption process the reverse process has been taking place to retrieve the original image. Using this proposed method there is no pixel expansion and original quality of image is reconstructed and proved in the experimental result. Based on this techniques recovered image quality and also security.

ACKNOWLEDGEMENT

I would like to thank reference authors and also like to thank the anonymous reviewers, whose comments and suggestions have helped them to improve the quality of the original manuscript.

REFERENCES

[4] Rezvan dastanian et al “Multi secret sharing scheme for encrypting Two secret images into two shares” 2011 International conference on information and electronics engineering” IPCSIT vol.6 (2011)
[11] Hiroki koga “A general formula of the (t; n)-threshold visual secret sharing scheme,” University of Tsukuba.
[12] Eric r. verheul et al “Constructions and properties of k out of n visual secret sharing schemes” Design and cryptography.
Table 1. The details of pixel in secret image (share1, share2, share3) and reconstructed table

<table>
<thead>
<tr>
<th>Image</th>
<th>Total columns In Image</th>
<th>Total Rows In Image</th>
<th>Total Black pixels In Image</th>
<th>Total white Pixels In Image</th>
<th>Total pixels In Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original image</td>
<td>245</td>
<td>50</td>
<td>3928</td>
<td>8322</td>
<td>12250</td>
</tr>
<tr>
<td>Share1 Of the image</td>
<td>245</td>
<td>50</td>
<td>5998</td>
<td>6252</td>
<td>12250</td>
</tr>
<tr>
<td>Share2 Of the image</td>
<td>245</td>
<td>50</td>
<td>6042</td>
<td>6208</td>
<td>12250</td>
</tr>
<tr>
<td>Share3 of the image</td>
<td>245</td>
<td>50</td>
<td>6104</td>
<td>6116</td>
<td>12250</td>
</tr>
<tr>
<td>Share1+2+3 of the image</td>
<td>245</td>
<td>50</td>
<td>10116</td>
<td>2134</td>
<td>12250</td>
</tr>
</tbody>
</table>
LEAD/LAG COMPENSATOR DESIGN FOR UNSTABLE DELAY PROCESSES FOR LOW AND HIGH ORDER SYSTEM USING GAIN AND PHASE MARGIN SPECIFICATIONS

SACHIN B. CHARBHE1 & SHARAD P. JADHAV2

1Electronics Department, Mumbai University, Ramrao Adik Institute of Technology, Dr. D. Y. Patil Vidyhanagari, Nerul Sec-7, Navi-Mumbai.
2Instrumentation Department, Mumbai University, Ramrao Adik Institute of Technology, Dr. D. Y. Patil Vidyhanagari, Nerul Sec-7, Navi-Mumbai.

Abstract - This project considers Lead/Lag compensator design problem for a class of unstable delay processes based on a new set of gain and phase margin specifications. Due to the nature of the unstable system, both upper and lower gain margins are required to measure the true stability robustness with regard to gain change. In addition to phase margin, such combined margin problem leads to set of nonlinear and coupled equations which have no analytical solution. Thus, an effective graphical method is developed such that the solution is determined from the intersections of the curves constructed by transformed set of nonlinear equations. The resulting design problem leads to three nonlinear and coupled complex equations which have no analytical solution in general. To find the solution from some intersection points of suitable constructed curves based on the frequency response of the process. The proposed method is given with Low and High order unstable time delay process with example and results.

Keywords - Control system, Unstable time delay process, Nonlinear equation, Graphical tuning method, Gain margin and phase margin.

I. INTRODUCTION

One of the prime concerns of our project is to lead/lag compensator design for a class of unstable delay processes based on new set of Gain and Phase margin. Automatic control System has played a vital role in the advance of Engineering and science. In addition to its extreme importance in Space-Vehicle System, missile-guidance Systems, Robotic Systems. And the like, Automatic control has become an importance and integral part of modern manufacturing and industrial processes. It is also essential in such industrial operations as controlling Pressure, Temperature, Humidity, Viscosity and Flow of Process industries. Generally unstable processes are encountered in industry, and examples include Batch chemical reactors and the combination of Feed/Effluent heat exchanger. They are known to be very difficult to control, because of the nature of the open loop instability.

Generally the purpose of the Lead-Lag compensator is to create a controller which has an overall magnitude of approximately 1. The lead-lag compensator is largely used for phase compensation rather than magnitude. A pole is an integrator above the frequency of the pole. A zero is a derivative above the frequency of the zero. Adding a pole to the system changes the phase by \(-90^\circ\) and adding a zero changes the phase by \(+90^\circ\) deg. So if the system needs \(+90^\circ\) added to the phase in a particular frequency band then you can add a zero at a low frequency and follow that zero with a pole at a higher frequency[19].

Stability robustness is a key issue in the process control and is often measured by gain and phase margins. The controller design based on gain and phase margin specifications is common and is widely used in practice. There have been reported works on gain and phase margin design methods for the tuning of lead/lag compensators, but they were developed only for the stable processes. figure 1.2 Shows Stable system [13][8].

Unstable processes are encountered in industry, and examples include batch chemical reactors and the combination of a feed/effluent heat exchanger. They are known to be very difficult to control, because of the nature of the open loop instability. A common control scheme for such processes consists of inner and outer loops, which is shown in figure 1.1. The PD controller in the inner loop plays the role of stabilizing a given unstable process. Then, the PI controller of the outer loop is tuned for performance. As all know, the PD controller (kp + kds), is not a proper transfer function and thus not physically realizable.
In order to obtain a proper and thus physically realizable controller, one has to multiply with a first-order filter $1/Nkds+1$ where $0 < N \leq 0.01$, to the derivative term of the PD controller, so that the practical D action becomes $kds/Nkds+1$. Then the resulting implementable PD controller is given by $(Nkp+1)kds/Nkds+1$, which is actually in the form of a lead/lag compensator. It is more reasonable and accurate to use the lead/lag compensator in the inner loop and design it for proper stability margins without approximations, leaving the outer loop to cater to performance as usual. Note that for the control of an unstable process, the first and foremost problem is to stabilize it. No performance can be achieved without closed loop stability. Therefore, the stabilization of an unstable delay process by the lead/lag compensator is inadequate to measure stability robustness for an unstable process, see figure 1.3, where the left intersection point of the open-loop Nyquist curve with the real axis also contributes to the stability of the closed loop. Without knowing its location, no stability or margin can be determined. Hence, we here propose for the first time that the gain margin specifications. For an unstable process should be two-sided an upper limit ($\Gamma A$) for gain increase, which is the normal-sense gain margin, and a lower limit ($\Delta A$) for gain decrease, which is normally called gain reduction. This new formulation leads to a stabilizing gain interval ($\Gamma A, \Delta A$), meaning that if the open-loop transfer function is multiplied by a gain within this interval, the closed loop is guaranteed to remain stable. This is different from a normal stable process case, as shown in figure 1.3[19].

In this paper, we formulate controller design problem for Unstable Delay Processes for low and high order system based on Gain and Phase Margin specifications. The rest of the paper is organized as follows. Problem Formulation described with equation and design tuning Method for low and high order unstable time delay system developed in section II, MATLAB Implementation with design step given in section III, Results of example low and high order unstable time delay system IV, finally Conclusion and future scope.

II. DESIGN METHOD

we formulate a controller design problem for unstable delay processes based on specifications of a gain margin, a gain reduction, and a single phase margin and present an effective graphical tuning method for lead/lag compensators which is to be placed in the inner loop of the control system for stabilization purposes[19].
\[ G(jwGC)(jwGC) = e^{j\phi_m} \text{(PhaseMargin)} \]  
(2.5) 
\[ Kc[(1+jwp1T1)/(1+jwp1T2)]*G(jwp1) = -1/A^- \]  
(2.6) 
\[ Kc[(1+jwp1T1)/(1+jwp1T2)]*G(jwp1) = -1/A^+ \]  
(2.7)

where \( wp1 \) and \( wp2 \) are two phase crossover frequencies and \( wg \) is the gain crossover frequency.

Three equations involve six unknowns, that is three compensators parameters and three crossover frequencies. Since an unstable pole is present in the process equation 2.1, an anticlockwise encirclement of the critical point by the open-loop Nyquist curve is necessary for closed-loop stability, according to the Nyquist stability theorem. In figure 1.2, 1.3 that in order to form such an anticlockwise encirclement, the three frequencies should satisfy \( wp1 < wg < wp2 \).

Graphical Tuning Method: A graphical tuning method is useful to solve some nonlinear problems for which no analytical solution can be derived. The typical open-loop Nyquist curve of \( G(s)C(s) \) is given in figure 2.2, 2.3. There are two cases for its initial part from zero frequency. One is for \( m = 0 \) with \( wp1 = 0 \), as shown in figure 2.2. In this case, the Nyquist curve starts from the real axis and moves immediately downward in the anticlockwise direction to encircle the critical point[13].

\[ \text{Figure 2.2:} \; G(s)C(s) \text{ with } wp1 = 0. \]

Since the phase of \( G(jw)C(jw) \) is always larger than \( \pi \) at frequency \( w_\pi(wp1, wg) \), we have
\[ \angle G(j\phi_\phi C(j\phi_\phi) + \pi]) < \angle G(jwp1C(jwp1)) + \pi] > 0 \]  
(2.9)

The other case is \( wp1 = 0 \), as shown in figure 2.3, where the Nyquist curve first moves upward in the clockwise direction for \( m = 0 \), or \( m = 1 \) in equation 2.1. The phase of \( G(jwC(jw)) \) is always larger than \( \pi \) at frequency \( w_\pi(wp1, wg) \) which leads to
\[ \angle G(j\phi_\phi C(j\phi_\phi) + \pi]) \angle G(jwp1C(jwp1)) + \pi] < 0 \]  
(2.10)

consider the case of \( wp1 = 0 \). Submit \( wp1 = 0 \) in to equation 2.6 then we get, \( Kc = -1/\Delta \) \( G(0) \)  
(2.11)

Complex equation 2.7 and 2.8 are equivalent to the following four real equations

\[ (1 + T1wp2)/(1 + T2wp2) = -1/(AKcG(jwp2)) = a + jb \]  
(2.12)

\[ 1 + T1wp2 = (a + jb) \cdot (1 + T2wp2) \]  
(2.13)

\[ 1 + T1wp2 = (a + aT2wp2 + jb - bT2wp2) \]  
(2.14)

Compare real part.

\[ a = a - bT2wp2 \]  
(2.15)

\[ \text{Figure 2.3:} \; G(s)C(s) \text{ with } wp1 = 0. \]

Re\[ -1/(KcAG(jwp2)) - T2wp2Im\{1/(KcAG(jwp2)) \]  
(2.16)

Im\[ -1/(KcA \cdot G(jwp2)) + T2wp2Re\{-1/(KcA \cdot G(jwp2)) \]  
(2.17)

Re\[ \exp(\{jwp\}KcG(jwp)) \]  
(2.18)

Im\[ \exp(\{jwp\}KcG(jwp)) \]  
(2.19)

It shows that equation 4.16 to 4.19 are linear equations with regard to \( T1 \) and \( T2 \). We solve two equations each to express \( T1 \) and \( T2 \) as functions of \( wp2 \) and \( wg \) respectively, as follows

\[ T1 = (1/\{KcA\})Im\{-1/(G(jwp2))\} + Re\{-1/(G(jwp2))\}Re\{-1/(G(jwp2))\} \]  
(2.20)

\[ T2 = (1/\{KcA\})Re\{-1/(G(jwp2))\} - 1/(wp2/\{KcA\}) \]  
(2.21)

\[ T1 = (1/\{Kc\})Im\{-\exp(\{jwp\}/G(jwp))\}Re\{-\exp(\{jwp\}/G(jwp))\} \]  
(2.22)

\[ T2 = (1/\{Kc\})Re\{-\exp(\{jwp\}/G(jwp))\} \]  
(2.23)

One set of equations 2.20,2.21 gives a relationship between \( T1 \) and \( T2 \) via intermediate variables, \( wp2 \) and \( wg \), and can be plotted as a curve in a plane with \( T1 \) as the horizontal axis and \( T2 \) as the vertical axis. We do so for two sets in equation 2.20 and 2.21 and plot the resulting two curves in the same plane. The intersection point will meet both sets, which is the solution of our design problem. The values of designed \( T1 \) and \( T2 \) can be read off directly from the intersection point. The corresponding intermediate variables \( wp2 \) and \( wg \) are obtained as well. The computational efforts can be substantially reduced by limiting possible frequency ranges for \( wp2 \) and \( wg \) via which equations 2.20, 2.21 and 2.22, 2.23 are plotted. The controller equation 2.2 with
non-negative parameters can only give a phase change in the range from $-\pi/2$ to $+\pi/2$. It follows from equations 2.7 and 2.8 that the solution will occur only in the frequency ranges of $w_2$ and $w_g$ bounded implicitly by
\[-3\pi/2 < G(jw_g) < -\pi/2\] (2.24)
\[-3\pi/2 < \phi_m < G(jw_2) < -\pi/2 + \phi_m\] (2.25)

Thus, we plot equations 2.20, 2.21 and 2.22, 2.23 with respect to $w_2$ and $w_g$ in the ranges given in equations 2.24 and 2.25.

Similarly, from equations 2.7 and 2.8 as done above, we transform the complex equation 4.6 in to the two real equations
\[Re[-1/(KcAG(jwp))] - T2wp11m[-1/(KcAG(jwp))] = 1\] (2.26)
\[Im[-1/(KcAG(jwp))] + T2wp1Re[-1/(KcAG(jwp))] = wp1T1\] (2.27)
and solve for $T1$ and $T2$ in terms of $wp1$ as
\[T1 = [(1/(KcA)^2[−1/G(jwp1)] + Re^2[−1/G(jwp1)]) - Re[−1/] G(jwp1)])/(wp1Im[−1/G(jwp1)]\] (2.28)
\[T2 = [(1/(KcA)Re[−1/G(jwp1)] - 1)]/(wp1/KcA)Im[−1/G(jwp1)]\] (2.29)

And the possible frequency range of $wp1$ for the solution is limited by $−3\pi/2 < G(jwp1) < -\pi/2$ (2.30)

However, unlike the case of $wp1 = 0$, this new case leaves Kc unknown yet. Thus, to find a suitable Kc in order to plot equations 2.20, 2.21, 2.22, 2.23 and 2.28, 2.29 in the same plane to enable us to find the solution if any. To this end, we choose a stabilizing gain as a good initial value and then refine it through some simple iteration. To set the initial value for Kc, one sees in the Introduction that the PD controller is chosen Kc=kp is taken as the initial value.

Theoretically, we can design a PD controller $kp(1 + kds)$ to stabilize the process equation 2.1. Then choose $Kc=Kp$ is taken as the initial value.

With the above chosen value of Kc, one can plot the graphs of equations 2.20,2.21,2.22, 2.23 and 2.28, 2.29 in the same diagram. Three curves usually do not have a common intersection but yield three intersections each for a pair of curves. Thus, we need to adjust Kc in such a way that the three intersections move close to each other and eventually merge to one point, which is the desired solution. It is easily seen that the functions defined by equations 2.20, 2.21, 2.22, 2.23 and 2.28, 2.29 have the same structure but with different parameters $A$, $A$ and $exp(j\phi_m)$, implying that the three curves will move in the same direction when Kc changes.

In addition, since the parameters satisfy $1/A < |exp(j\phi_m)| < 1/A$ the curve for equations 2.28, 2.29 will move faster than other two with the same amount of change of Kc. In the $T1/T2$ plane, the curve for equations 2.28, 2.29, which achieves gain reduction $A$ exactly, will divide the plane into two parts. One part is larger than $A$, denoted by $A+$, and the other part is smaller than $A$, denoted by $A−$. Note that changes of Kc and have the A same effect on equations 2.28, 2.29. Thus, the curve for equations 2.28, 2.29 will move to the $A−$ side when Kc decreases. If the intersection of equations 2.20, 2.21 and 2.22, 2.23 lies on the $A+$ side, one should reduce Kc, otherwise one should increase it. In this way, the curve for equations 2.28, 2.29 will move closer to the intersection of equations 2.20, 2.21 and 2.22, 2.23 until the three intersections merge to one.

For the stabilizable case, we find that the phase margin between 15 and 30 is appropriate for an unstable process, that is, half for the stable case. Calculate the stabilizing gain range for a given unstable process and choose suitable gain margins for the given controller. With the chosen gain and phase margins our design method is invoked to find the controller solution[19]

**Lead/Lag Compensator Design:**

Consider the PI controller given by the Transfer function:
\[G(s) = Kp + Ki/s\] (2.31)

The closed loop transfer function $TF$ of inner loop of fig1.3 is defined by the plant $P(s)$. Assume Gain margin ($Gm$) of outer loop as $Am1$ and Phase margin as $exp(−j\phi_m)$. The gain and phase crossover frequency are defined by variable $wgc1$ and $wp1$. The open loop transfer function at $wgc1$ is
\[G1(jwgc) = P1(jwgc1) = exp(−j\phi_m)\] (2.32)

The open loop transfer function at $wp1$ is
\[G1(wpj) = P1(wpj1) = −1/Am1\] (2.33)

The frequency response of PI controller is
\[G1(jw) = Kp + Ki/jw\] (2.34)
\[G1(jw) = Kp − jKi/w\] (2.35)

Replace equation 4.35 in equation 4.32 and equation 4.33.

\[(Kp − jKi/wgc1)P1(jwgc1) = exp(−j\phi_m)\] (2.36)
\[(Kp − jKi/wp1c1)P1(wp1c1) = −1/Am1\] (2.37)

separate real and imaginary part of equation 4.36.
\[Kp = Re[exp(−j\phi_m)/P1(jwgc1)]\] (2.38)
\[Ki = −wgc1Im[exp(−j\phi_m)/P1(jwgc1)]\] (2.39)

Similarly separating real and imaginary part of equation 4.37.

\[Kp = Re[−1/Am1P1(jwgc1)]\] (2.40)
\[Ki = −wp1c1Im[−1/Am1P1(jwgc1)]\] (2.41)

In equation 2.38 and 2.41, the unknown are $Kp$, $Ki$ and $wgc1$.

By comparing equation 2.38 and 2.40, we get $wgc1$. By replacing $wgc1$ in equation 2.39 and 2.41, we gate $Kp$ and $Ki$.
compensator design for a class of unstable delay processes based on new set of Gain and Phase margin specifications. Consider the process given by $G(s) = \frac{1}{(0.5s + 1)(2s - 1)}e^{-0.5s}$ (2.42).

The gain and phase margins are set to (0.5, 2) and 37°, respectively. Assume the case of $\omega_p = 0$ first. $K_c = 2$ is computed by equation 2.11. The graphs of equations 2.20, 2.21 and 2.22 are plotted with respect to $\omega_p$ and $\omega_g$, which gives the resultant compensator as $C(s) = 2[(0.74385 + 1)/(0.14075 + 1)]$ (2.43).

Obviously, the solution satisfies $T_1 > 0, T_2 > 0$, and $\omega_p1 < \omega_g < \omega_p2$, and the process frequency response has a monotonic decreasing magnitude over the frequency. Under the control structure in figure 1.1, we use $C(s)$ as the inner loop controller and design the PI controller for the outer loop. Which is also based on the performance specifications in terms of gain and phase margins. Set the gain and phase margins to be 3 and 60°, respectively. The proposed method gives a much better response in terms of a smaller overshoot, shorter settling time, and better disturbance rejection [9][12].

**Design Example High-order unstable time delay process.** The Lead/Lag compensator design for High-order unstable time delay process is carried out in two stages. The inner, outer loop feedback control as shown in figure 1.1. We use $C(s)$ as the inner loop Lead compensator (PD) and design the Lag compensator (PI) for outer loop and overall transfer function design lead-lag compensator. Consider a high-order unstable time delay process $G(s) = \frac{(0.5s + 1)(s - 1)}{(5s + 1)(5s - 1)}e^{-0.2s}$ (2.44). The existing control design methods for unstable processes in the literature usually assume a low-order one and need to use model reduction for their applications. The proposed method has no restriction on process order and it utilizes frequency response directly without a need for any approximation. This example demonstrates such an advantage. Suppose the gain and phase margins to be (0.7, 2), and 20°, respectively. Assume the case of $\omega_p1 = 0$ first. We obtain $K_c$ from equation 2.11. Plot the graphs of equation 2.20, 2.21 and 2.22, 2.23 which gives the intersection point. The two additional crossover frequencies are $\omega_g$ and $\omega_p$, which satisfy $\omega_p1 < \omega_g < \omega_p2$. It is indeed the case of $\omega_p1 = 0$ and the above intersection produces the solution controller. The outer-loop PI controller is designed under gain and phase margins of 2 and 50°. The proposed method gives a much better response in terms of a smaller overshoot, shorter settling time, and better disturbance rejection [7][5].

