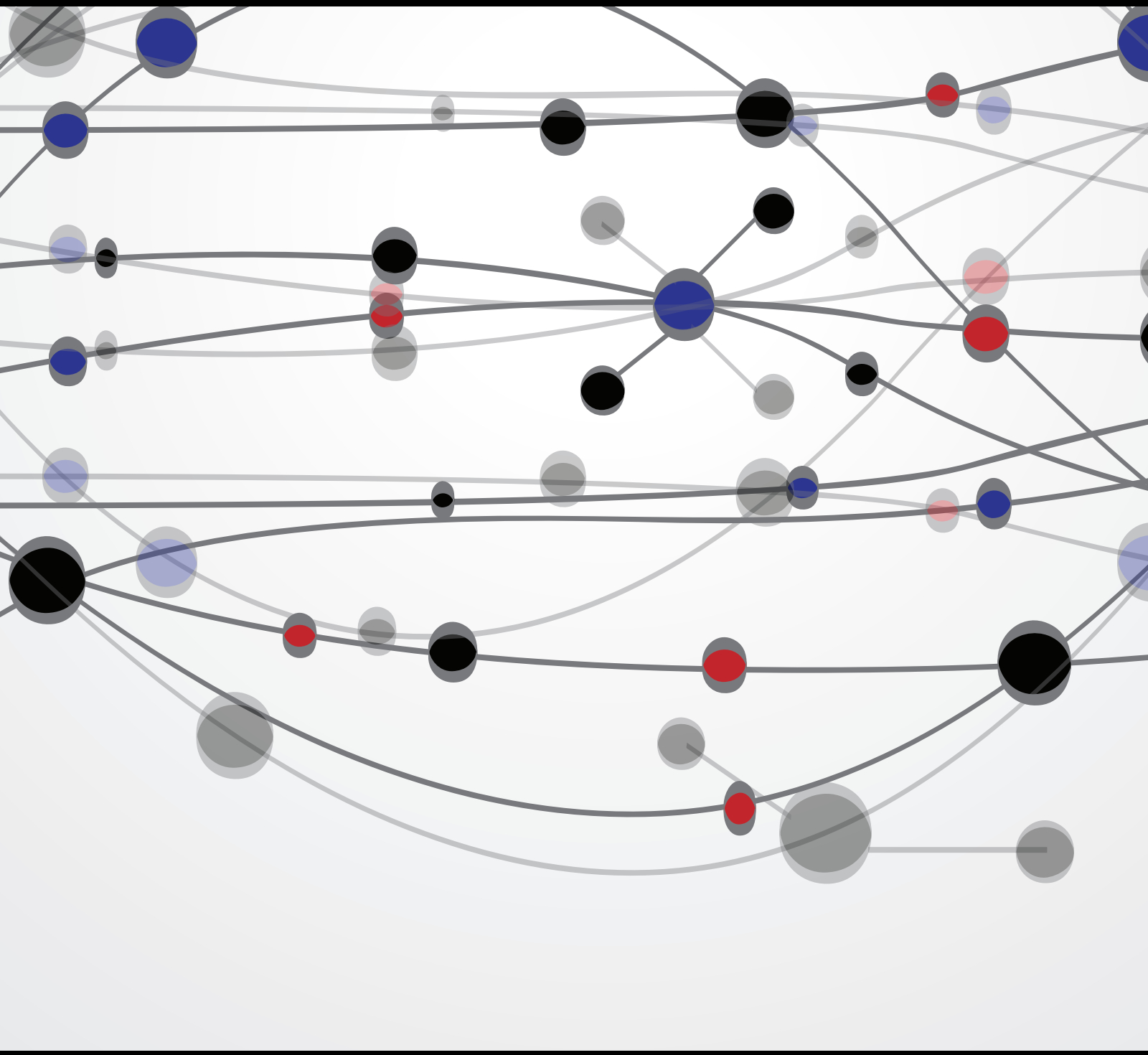
The Scientific World Journal

Security of Information and Networks

Guest Editors: Iftikhar Ahmad, Aneel Rahim, Adeel Javed, and Hafiz Malik



Security of Information and Networks

The Scientific World Journal

Security of Information and Networks

Guest Editors: Iftikhar Ahmad, Aneel Rahim, Adeel Javed, and Hafiz Malik

Contents

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2015, Article ID 150640, 2 pages
http://dx.doi.org/10.1155/2015/150640

Editorial

Security of Information and Networks

Iftikhar Ahmad,[1] Aneel Rahim,[2] Adeel Javed,[3] and Hafiz Malik[4]

[1] *King Saud University, Saudi Arabia*
[2] *Waterford Institute of Technology, Waterford, Ireland*
[3] *University of Otago, Dunedin, New Zealand*
[4] *University of Michigan, Dearborn, USA*

Correspondence should be addressed to Iftikhar Ahmad; wattoohu@gmail.com

Received 22 March 2015; Accepted 22 March 2015

This special issue aims to make people aware and up to date of the innovative research in the area of emerging techniques and methods in security of information and network. The majority of organizations in the commercial and government sectors are relying completely on their computer and network systems. Extensive attacks can cause heavy loss in a few seconds. Therefore, securing their computers, networks, and information is imperative. Based on the importance of security and interest of the researchers, this special issue is focused on attracting good work on innovative methods and techniques in order to address unique security issues which are introduced by new computing paradigms and techniques.

This special issue presents the following papers in the most active areas of research in security of information and networks. The brief introduction of the selected papers is provided.

O. Tayan et al. proposed the paper "A Hybrid Digital-Signature and Zero-Watermarking Approach for Authentication and Protection of Sensitive Electronic Documents." This paper addresses the problems and threats associated with verification of integrity, proof-of-authenticity, tamper-detection, and copyright protection for digital-text content. The proposed algorithm was implemented and shown to be robust against undetected content modifications and is capable of confirming proof-of-originality whilst detecting and locating deliberate/nondeliberate tampering. Additionally, enhancements in resource-utilization and reduced redundancies were achieved in comparison to traditional encryption-based approaches.

Y. Liu et al., worked on the paper "Network Anomaly Detection System with Optimized DS Evidence Theory" in which a novel network anomaly detection system is proposed with Optimized Dempster-Shafer (ODS) evidence theory and Regression Basic Probability Assignment (RBPA) function. In this model, the authors add weights to each senor to optimize DS evidence theory according to its previous predicted accuracy and RBPA employs sensor's regression ability to address complex network. They proved that this network anomaly detection model has a better detection rate, and RBPA and ODS optimization methods can improve system performance significantly.

A. Soleymani et al. proposed the paper "A Chaotic Cryptosystem for Images Based on Henon and Arnold Cat Map." In this paper, an encryption scheme is presented for securing images based on Arnold cat and Henon chaotic maps. The scheme uses Arnold cat map for bit and pixel-level permutations on plain and secret images, while Henon map creates secret images and specific parameters for the permutations. Both the encryption and decryption processes are explained, formulated, and graphically presented. The results of security analysis of five different images demonstrate the strength of the proposed cryptosystem against statistical, brute force and differential attacks. The evaluated running time for both encryption and decryption processes guarantees that the cryptosystem can work effectively in real-time applications.

S. M. Al-Saleem and H. Ullah worked on the paper "Security Considerations and Recommendations in Computer-Based Testing" in which the security considerations associated with CBT are investigated and some recommendations are given for the security of tests. A palm-based biometric authentication system is proposed and incorporated in

basic authentication system (username/password) in order to check the identity and authenticity of the examinee.

Z. Moghaddasi et al. worked on the paper "Improving RLRN Image Splicing Detection with the Use of PCA and Kernel PCA." This study focuses on improving one of the image splicing detection algorithms, that is, the Run-Length Run Number (RLRN) algorithm, by applying two dimension-reduction methods, namely, principal component analysis (PCA) and kernel PCA. Support vector machine is used to distinguish between authentic and spliced images. Results show that kernel PCA is a nonlinear dimension-reduction method that has the best effect on R, G, B, and Y channels and gray-scale images.

M. A. Saleh and A. A. Manaf proposed the paper "A Novel Protective Framework for Defeating HTTP-Based Denial of Service and Distributed Denial of Service Attacks." This work proposes a Flexible, Collaborative, Multilayer, DDoS Prevention Framework (FCMDPF). The innovative design of the FCMDPF framework handles all aspects of HTTP-based DoS/DDoS attacks through the following three subsequent framework's schemes (layers). Firstly, an Outer Blocking (OB) scheme blocks attacking IP source if it is listed on the black list table. Secondly, the Service Trace Back Oriented Architecture (STBOA) scheme validates whether the incoming request is launched by a human or by an automated tool. Then, it traces back the true attacking IP source. Thirdly, the Flexible Advanced Entropy Based (FAEB) scheme eliminates High Rate DDoS (HR-DDoS) and Flash Crowd (FC) attacks. The proposed framework's design provides an efficient protection for web applications against all sorts of DoS/DDoS attacks.

H. Y. Lee, proposed the paper "Method for Detecting Manipulated Compilation of Sensing Reports in Wireless Sensor Networks." In the proposed method, every sensing report is collaboratively generated and verified by cluster nodes based on very loose synchronization. Once a cluster node has detected an MCA for a real event, it can reforward a legitimate report immediately. Therefore, the event can be properly reported to the users. The performance of the proposed method is shown with analytical and experimental results.

V. Jaganathan et al. worked on the paper "Using a Prediction Model to Manage Cyber Security Threats." As cyberattacks are an important issue faced by all organizations, securing information systems is critical. Organizations should be able to understand the ecosystem and predict attacks. Predicting attacks quantitatively should be part of risk management. The cost impact due to worms, viruses, or other malicious software is significant. This paper proposed a mathematical model to predict the impact of an attack based on significant factors that influence cyber security. This model also considers the environmental information required. It is generalized and can be customized to the needs of the individual organization.

M. M. Al-Dabbagh et al. worked on the paper "Intelligent Bar Chart Plagiarism Detection in Documents." Plagiarism is considered as heinous electronic crime and intellectual theft. This paper introduced a new technique for extracting the features from documents which cannot be mined via Optical Character Recognition (OCR). By identifying the intimate relationship between the text and graphical components, the present technique pulls out the start, end, and exact value for each bar. Furthermore, the Word 2-gram and Euclidean distance methods are used to accurately detect and determine the plagiarism. The efficient detection of various plagiarized patterns is demonstrated. The system not only identifies copy-move forgery of bar charts but also distinguishes any possible modification applied on these images such as change of scales, colors, swapping among bars location, and even alteration on caption including summarizing and restructuring. This technique can constitute a basis for intelligent forgery detection in documents.

Iftikhar Ahmad
Aneel Rahim
Adeel Javed
Hafiz Malik

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2015, Article ID 493162, 11 pages
http://dx.doi.org/10.1155/2015/493162

Research Article

Method for Detecting Manipulated Compilation of Sensing Reports in Wireless Sensor Networks

Hae Young Lee

Department of Information Security, Seoul Women's University, 621 Hwarang-ro, Nowon-gu, Seoul 139-774, Republic of Korea

Correspondence should be addressed to Hae Young Lee; whichmeans@gmail.com

Received 27 August 2014; Accepted 2 November 2014

Academic Editor: Aneel Rahim

In cluster-based wireless sensor networks (WSNs), a few sensor nodes, including cluster heads (CHs), can be physically compromised by a malicious adversary. By using compromised CHs, the adversary can intentionally attach false message authentication codes into legitimate sensing reports in order to interrupt reporting of the real events. The existing solutions are vulnerable to such a type of security attacks, called *manipulated compilation attacks* (MCAs), since they assume that CHs are uncompromised. Thus, the reports manipulated by compromised CHs will be discarded by forwarding nodes or rejected at base stations, so that real events on the fields cannot be properly reported to the users. In this paper, the author proposes a method for the detection of MCAs in cluster-based WSNs. In the proposed method, every sensing report is collaboratively generated and verified by cluster nodes based on very loose synchronization. Once a cluster node has detected an MCA for a real event, it can reforward a legitimate report immediately. Therefore, the event can be properly reported to the users. The performance of the proposed method is shown with analytical and experimental results at the end of the paper.

1. Introduction

A wireless sensor network (WSN) [1] consists of a large number of small sensor nodes reporting real events (e.g., the appearances of vehicles) on the field and a few base stations (BSs) that collect sensing reports of the nodes. In many applications, these nodes are deployed in hostile environments, such as battlefields, and not attended by the users, so that a malicious adversary can physically capture some of them and thus potentially compromise the whole of their information, including keying material [2]. By using such compromised nodes, the adversary can inject *fabricated sensing reports*, which represent nonexistent events in the field, into the network with the goal of deceiving the BSs or depleting the limited energy resources of forwarding nodes [3, 4]. To prevent them from fabricating reports, every sensing report should be endorsed by multiple nodes, through attaching multiple message authentication codes (MACs) generated by them using different cryptographic keys, as in [3–8].

On the other hand, such a collaborative process has given rise to another security vulnerability in which some comprised nodes can intentionally attach false MACs to legitimate reports in order to interrupt reporting of the real events [8]. Although several security solutions [9–12] have been recently proposed to deal with such a type of security attacks, called *false-endorsements insertion attacks* (or *false-endorsement-based denial-of-service (DoS) attacks*) (FEIAs), they have a common vulnerability; a real event cannot be reported if a delegate, usually a cluster head (CH), for the event is compromised since they assume that the node is honest.

Let us suppose a cluster having a compromised CH, n_4, as shown in Figure 1. When n_4 has detected an event, it first announces the detection of the event. Once the announcement has been accepted by the other nodes in the cluster, n_1, n_2, n_3, and n_5, they send legitimate MACs, generated using their keys, for the event to the CH, in order to endorse the announcement. Although n_4 has collected a large enough number of legitimate MACs (i.e., M_1, M_2, M_3, and M_5), it can, in order to suppress reporting of the event, attach false MACs (X_1, X_2, X_4, and X_5) to a report or modify the contents of the report or even discard the report. Such a report would

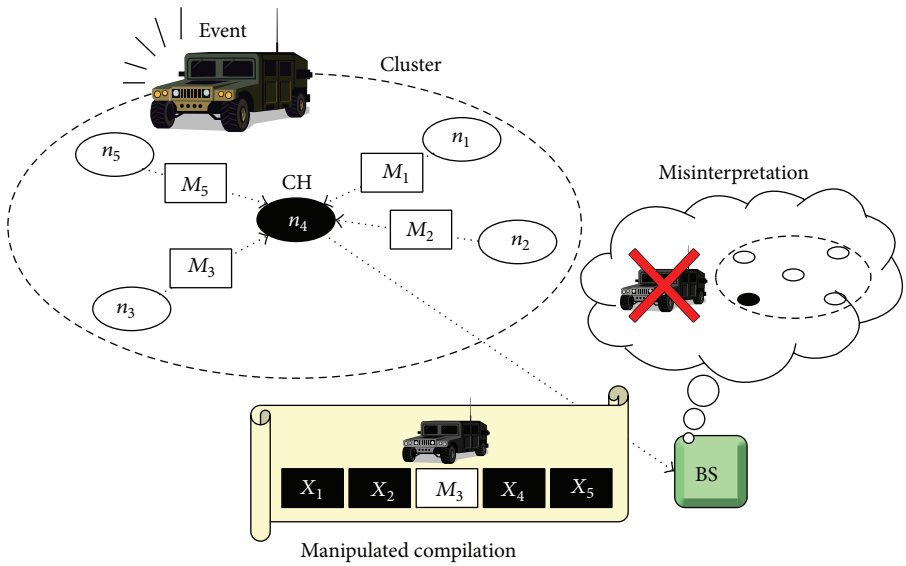


Figure 1: An example of manipulated compilation attacks (MCAs).

be filtered out by a forwarding node since the node considers the report to be a fabricated one. Moreover, if the report is delivered to a BS, innocent nodes could be misinterpreted as compromised ones. For example, in Figure 1, the BS could consider the report to be a fabricated one injected by a node n_3, which might be compromised, since every MAC in the report is incorrect except M_3.

In this paper, the author proposes a method for detecting such a type of security attacks, called *manipulated compilation attacks* (MCAs), in cluster-based WSNs. In the proposed method, for an event, a report is collaboratively generated by nodes in a cluster and forwarded by the CH, as in the existing solutions. In contrast to them, the method allows the report to be verified by the other nodes in the same cluster; if one of them has detected an MCA, the node can reforward a legitimate report generated by it. Therefore, the event can be properly reported to the users. In order to facilitate such a collaborative verification process, the local clocks of nodes in a cluster are very loosely synchronized while each report is being generated and verified. The performance of the proposed method is analyzed by analytical and experimental methods.

The remainder of the paper is organized as follows: Section 2 introduces the proposed method in detail, under the assumption of an ideal environment. Section 3 reviews the results of the performance analysis on the proposed method. In Section 4, some considerations for applying the proposed method to real-world WSNs are discussed. Related work on MCAs is surveyed in Section 5. Finally, conclusions and future work are then discussed in Section 6.

2. Detection of Manipulated Compilation Attacks (MCAs)

In this section, the proposed method is described in detail, under the assumption of an ideal environment. Some considerations for real-world applications are discussed in Section 4.

