

**International Journal of
Computer Science and Security
(IJCSS)**

ISSN : 1985-1553



VOLUME 4, ISSUE 5

PUBLICATION FREQUENCY: 6 ISSUES PER YEAR

**International Journal of
Computer Science and Security
(IJCSS)**

Volume 4, Issue 5, 2010

Edited By
Computer Science Journals
www.cscjournals.org

Editor in Chief Dr. Haralambos Mouratidis

International Journal of Computer Science and Security (IJCSS)

Book: 2010 Volume 4, Issue 5

Publishing Date: 20-12-2010

Proceedings

ISSN (Online): 1985-1553

This work is subjected to copyright. All rights are reserved whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication of parts thereof is permitted only under the provision of the copyright law 1965, in its current version, and permission of use must always be obtained from CSC Publishers. Violations are liable to prosecution under the copyright law.

IJCSS Journal is a part of CSC Publishers

<http://www.cscjournals.org>

© IJCSS Journal

Published in Malaysia

Typesetting: Camera-ready by author, data conversion by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers

Editorial Preface

This is fifth issue of volume four of the International Journal of Computer Science and Security (IJCSS). IJCSS is an International refereed journal for publication of current research in computer science and computer security technologies. IJCSS publishes research papers dealing primarily with the technological aspects of computer science in general and computer security in particular. Publications of IJCSS are beneficial for researchers, academics, scholars, advanced students, practitioners, and those seeking an update on current experience, state of the art research theories and future prospects in relation to computer science in general but specific to computer security studies. Some important topics cover by IJCSS are databases, electronic commerce, multimedia, bioinformatics, signal processing, image processing, access control, computer security, cryptography, communications and data security, etc.

This journal publishes new dissertations and state of the art research to target its readership that not only includes researchers, industrialists and scientist but also advanced students and practitioners. The aim of IJCSS is to publish research which is not only technically proficient, but contains innovation or information for our international readers. In order to position IJCSS as one of the top International journal in computer science and security, a group of highly valuable and senior International scholars are serving its Editorial Board who ensures that each issue must publish qualitative research articles from International research communities relevant to Computer science and security fields.

IJCSS editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJCSS. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review

process. IJCSS provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members

International Journal of Computer Science & Security (IJCSS)

Editorial Board

Editor-in-Chief (EiC)

Dr. Haralambos Mouratidis
University of East London (United Kingdom)

Associate Editors (AEiCs)

Dr. Padmaraj M. V. nair
Fujitsu's Network Communication division in Richardson, Texas (United States of America)

Dr. Blessing Foluso Adeoye
University of Lagos (Nigeria)

Associate Professor. Azween Bin Abdullah
Universiti Teknologi Petronas (Malaysia)

Editorial Board Members (EBMs)

Dr. Alfonso Rodriguez
University of Bio-Bio (Chile)

Dr. Srinivasan Alavandhar
Glasgow Caledonian University (United Kingdom)

Dr. Debotosh Bhattacharjee
Jadavpur University (India)

Professor. Abdel-Badeeh M. Salem
Ain Shams University (Egyptian)

Dr. Teng li Lynn
University of Hong Kong (Hong Kong)

Dr. Chiranjeev Kumar
Indian School of Mines University (India)

Professor. Sellappan Palaniappan
Malaysia University of Science and Technology (Malaysia)

Dr. Ghossoon M. Waleed
University Malaysia Perlis (Malaysia)

Dr. Srinivasan Alavandhar
Caledonian University (Oman)

Dr. Deepak Laxmi Narasimha
University of Malaya (Malaysia)

Professor. Arun Sharma
Amity University (India)

Professor Mostafa Abd-El-Barr
Kuwait University (Kuwait)

Table of Content

Volume 4, Issue 5, December 2010

Pages

- 436 - 450 A Proposed Security Model for Web Enabled Business Process Management System
M. S. Kandil, Mohamed Abu El-Soud , A. E. Hassan , Abd Elghafar M. Elhady
- 451 - 465 An Approach for Managing Knowledge in Digital Forensics Examinations
April Tanner, David Dampie
- 466 - 481 Securing Image Transmission Using in- Compression Encryption Technique
Shaimaa Ahmed Elsaid, Khalid F. A. Hussein, Mohamed M. Fouad
- 482 - 490 Building a Distributed Secure System on Multi-Agent Platform Depending on the Principles of Software Engineering Life Cycle
Dr.Ghossoon Mohammed Waleed Al-Saadoon
- 491 - 496 Performance Variation of LMS And Its Different Variants
Sanjeev Kumar Dhull, DR .O. P SAHU

A Proposed Security Model for Web Enabled Business Process Management System

M. S. Kandil

*Faculty of Engineering
Mansoura University*

arwaahmed1@gmail.com

Mohamed Abu El-Soud

*Faculty of Computer & Information Sciences
Mansoura University*

a_m_elhady@yahoo.com

A. E. Hassan

*Faculty of Engineering
Mansoura University*

arwaahmed1@gmail.com

Abd elghafar M. Elhady

*Faculty of Computer & Information Sciences
Mansoura University*

a_m_elhady@yahoo.com

Abstract

Business Process Management systems (BPMS) and technologies are currently used in many organizations' IT applications. This could lead to a dramatic operational efficiency improvement on their business and administrative environments. With these atmospheres, the security issue is becoming a much more important challenge in the BPMS literature. The Role-Based Access Control (RBAC) model has been accepted as a promise security model solution and standard. RBAC is able to accomplish the central administration of an organizational specific security policy. It is also able to meet the secure processing needs of many commercial and civilian government organizations. In spite of these facts, RBAC model is not reliable when applying to the BPMS without further modifications and extensions. RBAC is modified to fit with Service oriented (SRBAC), but still not reliable enough to handle BPMS.

Authors of that research proposed a security model based on SRBAC model to be more reliable when using with BPMS. Authors of that research named that proposed security model as Improved Role Based Access Control (IRBAC). The IRBAC model is directly applicable to the BPMS.

Authors defined a graphical representation and technical implementation of the IRBAC model.

This IRBAC model is tested using simple case study. The test compares between the IRBAC model and SRBAC model where IRBAC is implemented in two cases (IRBAC with caching and IRBAC with no caching). The test results show the validity and performability of the IRBAC model.

1. INTRODUCTION

Since the beginning of the shift from a functional to a process-centered view of business activities in the 80s [2], business processes play a major role in today's companies. BPMS is applied to "analyze and continually improve fundamental activities such as manufacturing, marketing, communications and other major elements of a company's operations" [3]. In other words, it is applied to engineer lean and streamlined business processes [2]. The introduction

of BPMS has several benefits such as cost reduction, quality improvements and error reduction, visibility gain, and process step automation [4]. In recent years, business processes are often the target of security hazards, such as viruses, hacker attacks, or data theft [5,6].

Because business processes generate valuable information and knowledge as output, decision makers and security experts need to improve methods to secure them against external or internal attacks. These attacks could result in demand and loss of value for system and organization. These damages can be monetary loss (e.g., loss of profit due to the interruption of business activities) and/or intangible value loss (e.g., loss of reputation).

The Data stores detailed information of a organization, and Business Processes that are Performed in the organization's System should be protected. When a user connect to the system, the environment (Data/Business Processes) Created For the user should be ensured in. In order to solve the above issues, adaptive access control is necessary to make sure of the information security of Business Process.

RBAC has become a widely accepted mechanism for security management [7]. RBAC uses the assignment between users, roles and permissions to provide a more convenient access control management model. However, the traditional RBAC does not consider the user's current environment. It merely bases on the predefined role and permission plan. Some research has combined RBAC with BPMS to achieve dynamic authorization [8,9,10,11]. Nevertheless, most of research with BPMS adopts a Model to use RBAC Methodology with BPMS. These Models have some shortages. Examples of these shortages are that some of these models didn't present the most optimum solution of applying RBAC with BPMS. Also, they didn't present a complete implantation of this combination.

Traditional security systems with BPMS didn't secure the system. Dey et al in [9] stated that in February 2000, a Denial of Service (DoS) attack caused access problems of Yahoo's website, costing an estimated half a million US Dollars in just three hours. The consequence is an ever increasing amount of money on improving security (from 1999 to 2000, the number of organizations spending more than \$ 1 million annually on security nearly doubled, representing 12% of all organizations in 1999 to 23% in 2000 [12]). The main problem with security - in this context information security is the lacking integration of security considerations into business processes [13].

Therefore, appropriate access control will improve the feasibility of using BPMS technology in Organizations.

Authors of that research proposed a hybrid model which modified SRBAC model to achieve a dynamic authorization security model (IRBAC).

IRBAC model is proposed in two cases. First case when IRBAC is combined with caching. And the second case when IRBAC is proposed with no caching. The proposed model is tested in the two cases and results are compared with results of SRBAC model.

This proposed model is a generic security model. This model could be added to any BPMS and handle the authorization of system's users.

2. RELATED WORK

Access control and authorization concerns are one of the key challenges preventing BPM gaining widespread recognition. Firstly, it is not realizable to apply role based model to business process systems directly. Moreover, the inter-organization business process scenario becomes more complicated. For instance, the inherited roles might be stored remotely and permissions constraints will consequently require several remote invocations [14].

Although the concept of role has existed for a long time in systems security, the work presented by Sandhu et al in [15] has prompted a renewed interest in this approach. But proposed model that greatly simplifies security management is presented in [16]. RBAC

model is now adopted in many commercial products to different degrees since access control is an important requirement of information systems. RBAC was found to be the most attractive solution for providing security characteristics in inter-organizational business systems [17]. Moreover, it would be much easier for organizations to enhance security protection from existing RBAC based policies.

David F.Ferraiolo et al in [18] and Ravi S.Sandhu et al in [15] define **Traditional RBAC Model** as a model composed of three components:

- A user is a human being belongs to an organization.
- A role is a named job function within the business process context that regards the authority and responsibility.
- A permission is an approval of actions granted to specific roles. A constraint regulates the relations between different elements.

In this model, the central notion is that permissions are associated with roles, and users are assigned to appropriate roles. This greatly simplifies management of permissions. It is suitable for simple Web applications. But in more advanced web applications such as BPMS and Service Oriented Architecture (SOA) applications, traditional RBAC is not suitable for them. Moreover traditional RBAC can not completely express dynamic characters of role according to what is mentioned in [1].

Xin Wang et al in [19] added a service element to original RBAC model and proposed a new model called Extended RBAC Model, which indicates the Web service deployed within the enterprise system and divided roles into human role and computer role. The human role indicates the tasks to be performed by human users, while the computer role indicates the tasks to be performed by Web services. This model extension addresses the SOA upgrade in this kind of progressive manner.

In [19], authors rely on role hierarchy which causes shortages in system performance. To access a specific service, role server could be accessed more than one time to get role which contain permissions for that user on this service, which causes more network traffic and less overall system performance. Authors divided the system operations into two types, one is performed by users and other is performed by web services. Also, Authors ignore the relation between web services and users of the system, in other words authors didn't define how user can fire web services that perform specific functions in the system.

Another system proposed in [1] is called a Service-oriented Role Based Access Control (SRBAC) model in which, traditional protected objects are replaced by services, and a new notion of actor is introduced. An Actor is a dynamic object which is created when a user activates a role. Its condition and action may present the characters of the role activated.

In this model, Roles are organized in role Hierarchy. This causes system performance decreasing by causing more network traffic and less overall system performance as mentioned in previous model. Moreover, authors were rely on creating actor for user each time he accesses new role that contains the services he needs. This makes user has to switches among actors to manages services that spreading across more than one role. This scenario was designed to reflect the dynamic execution process of the role. They proposed that roles are dynamic continuously but in most systems, this state can exists at beginning of system building and deployment and rarely happened after that, along system life.

Authors of that research used SRBAC model after modifying it and proposed a new security model called IRBAC. The IRBAC model is the modified SRBAC that has two cases, first combines it with caching technique and second case uses no caching.

3. PROPOSED IRBAC MODEL

In this section, the IRBAC model will be presented. Several IT technologies are combined to provide a dynamic, fast, and secured mechanism for accessing system processes in the model. The implementation of the IRBAC model is presented. The IRBAC model is considered as a Generic Security model which used BPMS principles and could be applicable on any BPMS to manage the authentication and authorization of users on BPMS.

The IRBAC model rely on using the RBAC model in BPMS to improve the security of the system and provide a dynamic management environment for roles /permissions / users assignment which enable system user to adapt role and permission according to any changes happened in the system authorization.

The IRBAC model has two cases:

- First case uses caching strategy to decrease the overall response time experienced by the user when he/she is interacting with the system thus increase system Performance. Where authors utilized from the tests have been made by Kohler et al in [20] on using caching strategy in Business Process-driven Environments which results that using caching in Business Process-driven Environments decrease the response time of user requests significantly thus improve increase the overall system performance.
- The second case depends on that there are some systems has many changes happened to roles' permissions during the operating of users on the system. In this case caching technique is not suitable with the system needs whoever it is better in performance. So IRBAC model uses no caching to meet the operational needs of these systems.

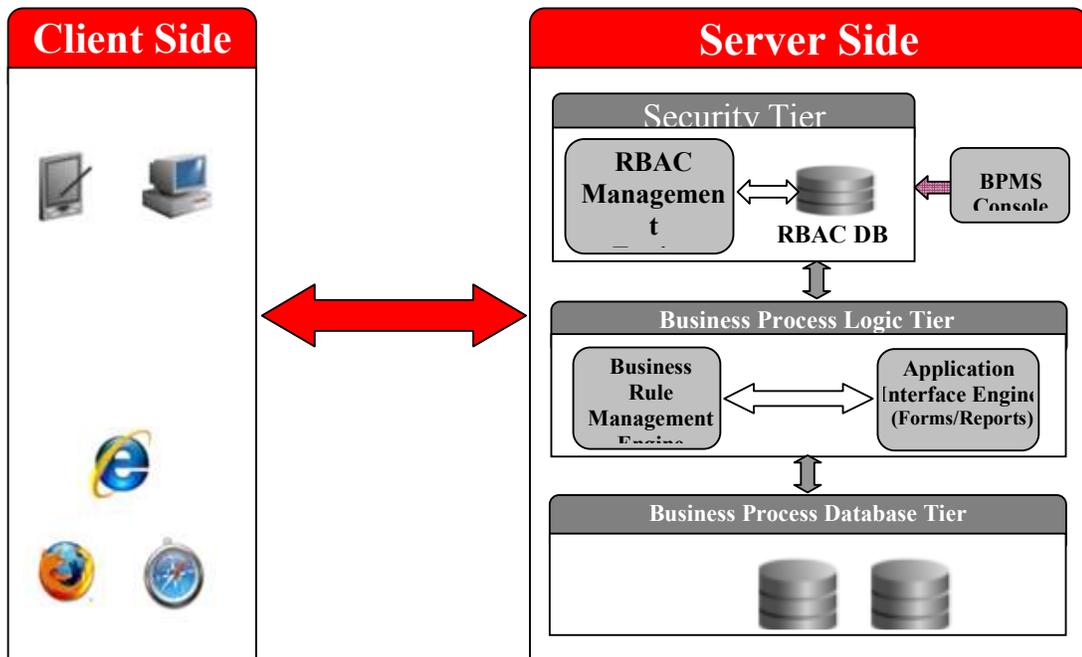


FIGURE 1: Proposed Model Architecture Diagram (Client / Server) N-Tier

The client / server architecture of proposed model is presented in figure (1), in which **client Side** can be computer with browser from which user can access to the BPMS .the Server Side Consists of three main components.

- **Security Tier:** which responsible on verifying the authentication and authorization of users which are dealing with the BPMS. It also detects if any changes happened to the system's processes and perform appropriate action to adapt the security tier of the BPMS.

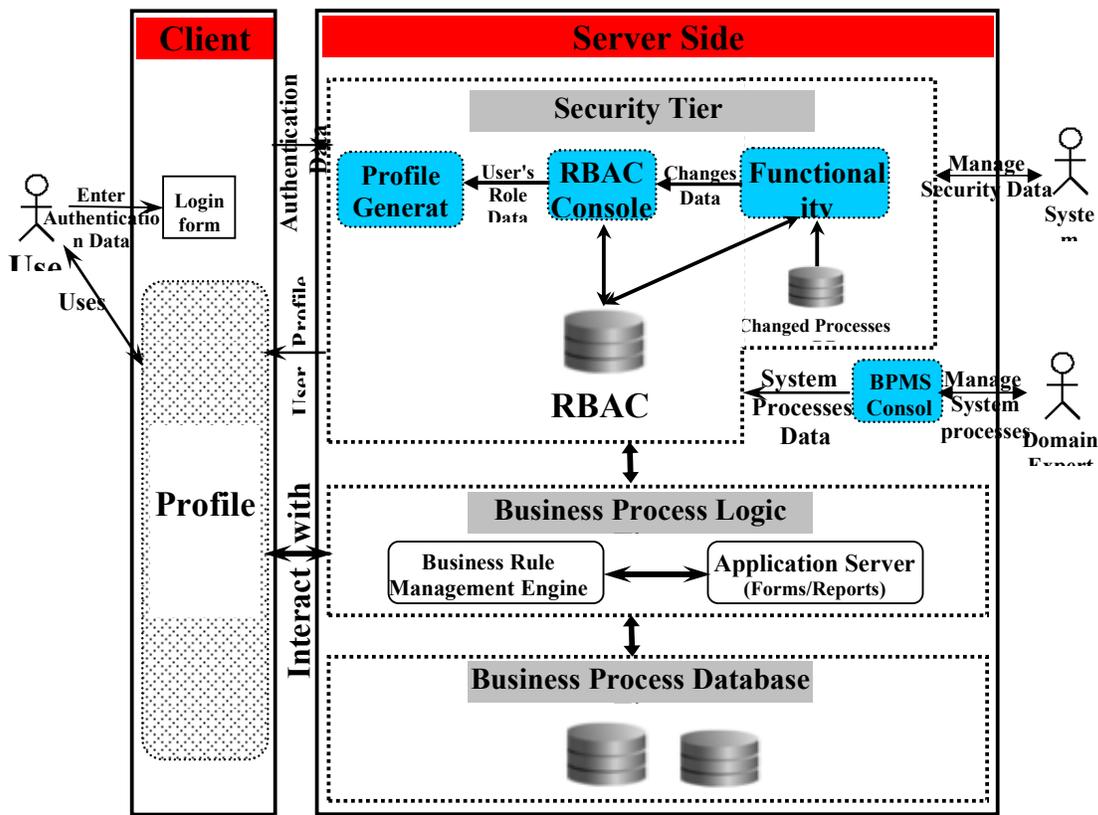
- **Business Process Logic Tier:** This maintains logic of the business system and rules which organize it in Business Rule Management Engine and Application forms and reports in Application Interface Engine.
- **Business Process Database Tier:** In which all Organization's data and information used in the system are stored in Databases.

In addition to these three tier authors design additional component called BPMS Console. By this component any system processes can be configured and/or reconfigure and the output file is delivered to the actual BPS to activate the changes.

3.1 Schematic Diagram of the Proposed Model (N-tier)

In previous section authors present a general view of the proposed model and its components in brief as client/server architecture model. Here, Authors represent the proposed model in more details as schematic diagram. as shown in figure (2), this model consists of four tiers; Client Tier, Security Tier, Business Process Logic Tier, and Business process Database Tier in addition to BPMS Console component. Authors satisfy with what they presented about Process Logic Tier, and Business process Database Tier and will focus in this section on other two Tiers which compromise the core of their work in this research.

FIGURE 2: Schematic Diagram of the Proposed Model



- **Client Tier:** through which any User of the system can access the BPMS According to his authorization where user enter his authenticated data which it send to security tier and accept his profile on his client machine . With this profile, user can deal with the system processes without any need to access security tier to get his authorization data on called processes in case of using caching technique. But in the other case, with no caching, user profile has to connect to security tier to get the last permissions of user on the calling process.
- **Security Tier:** which is responsible for applying RBAC model to BPMS and it consists of three main components :

operation stage. In this type, caching the user's roles in a complete profile is better in performance.

Authors of this research change the SRBAC model As in figure (4) by replacing services with processes, removing role hierarchy and assign any user to only one role which maintaining permissions on all system processes. The role can be assigned to many users. When user login to the system his role is captured and user profile is created on his client machine using caching technique. However in no caching case, when user login to the system his processes list available in his role is captured and user profile is created on his client machine and the process's permissions are checked when user is calling that process.

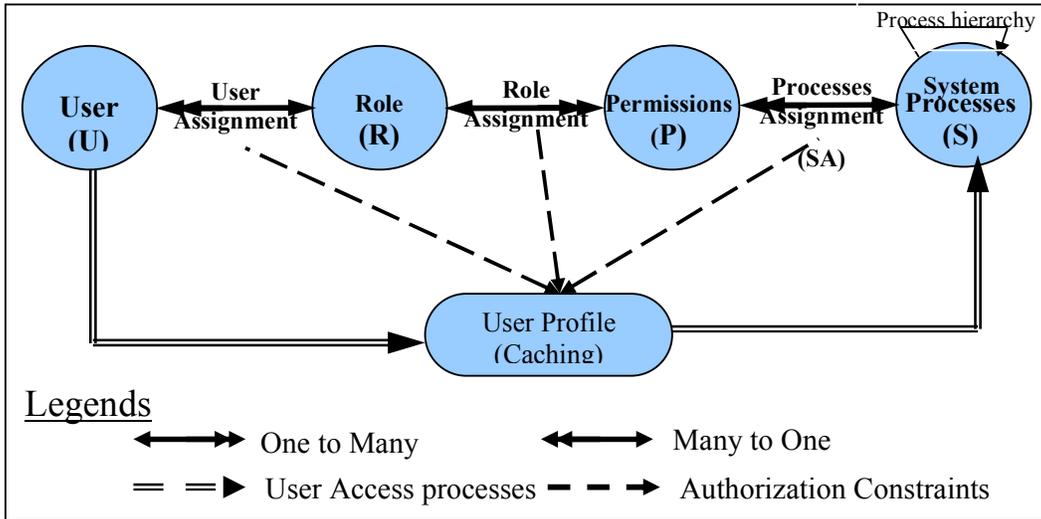


FIGURE (4) Proposed security Model

Moreover, it is better to collate all operations of user in one role and display it to his than split them into more than one role like it is in SRBAC model.