**III. MATLAB IMPLEMENTATION.**

The Lead/Lag compensator design for Low-order unstable time delay process is carried out in two stages. The inner, outer loop feedback control as shown in figure 1.1. We use $C(s)$ as the inner loop Lead compensator (PD) and design the Lag compensator (PI) for outer loop and overall transfer function design lead-lag compensator.

**Design Steps.**
1. Design of inner lead compensator which will guarantee about stability of the system.
2. Design of outer lag compensator to satisfy the designed specifications.
3. Obtain the intersection point of both the curves $T_1$ and $T_2$.
4. Obtain the corresponding Gain crossover $\omega_{gc}$ and Phase crossover $\omega_{pc}$.
5. Design lead compensator given by equation 2.2 using above designed values.
6. Simulate the designed lead compensator with simulink environment of MATLAB.

**Design specification.**
1. Gain margin specification. The Gain margins are set to (0.5, 2).
2. Phase margin specification. The Phase margins are set to 37°. The design steps for Outer Lag compensator (PI) are as follows
1. Obtain closed loop transfer function of inner loop.
2. Assume Gain and Phase margin for outer loop.
3. Find Gain crossover frequency ($\omega_{gc}$) by comparing equations 2.38 and 2.40.
4. Obtain $K_p$ by equation 2.37.
5. Obtain $K_i$ by equation 2.41.
6. Simulate the designed lag compensator with simulink environment of MATLAB.

The design steps for Lead-Lag compensator are as follows
1. Obtain overall transfer function.
2. The Gain margins are set to (0.5, 2).
3. The Phase margins are set to 37°.
4. Find Gain crossover frequency ($\omega_{gc}$).
5. Find Phase crossover frequency ($\omega_{pc}$).
6. Obtain $K_c$ by equation.
7. Obtain $dc$ by equation.
8. Simulate the designed lead-lag compensator with simulink environment of MATLAB.
IV. RESULTS

Figure 4.1: Plot of function equations 2.20, 2.21, 2.22 and 2.23

Figure 4.2: Step responses for Design example (PD)

Figure 4.3: Step responses for Design example (PD)

Figure 4.4: Step responses for Design example lead-lag compensator

Figure 4.5: Block diagram for Design example lead-lag compensator

Figure 4.6: Plot of function equations 2.20, 2.21, 2.22 and 2.23 lead compensator(PD)

Figure 4.7: Step responses for Design example lead compensator(PD)

Figure 4.8: Step responses for Design example lead-lag compensator

V. CONCLUSIONS

A simple graphical method is proposed for the design of lead/lag compensator for the unstable system. This method does not require solution of nonlinear equation to obtain the compensator parameters. From the results, it is observed that the designed lead (PI) compensator guarantees about stability and lag (PI) compensator is used to achieve design specifications. The design method is successfully applied to lower and higher order unstable time delay systems. The designed compensator guarantees about stability of the system. MATLAB implementation of Design Example Low and High-order unstable time delay process are shown using Gain margin and phase margin specification.

Future scope
1. The method can be extended to non minimum phase and unstable time delay system.
2. The method can be extended to MIMO unstable system.
REFERENCES


Abstract— Educational knowledge is broadly distributed on Internet. In this project, I give an overview of using Natural language processing on ontology based distributed Educational knowledge management. Natural language processing (NLP) is becoming much more robust and applicable in realistic applications. One area in which NLP has still not been fully exploited is information retrieval (IR). In particular users are interested in search over intranets and other local Web sites. User see dialogue-driven search which is based on a largely automated knowledge extraction process as one of the next big steps. Instead of replying with a set of documents for a user query the system would allow the user to navigate through the extracted knowledge base by making use of a simple dialogue manager. I found that users are willing to interact and use those visual interfaces. I also found that users preferred such a system that guides him through the result set over a baseline approach. The information can change any day and the idea is to have always the most up-to-date facts and relations available to assist a searcher. Currently, we do not have systems which support this type of interaction. However, my aim is to automatically acquire knowledge (a domain model) from the document collection and employ that in an interactive search system. This research combines the advantages of both NLP and IR, which will represent an important first step towards full NLP-driven intranet search.

Keywords—Natural language processing, ontology, information retrieval, artificial intelligence.

I. INTRODUCTION

Natural language processing (NLP) is a subfield of artificial intelligence and linguistics. It studies the problems of automated generation and understanding of natural human languages. Natural language generation systems convert information from computer databases into normal-sounding human language, and natural language understanding systems convert samples of human language into more formal representations that are easier for computer programs to manipulate. Natural Language Processing is the artificial intelligent concept where the machines understand Natural Languages like English, Korean, French, Telugu, Hindi etc. We are going to develop a tool that will take the Database queries in the form natural language and then processes it and gives the result. This includes many sub components like Language Analyzer, Query Builder and Viewer. The system will first parses the query in natural language and finds the major parts in the string. Then first it will look for the table name and then it parses the string for the where clause and then for the order by clause. After parsing it will construct the query string based on the data available. The generated SQL query is posted to the database to fetch the results.

Motivation for making use of NLP and ontology model for IR:
Imagine we could interact with a university intranet search engine just like with a human person in a natural dialogue. The search engine would automatically extract knowledge from the Web site so that a searcher can be assisted in finding the information required. A student who asks for a particular course can be directed to the most recent lecture notes or the contact details of the lecturer. An external searcher typing in “PhD NLE” could be assisted by allowing him to explore the space of experts and projects available in the area of natural language engineering. Obviously, this information can change any day and the idea is to have always the most up-to-date facts and relations available to assist a searcher. Currently, we do not have systems which support this type of interaction. However, our aim is to automatically acquire knowledge (a domain model) from the document collection and employ that in an interactive search system.

One motivation for a system that guides a user through the search space is the problem of “too many results”. Even queries in document collections of limited size often return a large number of documents, many of them not relevant to the query. Part of the problem is the fact that both on the Web and in intranet search queries tend to be short and short queries always pose ambiguity and uncertainty issues for information retrieval systems [1]. Some form of dialogue based on feedback from the system could be very useful in helping the user find the right results. This combination of NLP and IR we assume is particularly promising and scalable in smaller domains like university intranets or local Web sites.

II. LITERATURE SURVEY

A. Conventional IR V.S. Intelligent IR

Traditionally, information retrieval emphasizes document retrieval which is very much dependent on
human classification and the use of humanly prepared searching strategies, while intelligent information retrieval emphasizes automatic extraction of useful information and facilitates interaction between the user and facilitates interaction between the user and system by giving the user natural language access tools.

B. Conventional Information Retrieval:
In principle, the conventional information retrieval (CIR) systems based on determining the relationships relevant or no relevant between the information need of the users and the information in the documents. The outline of conceptual IR is shown in Fig.1 (a). The stored documents need to be organized and controlled. Organization and control activities include classification, cataloging, subject indexing and abstracting [11]. In the early CIR, those activities is very much dependent on information specialists, who need both an understanding of what the document is about, that is, some comprehension of its subject matter and a good knowledge of the user’s need. An expanded view of the retrieval operations shown in Fig.1 (b) As depicted in Fig.1 (b), queries or information need of the users need to be analyzed and translated into particular vocabularies. Usually, the output of CIR consists of one or more bibliographic references with some added information such as an abstract or full text of documents. In operational environments, the stored documents are represented by sets of index term, sometimes called term vectors. Usually the terms are unweighted, although in some retrieval situation each term may be assigned a weight to reflect its relative importance. Queries may similarly be expressed by using sets of unweighted or weighted terms. In many practical systems, the query terms are joined by Boolean operator. Based on an inverted index, the indexed document set corresponding to a given query formulation is easily determined. Various extensions to the standard inverted-index technology have been proposed. Those extensions are distance constraints, term weights, synonym specification and term truncation. More recently, because of the explosion of non bibliographic databases and free style queries requested directly from naïve online searchers, the intelligent information retrieval which is controlled by automatic, machine performed procedures is needed. [10]

C. Intelligent Information Retrieval:
The existence of non bibliographic databases and online information retrieving by computer users pose many problems for information access system, such as, the language for document representation, command language, database selection, problems relating to “user friendliness and ease for use”. Since no bibliographic databases outnumber their bibliographic counterparts, the need for automatic IR (or Intelligent IR) increases.

![Figure 1: Conventional Information Retrieval](image)

Intelligent Information Retrieval (IIR) differs from the CIR systems in that they must be more flexible user-friendly and responsive, automatic indexing and classifying, possibly segmenting, combining or synthesizing a response rather than just retrieving documents, and possibly extracting useful information[10].

CIR is rigid, inflexible but fast, portable, relatively inexpensive, relatively ease to learn. IIR is highly expressive, flexible but potentially ambiguous, slow, brittle and expensive.[10]

D. Natural Language in IIR:
An information retrieval system without vocabulary control may be referred to as a “natural language” (NL) or “free-text” system. Natural language systems become both more prevalent and more feasible because of the explosion of non bibliographic databases. In addition to free text, IIR might concern with NL queries which are attractive to naïve users who don’t want to learn an artificial query language, which includes Boolean operators, proximity and truncation. Even though natural language provides highly degree of exhaustibility, highly flexible and user friendly IR system, it contains many problems, of language due to synonymy, unknown word, ill-formed sentences, syntactic ambiguities caused by structure, semantic ambiguities caused by homographs, and contextual ambiguity [11]. The
searching of NL is, then, more difficult than that of controlled vocabularies. However, NL searching offers a number of benefits. That is, it permits the conduct of searches of unlimited specificity.

Table 1 Advantages and Disadvantages of NL system.[12]

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>highly expressive</td>
<td>very difficult to make generic searches</td>
</tr>
<tr>
<td>permits a variety of access points</td>
<td>problem with synonyms</td>
</tr>
<tr>
<td>highly flexible</td>
<td>partial matches to language</td>
</tr>
<tr>
<td>highly representative in reality</td>
<td>partial listings of ideas</td>
</tr>
<tr>
<td>possesses long memory periods of ideas</td>
<td>ambiguous, lengthy, ill</td>
</tr>
<tr>
<td>requires no strong to get</td>
<td>not standardized</td>
</tr>
<tr>
<td>every input is mapped into simple concepts</td>
<td>not very explicit</td>
</tr>
<tr>
<td>freedom of expression</td>
<td>use mentioning of own search term, synonym etc.</td>
</tr>
<tr>
<td>high generalization</td>
<td></td>
</tr>
<tr>
<td>limited by necessity</td>
<td></td>
</tr>
</tbody>
</table>

E. Role of Natural Language Processing in IIR

Document and queries in CIR are generally represented by document description and artificial query language, respectively. Document descriptions are described by single words or sometimes groups of words, extracted from the document texts. Artificial query languages are processed by Boolean operators, proximity and truncation. The retrieval of relative items depends on matching computation. Accordingly, no refined language analysis and understanding are necessary to be included in the information retrieval system. However, documents are necessarily available in the form of natural language or free text which needs IIR for deriving IR’s power. In IIR environment, a linguistic component could indicate term relationships between indexing units, which might be used to generate more complete and representative document descriptions than can obtained from single terms alone. More refined content analysis might improve retrieval effectiveness, and enhance traditional retrieval techniques with question-answering and information extraction capabilities that provide specific responses to certain questions. Additionally, a natural language capability might facilitate interaction between user and system by giving the user natural language access tools to replace conventional formalized query languages. To implement IIR, the marriage between IR system and AI techniques, which include Natural Language Processing, knowledge representation, reasoning and inference, is then natural and potentially important.

An overview of AI-based IR is given in[12,13,14,15]

F. Why would we use Ontology over a relational database (RDB)?

The use of a relational database is to make the following assertion – I understand the data that exists in my domain completely, and that data is relatively static. That is not to say that changes may never happen, but the design of an ERD must remain relatively static for applications to effectively build on top of it. The use of an Ontology model (or models, there is no constraint toward using a single Ontology in an enterprise or industry) is to make the opposite assertion – I do not fully understand the data that exists in my domain; I know that I’ll never understand it completely, and far from being static, the data changes constantly.

The relational model has relations between entities established through explicit keys (primary, foreign) and, for many-to-many relationships, associative entities. Changing relationships in this case is cumbersome, as it requires changes to the base model structure itself, which can be difficult for a populated database. Querying for this kind of data based on a relational model can also be cumbersome since it can result in very complicated where clauses or significant table joins.

Hierarchical models have similar limitations when it comes to real world updates and are not very flexible when it comes to trying to traverse the model “horizontally”.

The graph model, which is how semantic models are implemented, makes it much easier to both query and maintain the model once deployed. For example, if a new relationship is needed to be represented that had not been anticipated during design. With a triple store representation that additional representation is easily maintained. A new triple is simply added to the data store. A critical point is the relations are part of the data, not part of the database structure.
II. PROPOSED WORK

To provide an interface with a university intranet search engine just like with a human person in a natural dialogue. The search engine would make use of NLP and ontology model to automatically extract knowledge from the database or Web site so that a searcher can be assisted in finding the information required. A student who asks for a particular course can be directed to the most recent lecture notes or the contact details of the lecturer. A simplified searching technique will be available for the user so that he can make use of natural language to interact with educational knowledge management database without unnecessary search results. Obviously, this information can change any day and the idea is to have always the most up-to-date facts and relations available to assist a searcher by making use of ontology schema of database. Currently, we do not have systems which support this type of interaction. However, our aim is to automatically acquire knowledge (a domain model) from the document collection and employ that in an interactive search system.

The explosion in unstructured text documents has presented a formidable challenge in computer-aided knowledge extraction to professionals in all fields. And simple keyword searches, advanced Boolean searches and statistically-based methods, have long been inadequate - retrieving a list of documents that might – just maybe – contain potentially relevant information.

For innovation and technical problem solving, in particular, these conventional search technologies are a failed strategy because the process of matching keywords cannot understand the context of the user’s request – their design intent and need. The result of a traditional search technologies are piles of mostly irrelevant documents, when what engineers and scientists need are answers to questions critical to their projects that require quick, informed decisions.

Hence NLP combined with Ontology model provides an interface with a university intranet search engine just like with a human person in a natural dialogue. The search engine would make use of NLP and ontology model to automatically extract knowledge from the database or Web site so that a searcher can be assisted in finding the information required.

IV. RESEARCH METHODOLOGY

A. Steps to acquire and process

Use of PROTAGE software for creating ontology schema:

PROTAGE is a free, open source ontology editor and a knowledge acquisition system. Like Eclipse, Portage is a framework for which various other projects suggest plug-in. This application is written in Java and heavily uses Swing to create the rather complex user interface.

Firstly sub-domain ontology model of Educational Information (e.g. university management ontology) is constructed. Based on this model, I developed the semantic retrieval system. The relationships of domain concepts can be parsed from the ontology model. Retrieval results can be got from the second step retrieval from the distributed knowledge on the Internet.

B. Proposed System mapping

A prototype mapping system that support the process of semi-automatic ontology mapping for the purpose of improving semantic interoperability in heterogeneous systems. The approach is based on the idea of semantic enrichment, i.e. using instance information of the ontology to enrich the original ontologies and calculate similarities between elements in two ontologies.[17]

The word ontology has been used to describe artifacts with different degrees of structure. These range from simple taxonomies (such as the Yahoo! hierarchy), to metadata schemes (such as the Dublin Core), to logical theories. In our context, the scope and assumption of my work are the following:

(1) An ontology specifies a conceptualization of a domain in terms of concepts, attributes and relations. Concepts are typically organized into a tree structure based on subsumption relationship among concepts. Ad hoc relations further connect concepts and formulate a semantic net structure in the end.

(2) In next stage, I will focus on finding mappings between concepts and between relations. This is because they are central components of ontologies and matching them successfully would aid in matching the rest of the ontologies.

(3) Ontologies can be expressed in different representational languages. Here, we assume that it is possible to translate between different formats. In practice, a particular representation must be chosen for the input ontologies. [16]

Definition 1 (ontology mapping model)

A 5-tuple [S, T, F, R(si, tj ), A] where

• S is a set composed of logical views (representation) for the elements in source ontology.
• T is a set composed of logical views (representation) for the elements in target ontology.
• F is a framework for representing ontology elements and calculating relationships between elements in the two ontologies.
• R(s I, t j) is ranking functions which associate a real number with an element s I ∈ S and an element t j ∈ T. Such ranking defines an order among the elements in source ontology with regard to one element t j in the target ontology.
• A is a set composed of mapping assertions. A mapping assertion is a formal description of the mapping result, which supports further description of the exact nature of the derived mappings. It has the following components:
A. A pair of ontology elements,
B. A type of correspondence,
C. A degree of correspondence, and
D. A set of sources of assertion.[16]

C. Mapping Technique to be used in future development

I will propose or develop an algorithm that takes as input semantically enriched elements of ontologies and produces as output suggestions to the user for possible correspondences. The mapped performs a computation of correspondence measures for the pairs of compared ontology elements, based on the similarity of their enriched structures.

- The enhancer utilizes an electronic lexicon to adjust the similarity values that have been computed by the mapper, with the intention of re-ranking the mapping assertions in the result list.
- The presenter determines which recommendations to suggest to the user, based on the partial ordering of correspondence measures and the current configuration profile.
- The exporter translates and exports the mapping results to a desired format so that other follow up applications can import and use the results in a loosely coupled way.
- The configuration profile is a user profile to assign individual variable values for different tuning parameters and a threshold value for exclusion of mappings with low similarity.

V. ARCHITECTURE OF ONTOLOGY-BASED APPLICATION

One of the important applications of ontology models is to provide semantic query of domain knowledge. A general architecture of ontology based applications is described as figure 3. Where ontology model is key part of applications, after formal representation of ontology models, the model data were stored in the ontology persistent storage facilities. Instead of Semantic server NLP is used to parse, reason or transfer ontology models. The retrieval application is provided to the end users through interface.

Workflow:

Experiment 1: Carried out on NLP based search engine for parsing into ontology schema of temporary database.

Experiment 2: Carried out by entering user query.

Experiment 3: Parsed query using NLP and the expected result.
VI. CONCLUSION

Educational knowledge is broadly distributed on the Internet. In my Project Phase I, I developed a GUI search engine using NLP for parsing and make use of ontology schema for distributed Educational knowledge management. Key steps to implement Semantic Web to manage Educational knowledge are proposed. Using Natural language processing (NLP) is becoming much more robust and applicable in realistic applications. This work can obviously only be a first step. There are number of limitations in such a study and I will take the findings as a guideline for future work.

REFERENCES:


[16] Xiaomeng Su, Jon Atle Gulla, Department of Computer and Information Science, IDI, Norwegian University of Science and Technology (NTNU), N-7491, Trondheim, Norway, 2005.

STUDY ON INTRUSION DETECTION USING WIRELESS NETWORK

HITESH GUPTA¹, SHIVSHAKTI SHRVASTAV² & SANJANA SHARMA³

¹²³Patel Institute of Technology, Bhopal

Abstract—Now a days it is very easy to execute the network attack and tools for execute the attack are freely available on the internet with the help of basic knowledge of network and software technology. To overcome this problem intrusion detection system has been used. Number of scheme has been proposed in the research for detecting and reporting of malicious activity in wireless network. The resources consumed by ad hoc network member nodes to monitor, detect, report, and diagnose malicious activity, however, may be greater than simply rerouting packets through a different available path. This paper in view of the problems of intrusions in the wireless network.

Keywords: wireless network, IDS, MANET, intrusion detection, worm model.

I. INTRODUCTION

Wireless communication offers many benefits to users and organization such as portability and flexibility and lower installation cost. Wireless local area network allow to move your laptop place to place in your office without network connection and reduced wiring cost. Ad hoc networks, such as those enabled by Bluetooth, allow data synchronization with network systems and application sharing between devices. Handheld devices such as personal digital assistants (PDA) and cell phones allow remote users to synchronize personal databases and provide access to network services such as wireless e-mail, Web browsing, and Internet access.