2.1. Network Model. A highly dense cluster-based WSN is considered since dense WSNs are resilient against false data injection attacks (FDIAs) [3–8], in which malicious adversaries inject fabricated sensing reports into the networks using a few compromised nodes with the goal of energy draining and/or false alarms. It is assumed that sensor nodes are similar to the current generation of sensor nodes (e.g., 802.15.4 Motes produced by [13]), in terms of computational and communication capabilities, and energy resources. When an event is detected by a node in a cluster, it is also detected by the other nodes in the cluster simultaneously. All nodes in a cluster can directly communicate with each other. A message transmitted by one of them can be heard by the other nodes in the cluster due to the broadcast nature of the wireless communications. Each node has some keys shared with the BSs and some other keys shared with all nodes in the same cluster. Some of the former keys can be shared with forwarding nodes in order to facilitate en-route verification (e.g., by [9] or [10]) of sensing reports. The latter keys are used to verify the origins of messages within the cluster.

2.2. Threat Model and Design Goal. It is assumed that a malicious adversary can physically compromise a few nodes, including CHs. However, the adversary cannot compromise the BS and half of the nodes in each cluster without being detected. The goal of the adversary is to suppress reporting of real events on the field, so that the adversary uses compromised nodes to launch MCAs against real events. The goal of the proposed method is to detect such misbehavior without any special equipment. Even if an MCA has been launched against a real event, a legitimate report for the event must be restored by honest nodes in order to report the event to the users. The method should be simple enough to be

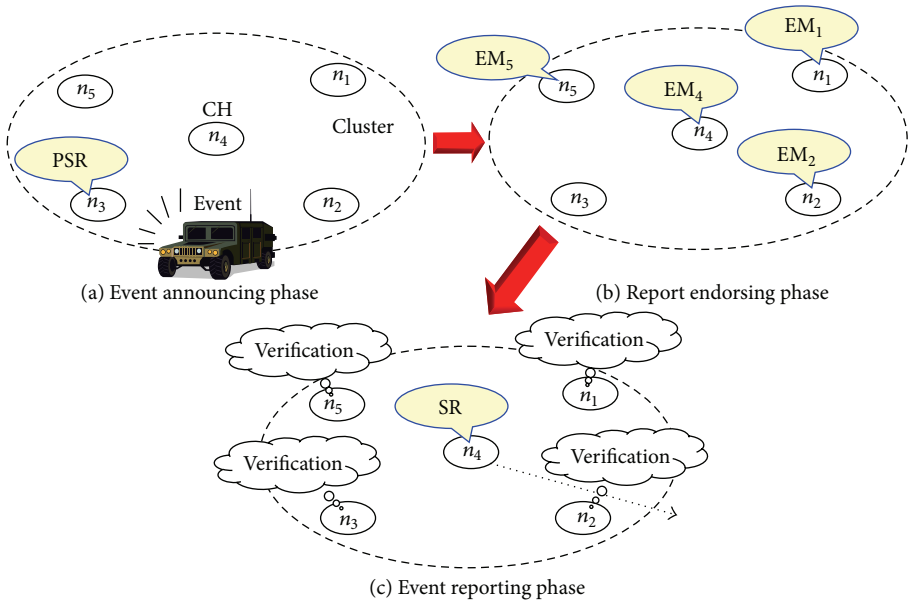


Figure 2: Basic procedure of the proposed method.

implemented on low-end node devices. It is also assumed that compromised nodes cannot forge their identities; such forgeries can be detected by using [14]. Please note that the identification of compromised nodes (e.g., as in [15]) is a separate issue and beyond of the scope of this paper.

2.3. Basic Procedure. Figure 2 shows the basic procedure of the proposed method. In the method, every sensing report is generated through the three phases: (a) *event announcing phase*, (b) *report endorsing phase*, and (c) *event reporting phase*. In the event announcing phase, a node in a cluster announces the detection of an event to the other nodes in the same cluster. If the nodes accept the announcement, their local clocks are very loosely synchronized by the receiving of the announcement. In the report endorsing phase, the nodes send messages to endorse the announcement to the CH. Due to the broadcast nature of wireless communications, all the nodes in the cluster can collect these endorsement messages. Finally, a report is forwarded by the CH and verified by the other nodes in the cluster within the event reporting phase. If one of them finds an MCA, it reforwards a report prepared by it. Thus, the event would be properly reported to the users. Please note that other loose synchronization methods [16, 17] or channel access methods (e.g., FDMA) could be used for very prompt notifications of events (see Section 4.1).

2.4. Event Announcing Phase. When a node in a cluster detects a real event, the event is also detected by the other nodes in the same cluster simultaneously. Each of the detecting nodes in the cluster (i.e., all the nodes in the cluster) then prepares a *preliminary sensing report* (PSR) for the event and sets a random timer. The PSR would contain the ID of the node, the contents for the event, and two MACs over the contents, generated using two of the node's keys, one for en-route verification and another for source verification. Upon the expiration of the timer, the node broadcasts its PSR and then resets its local clock to 0. When one of the other detecting nodes has received the PSR, the node checks if the PSR states the same event detected by the node and was generated by one of the nodes in the cluster. If so, the node accepts the PSR. Once a node in the cluster has accepted the PSR, the node cancels its timer and resets its local clock to 0. Therefore, if all the nodes in the cluster accept the PSR, their lock clocks are very loosely synchronized.

Each of the accepting nodes then prepares an endorsement message (EM), usually containing the ID of the node and two MACs over the contents, generated using two of the node's keys (as in PSR). A compromised node may not broadcast a PSR, in order to suppress reporting of the event. However, the detection of the event will be announced by an honest node in the cluster, so that the process will continue. From the synchronization point (upon accepting the PSR), time is divided into 2 rounds, each composed of N time slots of a fixed size, where N is the number of the nodes in the cluster. The first slot of each round is always assigned to the CH, and the rest of the slots are assigned to the other nodes in the cluster. The determination of the slot size, CH selection, and slot assignment are separate issues and beyond of the scope of the paper. We might use some existing solutions for these issues (e.g., [18, 19] for the assignment), possibly with slight modifications.

As shown in Figure 3, let us suppose that there is a cluster consisting of 5 nodes, $n_0, \ldots, n_5$, and n_4 which is the current CH. For an event, n_3 has first broadcasted its PSR ("red" dotted arrows within event announcing), and set its local clock to 0. n_1, n_2, n_4, and n_5 accept the PSR if it states the event detected by them. Then, by setting their local clocks to 0 upon accepting the PSR, the clocks are very loosely synchronized

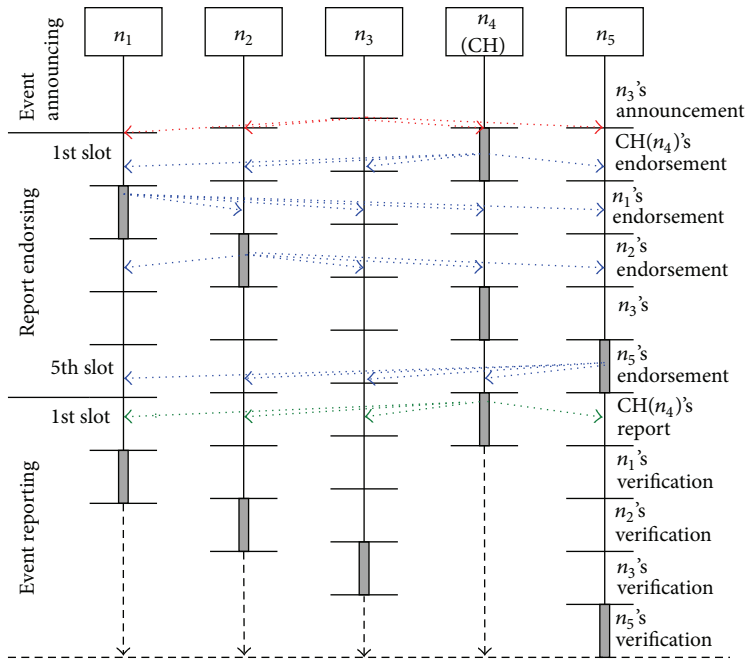


Figure 3: Report generation and verification in the proposed method.

(the uppermost horizontal lines). They also prepare their EMs. Even if a compromised node has not announced the detection of the event, one of the other honest nodes will broadcast a PSR. From the synchronization point, time is divided into 2 rounds, each composed of 5 time slots (i.e., $N = 5$). The first slot of each round has been assigned to the current CH, n_4. The second, third, fourth, and fifth slots of each round have been assigned to n_1, n_2, n_3, and n_5, respectively.

2.5. Report Endorsing Phase. In the report endorsing phase, each of the accepting nodes broadcasts its EM, within the slot assigned to it. Meanwhile, every node in the cluster collects these EMs and prepares a sensing report (SR). The SR should include the IDs of the cluster and node, the contents, and N MACs collected from the PSR and EMs. Some compromised nodes may not broadcast their EMs. In this case, a node simply puts *blank MACs*, which are filled in 0 s, on the positions of their MACs in the SR. Compromised nodes may also broadcast EMs containing false MACs. However, unless $\lceil N/2 \rceil$ or more nodes are compromised, $\lceil N/2 \rceil$ or more legitimate MACs can be collected and attached into the SR.

In the example shown in Figure 3, the current CH, n_4, first broadcasts its EM within the first slot. n_1 and n_2 then broadcast their EMs within the second and third slots, respectively. Although the fourth slot is assigned to n_3, the other nodes have already collected n_3's MAC from the PSR. Thus, n_3 need not broadcast an EM. Finally, n_5 broadcasts its EMs within the last slot. These broadcasts of EMs are represented in "blue" dotted arrows within report endorsing in the figure. Meanwhile, each of them collects these EMs and prepares an SR. Even if two of the nodes (e.g., n_1 and n_2) are compromised, 3 legitimated MACs (e.g., from n_3, n_4, and n_5) can be collected. If they have not broadcasted their EMs, blank MACs are put on the positions of their MACs in the SR.

2.6. Event Reporting Phase. In this phase, the CH forwards its SR on behalf of the other nodes in the cluster, within the first slot. Due to the broadcast nature of wireless communications, theses nodes can also hear the SR forwarded by the CH. Each of them, within its slot, compares the SR lastly forwarded with its SR (the SR prepared by it). If they match exactly, the former is considered to be legitimate. But if not, the former is considered manipulated, so that the node forwards the latter immediately within its slot. The SR newly forwarded can be also verified by the *remaining nodes* that have not compared them yet in the cluster. A compromised CH may forward a SR manipulated, or even nothing. However, such an MCA can be detected by the remaining honest nodes in the cluster. If the last slot has been assigned to a compromised node, no one can verify a SR of the node. Nevertheless, a legitimate SR would have been already forwarded by an honest node in the cluster before the last slot. Therefore, unless $\lceil N/2 \rceil$ or more nodes are compromised, a legitimate SR, which has $\lceil N/2 \rceil$ or more legitimate MACs, can be delivered to a BS.

In the example shown in Figure 3, the current CH, n_4, first forwards its SR toward a BS within the first slot of the second round. Within the second slot, n_1 compares the SR forwarded by n_4 with its SR. If they do not match, or the CH did not forward a SR before the second slot, it reforwards the latter immediately toward the BS. In this case, n_1's SR is also verified by the remaining nodes, n_2, n_3, and n_5, within their slots. Please note that n_2 will verify n_4's SR within the slot if n_1

Table 1: The sizes of the messages used to generate a report.

Size	The proposed method		The existing solutions	
	Not using key indices	Using key indices	Not using key indices	Using key indices
PSR	$S_{NID} + S_C + 2 \cdot S_{MAC}$	$S_{NID} + S_C + S_{KID} + 2 \cdot S_{MAC}$	$S_{NID} + S_C + S_{MAC}$	$S_{NID} + S_C + S_{MAC}$
EM	$S_{NID} + 2 \cdot S_{MAC}$	$S_{NID} + S_{KID} + 2 \cdot S_{MAC}$	$S_{NID} + 2 \cdot S_{MAC}$	$S_{NID} + S_{KID} + 2 \cdot S_{MAC}$
SR	$S_{CID} + S_{NID} + S_C + N \cdot S_{MAC}$	$S_{CID} + S_{NID} + S_C + N(S_{KID} + S_{MAC})$	$S_{CID} + S_{NID} + S_C + N \cdot S_{MAC}$	$S_{CID} + S_{NID} + S_C + N(S_{KID} + S_{MAC})$

has forwarded nothing. If n_2 has found an MCA, or n_4 and n_1 did not forward a SR before the third slot, it reforwards its SR. The remaining nodes, n_3 and n_5, will also verify the lastly forwarded SR. Therefore, unless three or more nodes are compromised, at least one legitimate SR would be forwarded by other honest nodes.

2.7. En-Route Verification of Sensing Reports. While a SR is being forwarded toward a BS, forwarding nodes can check the legitimacy of the SR based on a FEIA countermeasure, such as [9]. Blank MACs are considered to be false ones. In order to enable every SR to be delivered to BSs, a centralized detection solution, such as [20], can be applied to the network. However, in general, such solutions cannot provide a mechanism to filter fabricated reports out en-route, so that they can be vulnerable to injecting fabricated reports in terms of energy saving.

3. Performance Analysis

The performance of the proposed method is analyzed in this section.

3.1. Communication Overhead. The proposed method can be employed to generate sensing reports in the existing FEIA countermeasures [9–12]. In the method, a PSR usually consists of the ID of a node, the contents for an event, and two MACs over the contents generated using a key shared with BSs and another one shared with the other nodes in the same cluster. An EM is usually comprised of the ID of a node and two MACs over the contents generated using two keys. A SR usually includes the ID of a cluster, the ID of a node, the contents for an event, and N MACs generated using different keys from different nodes in the cluster. In case of using a global key pool (e.g., when the multipath-based en-route filtering scheme (MEF) [10] is used), the indices of keys used to generate MACs are included in the messages. Also in the existing solutions, messages in a similar format are used to collaboratively generate a report.

The sizes of these messages in bytes are summarized in Table 1, where S_{CID} is the size of a cluster ID, S_{NID} is the size of a node ID, S_C is the size of the contents for an event, S_{KID} is the size of a key ID, and S_{MAC} is the size of a MAC. In terms of message sizes, the difference between the method and them is S_{MAC} in a PSR.

Both in the method and existing solutions, a report is collaboratively generated through the same number of message transmissions and receptions within a group (e.g., a cluster). Therefore, their communication overhead is

$$\begin{aligned} O_{comm} = {} & |\mathrm{PSR}| \times \left(O_{trans} + (N-1)\, O_{recept}\right) \\ & + |\mathrm{EM}| \times \left((N-1)\, O_{trans} + (N-1)^2\, O_{recept}\right) \\ & + |\mathrm{SR}| \times \left(O_{trans} + (N-1)\, O_{recept}\right), \end{aligned} \quad (1)$$

where is $|M|$ is the size of message M, O_{trans} is the communication cost for a transmission, and O_{recept} is the communication cost for a reception.

3.2. Computation Overhead. The computation overhead of the proposed method has been analyzed based on the number of the MAC computations (generation and verification) required to generate a report, as in [3]. In the event announcing phase of the method, at most $2 \cdot N$ MACs are generated by all nodes in a cluster. In case of deterministic announcement (e.g., a preassigned node always announces the detection of an event), 2 MACs are generated by a node. Once a node has broadcasted a PSR, a MAC of the PSR is verified by the other nodes (i.e., $N - 1$ nodes) in the cluster to check the origin of the PSR. In the report endorsement phase, $2(N-1)$ MACs are generated by them. Each of $N - 1$ MACs attached in EMs is then verified by $N - 1$ nodes to check the origins of the EMs. No MAC is generated or verified in the event reporting phase.

The existing solutions also require a report to be generated through multiple MAC computations: (1) when an event has been detected, at most N MACs are generated by all participating nodes before the announcement of the detection. Once the detection has been announced by a node (announcing node), the MAC attached in the announcement is verified by the other participating nodes. In case of deterministic announcements, a MAC is generated by an announcing node and then verified by the other participating nodes. (2) In order to endorse the announcement, a single MAC is generated by the announcing node and $2(N - 1)$ MACs are generated by the other participating nodes. The announcing node collects them and verifies $N - 1$ MACs to check origins of the endorsements. (3) No MAC is computed when the announcing nodes has forwarded a report.

The number of MAC computations is summarized in Table 2, where EAP, REP, and EFP are the event announcing phase, report endorsement phase, and event forwarding phase. In the method, every node in a cluster prepares a SR with the verification of EMs, so that the computation overhead of the method is larger than that of the existing ones,

Table 2: The number of MAC computations in the proposed method and the existing solutions.