3.2 Conceptual model of IRBAC model's Security Tier Functionality.

In this section, Authors will explain the functionality of Security Tier, its objects, the relations between them, and the interaction between its components with each others and between them from one side, users and other model components on another side.

Authors of this research modify the model proposed in [1] by adding set of objects and components to meet their vision of new security model. According to figure (5), the proposed security model will contain six main objects:

- **System Processes (S):** This represents a list of system's Processes or Services which was created using BPMS Console component by Domain expert.
- **Changed Processes (CP):** This represents a list of changes happened to system's processes. This object is used by Functionality Adapter Component to update system processes(S) object with the last changes of system processes data.
- **Role (R):** This represents all the permissions of a specific type of system users on whole system's processes.
- **Permission (P):** This represents the access rights of one of system's processes for a specific Role (R).
- **User (U) :** which represents the users of BPS.
- **User Profile (UP):** This is generated by Profile Generator component. It contains all the user's permissions on the whole BPMS processes in caching case and contains list of BPMS processes available to user in no caching case.

First list of system processes or services (S) is created. Then RBAC Console is used by system administrator to perform the following steps in sequence:

1. Creates new Role and then creates permissions for all system processes and assign them to this role in **Permission Assignment (PA)** step,
2. For each permission of that role's permissions, specifies access rights of one system process in **System Processes Assignment (SA)** step.
3. Creates new users and assigns them to their appropriate Role in **User Assignment (UA)** step, where one or more user can share the same role but a user can't assign to more than one role in the system.
4. When the user login to BPMS:
 - **In caching case**, RBAC Console check user authentication. Then it captures all user authorization data which contains all constraints of the login user on that BPMS. The authorization data sent from RBAC Console to Profile Generator which uses it in creating complete User Profile. User Profile is sent to user client machine. User uses that profile which it creates on his client machine to interacts directly with the BPMS without needing to connect to the security server to capture his privileges on any service as long as his session is alive.
 - **In no caching case**, Console check user authentication. Then it captures all BPMS processes available to user and sent to Profile Generator which uses it in creating summarized User Profile and sent to user client machine. When user calls one of BPMS processes, user profile asks RBAC console to get the last permissions of user on that processes and then call that process under these permissions.

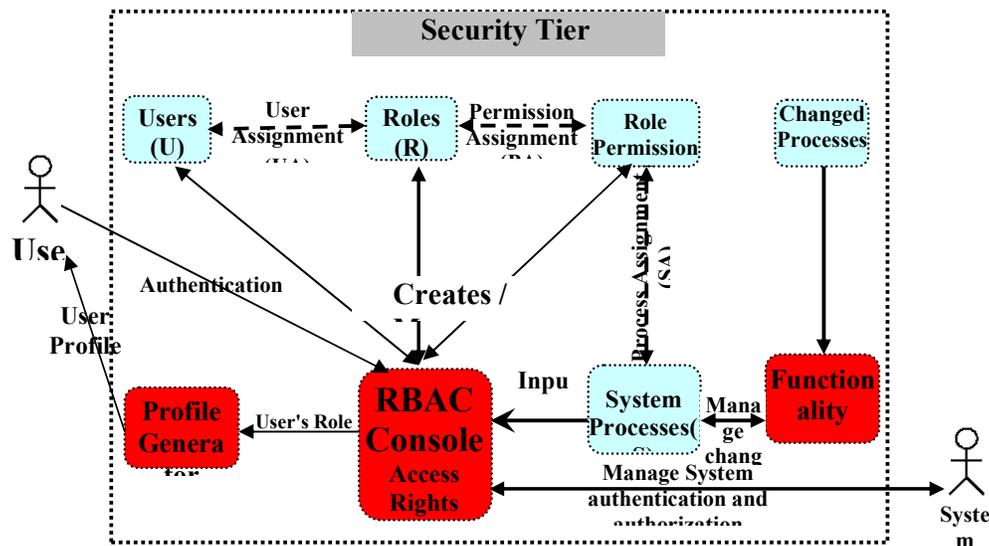


FIGURE 5: Conceptual Model of Proposed Security Model

But what if any changes happened in BPMS processes after BPMS had been deployed. How can these changes deployed to the running system? Authors use plug and play mechanism to do that. Where domain expert uses BPMS Console component to specify the changes happened to system processes. BPMS Console creates change processes object that maintain these changes. Then change processes object is plug into the BPMS. When system administrator login to the system, the Functionality Adapter component in the security tier of BPMS check the change processes object and executes all changes to the system.

3.3 Proposed Security Model Analysis.

According to scenario of model from authors point of view, there are three types of users will deal with the proposed model. These users are:

- **System Domain Expert:** He is a person who specifies and manages BPMS processes.
- **System Administrator:** is a person who creates/ manage users accounts, system's roles and permissions, and assign users to roles.
- **System user:** is a person who uses or operate BPMS.

In the following section authors will demonstrate the proposed model analysis by explain the model flow chart and the model use case.

3.3.1 Proposed Security Model Flow Chart

Figure (6) shows the proposed model flow chart. The user logs in to the System by entering his username/password. These data is checked by RBAC Management Engine. If authentication data is correct, Profile Generator component generates the profile of the user according to his type. If user is Domain Expert, the Profile Generator creates BPMS Domain Expert profile which contains BPMS Console of BPMS. If user is system Administrator, the Profile Generator creates BPMS System Administrator profile which contains RBAC Management Engine of BPMS. If user is Regular System User, the Profile Generator creates BPMS User profile which contains BPMS's processes available to the login user and his permissions on these processes according to Role he belongs to in caching case. However the profile generator creates only a list of all system processes available to user and displays it to him in summarized profile on his machine.

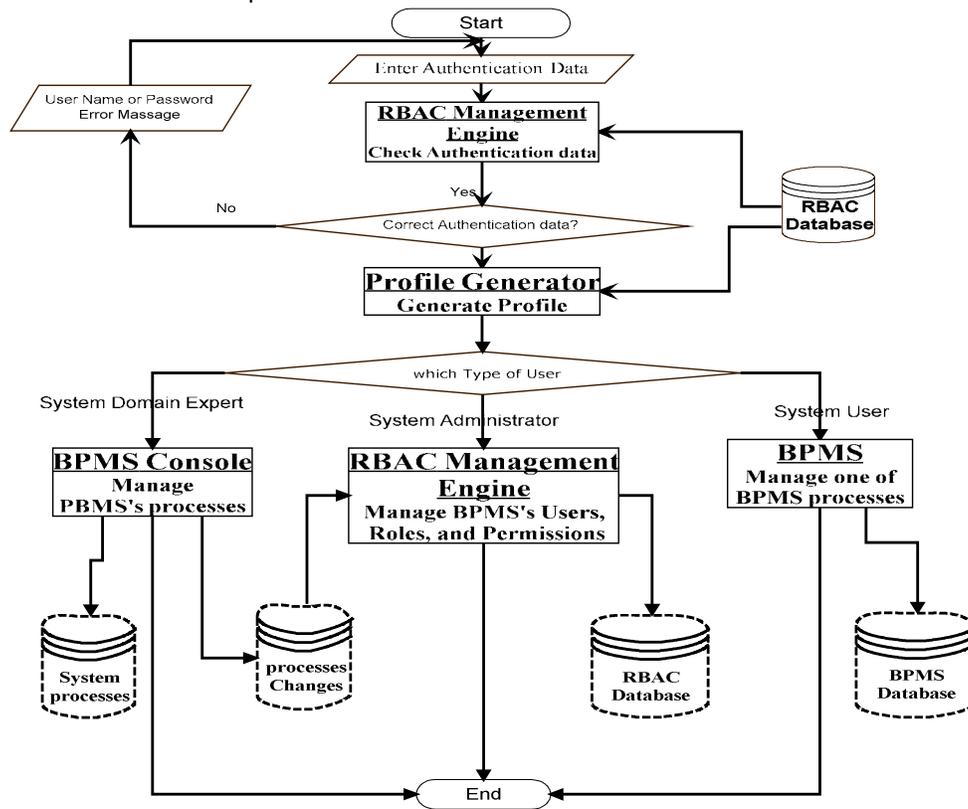


FIGURE 6: System Flow Chart of Proposed Security Model

3.3.2 Proposed Security Model Case Study

According to the authors' vision of the IRBAC model, use case consists of seven actors (System domain expert, System Administrator, System User, system processes DB, BPMS Database, and RBAC Database, and Business Process System) and nine use cases (Manage system processes, Manage role and specify access permissions, Manage User and assign them to appropriate Role, Authenticate to System, Create profile, ask for system process, check user permission, and call process under specific permissions) as shown in figure(7).

System Domain Expert uses "Manage all system's processes to create all system's processes at first or to modify these processes or add new processes next. This operation's data is store in system processes DB, which delivers to actual system to deploy the changes in system processes by updating system's processes which stored in BPMS Database of the

and create summarized user profile and deliver it to him. When user needs to perform one of system processes, he selects the process from his profile displayed to him. User profile check user permission on the selected process by getting it from RBAC DB and call this process under user's permissions on that process.

4. Proposed model Simulation and Validation

In this validation, authors compare between the performance in the two cases (with caching and no caching) and SRBAC model performance that was mentioned previously as a similar model to the IRBAC model and it was described well. This performance comparison has made on a small case study that simulate the IRBAC model in the two cases and SRBAC model.

Authors will propose the following three scenarios for two models:

- For SRBAC model, after user authentication to one of his roles, an actor is created. The services of that role that he can access will appear to him in this actor with role's permissions. When he want to access any services from list of services appear to him, the actor will check the access rights of on that service and deal with that service. When user wants to access another role he belongs to, he must authenticate again, but this time to the other role which will create a new actor for that role. Through the new actor, user can deals with the system with another manner.
- For the proposed model, there are two scenarios:
 - With caching technique, after user authentication, the Profile Generator will check the all processes available to user and his permissions on these processes and caching all of them together in complete profile generated to user on his client terminal. When user wants to access specific process, his profile which was cached on his client terminal get access rights of that service from list of access rights stored within the profile which is generated at login without need to connect to the security server to access the process permissions, then connect to application server to get the process under his permissions.
 - With no caching technique, after user authentication, the Profile Generator will check the all processes available to user display a list of all of them in profile generated to user on his client terminal. When user wants to access specific process, his profile checks calling process's permissions for that user from security server. Then connect to application server to get the process under his permissions.

From scenarios that has been stated, there are two stages will be take in consideration in validation process .first stage is at login stage, and the second stage is at process calling.

Our test contains 10 users that are connecting to BPMS which consists of 40 processes. Each one of user can access only 30 processes with different permissions. Then 10 times of process call has been performed and measure the response time for each process call in the SRBAC model and the proposed model with and without caching and drew statically graphs which illustrate the results. In SRBAC model, Authors proposed that login user has three roles and the 10 processes he needs to access spread across these roles. Then to make 10 process calls across three roles, he needs to login to each role separately and make process call to required processes in this role.

Figure (8) shows the response time of login stage in the SRBAC model and the proposed model with and without caching. Figure (9) shows the response time of process calling in the SRBAC model and the proposed model with and without caching.

The results show that the proposed model without caching is better than the SRBAC model and proposed model with caching in login stage where the average response time for the proposed model no caching is $(80.29 \times 10^{-11} \text{ s})$ but it is $(112.76 \times 10^{-11} \text{ s})$ in the proposed model with caching and average of three times login for 10 user of the SRBAC model is $(447.65 \times 10^{-11} \text{ s})$. Whereas the proposed model with caching and SRBAC model is better than the proposed model without caching after login stage, along session life between user and BPMS. The average of response time of the proposed model with caching is $(2.76 \times 10^{-11} \text{ s})$ and it is $(2.1 \times 10^{-11} \text{ s})$ in SRBAC model, but it is $(44.51 \times 10^{-11} \text{ s})$ in the proposed model without caching for each process calling.

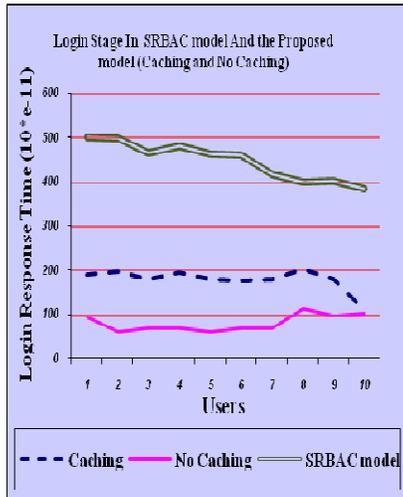


FIGURE 8: Login stage Response time

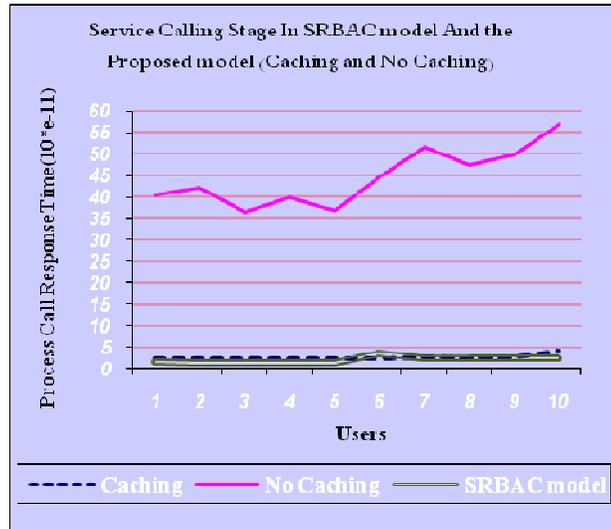


FIGURE 9: Process calling Response time

Authors combines the two stages (login & process access) in complete scenario for the proposed model with (caching & no caching) and SRBAC model and make 10 users perform all scenarios. The results can be seen in figure (10). The average of total response time using proposed model with caching is (163.07*10e-11s) whereas it is (1144.7*10e-11s) when using the proposed model without caching and it is (454.59*10e-11s) when using the SRBAC model.

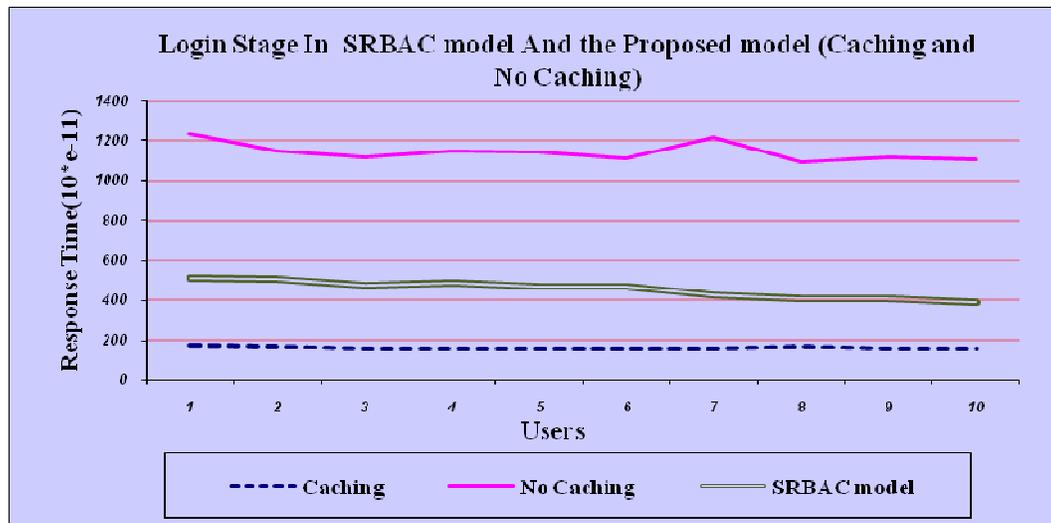


FIGURE 10: Total Response time of proposed model and SRBAC model

From results that have been reached, Authors conclude that when system's roles are rarely changing, the proposed model with caching is the best solution for managing user's authorization. But, when system's roles are continuously dynamic, the proposed model without caching is better solution whoever is the lowest in performance but it grants that roles' permissions are up-to-date when user calls BPMS processes.

5. CONCLUSION

In this paper, authors proposed a generic security model (IRBAC) which modified SRBAC model to achieve a dynamic authorization security model when applying on any BPMS. The IRBAC model is more reliable when directly applied on the BPMS.

IRBAC model is compared with SRBAC in two cases. First case when IRBAC is combined with caching. And the second case when IRBAC is proposed with no caching.

Authors of that research presented a client/server N-tier architecture diagram of the proposed model. The client side represents the computer with browser from which system user interacts with the BPMS. The Server side consists of three tiers. First tier represents security tier and is responsible on manage the authentication and authorization of the BPMS. Second tier is business process logic tier which maintains all business logic of the BPMS and consists of business rule management engine and Application Interface Engine. Last tier is database tier, in which all BPMS data is maintained and managed.

Then authors presented a schematic diagram of the proposed model. It displayed the three types of users that deal with the security model, what component of the model user interacts with and interaction between all system components. The security tier, consists of three components. First component is RBAC Console which responsible on managing the authentication and authorization of all BPMS users on the system. Second component is profile generator component which captures all system processes available to login user and his permissions and creates his profile and send it to his machine (client side). When BPMS processes are changed, the functionality component is responsible on applying all changes on the actual system at system administrator login. In addition to these components, authors displayed DBPM Console component which enable BPMS domain expert from managing all system processes.

The modifications that made on the SRBAC model which modifications lead to a more reliable security model is presented.

Analysis of the proposed security model is done. That analysis is presented by presenting the proposed model system flow chart and use cases. This analysis presents how the three types of users (BPMS Domain expert, System Administrator, System User) interact with the system through the proposed security model.

Finally implementation of the proposed model using a simple case study is done. The case study is the Cultural Affairs System of Mansoura University. This model is implemented entirely in PHP language and MySQL. The performance of the proposed model in tested in two cases and compared the SRBAC model. This comparison had made in two stages. First stage is at login stage which appeared that the proposed model with no caching is better than SRBAC model figure(8). Second stage is at process calling stage which appeared that the proposed model with caching is better than SRBAC. But SRBAC is better than the proposed model with no caching figure(9).

Then authors combined the two stages in figure (10) which showed that the proposed model with caching technique is better solution for managing authorization of system's users.

In the future, authors of that research will apply the proposed model on another real BPMS case study in details.

6. REFERENCES

1. Xu Feng ,Lin Guoyuan , Huang Hao , Xie Li;"Role-based Access Control System for Web Services"; In Proceedings of the 4th IEEE International Conference on on Computer and Information Technology ,2004
2. Ateniese, G., Camenisch, J., and Madeiros, B. de, "Untraceable RFID tags via insubvertible encryption", Proceedings of the 12 ACM conference on Computer and communications security, November, pp.92-101, 2005.

3. Barkley, J., Beznosov, K., and Uppal, J., "Supporting Relationship in Access Control Using Role Based Access Control", Proceedings of ACM Role-Based Access Control Workshop, Fairfax, Virginia, USA, pp. 55-65, 1999.
4. Bernardi, P., Gandino, F., Lamberti, F., Montrucchio, B., Rebaudengo, M., and Sanchez, E.R., "An Anti-Counterfeit Mechanism for the Application Layer in Low-Cost RFID Devices", In International Conference on Circuits and Systems for Communications, IEEE, July, pp.207-211, 2006.
5. T. Neubauer, M. Klemen, and S. Biffi. Secure Business Process Management: A Roadmap. In Proceedings of the First International Conference on Availability, Reliability and Security ARES, pages 457–464. IEEE Computer Society, 2006.
6. T. Neubauer and J. Heurix : Objective Types for the Valuation of Secure Business Processes. In Proceedings of the Seventh IEEE/ACIS International Conference on Computer and Information Science, page 231. IEEE Computer Society, 2008.
7. M. Wu and Y. Fong : Applying Role-Based Access Control in Combining the Chinese and Western Medicine Systems. In Proceedings of the 19th International Conference on Systems Engineering . IEEE Computer Society, 2008.
8. Chen, G., and Kotz, D., "A Survey of Context-Aware Mobile Computing Research", Technical Report 2000-381, Dept. of Computer Science, Dartmouth College, Hanover, N.H, 2000.
9. Dey, A. K., and Abowd, G. D., "Towards A Better Understanding of Context and Context-awareness", GVU Technical Report GITGVU-99-22, pp.304-307, 1999.
10. Schilit, B. N., Adams, N., and Want, R., "Context-Aware Computing Applications", In Proceedings Workshop on Mobile Computing Systems and Applications, IEEE, pp.85-90, December, 1994.
11. Wolf, R., Keinz, T., and Schneider, M., "A Model for Context-dependent Access Control for Web-based Services with Role-based Approach", Proceedings of the 14th International Workshop on Database and Expert Systems Applications, September, pp.209-214, 2003.
12. Heiko, K., and Hartmut, P., "RFID Security", Information Security Technical Report, December, Volume 9, Issue 4, pp.39-50, 2004.
13. Li, Y.Z., Jeong, Y.S., Sun, N., and Lee, S.H., "Low-Cost Authentication Protocol of the RFID System Using Partial ID", In Computational Intelligence and Security, IEEE, pp.1221-1224, November, 2006.
14. M. Sloman and E. Lupu. Security and management policy specification. Network, IEEE, 16(2):10–19, 2002.
15. R. Sandhu, E. Coyne, H. Feinstein, and C. Youman. Rolebased access control models. IEEE Computer, 29(2):38–47, 1996.
16. T. Neubauer, M. Klemen, and S. Biffi. Secure Business Process Management: A Roadmap. In ARES'06, pages 457– 464, 2006

17. C. Yang. Designing secure e-commerce with role-based access control. *International Journal of Web Engineering and Technology*, 3(1):73–95, 2007.
18. David F. Ferraiolo, John F. Barkley, and D. Richard Kuhn. A role based access control model and reference implementation within a corporate intranet. In *ACM Transactions on Information Systems*
19. Xin Wang, Yanchun Zhang, Hao Shi ;" Access Control for Human Tasks in Service Oriented Architecture "; in *IEEE/ the Fourth International Conference on Computer and Information Technology (CIT'04)*;2004 *IEEE Computer*, 29(2):38–47, 1996.
20. Mathias Kohler and Andreas Schaad . ProActive Access Control for Business Process-driven Environments.in *IEEE/ Annual Computer Security Applications Conference* 156 .2008.