Fig. 1 architecture of wireless network

Adhoc network is wireless network that have not any fixed infrastructure and have not any base station, so it can be setup and used easily. All the computers in such a network should be in the radio range of each other to make communication.

mobile adhoc network (MANET) is a group of wireless node that are distributed without relying on any standing network infrastructure. Mostly MANET is a collection of two or more devices or nodes with wireless communication and networking capabilities. these devices includes laptop, PDA and wireless phones that have some limited transmission range. Like a wireless ad-hoc network that is an infrastructure less, self-organized and not does not require any centralized administration.

The paper is organized as follows: section 2 intrusion detection technology, section 3 work on IDS, section 4 introduction of critical path detection and worm model, section 5 conclusion, section 6 references.

II. INTRUSION DETECTION TECHNOLOGY

The rapid development of internet the network has become important source of information. People rely heavily on computer and network system which makes it necessary to ensure the safety of computer and network system. Intrusion detection system secure to the network environment. Intrusion detection (ID) is the process of monitoring the events occurring in a computer system or network and analyzing them for sign of possible incidents which are unaccepted or
which are about to cause threads of violation of computer security policies/standard security attackers gaining unauthorized access to system from the internet and authorized users of systems who misuse their privileges or attempt to gain added priviledges for which they are not authorized.

An IDS is a software that automates the intrusion detection process.

**A. Types of intrusion detection system**

In the network environment there are three main types of intrusion detection system that are host based intrusion detection system, network based intrusion detection system, distributed based intrusion detection system, we can define them with the help of figure.

**Host-based IDS** - the first implementation of the intrusion detection model was classified as a host-based IDS. A host monitor monitors system logs for evidence of malicious or suspicious application activity in real time. Host-based IDS require small program/agent to be installed on individual system to be monitored. The agent supervise the os and write data to log file and active alarm. Host-based IDS can only monitor the host system on which the agents are installed. It does not monitor the entire network.

**Network-based IDS** - it consist of a network appliance/sensor with a network interface card operating in promiscuous mode (which is a configuration of a network card that makes it pass all traffic received to the kernel rather than just forms addressed to it) and a separate management interface. This IDS is placed along a network segment or boundary and monitors all the traffic on that segment. According to their adoption of different technology, intrusion detection system can be divided into two categories: anomaly detection and signature based detection.

**Anomaly detection** - anomaly based detection examines ongoing traffic, activity, transaction and behaviours in order to identify intrusion by detecting anomalies. The system administrator define the baseline of normal behavior (state of the network’s traffic load, breakdown, protocol, typical packet size etc) anomaly detector monitor network segments, compare their state with define baseline and look for current behavior that deviates from the normal.

**Signature based detection** - signature based detection uses a rule set to identify intrusion by watching for pattern of events specific to known and documented attack. It compare the information gathered against those attack signature (stored in database) to detect a match. The disadvantage of this type of detection is that if the database is not updated with regularity, new attacks would slip through.

**III. BACKGROUND**

A number of IDS techniques have been proposed in the research literature. Moreover, a number of trust building and cluster-based voting schemes have been proposed to enable the sharing and vetting of messages, and data, generated and gathered by IDS systems. Zhang and Lee describe a distributed and collaborative anomaly detection-based IDS for ad hoc networks [3, 4]. Tseng et al. describe an approach that involves the use of finite state machines for specifying correct AODV routing behavior and distributed network monitors for detecting run-time violation of the specifications [5]. Pirzada and McDonald present a method for building confidence measures of route trustworthiness without a central trust authority. The authors also present a concise summary of previous work in the area of establishing trust in ad-hoc networks [6]. Theodorakopoulos and Baras present a method for establishing trust metrics and evaluating trust [7]. Michiardi and Molva assign a value to the “reputation” of a node and use this information to identify misbehaving nodes and cooperate only with nodes with trusted reputations [8]. Albers and Camp couple a trust-based mechanism with a mobile agent based intrusion detection system, but do not discuss the security implications or overhead needed to secure the
network and individual nodes from the mobile agents themselves [9]. Sun, Wu and Pooch introduce a geographic zone-based intrusion detection framework that uses location-aware zone gateway nodes to collect and aggregate alerts from intra-zone nodes. Gateway nodes in neighboring zones can then further collaborate to perform intrusion detection tasks in a wider area and to attempt to reduce false positive alarms [10]

IV. INTRODUCTION OF CRITICAL PATH DETECTION AND WORM MODEL

Firstly in this section we know that what is critical path. It is a path whose failure or malicious behavior disconnects or significantly degrades the performance of the network. Once identified, a critical path can be the focus of more resource intensive monitoring or other diagnostic measures. If a path is not considered critical, this metric can be used to help decide if the application or the risk environment warrant the expenditure of the additional resources required to monitor, diagnose, and alert other path about the problem.

Now we define the work of IDS with critical path with the help of diagram.

![Fig. work of IDS](image)

the architecture of intrusion detection system, we first generate the test traffic using tcl script and find out the critical link in the network and then block the path, after that worm propagation model is injected in the network through critical link and find out the type of attack, tcp, udp and cbr comparison before intruder and after intruder and we also find the node who spread the malicious activity

Worm Propagation Model

Worm propagation model can be described as detailed network and abstract network. The detailed network could be an enterprise network, the whole network can be considered as detailed network in our simulation. Abstract network is also a part of the detailed network but the only difference is that it can be worm node where the infection occurrence can be assumed, so an abstract network can be called as susceptible infectious removal model. Worm node takes place in the network were we have blocked the path. The communication between detailed network and abstract network is done through actual packet transmission that is the probing packet generated by compromised node in both parts.

V. CONCLUSION

This paper based on wireless network, ad-hoc network and intrusion detection. We proposes a method to find the intruder in wires network with the help of critical path and worm model, which is main point of the study. Of course there is no doubt that it needs further more study to improve this method and to achieve intelligent intrusion detection.

REFERENCES


[2] Ketan Nadkarni, Amitabh Mishra,”A Novel Intrusion Detection Approach for Wireless Ad hoc Networks”. Available: {ketann@vt.edu, mishra@vt.edu}


SURVEY ON SECURITY ISSUES IN CLOUD COMPUTING

VISHESH P. GAIKWAD¹ & S.M NIRKHI²

¹²Computer Science and engineering Department
G.H.R.C.E, Nagpur University

Abstract— The data security in cloud is an important issue. The important data can be stored in cloud and the security of that data is totally dependent on cloud. The data might be uncovered by the malicious third party user because of wireless connection between client and cloud without proper authentication and protection. In this paper we figure out the different security issues with the cloud. When the data is stored in cloud the data should be properly managed and cloud have to provide a proper security to the data. In this paper discussing the different type of issues with the cloud and also possible policies are mentioned here that we can take care of those issues while discussing about the security provided by the cloud.

Keywords- Cloud storage, Security in cloud.

Cloud computing [1] is the use of computing resources that are delivered as a service over a network. The name comes from the use of a cloud-shaped symbol as an abstraction for the complex infrastructure it contains in system diagrams. It is cost effective development of a scalable web portals on highly available and fail safe infrastructure. The evolution of cloud can handle a massive data as per on demand service so this is a cost effective technique to store data on cloud.

Cloud [5] computing can be seen as a service-oriented architecture (SOA) explore every computing environment for providing security trust issues. Facing cloud computing, a prerequisite control measure is to ensure that a concrete Cloud computing Service Level Agreement (SLA) is put in place and maintained when dealing with outsourced cloud service providers and specialized cloud vendors. Due to the nature and demand of emerging cloud technologies, there is a certain degree of inexperience when dealing with cloud security. Currently Cloud computing clients have to trust 3rd party cloud providers on many fronts, especially on the availability of cloud service as well as data security. Therefore the SLA forms an integral part of a client’s first line of defense. The SLA thus becomes the solitary legal agreement between the service provider and client.

I. TYPES OF CLOUD

For providing a secure cloud computing solution [5], we have to decide which type of cloud are to be implemented. Mainly three types of cloud modes which are public, private and hybrid.

A. Public Cloud:

A public cloud is a model which allows users’ access to the cloud via web browsers. It is based on a pay-per-use model. It is similar to prepaid balance of mobile you can make a call if the prepaid balance is available if your balance is finished then you are not able to make a call. Cloud clients are happy pay per use manner and now this cloud is used widely in the IT infrastructure [9]. This type of cloud is less secured than the other types of cloud. Therefore trust and privacy concerns are rife when dealing with Public clouds with the Cloud SLA at its core. A key management consideration, which needs to be answered within the SLA deals with ensuring that ample security controls are put in place. One option is for both the cloud vendor and client mutually agree in sharing joint responsibility in enforcing cloud checks and validation are performed across their own systems.

B. Private Cloud
A private cloud [5] is set up within an organization’s internal enterprise datacenter. It is easier to align with security, compliance, and regulatory requirements, and provides more enterprise control over deployment and use. In the private cloud, scalable resources and virtual applications provided by the cloud vendor are pooled together and available for cloud users to share and use. It differs from the public cloud in that all the cloud resources and applications are managed by the organization itself, similar to Intranet functionality. Utilization on the private cloud can be much more secure than that of the public cloud because of its specified internal exposure. Only the organization and designated stakeholders may have access to operate on a specific Private cloud.

C. Hybrid Cloud

Hybrid cloud [5] is a combination of two or more cloud. The other clouds are bound together offering the benefits of others into one hybrid cloud. It provides a lack of flexibility, security for in-house application. It provides flexibility in in-house application with fault tolerance and therefore it is not scalable.

II. SECURITY BETWEEN CLOUD CLIENT AND CLOUD SERVICE PROVIDER

Knowing that there are possibilities for security and trust issues on both sides of the cloud customer-provider relationship allows us to separate what each side should do to build a secure system. This is a paradigm shift from a traditional model where software and computing resources were both provided in-house. While in the internal model, system and network security was mostly handled by the system and network engineers, so even an insecure piece of software would only be accessible to people within the company and particularly [12]Layers and Obligations for Cloud Security:

In the cloud model [9], the network engineers are not concerned with these problems and it is up to the cloud customer to protect their data.

Note that there is some overlap where the two meet Take Figure 2 as an example. Here the provider is providing the platform as a service to the customer. The customer is responsible for writing software that runs on top of the platform and for ensuring the security of data up to the point it is given to the platform service. This should include encrypting the data if the cloud provider does not do a satisfactory job of protecting the data (i.e. using a weak cipher, the details of which should be disclosed when signing up for the service). The provider is responsible for securing the infrastructure (network connectivity, physical machines, and platform environment). When the underlying provider services meet the customer software implementation there should be a clear, well defined interface for transferring data

III. RELATED WORKS

SECURITY ISSUES & POLICIES IN CLOUD COMPUTING

Cloud computing is a new computing model, this model have a different cloud computing issues:

A. Security issue in cloud computing

a) In Cloud computing boundaries are not defined clearly for protecting the device user, the traditional computing model can protected device user by dividing physical and logical security zones.

b) Service security issues. [6]

The service provider controlled the data and other resources. If the providers not providing proper security to data stored in cloud. Then the confidentiality of data should be reduce.

c) Protection for user data:[1]this issue is about to protecting data storage, data recovery, data integrity data encryption and the address of the user.

d) The number of user’s changes dynamically, as well as user uses the different services, leading the user can not be classified.

e) The cloud service provider[3]has more rights than the user, therefore we have to manage the balance between the service provider and service provider becomes a problem

f) Cloud computing [11] have a complex structure and user can dynamically changes in the cloud environment, thus the security and integrity is an important issue to be considered.

Figure 2: Example of the separation of security concerns between a PaaS customer and provider.
g) A multi-tenancy trusted computing environment model (MTCEM)[2] is designed for IaaS delivery model, main aim is to assure a trusted infrastructure to a client. It provide a two level hierarchy chain to provide a cloud customers control and provide computing platform in cloud. Its also provide a mechanism that provide assure the IaaS platform by cloud service provider is trusted.

h) There are two main approaches. First approach is the main focus on the gap that is slowing down the cloud adaption and reviewing the threat challenges. And second approach is discussing about to address some of the widely attacks using machine learning techniques. Cloud server and the customer’s tool that protect themselves from known or unknown security issues.

B. Security policy in cloud computing

To solve this problem there are some important policies as we shown below.

a) We can divide cloud into multiple domains, and each domain has a different kind of security according to domain the security is to be provided. the domain is to be divided as a global or local.

b) Check that the user’s connection and communications security with the SSL, VPN, PPTP, etc. Using license and allowing and providing multiple authorizations among user, it should be ensure that the data travelled securely between a consumer and service provider.

c) User data security assurance: According to requirement the security is to be providing to a client or consumer. It must be assured that the security is important.

d) Using a series of measure to solve the user dynamic requirements, including a complete single sign-on authentication, proxy, collaborative certification, and certification between security domains.

e) Establishment of third-party monitoring mechanism to ensure that operation of cloud computing environment is safe and stable.

f) The computing requested by service requestor, should carry out the safety tests, it can check whether they contain malicious requests to undermine the security rules.

IV. CONCLUSION

In this paper we are discussion different issues and also provide some of the important policies. This policies is useful for maintaining the data on the cloud. This policies is helpful to provide a provide security to the cloud data. These issues are also have to keep in mind while data are stored in the cloud and also apply the possible solutions for those issues with the cloud.

REFERENCES

[7] Shahryar Shaﬁque Quersh1 , Toufeeq Ahmad1, Khalid Raﬁque2, Shuja-ul-Islam3 “Mobile cloud computing as future for mobile applications – implementation methods and challenging issues”-2011.

***
EYE RECOGNITION AND HAND GESTURE IDENTIFICATION FUSION SYSTEM

S. A. CHHABRIA1, SHRUNKHALA SATISH WANKHEDE2 & R. V. DHARASKAR

1Head Of Information Technology Department, G. H. Raisoni College Of Engineering, Nagpur, Maharashtra
2Computer Sciences And Engineering, G. H. Raisoni College Of Engineering, Nagpur, Maharashtra
3Matoshri Pratishthan's Group of Institutions, Nanded, Maharashtra

Abstract-In this paper, an individual human computer interface system using eye motion and hand gestures is introduced. Traditionally, human computer interface uses mouse, keyboard as an input device. This paper presents interface between computer and human. This technology is intended to replace the conventional computer screen pointing devices for the use of disabled. The paper presents a novel idea to control computer mouse cursor movement with human eyes and hand gestures. Hand gesture is used as a mechanism for interaction with the computers.

Keywords- Eye tracking, mouse movement, Eye-blinking detection, Hand gesture recognition, Hand tracking

1. INTRODUCTION

Recently there has been a growing interest in developing natural interaction between human and computer. Several studies for human-computer interaction in universal computing are introduced. [1] The vision-based interface technique extracts motion information without any high cost equipments from an input video image. Thus, vision-based approach is taken into account an effective technique to develop human computer interface systems. For vision-based human computer interaction, eye tracking is a hot issue. Eye tracking research is distinguished by the emergency of interactive applications. However, to develop a vision-based multimodal human computer interface system, an eye tracking and their recognition is done. Real-time eye input has been used most frequently for disabled users, who can use only their eyes for input.

Hand gesture has been one of the most common and natural communication media among human being. The keyboard and mouse are currently the main interfaces between man and computer. Hand gesture recognition research has gained a lot of attentions because of its applications for interactive human-machine interface and virtual environments. Most of the recent works related to hand gesture interface techniques [1] has been categorized as: glove-based method [2, 3] and vision-based method. There are many vision-based techniques, such as model-based [4] and state-based [5] for locating objects and recognizing gesturers. Recently, there have been an increasing number of gesture recognition researches using vision-based methods. This paper introduces an eye and hand gesture recognition system to recognize 'dynamic gesture'.

“Design and implementation of human computer interface tracking system based on multiple eye features”. For human eye (Iris) detection, batch mode is employed. Iris tracking technique is implemented on static images. This technique simply works when the direction of iris is left, right or center. If the position of iris is up or down, it does not work. The system not works in real time. It is not expert to handle blinks and close eyes. [6]

This paper is aimed for designing and implementing a human computer interface system that tracks the direction of the human eye. The particular motion as well as direction of the iris is employed to drive the interface by positioning the mouse cursor consequently. The location of the iris is completed in batch mode. This means that the frames are stored in a permanent storage device and are retrieved one by one. Each of the frames is processed for finding the location of the iris and thereby placing the mouse cursor consequently. Such a system that detects the iris position from still images provides an alternate input modality to facilitate computer users with severe disabilities.

“Statistical models of appearance for eye tracking and eye blink detection and measurement”.[7,8] Active Appearance Model (AAM) a proof-of- concept model for the eye region is created to determine the parameters that measure the degree of eye blinks. After developing an eye model, a blink detector is projected. The main advantage of using AAM technique is that the detailed description of the eye is obtained and not just its rough location. The main drawback of AAM technique is that it is designed to work for a single individual and additionally the blink parameters have to be identified in advance.

“Simultaneous eye tracking and blink detection with interactive particle filters”. [9] Eye position is found using eye recognition algorithm. Then these filters are used for eye tracking and blink detection. For
Eye Recognition and hand Gesture Identification fusion System

...describing state transition, auto regression models are used. A statistical active appearance model (AAM) is developed to track and detect eye blinking. The model has been designed for variations of head pose or gaze. During this paper, the model parameters which encode the variations caused by blinking are analyzed and determine. This international model is further extended using a series of sub-models to enable independent modeling and tracking of the two eye regions. Many techniques to enable measurement and detection of eye-blink are proposed and evaluated. The results of various tests on completely different image databases are presented to validate each model.

“Communication via eye blinks- Detection and duration analysis in real-time” [10] Initial eye blink is employed to find the eyes. The algorithm detects the eye blinks. The “Blink link” prototype can be used in order to get in touch with the device. Simply by considering the motion information among two consecutive frames and determining that if this motion is caused by blink, eyes are tracked and monitored constantly. This system is a real-time system. The disadvantage of this system is that it can only handle long blinks and is not able to handle short blinks. In case of short blinks it just simply avoids the blinks.

“MouseField: A Simple and Versatile Input Device for Ubiquitous Computing”. [11] “MouseField” is a individual personal laptop or human computer interaction system that uses RFID reader and motion sensor. Especially the vision based face and hand motion tracking and gesture recognition is an attractive input mode for better human-computer interaction. Human gesture information has been variously employed in the game, virtual reality and other applications. Such gesture information is classified into the static gesture which uses spatial information only and the dynamic gesture which uses the spatial information and time information together. Since, the dynamic gesture can presents various expressions and it is considered as a natural presenting technique. Such motion information can be acquired by both using device-based interface and vision-based interface. The device-based interface technique gets motion information by motion capture devices and marker. However, the vision-based interface technique extracts motion information from input video image without any high cost equipments. Thus, vision-based approach is considered an effective technique to develop human computer interface systems. For vision-based human computer interaction, eye and hand tracking is hot issue. Eye tracking search is distinguished by the emergence of interactive applications. Although various interaction technologies for handling information in the present computing atmosphere have been proposed, some techniques are too easy for performing human computer interaction, and others require special expensive equipments to be set up everywhere, and cannot quickly be accessed in our daily environment. In this, a new simple and versatile input device called the MouseField that enables users to control various information appliances easily without large amount of expenses. [11] MouseField consists of an identification recognizer and motion sensors that can detect an object and its movement after the object is placed on it. The system can easily translate the user’s actions as a command to control the flow of information. A robust and versatile input device called the MouseField that can be used at almost any place for controlling information appliances. MouseField is a device that combines ID reader and motion sensing devices into one package.

HAND GESTURE IDENTIFICATION

Glove-based gesture interfaces require the user to wear a device, and generally carry a load of cables that connect the device to a computer.

Huang et al. [12] use 3D neural network method to develop a Taiwanese Sign Language (TSL) recognition system to recognize 15 different gestures. This paper presents sign language recognition system which consists of three modules: model-based hand tracking, feature extraction, and gesture recognition using 3D Hopfield neural network (HNN).

David and Shah [13] propose a model-based approach by using a finite state machine to model four qualitatively distinct phases of a generic gesture. Hand shapes are described by a list of vectors and then matched with the stored vector models. The seven gestures are representatives for actions of Left, Right, Up, Down, Grab, Rotate, and Stop.

Darrell and Pentland [14] propose space-time gesture recognition method. This paper presents a method for learning, tracking, and recognizing human gestures using a view-based approach to model both object and behavior. Signs are represented by using sets of view models, and then are matched to stored gesture patterns using dynamic time warping.