MAC computations	The proposed method		The existing solutions	
	Random	Deterministic	Random	Deterministic
EAP	$3 \cdot N - 1$	$N + 1$	$2 \cdot N - 1$	N
REP	$N^2 - 1$	$N^2 - 1$	$2 \cdot N - 1$	$2 \cdot N - 1$
ERP	0	0	0	0

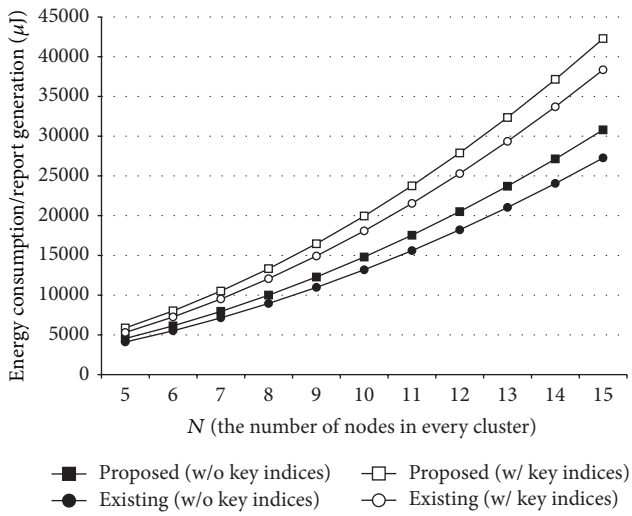


Figure 4: Energy efficiency of the proposed method in case of random announcement.

especially in REP. However, the method can provide MCA resilience for WSNs, whereas the others are vulnerable to MCAs.

3.3. Energy Efficiency. The energy efficiency of the proposed method has been analyzed based on these communication and computation overheads. It is assumed that $S_{\mathrm{CID}} = 2$, $S_{\mathrm{NID}} = 2$, $S_{\mathrm{C}} = 8$, $S_{\mathrm{KID}} = 2$, and $S_{\mathrm{MAC}} = 2$. It is also assumed that $O_{\mathrm{trans}} = 16.25\,\mu\mathrm{J}$, $O_{\mathrm{recept}} = 12.5\,\mu\mathrm{J}$, and $15\,\mu\mathrm{J}$ is consumed for each MAC computation, as in [3]. The amount of energy resources consumed by the generation of a report has been measured. Since SR in the proposed method is identical to that in the existing countermeasures, the energy consumption due to the forwarding of a report is not considered.

Figures 4 and 5 show the energy efficiency of the proposed method when N is between 5 and 15. A report is generated through a random announcement in Figure 4, whereas an announcing node is deterministically chosen in Figure 5. As shown in the figures, the proposed method (rectangles) consumed more energy resources than the existing solutions (circles) due to the mutual verification among cluster nodes. However, it could provide WSNs with the resilience to MCAs, so that real events on the field could be properly reported to the users. In case of using a global key pool especially, extra energy consumption could be reduced to 9~10%.

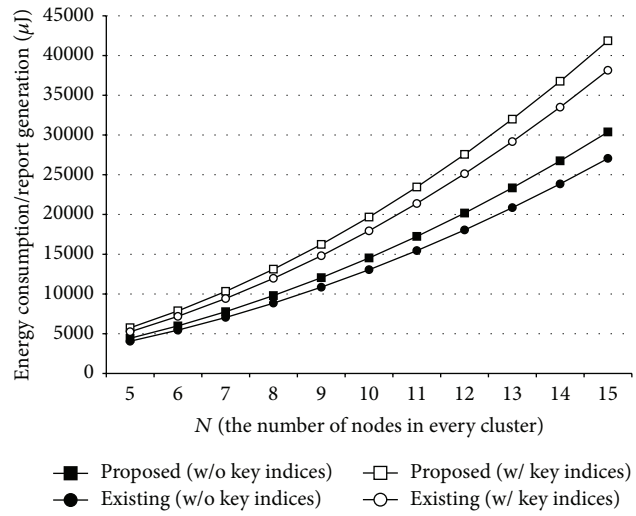


Figure 5: Energy efficiency of the proposed method in case of deterministic announcement.

3.4. Resilience to MCAs. If C ($\leq N$) nodes in a cluster have been compromised by an adversary, a SR generated by an honest node for a real event could include at most C false MACs in the proposed method. At a BS, a report can be considered to be legitimate if at least half of the MACs attached in the report are correct. Therefore, unless $\lceil N/2 \rceil$ or more nodes in a cluster are compromised, SRs generated in the cluster will be properly interpreted at BSs.

The resilience of the method against MCAs has been also evaluated through simulation. A field size of $200 \times 200\,\mathrm{m}^2$ is used and a single BS is located at the end of the field. The network consists of 400 clusters and each covers $10 \times 10\,\mathrm{m}^2$. Note that the proposed method deals with in-cluster generation and verification of reports. Thus, it is not virtually affected by the size of the network. In each cluster, $N/2 - 1$ nodes, including the CH, are compromised by an adversary (for the performance evaluation). 2,000 real events occurred on random locations. The method is employed to generate reports in two existing FEIA countermeasures: the probabilistic voting-based filtering scheme (PVFS) [9] and MEF [10].

Figure 6 shows the percentage of the reports delivered to the BS (PRD) on a PVFS-based WSN when N is $6, 8, \ldots, 14$ and the number of false MACs per report, inserted by a CH (i.e., due to an MCA), is between 3 and 14. A forwarding node discards a report if the number of false MACs in the report reaches $N/2$ (i.e., half of them are incorrect). As shown in the figure, the method (filled rectangles) employed for PVFS could guarantee that an event would be always delivered to a BS regardless of MCAs until $N/2 - 1$ compromised nodes since every report would have at most $N/2 - 1$ false MACs. In contrast, PVFS (diamonds) without the method was totally vulnerable to MCAs; even a single compromised CH could make legitimate reports filtered out through the manipulation of their MACs. Moreover, even if such reports could be delivered to the BS, they would be rejected by the BS. Therefore, reporting of the events occurring around

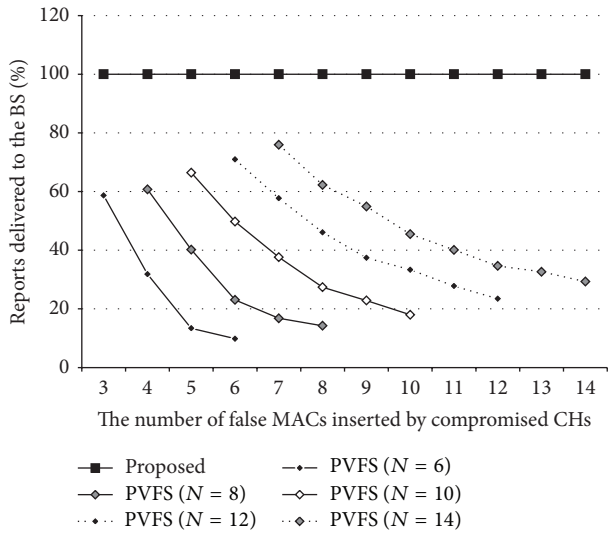


Figure 6: MCA resilience of the proposed method on PVFS.

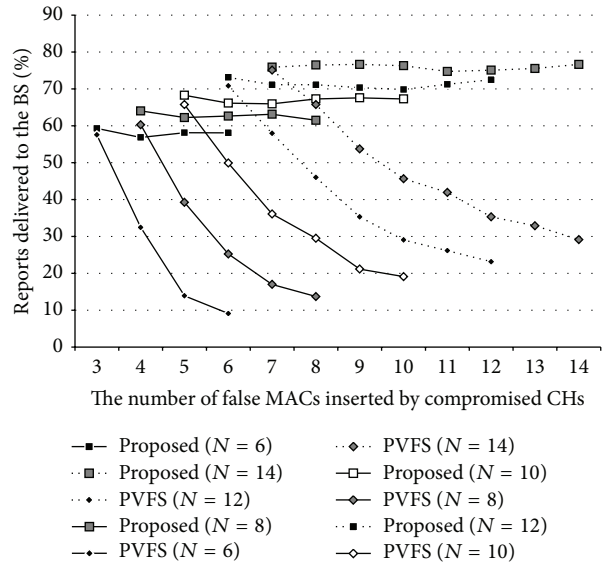


Figure 7: MCA resilience of the proposed method on PVFS when half of the nodes in a cluster are compromised.

the cluster could be completely suppressed by the CH with ease. The method would be enough worth considering since it can provide MCA resilience for the network with 10% overhead.

Figure 7 shows PRD on a PVFS-based WSN when half of the nodes in a cluster are compromised. As shown in the figure, the method (filled rectangles) could provide the network with MCA resilience to a certain degree. If an event could be detected by two or more clusters (e.g., due to its movement), the event would be properly reported to the users with the method. Please note that the acceptance of them at the BS is another issue. In contrast, the PRD of PVFS severely decreased with the number of false MACs, which might result in a malfunction of the network. Note that reports were discarded by the verification mechanism of PVFS, not by the proposed method. In PVFS (and also in MEF), PRD would decrease as reports travel more hops. Thus, for ultra-large-scale WSNs, we should slightly alter some PVFS parameters to achieve a sufficient level of PRD. However, such alterations also decrease the resilience of the networks against to FDIAs (see Section 5).

Figure 8 shows PRD when MEF using dual-path routing is applied to the network, N is $6, 8, \ldots, 14$, and the number of the false MACs per report is between 3 and 14. There are 2,000 keys in the global key pool maintained by the BS and each node loads 50 keys randomly chosen from the pool. In MEF, a forwarding node drops a report immediately if the report carries any false MAC. Thus, even the method (rectangles), as shown in the figure, could not guarantee that every legitimate report would be delivered to the BS. However, the PDR of the method, while being affected by the number of compromised nodes, was not affected by the number of false MACs inserted by compromised CHs, so that the method is resilient to MCAs. In contrast, the PDR of MEF (triangles) was seriously affected by the number of the false MACs. Also, a report manipulated by a compromised CH, while stating a real event, would be eventually rejected by the BS even if it could be delivered to the BS. Thus, method would be well worth considering despite its overhead.

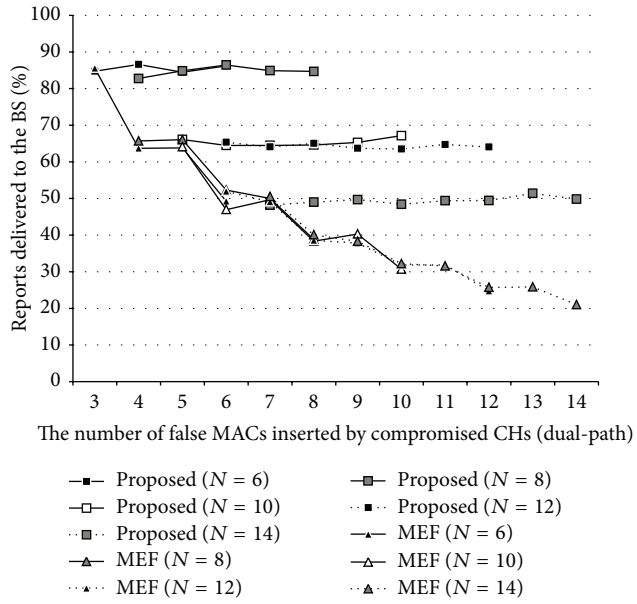


Figure 8: MCA resilience of the proposed method on MEF w/dual-path routing.

Figure 9 shows PRD when MEF using triple-path routing is applied to the network, N is $6, 9, \ldots, 15$, and the number of the false MACs per report is between 3 and 15. As shown in the figure, the PRDs of the method and MEF increased in triple-path routing. The PRD of the method especially (rectangles) was largely enhanced with the routing although more energy resources would be consumed for event reporting. In contrast, the vulnerability of MEF (triangles) was not alleviated regardless of using more energy resources.

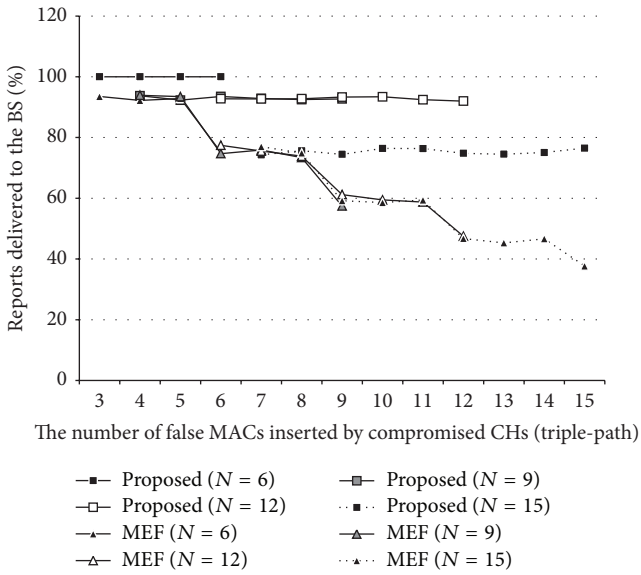


Figure 9: MCA resilience of the proposed method on MEF w/triple-path routing.

4. Consideration for Real-World Applications

This section discusses some considerations for applying the proposed method to real-world WSNs.

4.1. Collaboration Problems. The proposed method employs a loose time synchronization mechanism for collaborative reporting on events, in which local clocks in a cluster are very loosely synchronized by receiving a PSR. A benefit of such a simple mechanism would be that communication and computing overheads could be reduced, so that it could save energy resources and be implemented on low-end node devices. On the other hand, due to its imprecision in time synchronization, the size of a slot, T_S, must be large enough to minimize synchronization errors. That is, its applications must allow a large latency in reporting events. The author argues that such a large latency would be acceptable for real-world responses in large-scale WSNs since it might be much smaller than delays in real-world responses; even $T_S = 1$ s might be acceptable for ultra-large-scale WSNs.

For some other applications that require very prompt notifications of events, the following approaches could be considered.

(i) Use of other loose synchronization methods: other loose synchronization solutions [16, 17] that can provide much more precision (e.g., 7.24 μs [16]) could be employed for the collaboration. For example, one of the detecting nodes is elected as a CH in the initial round comprised of N slots. In the next round, endorsements for the event are collected. A report for the event is forwarded and verified in the final round. A problem of the use of such precise solutions is that they could involve more communication and computation overheads.

(ii) Use of other channel access methods: we could use alternative channel access methods, such as FDMA, for the collaboration. For example, nodes in a cluster use different channels, so that endorsements could be collected simultaneously through these channels. If one of them finds an MCA on a SR, it reforwards its report through the channel assigned to it. However, due to additional circuitry requirements to dynamically communicate with different radio channels, the cost of sensor nodes is increased, which is contrary to the objective of sensor network systems [21].

4.2. Handling Multiple Events. While a report for an event is being generated by nodes in a cluster, another event can be also detected by them, which could interrupt the generation of the report. This problem might be a security vulnerability since malicious adversaries could misuse it as "physical" denial-of-service attacks. The following approaches could alleviate the problem.

(i) Shortening of T_S: a most simple approach might be to shorten T_S, enough to handle multiple events. However, it might increase synchronization errors.

(ii) Buffering of events: if another event occurs during the generation of a report, each node stores the event in a buffer. For the generation of a report on the new event, local clocks need not be reset again; the generation is begun with the third round, which leads to the election of a CH. The report, which includes endorsements collected in the fourth round, would be forwarded and verified in the fifth round. A problem of event buffering is that some events might be missed if events occur very frequently in WSNs.

(iii) Prompt notifications of events: we could handle multiple events by making them being reported promptly, which is discussed in Section 4.1. However, it requires high-performance hardware or additional circuitry.

4.3. Missing Events. The proposed method assumes that an event can be detected by all nodes in a cluster simultaneously. However, in the real world, each sensor could miss some events, which would lead to the production of SRs with many blank MACs; such SRs might be considered to be fabricated ones. Thus, PRD would decrease with missing (false) MACs, as shown in Figures 6, 7, 8, and 9. This problem could be alleviated with the following approaches.

(i) Assignment of multiple nodes to each time slot: for very-dense WSNs, a simple (but effective) approach might be to assign multiple nodes to each time slot. For example, we could deploy $2 \cdot N$ nodes to each cluster and assign two nodes to each time slot. When an event has occurred in a cluster, one of the detecting nodes is elected as a CH through the same procedure. For each slot of the first round, one of the two nodes assigned to the slot first broadcasts its EM by the expiration of its (another) random timer. Upon receiving the EM, the other node checks the EM.

If the EM is "correct," the node cancels its timer and quits from participating in the remaining procedure. Then, an SR would be compiled with the EM and verified by the former node. If not, the latter node broadcasts its EM within the slot. Then, the CH would choose a more "trustable" one to endorse its SR. There is still a probability that the whole nodes assigned to a slot could miss events. However, it might be low enough to make every SR carry few or no blank MACs. Theoretically, the probability is $(P_{\mathrm{ME}})^{N_{\mathrm{NPS}}}$, where P_{ME} is a probability of missing events and N_{NPS} is the average number of nodes assigned to a time slot.

(ii) Attachment of partial MACs: the problem could be alleviated by allowing a report carry partial MACs. For example, a report is allowed to carry just $N/2$ MACs, excluding blank MACs. This approach requires a slight modification on PVFS or MEF and very-dense WSNs.

(iii) CH-level cooperation with neighboring clusters: another potential approach might be to make CHs cooperate with each other. In many applications, an event can be detected by multiple clusters simultaneously or a gap of time. Thus, if a CH could not collect enough number of MACs, it might cooperate with its neighboring CHs. A single SR might be produced and forwarded by them. The cooperation of clusters is a separate issue and should be investigated in future.

4.4. Packet Loss. The proposed method does not consider packet loss (including due to CRC errors) although wireless links are usually unreliable. Loss of some packets might interrupt the generation of reports. The following discusses potential problems of packet loss for each type of message.