An Approach for Managing Knowledge in Digital Forensic Examinations

April L. Tanner

*Department of Computer Science and Engineering
Mississippi State University
Mississippi State, 39762, USA*

alb117@msstate.edu

David A. Dampier

*Associate Professor of Computer Science and Engineering
Mississippi State University
Mississippi State, 39762, USA*

dad6@msstate.edu

Abstract

Computers and digital devices are continuing to evolve in the areas of storage, processing power, memory, and features. Resultantly, digital forensic investigations are becoming more complex due to the increasing size of digital storage reaching gigabytes and terabytes. Due to this growth in disk storage, new approaches for managing the case details of a digital forensics investigation must be developed. In this paper, the importance of managing and reusing knowledge in digital forensic examinations is discussed, a modeling approach for managing knowledge is presented, and experimental results are presented that show how this modeling approach was used by law enforcement to manage the case details of a digital forensic examination.

Keywords: Digital Forensics, Concept Mapping, Case Domain Modeling, Digital Forensic Examinations

1. INTRODUCTION

Of the many issues associated with computer forensics, knowledge management strategies are also important to the future of not only computer forensics, but digital forensics as well. Several models have been developed [1] [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] [12] [13] [14] [15] [16] [17]. These models are extensions of the DFRWS model which served as the basis for digital forensic modeling approaches. These models focused on the investigative process and the different phases, they addressed the complexity of an investigation and the features and functionality of devices, and the concrete principles of an investigation. Of the models listed, one focused on a specific phase and produced empirical results. Empirical results of actual application and usage of modeling approaches by digital forensic investigators are lacking significantly. Research involving investigators is extremely limited in digital forensic research, especially when focusing on the examination phase of a digital forensic investigation. Reasons for this may be that investigators can not understand the modeling approach, investigators may be hesitant to learn a new method or model and may rely on their own departmental or organizational procedures, and/or investigators may be unaware of the different modeling approaches. In either of the cases, research is lacking to determine if, in fact, modeling approaches are being used at all in digital investigations. Furthermore, research is also needed to address knowledge management strategies in computer forensics. According to [18], "Effective knowledge management maintains the knowledge assets of an organization by identifying and capturing useful information in a usable form, and by supporting refinement and reuse of that information in service of the

organization's goals. A particularly important asset is the internal knowledge embodied in the experience of task experts that may be lost with shifts in projects and personnel." There is a need for knowledge management in digital forensics due to the increased usage of the Internet, the increase in digital crimes using different types of digital media, and the constant advances in technology. A simplified method for capturing and reusing digital crime knowledge could prove to be invaluable to the law enforcement community.

Tacit knowledge or expert knowledge is basically an internal knowing of what needs to be done and how it should be done [18]. Computer crimes are increasing, and there is a great need for knowledge sharing amongst the local, state, and federal authorities to further combat these crimes. When computer forensic examiners perform examinations, their specialized skills may not be recorded. These specialized skills could be very useful for external reviews and training. Skilled and experienced personnel know what to look for, where to look, and how to look without compromising the evidence. Externalizing this knowledge could assist novice examiners in investigations and could potentially lead to the creation of a knowledge repository. In most cases, digital forensic examiners must search through large amounts of data to find evidence. With digital storage capacities becoming increasingly larger, this task is becoming even more complex and time consuming. Knowledge management methodologies in the computer forensics domain have been addressed in [19] [20]. Bruschi, Monga, and Martignoni [19] proposed a model that organizes forensic knowledge in a reusable way. This model uses past experiences to train new personnel, to enable knowledge sharing among detective communities, and to allow third parties to assess the quality of collected information. They also suggested that disciplined methodologies should be created that provide the possibility of archiving digital forensic knowledge that would aid in training and best practice guidelines.

A method for effectively reusing and managing knowledge could greatly improve the digital forensic process. According to [20], the practice of digital forensics could be enhanced by developing "knowledge management strategies specific to law enforcement that will operate within the specific context of criminal investigations". In [19], their approach aims to provide a "methodology for archiving, retrieving, and reasoning about forensic knowledge, in order to incrementally improve the skills and the work of a team of detectives." Their proposed software tool and approach will produce reusable forensic knowledge as support during investigations, will organize past experience to encourage knowledge sharing among forensic experts, and will record collected information in a way that eases quality assessment. In order to demonstrate the importance of capturing and reusing knowledge, Kramer utilized concept maps to provide a method for capturing the tacit knowledge of design process experts.

Kramer's [21] research project attempted to collect, understand, and reuse the knowledge of multiple domain experts on design processes that drive initial design decisions associated with translating "Requirements on Orbit" to "Design Requirements." Concept maps were utilized as a knowledge acquisition and representation tool among multiple domain experts in the translation from a statement of requirements to design requirement specifications. Three specific goals for this research were as follows: demonstrating how concept maps can be used for knowledge acquisition among multiple domain experts; developing a prototype knowledge representation model from the concept maps for guiding the development of design requirements from "Statements of Requirements on Orbit"; and assessing the utility of that prototype knowledge acquisition and representation model by examination of a limited problem set. Kramer was able to effectively show the usefulness of concept maps in eliciting and representing expert knowledge; consequently, this paper explores the possibility of utilizing concept maps in the digital forensics domain. A possibility exists for incorporating concept maps into every phase of a digital investigation; however, in this research, concept mapping will be applied only to the examination phase of an investigation.

2. THE CONCEPT MAPPING CASE DOMAIN MODELING APPROACH

Conceptual models are suitable for representing the information domain of a computer forensics examination. Concept maps are a type of conceptual model that organizes and represents knowledge hierarchically by showing the relationships between concepts. Concept maps were first used in 1972 to track and better understand children's knowledge of science [22]. Since then, researchers and practitioners from various fields have used them as evaluation tools, to plan curriculums, to capture and archive expert knowledge, and to map domain information [21] [22][23]. Novak and Cañas stated that "concept mapping has been shown to help learners learn, researchers create new knowledge, administrators to better manage organizations, writers to write, and evaluators assess learning." Furthermore, a concept map can be viewed as a "simple tool [that] facilitates meaningful learning and the creation of powerful knowledge frameworks that not only permit utilization of the knowledge in new contexts, but also the retention of knowledge for long periods of time" [22]. In other words, information that is learned through the use of concept maps allows one to relate this information to previous and potentially new information and retain this information longer. Concept mapping is suitable for modeling the case domain because concept maps are easy to understand, can be used to organize information, has a semi-automated tool available, can be shared, has the ability to create new knowledge and uncover gaps in a person's knowledge.

The concept mapping case domain modeling approach (CMCDMA) was developed from Bogen's [24] case domain model and the concept mapping model used by Novak and Canãs [22]. Bogen's [24] case domain model provided a framework for analyzing case details by filtering important forensic-relevant case information; in addition, it provided a foundation for organizing knowledge and focusing a forensics examination plan, and it utilized established ontology and domain modeling methods to develop the framework of the model, and artificial intelligence and software engineering concepts, such as Unified Modeling Language (UML) conceptual diagrams, were used to represent the model. The concept mapping model provides a way to organize the case details of an examination, which could be useful later for analyzing the evidential findings. Elements of both models were used to create a five phase, non-linear process for modeling the information domain consisting of the following steps: identifying a focus question, identifying the case concepts, identifying the attributes, identifying the relationships, and instantiating the model.

First, the focus question is created. The focus question helps provide the context for the map to aid in searching for evidence and searching for additional evidence. Second, the case concepts or keywords are identified. Nouns and noun phrases or objects or events are generally chosen to represent the case information. General and specific concepts can be created and used in future investigations. Concepts can be reused from previous cases/models; reusing the concepts can save time when developing future cases/models. Figure 1 provides a representation of the concept mapping case domain model for a murder-gambling case. The case scenario for murder-gambling is as follows:

May Doe was involved in a fatal car accident at 12:25 pm, Wednesday, February 11, 2009. She was driving a 2001 Honda Accord. Her death was initially labeled an accident. However, May's parents strongly feel that she was murdered by her husband, Jim Doe. According to John, May and Jim's twenty-five year old son, his father proposed to a woman named Pam Dean one week after his mother's death. John also stated that his father received \$500,000 from his mother's life insurance policy with AcciLife Insurance Company.

Upon further review of the May's totaled vehicle, it was found that the car did not contain any brake fluid (or oil) and several holes were found in the brake line. Six days prior, May had her engine serviced as a result of the appearance of the engine service light coming on in her vehicle. A receipt taken from her purse showed that her brakes were checked, her brake fluid was refilled, and the oil was changed. In addition, a thumb drive was also found in the arm rest of May's car. Family members, friends of the family, and

neighbors were interviewed by the police; however, no one noticed anything out the ordinary between them. Everything seemed fine according to the son, but John told police that his parents had been arguing a lot lately about his father's gambling.

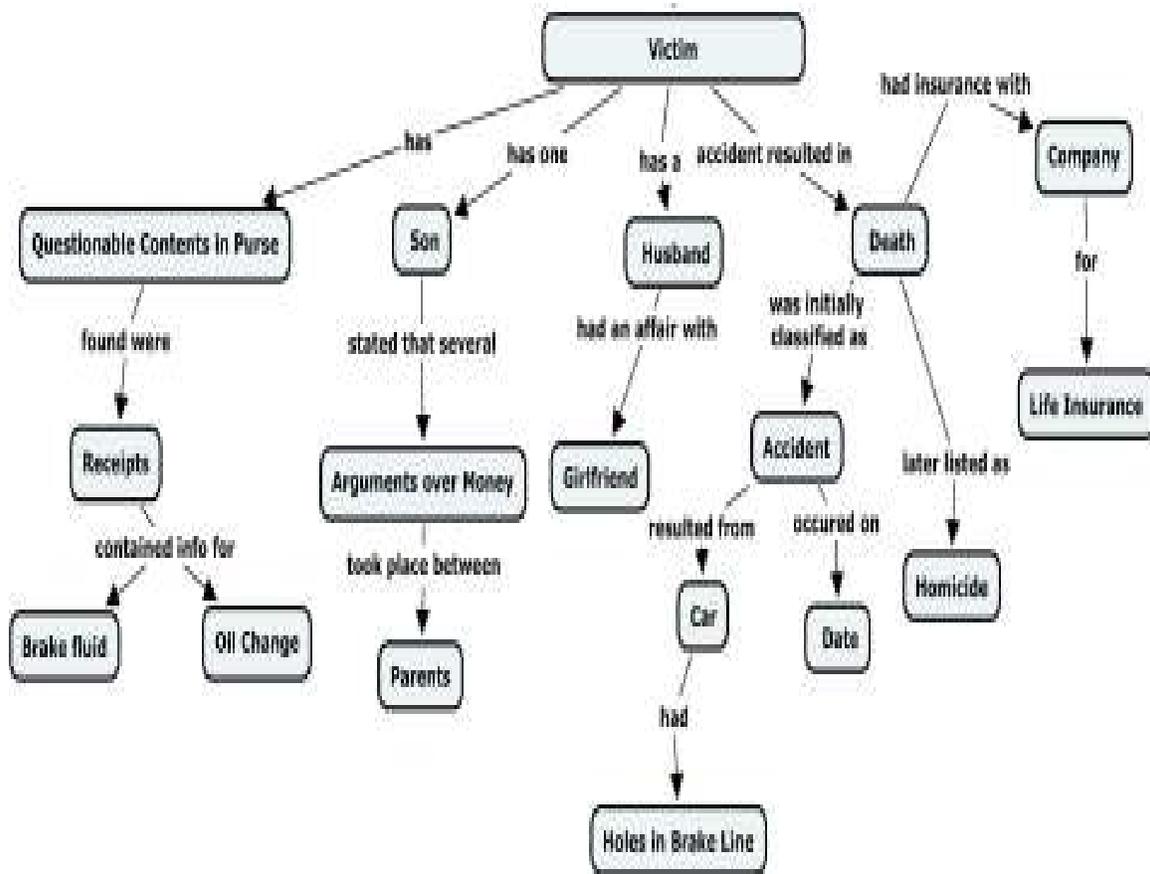


FIGURE 1: Keyword Concept Map for Murder-Gambling Case Scenario

In Figure 1, the general concepts and their relationships are shown. From this concept map, a general, quick overview of the case is shown. After the preliminary map has been created, the attributes from the case scenario should be established. Attributes help clarify the concepts' meanings, represent specific events or objects, and can be used for constructing keyword searches, examining documents, examining network logs, and linking other concepts [24]. Next, the relationships are identified. They show how the concepts are related to one another and consist of verb, verb phrases, numbers, and symbols. In the last phase of the CMCDMA, the model is instantiated by adding the attributes, or the specific information, to the map such as the name of the victim, the type of car driven by the victim, and the date the last oil change was performed as shown in Figure 2. Attributes can also include icons such as photos, documents, video, audio clips, and other digital media. Figure 3 represents an instantiated keyword concept map containing the attributes of the murder-gambling case scenario with icons displayed for the May Doe and Honda Accord Concepts. Each of the figures was created using concept mapping software, CmapTools. The concept mapping case domain model is not reliant on the CmapTools software. This model can be constructed without the use of CmapTools. However, it would be very beneficial in the law enforcement community for including additional resources such as photos, subpoenas, search warrants, and examination search procedures used. Keyword concept maps can provide an examiner with a quick way to view the evidence that was collected based on specific keywords or can be used to store documents associated with the case within

subjects' experience with computer forensic examinations had on their abilities to use the concept mapping case domain modeling approach to plan, search, and identify evidence in the digital forensic examination. The overall amount of evidence found and time spent in the phases was compared between those with little or no experience and those with experience.

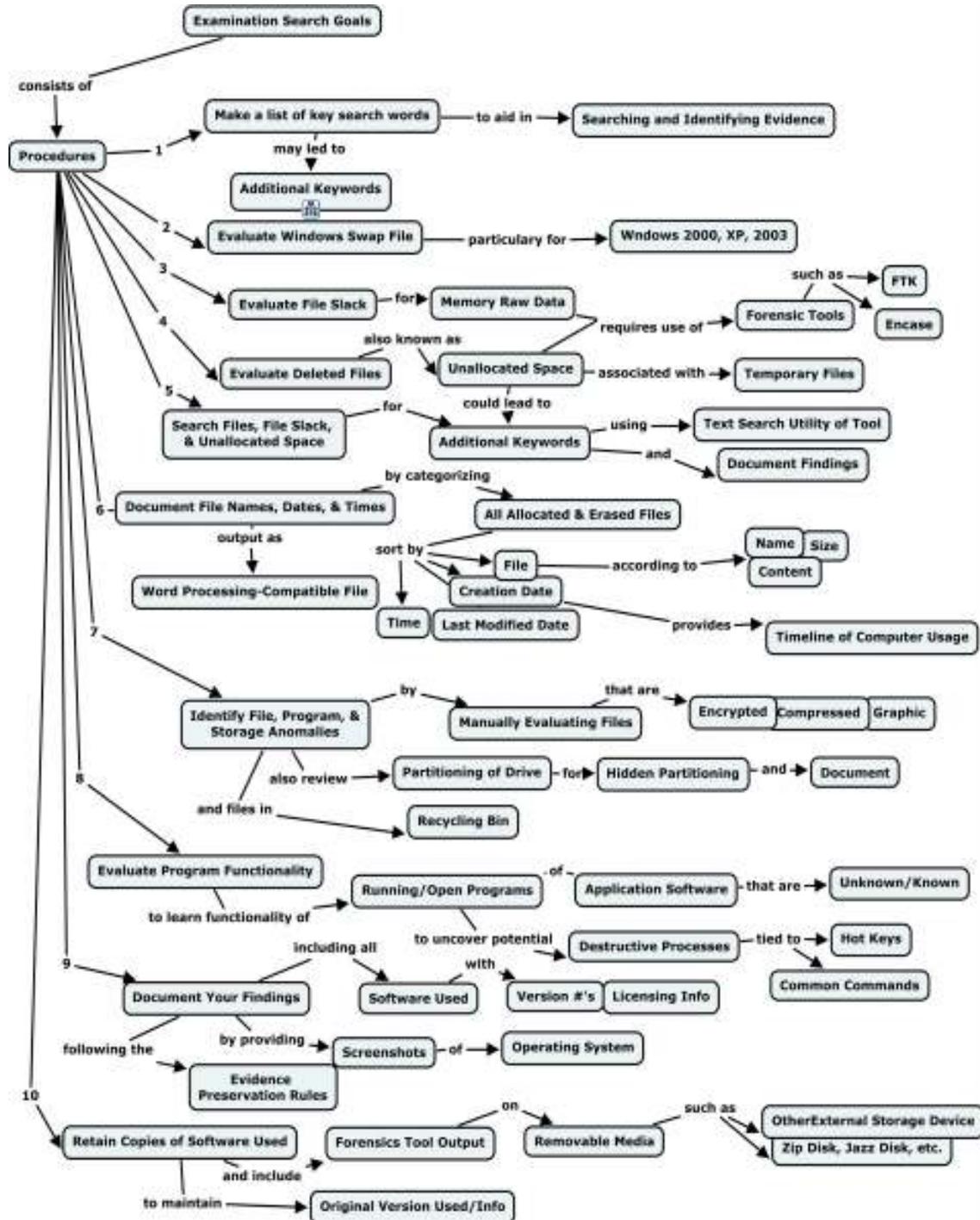


FIGURE 4: A General Examination Concept Map

Survey questions were given in an effort to obtain both qualitative and quantitative data about the concept mapping case domain modeling approach. The responses for the discussion questions

are not included; however, the analysis section includes insightful discussion responses given by the groups.

Table 1 provides the level of experience for subjects in the experimental groups for each of the four experiments based on the answers that the subjects provided voluntarily. At the beginning of the seminar course, the subjects were asked to rate their level of expertise with respect to computer forensic examinations. The experience levels were as follows:

- [A] No Experience (0-1 years) consists of knowledge of the computer forensic investigation process.
- [B] Little Experience (1-2 years) which consists of the previous experience level and attended seminars/courses/workshops in computer forensics.
- [C] Some Experience (2-3 years) consists of the previous experience levels, securing the computer/digital evidence, and notifying forensics lab, knowledge of computer forensic software and hardware.
- [D] More experience (3-4 years) consists of the previous experience levels, used digital forensic software and hardware tools to authenticate or copy evidence in an actual digital forensic investigation.
- [E] Expert/Experienced (4-5+ years) consists of the previous experience levels, performed digital forensic examinations, created reports using digital forensics software.

Experiment	[A]	[B]	[C]	[D]	[E]
Experiment 1					
E1E-1			X		
E1E-2					X
E1E-3			X		
Experiment 2					
E2E-1			X		
E2E-2				X	
E2E-3					X
Experiment 3					
E3E-1			X		
E3E-2		X			
E3E-3		X			
Experiment 4					
E4E-1		X			
E4E-2		X			

TABLE 1: Experience Level of Subjects in Experimental Groups for Experiments 1-4

In order to conduct the examination, forensics software was used to search and identify evidence utilizing the keywords and concept maps created from the concept mapping case domain modeling approach and the examination concept map. This evidence was bookmarked and included in the final report. Computer forensic software, such as FTK, allowed the case examiner to provide additional/important notes about the bookmarked evidence in addition to time and date information and the location of the evidence. For this approach, the bookmarked information was used to indicate what evidence was found and where the evidence was found. Once all the keywords had been searched and the examiner had completed his/her examination of the evidence drive, a report was generated including all of the bookmarked items created by the examiner. After the report had been created, a summary report was filled out. The summary

report aided in analyzing the evidence findings and was useful in presenting new information about the case that was unknown by the subject before the examination.

The murder-gambling case scenario discussed previously was used by the subjects in the experiment during the examination. The evidence drive consisted of a 2 gigabyte (GB) thumb drive that contained a total of 2572 files (counts were determined by Forensic Toolkit's count of file items), including 59 evidence files.

4. EXPERIMENTAL ANALYSIS

The data for these statistical analysis tests were taken from the experimental groups of the four experiments. The experimental group data was grouped into two categories: Little or No Experience (LNE) and Experienced (E). The LNE group consisted of four subjects and the E group consisted of seven subjects. The data for experiments 1-4 was combined and analyzed according to the groups. For instance, in Table 2, E3E-2 represents experiment 3 and experiment group subject 2. Table 2 represents the data collected during the planning and examination efforts in Experiments 1-4, where time is expressed in minutes. Time data information was provided for each subject in the experimental groups (concept mapping case domain modeling approach) for each experiment. In this research, subjects with little or no experience had 0-2 years experience in computer forensic examinations; in addition, those subjects with more than 2 years experience in computer forensics examinations were considered experienced.