Starner et al. [15] describe an extensible system which uses one color camera to track hands in real time and interprets American Sign Language (ASL). They use hidden Markov models (HMMs) to recognize a full sentence and demonstrate the feasibility of recognizing a series of complicated series of gesture. Instead of using instrumented glove, they use vision-based approach to capture the hand shape, orientation and trajectory. The vision-based method selects the 3-D input data as the feature vectors for the HMM input, other HMM-based [16, 17] hand gesture recognition systems have also been development. Liang et al. [118] develop gesture recognition of TSL by using Data-Glove to capture the flexion of 10 finger joints, the roll of palm and other 3-D motion information.

Cui and Weng [19] develop a non-HMM-based system which can recognize 28 different
gestures in front of complex backgrounds. Nishikawa et al. [20] propose a new technique for description and recognition of human gestures. The proposed method is based on the rate of change of gesture motion direction that is estimated using optical flow from monocular motion images.

Nagaya et al. [21] propose a method to recognize gestures using an approximate shape of gesture trajectories in a pattern space defined by the inner-product between patterns on continuous frame images. Heap and Hogg [22] present a method for tracking of a hand using a deformable model, which also works in the presence of complex backgrounds. The deformable model describes one hand posture and certain variations of it and is not aimed at recognizing different postures.

Zhu and Yuille [23] develop a statistical framework using principal component analysis and stochastic shape grammars to represent and recognize the shapes of animated objects. It is called flexible object recognition and modeling system (FORMS). Lockton et al. [24] propose a real-time gesture recognition system which can recognize 46 ASL letter spelling alphabet and digits. The gestures that are recognized by [25] are ‘static gestures’ of which the hand gestures do not move. Different from [25], this paper introduces a hand gesture recognition system to recognize ‘dynamic gesture’.

3. GOAL OF THE SYSTEM:

1. Facilitating the handicapped in using the computer
2. Controlling the mouse pointer through eye and hand gesture
3. Human computer interaction provides real time eye tracking and hand gesture estimation

4. OBJECTIVES OF THE SYSTEM:

1. Easy interaction with computer without using mouse
2. Limitation of stationary head is eliminated.
3. Pointer of the mouse will move on screen where the user will be looking & the clicks will be performed by blinking.

5. PROPOSED SYSTEM

Controlling mouse cursor by using eye and hand fusion technique. Chess playing is an application of this system.

<table>
<thead>
<tr>
<th>VARIOUS GESTURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ = Move the Knight Right two step and then up</td>
</tr>
<tr>
<td>+ = Move the Knight Left two step and then up</td>
</tr>
<tr>
<td>+ = Move the Knight Left two step and then down</td>
</tr>
<tr>
<td>+ = Move the Knight Up two step and then right</td>
</tr>
<tr>
<td>+ = Move the Knight Up two step and then left</td>
</tr>
<tr>
<td>+ = Move the Knight down two step and then left</td>
</tr>
<tr>
<td>Eye Recognition and hand Gesture Identification fusion System</td>
</tr>
<tr>
<td>-------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Move the Knight</strong> Down two step and then right</td>
</tr>
<tr>
<td><strong>Move Pawn</strong> Upward</td>
</tr>
<tr>
<td><strong>Pawn</strong> diagonally Right to kill</td>
</tr>
<tr>
<td><strong>Pawn</strong> diagonally Left to kill</td>
</tr>
<tr>
<td>diagonally leftward - UP</td>
</tr>
<tr>
<td><strong>Move Bishop</strong></td>
</tr>
<tr>
<td><strong>Move Bishop</strong> diagonally rightward - UP</td>
</tr>
<tr>
<td><strong>Move Bishop</strong> diagonally leftward - down</td>
</tr>
<tr>
<td><strong>Move Bishop</strong> diagonally rightward - down</td>
</tr>
<tr>
<td><strong>Move Rook</strong> left</td>
</tr>
<tr>
<td>= Move <strong>Rook</strong> Right</td>
</tr>
<tr>
<td>= Move <strong>Rook</strong> Up</td>
</tr>
<tr>
<td>= Move <strong>Rook</strong> Down</td>
</tr>
<tr>
<td>= Move <strong>Queen</strong> Right</td>
</tr>
<tr>
<td>= Move <strong>Queen</strong> Left</td>
</tr>
<tr>
<td>= Move <strong>Queen</strong> Up</td>
</tr>
<tr>
<td>= Move <strong>Queen</strong> Down</td>
</tr>
<tr>
<td>= Move <strong>King</strong> Right</td>
</tr>
<tr>
<td>= Move <strong>King</strong></td>
</tr>
</tbody>
</table>
CONCLUSION

This paper focused on the analysis of the development of PC using human eyes and hand gesture. The mouse pointer is operated using eye and hand gesture. The most unique aspect of this system is that it does not require any wearable attachments. This makes the interaction more efficient and enjoyable. A user interface is the system by which human interact with a computer.

REFERENCE


Eye Recognition and hand Gesture Identification fusion System


[18] Abdul Wahid Mohamed, “Control of Mouse Movement Using Human Facial Expressions” 57, Ramakrishna Road, Colombo 06, Sri Lanka.


IMPROVEMENT OF QUALITY OF AN IMAGE BY USING GABOR WAVELET TEXTURE

GRISHMA Y. BOBHATE¹ & USHA A. JOGLEKAR²

Department of Computer Engg, Smt. Kashibai Navale College of Engg, off Sinhgad Road, Vadgaon (Bk), Pune, India.

Abstract—The aim of this paper is to improve the quality of a particular image by using wavelet texture and provide a better efficient image retrieval technique. Since the wavelet functions decompose the images into different frequency components and analyses them by calculating three moments of an image as mean, variance and skewness. This will provide fruitful information of an image and these feature values are used to retrieve the correct results from database. Wavelet technique is an efficient method for texture retrieval and classification. In this we introduced an important term Gabor Wavelet texture that satisfy certain mathematical requirements and are used in representing data or other function for image processing.

Keywords- Gabor Wavelet Texture, Mean, Variance, Skewness, Image Processing.

I. INTRODUCTION

In recent years many researches have been made on wavelet. An excellent approach have brought wavelet in various fields such as biomedical applications, wireless communications and computer graphics. Wavelets can be used to extract information from many different kinds of data, including – but certainly not limited to – audio signals and images [3]. Sets of wavelets are generally needed to analyze data fully.

Nowadays there is a rapid growth in image analysis and retrieving techniques, which required systematic design of the system and provide better quality of an image. In this work, an image sample is analyzed using wavelet transform which is the most appropriate texture feature extraction [6]. Wavelet Analysis is a set of tools and gives an effective way to represent time-frequency transformation of a signal using Wavelet Transform [7].

As a definition, Wavelet is a waveform of a signal for limited duration effectively that has an average value zero. Wavelet Transform decomposes an image into different scale levels and improve quality of an image by analyzing texture feature.

The rest of the paper is organized as Section II explains the concept of texture analysis. Section III describes types of transforms. Section IV presents Experimental results based on Gabor wavelet texture features. Section V gives Applications Section VI concludes the paper with a future work.

II. TEXTURE ANALYSIS

Texture is a regular repetition or pattern on a surface and can be defined as a set of local statistical properties of coefficients which constitute the image representation. Texture analysis [8] plays a vital role in image processing and pattern recognition tasks such as remote sensing, medical imaging and robot vision and query by content in large image databases. Purpose of texture analysis:

1) To identify different textured and non-textured regions in an image.
2) To classify or segment different texture regions in an image.
3) To extract boundaries between major texture regions.
4) To describe the Texel unit.
5) To draw 3-D shape from texture.

III. TYPES OF TRANSFORMS

Various transforms techniques have been proposed during last decades but the texture analysis problem remains difficult and subject to intensive research. In this work we will study different types of transforms and estimates the transformation.

1. Discrete Fourier Transform

The Fourier transform's utility lies in its ability to analyze a signal in the time domain for its frequency content. The transform works by first translating a function in the time domain into a function in the frequency domain. The discrete Fourier transform (DFT) [7] estimates the Fourier transform of a function from a finite number of its sampled points. DFT has an efficient implementation through the FFT. It is also separable. It has same symmetry properties as in continuous Fourier transform.

2. Discrete Cosine Transform

Discrete Cosine Transform (DCT) [7] has emerged as the de-facto image transformation in most visual systems. DCTs are equivalent to DFTs of roughly twice the length, operating on real data with even symmetry (since the Fourier transform of a real
and even function is real and even). The Discrete Cosine Transform (DCT) attempts to decorrelate the image data. After decorrelation, each transform coefficient can be encoded independently without losing compression efficiency. They are real, orthogonal, and separable with fast algorithms for its computation. They have a great relevance to data compression.

3. Wavelet Transform

The Wavelet transform [3][8] is a method that allows us to compress, transmit and analyze one-dimensional signals, but also bi-dimensional functions, such as images.

The continuous wavelet transform of a 1-D signal \( f(x) \) is defined as,

\[
W_f(b) = \left[ f(t) \Psi_{a,b}(t) \right] dt = (f, \Psi_{a,b}), \, a \neq 0 \tag{1}
\]

Where the wavelet \( \Psi_{a,b} \) is computed from the mother wavelet \( \Psi \) by scaling 'a' and translation 'b' (described below).

In the Fourier analysis, the base function were sinusoidal, in order to provide only frequency information of the initial signal. The wavelet transform, however, is based on the use of elemental functions, called wavelets and scaling, which are created by scalings and translations of a base function, known as the mother wavelet:

\[
\Psi_{a,b}(x) = (1/\sqrt{a}) \Psi((x-b)/a) \tag{2}
\]

In the case of two dimensions, it can be represented as

\[
\Psi_{a,b}(x,y) = (1/\sqrt{a_1a_2}) \Psi(((x-b_1)/a_1), ((y-b_2)/a_2)) \tag{3}
\]

In other words, the wavelet transform of an image \( f(x,y) \) is a set of coefficients \( c_{ij} \) that depend on the scale and the translation of the signal (Materka and Strzelecki, 1998):

\[
c_{ij} = \int f(x,y) \Psi_{a,b}(x,y) \, dx \, dy \tag{4}
\]

Wavelet transform (WT) represents an image as a sum of Wavelet functions (wavelets) with different locations and scales [17]. Any decomposition of an image into wavelets involves a pair of waveforms: one to represent the high frequencies corresponding to the detailed parts of an image (wavelets function or mother wavelet \( \Psi \)) and one for the low frequencies or smooth parts of an image (scaling function \( \Phi \)).

In this way we have seen various types of transform for image processing and extracting the texture features. In this work, we are using one of the types of extracting texture feature as Gabor filter [2][5].

Gabor Filter

Gabor filters [3] are a group of wavelets, with each wavelet energy is captured at a specific frequency with a specific direction. By expanding a signal using this basis it provides a localized frequency description. Thus the local features i.e. energy of the signal is provided. Texture features can then be extracted from this group of energy distributions. The scale (frequency) and orientation are the tunable property of Gabor filter [5] which makes it especially useful for texture analysis.

Manjunath and Ma [4] have shown that image retrieval using Gabor features outperforms that using pyramid-structured wavelet transform (PWT) features, tree-structured wavelet transform (TWT) features and multiresolution simultaneous autoregressive model (MR-SAR) features.

For a given image \( I(x, y) \) with size \( P \times Q \), its discrete Gabor Wavelet transform is given by a convolution:

\[
G_{mn}(x, y) = \sum_{s} \sum_{t} I(s-x, y-t) \Psi^{*}_{mn}(s,t) \tag{5}
\]

Where, \( s \) and \( t \) are the filter mask size variables, and \( \Psi^{*}_{mn} \) is the complex conjugate of \( \Psi_{mn} \) which is a class of self-similar functions generated from translation and rotation of the following mother wavelet:

\[
\Psi(x,y) = (1/2\pi \sigma_x \sigma_y) \exp[-1/2((x^2/\sigma_x^2)+(y^2/\sigma_y^2))] \exp(2\pi W x) \tag{6}
\]

Where \( W \) is called Modulation Frequency. The self-similar Gabor wavelets [1] are obtained through the generating function:

\[
\Psi_{mn}(x, y) = a^m \Psi(\tilde{x}, \tilde{y}) \tag{7}
\]

Where \( m \) and \( n \) specify the scale and orientation of the wavelet respectively with \( m= 0, 1 \ldots M-1, n= 0, 1 \ldots N-1 \) and

\[
\tilde{x} = a^m (x \cos \theta + y \sin \theta) \\
\tilde{y} = a^m (-x \sin \theta + y \cos \theta)
\]

Where \( a > 1 \) and \( \theta = n\pi/N \). \( N \) denotes the number of orientation and \( a^m \) denotes that energy is independent of \( m \).
Above variables can be described as follows:
\[ a = (U_h/U_l)^{1/M-1} \]
\[ W_{mn} = a^m U_l \]
\[ \sigma_{x, mn} = ((a+1) / (\sqrt{2} \ln 2)) / 2 \Pi a^m (a - 1) U_l \]  
(8)
\[ \sigma_{y, mn} = 1 / (\sqrt{2} \Pi \tan(\Pi/2N) \sqrt{((U_h/\sqrt{2} \ln 2)) - (1/2 \Pi \sigma_{x, mn})^2}) \]  
(9)

Where \( U_h \) and \( U_l \) denote the upper and lower center frequencies of filtered image.

The distance between two textures (i: query image and j: target image) in the database can be defined as:
\[ d(i, j) = \sum \sum d_{mn} (i, j) \]  
(10)

Where, \( d_{mn} = ((\mu_{mn} - (\mu_{mn}))^2 + (\sigma_{mn} - (\sigma_{mn})))^{1/2} \)  
(11)

\( \mu_{mn} \) and \( \sigma_{mn} \) are mean and standard deviations of the magnitude of the transformed coefficient which are used to represent the region for classification.

**IV. EXPERIMENTAL RESULTS**

In this work, we first convert an image into grayscale image as shown in figure (a) to extract Gabor texture features, and then divide into 64×64 pixels. The Gabor filter is applied on that grayscale image with five scales and 8 orientation which results into 40 subimages. This mean for each scale eight orientation are calculated which in turn 40 matrices values are obtained.

For each subimage, three moments are calculated as mean, variance and skewness. Thus, 120-dimensional vector are used for Gabor texture features extraction which help to retrieve the similar images from the database.

The filter define in equation (6) explains in the spatial domain for real part (with cosine shape due to symmetric component) and imaginary part (with sine shape due to the asymmetric component). Both parts can be yield in 5 scales and 8 orientations as shown in figure (b). Here, each scale shown in figure (b) define various filtering techniques for an image features. First scales captures antisymmetrical features from horizontal, vertical and oblique filter of these scales by defining central symmetry of real part. Second scales are folded by periodicity with attenuation \( \sigma \) and the same definition is used for all orientations. Similarly, third and fourth scale are calculated for all orientation.

Finally, Low pass filter smooth an image as shown in figure (c) by keeping the low frequencies. The multi resolution scheme is completed with a low-pass filter (DC) for recovering the luminance information. The results are summarized in table 1 for scale one and eight orientation. Three moment: mean, variance and skewness are calculated when scale=1 for 8 orientation. These values are represented in the form of graph for retrieving the images from database. As this values are calculated for all orientation it improves the process of texture discrimination. This proves a better effective way to improve quality of an image by analyzing scales and orientation in Gabor filter scheme. This results produces the effectiveness for image processing [2].

<table>
<thead>
<tr>
<th>Sr. no.</th>
<th>Scale</th>
<th>Orientation</th>
<th>M</th>
<th>V</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>6</td>
<td>0.0641</td>
<td>0.0274</td>
<td>0.0071</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>7</td>
<td>0.0711</td>
<td>0.0277</td>
<td>0.0071</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>8</td>
<td>0.0796</td>
<td>0.0280</td>
<td>0.0071</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>9</td>
<td>0.0714</td>
<td>0.0277</td>
<td>0.0071</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>10</td>
<td>0.0645</td>
<td>0.0274</td>
<td>0.0071</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>11</td>
<td>0.0714</td>
<td>0.0277</td>
<td>0.0071</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>12</td>
<td>0.0798</td>
<td>0.0280</td>
<td>0.0071</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>13</td>
<td>0.0714</td>
<td>0.0277</td>
<td>0.0071</td>
</tr>
</tbody>
</table>

Table 1: The 1- Dimensional vector for one scale and one orientation value. Three Moments are: M-mean, V-Variance, S-Skewness.

![Figure 1. Gabor filtered image. (a) Texture image is converted into grayscale image. (b) Image with 5 scales and 8 orientations. (c) Low pass filter is applied to smooth the image.](image1)

![Figure 2. Efficient graph of Gabor filter](image2)
This figure shows 120-dimensional vector of Gabor filter ranges from 4.4577e-04 to 0.7264. These magnitude feature values are normalized to obtain maximum magnitude in each set of feature vector of an image. Here, Image (a) is given as query image; it calculates feature values as shown in graph to retrieve correct result from database. Hence, It is clear that Gabor wavelet texture is more effective for texture classification. Hence this method can be widely used for texture analysis and segmentation.

V. APPLICATIONS

1. Gabor filter can be used for Face recognition is one of the most important application. The face image is convolved with a set of Gabor wavelets and the resulting images are further processed for recognition purpose.

2. Gabor wavelets networks are used for face reconstruction by a set of Gabor wavelet coefficients with a certain quality. The quality of the reconstruction depends on the number of wavelets used.

3. Gabor wavelet texture is used for other images such as textures, remote sensing images and landscape images and also for fingerprint recognition, facial landmark location and iris recognition.

VI. CONCLUSION

Gabor Wavelet performs better characteristics for texture classification which makes helpful for improving image quality. Gabor filter eliminates DC-component of input image and achieve a high accuracy rate. This wavelet texture method is used to reduce the redundancy of filtered images which provides a better desired result. In our future work, texture segmentation will be incorporated into our system to facilitate texture-based retrieval.

REFERENCES

LKM FOR BUILDING EFFICIENT FILE SYSTEM INTRUSION DETECTION SYSTEM WITH SECURE UPDATE MECHANISM

RATAN DEOKAR1, SHRADDA PATHAK2, ATUL MAINDAD3, ANIRUDDHA PORE4 & MINAL PIMPALE5

1Department of Information Technology, MIT, University of Pune, India. 2,3,4,5Department of Computer Engineering, PCCOE, University of Pune, India.

Abstract— Access control mechanism is generally used to protect the activities on file system specific to the operating system. This directly implies that access protection mechanism is generally tightly coupled with almost all operating systems. Still the phenomenon of intrusion exists. So, here we are not only detecting intrusion but also preventing it by building Loadable Kernel Module. LKM prevents the unauthorized access and helps to build efficient file system.

In order to achieve this we are maintaining the MAC DTS (Modification, Access and Creation Date and Time stamp) in our log file and providing security by hiding log file. Along with this we are providing secure update mechanism which will prevent the changes done by intruder from being reflected in original file system.

Keywords: Host Based Intrusion Detection System (HIDS); Virtual File System (VFS); System calls; Modification, Access and Creation Date and Time stamp (MAC DTS); Loadable Kernel Module (LKM).

I. INTRODUCTION

A File system is the way the files are organized on the disk i.e. data structure. In Linux everything is considered as file. Every user has its own privilege to access the file and for root all the rights are authorized. So by getting the root access malicious attacker can break into the system. Malicious attacker will try to get the root access by finding and cracking the root password and they will try to evade by disabling the logging feature. So it is important to protect the log file from malicious activities.

The Kernel can be altered and this is done in order to bring it down or take control of it. So it is important to keep the kernel secure in Linux Operating System. This process known as Kernel Hardening. In Kernel Hardening additional kernel-level security is provided to improve the security of the system, while keeping system close to traditional Linux.

An intrusion detection system (IDS) is a device or software application that monitors network or system activities for malicious activities or policy violations and produces reports to a Management Station. It primarily focused on identifying possible incidents, logging information about malicious activities, and reporting attempts. One issue related to Intrusion detection system is, it provides information about intrusion but does not provide any security of data which is modified or deleted by intruder. Here we need to consider the security of modified or deleted files. This paper deals these issues of file system.

II. THEORETICAL BASIS

Linux views all file systems as a set of objects such as superblock, inode, dentry, and file. Superblock describes and maintains state for the file system and it is at root of every file system. Inode is an object that is managed within a file system (file or directory) in Linux. The inode contains inode number of file and all the information regarding a file except filename. This metadata is used to manage objects in file system. Dentries is set of structure which is used to translate between names and inodes and a directory cache is maintained to keep the track of most-recently used dentries. The relationships between directories and files are maintained by the dentry for traversing the file system.