(i) Loss of some PSRs: loss of PSRs might not be a serious problem since any node who has first broadcasted a "correct" PSR would become a CH; although a few PSRs have been lost, each of the remaining nodes has still a chance to broadcast its PSRs. In very-dense WSNs, most PSRs for an event could be lost due to collisions. Such collisions could be minimized by extending the interval of random timers (i.e., the size of a round) or by using other collaboration methods described in Section 4.1.

(ii) Loss of some EMs: we could consider loss of some EMs to be missing events, so that it would cause the production of SRs with many blank MACs. Thus, it could be alleviated with the approaches described in Section 4.3.

(iii) Loss of some SRs: SR loss might not be a significant problem since the "remaining" nodes would attempt to forward their SRs.

5. Related Work

Ye et al. [3] first addressed FDIAs in which fabricated sensing reports are injected through a few compromised nodes in order to make false alarms or energy consumption and then proposed the statistical en-route filtering scheme (SEF) as a countermeasure. In SEF, every sensing report must carry a certain number of MACs, generated by different detecting nodes, using keys from different partitions of the global key pool. A report is, while being forwarded toward a BS, verified by some forwarding nodes and discarded immediately if the verification fails. BSs maintain the key pool, so that a report is finally verified by a BS. Thus, unless an adversary has compromised a large number of nodes, the adversary has no choice but to forge some MACs in order to fabricate a report. Such a report having forged MACs will be discarded by a forwarding node or rejected at a BS. Such a collaborative generation and en-route verification mechanism have been served as a foundation for providing resilience against FDIAs in other countermeasures [3–8]. In the interleaved hop-by-hop authentication scheme proposed by Zhu et al. [5], for example, a report must be endorsed by all the nodes in a cluster and verified by every forwarding node in an interleaved fashion. Polynomials instead of MACs can be used to endorse reports since it could improve resilience to forgeries of identities [22].

Li et al. [9] first found that such a mechanism has given rise to a vulnerability that a few compromised nodes can insert some false MACs into a report for a real event in order to suppress reporting of the event (i.e., FEIAs). Thus, they proposed PVFS that allows a report to be verified multiple times by forwarding nodes. Once the number of successful verifications has reached a predefined threshold value, the report is considered to be legitimate. The report is then forwarded to the BS without en-route verification, so that some resilience to FEIAs can be provided. If the number of failed verifications exceeds another threshold value, the report is dropped immediately in order to provide some resilience to FDIAs. MEF proposed by Kim and Cho [10] can be also used to defend FEIAs. In MEF, a forwarding node instantly drops a report if the verification fails, as in SEF. However, multiple copies of a report are generated for each event and are delivered to BSs through different paths. The keys used to endorse them differ from each other. Thus, the event can be reported to the users even if some of them were dropped by forwarding nodes. Another countermeasure proposed by Krauß et al. [11] enables a report generating node (e.g., a CH) to detect a node that has sent a false MAC, by making the node prove the correctness of the MAC. If the proof fails or the node does not perform the proof, the node is excluded in the report generation process. In [12], Yu et al. adopt polynomials instead of MACs for verification, which could increase resilience against forging identities.

All of these countermeasures are vulnerable to MCAs since they assume that a report generating node will compile a report with the MACs sent by other detecting nodes, without any manipulation. A compromised one, however, can compile a report for a real event with forged MACs although it has collected legitimate MACs from other detecting nodes. The real event will not be reported to the users since the report will be considered to be fabricated one representing nonexisting event.

6. Conclusions and Future Work

In this paper, the author proposed a method for detecting MCAs in cluster-based WSNs. In the proposed method, every report is collaboratively generated and verified by all nodes in a cluster based on very loose synchronization. Unless $\lceil N/2 \rceil$ or more nodes in a cluster are compromised, reports generated in that cluster could be delivered to the BSs. The performance of the proposed method was analyzed with analytical and experimental results. With only 10% overhead, the method could provide WSNs with MCA resilience. Therefore, it would be well worth considering for WSN security.

Although the method is designed for cluster-based WSNs, it could be employed for flat WSNs with a modification. Several issues not covered in the paper, including secure CH selections, CH-level collaboration, and slot assignment with the consideration of clock synchronization, will be studied. For the verification purpose, the author will try to formally prove the proposed method, for example, through model checking. In order to enhance the performance of the proposed method, the author will also try to apply other synchronization solutions or channel access methods to the method. By extending the method, en-route detection of report manipulation will be further investigated. All of these will be implemented on real nodes, in order to provide guidelines on the selection of design parameters, such as N and T_S.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by a special research grant from Seoul Women's University (2014). This research was partially supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (NRF-2013R1A1A1006542).

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–105, 2002.

[2] C. Krauß, M. Schneider, and C. Eckert, "On handling insider attacks in wireless sensor networks," *Information Security Technical Report*, vol. 13, no. 3, pp. 165–172, 2008.

[3] F. Ye, H. Luo, S. Lu, and L. Zhang, "Statistical en-route filtering of injected false data in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 839–850, 2005.

[4] H. Yu, J. He, R. Liu, and D. Ji, "On the security of data collection and transmission from wireless sensor networks in the context of internet of things," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 806505, 13 pages, 2013.

[5] S. Zhu, S. Setia, S. Jajodia, and P. Ning, "Interleaved hop-by-hop authentication against false data injection attacks in sensor networks," *ACM Transactions on Sensor Networks*, vol. 3, no. 3, article 14, 2007.

[6] T. P. Nghiem and T. H. Cho, "A fuzzy-based interleaved multi-hop authentication scheme in wireless sensor networks," *Journal of Parallel and Distributed Computing*, vol. 69, no. 5, pp. 441–450, 2009.

[7] Z. Yu and Y. Guan, "A dynamic en-route filtering scheme for data reporting in wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 18, no. 1, pp. 150–163, 2010.

[8] Z. Liu, J. Wang, and X. Zhang, "A false data filtering scheme using cluster-based organization in sensor networks," in *Proceedings of the IEEE International Conference on Communications (ICC '11)*, pp. 1–5, Kyoto, Japan, June 2011.

[9] F. Li, A. Srinivasan, and J. Wu, "PVFS: a probabilistic voting-based filtering scheme in wireless sensor networks," *International Journal of Security and Networks*, vol. 3, no. 3, pp. 173–182, 2008.

[10] M. S. Kim and T. H. Cho, "A multipath en-route filtering method for dropping reports in sensor networks," *IEICE Transactions on Information and Systems*, vol. 90, no. 12, pp. 2108–2109, 2007.

[11] C. Krauß, M. Schneider, and C. Eckert, "Defending against false-endorsement-based DoS attacks in wireless sensor networks," in *Proceedings of the 1st ACM Conference on Wireless Network Security (WiSec '08)*, pp. 13–23, March 2008.

[12] C.-M. Yu, Y.-T. Tsou, C.-S. Lu, and S.-Y. Kuo, "Constrained function-based message authentication for sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 2, pp. 407–425, 2011.

[13] Advanticsys, http://www.advanticsys.com/.

[14] J.-W. Ho, M. Wright, and S. K. Das, "Fast detection of mobile replica node attacks in wireless sensor networks using sequential hypothesis testing," *IEEE Transactions on Mobile Computing*, vol. 10, no. 6, pp. 767–782, 2011.

[15] J.-W. Ho, M. Wright, and S. K. Das, "Zone trust: fast zone-based node compromise detection and revocation in wireless sensor networks using sequential hypothesis testing," *IEEE Transactions on Dependable and Secure Computing*, vol. 9, no. 4, pp. 494–510, 2012.

[16] A. S. Uluagac, R. A. Beyah, and J. A. Copeland, "Secure SOurce-BAsed Loose Synchronization (SOBAS) for wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 4, pp. 803–813, 2013.

[17] J. Deng, R. Han, and S. Mishra, "Secure and robust loose time synchronization mechanism for wireless sensor networks," in *Proceedings of the the 13th USENIX Security Symposium*, August 2004.

[18] A. Y. Barnawi, "Adaptive TDMA slot assignment using request aggregation in wireless sensor networks," *Procedia Computer Science*, vol. 10, pp. 78–85, 2012.

[19] T. Herman and S. Tixeuil, "A distributed TDMA slot assignment algorithm for wireless sensor networks," in *Algorithmic Aspects of Wireless Sensor Networks*, vol. 3121 of *Lecture Notes in Computer Science*, pp. 45–58, Springer, Berlin, Germany, 2004.

[20] H. Y. Lee and T. H. Cho, "A scheme for adaptively countering application layer security attacks in wireless sensor networks," *IEICE Transactions on Communications*, vol. 93, no. 7, pp. 1881–1889, 2010.

[21] I. Demirkol, C. Ersoy, and F. Alagöz, "MAC protocols for wireless sensor networks: a survey," *IEEE Communications Magazine*, vol. 44, no. 4, pp. 115–121, 2006.

[22] X. Yang, J. Lin, P. Moulema, W. Yu, X. Fu, and W. Zhao, "A novel en-route filtering scheme against false data injection attacks in Cyber-Physical Networked Systems," in *Proceedings of the 32nd IEEE International Conference on Distributed Computing Systems (ICDCS '12)*, pp. 92–101, June 2012.

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2015, Article ID 238230, 19 pages
http://dx.doi.org/10.1155/2015/238230

Research Article

A Novel Protective Framework for Defeating HTTP-Based Denial of Service and Distributed Denial of Service Attacks

Mohammed A. Saleh[1] and Azizah Abdul Manaf[2]

[1] *Faculty of Computing, Universiti Teknologi Malaysia (UTM), 81310 Skudai, Johor, Malaysia*
[2] *Advanced Informatics School, Universiti Teknologi Malaysia, 54100 Kuala Lumpur, Malaysia*

Correspondence should be addressed to Mohammed A. Saleh; mabbsaleh@gmail.com

Received 24 June 2014; Revised 29 August 2014; Accepted 7 September 2014

Academic Editor: Adeel Javed

The growth of web technology has brought convenience to our life, since it has become the most important communication channel. However, now this merit is threatened by complicated network-based attacks, such as denial of service (DoS) and distributed denial of service (DDoS) attacks. Despite many researchers' efforts, no optimal solution that addresses all sorts of HTTP DoS/DDoS attacks is on offer. Therefore, this research aims to fix this gap by designing an alternative solution called a flexible, collaborative, multilayer, DDoS prevention framework (FCMDPF). The innovative design of the FCMDPF framework handles all aspects of HTTP-based DoS/DDoS attacks through the following three subsequent framework's schemes (layers). Firstly, an outer blocking (OB) scheme blocks attacking IP source if it is listed on the black list table. Secondly, the service traceback oriented architecture (STBOA) scheme is to validate whether the incoming request is launched by a human or by an automated tool. Then, it traces back the true attacking IP source. Thirdly, the flexible advanced entropy based (FAEB) scheme is to eliminate high rate DDoS (HR-DDoS) and flash crowd (FC) attacks. Compared to the previous researches, our framework's design provides an efficient protection for web applications against all sorts of DoS/DDoS attacks.

1. Introduction

Historically, a series of DDoS attacks that occurred in February 2000 against Amazon, Yahoo, and eBay websites had caused an estimated cumulative loss of 1.2 billion USD. Analysts estimated that during the three hours Yahoo web site was down; it lost about 500,000 USD. According to the bookseller Amazon, the DDoS attack was a reason for losing 600,000 USD during the 10 hours of downtime. Likewise, during the DDoS attacks against eBay, eBay.com availability was degraded from 100% to only 9.4%. In January 2001, Microsoft lost approximately 500 million USD over the course of a few days from a DDoS attack on its site. In 2011, DDoS attacks devastated five high-profile websites, namely, Visa, MasterCard, Sony, WordPress, and the CIA.

Nowadays, DDoS attacks are able to launch a destructive power in a single attack. The biggest peak bandwidth of DDoS in 2010 exceeded 100 Gbps, which represents 100% increase over the size of attack in 2009. As well, peak bandwidth of DDoS in 2013 exceeded 300 Gbps, which represents over three times that of 2010 [1]. Estimates expect that the cost of a 24-hour outage for a large e-commerce company would approach 30 million USD [2, 3].

Denial of service (DoS) attack is an effort by a single machine, namely, an attacker to make a target (server or network) unavailable to its customers, which yields to forbid customers from accessing the service. DoS attack consists of highly damageable attacks to collapse or degrade the quality of service in hardly unexpected manner [4].

Distributed denial of service (DDoS) attack is an attempt to flood a victim, whether is a machine or network, through a volume of traffic that is generated by large number of machines. Furthermore, to diffuse source of attack, these machines are combined from different networks, so it is hard to trace back IP sources of attacks and then to block attacks accordingly [5, 6]. Usually, DDoS attack uses a large

number of compromised hosts called zombies or bots that are collected from unprotected computers by planting malicious software on these unprotected computers. Then, these hosts, namely, zombies or bots, are grouped together to shape one huge network called a Botnet, which awaits a command from the attacker to launch the DDoS attack [7–12].

Flash crowd (FC) is a sudden high request in a service caused by legitimate users who simultaneously request the server at the same period. Flash crowd (FC) eventually forces the server to decease its performance and takes it down completely. It occurs due to unexpected big amount of service accesses at the same time. Flash crowd (FC) overwhelms the server, and therefore it causes a denial of service (DoS) attack, which results in either a delay of response or a complete takedown. Flash crowd (FC) could happen due to some exciting event that has just occurred. Likewise, it could be due to the broadcasting of a new service or a free hot software download [2, 13, 14]. From perspectives of service requesters, regardless of whether they are legitimate or illegitimate, flash crowd (FC) may not be counted as an attack. On the contrary, it is counted as an attack from the perspectives of victim or service provider, since it has affected the web server negatively.

Low rate distributed denial of service (LR-DDoS) attack is an intelligent attack that saturates the victim with packets adequately in low rate, in order to avoid the current anomaly-based detection schemes. LR-DDoS attack has an ability to conceal its traffic, since it is identical to normal traffic. LR-DDoS attack is widely used in a large size DDoS attack, which joins several low rate attacks, such as a Botnet to initiate a low rate DDoS attack. LR-DDoS attack produces network traffic similar to the normal network traffic, and, therefore, it is difficult to be detected and mitigated [2, 10, 15].

A high rate distributed denial of service (HR-DDoS) attack is a synonym for the traditional DDoS attacks when attackers exceed and violate the adopted threshold value [15, 16].

Attacker tracing back (TB) can be defined as a method for finding out the exact true IP source of the attacker who launched DoS/DDoS attacks. Client validation (CV) is a method for verifying the validity of the service's requester to validate its legitimacy and illegitimacy and therefore to pass the former and to deny the latter [17].

Outer blocking (OB) is a mechanism for blocking (denying) attackers at the network entrance, more precisely at the Edge Router, which provides network connectivity that is resistant to spoofing attacks. In addition, it helps to save the server's resources, since the attacking IP source is blocked at the outer layer [18].

This research proposed and designed an alternative solution called a flexible, collaborative, multilayer, DDoS prevention framework (FCMDPF), which handles all aspects of HTTP-based DoS/DDoS attacks. FCMDPF framework is flexible because it eliminates the impact of flash crowd (FC) attacks gradually, while it blocks high rate HTTP DoS/DDoS (HR-DDoS) attacks immediately. In addition, it is a collaborative multilayer DDoS prevention framework because it is protecting web server against HTTP DoS/DDoS attacks at different collaborative points through which packets had gone. Each point at different framework's layer collaborates to protect web server from HTTP DoS/DDoS attacks by performing its special tests, and then it forwards the packet to the next framework's layer if it succeeds, or otherwise it will be dropped. In the same manner, the next framework's layer performs its special tests, and then it forwards the packet to the next point if it succeeds, or otherwise it will be dropped, until packet reaches the final target. FCMDPF framework comprises three subsequent multilayer points for detecting and preventing HTTP DoS/DDoS attacks. The first layer of FCMDPF framework is an outer attack blocking (OB) at the edge router, while the second layer of FCMDPF framework is service traceback oriented architecture (STBOA). The third layer of FCMDPF framework is flexible advanced entropy based (FAEB) layer.

The first layer of FCMDPF framework is an outer attack blocking (OB) scheme, which is deployed at the edge router, since it is the most nearest point to the IP attacking source. An outer blocking (OB) scheme first compares and examines the IP source of the incoming request according to its blacklist database table. Then OB scheme blocks or forwards it to the next layer of FCMDPF framework based on whether the incoming request's IP source is listed in blacklist database table at the edge router or not. In case this IP source of the incoming request is not listed on blacklist database table, OB scheme forwards it to the next layer of FCMDPF framework. Otherwise, if it is listed on the blacklist database table, OB scheme blocks it immediately, and host unreachable message will be sent to the requester. This layer provides a helpful service to the web server that all blocking processes will be done at an outer blocking layer, which helps the web server to save its recourses.

An outer attack blocking (OB) scheme is constructed by two integrated components as follows. The first component is blacklist database table, which is used by the OB scheme to record IP sources those are classified as attacking IP sources by STBOA scheme and FAEB scheme in case these IP sources failed to pass their tests. The blacklist database table is created and deployed as well on the edge router, more precisely on Quagga router, which is a part of OB_Shield subsystem. The second component is signaling technique that is used by STBOA scheme and FAEB scheme to report attacking IP sources to OB scheme and therefore to update its blacklist database table and hence to block these IP sources on upcoming requests.