Little or No Experience	Planning Time (minutes)	Examination Time (minutes)	Total Time (minutes)
E3E-2	38	33	71
E3E-3	5	87	92
E4E-1	7	104	111
E4E-2	55	72	127
AVERAGE	26.25	74.00	100.25
Experience			
E1E-1	10	140	150
E1E-2	44	131	175
E1E-3	30	123	153
E2E-1	13	114	127
E2E-2	40	80	120
E2E-3	40	87	127
E3E-1	27	65	92
AVERAGE	29.14	105.71	134.86

TABLE 2: Planning and Examination Effort for Experimental Groups in Experiments 1-4

Table 3 represents the amount of evidence, which is expressed as percentages, found by each subject in the experimental groups in each experiment. The evidence was classified into seven groups: Emails, May, Jim, Life Insurance, Gambling, Vehicle, and Other. The group names of the evidence represented the types of evidence and the names of the victim and suspect who had files on the evidence drive. In addition, the overall or total percentage of the evidence found by each subject and each group are provided in the last column.

Little or No Experience	% of Emails	% of May	% of Jim	% of Life Insurance	% of Gambling	% of Vehicle	% of Other	Overall %
E3E-2	33.33	100.00	14.29	100.00	91.67	50.00	45.45	54.24
E3E-3	58.33	100.00	14.29	80.00	16.67	70.00	63.64	47.46
E4E-1	58.33	50.00	14.29	60.00	58.33	50.00	81.82	55.93
E4E-2	75.00	100.00	28.57	100.00	83.33	70.00	90.91	76.27
AVERAGE	56.25	87.50	17.86	85.00	62.50	60.00	70.46	58.48
Experience								
E1E-1	91.67	100.00	28.57	80.00	83.33	70.00	63.64	79.66
E1E-2	50.00	100.00	42.86	80.00	33.33	40.00	54.55	52.54
E1E-3	58.33	100.00	28.57	100.00	50.00	40.00	63.64	57.63
E2E-1	58.33	100.00	14.29	100.00	75.00	40.00	45.45	55.93
E2E-2	50.00	100.00	42.86	80.00	83.33	50.00	54.55	62.71
E2E-3	58.33	50.00	14.29	80.00	25.00	50.00	45.45	42.37
E3E-1	58.33	100.00	57.14	80.00	58.33	30.00	36.36	52.54
AVERAGE	60.71	92.86	32.66	85.71	58.33	45.71	51.95	57.63

TABLE 3: Amount of Evidence Found in Experiments 1-4 by Experimental Groups

The chosen method of statistical analysis for testing the hypotheses in the experiment data was the independent, one-sided t-test. The t-test was used to compare the differences or means of the two independent groups. When the t-test's criteria were not met, the non-parametric Kolmogorov-Smirnov (K-S) test was used to evaluate the difference between the means of the two groups. Each of the alternative hypotheses was evaluated based on the 95% confidence interval. The alternative hypotheses were accepted and recognized as having a statistically significant difference when the probability of the null hypothesis was less than or equal to 5% or .05. Otherwise the alternative hypotheses were rejected.

The results of the t-tests and K-S tests were appropriately applied to the effort/time data, expressed in minutes, for both the LNE and E groups as shown in Table 4. If t-tests were used to evaluate the data, then the field for t-values contained a value for the test; otherwise, the K-S tests were used and the fields were marked with "-". Based on the results of the statistical tests, the concept mapping case domain modeling approach resulted in the LNE group spending a significantly less amount of time in the total experimental activity than the E group. Although no significant difference was observed during the planning and examination phases, the LNE group did spend less time in the planning and examination phases than the E group.

Hypothesis	Little or No Experience Mean (\bar{x})	Experienced Mean (\bar{y})	t	p	Result
h_{e1}	$\bar{x} = 26.25$	$\bar{y} = 29.14$	-0.258	0.401	Reject h_1
h_{e2}	$\bar{x} = 74.00$	$\bar{y} = 105.71$	-1.741	0.058	Reject h_2
h_{e3}	$\bar{x} = 100.25$	$\bar{y} = 134.86$	-2.120	0.032	Accept h_3
Hypothesis Legend					
h_{e1} = The group having little or no experience spent a significantly less amount of time in the planning phase/session than the experienced group.					
h_{e2} = The group having little or no experience spent a significantly less amount of time in the examination phase/session than the experienced group.					
h_{e3} = The group having little or no experience spent a significantly less amount of time on the total experimental activity than the experienced group.					

TABLE 4: Statistical Results for Effort Based on Experimental Group Experience Level

Table 5 provides the results of the *t*-tests and K-S tests that evaluated whether the amount of evidence found by the LNE and E groups were statistically significant. The amount of evidence found data is expressed in percentages. Based on the statistical tests, the LNE group found a significantly greater amount of evidence containing Other files than the E group. Although no other significant differences were found between the groups, the LNE group’s mean amount of evidence found was slightly higher for Gambling files, Vehicle files, and total overall evidence.

All the subjects from both groups indicated that the model was helpful in understanding the case concepts and examination tasks. The investigators all indicated that they were confident or extremely confident in their abilities to apply the modeling approach during an investigation/examination. The results of the experiment indicated that the concept mapping case domain modeling approach was useful for typical law enforcement involved in computer forensic cases. Furthermore, this experiment showed that subjects with experience or little or no experience in computer forensic examinations were able to properly use the concept mapping case domain modeling approach to plan, search for, and identify evidence. According to the post-experiment discussion survey responses, a majority of the subjects felt that the concept mapping case domain modeling approach and graphical representation would be beneficial to law enforcement during examinations, for training, and for presenting information to jurors. The subjects also stated that the CMCDMA made it easier to organize the details of the case, it offered a graphical representation of what occurred and what was discovered, and it helped them to focus and limited the amount of data to search/analyze/review. On the other hand, the subjects also felt that the CMCDMA was time consuming, the examination map was cluttered and hard to follow, and the concept map duplicated the investigator’s notes.

Hypothesis	Little or No Experience Mean (\bar{x})	Experienced Mean (\bar{y})	t	p	Result
h_{e4}	$\bar{x} = 56.25$	$\bar{y} = 60.73$	--	0.997	Reject h_4
h_{e5}	$\bar{x} = 17.86$	$\bar{y} = 32.66$	--	1.000	Reject h_5
h_{e6}	$\bar{x} = 87.50$	$\bar{y} = 92.86$	--	0.643	Reject h_6
h_{e7}	$\bar{x} = 85.00$	$\bar{y} = 85.71$	--	0.997	Reject h_7
h_{e8}	$\bar{x} = 62.50$	$\bar{y} = 58.33$	0.243	0.407	Reject h_8
h_{e9}	$\bar{x} = 60.00$	$\bar{y} = 45.71$	--	0.377	Reject h_9
h_{e10}	$\bar{x} = 70.46$	$\bar{y} = 51.95$	2.069	0.035	Accept h_{10}
h_{e11}	$\bar{x} = 58.48$	$\bar{y} = 57.62$	0.946	0.179	Reject h_{11}
Hypothesis Legend					
h_{e4} = The group with little or no experience found a significantly different amount of evidence files containing Emails than the experienced group.					
h_{e5} = The group with little or no experience found a significantly different amount of evidence containing May files than the experienced group.					
h_{e6} = The group with little or no experience found a significantly different amount of evidence containing Jim files than the experienced group.					
h_{e7} = The group with little or no experience found a significantly different amount of evidence containing Life Insurance files than the experienced group.					
h_{e8} = The group with little or no experience found a significantly greater amount of evidence containing Gambling files than the experienced group.					
h_{e9} = The group with little or no experience found a significantly different amount of evidence containing Vehicle files than the experienced group.					
h_{e10} = The group with little or no experience found a significantly greater amount of evidence containing Other files than the experienced group.					
h_{e11} = The group with little or no experience found a significantly greater amount of overall evidence than the experienced group.					

TABLE 5: Statistical Results for Amount of Data Found Based on Experience Level

The concept mapping case domain modeling approach (CMCDMA) was created to improve upon the weaknesses of Bogen’s case domain model (CDM). The goals of both models were to create a model that could be used to share and capture knowledge, to create an approach that was domain specific and could be used during the examination phase of a digital forensic investigation, to create an approach that could reduce the time spent planning and examining evidence, and to create a modeling approach that could be used to recover more evidence than when using an ad hoc approach. Section 3 discussed the experimental design and implementation of the CMCDMA; the experimental designs of both models are very similar, and Table 6 provides a brief overview of the evidence disk characteristics and maximum time allotted for the experiment. Although the size of the evidence drive in the CMCDMA was smaller than the sizes of the drives in Bogen’s CDM experiments, both modeling approaches utilized similar techniques, such as keyword searches, that saved time and eliminated the needed to search through every file on the evidence drive. This technique allowed the subjects to utilize forensic software that would find specific key terms on the entire evidence drive very quickly. In addition, the subjects in the CDM experiments were given four hours for planning and a hour and a half more time to search for evidence, while subjects in the CMDCMA experiment were given a maximum of two hours for both planning and examination.

Experiment Approach Used	# of Evidence Files	Total # of Files on Evidence Disk	Size of Evidence Disk (GB)	Maximum Time Allowed in Experiment
CMCDMA	59	2572	1	2.5 hours
Bogen (CDM)				
Experiment 1	99	2981	40	4 hours
Experiment 2	29	58,459	40	4 hours
Experiment 3	33	58,894	10	4 hours

TABLE 6: Comparison of CMCDMA and CDM Experimental Design Data

Experiment Approach Used	Mean Planning Time (min)	Mean Examination Time (min)	Mean Total Time (min)	Overall % of Evidence Found
CMCDMA	28.09	94.18	122.28	57.94
Bogen (CDM)				
Experiment 1	162.83	167.00	329.83	49.33
Experiment 2	134.14	137.71	271.86	35.47
Experiment 3	78.67	89.17	167.83	25.50

TABLE 7: Comparison of CMCDMA and CDM Experimental Group Data

In Table 7, the data for the CMCDMA experiment was combined from Table 2 and Table 3 in order to determine the mean planning time, mean examination time, the mean total time spent in the experiment, and the overall percentage of evidence found by the experimental group, which consists of an aggregation of the LNE and E groups data. Table 7 also provides the results of Bogen’s CDM experiments as well. The data shows that subjects using the CMCDMA spent less time planning, less total time in the experiment, and found at least 7% more evidence than those subjects using Bogen’s CDM method. In all but one of the experiments, the subjects in the CDM experiment spent more time in the examination phase of the experiment than those subjects in the CMCDMA. Reasons for the large amount of time differences in the planning times of the CMCDMA and the CDM method are that the CDM method was more paper intensive and required the subjects to fill out forms and transfer the information to other forms; also, the subjects were required to complete four activities, which consisted of modeling the information domain of the case utilizing UML conceptual diagrams, developing search goals, specifying search methods for each search goal, and finally conducting the examination. Furthermore, each of these activities required additional tasks to be performed. In relation to the CDM method, the CMCDMA consisted of only modeling the information domain of the case utilizing concept maps, which was one process composed of five tasks.

Although the CDM approach was successful in allowing the subjects to recover more evidence than when using an ad hoc approach, several of the CDM subjects indicated that Bogen’s method felt more like paperwork; in addition, they indicated that the availability of semi-automated software would have allowed them to model the details of the case and document their findings. The CMCDMA was developed to provide a simpler way represent the case details of the investigation using concept maps that could contain evidence items specific to the case such as photos, documents, video, and other files. All the information related to the case could be accessed in one location, including other important documents such as subpoenas, search warrants, and other critical documents. The concept maps also provided a quick way to review the case details, to locate keywords that could be used to search the evidence drive, and to manage knowledge gained for a particular type of case. General concepts and concept maps created for that particular case could be used in future cases and altered to include specific

attributes for each of the different cases. Managing knowledge in this way could greatly reduce the time needed to investigate and examine digital forensic cases in the future.

5. CONCLUSION AND FUTURE WORK

This paper described the need for managing knowledge in a digital forensic investigation. The concept mapping case domain modeling approach was presented as a method for managing knowledge acquired during a digital forensic examination. The approach provided a way to visually represent the knowledge gained during an investigation and also discussed how the approach can be applied in real digital forensic cases, for training and during an examination. Empirical evidence was provided that showed how novice and experienced law enforcement officers used the approach to plan, search, and identify digital evidence in a digital forensic examination. More research is needed to address ways to model the knowledge obtained during a digital forensics investigation due to the ever increasing sizes of digital storage. Domain modeling in digital forensics is still an emerging area. Several modeling approaches have been proposed, however, little or no empirical data is available for comparing the applicability and usability of these approaches by law enforcement and forensic practitioners. From researching several methods, no other experimental data is available in domain modeling other than Bogen's experimental results. Resultantly, there is a substantial need for experimental data produced by modeling approaches in order to determine if these modeling approaches can be applied by law enforcement in real world cases, if they are useful for managing knowledge, and if these approaches can improve digital forensics investigations by reducing the amount of time needed to examine digital forensic evidence. Future research endeavors include automating the concept map creation process after the examination of digital forensic evidence has occurred. Most digital forensic examiners use computer forensic tools to examine digital evidence, such as Encase and AccessData's Forensic Toolkit (FTK). An automated process could be developed that creates and positions the concepts based on the data in the digital forensic examiner's Encase or FTK report and categorize the evidential findings based on their file extensions, date and time, and etc.

6. REFERENCES

1. V. Baryamureeba, F. Tushabe. "The Enhanced Digital Investigation Process Model". In Proceedings of the 4th Annual Digital Forensic Research Workshop, Baltimore, MD, 2004
2. N. Beebe and J. Clark. "A Hierarchical, Objectives-Based Framework for the Digital Investigations Process". In Proceedings of the 4th Annual Digital Forensic Research Workshop, Baltimore, MD, 2004
3. B. Carrier and E. Spafford. "An Event-Based Digital Forensic Investigation Framework". In Proceedings of the Fourth Annual Digital Forensic Research Workshop, Baltimore, MD, 2004
4. S. Ciardhuáin. "An Extended Model of Cybercrime Investigations". International Journal of Digital Evidence, 3(1):1-22, 2004
5. M. Reith, C. Carr, G. Gunsch. "An Examination of Digital Forensic Models". International Journal of Digital Evidence, 1(3):1-20, 2002
6. G. Ruibin, T. Yun, M. Gaertner. "Case-Relevance Information Investigation: Binding Computer Intelligence to the Current Computer Forensic Framework". International Journal of Digital Evidence, 4(1):1-13, 2005
7. J. Venter. "Process Flow Diagrams for Training and Operations". Advances in Digital Forensics II, Springer, pp. 331-342 (2006)
8. Tanner and D. Dampier. "Concept Mapping for Digital Forensics Investigations". Advances in Digital Forensics V, Springer, pp. 201-300 (2009)

9. Tanner and D. Dampier. *"Improving Digital Forensics Investigations with Concept Mapping"*. In Proceedings of the Fifth International Conference on Digital Forensics, Orlando, FL, 2009
10. S. Peisert, M. Bishop, S. Karin and K. Marzullo. *"Toward Models for Forensic Analysis"*. In Proceedings of the Second International Workshop on Systematic Approaches to Digital Forensic Engineering. Bell Harbor, WA, 2007
11. M. Khatir, S. M. Hejazi and E. Sneider. *"Two Dimensional Evidence Reliability Amplification Process Model for Digital Forensics"*. In Proceedings of the Third International Workshop on Digital Forensics and Incident Analysis. Malaga, Spain, 2008
12. Y. Shin. *"New Digital Forensics Investigation Procedure Model"*. In Proceedings of the Fourth International Conference on Networked Computing and Advanced Information Management. Gyeongju, Korea, 2008
13. Carrier, E. Spafford. *"Getting Physical with the Digital Investigation Process"*. International Journal of Digital Evidence, 2(2):1-20, 2003
14. National Institute of Justice. Electronic Crime Scene Investigation: A Guide for First Responders 2001 [Online]. Available at: <http://www.ncjrs.gov/pdffiles1/nij/187736.pdf>, 2001
15. M. Pollitt. *"An Ad Hoc Review of Digital Forensic Models"*. In Proceedings of the Second International Workshop on Systematic Approaches to Digital Forensic Engineering. Bell Harbor, WA, 2007
16. R. Rowlingson. *"A Ten Step Process for Forensic Readiness"*. International Journal of Digital Evidence, 2(3):1-28, 2004
17. P. Stephenson. *"Modeling of Post-Incident Root Cause Analysis"*. International Journal of Digital Evidence, 2(2):1-16, 2003
18. Cañas, D. Leake, and D. Wilson. *"Managing, Mapping, and Manipulating Conceptual Knowledge"*. IHMC, 2007
19. Bruschi, M. Monga, and L. Martignoni. *"How to Reuse Knowledge about Forensic Investigations"*. In Proceedings of the 4th Annual Digital Forensic Research Workshop. Baltimore, MD, 2004
20. M. Pollitt and A. Whitley. *"Exploring Big Haystacks: Data Mining and Knowledge Management"*. Advances in Digital Forensics II, Springer, pp. 67-76 (2006)
21. M. Kramer. *Using Concept Maps for Knowledge Acquisition in Satellite Design: Translating "Statement of Requirements on Orbit" to "Design Requirements"*. PhD Thesis, Nova Southeastern University, 2005
22. J. D. Novak and A. J. Cañas. *"The Theory Underlying Concept Maps and How to Construct Them"*. Technical Report IHMC Cmap Tools 2006-01, Florida Institute for Human and Machine Cognition, 2006
23. S.O. Tergan, *"Digital Concept Maps for Managing Knowledge and Information: Searching for Synergies"*. Knowledge and Information Visualization, Springer, pp. 185-204 (2005)
24. C. Bogen. *"Selecting Keyword Search Terms in Computer Forensics Examinations using Domain Analysis and Modeling"*, PhD Thesis, Department of Computer Science and Engineering, Mississippi State University, 2006

Securing Image Transmission Using In- Compression Encryption Technique

Shaimaa A. El-said

*Faculty of Engineering / Electronics
and Communication Department
Zagazig University
Zagazig, 44519, Egypt.*

Eng.sahmed@windowlive.com

Khalid F. A. Hussein

*Electronics research institute
Microwaves Department
Researches National Institute
Dokki, Egypt*

khalid_elgabaly@yahoo.com

Mohamed M. Fouad

*Faculty of Engineering / Electronics
and Communication Department
Zagazig University
Zagazig, 44519, Egypt.*

fouadzu@hotmail.com

Abstract

Multimedia is one of the most popular data shared in the Web, and the protection of it via encryption techniques is of vast interest. In this paper, a secure and computationally feasible Algorithm called Optimized Multiple Huffman Tables (OMHT) technique is proposed. OMHT depends on using statistical-model-based compression method to generate different tables from the same data type of images or videos to be encrypted leading to increase compression efficiency and security of the used tables. A systematic study on how to strategically integrate different atomic operations to build a multimedia encryption system is presented. The resulting system can provide superior performance over other techniques by both its generic encryption and its simple adaptation to multimedia in terms of a joint consideration of security, and bitrate overhead. The effectiveness and robustness of this scheme is verified by measuring its security strength and comparing its computational cost against other techniques. The proposed technique guarantees security, and fastness without noticeable increase in encoded image size.

Keywords: Image encryption and compression, optimized multiple Huffman tables, OMHT performance analysis, computational cost analysis

1. INTRODUCTION

With the rapid development of multimedia and network technologies, the security of multimedia becomes more and more important, since multimedia data are transmitted over open networks more and more frequently. Typically, reliable security is necessary to content protection of digital images and videos. Encryption schemes for multimedia data need to be specifically designed to

protect multimedia content and fulfill the security requirements for a particular multimedia application. For example, real-time encryption of an entire video stream using classical ciphers requires heavy computation due to the large amounts of data involved, but many multimedia applications require security on a much lower level, this can be achieved using selective encryption that leaves some perceptual information after encryption.

As an important way of designing a secure video encryption schemes, secret Multiple Huffman Tables (MHT) have been suggested in some designs. The major advantage by using this kind of joint compression-encryption approach is that high compression ratio and high encryption degree can be achieved in one single step, which simplifies the system design and makes it flexible for some advanced multimedia processing [1] in addition to the reduction of time required to perform compression followed by encryption. After re-studies the security of multimedia encryption scheme based on secret Huffman tables, the present cryptanalysis shows presence of drawbacks in MHT technique.

To overcome the drawbacks of MHT technique, a new scheme for more general and efficient secure multimedia transmission, OMHT, is proposed. OMHT depends on using statistical-model-based compression method to generate different tables from a training set has the same data type as images or videos to be encrypted leading to increase compression efficiency and security of the used tables. Using known fixed tables in MHT technique generated by mutation (a method introduced in [1]) for compressing and encrypting images causes degradation in both compression ratio and security. We focus our research attention to enhancing multiple Huffman tables coding techniques. It is a challenging problem to verify joint consideration of security, bitrate overhead, and friendliness to delegate processing. Performance analysis of the newly proposed scheme OMHT shows that it can provide superior performance over both generic encryption and MHT in the security and compression.

This paper is organized as follows: Section 2 shows an overview of multimedia encryption techniques. A new proposed scheme, Optimized Multiple Huffman tables coding technique (OMHT) is described in section 3 with a detailed description for proposed adaptive quantization technique. Section 4 presents a performance analysis of the proposed scheme OMHT technique. The computational cost of the proposed technique is analyzed in section 5. Conclusion is given in section 6.

2. OVERVIEW of MULTIMEDIA ENCRYPTION TECHNIQUES

When dealing with still images, the security is often achieved by using the naïve (traditional) approach to completely encrypt the entire image, traditional encryption, with a standard cipher [2] (DES, AES, IDEA, etc.). As shown in Fig. (1), assuming that the plaintext and the ciphertext are denoted by P and C , respectively, the encryption procedure in a cipher can be described as $C = E_{K_e}(P)$, where K_e is the encryption key and $E(\cdot)$ is the encryption function. Similarly, the decryption procedure is $P = D_{K_d}(C)$, where K_d is the decryption key and $D(\cdot)$ is the decryption function. When $K_e = K_d$, the cipher is called a private-key cipher or a symmetric cipher. For private-key ciphers, the encryption-decryption key must be transmitted from the sender to the receiver via a separate secret channel. When $K_e \neq K_d$, the cipher is called a public-key cipher or an asymmetric cipher. For public-key ciphers, the encryption key K_e is published, and the decryption key K_d is kept private, for which no additional secret channel is needed for key transfer. Ciphering the complete compressed file may result in excessive computational burden and power consumption at the decoder and perhaps even the server/ encoder.