The Linux architecture supports a large variety of file systems on a large variety of storage devices. This is done by using common set of API functions which contained in VFS layer. For example, the write function call which allows some number of bytes to write into a given file descriptor. The write function is unaware of file system types, such as ext2 or NTFS. It is also unaware of the particular storage medium upon which the file system is mounted. Yet, when the write function is called for an open file, the data is returned as expected for math, etc.

III. RELATED WORK

For enhancing security of the file system in Linux a prototype system STEFS (Security and Trusted Enhanced File System) [4] of dependability based on trusted Computing platform is presented in this paper. STEFS provides three dynamic loading modules which consist of file encryption and integrity checking measures, multilevel access control mechanism to strengthen the security protection of system sensitive data. More reliability can be achieved by considering these three modules.
Das et al.[1] is emphasizing on preserving MAC DTS of files on an event of file system access, by using the concept of system call trapping at different levels. First, security administrator identifies a set of system critical files which are important from the system security viewpoint. When the intruder violates the current access control policy and try to view any file implies a call to sys_open () system call, we trap sys_open (), take the path name argument to the modified system call, get the MAC DTS information from inode (of the pathname specified) and redirect to the normal operation of sys_open ()..

Barik et al. [3] is focusing on an efficient technique for enhancing forensic capabilities of Ext2 file system. The forensic information is required to get the evidence of intruder’s malicious activities. It is file system specific solution, but generic solution is required.

Guo Jin, Li Bo, [4] has discussed design and implementation of a cryptographic file system in this paper. By encryption and decryption techniques file security can be achieved. This paper do not deals with IDS.

Kashif Ahmad Khan in “SELinux IN and Out”[5] has discussed various SELinux policies. The SELinux supports DAC (Discretionary Access Control) and its checks are performed after DAC’s.

Vigna et al. [7] has discussed network-based intrusion detection systems and host-based intrusion detection systems. Network based solutions provides the ease of maintenance and the possibility to monitor several targets with a single IDS. Host based system analyses all part of the computer system and the network packets on its network interfaces. Authors analyzed the advantages and limitations of host-based solutions with regard to network-based techniques and concluded that network based system dominates the host based system.

IV. ARCHITECTURE

Interface between user programs and kernel is provided by the VFS layer. It reduces complexity by providing an abstraction within the kernel which allows different file system implementations to coexist. This is done by combining common set of API functions. Various system calls such as read, write, open, stat are implemented by VFS. It provides a fast lookup mechanism by taking a pathname parameter to search through the directory entry cache and translate it into a specific dentry. Dentries are used for the performance purpose and they live on RAM only. Inodes are file system objects such as regular files, directories. Every individual dentry has a pointer to inode. Inodes resides either on the disk (for block device file systems) or in the memory (for pseudo file systems). Inodes, which resides on the disk, are brought into the memory when required, and inode changes are written back to disk. Several dentries can point to a single inode (hard links). And single dentry can also be point to single inode (soft links). Each file has its own inode structure. Each inode structure has its inode number that is unique for the currently mounted file system.

Our driver will control inode and file operations for the specified file and its parent directory. For this purpose, we will change pointers on inode and file operation structures to our own functions. That will allow us to hide file even from the system.

![Figure 1. Working of Dentry Cache.](image1)

![Figure 2. Architecture.](image2)
individual file system. The main Objective of our system is building Loadable Kernel Module (LKM). LKM consists of File Hide Mechanism, Changes in File System, Changes Stored in Module Only, and Password Protected Update Mechanism. File Hide Mechanism is used for hiding files in our system by using the working of Dentry Cache. Changes in file system are stored in module only i.e. even if intruder tries to violate the access policies, those changes are not reflected in the original file system and are saved in LKM, for this purpose Password Protected Update Mechanism is used. Our paper mainly focuses on Secure Update Mechanism.

A. Secure update mechanism
Following are the steps followed for secure update mechanism:

- We are defining access policies for group of users.
- Maintaining list of important files.
- Administrator will be provided with Save and Discard option.

1. Defining access policies
For different group of users we are defining access policies such as read/write/execute. And those groups of users will be allowed to perform only those operations.

   ![Image](image3.png)

   **Figure 3. Defining access policies**

   In the above diagram group G1 is provided with the read and execute operations and similarly group G2 is provided with the read and write operations. So G1 users are allowed to perform only read and execute operation whereas G2 can perform read and write operations. They are not allowed to perform the operations which are not defined in access policies.

2. Actions for access violation
The group of users is provided with access policies. If those users try to execute the operations that are not defined in their access policies then access policies gets violated. Then the changes made by the users who are not allowed to execute operations other than their access policies, are loaded in LKM memory with Save and Discard options. Administrator will monitor those changes and decide whether to Save or Discard those changes. We are providing the password mechanism for it i.e. only after providing the password admin will be able to save or discard the changes. If admin will save the changes then those changes will get saved permanently in file system else if he choose discard option then changes will be discarded.

   ![Image](image4.png)

   **Figure 4. Actions for Access Violation**

3. Administrator will be provided with Save and Discard option
The changes made by the users who are not allowed to execute operations other than their access policies, are loaded in LKM memory with Save and Discard options. Administrator will monitor those changes and decide whether to Save or Discard those changes. We are providing the password mechanism for it i.e. only after providing the password admin will be able to save or discard the changes. If admin will save the changes then those changes will get saved permanently in file system else if he choose discard option then changes will be discarded.

   ![Image](image5.png)

   **Figure 5. Contents of LKM memory**

B. Advantages
- System provides the effective mechanism for unauthorized access and modification of the important files and secure updating of data of those files.

---

**International Conference on Computer Science and Information Technology (ICCSIT), 3rd Feb-2013, Nagpur, ISBN: 978-93-82208-60-0**
• It doesn't require any change in existing installations.
• MAC DTS log can serve as a building block for modeling anomaly based HIDS, as far as File System based intrusion detection is concerned.
• System will work with any underlying file system if it is registered with the VFS and mounted.

V. CONCLUSION

Every operating system has its own set of critical files, whose access is generally protected by access control mechanisms, native to the operating system. Because of the importance of the files they invite unauthorized access and modification. This system provides the effective mechanism for such types of intrusion and secure updating of data. It doesn't require any change in existing installations. It will work with any underlying file system if it is registered with the VFS and mounted. Using password protected secure update mechanism it’s possible to prevent the changes done by intruder from being reflected in original file system.

REFERENCES


CLASSIFIER BASED FEATURE SELECTION FOR DISEASE PREDICTION

V.VAITHIYANATHAN1, K.RAJESWARI2 & SHAILAJA V.PEDE3

1Research, School of Computing and CTS Chair Professor, SASTRA University, Tanjore, India
2PCCOE, Pune University, & Ph.D Research Scholar, SASTRA University, Tanjore, India.
3(ME Computer) PCCOE, Pune University, Pune, India

Abstract: The wide applications of data mining are like e-business, marketing, medicine and other industries and sectors. The medical domain is filled with voluminous information stored in medical databases. Extracting useful knowledge from the database and providing support to make scientific decisions for the diagnosis and subsequent treatment is now becoming significant. Data mining in medicine can also improve the management level of hospital information and promote the development of telemedicine and community medicine. As medical information is with multiple attributes, incomplete and time variant, medical data mining differs from other fields and hence more efforts for knowledge extraction is mandatory. Diagnosis of most of the diseases is expensive as many tests are required to predict the disease. By using data mining techniques we can reduce the number of expensive tests to by selection of those attributes which are really important for prediction of disease. Dimensionality reduction plays an important role in the field of medicine as it contains multiple attributes. In this paper the important attributes are identified using classifiers such as decision tree, SVM and neural networks on the standard benchmark datasets such as Heart, Breast cancer and Diabetes disease from UCI repository. It is observed that removing some of the attributes improves the accuracy of all classifiers and Support Vector Machines (SVM) gives better accuracy than the other classifiers.

Keywords: Feature selection, Medical Data mining, Disease diagnosis, Classifiers, Dimensionality reduction.

I. INTRODUCTION

Large amount of heterogeneous data is generated in medicine field. For example, medical data may contain SPECT images, signals like ECG, clinical information like temperature, cholesterol levels, etc., as well as the physician’s interpretation. Those who deal with such data understand that there is a widening gap between data collection and data comprehension. Computerized techniques are needed to help for analyzing data for discovery of knowledge not only to help physician also for prediction of disease and satisfaction of customer in minimum cost.

In this paper we have considered three datasets of different diseases such as Heart disease, Breast cancer and Diabetes disease. All these datasets are downloaded from standard benchmark UCI repository. Now there are high probability of heart attacks, diabetes because of hyper tension and today’s lifestyle. And many tests are needed for diagnosis of that disease. By making use of data mining techniques and reducing the number of attributes or selecting only important features of that disease among all available attributes we can try to diagnosis that disease using classifiers and help to clinical decision support system.

As heart disease and Heart attack is one of major disease in todays life so we have considered Heart Cleveland dataset contains 14 attributes including class attributes and the number of instances are 303. The available dataset is initially discretized based on some assumptions after studying various papers and consulting with medical practitioners. The detail description of dataset is given in following Table 1.

Breast cancer is another problem observed in females. This dataset contains total 10 attributes including class attribute and number of instances are286. The detail dataset description is available in Table 2.

Diabetes most common disease found in any person of any kind of age. Diabetes diagnosis dataset contains 9 attributes including class attribute. Numbers of instances are 768 where last attributes shows sick or healthy.

Table 1. Heart Cleveland dataset description

<table>
<thead>
<tr>
<th>Srv. No.</th>
<th>Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Age</td>
<td>Values [36,39,46,48]</td>
</tr>
<tr>
<td>2</td>
<td>Sex</td>
<td>Value 1: Male, 0: Female</td>
</tr>
<tr>
<td>3</td>
<td>Cp</td>
<td>1: typical type 1 angina, 2: typical type angina, 3: non angina pain, 4: asymptotic</td>
</tr>
<tr>
<td>4</td>
<td>Trestbps</td>
<td>0:120, 1:120-139, 2:140-159, max men having admission to the hospital</td>
</tr>
<tr>
<td>7</td>
<td>Resting</td>
<td>0: normal, 1: low, 2: high</td>
</tr>
<tr>
<td>8</td>
<td>Thal</td>
<td>0:0, 1:1, 2:2, 3:3</td>
</tr>
<tr>
<td>9</td>
<td>Exng</td>
<td>Value 1: yes, 0: no</td>
</tr>
<tr>
<td>10</td>
<td>Oldpeak</td>
<td>0:0, 1:5, 2:5, 3:5, 4:5, 5:5</td>
</tr>
<tr>
<td>11</td>
<td>Slope</td>
<td>Value 1: flat, 2: upsloping, 3: downsloping</td>
</tr>
<tr>
<td>12</td>
<td>Ca</td>
<td>Number of major vessels colored by fluoroscopy: 0: none, 1: one, 2: two, 3: three</td>
</tr>
<tr>
<td>13</td>
<td>Thal</td>
<td>0: normal, 1: fixed defect, 2: reversible defect</td>
</tr>
<tr>
<td>14</td>
<td>Num</td>
<td>Value 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100</td>
</tr>
</tbody>
</table>
II. BACKGROUND

Healthcare professionals store significant amounts of patients’ data that could be used to extract useful knowledge. Researchers have been investigating the use of statistical analysis and data mining techniques to help healthcare professionals in the diagnosis of heart disease. Statistical analysis has identified the risk factors associated with heart disease as age, blood pressure, smoking habit [1], total cholesterol [2], diabetes [3], hypertension, family history of heart disease [4], obesity, and lack of physical activity [5]. Knowledge of the risk factors associated with heart disease helps health care professionals to identify patients at high risk of having heart disease. Researchers have been applying different data mining techniques such as decision tree, naïve bayes, neural network, bagging, kernel density, and support vector machine over different heart disease datasets to help health care professionals in the diagnosis of heart disease [6]-[11]. In [12] showed correct classified accuracy of approximately 77% with logistic regression. Another model R-C4.5 is applied which is based on C4.5 shows better results, rules created by R-C4.5 can give health care experts and useful explanations [13]. CLASSIT conceptual clustering system [14] achieved a 78.9% accuracy on the Cleveland database. In [15] Neural network based method obtained 89.01% classification accuracy from the experiments made on the data taken from Cleveland heart disease database. The ANFIS classification [16] with PCA of diabetes disease was classified due to training and test of all the diabetes disease dataset. The obtained test classification accuracy was 89.47% by using the 10 fold cross validation.

III. RESEARCH OBJECTIVE

The main objective of this research is to develop a decision support system by selecting important features and classifiers which reduce the number of clinical test. The System can discover and extract hidden knowledge associated with diseases (heart attack, cancer and diabetes) from a historical disease database. It show the result after diagnosing disease and thus support to healthcare practitioners to make intelligent clinical decisions which traditional decision support systems cannot. By providing effective treatments, it also helps to reduce treatment costs.

IV. SCOPE OF THE WORK

Here the scope of this work is that integration of clinical decision support with computer based techniques which may keep patient records that helps to reduce errors in multiattribute medical field, cares about patient safety, remove unwanted practice variation, and improve patient outcome. Using data mining and generating useful knowledge which can help to improve the quality of clinical decisions. The main objective of this research is to develop a useful clinical decision support system for accurate diagnosing of major diseases like Heart Disease using classifiers, namely, Decision Trees, Neural Network and SVM with feature selection. So it provides effective treatments, it also helps to reduce treatment costs.

V. METHODOLOGY

This paper shows the analysis of various data mining techniques which can be helpful for medical analysts or practitioners for support in correct diagnosis of any major disease. The main methodology used for our work was by examining the standard datasets downloaded from UCI and executed through WEKA 3.7.7

A. Classification of dataset using different classifiers:

We have used widely popular data mining tool WEKA 3.7.7 for checking the accuracy of classifiers on different datasets. Initially data is preprocessed while preprocessing data discretization is done and then classified using multilayer perceptron, J48 decision tree and Support Vector machine as SMO classifier. Following table shows results of classifiers among all three datasets considering all attributes.

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Class</td>
<td>no-recurrence-events, recurrence-events</td>
</tr>
<tr>
<td>3</td>
<td>Menopause</td>
<td>&lt;40, &gt;40, premeno</td>
</tr>
<tr>
<td>4</td>
<td>Tumor size</td>
<td>0-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59</td>
</tr>
<tr>
<td>5</td>
<td>Inv-nodes</td>
<td>0-2, 3-5, 6-8, 9-11, 12-14, 15-17, 18-20, 21-23, 24-26, 27-29, 30-32, 33-35, 36-39</td>
</tr>
<tr>
<td>6</td>
<td>Node-axis</td>
<td>Yes, No</td>
</tr>
<tr>
<td>7</td>
<td>Diam-nilig</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>8</td>
<td>Breast</td>
<td>left, right</td>
</tr>
<tr>
<td>9</td>
<td>Breast quad</td>
<td>left-up, left-low, right-up, right-low, central</td>
</tr>
<tr>
<td>10</td>
<td>Implant</td>
<td>Yes, No</td>
</tr>
</tbody>
</table>
Dimensionality reduction is one of the major issues for dataset containing large volume of data, in above table it shows accuracy of all classifier including all attributes of that datasets. Feature selection is one of the useful methods for dimensionality reduction. By selecting only important features or rectifying unwanted attributes of that dataset we can predict the class more accurately so that it becomes the cheapest way by performing minimum tests for prediction of disease. For testing the accuracy of all classifiers after removal of attributes on above all dataset we have tried removal of attributes one by one and by considering the combination of attributes and the accuracy is measured. Following table shows the results of all classifiers after removal of one, two, three and four attributes from the actual attributes. In this case we have considered stopping criterion as removal of minimum 50% attributes. We have tested by removing more attributes but the later on the accuracy of classifier was less. In this paper we have shown the results of classifiers after removing four attributes.
As shown in the graph there is an improvement in the accuracy of classifier after removal of attributes. Table 9 shows the results after attribute removal from that it is clear that if we remove the attribute number 3,9 and 12 the accuracy of classifier is minimum which shows an importance of these three attributes for heart disease dataset using SVM classifier. By selection of these important features it helps to avoid expensive tests for diagnosis of disease.

VII. CONCLUSION AND FUTURE WORK

From above analysis we conclude that for the dataset Heart disease there is an improvement in accuracy of all classifiers after removal of few attributes. This shows that these attributes are not much important for prediction of disease. Other than considering all those attributes we may consider only those attributes those are really helpful for disease prediction so that dimensionality of the dataset also decreased without degrading the performance of classifier. This survey also conclude that the accuracy of SVM classifier is better than other classifiers. This works extends to select the exact features of any dataset without degrading the performance of the clinical decision support system in medical data mining.

REFERENCES


[17] Jos eR. Quevedo,AntonioBahamonde,MiguelPerezEnciso,OscarLuaces ”, Disease liability prediction from large scale genotyping data using classifiers with a reject option”, IEEE/ACM transactions on computational biology and bioinformatics, month2010


International Conference on Computer Science and Information Technology (ICCSIT), 3rd Feb-2013, Nagpur, ISBN: 978-93-82208-60-0
ANALYSIS OF NETWORK-LAYER ATTACKS IN MANET

POOJA K. MAHAJAN¹ & E. JAYANTHI²

¹²Department of Computer (C.N.) Sinhagad College of Engg. Pune, India

Abstract— MANET is a mobile ad hoc network which is capable of working without any fixed infrastructure. It is a temporary network which is formed by collection of mobile hosts without any centralized authority. In MANET the network topology changes dynamically. The nodes need to be more intelligent as there is no centralized authority. Also there are no routers or base stations in case of MANET. Hence nodes are responsible to handle incoming and outgoing data packets. Considering these characteristics of MANET it is susceptible to variety of attacks. In this paper we provide an analysis of network layer attacks in MANET.

Keywords- MANET; attacks; malicious nodes; etc.

I. INTRODUCTION

MANET is a mobile ad hoc network in which multi hop communication takes place. It is infrastructure-less and nodes are free to move from one place to another. Hence route changes frequently because of dynamic network topology. Each node in MANET acts as a router in order to handle data traffic in a network. These characteristics of MANET make it vulnerable to variety of insider attacks. Insider attacks are those that are caused by malicious node inside the network. Most popular reactive (On demand) routing protocols in MANET are Ad hoc On Demand Distance Vector (AODV) and the Dynamic Source Routing (DSR) which are susceptible to variety of attacks [12].

In multi-hop communication a source node sends packets to the destination node which may be far away from it through intermediate nodes. This mechanism helps for better connectivity and low power consumption than in direct transmission over a long distance [10]. Hence every node should cooperate with other in order to provide proper communication in a network. But sometimes they do not cooperate because of maliciousness and selfishness. This leads MANET vulnerable to variety of attacks.

II. DIFFERENT TYPES OF NETWORK-LAYER ATTACKS

In case of MANET routing is most important thing for proper communication in a network. Sometimes due to the wrong attitude of the malicious nodes different types of routing attacks are occurred in a network. The network services can be disturbed by an attacker using different techniques.

A. Blackhole Attack

In black hole attack the malicious node advertise itself for having shortest path towards the destination node. When the malicious node receives the RREQ packet from the neighbourhood node; it immediately sends a false RREP. In this way the malicious node advertise itself for having fresh routes without checking its routing table.

In the above figure; the node 1 wants to send data packets to node 4. Hence it broadcasts the RREQ packet from the neighbourhood node; it immediately sends a false RREP. In this way the malicious node advertise itself for having fresh routes without checking its routing table.

In the above figure; the node 1 wants to send data packets to node 4. Hence it broadcasts the RREQ packets in the network. Consider node 2 is a malicious node, hence it will immediately reply to the RREQ packet without forwarding the RREQ packet to the other nodes. In this way node 1 will think that the node 2 is having active route and it will ignore the RREP packets coming from the other desired nodes. Now, node 1 will send the data packets to node 2. In this all data packets will be lost and will not reach towards the destination node.
B. Grayhole attack

In this type of attack a malicious node (called as gray hole) first agrees to forward packets but later it may drop packets destined for specific node. In another case a gray hole may drop packets for some period but later it behaves normally [2]. This uncertain behavior of a node can cause a big damage to the communication in a network.