The second layer of FCMDPF framework is service traceback oriented architecture (STBOA) scheme that is designed to validate whether the incoming request is launched by a human (real web browser) or by an automated tool (bots). Then, it traces back the incoming request in order to find out the true IP attacking source. Service traceback oriented architecture (STBOA) scheme is designed based on service traceback oriented architecture (STBOA) algorithm.

The third layer of FCMDPF framework is flexible advanced entropy based (FAEB) scheme, which is employed to detect anomalies in HTTP network traffic and to differentiate whether it is high rate DDoS (HR-DDoS) attacks or flash crowd (FC) attacks. Flexible advanced entropy based (FAEB) scheme is designed based on flexible advanced entropy based

(FAEB) algorithm. In case FAEB classifies that the incoming HTTP network traffic is high rate HTTP DoS/DDoS (HR-DDoS) attacks, it blocks it immediately. Whereas if FAEB classifies that the incoming HTTP network traffic is flash crowd (FC) attacks, it decreases the maximum connection's timeout value and decreases the maximum allowed request per this timeout, until these two values reach zero. Once the values of timeout and the maximum allowed requests reach zero, FAEB scheme disables KeepAlive feature of HTTP connection. Therefore, the mode is exchanged from detecting and preventing flash crowd attack to detecting and preventing high rate DDoS attack. In addition, FAEB scheme uses signaling technique to update the edge router's blacklist database.

Moreover, Section 3 in this paper provides full and granular details of FCMDPF framework. Lastly, FCMDPF framework is evaluated based on the analysis of experimental simulations, as is described in Section 4.

This paper is organized as follows. First, Section 1 introduced the interested topic, defined the relevant terms, and provided high-level description of FCMDPF framework. Then, Section 2 reviewed the previous related works. It classified existing frameworks and schemes that protect web applications from HTTP-based DoS and DDoS attacks, conducted survey on them, and identified the optimal specifications that should be offered by a protective framework to protect web applications from all sorts of HTTP-based DoS and DDoS attacks. After that, Section 3 provided full and granular details, or low-level description, of FCMDPF framework. As well, it explained the systematic procedures of the framework's evaluation. Next, Section 4 presented discussion and analysis. Finally, Section 5 concluded this paper.

2. Literature Review

This section identifies problems in the current related works and also describes the optimal framework specifications. The literature review began by classifying existing schemes and frameworks into different categories. Secondly, it conducted a comprehensive survey of detection and prevention schemes and frameworks for all sorts of HTTP-based DoS and DDoS attacks in order to show the problems with each related work. Finally, it described the optimal specifications for a protective framework against HTTP-based DoS and DDoS attacks, which fix all the shortcomings found in previous related works, as are set out below.

2.1. Classifying Existing Schemes and Frameworks. Based on extensive studies and analysis of the related works, existing schemes and frameworks can be classified into one or more of five categories. The five categories are high rate DDoS (HR-DDoS) attacks, low rate DDoS (LR-DDoS) attacks, flash crowd (FC) attacks, outer blocking (OB), and traceback and client validation (TB and CV). Various researchers [1, 19–23] highlighted that the protective scheme or framework should protect web applications from high rate DDoS (HR-DDoS) attacks, whilst other researchers suggested it should provide a protection for web applications from Low Rate DDoS (LR-DDoS) attacks [24, 26]. Other researchers [26, 27] claimed that it should provide protection against flash crowd (FC) attacks. Likewise, another group [20, 28, 29] emphasized it should be able to trace back (TB) to the true source of the attack, verify the client's validity (CV), and block it at the edge router as well (OB).

2.2. Survey on the Detective and Protective Schemes and Frameworks That Protect Web Applications from HTTP-Based DoS and DDoS Attacks. Table 1 presents a comprehensive survey of the detective and preventive schemes and frameworks that handles all sorts of HTTP-based DoS and DDoS attacks. The survey is conducted according to the five categories that are identified in the previous subsection.

2.3. Optimal Specifications for Detective and Protective Framework to Protect Web Applications from All Sorts of HTTP-Based DoS and DDoS Attacks. Based on the survey unveiled in Table 1 above and a review of the related works, the optimal specifications for a protective framework against HTTP-based DoS and DDoS attacks are required to provide full support to all of the following features.

(1) The framework should provide a protection against both HR-DDoS and FC attacks. Due to similarities between HR-DDoS and FC attacks, the framework should be able to differentiate between them clearly to block the former immediately and block the latter gradually.

(2) The framework should provide a protection against LR-DDoS attacks.

(3) The framework should provide a mechanism to verify the validity of the incoming requests whether they are legitimate (normal web browser) or illegitimate (botnet). In addition, the mechanism should be able to pass the former and block the latter. As well, the framework should provide a mechanism to find out the true attacking IP source. The mechanism here should not be designed in a way that annoys the requesters (clients) by performing extra tasks such as CAPTCHA.

(4) The framework should provide a mechanism to block the attacking IP sources at the edge router (network entrance) near to the attacking source. The benefit of this technique helps to save the resources of web servers, since the blocking occurs before incoming requests reach the web server.

(5) The framework should be designed in a way that supports the concept of separation of duties in order to prevent a single point of failure problem. The framework's components should be deployed on different layers and collaboratively work together to protect web applications from HTTP-based DoS and DDoS attacks.

(6) The framework should be compatible with existing protocols.

Table 1: Comprehensive survey on detective and preventive schemes and frameworks to all sorts of HTTP-based DoS And DDoS Attacks.

Number	Scheme/framework	Objective	Providing/protecting				
			HR-DDOS	LR-DDOS	FC	OB	TB & CV
1	Service oriented traceback architecture (SOTA) [1]	Identify attack source.	✗	✗	✗	✗	√
2	Filtering tree [2]	Protect Cloud Computing against XML and HTTP DDoS attacks.	✗	✗	✗	√	√
3	Attack source identification at router level using marking algorithm [3]	Overcome IP spoofing.	✗	✗	✗	√	√
4	Confidence based filtering (CBF) [4]	Firewall web application.	√	✗	✗	√	✗
5	A New algorithm for detecting and defending CC attacks [5]	Protect web server from CC attacks.	✗	✗	✗	✗	√
6	Intelligent decision prototype (IDP) [6]	Identify and defend attack source.	✗	✗	✗	√	√
7	Defense system for cloud computing [7]	Trace and identify the real source of DDoS attacks.	✗	✗	✗	√	√
8	EDoS-Shield [8]	Mitigate the economic denial of sustainability (EDoS) attack in the cloud computing.	✗	✗	✗	√	√
9	Diagnosis of network anomaly based on statistical traffic analysis [9]	Spot the anomalies of network Traffic.	√	✗	✗	√	✗
10	Dynamic hidden semi-Markov model (HTTP) [10]	Model the time varying user to detect web DDoS attacks.	√	✗	✗	✗	✗
11	Enhanced fast-SCTF [11]	Detect and filter distributed reflection denial of service (DRDoS) attacks.	✗	✗	✗	√	√
12	IP to hop count mapping table (IP2HC) filtering technique [12]	Defend against IP spoofing attack.	✗	✗	✗	√	√
13	Transition matrix [13]	Detect HTTP application based DDoS attacks.	√	✗	✗	✗	✗
14	Relative entropy based HTTP application DDoS detection [14]	Detect HTTP application based DDoS attacks.	√	✗	✗	✗	✗
15	Analysis of network's traffic by using IP addresses correlation [15]	Detect network DDoS attacks.	√	✗	✗	✗	✗
16	Large deviation measuring click ratio based web access behavior (LD-IID) scheme and large deviation measuring web access behavior based on Markov process (LD-MP) scheme [16]	Detect HTTP application based DDoS attacks.	√	✗	✗	✗	✗
17	Chi-square based space (CSDM) Davison method [17]	Enhance anomaly detection system accuracy.	√	✗	✗	✗	✗
18	An advanced entropy based DDoS detection scheme [18]	Determine the most suitable threshold value for detecting DDoS attacks accurately.	√	√	√	✗	✗
19	HTTP reject [19]	Block user's requests on the IP layer during DDoS attacks and keep the end user to be notified as well.	√	✗	✗	✗	✗
20	CALD [20]	Protect web server from flash crowd.	√	✗	√	√	✗
21	VicSifter [21]	Detect DDoS attacks and determine the attack's victims at an early stage.	√	✗	✗	✗	√
22	Throttling DDoS attacks [22]	Eliminate and slow down the impact of DDoS attacks against web server.	✗	✗	✗	√	√
23	An early DoS/DDoS attacks detection method based on the concept of short-term entropy [23]	Focus on the early DoS/DDoS attacks detection.	√	√	✗	✗	✗
24	A real time DDoS attacks detection and prevention system based on the analysis of per IP traffic behavior [24]	Monitor and detect DDoS attacks near to the attack's source.	√	✗	✗	√	✗

(7) The framework should be designed explicitly for processing the web application layer; HTTP protocol, rather than only the network layer; IP and ICMP protocols or the transport layer; TCP and UDP protocols.

(8) The framework should be easy to implement without causing processing and bandwidth overheads.

(9) The framework design should be dynamically able to adopt and update itself, once needed.

(10) The framework design should provide support to the hybrid scheme, which comprises the proactive and reactive schemes. The proactive scheme is required for client validation and traceback (CV and TB), whilst the reactive scheme is required for protecting against high rate DDoS (HR-DDoS) and flash crowd (FC) attacks.

(11) The framework should consume low storage memory.

(12) The framework should be resistant to IP source spoofing attacks, especially when finding out the true attacking IP sources.

3. Flexible, Collaborative, Multilayer, DDoS Prevention Framework (FCMDPF)

3.1. Design of FCMDPF Framework. As is shown previously in the literature review, all of the related works failed or at least could not protect web applications from HTTP DoS/DDoS attacks properly. The DoS/DDoS attacks are varying from high rate DoS/DDoS (HR-DDoS) attacks, low rate DoS/DDoS (LR-DDoS) attacks, and flash crowd (FC) attacks. The perfect protective solution should provide a protection from all of these mentioned attacks. In addition, it should be able to trace back attacking IP sources of the DoS/DDoS attacks in order to block them. From there, this research paper proposed and designed a comprehensive protective solution that handles all sorts of HTTP DoS/DDoS attacks called a flexible, collaborative multilayer, DDoS prevention framework (FCMDPF).

The FCMDPF framework is flexible because it eliminates the impact of flash crowd (FC) attacks gradually by decreasing the maximum connection's timeout value and decreasing the maximum allowed request per this timeout, until these two values reach zero. Once the values of timeout and the maximum allowed requests reach zero, FAEB scheme disables KeepAlive feature of HTTP connection. Therefore, the mode is exchanged from detecting and preventing flash crowd attack to detecting and preventing high rate DDoS attack. In the meanwhile, FCMDPF framework blocks high rate HTTP DoS/DDoS attacks immediately. In addition, it is a collaborative multilayer DDoS prevention framework because it protects web server against HTTP DoS/DDoS attacks at different collaborative points through which packets had gone. Each point at different framework's layer collaborates to protect web server from HTTP DoS/DDoS attacks by performing its special tests and then it forwards the packet to the next framework's layer (point) if it succeeds, or otherwise it will be dropped. In the same manner, the next framework's layer performs its special tests, and then it forwards the packet to the next point if it succeeds, or otherwise it will be dropped, until packet reaches the final target, which is the web application.

The FCMDPF framework comprises three subsequent multilayer points for detecting and preventing HTTP DoS/DDoS attacks. The first layer of FCMDPF framework is an outer attack blocking (OB) at the edge router while the second layer of FCMDPF framework is service traceback oriented architecture (STBOA). The third layer of FCMDPF framework is flexible advanced entropy based (FAEB) layer. Figure 1 illustrates the components of flexible collaborative multilayer DDoS prevention (FCMDPF) framework. As well, a protective system, namely, AntiDDoS_Shield, is developed based on FCMDPF framework.

The first layer of FCMDPF framework is an outer attack blocking (OB) scheme, which is deployed at the edge router, since it is the nearest point to the attacking IP source. An outer blocking (OB) scheme first compares and examines the IP source of the incoming request according to its blacklist database table. Then, OB scheme blocks or forwards it to the next layer of FCMDPF framework based on whether the incoming request's IP is listed in blacklist database table at the edge router or not [26–30]. In case this IP source of the incoming request is not listed on blacklist database table, OB scheme forwards it the next layer of FCMDPF framework, but if it is listed on the blacklist database table, OB scheme blocks it immediately and host unreachable message will be sent to the requester. This layer provides a helpful service to the web server, since all blocking processes will be done at an outer blocking layer, which helps the web server to save its recourses [6, 11, 21]. In addition, this research developed protective subsystem called OB_Shield, which is part of AntiDDoS_Shield system, based on OB scheme. OB_Shield subsystem employed ready configured Quagga and iproute2 routing suites. In addition, it integrated signaling technique that is used by STBOA_Shield subsystem and mod_antiddos subsystem to report back attacking IP sources to OB_Shield subsystem in order to update its blacklist database table.

An outer attack blocking (OB) scheme is constructed by two integrated components as follows. The first component is blacklist database table, which is used by the OB scheme to record IP sources that are classified as attacking IP sources by STBOA scheme or FAEB scheme in case if these IP sources failed to pass the tests of STBOA scheme or FAEB scheme. The blacklist database table is created and deployed as well on the edge router, Quagga router, which is a part of OB_Shield subsystem. The second component is signaling technique that is used by STBOA scheme and FAEB scheme to report attacking IP sources to OB scheme. Therefore, STBOA scheme and FAEB scheme update OB scheme's blacklist database table and hence block these IP sources on upcoming requests.

In fact, the outer blocking (OB) scheme, that is deployed at the edge router, besides using of signaling technique, is a new novel scheme in reference to the previous related works. The edge router's blacklist database will be updated with new IP sources in case the incoming request failed to satisfy STBOA scheme or FAEB scheme. The updating of

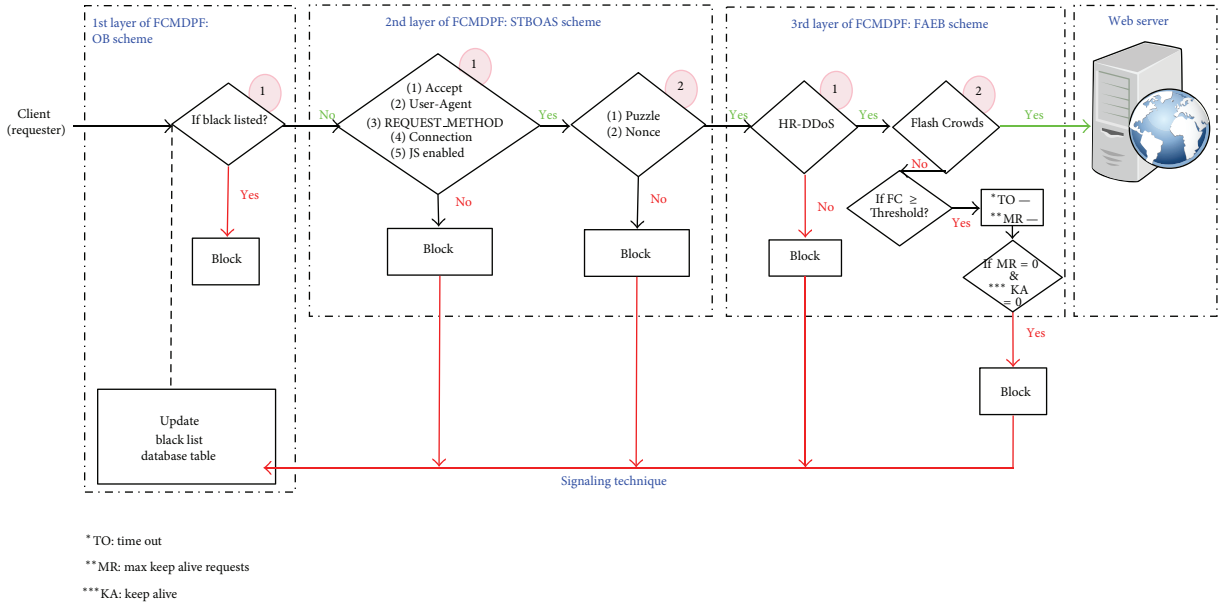


Figure 1: Components of flexible collaborative multilayer DDoS prevention framework (FCMDPF).

new IP source leads to judging that the incoming IP source is an attacker. Then, the FCMDPF framework uses signaling technique to classify the attacking IP source, so that IP source will be traced back, black listed, and blocked, as well.

The second layer of FCMDPF framework is service traceback oriented architecture (STBOA) scheme that is designed to validate whether the incoming request is launched by a human (real web browser) or by an automated tool (bots). Then, it traces back the incoming request in order to find out the true attacking IP source. Service traceback oriented architecture (STBOA) scheme is designed based on service traceback oriented architecture (STBOA) algorithm, as it is shown in Algorithm 1. Figure 2 demonstrates how STBOA scheme processes, treats, and validates the incoming requests. In addition, this research developed subsystem called STBOA_Shield, which is part of AntiDDoS_Shield system, based on STBOA scheme in order to validate and trace back the attacking source. STBOA_Shield subsystem is a web application that is developed by using SOAP and PHP scripting programming language.