However, there are number of applications for which the naïve based encryption and decryption represents a major bottleneck in communication and processing. Some recent works explored a new way of securing the content, named, partial encryption or selective encryption, soft encryption, perceptual encryption, by applying encryption to a subset of a bitstream. The main goal of selective encryption is to reduce the amount of data to encrypt while achieving a required level of security [3].

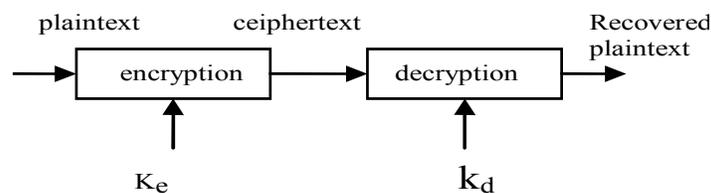


Figure 1: Traditional Encryption Techniques

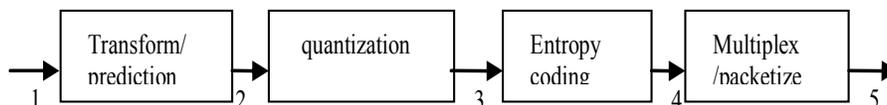


FIGURE 2: Candidate Domains Used to Apply Encryption to Multimedia.

According to Fig. 2, there are two straight forward places to apply generic encryption to multimedia. The first possibility is to encrypt multimedia samples before any compression, stages 1 and 2, Qiao et al. [4] and Uehara and Safavi-Naini [5] are examples of pre-compression selective encryption. The main problem with this approach is that the encryption often significantly changes the statistical characteristics of the original multimedia source, resulting in much reduced compressibility. Cheng and Li, 2000. The wavelet-based compression algorithm SPIHT [6] is an example of post-compression encryption scheme, stage 4 and 5. Wu et al proposed encryption scheme based on encoding with multiple Huffman tables (MHT) used alternately in a secret order [1]; is an example of in-compression selective encryption stages 3, and 4. The encryption with reasonably high level of security and unaffected compression can be achieved simultaneously, requiring almost negligible additional overhead. One of the major advantages by using this kind of joint encryption-compression approach is that encryption and compression can be achieved in one single step, which simplifies the system design and makes it flexible for some advanced multimedia processing such as scalability and rate shaping.

2.1. Multiple Huffman Tables (MHT) Technique

The MHT algorithms [1][7]-[9], aiming to increase the model space while maintaining the computational efficiency, keep the structure of the Huffman tree but enlarge the model space through tree mutation. The procedure of the basic MHT algorithm is described as follows:

- Step1: Train four original Huffman trees from different sets of training data. e.g. Huffman table of the JPEG DC coefficients.
- Step2: Based on the original trees, perform tree mutation, to create the whole Huffman tree space.
- Step3: Randomly select m different tables from the space, and number them from 0 to $m-1$.
- Step4: Generate a random vector $P = \{P_0, P_1, \dots, P_{m-1}\}$ each p is an Integer ranging from 0 to $m-1$.
- Step5: For the i^{th} encountered symbol, use table $P_{i \pmod{m}}$ to encode it.

MHT coding [1] makes use of standard coding tables. It is included in the final bit-stream for every image to be compressed. This approach presents disadvantages:

1. Visual degradation: very high-visual degradation can be achieved.

2. Cryptographic security: Gillman and Rivest [10] showed that decoding a Huffman coded bitstream without any knowledge about the Huffman coding tables would be very difficult. However, the basic MHT is vulnerable to known and chosen plaintext attacks as pointed out in [11].
3. It writes all codes of the corresponding tables in the final bitstream even if only some of them were used to encode the associated events of the particular input image.
4. It does not make use of any statistic about the distribution of the events of the image. To improve the security several kinds of enhanced MHT schemes have been proposed:
 - By inserting random bit in the encrypted bit stream or integrating with a stream cipher [8].
 - Recently another scheme via random rotation in partitioned bit streams has been reported [9].

3. OPTIMIZED MULTIPLE HUFFMAN TABLES (OMHT)

OMHT compression-encryption technique is a modification to the MHT scheme; it generates different Huffman tables for each type of images instead of using fixed Huffman tables for all images as in MHT technique. The main advantage of using OMHT technique over other lossy compression technique is that it produces a much smaller compressed file than any compression method, while still meeting the advantage of encryption. Remove small, invisible parts, of the picture is based on an accurate understanding of how the human brain and eyes work together to form a complex visual system. As a result of these subtle reductions, a significant reduction in the resultant file size for the image sequences is achievable with little or no adverse effect in their visual quality. As shown in Fig. 3 OMHT process takes two parallel paths A, and B, so it takes no additional time to add encryption to the compressed bitstream as both traditional and selective encryption techniques.

3.1. The Procedure of Compressing the Original Image

As shown in Fig. 3 (path A), The input $N \times M$ image is first converted into single vector by concatenating successive rows beside each other to form a long row that contains all the image pixels using matrix to vector converter. This vector is exposed to DCT to transform the image from spatial domain into frequency domain in which energy of the image information is concentrated in a few number of coefficients. The output of the DCT process is a vector that has the same length of the image (number of pixels in the image), but with many values approximated to zeros. After applying the DCT the output coefficients are arranged in a descending order according to its energy content. The energy content of the coefficients is summed from the beginning of the vector and toward the end till a specific energy percentage (EP) of the image energy is reached. Those coefficients that carry EP energy percent are chosen to be transmitted and the rest coefficients are neglected since they carry only very small energy that will not affect the visual quality of the recovered image. This EP value depends on image characteristics and it can be varied to achieve the desired compression ratio and the signal to noise ratio according the application: As we decrease the EP value, a higher compression ratio is obtained with slightly lower signal to noise ratio. Now the number of the transmitted coefficients (Tc) becomes very small. The reduced coefficients vector returned back to the spatial domain using IDCT to be processed by an efficient quantizer.

The proposed Least Probable Coefficients Approximation (LPCA) quantizer as shown in Fig. 4(a) and its flowchart in Fig. 4(b) reduces the output values of the IDCT by calculating their occurrence probabilities. The IDCT coefficients are arranged in a descending order according to their probabilities in a vector. The desirable quantization levels are taken as the most probable coefficients from the beginning of the arranged vector; if the required CR and SNR are achieved by transmitting only four quantization levels, those quantization levels are the first four coefficients in the arranged vector. The probability of the last QL is called neglecting probability (NP).

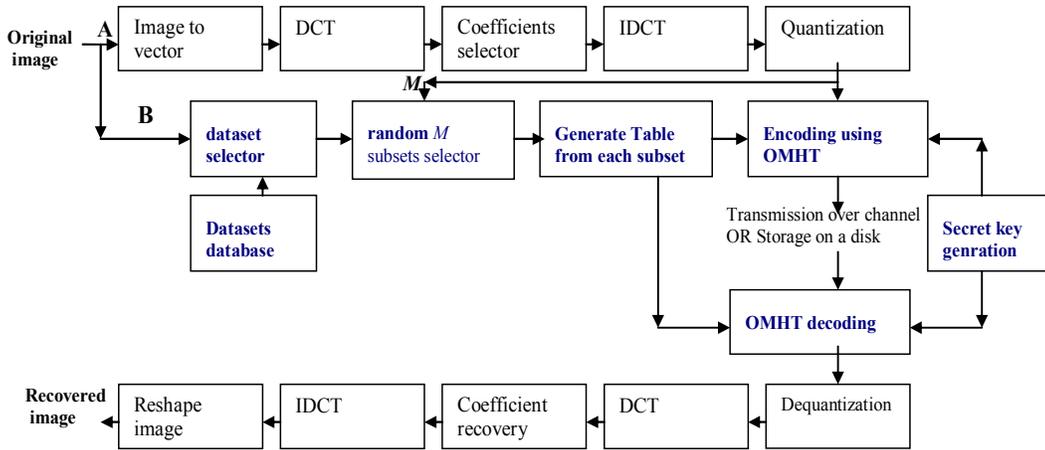


FIGURE 3: Optimized Multiple Huffman Table (OMHT) Coding System

All coefficients with probability less than Np are approximated to the nearest quantization level in value. This technique is irreversible; this means that the dequantized values can't be turned back to their original values leading to quantization losses. The quantization procedures are summarized as follows:

- IDCT coefficients are arranged in a descending order according to their probabilities in a vector.
- The desirable n quantization levels are taken as the n most probable coefficients from the beginning of the arranged vector.
- The probability of the last QL is called neglecting probability (NP).
- All coefficients with probability less than Np are approximated to the nearest quantization level in value. The proposed quantization reduces number of transmitted values but not the number of transmitted coefficients.

After the transmitted values are reduced by quantization, each quantized level is assigned a codeword using Huffman encoder that enables representing an image in a more efficient way with smallest memory for storage or transmission. Huffman coding is used to code the quantized values statistically according to their probability of occurrences. Short code words are assigned to highly probable values and long code words to less probable values. The average number L_{avg} of bits required to represent a symbol is defined as,

$$L_{avg} = \sum_{k=1}^L I(r_k)P(r_k)$$

Where, r_k is the discrete random variable for $k=1,2,\dots,L$ with associated probabilities $P(r_k)$. The number of bits used to represent each value of r_k is $I(r_k)$. The number of bits required to represent an image is calculated by number of symbols multiplied by L_{avg} [9].

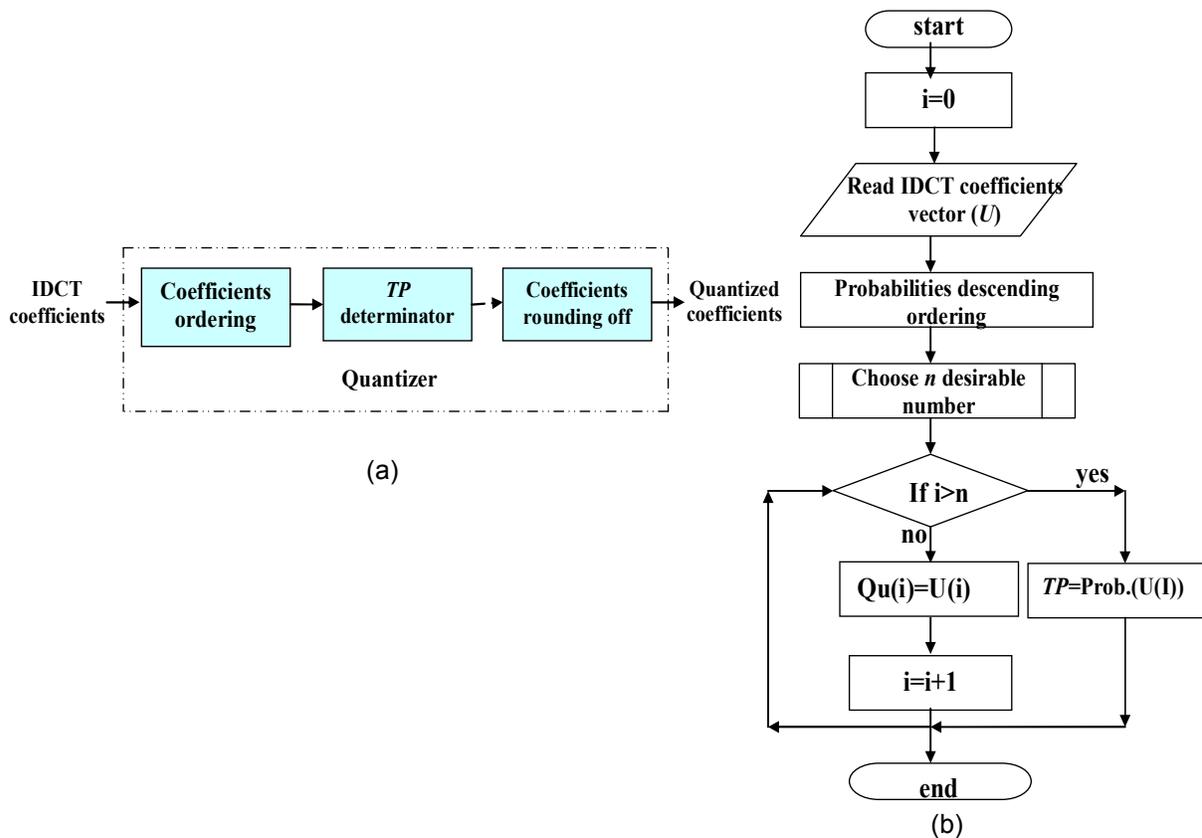


FIGURE 4: (a) Block Diagram of LPCA Quantizer (b) LPCA flowchart

3.2. Procedure of Preparing and Using the OMHT Tables

Following is the procedure of preparing and using the optimized multiple Huffman tables and how it is used to both encode and encrypt images as shown in Fig. 3 (path B).

Step 1: images training set are divided into L datasets. Each dataset's images have the same properties.

Step 2: each dataset contains N images.

Step3: The input image compared to datasets to select the dataset that has the same properties.

Step4: randomly choose M subsets each subset contains K images from the dataset. Concatenate all images of each subset and calculates the pixels probabilities. Then draw Huffman tree and find the Huffman table contains the different pixels' values and their associated variable codewords. Now we have M different tables to be used.

Step5: Tables are saved at each decoder, and the order by which the tables are generated and used is kept secret.

Step6: Number the generated M tables from 0 to M-1.

Step7: Generate a random vector P (the secret order) its length equal to the length of image under consideration. Each element value in P ranges from 0 to M-1.

Step8: For the *i*th encountered symbol (coefficient to be encoded), use table $P(i \bmod n)$ to encode it.

4. PERFORMANCE ANALYSIS of OMHT

For performance evaluation, the following experiments measure the compression performance and encryption strength of OMHT using test images that contains gray and colored images. The compression performance of OMHT is analyzed by calculating the compression ratio (CR), number of bits per symbol (BPP), the peak signal to noise ratio (PSNR), and the mean square error (MSE). A comparison between the proposed scheme based on generating tables based on statistical modeling of large dataset for each types of images, and a compression using fixed predetermined encoding tables, JPEG standard, on which the MHT technique based on done to show the effectiveness of the proposed scheme in compression. The encryption strength of the OMHT is tested and compared with other encryption techniques.

The achieved compression ratio can be calculated from the following equation:

$$CR = \frac{\text{original}}{\text{compressed}}$$

Where the original is the size of the original image and the compressed is the size of the Huffman encoder output compressed bitstream. Calculate the bit per pixel (BPP) is defined as:

$$BPP = \frac{B}{P}$$

Where P is the total number of pixels in an image and B is the total number of transmitted bits for this image. As a measure of reconstructed image quality, the peak signal-to-noise ratio (PSNR) in dB is used, this is defined as follows:

$$PSNR_{dB} = 20 \log_{10} \frac{2^n - 1}{\sqrt{MSE}}$$

Both mean square error (MSE) and the signal to noise ratio (SNR) for an $n \times n$ image are calculated from the following equations:

$$MSE = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n (\alpha[i,j] - \beta[i,j])^2$$

$$SNR = 10 \log_{10} \times \frac{\sum_{i=1}^n \sum_{j=1}^n (\alpha[i,j]^2)}{\sum_{i=1}^n \sum_{j=1}^n (\alpha[i,j] - \beta[i,j])^2}$$

Where, $\alpha[i,j]$ and $\beta[i,j]$ denote the original and decoded levels of the pixel $[i,j]$ in the image, respectively. A larger PSNR value means that the encoded image preserves the original image quality better.

Experiment 1 uses lossy OMHT to encrypt and compress the Lena image. It gives the ability to control the compression ratio and peak signal to noise ratio by either change number of QL while the amount of Tc is constant, or changing the amount of Tc while number of QL is constant.

As shown in Table 1, and Fig. 5, and 6, while the number of quantization levels is constant at $q=128$ and the amount of transmitted DCT's coefficients changes from $Tc=98.5\%$ of the image energy to $Tc= 99.9\%$. As Tc increases, the CR decreases providing an increase in PSNR.

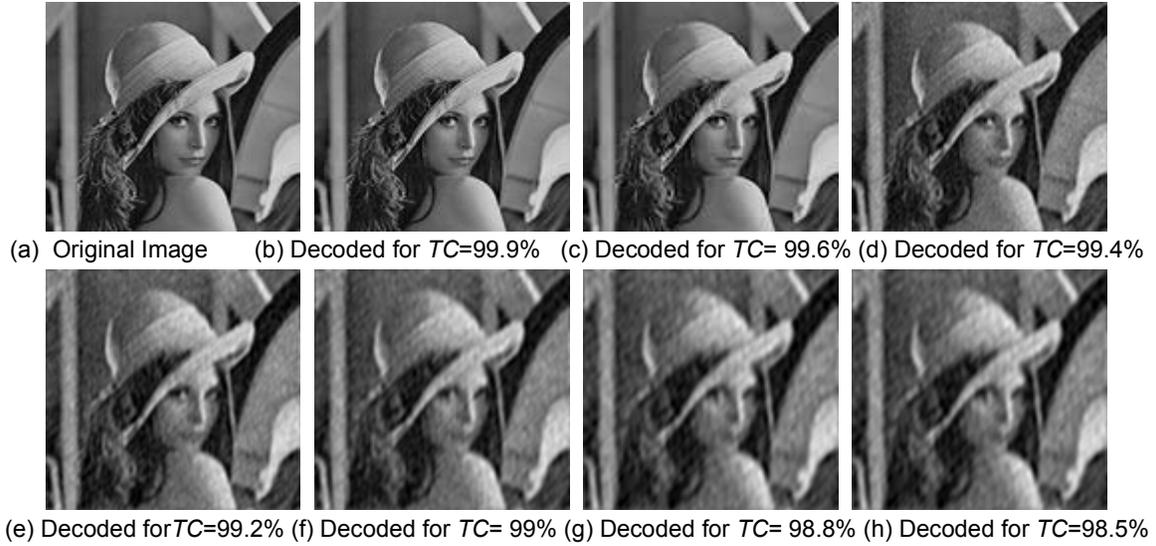


FIGURE 5: The Effect of Reducing Number of Transmitted Coefficients on Image Visual Degradation

	<i>Tc=99.9%</i>	<i>Tc=99.6%</i>	<i>Tc=99.4%</i>	<i>Tc=99.2%</i>	<i>Tc=99%</i>	<i>Tc=98.8%</i>	<i>Tc=98.5%</i>
CR	2.8179	7.0518	10.860	15.252	20.612	26.936	38.973
BPP	2.8390	1.1345	0.7366	0.5245	0.3881	0.2970	0.2074
PSNR	34.032	28.744	26.7053	25.929	24.794	24.065	23.125
MSE	20.848	94.222	116.673	179.287	200.80	290.092	360.09
SNR	26.947	21.669	19.6207	18.8446	17.719	16.9809	16.040

TABLE 1: Compression Performance of Applying OMHT with Different Number of Transmitted Coefficients

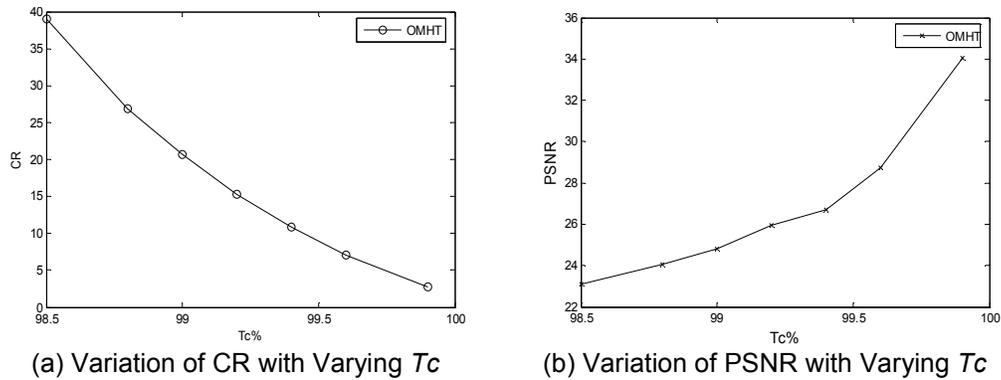


FIGURE 6: The Effect of Reducing OMHT Number of T_c on CR, and PSNR

OMHT provides the ability to maintain T_c constant and varies the number of QL . As shown in Table 2, and Fig. 7, and 8, while the amount of transmitted coefficients is constant at $T_c=99.5%$ and the number of quantization levels changes from using four quantization levels to using 256 quantization levels. As the number of quantization levels increases, the compression ratio decreases providing an increase in peak signal to noise ratio.