C. Wormhole attack

In wormhole attack two or more malicious nodes are connected with each other. The connection link between them is called as ‘wormhole link’. The attacker nodes capture packets at one end and replay them at another end via high speed network connection [11]. Wormhole attack is generally achieved by two nodes. As shown in the figure-3, node 6 and node 8 are attacker nodes which are connected via wormhole link. Now, the source node 1 wants to send data to the destination node 9. Hence it broadcasts RREQ packet to its neighborhood nodes. When malicious node 6 receives the RREQ packet forwarded by node 4 it records and tunnels this RREQ packet to malicious node 8 using wormhole-link. Then malicious node 8 forwards the RREQ packet to the destination node 9. As, the RREQ from 1-4-6-8-9 reaches first the destination node 9 will ignore the RREQ that reaches later and it unicasts the RREP packet to source node 1.

Though the wormhole-link is easy to set up it can badly disturb the network communication. The wormhole link can be use by an attacker with packet dropping attack in order to prevent destination node from receiving packets. Also, it prevents to discover any other routes in a network.

D. Byzantine attack

A malicious node or a group of malicious node work in collaboration in order to create routing loops, forwarding packets through non-optimal paths, or selectively drops packets in order to degrade quality of service [11].

E. Denial of Service Attack

When a node wants to establish a route towards destination node using AODV protocol then it broadcasts RREQ packets to the neighbourhood nodes. Nodes receiving the RREQ packet stores and forward the packet towards the destination node. If the route towards destination is broken then RERR packet is sent to the source node [12].

In Denial of Service attack the node avoids participation in the routing process though it is having active or shortest path towards the destination node. In figure 4 though the node 6 is having active and shortest path to the destination node 5 till it sends RERR packets to the source node 1. In such way it deny RREQ packet from reaching to the destination node.

F. Flooding Attack

The aim behind the flooding attack is to consume network resources. In this type of attack the malicious node broadcasts the RREQ packet towards the
destination node which is not actually present in the network. The large number of RREQ messages are broadcast in a short period of time. Because of this a network flooding takes place [3]. This cause unnessesary battery and resouce consumption of the nodes and could leads to Denial of Service attack as the nodes will not be able to receive other RREQ packet from legitimate node.

G. Modify and forward Attack

In this case of attack when malicious node receives the RREQ paket from neighbourhood node it modify one or more fields of the RREQ packet and then broadcast it. In such way the victim node will never be able to establish a route towards the destination node.

The possible fields can be modified in RREQ packet RREQ ID, hop count, Destination IP address, Source IP address, Destination sequence number, Sourcece sequence number, flages, etc [5].

![Figure 5. Modify and forward Attack](image)

H. Spoofing attack

Spoofing attack is also called as impersonation attack in which a malicious node pretends to be a authorized node. This allows the malicious node to take advantages of authorized services in a network [8].

III. CONCLUSION

Analysis of above attacks helps to understand the vulnerability of routing protocols. These attack characteristics should be considered in designing of mobile ad hoc network in order to prevent it against different attacks. Also one can use this analysis to find out solution in order to detect and prevent attacks.

ACKNOWLEDGMENT

I am heartily thankful to Prof. E. Jayanthi (Professor of Computer Engineering Department, Sinhgad College of Engineering Pune, Maharasthra, India) for her valuable guidance.

REFERENCES


LOSSLESS COMPRESSION OF MEDICAL IMAGES

MAYURI Y. THORAT¹ & VINAYAK K. BAIRAGI²

¹Researcher Electronics and Telecommunication Dept, Sinhgad Academy of Engineering, Kondhwa (Bk), Pune-411048
²Electronics and Telecommunication Dept, Sinhgad Academy of Engineering, Kondhwa (Bk), Pune-411048

Abstract – Telemedicine involves storage and transmission of medical images, popularly known as Teleradiology. Due to constraints on bandwidth and storage capacity, a medical image may be needed to be compressed before transmission/storage. The image compression is widely addressed research area. The method proposed in this paper is symmetry based technique for compression of 3D medical image data. The proposed technique takes an advantage of anatomic symmetries present in structures of medical images to reduce energy of sub-bands. The image is divided into two parts which are approximate replica of each other. The residual data is generated by taking difference of two parts. This data is then encoded using Run Length Encoding (RLE). The compressed image obtained by this method can be reconstructed back by applying reverse algorithm.

Keywords- Telemedicine, symmetry; 3D medical image compression, Run Length Encoding (RLE)

1. INTRODUCTION

The field of telemedicine is booming and needs the fast and error free communication of medical images to their destination to perform e-consultancy between various specialists to agree upon the correct diagnosis of patient [5]. So in the telemedicine world there is no scope of information loss and delay in transmission. We know the output of most of recently developed medical imaging devices is in digital format. But the issue with these images is that, they require large storage space, and large transmission time [3].

There is one important term regarding images as volumetric images. Volumetric or three dimensional, digital imaging now plays a vital role in many areas of research such as medicine and geology. Medical images acquire by tomographic scanners for instance are often given as a stack of cross sectional image slice. These images are in Digital Imaging and Communication in Medicine (DICOM) form [8]. Such images are called volumetric because they depict objects in their entire 3D extent rather than just a projection onto 2D image plane [9]. Since huge amount of volumetric data are continuously being produced in many places around the world, techniques for their analysis is more important.

We have seen that in recent years 3D medical images (volumetric images) like Magnetic Resonance Imaging (MRI), Computed Tomography (CT) are considered as important part of standard health care. These 3D medical images are now becoming an integral part of patients’ records. As the amount of 3D medical images generated increases, the storage, management, and access to these large repositories is becoming increasingly complex [3]. Typical clinical MRI scanners generate huge data daily. Because this data provides important diagnostic information, care must be taken in compressing it. Also lossy compression is generally avoided. As 3D medical images are large in volume and consist of valuable data, lossless compression is usually the standard in medical imaging to avoid any negative effects on image quality and diagnostic capabilities [9].

Some of the most desirable properties of any compression method for 3D medical images include: 1) high lossless compression ratios, 2) resolution scalability, which refers to the ability to decode the compressed image data at various resolutions, and 3) quality scalability, which refers to the ability to decode the compressed image at various qualities or signal-to-noise ratios (SNRs) up to lossless reconstruction [1].

One simple method to compress images to achieve maximum compression is Moving Picture Expert Group (MPEG). MPEG algorithm compresses data to form small bits that can be easily transmitted and then decompressed. It achieves high compression rate by storing one frame to another, instead of each entire frame. The image is then encoded using technique Discrete Cosine Transform (DCT) [4]. So the technique is very convenient for compression of images. But we cannot use this technique for compression of medical images, because MPEG uses lossy type of compression, so some of the data might be lost [11]. When we are dealing with medical images, we cannot afford to lose any kind of data. So we have to make use of technique which is lossless to perform compression of medical images.

All of we know that human body has vertical symmetry, which means one half part of body is approximate replica of other half part [9]. If we consider axial view of human brain, it is symmetric in nature. In this paper we are going to make use of this symmetry of human anatomy to achieve maximum compression of medical images without any loss. There are some examples of human body organs which possess bilateral symmetry like axial view of brain, pupil, labia, cervical, lumen, chest, thorax, larynx, lungs, etc.
II. PROPOSED 3D COMPRESSION METHOD

In this paper, we propose a novel scalable lossless compression method for 3D medical images that attains the three desired properties listed above and uses the symmetrical characteristics of the data to achieve a higher lossless compression ratio [1]. This method employs RLE based compression of slices within a 3D medical image. Specifically, it encodes slices by first applying principle of symmetry of human anatomy. After finding axis of symmetry it generates residual data by comparing both the parts. Then the residual data is compressed using RLE transform. This compressed data can be now easily used for storage and transmission in Telemedicine. At the receiver side the compressed data is reconstructed back to retrieve original input medical image. This image can be easily used by user which can be doctor or patient as we are dealing with Telemedicine.

III. BLOCK DIAGRAM

The basic block diagram of the system is illustrated in figure 1. The input image is three dimensional medical images like MRI or CT of brain. These images are having axial symmetry. First block performs symmetry detection. After finding axis of symmetry two parts as left part and right part are separated. These two parts are subtracted from each other to get residual data. This residual data is coded using RLE coding to get compressed data.

A. Finding axis of symmetry

Detecting axis of symmetry is an important issue. Several techniques have been proposed for detecting symmetry based on computation of Medical Axis of Transform (MAT) e.g. Hierarchical Voronoi Skeletons, Hamilton-Jacobi2 Transform [10]. The main aim of this project is to compress an image using the symmetry of the image. There are many techniques for extracting axes of symmetry of objects have been proposed, mainly for pattern recognition applications. One of simple and efficient of them is by finding the centroid of the image.

Symmetry is an important consideration in Gestalt laws for perceptual grouping. This describes how people see groups and objects in visual information. The reason is that it is very difficult for two completely unrelated objects to be placed in such a manner as to produce spatial symmetry [10]. It is much more likely that if a group of objects exhibit symmetry, then they are related, and are perceived in relation to one another (i.e., they are grouped together perceptually). This has motivated research in the detection of symmetries in images and shapes. To find the axis of symmetry, we have used the knowledge of the centroid and edges of the images. We have first found out the centroid and the edges of the images using different MATLAB functions. After finding the axis of symmetry of the image, we assume that the image is almost symmetrical, since we are mainly taking into consideration human medical images, and then aim at transmitting only that part of the image along the axis of symmetry which is required for further study or any other specific purpose.

The assumption is based on the fact that we are mainly using human medical images and that the human body is said to be largely completely symmetrical. Once only part of the image is obtained as specified above, we see that a large amount of compression is already obtained.

B. Finding residual data

Human body possesses vertical symmetry, i.e. one half part of human body is approximate replica of other half part. Due to the inherent symmetry of the human anatomy, cross-sections of the ROIs depicted in slices of 3D medical images are typically symmetrical [7]. There are two types of symmetries, global symmetry and local symmetry. Global symmetry refers to the symmetry of the whole sub-band that is nothing but a main axis of symmetry, while local symmetry refers to the symmetry of a small region within the sub-band as defined by a local axis of symmetry (see Figure. 2) [2].

As soon as we got centroide of the image, it is easy to plot axis of symmetry. Along the axis of symmetry the image is divided into two parts. By doing this we partition image in to two areas of equal size as left part (LHi-L) and right part (LHi-R). The part of image to the left side of the axis can be consider as left part and the part of the image to the right of axis as right part. We calculate the difference between the left part and right part that generates the residual data with less energy than the original data.
So we can consider following example of axial view of brain [1]. The figure 4 shown below gives high pass sub-band (LH) of an MRI slice. It is easy to find out main vertical axis of symmetry which is centred in the sub-band. So we can make partition of sub-band into areas along with axis of symmetry. Let’s denote area to the left of the axis as LHi-L and the area to the right as LHi-R. If LHi-L is to be flipped along the axis of symmetry, it would be expected to provide a good approximation to LHi-R and can therefore be used to predict LHi-R. Subband LHi may then be reduced to LHi-L and the prediction error (or residual) between LHi-R and G(LHi-R). So this data i.e. residual data obtained from right part and flipped left part and complete right part is used as input to the encoder which is RLE.

**C. Coding of residual data**

We are using RLE for encoding the data obtained from previous stage. Run-length encoding (RLE) is a very simple form of data compression in which runs of data (that is, sequences in which the same data value occurs in many consecutive data elements) are stored as a single data value and count, rather than as the original run. RLE works by reducing the physical size of a repeating string of characters [12]. This repeating string, called a run, is typically encoded into two bytes. The first byte represents the number of characters in the run and is called the run count. In practice, an encoded run may contain 1 to 128 or 256 characters; the run count usually contains as the number of characters minus one (a value in the range of 0 to 127 or 255). The second byte is the value of the character in the run, which is in the range of 0 to 255, and is called the run value. The main advantage of RLE is that, it performs lossless compression of data.

So at the output of this block we are getting data which is compressed bit stream. Now this data can be easily stored and transmitted.

**D. Reconstruction of original image**

At the receiver side we have to reconstruct image by following inverse of the algorithm which is used for encoding.

**IV. RESULTS**

Part I:

The results shown in table 2 are achieved by performing simple subtraction of 3D medical images.

<table>
<thead>
<tr>
<th>Name</th>
<th>Entropy</th>
<th>Size</th>
<th>Difference</th>
<th>Entropy</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>A83G L6G0</td>
<td>5.28</td>
<td>530</td>
<td>1</td>
<td>178</td>
<td>01</td>
</tr>
<tr>
<td>A83G L6G2</td>
<td>5.60</td>
<td>562</td>
<td>1</td>
<td>1.96</td>
<td>45</td>
</tr>
<tr>
<td>A83G L6G4</td>
<td>5.36</td>
<td>562</td>
<td>2</td>
<td>1.70</td>
<td>83</td>
</tr>
<tr>
<td>A83G L6G6</td>
<td>5.53</td>
<td>547</td>
<td>3</td>
<td>2.75</td>
<td>19</td>
</tr>
<tr>
<td>A83G L6G8</td>
<td>5.38</td>
<td>520</td>
<td>4</td>
<td>3.19</td>
<td>12</td>
</tr>
<tr>
<td>A83G L6G10</td>
<td>5.38</td>
<td>520</td>
<td>4</td>
<td>3.19</td>
<td>12</td>
</tr>
</tbody>
</table>

First of all we convert 3D image into 2D image. The first image is as the reference for reconstruction so it is stored as it is. Second image is now subtracted from first image and only residual part is stored as second image instead of complete second image. Then third image is subtracted from second image and only residual part is stored as third image just like did for second image. And hence so on. So by doing this we can achieve good compression. But this method results in some minor losses which we can not afford while working with medical images. The
same idea we used in our algorithm by making use of anatomic symmetry.

Part II:
Five sets of experimental results are obtained. The table shows the comparison between RLE and our proposed method.

Table 2. Comparison of compression ratios of MRI Images from database

<table>
<thead>
<tr>
<th>3D Medical Image Database (MRI)</th>
<th>Compression Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proposed Symmetry Based technique</td>
</tr>
<tr>
<td>Organ 1 - Brain</td>
<td>3.95:1</td>
</tr>
<tr>
<td>Organ 2 - L. Spine</td>
<td>5.52:1</td>
</tr>
<tr>
<td>Organ 3-Brain</td>
<td>4.47:1</td>
</tr>
<tr>
<td>Organ 4 - T. Spine</td>
<td>2.96:1</td>
</tr>
<tr>
<td>Organ 5 - C. Spine</td>
<td>4.33:1</td>
</tr>
</tbody>
</table>

The results shown in the table 2 gives comparison between the proposed method and the RLE. Five data sets of MRI images are taken. The images are compressed first by the proposed method of symmetry and then compression ratio is calculated. Then images are compressed by only RLE and again compression ratio is calculated. The table shows that our proposed method gives good results over RLE.

The following results are obtained by the method proposed in the paper.

The input images shown in figure 5 to which the above coding algorithm is applied for the generation of codes and then decompression algorithm is applied to get the original image back from the generated codes, which is shown in the figure 6. The output image is the decompressed image i.e. from the figure 6 it is clear that the decompressed image is approximately same as the input image. So our approach is to make use of this method which is proposed in the paper to reduce complexity in achieving compression of 3D medical images. Also we will try to increase compression ratio of medical images especially MRI of brain. So we will reconstruct the original image without any loss in an image that is using lossless compression of an image.

V. FUTURE WORK

In this paper we have got comparative results for proposed method and RLE. Our next approach is to use several different encoding techniques to compress and reconstruct medical images without any loss. The comparative study will help to choose better method.

VI. CONCLUSION

This paper focused on the evaluation of several commonly used algorithms for lossless compression. The proposed method is a 3D scalable lossless compression of medical image data which is based on global and local symmetries. This method uses anatomic symmetries present in the medical images and performs the compression. The compressed data is then decompressed to get original image back.

So our aim is to increase compression ratio with reducing complexity while achieving compression.

REFERENCES


[10] Gert J. van Tonder, Yoshimich Ejima, “Flexible computation of shape symmetry within the maximal disk paradigm” IEEE transaction on system, 2002


Abstract. The common handwriting recognition methods are Statistical, syntactical and structural, neural network and elastic matching. Support vector Machine (SVM) is a concept in statistics and computer science for a set of related supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis. This paper presents implementation of Support Vector Machine to recognize online handwriting Gurmukhi strokes. Preprocessing of Gurmukhi strokes consists of 5 basic algorithms for preprocessing. Prior to these algorithms, a basic step called Stroke Capturing is done. In this paper we have discussed a simple way to store data for Gurmukhi strokes and two cross validation techniques, namely, holdout and k-fold strategies for recognition of Gurmukhi strokes. These strokes are taken from the one hundred Punjabi words written by 2 writers.

Keywords: Online Handwriting Recognition, Recognition, Gurmukhi Strokes, Support Vector Machine.

I. INTRODUCTION

Nowadays, for devices like personal digital assistant or tablets, online handwriting recognition has become a useful feature. With this feature the demand of tablets increase as per the rate of their use has increase. To improve user and computer communication online handwriting recognition has great potential. Handwriting recognition is the most exciting and empirical research for image processing and pattern recognition. Handwriting recognition system is the activity through which user transmit the handwritten data into the computer through the sources of scanner, touch screens and other devices, where computer interpret that input into its language. Handwriting recognition systems are broadly divided into two: namely offline and online handwriting recognition systems. Statistical, syntactical and structural, neural network and elastic matching are the common handwriting recognition methods [1] [2]. Very few attempts have been made to build an online handwriting recognition system for Gurmukhi script. This paper has been done recognition using support vector machine using strokes written in Gurmukhi script written by two writers.

II. GURMUKHI SCRIPT

The word “Gurmukhi” literally means “from the mouth of guru”. Gurmukhi script is used primarily for the Punjabi language, which is the world’s 14th most widely spoken language [3]. Gurmukhi characters are shown in Fig. 1. Gurmukhi script is written in left to right direction and in top down approach [4]. Most of the Characters have a horizontal line at upper part. The characters of words are connected mostly by this line called head line.

Fig. 1: Gurmukhi Characters.

The next section discusses about recognition of Gurmukhi strokes using support vector machine of Gurmukhi strokes.

III. RECOGNITION

Preprocessing phase in handwriting recognition is applied to remove noise or distortions. In the process of handwriting recognition, it is important to identify correct features [5]. Computational complexity of a classification problem can also be reduced if suitable features are selected. In this paper Points generated after preprocessing phase are used as a feature for recognition; these points are always fixed for each stroke and are equidistant as far as possible. In this work number of points in a stroke is fixed to 64. Different strokes of Gurmukhi characters from two different writers are collected using the I-Ball electromagnetic digital pen Tablet which is having 1024 levels of pressure sensitivity, resolution 400 lpi, 200 rps report rate, and have ±0.01 inch accuracy. Each of the two writers has written same 100 words of Gurmukhi script. The words are stored as a collection of strokes in an xml format and the snapshot of an xml file is shown below in Fig. 2. There is a single xml file for each word and the strokes of each word are stored as a collection of points and only the stroke are considered for
3.1 Support Vector Machine

A Support Vector Machine (SVM) is a concept in statistics and computer science for a set of related supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis [6]. The advantage of SVM, is that it takes into account both experimental data and structural behavior for better generalization capability based on the principle of Structural Risk Minimization (SRM). The basic SVM formulation is for linearly separable datasets. It can be used for non-linear datasets by indirectly mapping the nonlinear inputs into to linear feature space where the maximum margin decision function is approximated. The mapping is done by using a kernel function. A Support Vector Machine includes Support Vector Classifier (SVC) and Support Vector Regressor (SVR).

3.2 Recognition using MATLAB

MATLAB is a high performance language for technical computing created by The MathWorks. It features a family of add-on application-specific solutions called toolboxes (i.e., comprehensive collections of functions) that extend the MATLAB environment to solve particular classes of problems. In this work MATLAB version 7.10.0.499(R2010a) with bioinformatics toolbox is used for recognition of Gurmukhi stroke. Bioinformatics toolbox offers an integrated software environment for genome and proteome analysis. In this work two methods holdout and k-fold cross validation is applied and are discussed in next sections.