First off, service traceback oriented architecture (STBOA) scheme validates the incoming request to determine whether the request is legitimate, which is launched by legal user, or illegitimate one that is executed by an automated tool, such as a bot. The purpose of this process is to identify illegitimate requests that are launched by IRC bots in order to block them immediately. After that, STBOA scheme traces back the incoming request in order to find out the true attacking IP source. STBOA scheme accomplishes these missions through different two subsequent stages.

In the first stage, STBOA scheme validates the incoming request by checking the request's header looking for unique header's values, which are carried out only by legitimate request, such as web browsers [31–37]. Algorithm 2 presents the unique header's values that are checked by STBOA scheme. STBOA scheme checks for "User-Agent", "Accept", and "Host" headers in HTTP connection, and it makes sure that the requester (client) has enabled its own Javascript language engine. Furthermore, it checks for "REQUEST_METHOD" header's value whether it is GET, HEAD, or POST [31, 38–44]. If the incoming request passes all of these checks, it will proceed to the next test. Otherwise, it will be terminated and blocked immediately, and the signal is sent to the edge router through signaling technique, so that the client will be blocked at the edge router.

In the second stage, STBOA scheme utilizes web service technology to formulate and generate a puzzle, random number, and nonce value. Then, it sends them back to the client or the requester [32–37, 45–49]. The client has to solve a puzzle by using a random number that is sent by web server (web application). Then, the client sends back the solved puzzle (puzzle's answer), along with the nonce value. After that, the web server (web application) will verify puzzle's answer and nonce value that are sent by the client whether they are correct or not. If both numbers are correct, the request will be forwarded for the next test. Otherwise it will be blocked immediately and a signal is sent back to the edge router to update its black list. Figure 3 demonstrates how STBOA scheme utilizes web service technology to validate a client.

In this research paper, formulating and generating a puzzle and nonce value are done based on web service by STBOA scheme. It is the most appropriate and preferred solution compared with the other similar solutions, such as CAPTCHA, which annoys the clients [13]. It is preferred because it does not burden a client to solve a puzzle and send the answer back to web server, since a legitimate web browser does this mission without client interception. In addition, the nonce value plays a significant role, since it is used for an extra verification purpose to ensure that the right

```
Begin:
Declare string variable called Usr
Declare string variable called Acpt
Declare string variable called Conn
Declare string variable called Cokie
Declare string variable called Rmeth
  Declare an integer variable called X
set Usr to $hdr['User-Agent']
set Acpt to $hdr['Accept']
set Conn to $hdr['Connection']
set Cokie to $_COOKIE['cookie']
set Rmeth to $_SERVER['REQUEST_METHOD']
set X to GenX()
  repeat
    set Usr not null, set Acpt not null and set Conn not null
    set resp_code 200
    repeat
     Cokie set not null
        repeat
         set Rmeth not null
         calculate Puzz_Solve()
         calculate Puzz_verify()
         calculate GenNonce()
         calculate VerifyNonce()
         if Puzz_Solve <> Puzz_verify
                          EdgeRouter_Bsn(src_ip)
          set resp_code 403
         if GenNonce <> VerifyNonce
                          EdgeRouter_Ban(src_ip)
         set resp_code 403
        until Rmeth set null
    until Cokie set null
until Usr set null or Acpt set null or Conn set null
End.
```

Algorithm 1: Service traceback oriented architecture (STBOA) algorithm.

```
$hdr['User-Agent']
$hdr['Accept']
$hdr['Host']
$_COOKIE['cookie']
$_SERVER['REQUEST_METHOD']
```

Algorithm 2: Request's headers checked by the STBOA algorithm.

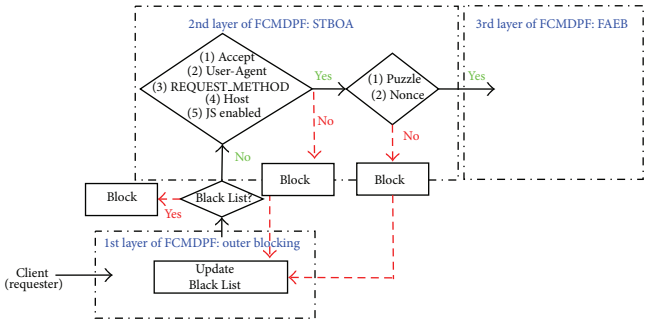


Figure 2: STBOA scheme processes and validates the incoming requests.

client answered the puzzle. STBOA scheme formulates and generates a puzzle, random number, and nonce value based on formulas in Algorithm 3.

Service traceback oriented architecture (STBOA) scheme, that is deployed in the second layer of FCMDPF framework, is an expansion and modification to the previous works done by Subbulakshmi et al. [1], Yang et al. [7], Mohan and Raji Reddy [12], Wang et al. [21], and Darapureddi et al. [22] in order to validate whether the request is launched by a human or an automated tool, such as a bot. STBOA scheme adopted and employed extra specifications such as supporting cookie by the client, calculating, solving some random puzzles, and sending valid requests to ensure the request launched by a human.

The third layer of FCMDPF framework is flexible advanced entropy based (FAEB) scheme, which is employed to detect anomalies in HTTP network traffic and to differentiate whether it is high rate DDoS (HR-DDoS) attack or flash crowd (FC) attack. Flexible advanced entropy based (FAEB)

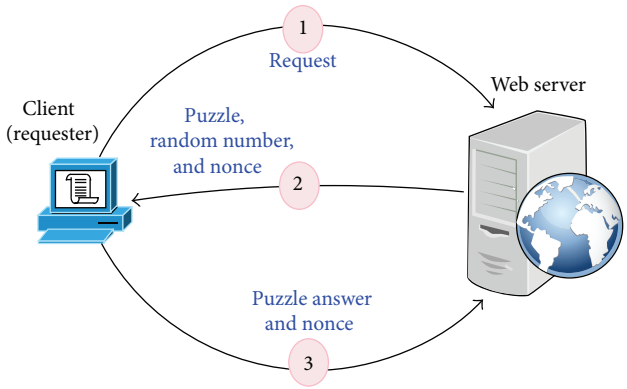


Figure 3: Conceptual steps of the STBOA web service puzzle for validating clients.

```
x = rand()
puzzle = (2 ∗ pow (2, x)) + (2 ∗ x) − 9
nonce = md5('fgwlcx'. time())
```

Algorithm 3: STBOA algorithm formulas to generate puzzle, random number, and nonce value.

scheme is designed based on flexible advanced entropy based (FAEB) algorithm, as it is shown in Algorithm 4. In case FAEB scheme classifies that the incoming HTTP network traffic is high rate HTTP DoS/DDoS (HR-DDoS) attack, it blocks it immediately, whereas if FAEB scheme classifies that the incoming HTTP network traffic is flash crowd (FC) attack, it decreases the maximum connection's timeout value and it decreases the maximum allowed request per this timeout, until these two values reach zero. Once the values of timeout and the maximum allowed requests reach zero, FAEB scheme disables KeepAlive feature of HTTP connection. Therefore, the mode is exchanged from detecting and preventing flash crowd attack to detecting and preventing high rate DDoS attack. In addition, FAEB scheme uses signaling technique to update the edge router's blacklist database. Figure 4 demonstrates how FAEB scheme processes, treats, and verifies the incoming requests.

First off, FAEB scheme examines the incoming request to determine whether the request belongs to white list table or blacklist table of Apache web server, as it is explained in Algorithm 5. In the former case, if the IP source of the incoming request belongs to IP sources in white list, it will be excluded from checking, and it always be allowed for accessing the web server (web application), while in the latter case, if the IP source of the incoming request belongs to IP sources in black list, it will be blocked immediately, and a signal is sent to update blacklist database table of the edge router.

After that, FAEB scheme examines an Apache web server to figure out whether it is under high rate DDoS attack and flash crowd attack or in normal situation. FAEB scheme does so periodically based on the adopted time in the module's configuration by calculating entropy of overall requests through the following formulas [50–55]:

$$\text{entropy} = \text{Pi} * \log_2 \text{Pi},$$
$$\text{Pi} = \frac{\text{uri\_counts}}{\text{total\_counts}}. \quad (1)$$

Then, FAEB scheme compares the computed result of entropy with the thresholds value of high rate DDoS (HR-DDoS) and flash crowds (FC) attacks that are adopted during the system's profiling. If FAEB scheme determines that Apache web server is under high rate DDoS attacks, it then blocks all requests that shared and participated in attacks, and it reports them to the edge router in order to update its black list, while if FAEB scheme determines that an Apache web server is under flash crowds attack, it decreases the maximum connection's timeout value and it decreases the maximum allowed request per this timeout, until these two values reach zero. Once the values of timeout and the maximum allowed requests reach zero, FAEB scheme disables KeepAlive feature of HTTP connection. Therefore, the mode is exchanged from detecting and preventing flash crowd attack to detecting and preventing high rate DDoS attack. Once the calculated entropy exceeds the maximum threshold's value of flash crowd attack, it then blocks all incoming requests that participated in attack. Then, it reports them to the edge router in order to update its black list through signaling technique. Whereas the Apache web server is neither under high rate DDoS, nor under flash crowds, it is considered under normal situation. Therefore, the incoming requests are treated as legitimate requests.

The FAEB scheme detects and prevents flash crowd (FC) attack by calculating the entropy of incoming requests that are launched towards hot pages of the website. The reason behind choosing hot web pages to simulate the flash crowd (FC) attack is that the legitimate users suddenly launch a large number of requests towards the web application. Indeed, these requests eventually overwhelm the server, and, therefore, they cause a denial of service (DoS) attack, which results in either a delay of response or a complete takedown [2, 13, 14, 29, 38]. The FAEB scheme detects and prevents flash crowd (FC) attacks by calculating the entropy based on the flash crowd attack entropy algorithm, as is shown in Algorithm 6.

The flash crowd (FC) attack entropy algorithm first calculates clicks' average of the hot web pages, and if it exceeds 10000, as depicted in Figure 5, it starts to calculate the entropy [43]. The reason behind 10000 clicks' average on web pages condition is that the entropy compares the calculated value to threshold, which is classified as long-term entropy based on [23]. If the calculated entropy is outside the range $-0.5 < H < +0.5$, it indicates that a flash crowd (FC) attack is taking place. Otherwise, it is not flash crowd (FC) attack.

Likewise, the FAEB scheme detects and prevents the HR-DDoS attack by calculating the entropy of incoming requests that are launched towards cold pages of a website [2, 13, 14, 29, 38]. The FAEB scheme detects and prevents the HR-DDoS attack by calculating the entropy of incoming requests based

```
Begin:
  Declare flag called AntiDDoSEnabled
  Declare flag called AntiDDoSForced
  Declare an integer called AntiDDoSPeriod
  Declare an integer called AntiDDoSHDDoS
  Declare an integer called AntiDDoSFlashCrowd
  Declare an integer called AntiDDoSLDDoS
  Declare an integer called AntiDDoSNormal
  Declare string called AntiDDoSCommand
  Declare string called WhiteList
    WHILE uri <> NULL
      IF main_request <> NULL OR prev_request <> NULL
        return DECLINED
      ENDIF
      IF AntiDDoSEnabled = ANTIDDOS_DISABLED
        return DECLINED
      ENDIF
      IF AntiDDoSForced = ANTIDDOS_DISABLED
        return DECLINED
      ENDIF
      IF handler = modantiddos-handler
        return antiddos_viewer
      ENDIF
      IF is_WhiteListed(remote_ip)
        return OK
      ENDIF
      IF is_BlackListed(remote_ip)
        return HTTP_FORBIDDEN
                        EdgeRouter_Ban(remote_ip)
      ENDIF
      set entropy_result = OK
      IF AntiDDoSForced = ANTIDDOS_DISABLED
                        t = time_now − AntiDDoSPeriod
        IF (uri_time < t)
         Calculate entropyUri (uri)
                        ENDIF
      IF entropyUri = AntiDDoSHDDoS
        return HTTP_FORBIDDEN
                        Blacklist_Table(remote_ip)
                        EdgeRouter_Ban(remote_ip)
      ENDIF
      IF entropyUri = AntiDDoSFlashCrowd
                        decrease timeout
                        decrease MaxKeepAliveRequests
                                IF (timeout == 0 && KeepAlive == Off)
                    return HTTP_FORBIDDEN
                                Blacklist_Table(remote_ip)
                                EdgeRouter_Ban(remote_ip)
                                ENDIF
      ENDIF
    ENDIF
    ENDIF
    return entropy_result
  ENDWHILE
End.
```

Algorithm 4: Flexible advanced entropy based (FAEB) algorithm.

```
Begin:
  IF is_WhiteListed(remote_ip)
    return OK
  ENDIF
  IF is_BlackListed(remote_ip)
    return HTTP_FORBIDDEN
            EdgeRouter_Ban(remote_ip)
  ENDIF
End.
```

Algorithm 5: Whitelist and blacklist checking in the FAEB algorithm.

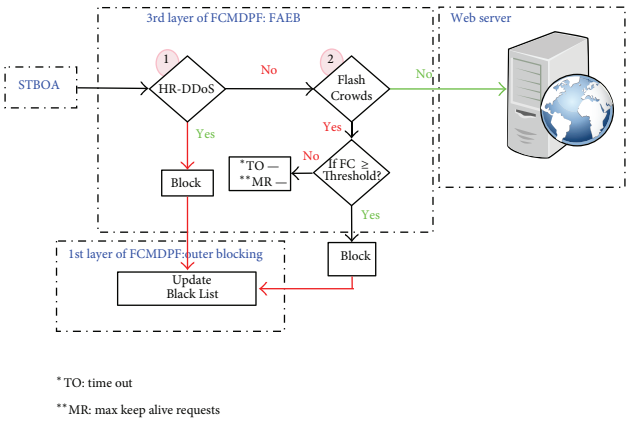


Figure 4: FAEB scheme verifies the incoming requests.

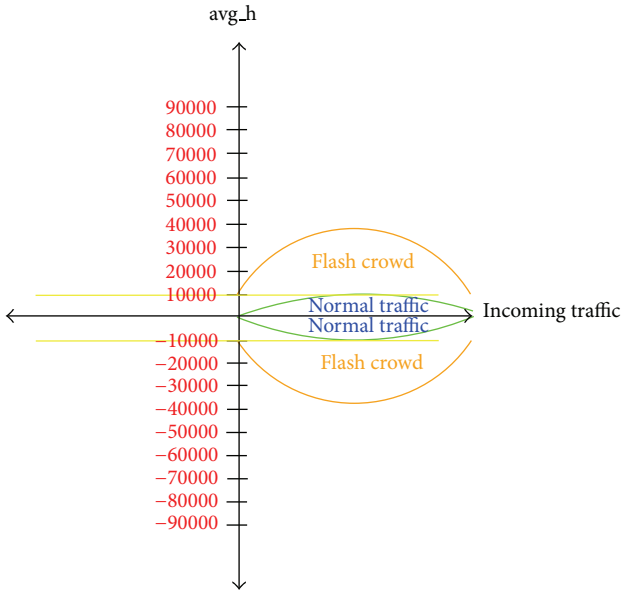


Figure 5: Clicks' average of the hot web pages.

on the high rate attack entropy algorithm, as is shown in Algorithm 7.

The high rate attack entropy algorithm first calculates the clicks' average for the cold web pages, and if it exceeds 10000, as shown in Figure 6, it starts to calculate the entropy [43]. The reason behind 10000 clicks' average on web pages condition is that the entropy compares the calculated value to threshold, which is classified as long-term entropy based on [23]. If the calculated entropy is outside the range $-1.36 < H < +1.36$, it indicates that a HR-DDoS attack is taking place. Otherwise, it is not high rate attack.

Finally, the FAEB scheme blocks incoming requests that represents high rate DDoS attack (HR-DDoS) immediately, while it blocks incoming requests that represents a flash crowd (FC) attack gradually by decreasing the maximum connection's timeout value and decreasing the maximum allowed request per this timeout, until these two values reach zero. Once the values of timeout and the maximum allowed requests reach zero, FAEB scheme disables KeepAlive feature of HTTP connection. Therefore, the mode is exchanged from detecting and preventing flash crowd (FC) attack to detecting and preventing high rate DDoS (HRDDoS) attack. In addition, the FAEB scheme feeds back those blocked IP sources to outer attack blocking (OB) scheme's blacklist database table through signaling technique. Otherwise, the FAEB scheme passes incoming requests that represent a normal traffic. Besides that, this research developed subsystem called mod_antiddos Apache module, which is part of AntiDDoS_Shield system, based on FAEB scheme. The mod_antiddos subsystem is an Apache web server module, which is programmed by using Apache APR library and C programming language.

Figure 7 presents the overall entropy for the three different cases: flash crowd attack case, high rate DDoS attack case, and normal traffic case.