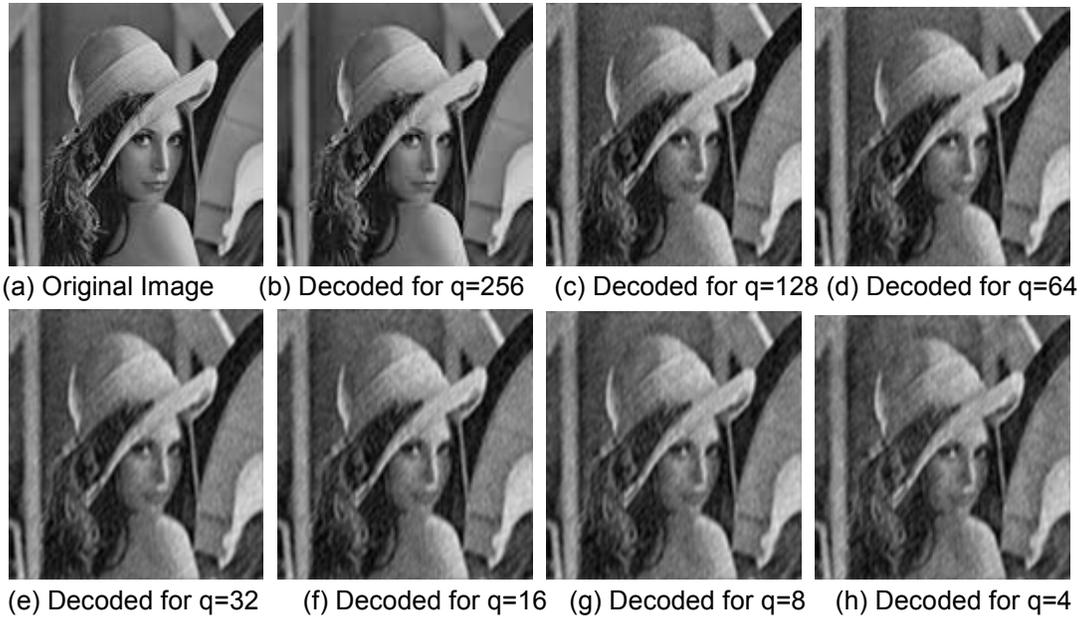


FIGURE 7: The Effect of Reducing Number of Quantization Levels on Image Visual Degradation

	$Q=4$	$Q=8$	$Q=16$	$Q=32$	$Q=64$	$Q=128$	$Q=256$
CR	24.08	18.437	15.99	13.25	11.45	10.49	10.14
BPP	0.332	0.4339	0.500	0.604	0.699	0.762	0.789
PSNR	22.09	23.295	24.42	26.14	27.59	27.75	27.92
MSE	382.2	305.51	222.2	119.67	102	98.12	80.94
SNR	15	16.21	17.33	19.06	20.5	20.66	20.84

TABLE 2: Compression Performance of Applying OMHT with Different QL on Lena Image

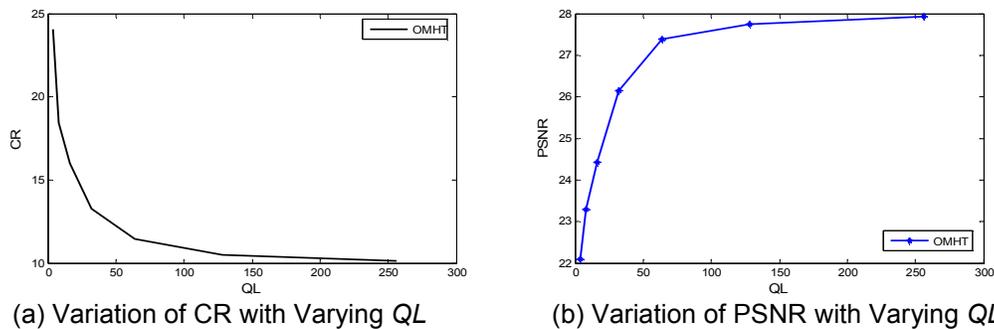


FIGURE 8: The Effect of Reducing OMHT Number of Quantization Levels on CR, And PSNR

Experiment 2 compares the compression performance of lossy OMHT with that of lossy JPEG technique to prove that the proposed technique adds security without affecting the compression ratio or the PSNR. Table 3 provides a comparison of CR between OMHT, and JPEG at Different BPP on Lena Image, while Table 4 provides a comparison of PSNR between OMHT, and JPEG at Different BPP on Lena Image. From Tables 3 and 4 it is obvious that using lossy OMHT technique provides higher PSNR and storage space and transmission bandwidth required than JPEG especially at low bitrates.

BPP	OMHT	JPEG
0.2	39	39.01
0.18	44.4	43
0.16	49.4	49
0.14	57.4	58.02
0.12	65.6	65.3
0.1	73.9	73

TABLE 3: Comparison of CR between OMHT, and JPEG at Different BPP on Lena Image

BPP	OMHT	JPEG
0.2	22.6	21.14
0.18	22.2	20
0.16	21.9	19.4
0.14	21.6	18
0.12	21.3	16.7
0.1	21	15

TABLE 4: Comparison of PSNR between OMHT, and JPEG at Different BPP on Lena Image

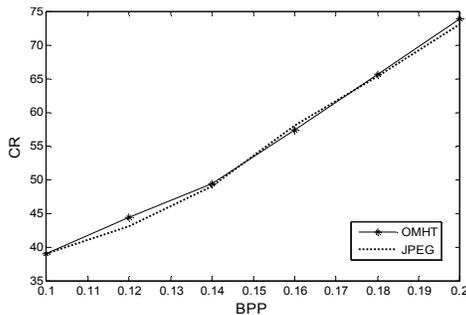


FIGURE 9: CR of both OMHT technique and JPEG for Lena image

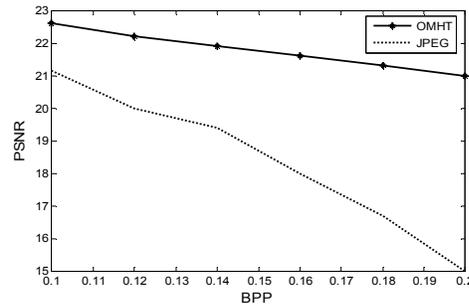


FIGURE 10: PSNR of both OMHT technique and JPEG for Lena image

Fig. 9 shows that both OMHT technique and JPEG technique have nearly the same compression levels at the same number of bits per pixel for Lena image. Fig. 10 shows that PSNR of OMHT technique is higher and more stable at low bitrate than that of JPEG for Lena image.

Experiment 3 measures the encryption strength performance of the proposed OMHT technique, colored football image in RGB (288x352x3) shown in Fig. 11(a) with its histogram shown in Fig. 11(b) is compressed and encrypted using the OMHT uses multiple Huffman tables, generated from a large set of training images that have the same type of the test image used in a secret order (secret key). Fig. 11(c) and 11(d) shows the test image and its histogram after decoding it with another technique as JPEG. While Fig. 11(e) and 11(f) shows the test image and its

histogram after decoding it with OMHT technique and the same encoding tables but without knowing the secret order (secret key). It is obvious that OMHT provides high perceptual security

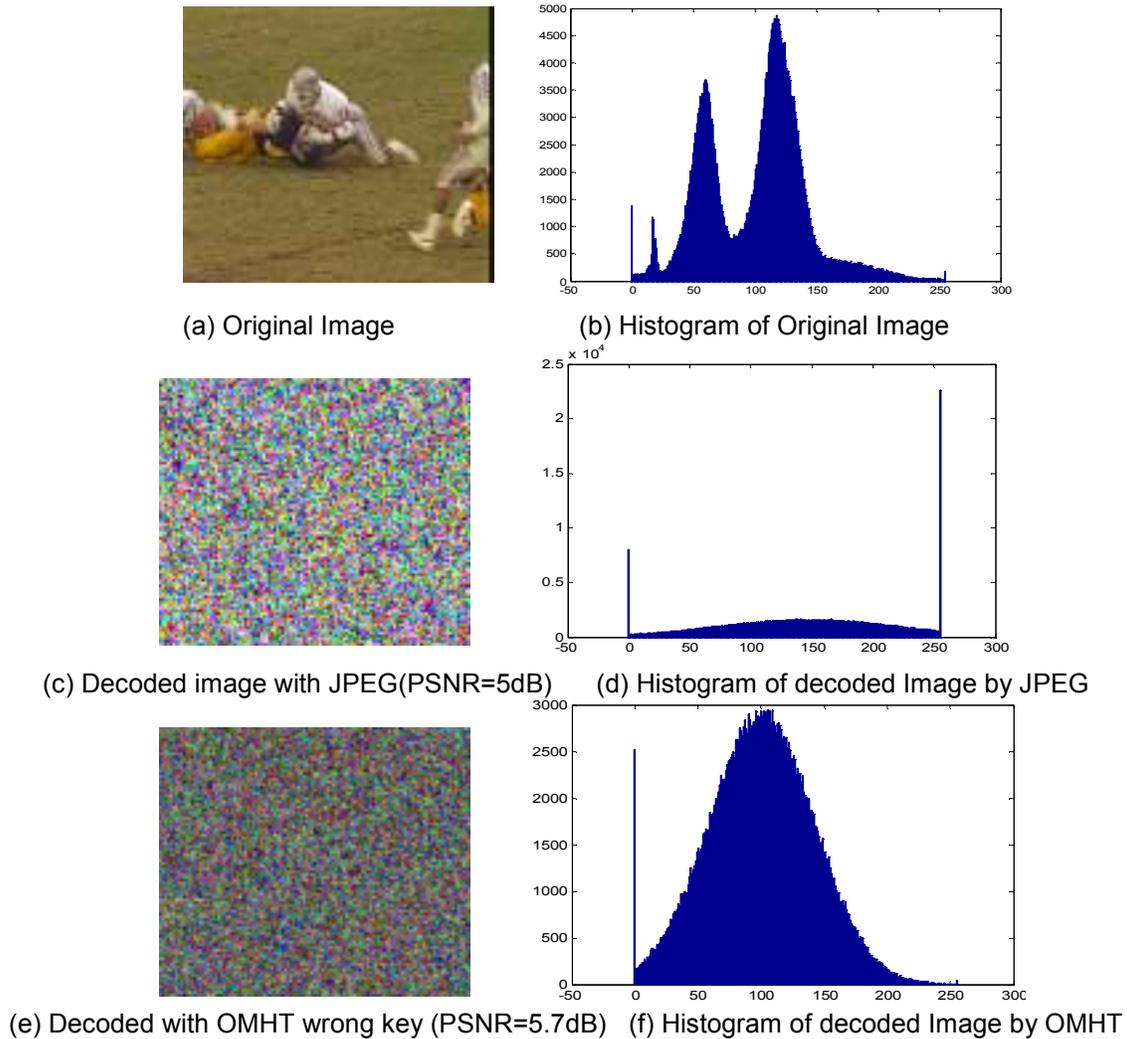


FIGURE 11: The Effect of Decoding Football Image without the Secret Order

Fig.12 shows the perceptual performance comparison between OMHT and other different encryption techniques used to encrypt Lena image. Fig.12(a) shows the original Lena image, Fig.12(b) shows the decoded image that was encrypted by OMHT, Fig.12(c) shows the decoded image that was encrypted by building a three level pyramid and encrypting the lowest resolution plus the first residual (HP Mode 30% encryption), Fig.12 (d) shows the decoded image that was encrypted by encrypting only the DC coefficients with the first AC coefficient of each block (SS Mode 30%), Fig.12 (e) shows the decoded image that was encrypted by scrambling the DC coefficients and one bitplane or three bitplanes (MM Mode 30%), Fig.12(f) shows the decoded image that was encrypted by encrypting the most significant bits of all coefficients (SA Mode 30%), Figs.2.band3.b clearly show that there can be still information left in the unencrypted parts of the data after selective encryption has been applied, Fig.12 (g) shows the decoded image that was Encrypted by Run-length, Fig.12 (h) Encrypted by sign bit encryption, Fig.12 (i) shows the decoded image that was Encrypted by band permutation(10 bands), Fig.12(j) shows the decoded image that was encrypted by bitplane permutation(n=6), Fig.12(k) shows the decoded image that was encrypted by bitplane permutation (n=7), and Fig.12(l) shows the decoded image that was encrypted by MHT. It is obvious that the PSNR of the decoded image that is encrypted by OMHT

is smaller than it is in all other techniques. So, the perceptual security strength of the OMHT technique is higher than other techniques.

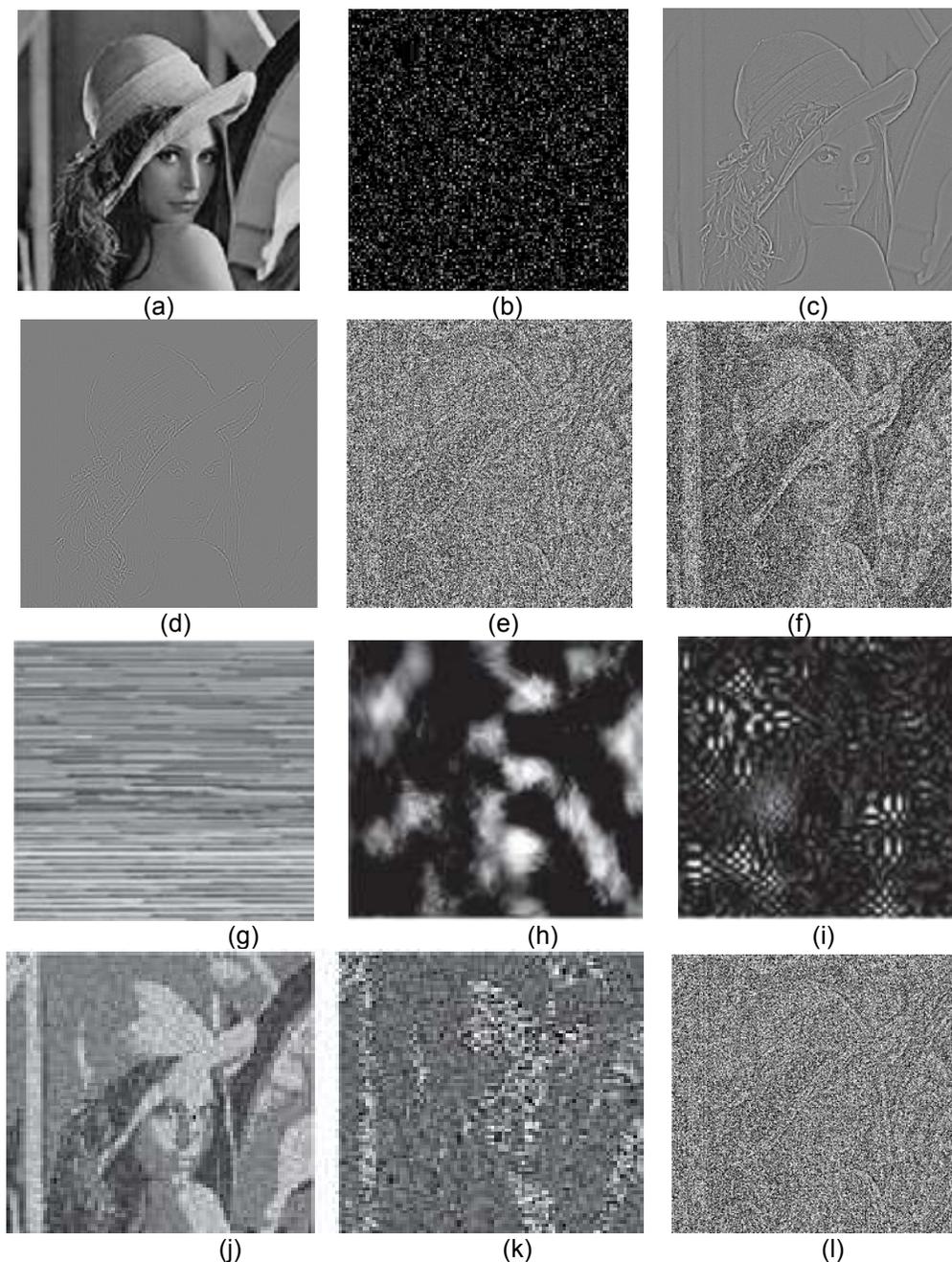


FIGURE 12: The Effect of Decoding Lena Image Encrypted with Different Techniques by JPEG: (a) Original Image, (b) Encrypted by OMHT(PSNR=4.8 dB), (c) Encrypted by HP Mode (PSNR=14.7 dB), (d) Encrypted by SS Mode(PSNR=14.2 dB), (e) Encrypted by MM Mode (PSNR=6.2 dB), (f) Encrypted by SA Mode(PSNR=6.4 dB), (g) Encrypted by Run-length (PSNR=6.5 dB), (h) Encrypted by sign bit encryption (PSNR=6.1 dB), (i) Encrypted by band permutation(10 bands) (PSNR=7.23 dB) (j) Encrypted by bitplane permutation (n=6) (PSNR=13.8 dB), (k) Encrypted by bitplane permutation (n=7) (PSNR=9.18 dB), (l) Encrypted by MHT (PSNR=6.4 dB),

5. COMPUTATIONAL COST ANALYSIS

The evaluation of the computational speed of ciphers usually consists of the analysis of the key-setup cost, the encryption cost and the decryption cost [16]. The encryption and the decryption costs are usually similar, and they are more important than the key-setup cost because one single key-setup can often be followed by thousands of encryption/decryption operations. In the following, we analyze these costs of our OMHT encryption scheme, and compare them with those of MHT and modern ciphers.

a) Key-Setup cost: The key-setup process includes all the computation and memory allocation operations prior to actual encryption of the first bit in the plaintext. The computational cost of OMHT key-setup is dominated by the construction of optimized multiple Huffman tables, generation of the secret order by which those tables are used, and comparing the test image with datasets. OMHT takes about 10 operation per table generation, single operation for secret key generation, and L operation for comparison. The total number of operations equal $10XMXL+1+L$, where L , M is number of datasets and number of subsets respectively. For $L=4$, $M=20$, the net Key-Setup cost =805 operations. For MHT technique it takes 20 operations per table entry, the total cost would be $20 \times t \times m$, where t and m are the table size and the number of selected tables, respectively. For the example of JPEG dc coefficient encryption as shown in the previous subsection, the key-setup cost would be around 2000 operations ($t=13$ and $m=8$). Compared with the ciphers listed in Table 6, the key-setup cost of OMHT encryption is much smaller than MHT and other ciphers.

b) Encryption/Decryption cost: The net computational cost of the OMHT is the same as the basic MHT-encryption scheme [1] is less than one CPU operation per encrypted bit as explained below. When a symbol is to be encoded with a normal Huffman coder, the shift amount is added to the base address of the table to obtain the address of the desired Huffman code. This process is illustrated in Fig.13 (a). In the basic MHT system, we store the base addresses of the tables in a cyclic queue according to the order that they are used. When a symbol is to be encoded/encrypted, the base address is first loaded from the memory, and then the shift-amount is added to it. Afterwards, the index to the cyclic queue of base addresses should be increased by one. Then, the index should be compared with the end of the queue in order to decide whether it should be reset to the beginning of the queue. Therefore, the computational difference between our cipher/encoder and a normal Huffman coder is one memory-load, one addition and one comparison operation for each symbol encoded. The encoding process of the proposed cipher/encoder is shown in Fig.13 (b). Since each symbol in the original data usually corresponds to more than 3 bits in the Huffman bitstream, then encryption cost of our algorithm is less than one CPU operation per encrypted bit, which is around 20 times smaller than the well-known AES as listed in Table 6.

Recently, a new cryptographic cipher named COS [18] with a very fast speed is gaining popularity. It is around 4–5 times faster than AES. Compared to COS, the encryption cost of OMHT is still several times smaller.

Cipher Type	Key-setup Cost (CPU instructions)	Encryption Cost (CPU instructions/bit)
MARS	9416	25
RC6	10372	22
Rijndael	35484	20
Serpent	26308	28
Twofish	37692	20

TABLE 6: Computational Costs of AES Finalists on a Pentium-MMX Machine. The Figures in This Table are Translated from [17] by Assuming Two CPU Instructions are Executed in Every Clock Cycle in a Pentium-MMX CPU

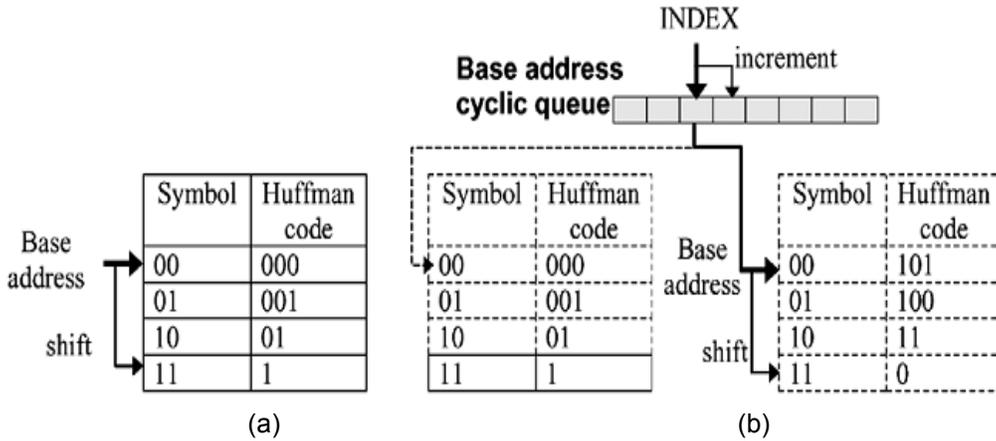


FIGURE 13: (a) Normal Huffman Coder Adds the Shift Amount to the Base address of the Table to Obtain the Address of the Desired Huffman Code. (b) OMHT Loads the Base Addresses of Huffman Tables from a Cyclic Queue, and the Index to the Queue is Increased by One After Coding of Each Symbol.

6. CONCLUSIONS

The experiments' results reveal that the proposed OMHT technique achieves better compression and security performance than that of MHT, and JPEG Image Compression Standard especially at low bitrate. The OMHT scheme provides

- **High security:** resistance against various types of attacks, including the ciphertext-only attack and the known/chosen plaintext attack[19].
- **Low encryption cost:** the encryption cost not exceed very small portion of the total computation cost of compression
- **No harm to the compression ratio:** The increase of the final bit stream size due to encryption is not higher than 0.5% of the original coded bitstream.
- Joint compression-encryption OMHT technique achieves both high security and compression performance in one single step, which simplifies the system design and reduces time required to perform compression followed by encryption.
- Since images have different statistics, using the same fixed JPEG standard predefined coding tables as suggested in MHT technique will not be effective in encoding all image and video types.

- The OMHT method obtains better performance in terms of storage space use and more stable peak signal to noise ratio than that of JPEG in encoding an image with small and great gray-level variations among adjacent pixels.
- Receivers haven't the secret order cannot decode the encoded images successfully.
- Further, the proposed new compression-encryption technique could be applied on any source data, not only images, which uses Huffman coding to achieve better compression ratio. Therefore, the proposed technique will be suitable for compression of text, image, and video files.