3.2.1 Hold out cross validation for training and testing

The holdout method is the simplest kind of cross validation. The data set is separated into two sets, called the training set and the testing set. The errors it makes are accumulated as before to give the mean absolute test set error, which is used to evaluate the model. The advantage of this method is that it is usually preferable to the residual method and takes no longer to compute. However, its evaluation can have a high variance. In MATLAB crossvalind function is used. This function is used to generate cross validation indices. Prototype of this cross validation function is crossvalind('HoldOut', N, P). It returns logical index vectors for cross-validation of N observations by randomly selecting P*N (approximately) observations to hold out for the evaluation set. P must be a scalar between 0 and 1. P defaults to 0.5 when omitted, corresponding to holding 50% out. 3.2.2 k-Fold cross-validation for Training and Testing   K-fold cross validation is one way to improve over the holdout method. The data set is divided into k subsets, and the holdout method is repeated k times. Each time, one of the k subsets is used as the test set and the other k-1 subsets are put together to form a training set. Then the average error across all k trials is computed. The advantage of this method is that it matters less how the data gets divided. Every data point gets to be in a test set exactly once, and gets to be in a training set k-1 times. The variance of the resulting estimate is reduced as k is increased. In MATLAB crossvalind('Kfold', N, k) function is used for Gurmukhi stroke recognition. It returns randomly generated indices for a k-fold cross-validation of N observations. Indices contain equal (or approximately equal) proportions of the integers 1 through k that define a partition of the N observations into k disjoint subsets. Repeated calls return different randomly generated partitions. k defaults to 5 when omitted. In k-fold cross validation, folds are used for training and the last fold is used for evaluation. This process is repeated times, leaving one different fold for evaluation each time.

In the next section the result after applying two partitioning strategies of cross validation in SVM is applied on the data collected from two different writers consisting 1685 strokes is discussed.

IV. DISCUSSIONS AND RESULTS

This section contains the results of the experiments carried out for recognition of Gurmukhi strokes written by 2 writers. Each writer has written 100
words in Gurmukhi script. These words are stored in an xml format as described in Section 3. Table 1, given below, shows the frequency of each stroke written by two writers. Each stroke is identified by a unique ID in this Table.

Table 1: Frequency of strokes for two writers

<table>
<thead>
<tr>
<th>Stroke ID</th>
<th>Writer 1</th>
<th>Writer 2</th>
<th>Stroke ID</th>
<th>Writer 1</th>
<th>Writer 2</th>
<th>Stroke ID</th>
<th>Writer 1</th>
<th>Writer 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>40</td>
<td>40</td>
<td>148</td>
<td>9</td>
<td>9</td>
<td>173</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>102</td>
<td>19</td>
<td>20</td>
<td>149</td>
<td>10</td>
<td>9</td>
<td>174</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>104</td>
<td>-</td>
<td>1</td>
<td>151</td>
<td>37</td>
<td>40</td>
<td>179</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>105</td>
<td>2</td>
<td>-</td>
<td>154</td>
<td>17</td>
<td>-</td>
<td>181</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>121</td>
<td>91</td>
<td>06</td>
<td>155</td>
<td>20</td>
<td>29</td>
<td>182</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>122</td>
<td>15</td>
<td>15</td>
<td>156</td>
<td>-</td>
<td>-</td>
<td>183</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>124</td>
<td>2</td>
<td>2</td>
<td>157</td>
<td>23</td>
<td>26</td>
<td>185</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>126</td>
<td>10</td>
<td>10</td>
<td>158</td>
<td>4</td>
<td>1</td>
<td>186</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>127</td>
<td>12</td>
<td>12</td>
<td>159</td>
<td>-</td>
<td>2</td>
<td>189</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>128</td>
<td>12</td>
<td>12</td>
<td>161</td>
<td>18</td>
<td>17</td>
<td>191</td>
<td>25</td>
<td>1</td>
</tr>
<tr>
<td>129</td>
<td>3</td>
<td>10</td>
<td>162</td>
<td>35</td>
<td>144</td>
<td>192</td>
<td>-</td>
<td>28</td>
</tr>
<tr>
<td>141</td>
<td>-</td>
<td>23</td>
<td>163</td>
<td>0</td>
<td>0</td>
<td>193</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>142</td>
<td>20</td>
<td>-</td>
<td>164</td>
<td>26</td>
<td>34</td>
<td>194</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>144</td>
<td>33</td>
<td>-</td>
<td>165</td>
<td>10</td>
<td>1</td>
<td>197</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>145</td>
<td>2</td>
<td>56</td>
<td>170</td>
<td>2</td>
<td>2</td>
<td>199</td>
<td>45</td>
<td>58</td>
</tr>
<tr>
<td>146</td>
<td>32</td>
<td>34</td>
<td>171</td>
<td>-</td>
<td>-</td>
<td>202</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>147</td>
<td>14</td>
<td>14</td>
<td>172</td>
<td>1</td>
<td>1</td>
<td>204</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stroke ID</th>
<th>Writer 1</th>
<th>Writer 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>205</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>212</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>213</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stroke ID</th>
<th>Writer 1</th>
<th>Writer 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>215</td>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>216</td>
<td>27</td>
<td>26</td>
</tr>
<tr>
<td>226</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stroke ID</th>
<th>Writer 1</th>
<th>Writer 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>229</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Graph 1: Recognition accuracy of three writers for preprocessed strokes

For the recognition process, we have used SVM classifier and this is implemented in MATLAB. SVM is applied to preprocessed data. Recognition is done for each writer separately. Writer 1 has used 810 strokes and writer 2 has used 875 strokes for writing 100 words in Gurmukhi scripts as shown in Table 1. Recognition has been done for only those strokes that are written by writer 1 only and are present in the strokes written by writer 2 also. Recognition has also been done using 1000, 2000 and 3000 strokes of preprocessed data for each writer. The subsequent section 4.1 gives the recognition accuracy for preprocessed data using holdout and k-fold cross validation techniques and

4.2 gives the result for recognition using 1000, 2000 and 3000 preprocessed strokes.

4.1 Recognition using holdout cross validation
For preprocessed data, three types of partitions (70% training, 60% training and 50% training) have been used for recognition. These partitions are used for recognition of stroke IDs, for individual writer.

From the above graph, it has been observed that the average recognition accuracy for 50% training for writer 1 is 98.38% and for writer 2 is 98.47%, for 60% training recognition accuracy for writer 1 is 98.44% and for writer 2 is 98.46% and for 70% training dataset recognition accuracy for writer 1 is 98.26% and for writer 2 is 98.29%. From Graph 1, it has been observed that writer 1 is the inconsistent writer and writer 2 is the best one. So here it can be determined that there is minimum variation in handwriting of writer 2. For preprocessed data another technique for 10-fold cross validation is used for recognition and After applying this cross
validation technique it has been observed that the average recognition accuracy for writer 1 is 98.75% and for writer 2 is 98.82%. So it can be concluded that writer 2 is the most accurate writer as compared to writer 1.

4.2 Recognition results on specific number (1000, 2000 and 3000) of strokes

In this study 1000, 2000 and 3000 preprocessed strokes are considered from each writer. These preprocessed strokes were taken randomly from the stroke database. The results are calculated only for those stroke IDs present in 1000 strokes. Graph 2 shows recognition accuracy of 1000, 2000 and 3000 preprocessed strokes for two different writers using 10-fold cross validation technique. 1000, 2000 and 3000 represents 1000, 2000 and 3000 strokes respectively.

Graph 2: Recognition accuracy of three writers for 1000, 2000 and 3000 preprocessed strokes

For each writer, recognition using 3000 preprocessed strokes gives maximum accuracy as compared to 1000 and 2000 preprocessed strokes. Out of the two writers, writer 1 gives highest accuracy for 1000, 2000 and writer 2 gives highest for 3000 preprocessed strokes.

V. CONCLUSIONS

This research work focuses on developing a stroke recognition system for Gurmukhi script using SVM. One hundred words from three different writers are considered and each of these words is divided into strokes. Writer 1 writes 810 strokes, writer 2 writes 875 strokes for the same 100 words. SVM is used to perform Gurmukhi stroke recognition. Recognition has been done using 1000, 2000 and 3000 strokes of preprocessed data for each writer, other than this two different partitioning techniques holdout and 10-fold cross validation have been used for generating training and testing sets. Recognition is done for 70%, 60% and 50% training sets in holdout cross validation. For 70%, 60% and 50% training set of preprocessed stroke in holdout cross validation method, writer 2 has better accuracy 98.47%, 98.46% and 98.29% respectively. For 10-fold cross validation writer 2 attains 98.53% for preprocessed strokes. For 1000, 2000 and 3000 preprocessed strokes are considered and recognition accuracy for each writer is examined. For writer 1 and writer 2 recognition accuracies 99.16%, and 99.26% are achieved with 3000 preprocessed strokes which are higher as compared to 1000 and 2000 preprocessed strokes. Writer 2 attains better recognition accuracy 99.26% amongst all the writers for 1000, 2000 and 3000 preprocessed strokes. From this, it can be concluded that writer 2 has achieved maximum accuracy for Gurmukhi strokes in each situation. So it can be concluded that writer 2’s handwriting is better writer as compared to others.

VI. ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge Dr. R.K. Sharma and Mayank Gupta for providing data, feedback for this research.

REFERENCES

PROFIT AND PENALTY AWARE SCHEDULING FOR REAL-TIME ONLINE SERVICES

POOJA PRIYA¹, SNEHAL SAOKAR² & PRIYA SHINDE³

¹,²&³“Department Of Information Technology, Bharati Vidyapeeth’s College of Engineering for Women, Pune, India

Abstract—As computer and Internet technology continue to advance, real-time online services are emerging. Different from traditional real-time applications for which the scheduling objective is to meet task deadlines the optimization goal for online service systems is to maximize profit obtained through providing timely services. For this class of applications, there are two distinctive characteristics. First, tasks are associated with a pair of time dependent functions representing accrued profit when completed before their deadlines and accrued penalty otherwise, respectively. Second, the service requests or tasks arrive periodically with execution time varying in a wide range. This paper presents a novel scheduling method and related analysis for such applications. Two scheduling algorithms, i.e., the non preemptive and preemptive Profit and Penalty aware (PP-aware) scheduling algorithms, are proposed with an objective to maximize system’s total accrued profit. Our simulation results clearly demonstrate the advantages of the proposed algorithms, with respect to the system total accrued profit, over other commonly used scheduling algorithms, such as Earliest Deadline First (EDF) and Utility Accrual (UA) algorithms.

I. INTRODUCTION

For real-time e-service applications, there is usually a service level agreement established between a client and a service provider. Service providers agree to guarantee services with certain levels of quality-of-service, such as timeliness: in return, clients pay for the services completed before their requested deadlines. However, missing a deadline normally does not cause a catastrophic result; instead, penalties are applied if timing requirements are not met. Furthermore, depending on how fast a request is served, or how soon a client is notified the termination of his/her request due to the inability to meet the required deadlines, the price paid by the client or the penalty suffered by the service provider may be different. To a client, besides the quality of the deliverables, the timeliness is a major criterion in judging the quality-of-service provided by the service provider while to a service provider, the goal is to maximize its long term profits attained under uncertain application workloads. For illustrative purpose, consider a real-time e-service request given below. Example 1: A traveler from city A to city B requests a travel plan from an online travel planning service provider and indicates a deadline he/she is willing to wait. Depending on when the traveler gets the requested travel information from the service provider, the payment is different—the longer it takes, the less he/she pays. In other words, the payment is given only when the request is successfully completed before its deadline and the amount is decided by a non increasing time dependent payment function. Furthermore, the longer the client waits fruitlessly, the more penalty the system suffers, either by paying more monetary compensation to the client, or losing future service requests from unsatisfied clients. Specifically, the penalty occurs when a request is unfulfilled and it is a non decreasing time dependent function.

II. PP-AWARE SCHEDULING OF SINGLE VEP TASK

Under the VEP model, task execution time is not known a priori. Therefore, before and during the execution of a VEP task, we face choices regarding whether to continue or abort the task. On one hand, the longer we execute the task, the closer we are to the completion point of the task and hence have a high possibility to make profit. On the other hand, due to the non increasing property as well as non decreasing property, the longer the task executes, the less profit we make even if the task finishes before its deadline. We may even have to pay higher loss if it cannot meet its deadline. Hence, in order to maximize the gain for executing a task, a judicious decision as to continue or to abort the task with the consideration of both gain and loss function.

III. NONPREEMPTIVE PP-AWARE SCHEDULING

When VEP tasks are accepted to the pending queue, the problem becomes how to make appropriate scheduling decisions to maximize the system’s profit. For non preemptive PP-aware scheduling, this scheduling decision is made at the time when:

1) A task successfully completes its execution before its deadline.
2) The current time reaches the critical time of the task being executed.

As penalty increases and gain decreases with time, our non preemptive PP-aware scheduling method selects a task with the highest expected gain and executes it only to its critical time, aborts and removes a task as soon as it has a higher-than-tolerable risk to cause a loss. Therefore, at each
scheduling point, we not only choose the next task to execute, but also further check whether the current selection raises the risks of other tasks beyond the tolerable level, and remove them as soon as possible.

IV. PREEMPTIVE PP-AWARE SCHEDULING

If preemption is allowed and a task with higher potential gain emerges, the current executing task will be preempted by the one with higher gains. For preemptive PP-aware scheduling, scheduling decisions are made and preemption points are calculated at the time when:
1) a new task arrives;
2) a task successfully completes its execution before its deadline;
3) the current time reaches previously calculated preemption point or critical time or its deadline, whichever is the earliest.

These time instances are called scheduling points. It is worth pointing out that task arrival times are also scheduling points for preemptive PP-aware algorithm which is different from the non preemptive PP-aware scheduling algorithm. The reason is that for non preemptive PP-aware scheduling algorithm, newly arrived tasks cannot preempt the current task even if they have higher priorities and hence do not call for scheduling decisions. While for preemptive algorithm, current task may be preempted by newly arrivals. Therefore, task arrival times are also scheduling points for the preemptive PP-aware scheduling algorithm.

V. RELATED WORK.

For conventional real-time systems, such as process control systems, the major concern is to meet task deadlines. Extensive research has been conducted in designing and analyzing scheduling algorithms for such purposes. The deadline monotonic and rate monotonic scheduling algorithms are among the most widely studied algorithms and form the basis against which other algorithms are compared. However, with these classes of algorithms, the task’s urgency, rather than its importance, is used as job scheduling criteria.

VI. CONCLUSION

This paper presents a simulation of a Profit and Penalty Aware Scheduling for Real-Time Online Services.

ACKNOWLEDGEMENT

First and foremost, I would like to thank my guide, Prof. Mrs S. M. Murgeon, for her guidance and support. I will forever remain grateful for the constant support and guidance extended by guide, in making this report. Through our many discussions, she helped us to form and solidify ideas. The invaluable discussions I had with her, the penetrating questions she has put to me and the constant motivation, has all led to the development of this project.

REFERENCES


Fig: System Architecture
“NETWORKING AND DATA AGGREGATION CONTROL FOR VANET”
AMEETA B RANGARI & Dr. GIRISH AGRAWAL
12RTMNU NAGPUR, INDIA

Abstract—Traditional data aggregation schemes for wireless sensor networks usually rely on a fixed routing structure to ensure data can be aggregated at certain sensor nodes. In-network data aggregation is a useful technique to reduce redundant data and to improve communication efficiency however. They cannot be applied in highly mobile vehicular environments. We will try to implement an adaptive forwarding delay control scheme, namely Catch-Up, which dynamically changes the forwarding speed of nearby reports so that they have a better chance to meet each other and be aggregated together. The Catch-Up scheme to be designed will be based on a distributed learning algorithm where each vehicle learns from local observations and chooses a delay based on learning results.

I. INTRODUCTION:

Vehicular Ad-Hoc Network, or VANET is a technology that uses moving cars as nodes in a network to create a mobile network. VANET turns every participating car into a wireless router or node, allowing cars approximately 100 to 300 meters of each other to connect and, in turn, create a network with a wide range. As cars fall out of the signal range and drop out of the network, other cars can join in, connecting vehicles to one another so that a mobile Internet is created. It is estimated that the first systems that will integrate this

Technology are police and fire vehicles to communicate with each other for safety purposes. VEHICULAR Ad hoc Networks (VANETs) have been regarded as an emerging and promising field in both industry and academia. It has potential to improve the efficiency and safety of future highway systems. One good example is traffic congestion detection. One good example is traffic congestion detection. Once a vehicle detects the number of its neighboring vehicles exceeds a certain limit, it will broadcast a warning to the vehicles following behind. This warning could travel a rather long distance so that vehicles, possibly several kilometers away, can have enough time to choose an alternate route. Since each vehicle in VANETs is able to detect traffic conditions and generate traffic reports distributed and independently.

II. LITERATURE SURVEY:

In sensor networks, researchers have proposed a number of structure-based aggregation schemes. Which rely on a fixed routing tree to ensure reports can be merged at the tree forks. Certainly, these schemes are not suitable for dynamic vehicular environments. Fan et al. proposed a structure-free aggregation protocol based on randomized waiting. However, this probabilistic approach cannot guarantee the aggregation of all reports from a single event source. Later, the authors presented a semi-structured approach to improve the aggregation degree, but this approach still needs to maintain a routing structure. Several VANET projects, such as Self-Organizing Traffic Information System (SOTIS) and Traffic View, use periodical broadcasting for data aggregation. As time elapses, neighboring vehicles can get an overview of current traffic conditions in the vicinity by periodically exchanging information. However, content exchange based on periodical broadcasting may not be an efficient way in terms of communication overhead, and what is an optimal broadcast interval is still an unsolved issue. VANET offers several benefits to organizations of any size. While such a network does pose certain safety concerns (for example, one cannot safely type an email while driving), this does not limit VANET’s potential as a productivity tool. GPS and navigation systems can benefit, as they can be integrated with traffic reports to provide the fastest route to work. A commuter can turn a traffic jam into a productive work time by having his email downloaded and read to him by the on-board computer, or if traffic slows to a halt, read it himself. It would also allow for free, VoIP services such as Google Talk or Skype between employees, lowering telecommunications costs. In VANET, or Intelligent Vehicular Ad-Hoc Networking, defines an intelligent way of using Vehicular Networking. In VANET integrates on multiple ad-hoc networking technologies such as Wi-Fi IEEE 802.11p, WAVE IEEE 1609, WiMAX IEEE 802.16, Bluetooth, IRA, Zig Bee for easy, accurate, effective and simple communication between vehicles on dynamic mobility. Effective measures such as media communication between vehicles can be enabled as well as methods to track the automotive vehicles. In VANET helps in defining safety measures in vehicles, streaming communication between vehicles, infotainment and telematics. In VANETs, several studies have implemented data aggregation mechanisms. In the Self-Organizing Traffic Information Systems, vehicles on a road segment
periodically send out reports containing the traffic information on the current road segment. During the broadcast interval, a vehicle collects and aggregates information received from neighboring vehicles. This approach helps generate an overview of current traffic conditions by periodical broadcasting.

III. PROPOSED SYSTEM:

In this paper, we propose an adaptive forwarding delay control scheme for VANETs, namely, Catch-Up. The basic idea of the Catch-Up scheme is to insert an adaptive delay before forwarding a report to the next hop. Since reports are propagated along a straight road segment, when different delays are applied to different reports, nearby reports can have a better chance to meet each other and be aggregated together. We formulate in-network data aggregation in VANETs as a distributed learning problem. Vehicles cooperate with each other to increase the chance of report encounters. We propose a new paradigm of distributed learning, “Learning-From-Others.” Traditional distributed learning encourages local information exchange to facilitate the learning process.

IV. RESEARCH OBJECTIVE:

A) To formulate an adaptive forwarding delay control scheme for VANETs;
B) To formulate a distributed Markov Decision Process (MDP) model to be designed for individual vehicles with the objective of improving global performance through distributed cooperation;
C) To formulate a Q-learning-based scheme to reduce the exchange of entire local knowledge base;
D) Finally we propose a fuzzy rule base function approximation to speed up the learning process;
E) Simulation of the scheme using NS2 simulator.

V. ALGORITHM:

We introduce the Catch-Up scheme—an adaptive forwarding delay control scheme for VANET data aggregation. The scheme is designed based on a customized distributed learning algorithm. The MDP model works as follows: Suppose that we have vehicles I and j. Before vehicle I forward a report to vehicle j, vehicle I attaches its local variable si; si; ai to the report (as shown in (1)). After vehicle j receives the report, it calculates the reward Ri with (2). Evidently, reward Ri is obtained based on vehicle j action a is. Vehicle j then uses s i; a i to update its local knowledge base and also calculates the optimal action (WALK or RUN) based on its local knowledge base. After a delay (WALK or RUN), vehicle j forward the report to the next hop. Please note that in this process vehicle j uses vehicle j’s action/reward pair to update its local knowledge base. In other words, vehicle j learns from other vehicles’ action/reward pairs. This important feature guarantees that little extra communication overhead is introduced by the learning process.