Flexible advanced entropy based (FAEB) scheme, that is deployed in third layer of FCMDPF framework, is an expansion to the previous related works done by Chonka et al. [6], Xie and Tang [10], (Zheng et al. [18]), Wen et al. [20], Schweizer [56], Ciufo [57], Monshouwer [58], Ye and Zheng [13], Sqalli et al. [8], and SpiderLabs ModSecurity [59]. FAEB scheme provides an ideal protective solution for high rate HTTP DoS/DDoS (HR-DDoS) and flash crowd (FC) attacks smoothly by blocking high rate DoS/DDoS attacks immediately, while blocking flash crowd attacks gradually. On the other hand, offering a protective solution against the low rate HTTP DoS/DDoS (LR-DDoS) attacks is ignored intentionally in this research, since such a protection now is available by default in all recent web servers. Despite this fact, FAEB is able to provide a protection against the low rate HTTP DoS/DDoS (LR-DDoS) attacks.

3.2. Evaluating FCMDPF Framework. In this research paper, evaluating FCMDPF framework is done based on simulation of practical experiments of the AntiDDoS_Shield system, which is developed based on FCMDPF Framework, and the analysis of corresponding experimental results. As is explained earlier, the AntiDDoS_Shield system is developed in this research, as well. In fact, four different types of experiments are launched for testing and evaluating the AntiDDoS_Shield system. The first type of experiment is to

```
calculate avg_h = (hot_uri1_clicks + hot_uri2_clicks + ··· + hot_uriN_clicks)/N
IF (avg_h >= 10000)
      calculate entropy H = −H(hot_uri1) + −H(hot_uri2) + ··· + −H(hot_uriN)
  IF (H <= −0.5 OR H >= +0.5)
        decrease Timeout
        decrease MaxKeepAliveRequests
              IF (Timeout == 0 && MaxKeepAliveRequests == 0)
                    return HTTP_FORBIDDEN
                    Blacklist_Table(remote_ip)
                    EdgeRouter_Ban(remote_ip)
              ENDIF
  ENDIF
ENDIF
```

Algorithm 6: Flash crowd attack entropy algorithm.

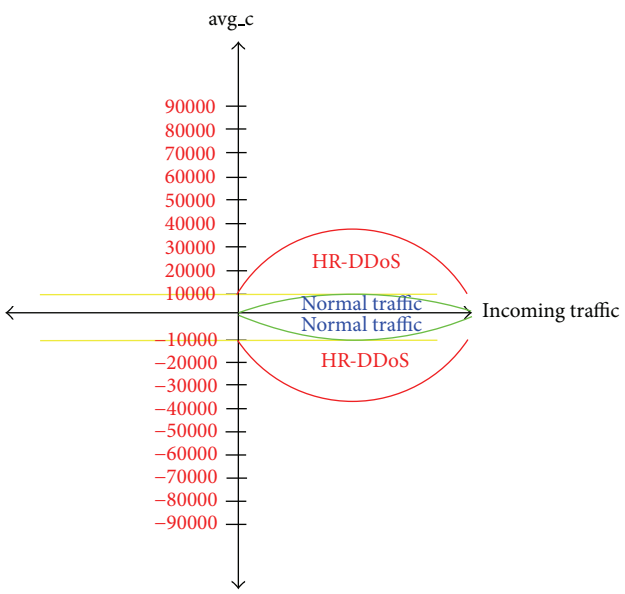


Figure 6: Clicks' average of the cold web pages.

test and evaluate the work for providing a protection against flash crowd (FC) attacks, while the second type of experiment is to test and evaluate the work for providing a protection against high rate DDoS (HR-DDoS) attacks. The third type of experiment is to test and evaluate the work for validating the client and tracing back the true IP source of attack. The fourth type of experiment is to test and evaluate the concerned work for blocking attacking IP source, as nearest as possible to attacking IP source, at the edge router, network entrance.

Simulation of flash crowd (FC) attack's experiment is accomplished by launching huge number of distributed incoming requests against hot pages of the website (hot web pages). Several studies proved that hot web pages represent about 10 percent of whole web pages [14, 16]. The reason behind choosing hot pages to simulate flash crowd (FC) attack is that the legitimate users launch sudden high requests for accessing them. It eventually overwhelms the server and, therefore, causes a denial of service (DoS) attack, which results in either a delay of response or a complete takedown [4, 60, 61], while, simulating high rate (HR-DDoS) DDoS attack's experiment is carried out by launching large number of distributed requests against cold pages of website, whereas the third and fourth experiments are accomplished by simulating incoming legitimate and illegitimate requests towards an Apache web server.

The simulation environment is constructed by using virtualization technology to include all of the needed vectors and players. As is explained in Figure 8 from right to left, it consists of:

(i) The Apache web server serves the incoming requests and responds to them accordingly. Two of developed subsystems, namely, STBOA_Shield subsystem and mod_antiddos subsystem, are installed on Apache and are configured, as well.

(ii) The Quagga and iproute2 routing suites software are employed on the edge router at the entrance of the network. The main objective of these two tools is to permit or deny network traffic routing to inside and outside of the network.

(iii) There are web application legitimate clients (customers) and attackers.

The ways to know and therefore to detect whether there is an attack or not are through one of the following ways.

(i) The incoming request fails to pass the validation tests successfully. The validation's tests require that the incoming request carries out on its HTTP header all of the following pair's values: User-Agent, Accept, Host, and REQUEST_METHOD, and Javascript language engine is enabled. As well, if the incoming request fails to answer the generated puzzle successfully, it is flagged as an attack.

(ii) The attacker launches huge volume of incoming requests that are headed against the hot web pages of the web application, which cause flash crowd (FC) attack.

(iii) The attacker launches huge volume of incoming requests that are headed against the cold web pages

```
calculate avg_c = (cold_uri1_clicks + cold_uri2_clicks + ··· + cold_uriN_clicks)/N
IF (avg_c >= 10000)
        calculate entropy H = −H(cold_uri1) + −H(cold_uri2) + ··· + −H(cold_uriN)
    IF (H <= −1.36 OR H >= +1.36)
            return HTTP_FORBIDDEN
            Blacklist_Table(remote_ip)
            EdgeRouter_Ban(remote_ip)
    ENDIF
ENDIF
```

Algorithm 7: High rate attack entropy algorithm.

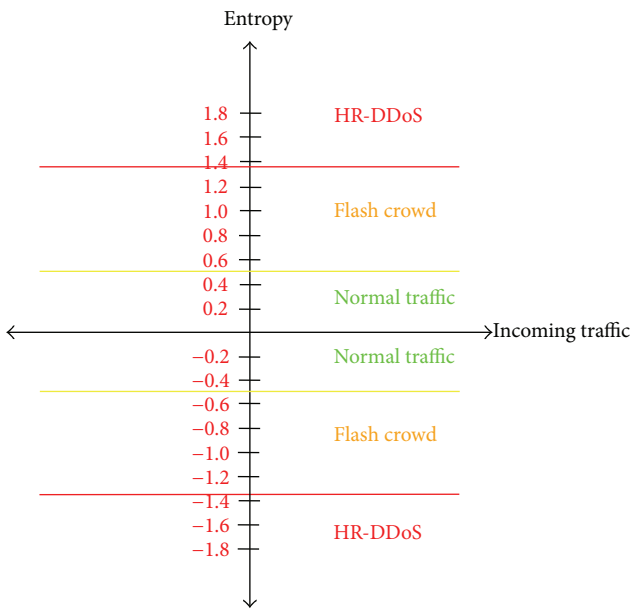


Figure 7: The overall entropy for the flash crowd (FC) attack case, high rate DDoS (HR-DDoS) attack case, and a normal traffic case.

of the web application, which cause high rate DDoS (HR-DDoS) attack.

(iv) The edge router at the network entrance fails to inspect the incoming requests. Therefore, it fails to detect and prevent the attacking IP sources at this point before the incoming request traverses or move to the inside (private) network.

The simulation model, as is shown in Figure 9, is conducted for each one of the four different simulations through the five following steps.

(1) Jmeter-Client opens the corresponding distributed testing plan, one plan out of four plans that are listed below for each simulation, and then it sends a command to Jmeter-Servers:

(i) client validation and traceback testing Plan.jmx;
(ii) flash crowd attack testing Plan.jmx;
(iii) high rate DDoS attack Plan.jmx;
(iv) edge router outer blocking testing Plan.jmx.

(2) Jmeter-Servers simulate the required incoming requests based on the distributed testing plan, which is sent by Jmeter-Client.

(3) The OB_Shield subsystem, Quagga Router, first checks and examines the IP source of the incoming request based on its blacklist database table. If this IP source is listed on the black list database table, OB_Shield blocks it immediately by responding back to the requester with host unreachable message. Otherwise, it forwards it to Apache web server.

(4) The STBOA_Shield subsystem validates the incoming request to ensure that is launched by a human not by an automated tool like botnet. If this incoming request succeeds to pass this test, which is launched by a human, it will proceed to the next test of STBOA_Shield subsystem. Otherwise, STBOA_Shield subsystem blocks it immediately by responding back to the requester with HTTP_FORBIDDEN message. Then, STBOA_Shield subsystem sends back a puzzle to the requester, which the requester needs to solve it correctly. If the requester passes this test too, it will proceed to the next test of mod_anitddos subsystem. Otherwise, STBOA_Shield subsystem blocks it immediately by responding back to the requester with HTTP_FORBIDDEN message. As well, when STBOA_Shield subsystem responds to the requester with HTTP_FORBIDDEN message, it reports the attacking IP sources to OB_Shield subsystem in order to update its blacklist database table through signaling technique.

(5) The mod_antiddos subsystem checks Apache web server to determine whether it is under high rate DDoS attack and flash crowd attack or in normal situation. If Apache is under high rate DDoS attack, the mod_antiddos subsystem blocks incoming requests immediately by responding back to the requesters with HTTP_FORBIDDEN message. While if it is under flash crowd attack, the mod_antiddos subsystem blocks incoming requests gradually by decreasing the maximum connection's timeout value and decreasing the maximum allowed request per this timeout, until these two values reach zero. Once the values of timeout and the maximum allowed requests reach zero, FAEB scheme disables

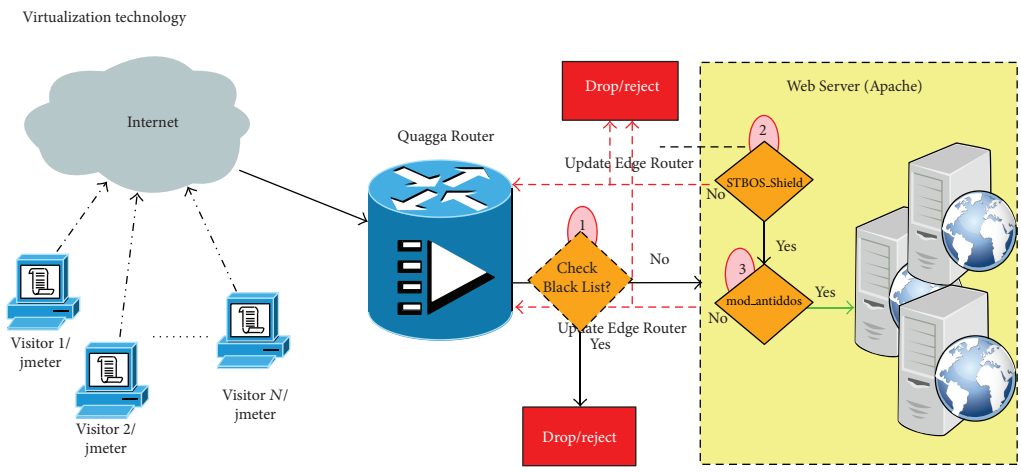


Figure 8: The construction simulation environment.

KeepAlive feature of HTTP connection. Therefore, the mode is exchanged from detecting and preventing flash crowd (FC) attack to detecting and preventing high rate DDoS (HRDDoS) attack. In the meanwhile, FCMDPF framework blocks high rate HTTP DoS/DDoS attacks immediately. Otherwise, the incoming request accesses its final target. As well, when the mod_antiddos subsystem responds to the requester with HTTP_FORBIDDEN message, it reports the attacking IP sources to OB_Shield subsystem in order to update its blacklist database table through signaling technique.

The parameters or variables that are measured during the simulation of the four experiments for testing and evaluating the work are as follows.

(1) There is the quantity of incoming requests that are simulated by Jmeter-Servers in each one out of the four different simulations.

(2) The quantity of the detected and prevented attacks that are launched against the web application: in the first three simulations, the quantity of the detected and prevented attacks is gathered from Apache log file for the work. The work should respond to the incoming request with HTTP response code number 403 (forbidden), 500 (internal server error), or 503 (service unavailable), depending on the behaviour of the work. On the contrary, once the work responds to the incoming request with HTTP response code number 200, it indicates that the work failed to detect and prevent the attack, while in the last simulation, the quantity of the detected and prevented attacks is gathered from network traffic that is captured and saved in.pcap format for the work. Once the work detects and prevents the attack, it responds to the requester with host unreachable message or otherwise the work failed to detect and prevent the attack.

The actual simulation model generated 420000 incoming requests for each simulation because it is the minimum required number that causes DoS/DDoS attack [2, 20, 23]. Therefore, it generated 420000 incoming requests in order to test and evaluate blocking and tracing back the attacking IP sources at the edge router. As well, it generated 420000 incoming requests in order to test and evaluate validation of the incoming requests. Besides that, it generated 420000 incoming requests in order to test and evaluate detecting and preventing high rate HTTP DoS/DDoS (HR-DDoS) attack. In addition, it generated 420000 incoming requests in order to test and evaluate detecting and preventing flash crowd (FC) attack.

The AntiDDoS_Shield system detected and prevented all high rate HTTP-based DoS/DDoS (HR-DDoS) attacks, which were 420000 high rate HTTP-based DoS/DDoS attacks, through mod_antiddos module subsystem. The mod_antiddos subsystem first calculated clicks' average of the cold web pages that exceeded 10000. Then, it calculated the entropy, which was out of $-1.36 < H < +1.36$ range. Hence, the mod_antiddos subsystem indicated that the web server is under high rate DDoS attack. Therefore, it blocked all incoming requests immediately by replying to them with HTTP_FORBIDDEN message or HTTP response code number 403. Then, it updated OB_Shield with these attacking IP sources, as well.

In addition, the AntiDDoS_Shield system detected and detected and prevented 369726 out of 420000 flash crowd (FC) attacks through mod_antiddos module subsystem. The mod_antiddos subsystem first calculated clicks' average of the hot web pages that exceeded 10000. After that, it calculated the entropy, which was out of $-0.5 < H < +0.5$ range. Hence, the mod_antiddos subsystem indicated that the web server is under flash crowd (FC) attack. Therefore, it blocked all incoming requests gradually by decreasing the maximum connection's timeout value and decreasing the maximum allowed request per this timeout to the half, until these two values reached zero. Then, the mod_antiddos subsystem disabled KeepAlive feature of HTTP connection, and, therefore, the detection and prevention mode is exchanged from flash crowd attack to high rate DDoS attack. Thus, the mod_antiddos subsystem blocked flash crowd attacks gradually by replying to them with HTTP_FORBIDDEN

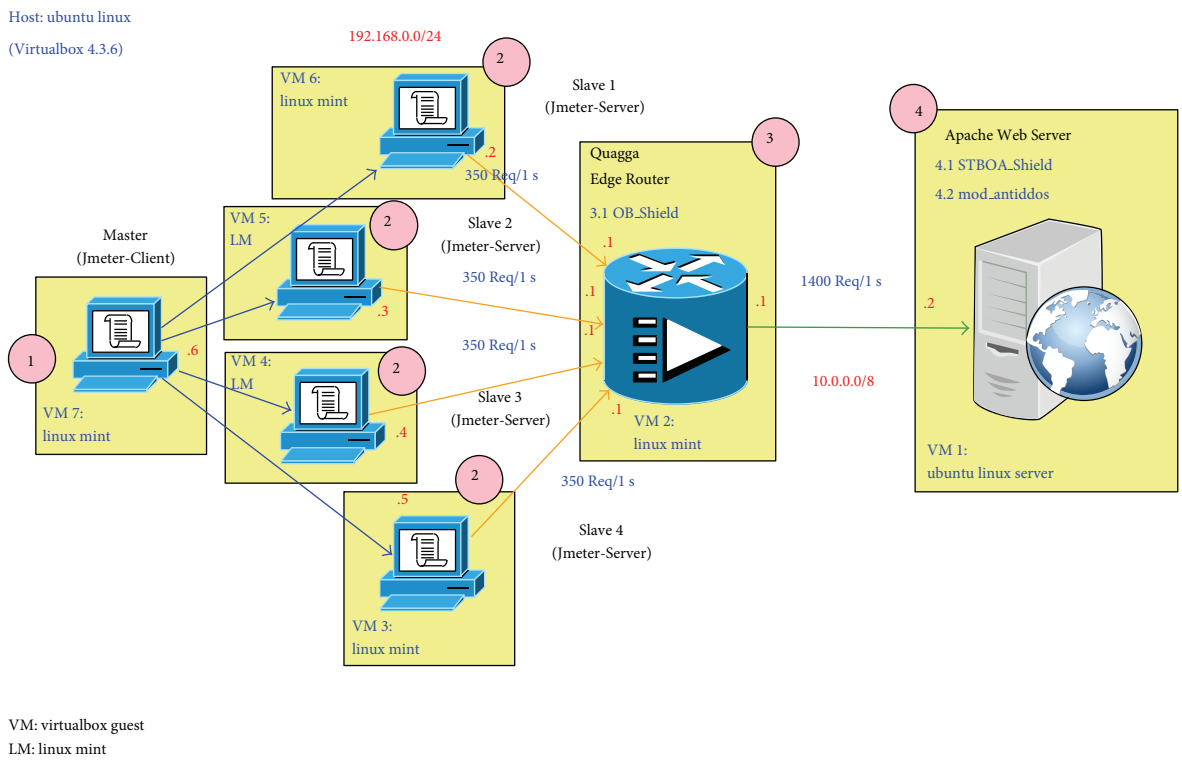


Figure 9: Sequential steps of simulation model.

message or HTTP response code number 403. Lastly, it updated OB_Shield with these attacking IP sources, as well.