7. REFERENCES

1. C.-P. Wu and C.-C. J. K. Kuo. "Design of integrated multimedia compression and encryption systems". IEEE Transactions in Multimedia, vol. 7, no. 5, pp. 828–839, 2005.
2. W. Stallings. "Cryptography and Network Security Principles and Practices", Upper Saddle River, NJ: Prentice Hall, 2003.
3. M. Van Droogenbroeck and R. Benedett. "Techniques for a selective encryption of uncompressed and compressed images". In Proceedings of Advanced Concepts for Intelligent Vision Systems (ACIVS '02), pp. 90–97, Ghent, Belgium, September 2002.
4. L. Qiao, K. Nahrstedt, and M.-C. Tam. "Is MPEG encryption by using random list instead of zigzag order secure?". in Proceedings of the IEEE International Symposium on Consumer Electronics (ISCE '97), pp. 226–229, Singapore, December 1997.
5. T. Uehara and R. Safavi-Naini. "Chosen DCT coefficients attack on MPEG encryption scheme". in Proceedings of IEEE Pacific Rim Conference on Multimedia, pp. 316–319, Sydney, Australia, December 2000.
6. H. Cheng and X. Li. "Partial encryption of compressed images and videos". IEEE Transactions on Signal Processing, vol. 48, no. 8, pp. 2439–2451, 2000.
7. C.-P. Wu and C.-C. Kuo. "Efficient multimedia encryption via entropy codec design". Proc. SPIE, vol. 4314, Jan. 2001.
8. D. Xie and C. J. Kuo. "Enhanced Multiple Huffman Table (MHT) Encryption Scheme Using Key Hopping". In Proceedings of IEEE International Symposium on Circuits and Systems, pp.568–571, May2004.
9. D. Xie and C. J. Kuo. "Multimedia Data Encryption via Random Rotation in Partitioned Bit Stream". In Proceedings of IEEE International Symposium on Circuits and Systems, pp.568–571, May2004.
10. D. W. Gillman and R. L. Rivest. "On breaking a Huffman code". IEEE Transactions on Information Theory, vol. 42, no. 3, pp. 972–976, 1996.
11. J. Zhou, Z. Liang, Y. Chen, and O. C. Au. "Security analysis of multimedia encryption schemes based on multiple Huffman table". IEEE Signal Processing Letters, vol. 14, no. 3, pp. 201–204, 2007.
12. W. Pennebaker and J. Mitchell. "JPEG Still Image Data Compression Standard", Van Nostrand Reinhold, New York, 1993.
13. <http://www.jpeg.org> (JPEG resources) [accessed at 4/8/2010]

14. <http://www.jpeg.org/public/jfif.pdf> (JPEG file interchange format) [accessed at 8/8/2010]
15. (independent JPEG group) <ftp.uu.net:/graphics/jpeg> [accessed at 8/8/2010]
16. C.-P. Wu and C.-C.J. Kuo. "*Efficient multimedia encryption via entropy codec design*". In Proc. SPIE Int. Symp. Electronic Imaging 2001, vol. 4314, Jan. 2001, p.128.
17. J. Nechvatal et al. "*Report on the Development of the Advanced Encryption Standard*". National Institute of Standards and Technology, U.S. Dept. Commerce, Tech. Rep., Oct. 2000.
18. E. Filiol and C. Fontain. "*A new ultra fast stream cipher design: COS ciphers*". In Proc. 8th IMA Conf. Cryptography and Coding, Dec. 2001.
19. Shaimaa A. El-said, Khalid F. A. Hussein, and Mohamed M. Fouad. "*Securing Multimedia Transmission Using Multiple Huffman Tables Technique*". Electrical and Computer Systems Engineering Conference (ECSE'10), Egypt, 2010.

Building a Distributed Secure System on Multi-Agent Platform Depending on the Principles of Software Engineering Life Cycle

Dr.Ghossoon M.Waleed Al-Saadoon

ghowaleed2004@yahoo.com

*College of Administrative Sciences,
Applied Science University
Kingdom of Bahrain ,Manama , Jufair, P.O.Box:5055
Tel : +(973) 17728777- 149, Fax: +(973)17728915*

Abstract

Today, applications in mobile multi-agent systems require a high degree of confidence that running code inside the system will not be malicious. Also any malicious agents must be identified and contained. Since the inception of mobile agents, the intruder has been addressed using a multitude of techniques, but many of these implementations have only addressed concerns from the position of either the platform or the agents. Very few approaches have undertaken the problem of mobile agent security from both perspectives simultaneously. Furthermore, no middleware exists to facilitate provisioning of the required security qualities of mobile agent software while extensively focusing on easing the software development burden. The aim is to build a distributed secure system using multi-agents by applying the principles of software engineering. The objectives of this paper is to introduce multi agent systems that enhance security rules through the access right to building a distributed secure system integrating with principles of software engineering system life cycle, as well as satisfy the security access right for both platform and agents to improve the three characteristics of agents adaptively, mobility and flexibility, which is the main problem that depending on the principles of software engineering life cycle.

There are 3 characteristics that satisfied using agent; mobility, adaptively and flexibility. Adaptively (which is the capability to respond to other agencies and/or environment to some degree). Mobility (the ability to transport itself from one environment to another) and Flexibility (can be defined to include the following properties; responsive, pro-active and social). This project based on the platform of PHP and MYSQL (Database) which can be presented in a website. The implementation and test are applied in both Linux and Windows platforms, including Linux Red Hat 8, Linux Ubuntu 6.06 LTS and Microsoft Windows XP Professional. Since PHP and MySQL are available in almost all operating systems, the result could be tested the platform as long as PHP and MySQL configuration is available. PHP5 and the MySQL (database) software are used to build a secure website. Multiple techniques of security and authentications have been used by multi-agents system. Secure database is encrypted by using md5. Also satisfy the characteristics for security requirements: confidentiality (protection from disclosure to unauthorized persons), integrity (maintaining data consistency) and authentication (assurance of identity of person or originator of data).

Keywords: Distributed Secure System, Platform, Multi Agent System, Access write, Lunix Read Hat , and DM5.

1. INTRODUCTION

Mobile agent technology offers a new computing paradigm in which a software agent can suspend its execution on a host computer, transfer itself to another agent-enabled host on the network, and resume execution on the new host. The area of mobile agent security is in a state of immaturity [1]. Numerous techniques exist to provide security for mobile agents, there is not at present an overall framework that integrates compatible techniques into an effective security model. The traditional host orientation toward security persists and focuses of protection mechanisms within the mobile agent paradigm remains on protecting the agent platform. However, emphasis is slowly moving toward developing techniques that are oriented toward protecting the agent, a much more difficult problem. Fortunately, there are many applications where conventional and emerging security techniques should prove adequate, if applied judiciously. The software was building the new platform using multi-agent system. The Unified Modeling Language (UML) used to build this prototype Model [2]. To make the software easier and systematic, the software engineer must incorporate a development strategy that encompasses the process, method and tool layers. This strategy is called Software Engineering Paradigm to develop Process Model. This paper reviews a web based system and the Prototype Model algorithm used for the system design [3, 4] as in Figure 1.

- The software designed a set of objectives to the users.[5,6].
- The software determines the requirements, and
- The user can review the existing software anytime.



FIGURE1: Software Engineering to Development Process Model

2. LITERATURE REVIEW

There are many literature sources that encourage over this paper. It has been significantly too interpreted about agents and in depth security issues itself that elaborates. These are related sources of the issue in the mobile agents. **Adam Pridgen & Christine Julien ,2006**, they introduce a mobile agent system that enhances security functionality by integrating core software and hardware assurance qualities, as well as addressing security concerns from the perspectives of both the platform and the agent [7]. **Loulou; Mohamed Jmaiel; Ahmed Hadj Kacem and Mohamed Mosbah,2006**, In order to facilitate analysis, design and specification of mobile agent systems, the possible attacks that may occur in a mobile agent system, they associate the specification of the basic concepts that ensuring security such as: agent authenticity, authority access, security policy and its various kind [8]. **Robert S. Gray, George Cybenko, David Kotz, Ronald A. Peterson and Daniela Rus ,2001**, the mobile agent systems involved the relocation of both code and state information. The area of mobile agent security is in a state of immaturity. While numerous techniques exist to provide security for mobile agents,

there is not at present an overall framework that integrates compatible techniques into an effective security model [9].

3. METHODOLOGY

System development methodology is a necessary process to develop software .The methodology consists of three main parts to build distributed Secure System using characteristics on multi agent system for this platform and depending on the principles of software engineering life cycle. From UML methodology, the software designer will able to identify the tasks on software developments to present the software architecture and the description of objects and their interactions with one another, as in Figure 2.The platform assigns a newly originated or incoming agent to a requested location or place, where it can compute and interact with other agents. Besides furnishing the engine on which an agent executes its code, typical services offered by an agent platform include the capability for an agent to clone itself, spawn or create new agents, terminate any spawned agents, locate other agents at the platform or a platform elsewhere, send messages to other agents, and relocate itself on another platform, all these process under the security rules and privilege from the management server – web server.

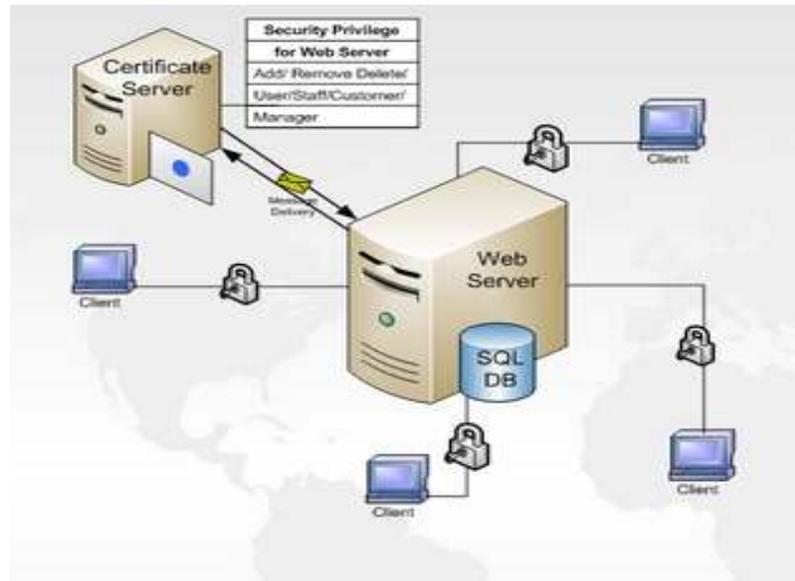


FIGURE 2: Security Platform of Server-Client Architecture

3.1 System Design

In this project has been assumed that client-server websites are used. The security levels have the rules and permission to access each data of clients-server.

The methodology will be divided into four main phases; Preliminary Requirements, Final Requirements, Analysis and finally Design. The 1st stage for the preliminary requirements includes activities which are defining the requirements, validate user requirements, define consensual requirements, establish keywords set and extract limits and constraints.

The 2nd phase in final requirements include some tasks for instance, characterize environment. The process is to determine the entities, define context and characterize environment. The determination is the use cases; it will draw up inventory of the use cases, identify cooperation failures and elaborate sequence diagrams. Follow up then is elaborate User Interface UI prototypes and validate it.

The 3rd phase is the analysis process to identify classes, study interclass relationship and construct the preliminary diagram to verify the global and local levels of mobile agents adequacy. These processes are to know-how study the entities in the domain context and determine agents between entities. All fields of study include the active-passive relationships, active entities relationships and agents relationship.

The final phase is the **design of the architecture and multi-agent mode** that determine packages, classes, design-patterns and elaborate component and class diagrams. **Figure 3** shows the principles of software engineering applied in the security management life cycle,



FIGURE 3 :software engineering applied in the security management life cycle

3.2 System Requirements

Web applications run in two locations: the server and the client. This means that both locations need to be developed to provide the best security for the user. The server needs to be developed in such a way that information being stored is not compromised; while the client needs to be developed to present and retrieved only the required information. A client that divulges too much information is not likely to be secure.

The server is where all the application's action is taken place. The PHP operates on a transitive level for the web page between the client and the server. Figure 4 shows the separation of the client and server.

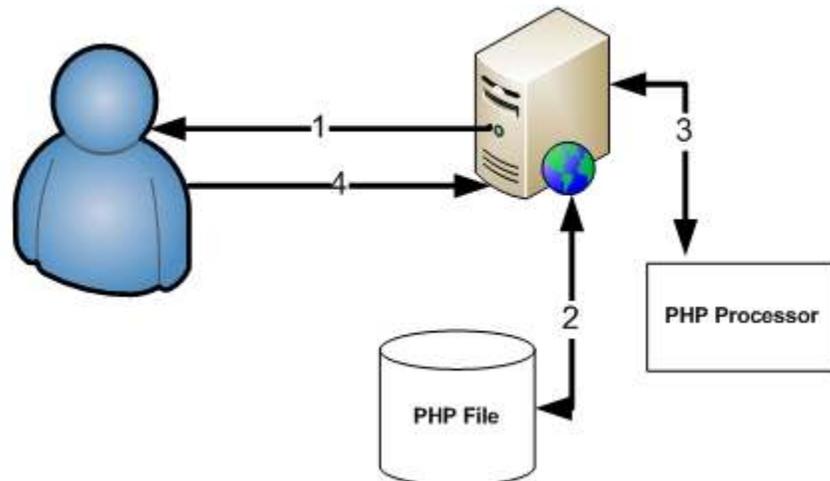


FIGURE 4: PHP Process for Distributed System

3.3 Security Analysis Phase

This phase includes security in PHP and PHP security audit.

I. Security in PHP : it includes these requirements.

- 1.Security is not an absolute.
- 2.Can always be more or less secure
- 3.Security is difficult to measure.
- 4.Security must be balanced with usability.
- 5.Security must be balanced with expense.
- 6.Security must be part of the design.
- 7.The filter input for the most input is obvious - form data (`$_GET` and `$_POST`), cookies (`$_COOKIE`), RSS feeds, etc. While the output Escaping is the process by which user escape any character that has a special meaning in a remote system. Unless user sending data somewhere unusual, there is probably a function that performs the escaping for. The two most common destinations are the client (use `htmlspecialchars()`) and MySQL (use `mysql_real_escape_string()`). If it must write down, make sure that it is exhaustive - find a reliable and complete list of all special characters.

II. PHP Security Audit

An audit is an examination and it does nothing should be off-limits. A PHP security audit primarily involves an examination of the source code. Other points of interest are software design, PHP configuration, and infrastructure security.

A. Setting the Bar

- How much security is needed?
- Start with a minimum level, and go from there. At the very least, a PHP application should have filter input and escape output.
- If a PHP application can't meet these minimum guidelines, it isn't worth a time.

B. Analyzing the Configuration

- The configuration of PHP is mostly dictated by `php.ini`.
- However, remember that PHP configuration directives can be modified in other places - `httpd.conf`, `.htaccess`, `ini_set()`.
- Things to avoid: `register_globals = On`, `allow_url_fopen = On`, `display_errors = On`, `magic_quotes_gpc = On`

C. Analyzing the Design

- Have the design explained it first. No one knows an application as well as the developers.
- A poor or unnecessarily complex design is a security risk.
- Is tracking data difficult?
- Is distinguishing between filtered and tainted data difficult?
- Stream-of-consciousness is why so many PHP applications are insecure. This is why so many PHP applications are insecure.

D. Searching the Source: Input

E. Searching the Source: Output

F. Searching the Source: Potential Problems

- Check for dynamic includes that use tainted data: include, require.
- Check for client-side restrictions: maxlength, radio, checkbox, select, Client-side filtering.

G. Searching the Source: Bad Habits

- Error suppression :@
- Misguided trust of HTTP request
- Headers: Referer, Host
- Unescaping: stripslashes(), etc.

The most common mistakes are sending tainted, unescaped data to the client or a database. There are others that most website frequently used: storing the authorization level in a cookie; passing the authorization level in the URL; storing the username and password in a cookie; and storing includes within document root.

3.4 COMMUNICATION AND ANALYSIS PHASE

First, all relation information, such as hotel management information, are collected. A discussion is completed with staff members and management and the requirements of system are identified:

- To computerize the current hotel management system.
- To enable the customer to get information about the hotel.
- To enable management to view which staff is members conduct the process with the customers.
- Accessibility to the program must be controlled. The users can't access to the application without authority.
- All levels of users have their own username and password.
- System administrator has the authority to add new user account.
- Staff members have the authority to add new customer records.

3.5 Quick Plan Phase

The system will display the login form when the web page is first accessed. The user needs to enter their username and password. If the username and password is matched with the login information stored in the database, the login is passed. Then the system will check the user level of the login user. There are four user levels in the system:

Administrator, Staff, Manager and Customer.

Each user level has different functionality. Users cannot access additional functionality without authority, as in Table 1.

Administrator	Staff	Manager	Customer
View, add or delete announcement.	View and add announcement.	View and add announcement.	View Result
View all the activities among other users.	Add, delete and update customers.	View customer results.	Change password
Add or remove system user.	View customer results.	Add, delete and update staff.	
View customer result.	Change password.	Change password.	

TABLE 1: privilege for each security level

4. SECURITY FOR SYSTEM DESIGN

First the websites creates specific task with access right for Administration ID to display User Login, Delete, Add new Group, and Add new User. The mobile agents are the interpreted language. The language has to be interpreted, because moving of a running object requires access to the global variables, or better to the current execution pointer and stack. The best way to achieve this is using a virtual machine which executes the interpreter language.

4.1 Security Agent Management

The security agent management architecture is used to design the heterogeneous Database Networks. The main activities of the administrator agent are the following:

- Roles Agent for creating the privileges for security and access control list.
- Creates administrator and local servers.
- Determines the group agents.
- Creating User that can deal in this platform.

In multi agents system based security agent management architecture, two main functionality agents are recognized: Global agents and Local agents.

The management system in this project is districted from public access where the registered users have authority to login the system. There are four user levels in this application: Administrator, Staff, Manager and Customer. Each user has own name and password.

The login process is only performed if the textbox for username and password is filled. Otherwise, a message will appear to ask the user to complete the login form. After the user completes the form, the system will check the login information. If login is successful, the page will redirect to the user home page. If login fails, a message will appear to prompt the user to enter the correct username and password.

Accessibility user level: the Administrator and Staff. If the login is accepted, the page will be redirected to the user home page, where all information posted by Administrator and Staff will be displayed. The user menu will be displayed in the left side of each page after the successful login. The announcement author is the same as the name of the logged in user. The date of the announcement is the current date of the server PC when the announcement is submitted.

Administrator is the only user able to delete announcements. In Figure.6, to delete announcements, simply choose "Edit Announcement" and click the "Delete" link that appear in the bottom of the desired announcement.

Only the Administrator can add a new system user (Staff) to control for security of the system. The user is only added if all textboxes filled and the requested username does not exist in the record. The Administrator will need to create an account for Staff. Then the Staff can add customer information (users) into the database. Password for new user is initially encrypted. The user can change the user password. The Administrator can also delete the user account in case user information is incorrect or user has resigned or write some comments about security encryption . Accessibility user level: System Administrator, Staff and Customer. The user needs to enter their current password and new password twice. The password will be changed if :

4.2 Security Level Local Agents

The security of an administrator constitute a sub-set of hosts in a local network. It is composed of a group of Local Agent (LA)s, which have specific functions. One can distinguish two kinds of LA:

- Intranet LA several Intranet Agents. The intranet agents manage the security of a local network. It controls LA s and analyzes the auditing events reported by these agents.
- Internet LA. In each level, notes agents communicate and exchange their information of heterogeneous DBs and analysis for detecting intrusive activities in a cooperative manner.

4.3 ROLES AGENT

The access key attribute of an agent is that it is able to act autonomously. Agents can then take on a wide range of responsibilities on behalf of users or other system entities including services and entering into agreements. Additionally, an agent will often perform some tasks on behalf of another entity. For example, (a software agent could perform a task on behalf of a person). It could also perform on behalf of another piece of software (another agent), an organization, or a particular role (manager, system administrator).

5. CONCLUSION

The software project was tested in both Linux and Windows platform, including Linux Red Hat 8, Linux Ubuntu 6.06 LTS and Microsoft Windows XP Professional. Since PHP and MySQL are available in almost all operating system,.

- The security platform used to verify and validate multi-agents build the access write. The security platform using mobile agents can satisfied the three characteristics (adaptively, mobility and flexibility).The adaptively can be modified and updated the rules and privileges to the system. While mobility, are the agents itself can be transport from one another to the others platform.
- The configuration of PHP, MySQL and also Apache web server is quite different in different platforms. In Windows XP, the installation of AppServ will install PHP, MySQL, Apache, phpMyAdmin at once. The configuration is done automatically. The PC can run as web server after reboot. Anyway, the firewall of Windows XP need to turn off, otherwise the client PC cannot access the web site hosted by server PC.
- In Linux platform, PHP, MySQL and Apache web server normally is ready and installed. If not, we can manually install them from software package. The directory of the web is located in /vary/www/html. The services of apache and MySQL maybe not start automatically after login to Linux platform depend on the system setting. If the services don't start, we can type a code in terminal window to manually start the services.

This system has the following strengthens, which will increase the commercialization potential:

- It can be run in many platforms such as Microsoft Windows, UNIX and Linux.
- Easy to configure and install in web server.
- The PHP and MySQL are free and open sourcing software's , so very easy to install and used in this project.