Design the simulation scenario and to generate mobility trace files using NS2.

REFERENCES:

Networking and data aggregation control for vanet


FleetNet–Internet on the Road, http://www.fleetnet.de/

http://vanet.info/projects

AN EFFICIENT AUDIO WATERMARK EMBEDDING & EXTRACTION DWT METHOD WITH NOISE REMOVING TECHNIQUE

MUNNA VERMA¹ & LAKHWINDER KAUR²

¹(Digital Electronics), RCET Bhilai(C.G)
²Reader, Electronics & Tele-communication Engg, RCET Bhilai(C.G)

Abstract-Digital Audio watermarking is now drawing attention as a new method of protecting multimedia content from unauthorized copying. This paper proposes a new watermarking system using LSB method for copyright protection of digital contents and noise suppression method for watermark recovery. In our proposed watermarking system, the original audio is segmented into non-overlapping frames. Watermarks are then embedded into the highest prominent peak in the magnitude spectrum of each frame. Watermarks are extracted by performing the inverse operation of watermark embedding process. Simulation results indicate that the proposed watermarking system is highly robust against various kinds of attacks such as noise addition, cropping, re-sampling, re-quantization, and MP3 compression, and achieves similarity values ranging from 13 to 20. In addition, our proposed system achieves low SNR.

Keyword: Copyright protection, digital watermarking, multimedia contents, LSB, Noise suppression, Compression & rarefaction region

1. INTRODUCTION

The recent growth in computer networks, and more specifically, the World Wide Web, copyright protection of digital audio becomes more and more important. Digital audio watermarking has drawn extensive attention for copyright protection of audio data. A digital audio watermarking is a process of embedding watermarks into audio signal to show authenticity and ownership. Audio watermarking should meet the following requirements:

(a) Imperceptibility: the digital watermark should not affect the quality of original audio signal after it is watermarked;
(b) Robustness: the embedded watermark data should not be removed or eliminated by unauthorized distributors using common signal processing operations and attacks;
(c) Capacity: capacity refers to the numbers of bits that can be embedded into the audio signal within a unit of time;
(d) Security: security implies that the watermark can only be detectable by the authorized person. All these requirements are often contradictory with each other. However, it should satisfy the important properties such as imperceptibility and robustness.

In this paper, we propose a new watermarking system using LSB for audio copyright protection and extraction with noise suppression method. The watermarks are embedded into the highest prominent peak of the magnitude spectrum of each non-overlapping frame. Experimental results indicate that the proposed watermarking system provides strong robustness against several kinds of attacks such as noise addition, cropping, re-sampling, re-quantization, and MP3 compression and achieves similarity values ranging from 13 to 20. In addition, our proposed system achieves low SNR.

2. LITERATURE REVIEW

A significant number of watermarking techniques have been reported in recent years in order to create robust and imperceptible audio watermarks. Lie et al. [1] propose a method of embedding watermarks into audio signals in the time domain. The proposed algorithm exploits differential average-of-absolute-amplitude relations within each group of audio samples to represent one-bit information. It also utilizes the low-frequency amplitude modification technique to scale the amplitudes in selected sections of samples so that the time domain waveform envelope can be almost preserved. In [2], authors propose a blind audio watermarking system which embeds watermarks into audio signal in time domain. The strength of the audio signal modifications is limited by the necessity to produce an output signal for watermark detection. The watermark signal is generated using a key, and watermark insertion depends on the amplitude and frequency of audio signal that minimizes the audibility of the watermarked signal. Ling et al. [3] introduce a watermarking scheme based on non uniform discrete Fourier transform (NDFT), in which the frequency points of embedding watermark are selected by the secret key. Zeng et al. [4] describe a blind watermarking system which embeds watermarks into DCT coefficients by utilizing quantization index modulation technique. In [5], the authors propose a watermarking system which
embeds synchronization signals in time domain to exist against several attacks. Pooyan et al. [6] introduce an audio watermarking system which embeds watermarks in wavelet domain. The watermarked data is then encrypted and combined with a synchronization code and embedded into low frequency coefficients of the sound in wavelet domain. The magnitude of quantization step and embedding strength is adaptively determined according to the characteristics of human auditory system. Wang et al. [7] proposes a blind audio watermarking scheme using adaptive quantization against synchronization attack. In addition, the multi resolution characteristics of discrete wavelet transform (DWT) and the energy compression characteristics of discrete cosine transform (DCT) are combined in this scheme to improve the transparency of digital watermark. Watermark is then embedded into low frequency components by using adaptive quantization according to human auditory system. In [8], authors propose a watermarking system in cepstrum domain in which a pseudo-random sequence is used as a watermark. The watermark is then weighted in the cestrum domain according to the distribution of cepstral coefficients and the frequency masking characteristics of human auditory system. Liu et al. [9] propose a blind watermarking system which takes the advantages of the attack-invariant feature of the cepstrum domain and the error-correction capability of BCH code to increase therobustness as well as imperceptibility of audio watermarking. In Cox’s method [10] watermarks are embedded into the highest m DCT coefficient of the whole sound excluding the DC component by the following equation:

\[ v'_i = v_i(1 + \alpha x_i) \]  

--- Ex 2.1

Where, m is the length of the watermark sequence, \( v_i \) is a magnitude coefficient into which a watermark is embedded; \( x_i \) is a watermark to be inserted into \( v_i \), \( \alpha \) is a Scaling factor and \( v'_i \) is an adjusted magnitude coefficient. The watermark sequence is extracted by performing the inverse operation of (1) represented by the following equation:

\[ x'_i = \frac{v'_i}{\alpha} - 1 \]  

--- Ex 2.2

3. PROPOSED ALGORITHM

3.1 Analysis of Problem

Commonly, data embedding has two general techniques which are, digital watermarking and steganography. According to the researchers; data embedding approaches have two main limitations, the size of the embedded data and the robustness of the watermark techniques.

3.2 Objectives

1. To analyze the features of audio file that can be used to implement the high rate data embedding;
2. To investigate the approaches used in audio watermarking domains, audio environment for implementing a secure, robust and high rate data embedding in the audio files i.e. try to reduce noise from audio signal;
3. To carry out intensive literature reviews of the existing techniques;
4. To identify the software metrics used to evaluate the audio watermarking approaches in data embedding.

(Watermark Hiding)

2. Check for PCM encoding of an audio wave file.
3. Divide an input into its compressed & rarefaction region using discrete wavelet Transform.
4. Select Watermark Text
5. Embed Watermark Text into rarefaction region of audio wave file.
6. Analyze noisy audio sample & noise.
7. Generate result audio stream with watermark bits embedded into it.
8. Save result Audio File.

(Watermark Extraction)

1. Select Input audio wave file.
2. Divide an audio into compression & rarefaction region using discrete wavelet Transform.
3. Analyze noisy samples of a rarefaction region.
4. Remove noise from noise samples using sample averaging method & passes audio stream from hearing test.
5. Read bit of corrected noisy samples & assemble it.
6. Decode watermark data.
7. Save result & Exit.

Figure 3.1 Data Flow diagram(Watermark Embedding)
The major advantage of Low-bit encoding are:
1. High watermark channel bit rate
2. Low computational complexity of the algorithm compared with others techniques
3. No computationally demanding transformation of the host signal, therefore, it has very little algorithmic delay. The major disadvantage is that the method are:
1. Low robustness, due to the fact that the random changes of the LSB destroy the coded watermark
2. It is unlikely that embedded watermark would survive digital to analogue and subsequent analogue to digital conversion

4. RESULT ANALYSIS

In this section, we evaluate the performance of our watermarking system for four different types of 16 bit mono audio signals sampled at 44.1 kHz:
- The song ‘Let it Be,’ by the Beatles;
- The beginning of Symphony No. 5 in C Minor, Op. 67, by Ludwig van Beethoven;
- An instrumental song ‘Hey Jude’ played by a Korean traditional musical instrument called the gayageum;
- A human voice providing TOEIC (Test of English for International Communication) listening test instruction. Each audio file contains 206,336 samples (duration 4.679 sec). By considering the frame size of 512 samples, we have 403 frames for each audio sample.

From each frame we detect 1 peak to embed watermark.

Thus, the length of the watermark sequence is 403×1=403.

In order to evaluate the performance of the proposed watermarking system in terms of watermark detection, the correlation coefficient between the original watermark X and the extracted watermark X* is calculated by the following similarity SIM(X, X*) formula:

\[
SIM(X, X^*) = \frac{X \cdot X^*}{\sqrt{X^2 \cdot X^2}}
\]

It is highly unlikely that X* will be identical to X. To decide whether X and X* match, we determine whether the SIM(X, X*) > T, where T is a detection threshold. Fig. 3 shows a qualitative evaluation of the original audio with a watermarked audio in which the watermarks are imperceptible using the proposed system.
Imperceptibility Test Informal listening using head reveals that the watermark embedded into the original audio using the proposed watermarking system does affect the quality of the sound up to little bit extent, which ensures the imperceptibility of the embedded watermark.

CONCLUSION

In this paper, we have presented a new watermarking system using LSB with noise suppression for copyright protection of sound contents. Experimental results indicate that our proposed watermarking system shows strong robustness against several kinds of attacks such as noise addition, cropping, re-sampling, re-quantization, and MP3 compression and achieves similarity values ranging from 13 to 20. In addition, our proposed system achieves SNR values ranging from 20 dB to 28 dB for different watermarked sounds. These results demonstrate that our proposed watermarking system can be a suitable candidate for audio copyright protection.

REFERENCES

[14] Levent M. Arslan and Umut Ulu dag “Audio watermarking using DC level shifting” Advanced topics in speech processing using DC level shifting, Jan 2001
NETWORK SECURITY THROUGH PENETRATION TESTING

PRIYANKA C. SHENDGE¹, SAMRUDHI L. KHAIRNAR² & PALLAVI S. MANE³

¹,² & ³ BVCOEW, Pune (India)

Abstract-Each organization needs to effectively deal with major security concerns, forming a security policy according to its requirements and objectives. In anonymizing networks which provides free access to network usually hides clients IP address. Unfortunately some users have misused such networks by using it for abusive purpose. To block such a users administrators relay on router blocking or IP address blocking for disabling the misbehaving user. To recover this problem we present penetration test in which server blacklists the misbehaving user for particular time period using clients MAC address. This proactive approach is usually interpreted wrongly in only up-to-date software and hardware. Regular updates are necessary, although, not enough, because potential mis-configurations and design flaws cannot be located and patched, making the whole network vulnerable to attackers. In this paper we present how a comprehensive security level can be reached through extensive Penetration Tests.

Keywords-Penetration testing, Network Security, Nymble.

I. INTRODUCTION

Every modern issue has to effectively deal with the security issues that arise from the technologies. An organization that truly wants to adapt a proactive approach, aggressively seeks out all types of vulnerabilities by using relevant methods with the actual hackers. This process of systematically and actively testing a deployed network to determine potential vulnerabilities is called penetration testing.

The penetration testing is divided into four phases: planning, discovery, exploitation and reporting.

- Planning phase- Initially the scope for the assignment is defined. The flow of the test is defined. After the management consent, the penetration testing team gathers crucial input about the organization operational procedures and security policies, towards defining the scope for the test.

- Discovery phase- It is also known as information gathering phase. During the information gathering process the penetration testing team launches scanning and enumeration procedures to gain as much information as possible about the target network and the participating systems and services. The gathering phase can be further divided into non-intrusive (public repositories, documents, mailing lists, web profiles etc) and intrusive (port scanning, firewall rules, matching OS fingerprints etc) inspection processes. Having adequate amount of information the testing team can profile the target network and enumerate possible exploitable vulnerabilities.

- Exploitation phase- Using as input the discovered vulnerabilities arriving from the previous phase, the penetration testing team revises matching proof-of-concept exploits that may lead to a network or service security bridge. While exploiting network vulnerabilities and mis-configurations, the testing team might discover additional information that can feedback the discovery phase. This interaction between the discovery and exploitation phases is continuous throughout the actual test.

- Reporting phase- The report writing can begin in parallel to the other three stages, although must finish after exploitation phase has been completed. A successful report details all the findings and their impacts to the organization by taking into account both the technical and management aspects in its format. It is very important to conduct a fully detailed and well documented report in order to inform the management about the security risks and provide technical details.

In earlier system instead of blocking the particular client for misbehavior the whole network associated with it is blocked. In our system we are blocking the particular client instead of blocking the whole network attached to it.

II. SYSTEM ARCHITECTURE

2.1 The Pseudonym Manager

The user must first contact the Pseudonym Manager (PM) and demonstrate control over a resource; for IP-address blocking, the user must connect to the PM directly. We assume the PM
has knowledge about Tor routers, for example, and can ensure that users are communicating with it directly. Pseudonyms are deterministically chosen based on the controlled resource, ensuring that the same pseudonym is always issued for the same resource. As we will explain, the user contacts the PM only once per linkability window (e.g., once a day).

2.2 The Nymble Manager
After obtaining a pseudonym from the PM, the user connects to the Nymble Manager (NM) through the anonymizing network, and requests nymbles for access to a particular server (such as Wikipedia). A user requests to the NM are therefore pseudonymous, and nymbles are generated using the users pseudonym and the servers identity. These nymbles are thus specific to a particular user-server pair. To provide the requisite cryptographic protection and security properties, the NM encapsulates nymbles within nymble tickets.

2.3 Blacklisting a user
If a user misbehaves, the server may link any future connection from this user within the current linkability window (e.g., the same day). A user connects and misbehaves at a server during time period within linkability window. The server later detects this misbehavior and complains to the NM in time period. Therefore, once the server has complained about a user, that user is blacklisted for the rest of the day. Even though misbehaving users can be blocked from making connections in the future, the users past connections remain unlinkable, thus providing backward unlinkability and subjective blacklisting.

2.4 Notifying the user of blacklist status
Users who make use of anonymizing networks expect their connections to be anonymous. If a server obtains a seed for that user, however, it can link that users subsequent connections. The user is notified of their blacklist status before they present a nymble ticket to a server. In our system, the user can download the servers blacklist and verify her status. If blacklisted, the user disconnects immediately.

1. LIMITATIONS
Need of refreshing
We need to refresh the database of abusive words in timely manner as the new words are entering system.
III. ACKNOWLEDGEMENTS

First and foremost, I would like to thank my guide, Prof. Miss Ketaki Naik, for her guidance and support. I will forever remain grateful for the constant support and guidance extended by guide, in making this report. Through our many discussions, she helped me to form and solidify ideas. The invaluable discussions I had with her, the penetrating questions she has put to me and the constant motivation, has all led to the development of this project.

I wish to express my sincere thanks to the Head of department, Prof. D.A.Godse also grateful thanks to Prof. Miss Ketaki Naik, Prof. S.M.Murge and the departmental staff members for their support.

I would also like to thank to my friends for listening to my ideas, asking questions and providing feedback and suggestions for improving my ideas.

REFERENCE


Abstract—In the past few years, a number of ideas have been proposed for indoor navigation systems. These ideas were not as widely implemented as outdoor positioning systems like GPS(Global Positioning Systems). We propose an indoor navigation assistance system using Bluetooth which is low cost and feasible to use in daily life. Our system enables users with handheld mobile devices to steer with ease through the indoor premises using the short range radio frequencies of Bluetooth. It also establishes user’s current location and the various paths leading to the destination. Dijkstra’s algorithm is used to determine the shortest path from the source to the required destination.

Keywords—Bluetooth, Indoor Navigation, Mobile module.

I. INTRODUCTION

The expedience and comfort provided by the existing outdoor navigation systems has facilitated the development of indoor navigation and location tracking. GPS positioning provided by mobile telephone operators are suitable for outside environments where clear line-of-sight with respect to the satellites or base stations is available. However, they suffer from multi-path effects within buildings, and therefore, in indoor they show poor performances [1].

In recent times there has been a growth in huge infrastructures like shopping malls, industrial complexes along with existing structures like hospitals and college campuses. Moving around in such premises can be difficult and finding the path to the desired location can be time consuming and tiresome. The ideal indoor navigation system should provide easy guidance to navigate through such areas.

We decided to use Bluetooth technology to estimate the location because of its main hardware features: low power consumption, low cost and low interference with devices that work on the same frequency range and its widespread use in typical mobile devices [2]. Desktop software built on Java platform would use Dijkstra’s algorithm to determine the shortest path to the desired destination.

II. RELATED WORK

- **3D indoor location and navigation system based on Bluetooth:** This paper presents the design and implementation on mobile device, of a 3D positioning and navigation system for indoor, based on the use of Bluetooth (BT) radio technology and implemented using Java and J2ME. This implementation is adaptable to many indoor environment (commercial centers, offices, museums, etc.) previously modeled and loaded. J2ME and Bluetooth technology are the main features used [2].

- **Low cost Bluetooth mobile positioning for location based application:** This system has 2 main components namely the Bluetooth sensor system and Central Navigation System. The Bluetooth Sensor System allows mobile devices whose Bluetooth mode is set to discoverable, to be scanned and detected, and they receive customizable text message and other relevant files (maps, sound files, video clips) of their positioning information, e.g. room identity. The positioning information is also sent to the Central Navigation System which in turn displays and updates the navigation map. The system is also used to track the movement of different BT mobile devices within the implemented environment [3].

- **Research of Indoor Local Positioning Based on Bluetooth Technology:** Bluetooth is a technology with low-power, short-distance, wireless systems, can be used to construction a wireless services of indoor. The development of local positioning services based on the Bluetooth technology is discussed in this paper. An approach based on the positioning information exchanged between the master and slave devices within the Bluetooth network [4].

- **Indoor Navigation Performance Analysis EPFL:** The computation of the best route is based on a Dijkstra’s algorithm which is specifically designed for a continuous and oriented graph. The algorithm has to estimate the shortest path between a start node and an end node given by a user who has a specific profile. The best route computed by the system is used as input for the navigation process and for providing guidance information to the user [5].
• Navizon I.T.S: This system locates Wi-Fi enabled devices. No application is required on the tracking device but the Wi-Fi radios should be on and within the coverage area of the network of nodes. The I.T.S(Indoor Triangulation system) site periodically reports the details about the Wi-Fi enabled devices. This system basically uses Wi-Fi to track devices using Navizon’s proprietary algorithm with the help information collected by the nodes [6].

III. PROPOSED SYSTEM

Our proposed system works in 2 main modules: desktop module and mobile module.

a. Desktop module: The desktop module that will be used by administrator ideally be doing the following:
   • Set map information
   • Save/Load map information to/from files
   • Set Bluetooth device information
   • Set paths
   • Find optimum paths
   • Auto compile mobile application.
   • Upload mobile application to dedicated website for public download access.

It will be developed on JDK platform.

b. Mobile module: The mobile application that can be downloaded from internet by any user on his phone will show the user overall map of the premise. Apart from this the application will allow the user to
   • See his current location
   • Select source point
   • Select destination point
   • Show all paths from source to destination
   • Show shortest path from source to destination
   • Update current location at regular intervals
   • Graphically show the path overlay on premise map
   • Navigate through premise map for additional information.

A set of Bluetooth devices placed at the various locations in the premises assist in determining the current location of the user.

IV. WORKING

1) Firstly all the application work i.e. the creation of the maps, managing all the paths, scaling of the maps, adding the Bluetooth devices, and managing the database is done by the administrator on the desktop module.

2) This whole J2ME application is compiled and a jar file is created and ready to transfer to the user.

3) Whenever a user enters a premise with this system it is necessary that his mobile’s Bluetooth is enabled. The server i.e. the desktop module will prompt the user a message through Bluetooth whether the latter wants the application or not, if yes the compiled J2ME application is sent to the user through Bluetooth.

4) Now the user has the jar file of the application. He has to run this application. The application helps the user to easily navigate through the premises. It shows the user his current position. The user can also find all possible paths from any source place to destination and out of all this he can also find the shortest path to the destination.
The Bluetooth devices help the application to find the users current position and also the position is continuously updated as the user moves.

V. CONCLUSION

There is a growing need for assistance while traversing complex establishments. Our proposed system presents an uncomplicated and cost effective approach for its users to navigate through indoor infrastructures. We also determine the shortest path using Dijkstra’s algorithm from source to destination and use Bluetooth for navigating and determining the user’s current location.

REFERENCES