6. REFERENCES

1. Dell'Acqua, P., M. Engberg, L.M. Pereira, " *An Architecture for a Rational Reactive Agent*", [online] available at: <http://centria.di.fct.unl.pt/~imp/publications/online-papers/epia03-agent.pdf>, 2003.
2. Marik, V., O. Stepankova, H. Krautwurmova, M. Luck., " *Multi agent systems and applications II*", Springer-Verlag Berlin Heidelberg, 2002.
3. K. Rabuzin, M. Malekovic, Miroslav Baca (2006), " *A SURVEY OF THE PROPERTIES OF AGENTS*" .pdf, University of Zagreb, Faculty of Organization and Informatics, Varaždin
4. " *An introduction to Agents, Todd Sundsted*" , JavaWorld.com, 06/01/98, [online] available at: <http://www.javaworld.com/javaworld/jw-06-1998/jw-06-howto.html>
5. Geppert, A., M. Kradolfer, D.Tombros: " *Realization of Cooperative Agents Using an Active Object-Oriented Database Management System*", Proc. 2nd Workshop on Rules in Databases (RIDS), Athens, Greece, 1995.
6. Hayes-Roth, B. " *An architecture for adaptive intelligent systems*, Artificial Intelligence Vol. 72, 1995.
7. A. Pridgen and C. Julien, " *A Secure Modular Mobile Agent System*", The University of Texas at Austin, TR-UTEDGE-2006-003.
8. M. Loulou, M. Jmaiel, A.Hadj Kacem and M.Mosbah, " *A Conceptual Model for Secure Mobile Agent Systems*", computational intelligence and security, Proceedings of the 3rd international conference on , Nov 3-6/2006. [online] available at: <http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=4072023>.
9. Robert S. Gray, G. Cybenko, D. Kotz, Ronald A. Peterson and Daniela Rus , " *D'Agents: Applications and Performance of a Mobile-Agent System*", Thayer School of Engineering / Department of Computer Science, Dartmouth College, November 28, 2001.

Performance Variation of LMS And Its Different Variants

Sanjeev Dhull

*Asst.Professor/ECE
G.J.U.S&T,Hisar
Hisar,125001 , India*

sanjeev_dhull_ap@yahoo.co.in

Dr.Sandeep Arya

*Chairman/Ece
G.J.U.S&T,Hisar
Hisar,125001 , India*

Dr.O.P Sahu

*Associate Professor/ECE
N.I,T Kurukshetra
kurukshetra, 136118, India*

Abstract

Acoustic echo cancellation is an essential and important requirement for various applications such as, telecasting, hands-free telephony and video-conferencing. Echo cancellers are required because of loud-speaker signals are picked up by a microphone and are fed back to the correspondent, resulting in an undesired echo. These days, adaptive filtering methods are used to cancel the affect of these echoes. Different variants of LMS adaptive algorithms have been. Implemented and they are compared based upon their performance according to the choice of step size.

Keywords: Echo, Algorithm, Adaptive, LMS, .

1. INTRODUCTION

Adaptive filters appear in many signal processing and communication systems for applications such as channel equalization, echo cancellation, noise reduction, radar and sonar signal processing, beam-forming,. Adaptive filters work on the principle of minimizing an error function, generally the mean squared difference (or error), between the filter output signal and a target (or desired) signal. Adaptive filters are used for estimation and identification of non-stationary signals, channels and systems. LMS algorithm and RLS and their variant are used to solve the problem. In today's scenario most of the systems are hands free, examples of these systems are hands-free telephones and video-conferencing these systems provide a comfortable and efficient way of communication. There is a major problem with these systems, signal degradation occurs when loudspeaker signals are picked up by a micro-phone and are sent back for processing. Therefore an undesired echo came in picture.In such hands-free systems, acoustic echo cancellers are necessary for full-duplex communication. Conventional techniques used in classical telephony such as clipping and voice controlled switching [1] have limited performance. More advanced adaptive filtering technique are expected to provide a better signal quality.

2. ADAPTIVE FILTERING TECHNIQUES

As Adaptive filters have applications in various applications such as identification, Acoustic echoes cancellation & inverse modeling, in this paper echo cancellation application is considered. Acoustic echoes are suppressed with the help of adaptive filtering techniques [2]. This algorithm basically adapts to a solution minimizing the mean-square error. It is based on the steepest-descent method. How filters weights are adapted, it is shown in fig an adaptive filter converges to an estimate of the impulse response of the acoustic path [3]. Of all existing adaptive algorithms the Least Mean algorithm is the best known. An FIR or IIR filter [6] is updated iteratively.

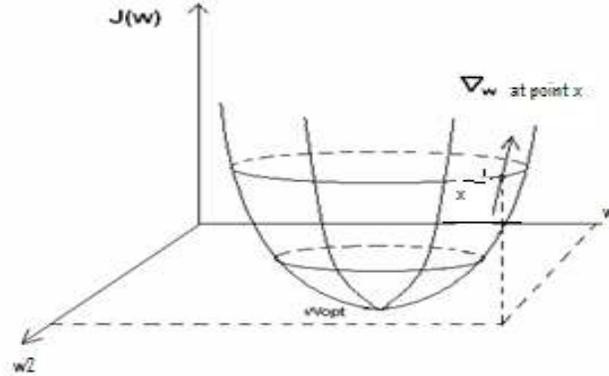


FIGURE1:Steepest Descent Method

Following equations explain that how filter weights are updated and error is minimized

$$W(n+1) = w(n) + 2\mu x(n)[d(n) - x^T(n)w(n)]$$

$$= w(n) + 2\mu x(n)[d(n) - w^T(n)x(n)]$$

$$= w(n) + 2\mu e(n)x(n)$$

Here $y(n) = w^T(n)x(n)$ is filter output

And $e(n) = d(n) - y(n)$ is error signal

$$W(n) = [w_0(n) \ w_1(n) \ \dots \ w_{M-1}(n)]^T \quad \text{Are filter taps which updated to find the minima?}$$

$$e(n+1) = y(n+1) - h^T(n+1)x(n+1)$$

LMS algorithm is a very simple and requires only $O(2Nz)$ multiplications and $O(2N)$. Another variant of LMS is NLMS.[6] The motivation of this algorithm is that the power of the input signal varies with time, so the step size between two adjacent filter coefficients will vary as well, then also the convergence speed. The convergence speed will slow down with small signals, and for the loud ones the over-shoot error would increase. So the idea is to continuously adjust the step size parameter with the input power. Therefore, the step size is normalized by the current input power, resulting in the Normalized Least Mean Square algorithm. The Normalized Least Mean Square (NLMS)[15] algorithm is a modified version of the LMS algorithm. In the LMS algorithm, the correction factor to the tap weight vector $W(n)$ is computed as $\mu U(n)e(n)$. Since this quantity is directly proportional to the tap input vector $U(n)$, the error in the gradient estimate gets magnified for large $U(n)$. This problem can be avoided by scaling the correction factor by the squared Euclidean norm of the tap input vector $U(n)$ (the average power of the input signal). This variant of the LMS algorithm, with the normalized correction factor, is called the Normalized LMS (NLMS) algorithm. The LMS [6][11][13] and their different variants can be driven using the following functions

Let us define an error signal $e(n+1)$ at time $n+1$ as

$$e(n+1) = y(n+1) - \hat{y}(n+1)$$

Here $y(n+1) = h^T x(n+1)$ is the output of a system and $h_t = [h_{t,0} \ h_{t,1} \ \dots \ h_{t,L-1}]^T$ Are responses of system
 And $\hat{y}(n+1) = h^T(n)x(n+1)$ is the model filter output and $h(n) = [h_0(n) \ h_1(n) \ \dots \ h_{L-1}(n)]^T$ is the model filter. One easy way to find adaptive algorithms that adjust the new weight vector $h(n+1)$ from the old one $h(n)$ is to minimize the following function $J[h(n+1)] = d[h(n+1), h(n)] + \eta \varepsilon^2(n+1)$

Here value of η plays an important role in updating the coefficients values. If η is very small that the algorithm makes very small updates. On the other hand, if η is very large, the minimization of $J[h(n+1)]$ is almost equivalent to minimizing $d[h(n+1)]$, Hence, the different weight coefficients $h_l(n+1)$, $l = 0, 1, \dots, L-1$, are found by solving the following equations:

$$\frac{\partial d[h(n+1), h(n)]}{\partial h_l(n+1)} - 2\eta x(n+1)\varepsilon(n+1) = 0$$

if the new weight vector $h(n+1)$ is close to the old weight vector $h(n)$, replacing the a *posteriori* error signal with the a *priori* error signal $e(n+1)$ is a reasonable approximation and equation. $\frac{\partial d[h(n+1), h(n)]}{\partial h_l(n+1)} - 2\eta x(n+1)e(n+1) = 0$ is much easier to solve for all distance

measures d . The LMS algorithm is easily obtained from above equation by using the squared Euclidean distance $d\varepsilon[h(n+1), h(n)] = \lVert h(n+1) - h(n) \rVert_2^2$. Using these equations and doing different mathematical operations we can find out different variants of the algorithm. Different variants of LMS [9][10] are NLMS, SIGN SIGN, SIGN DATA and SIGN ERROR. We will compare the performance of all these. Following are the comparison of different algorithms

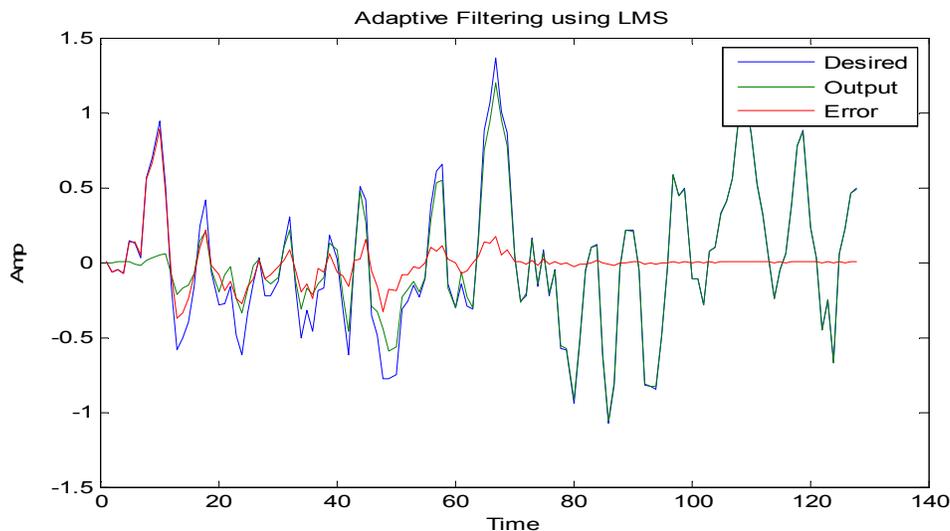


FIGURE 2 : LMS Adaptive Filtering With $\mu=0.5$

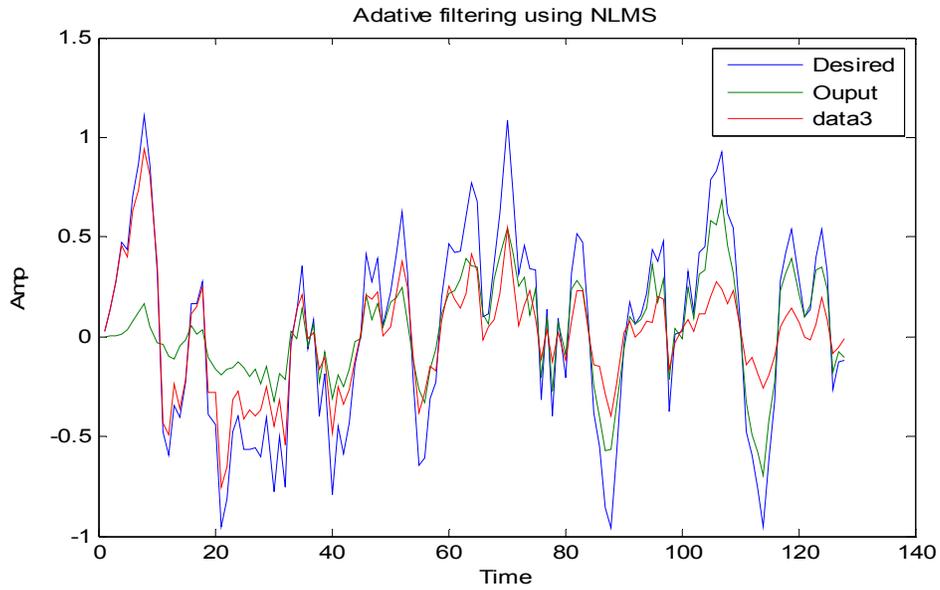


FIGURE 3: NLMS Adaptive Filtering with $\mu=.5$

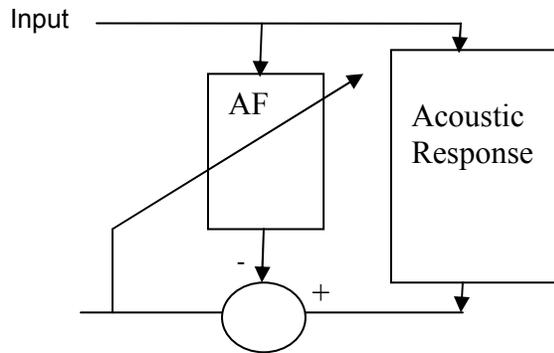


FIGURE 4: Adaptive System

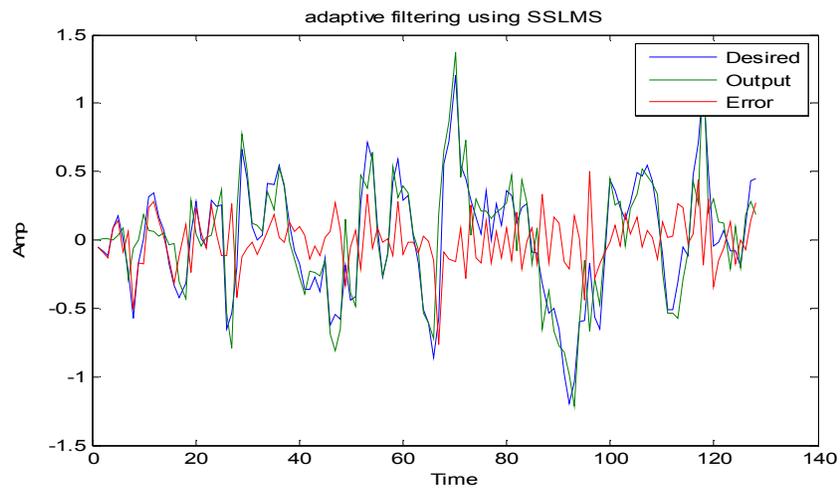


FIGURE 5: SLMS Adaptive Filtering With $\mu=.5$

All above results are for step size of .5.when we change the step size from .5 to near to 2 which is upper range of step size limit performance degrades

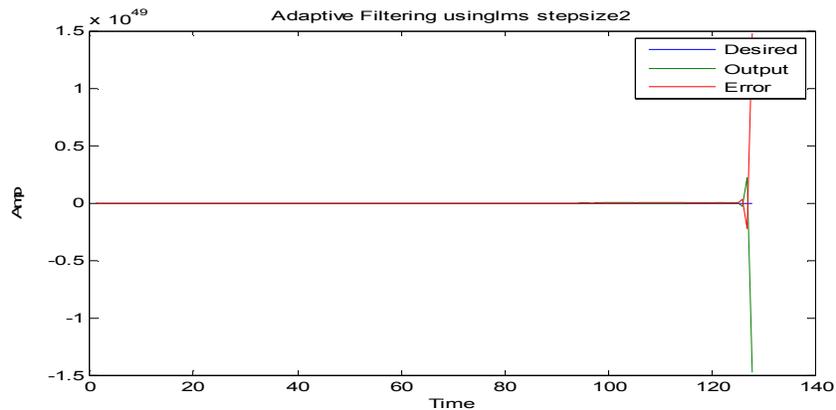


FIGURE 6: LMS Adaptive filtering with step size of 2

3. CONCLUSION

As seen from different graphs it is clear that the choice of step size between the specified range is very important .If it is too low or near to upper range convergence is poor and desired signal is not obtained as shown in different figure. Thus different variants of LMS indicate different performance properties according to the choice of step size.

4. REFERENCES

1. Haykin, S "Adaptive Filter Theory", Englewood Cliffs,N.J., Prentice-Hall, 1991.
2. Sondhi, M. M., Berkley, D.A. "Silencing Echoes on the Telephone Network", Proc. IEEE, 68(8): 948-963.,August1980,
3. Sondhi, M. M., Mitra, D. "New Results on the Performance of a Well-Known Class of Adaptive filter",IEEE, 1976, 64(11):1583-1597
4. Brehm, H., Stammers, W. "Description and generation of spherically invariant speech-model signals", Signal Processing 12(2):119–141, March 1987.
5. Moon, T. K., Stirling, W. C. "Mathematical Methods and Algorithms for Signal Processing", New Jersey: Prentice Hall, 2000.
6. Farhang-Boronjeny, B. "Adaptive Filters Theory and Application", New York: Wiley, 2000.
7. Widrow, B., Stearns, S.D. " Adaptive Signal Processing" Englewood Cliffs, NJ: Prentice. Englewood Cliffs, NJ: Prentice. Hall, 2001.
8. Mader.A., Puder, H., Schmidt, G.U. "Step-size control For acoustic cancellation filters, Signal Process". 80:1697–1719, 2000.
9. Ben Jebara, S., Besbes, H. "A variable step size filtered sign algorithm for acoustic echo cancellation", IEEEvol. 39(12):936-938, June 2003.
10. Sophocles J. Orfanidis. "Optimum Signal processing An Introduction", McGraw hill 1988.
11. Proakis, J.G., Melonakos, D.G. "Digital Signal Processing Principles, Algorithms and Applications", Prentice Hall,1996.

12. Brennan, R., Schneider. " *A flexible filter bank Structure for extensive signal manipulation in digital Hearing aids*",*Proc.IEEE Int. Symp. Circuits and Systems*, pp. 569-572,1998.
13. Vaseghi,S."*Theory and Application in Speech, Music and Communications*",*Wiley*2007.
14. Benesty,J.,Morgan, D. R.,Sondhi,,M. M. "A *Hybrid Mono/Stereo Acoustic Echo Canceller*",*IEEE Trans. Speech Audio Processing*.1997.
15. Ehtiati, Benoît Champagne,B. "*Constrained Adaptive Echo Cancellation for Discrete Multitone Systems*", *IEEE transactions on signal processing*, 57(1),JANUARY2009

CALL FOR PAPERS

Journal: International Journal of Computer Science and Security (IJCSS)

Volume: 4 **Issue:** 6

ISSN: 1985-1553

URL: <http://www.cscjournals.org/csc/description.php?JCode=IJCSS>

About IJCSS

The International Journal of Computer Science and Security (IJCSS) is a refereed online journal which is a forum for publication of current research in computer science and computer security technologies. It considers any material dealing primarily with the technological aspects of computer science and computer security. The journal is targeted to be read by academics, scholars, advanced students, practitioners, and those seeking an update on current experience and future prospects in relation to all aspects computer science in general but specific to computer security themes. Subjects covered include: access control, computer security, cryptography, communications and data security, databases, electronic commerce, multimedia, bioinformatics, signal processing and image processing etc.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJCSS.

IJCSS List of Topics

The realm of International Journal of Computer Science and Security (IJCSS) extends, but not limited, to the following:

- Authentication and authorization models
- Computer Engineering
- Computer Networks
- Cryptography
- Databases
- Image processing
- Operating systems
- Programming languages
- Signal processing
- Theory
- Communications and data security
- Bioinformatics
- Computer graphics
- Computer security
- Data mining
- Electronic commerce
- Object Orientation
- Parallel and distributed processing
- Robotics
- Software engineering

IMPORTANT DATES

Volume: 4

Issue: 6

Paper Submission: November 31, 2010

Author Notification: January 01, 2011

Issue Publication: January /February 2011

CALL FOR EDITORS/REVIEWERS

CSC Journals is in process of appointing Editorial Board Members for ***International Journal of Computer Science and Security (IJCSS)***. CSC Journals would like to invite interested candidates to join **IJCSS** network of professionals/researchers for the positions of Editor-in-Chief, Associate Editor-in-Chief, Editorial Board Members and Reviewers.

The invitation encourages interested professionals to contribute into CSC research network by joining as a part of editorial board members and reviewers for scientific peer-reviewed journals. All journals use an online, electronic submission process. The Editor is responsible for the timely and substantive output of the journal, including the solicitation of manuscripts, supervision of the peer review process and the final selection of articles for publication. Responsibilities also include implementing the journal's editorial policies, maintaining high professional standards for published content, ensuring the integrity of the journal, guiding manuscripts through the review process, overseeing revisions, and planning special issues along with the editorial team.

A complete list of journals can be found at <http://www.cscjournals.org/csc/byjournal.php>. Interested candidates may apply for the following positions through <http://www.cscjournals.org/csc/login.php>.

Please remember that it is through the effort of volunteers such as yourself that CSC Journals continues to grow and flourish. Your help with reviewing the issues written by prospective authors would be very much appreciated.

Feel free to contact us at coordinator@cscjournals.org if you have any queries.

Contact Information

Computer Science Journals Sdn Bhd

M-3-19, Plaza Damas Sri Hartamas
50480, Kuala Lumpur MALAYSIA

Phone: +603 6207 1607
 +603 2782 6991
Fax: +603 6207 1697

BRANCH OFFICE 1

Suite 5.04 Level 5, 365 Little Collins Street,
MELBOURNE 3000, Victoria, AUSTRALIA

Fax: +613 8677 1132

BRANCH OFFICE 2

Office no. 8, Saad Arcad, DHA Main Bulevard
Lahore, PAKISTAN

EMAIL SUPPORT

Head CSC Press: coordinator@cscjournals.org
CSC Press: cscpress@cscjournals.org
Info: info@cscjournals.org

COMPUTER SCIENCE JOURNALS SDN BHD
M-3-19, PLAZA DAMAS
SRI HARTAMAS
50480, KUALA LUMPUR
MALAYSIA