As well, the AntiDDoS_Shield system succeeded to detect and prevent all of those attacking IP sources, which were 420000 IP sources of incoming requests, at the edge router, namely, Quagga edge router by replying to the requester with host unreachable message.

Besides that, AntiDDoS_Shield system succeeded to validate and trace back 369726 out of 420000 incoming requests by STBOA_Shield subsystem. It validated incoming requests that are missing User-Agent header, Accept header, Host header, REQUEST_METHOD header, or disabled Javascript language engine. Besides that, it traced back incoming requests through a puzzle.

4. Discussion and Analysis

Finally, this section discusses and evaluates our proposed and designed protective framework for defeating HTTP-based DoS/DDoS attacks, namely, the flexible, collaborative, multilayer, DDoS prevention framework (FCMDPF). The FCMDPF framework is evaluated based on the optimal specifications for a protective framework to protect web applications from all sorts of HTTP-based DoS and DDoS attacks that are outlined above. Table 2 provides all evaluation details.

5. Conclusions

Although many researchers focused on proposing and designing robust schemes and frameworks for protecting web applications from all sorts of HTTP-based DoS/DDoS attacks, there are still open issues that need to be addressed, as are described previously in the literature review. This research paper proposes and designs a novel protective framework for defeating HTTP-based DoS/DDoS attacks, namely, the flexible, collaborative, multilayer, DDoS prevention framework (FCMDPF). The novelty of this framework's design fixes and overcomes all the shortcomings of the previous related works. It provides a novel alternative protective framework to protect web applications from all sorts of HTTP DoS/DDoS attacks, such as high rate DDoS (HR-DDoS) and flash crowd (FC). In addition, it is quite able to validate and trace back (TB and CV) the real attacking IP sources and block them at the edge router (OB), as well. Finally, the FCMDPF framework is evaluated based on the optimal specifications for a protective framework to protect web applications from all sorts of HTTP-based DoS and DDoS attacks, as are outlined above. The evaluation is unveiled in Table 2 above. It proved that the FCMDPF framework was successful in providing the full support to all the optimal specifications. In contrast, it suffers from low rate of false negatives, since it was not able to detect and prevent all of flash crowd (FC) attacks. As well, it failed to validate and trace back some of incoming requests. Future research will hopefully improve the accuracy rates of the FCMDPF framework.

Table 2: Evaluating FCMDPF framework based on the optimal specifications for a protective framework to protect web applications from all sorts of HTTP-based DoS And DDoS attacks.

Framework specifications/FCMDPF layers	OB layer	STBOA layer	FAEB layer	Remarks
(1) The framework should provide a protection against high rate DDoS (HR-DDoS) and flash crowd (FC) attacks, as well. It should be able to differentiate between them clearly to block the former immediately and block the latter gradually.	✗	✗	√	The FAEB scheme of the FCMDPF framework is quite able to differentiate between HR-DDoS and FC attacks precisely. Hence, it is able to provide the protection for web applications against them properly through the FAEB algorithm, flash crowd attack entropy algorithm, and high rate attack entropy algorithm, as are described previously. The mod_antiddos module subsystem, which is developed based on FAEB scheme, detected and prevented all high rate HTTP-based DoS/DDoS (HR-DDoS) attacks. As well, it detected and prevented 369726 out of 420000 flash crowd (FC) attacks.
(2) The framework should provide a protection against low rate DDoS (LR-DDoS) attacks.	✗	✗	√	Despite the FAEB scheme of the FCMDPF framework and therefore the mod_antiddos module subsystem, being able to protect web applications from LR-DDoS attacks, this protection is excluded intentionally in this research. It is excluded because the protection from LR-DDoS attacks is provided in all recent web servers by default.
(3) The framework should provide a mechanism to verify the validity of the incoming requests. As well, it should provide a mechanism to find out the true attacking IP source. Besides that, it should not be designed in a way that annoys the requesters by performing extra tasks, such as CAPTCHA.	√	√	✗	(i) The STBOA scheme of the FCMDPF framework and therefore STBOA_Shield subsystem, which is developed based on STBOA scheme, is quite able to verify the validity of an incoming request. The STBOA scheme verifies it through the STBOA algorithm to identify if it is legitimate or illegitimate and, therefore, subsequently to pass the former and block the latter. As well, it provides a mechanism that is quite able to trace back and find out the true attacking IP source in a way that does not burden or annoy the requester. In particular, the second phase of the STBOA scheme utilizes web services to send back a puzzle to the requester. In case the requester is a human using a real web browser (not a bot), he will answer this puzzle automatically by the browser itself without human interaction. Then, he will send back the answer to the web application. After that, the web application verifies (examines) the answer, and if it is correct, it passes it to the next layer or otherwise it blocks it immediately and updates the OB scheme blacklist database table with this attacking IP source. (ii) The STBOA_Shield subsystem, which is developed based on STBOA scheme, succeeded to validate and trace back 369726 out of 420000 incoming requests. (iii) The OB scheme then collaborates to block those updated attacking IP sources in upcoming incoming requests.
(4) The framework should provide a mechanism to block the attacking IP sources at the edge router near to the attacking source.	√	✗	✗	(i) The outer blocking (OB) scheme of the FCMDPF framework is quite able to block the attacking IP source that neither passes the STBOA scheme's tests, nor passes the FAEB scheme's tests, at the Edge Router (Network Entrance). (ii) The OB_Shield subsystem succeeded to detect and prevent all of those attacking IP sources, which were 420000 IP sources, at the edge router.

Table 2: Continued.

Framework specifications/FCMDPF layers	OB layer	STBOA layer	FAEB layer	Remarks
(5) The framework should be designed in a way that supports the separation of duties concept.	√	√	√	(i) The FCMDPF framework is a collaborative, multilayer, DDoS prevention framework because it protects web applications against HTTP DoS/DDoS attacks at the different collaborative points through which the incoming requests have gone. (ii) Each point at different framework layers collaborates to protect web applications from HTTP DoS/DDoS attacks by performing its special tests. Then, it forwards the request to the next framework's layer if it succeeds, or otherwise it will be dropped immediately. (iii) In the same manner, the next framework's layer performs its special tests, and then it forwards the packet to the next point if it succeeds, until it reaches the target. Otherwise, it will be dropped immediately.
(6) The framework should be compatible with the existing protocols.	√	√	√	(i) The entire FCMDPF framework's layers, the OB layer, the STBOA layer, and the FAEB layer are compatible with existing protocols. (ii) Indeed, the OB layer is compatible with the IP, TCP, and UDP protocols. The OB layer merely uses the IP protocol to pass or block IP source the incoming request based on its blacklist database table. (iii) As well, the STBOA and FAEB layers are compatible with the HTTP protocol. (iv) The STBOA layer checks HTTP protocol headers and then generates a mathematical puzzle in order to verify the validity of the requester. After that, it passes it to the next layer if it is legitimate, or it blocks it immediately if it is illegitimate. (v) The FAEB layer uses the HTTP protocol's relevant information in order to detect HR-DDoS and FC attacks and to block the former immediately, while it blocks the latter gradually.
(7) The framework should be deigned explicitly for processing web application layer; HTTP protocol, rather than only network layer; IP and ICMP protocols, or transport layer; TCP and UDP protocols.	√	√	√	The FCMDPF framework mainly concentrates on protecting the HTTP protocol from all sorts of DoS/DDoS attacks, such as HR-DDoS, LR-DDoS, and FC attacks. In addition, it traces back and finds out the true attacking IP sources.
(8) The framework should be easy to implement and does not cause huge processing and bandwidth overheads.	√	√	√	(i) In reality, the FCMDPF framework is simple to implement through collaborative multilayer; each layer is distributed and deployed at different point. (ii) The FCMDPF framework generates very low processing and bandwidth overheads, compared to those schemes and frameworks that use packet marking [25].
(9) The framework should be able to adopt and update itself dynamically, once needed.	✗	√	√	The FCMDPF framework can adapt and update itself once needed. In particular, when a new stealthy bot's feature is discovered, the relevant feature's pattern can be easily added to the STBOA scheme. As well, when a new or a different profile is in need, the relevant information such as HR-DDoS and FC threshold's values can be easily added to the FAEB scheme.

Table 2: Continued.

Framework specifications/FCMDPF layers	OB layer	STBOA layer	FAEB layer	Remarks
(10) The framework should provide support to the hybrid scheme.	√	√	√	In fact, the FCMDPF framework is designed in a way that supports the hybrid scheme that consists of proactive and reactive schemes. In particular, the OB and STBOA layers of the FCMDPF framework represent a proactive scheme, while the FAEB layer of the FCMDPF framework represents a reactive scheme.
(11) The framework should consume low storage memory.	√	√	√	In general, the FCMDPF framework's layers, the OB layer, STBOA layer, and FAEB layer, consume very low memory storage. In particular, the OB layer of the FCMDPF framework consumes very low memory to store its blacklist database table, while the STBOA layer of the FCMDPF framework does not consume any storage memory, since all of its transactions are done in the real time. As well, the FAEB layer of the FCMDPF framework consumes little memory to store the relevant information about web pages.
(12) The framework should be resistant to IP source spoofing attacks, especially during finding out the true attacking IP sources.	✗	√	✗	In fact, the FCMDPF framework is resistant to IP source spoofing attacks, since the STBOA scheme verifies whether the requester is legitimate or illegitimate by examining incoming request's headers and puzzle's answer. If the requester failed to satisfy these two tests, the requester is considered an attacker. Therefore, it will be blocked immediately.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This research paper took place in Universiti Teknologi Malaysia (UTM) with the financial support from the Ministry of Education of Malaysia, vote number R.K130000.7838.4F287.

References

[1] T. Subbulakshmi, I. A. A. Guru, and S. M. Shalinie, "Attack source identification at router level in real time using marking algorithm deployed in programmable routers," in *Proceedings of the International Conference on Recent Trends in Information Technology (ICRTIT '11)*, pp. 79–84, Chennai, India, June 2011.

[2] T. Karnwal, T. Sivakumar, and G. Aghila, "A comber approach to protect cloud computing against XML DDoS and HTTP DDoS attack," in *Proceedings of the IEEE Students' Conference on Electrical, Electronics and Computer Science: Innovation for Humanity (SCEECS '12)*, March 2012.

[3] A. Chonka, W. Zhou, and Y. Xiang, "Protecting web services with service oriented traceback architecture," in *Proceedings of the 8th IEEE International Conference on Computer and Information Technology (CIT '08)*, pp. 706–711, Sydney, Australia, July 2008.

[4] C. Qi, W. Lin, W. Dou, and S. Yu, "CBF: a packet filtering method for DDoS attack defense in cloud environment," in *Proceedings of the 9th International Conference on Dependable, Autonomic and Secure Computing (DASC '11)*, Sydney, Australia, December 2011.

[5] C.-T. Xia, X.-H. Du, L.-F. Cao, and H.-C. Chen, "An algorithm of detecting and defending CC attack in real time," in *Proceedings of the International Conference on Industrial Control and Electronics Engineering (ICICEE '12)*, pp. 1804–1806, August 2012.

[6] A. Chonka, W. Zhou, J. Singh, and Y. Xiang, "Detecting and tracing DDoS attacks by intelligent decision prototype," in *Proceedings of the 6th Annual IEEE International Conference on Pervasive Computing and Communications (PerCom '08)*, pp. 578–583, March 2008.

[7] L. Yang, T. Zhang, J. Song, J. Wang, and P. Chen, "Defense of DDoS attack for cloud computing," in *Proceedings of the IEEE International Conference on Computer Science and Automation Engineering (CSAE '12)*, pp. 626–629, Zhangjiajie, China, May 2012.

[8] M. H. Sqalli, F. Al-Haidari, and K. Salah, "EDoS-shield—a two-steps mitigation technique against EDoS attacks in cloud computing," in *Proceedings of the 4th IEEE International Conference on Cloud and Utility Computing (UCC '11)*, pp. 49–56, Victoria, Canada, December 2011.

[9] L. Liu, X. Jin, G. Min, and L. Xu, "Real-time diagnosis of network anomaly based on statistical traffic analysis," in *Proceedings of the 11th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom '12)*, pp. 264–270, June 2012.

[10] Y. Xie and S. Tang, "Online anomaly detection based on web usage mining," in *Proceedings of the IEEE 26th International Parallel and Distributed Processing Symposium Workshops (IPDPSW '12)*, pp. 1177–1182, May 2012.

[11] A. S. Sairam, L. A. Subramaniam, and G. Barua, "Defeating reflector based denial-of-service attacks using single packet filters," in *Proceedings of the 5th International ICST Conference on Communications and Networking in China (ChinaCom '10)*, August 2010.

[12] H. S. Mohan and A. Raji Reddy, "An effective defense against distributed denial of service in grid," in *Proceedings of the 1st International Conference on Integrated Intelligent Computing (ICIIC '10)*, pp. 84–89, Bangalore, India, August 2010.

[13] C. Ye and K. Zheng, "Detection of application layer distributed denial of service," in *Proceedings of the International Conference on Computer Science and Network Technology (ICCSNT '11)*, pp. 310–314, December 2011.

[14] J. Wang, X. Yang, and K. Long, "A new relative entropy based App-DDoS detection method," in *Proceedings of the 15th IEEE Symposium on Computers and Communications (ISCC '10)*, pp. 966–968, Riccione, Italy, June 2010.

[15] Z. Wang and X. Wang, "DDoS attack detection algorithm based on the correlation of IP address analysis," in *Proceedings of the 2nd Annual Conference on Electrical and Control Engineering (ICECE '11)*, pp. 2951–2954, September 2011.

[16] W. Jin, Y. Xiaolong, and L. Keping, "Web DDoS detection schemes based on measuring user's access behavior with large deviation," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '11)*, 2011.

[17] S. Oshima, T. Nakashima, and T. Sueyoshi, "The evaluation of an anomaly detection system based on Chi-square method," in *Proceedings of the 26th IEEE International Conference on Advanced Information Networking and Applications Workshops (WAINA '12)*, pp. 708–713, Fukuoka, Japan, March 2012.

[18] J. Zhang, Z. Qin, L. Ou, P. Jiang, J. Liu, and A. X. Liu, "An advanced entropy-based DDOS detection scheme," in *Proceedings of the International Conference on Information, Networking and Automation (ICINA '10)*, pp. V267–V271, October 2010.

[19] J. Schneider and S. Koch, "HTTPreject: handling overload situations without losing the contact to the user," in *Proceedings of the European Conference on Computer Network Defense (EC2ND '10)*, pp. 29–34, Berlin, Germany, October 2010.

[20] S. Wen, W. Jia, W. Zhou, and C. Xu, "CALD: Surviving various application-layer DDoS attacks that mimic flash crowd," in *Proceedings of the 4th International Conference on Network and System Security (NSS '10)*, pp. 247–254, September 2010.

[21] F. Wang, X. Wang, J. Su, and B. Xiao, "VicSifter: a collaborative DDoS detection system with lightweight victim identification," in *Proceedings of the 11th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom '12)*, pp. 215–222, Liverpool, UK, June 2012.

[22] A. Darapureddi, R. Mohandas, and A. R. Pais, "Throttling DDoS attacks using discrete logarithm problem," in *Proceedings of the International Conference on Security and Cryptography (SECRYPT '10)*, pp. 263–269, July 2010.

[23] S. Oshima, T. Nakashima, and T. Sueyoshi, "Early DoS/DDoS detection method using short-term statistics," in *Proceedings of the 4th International Conference on Complex, Intelligent and Software Intensive Systems (CISIS '10)*, pp. 168–173, Krakow, Poland, February 2010.

[24] Z. Yi, L. Qiang, and Z. Guofeng, "A real-time DDoS attack detection and prevention system based on per-IP traffic behavioral analysis," in *Proceedings of the 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT '10)*, vol. 2, pp. 163–167, Chengdu, China, July 2010.

[25] K. Kumar, A. L. Sangal, and A. Bhandari, "Traceback techniques against DDOS attacks: a comprehensive review," in *Proceedings of the 2nd International Conference on Computer and Communication Technology (ICCCT '11)*, pp. 491–498, September 2011.

[26] P. Du and A. Nakao, "DDoS defense deployment with network egress and ingress filtering," in *Proceedings of the IEEE International Conference on Communications (ICC '10)*, May 2010.

[27] H. Beitollahi and G. Deconinck, "A cooperative mechanism to defense against distributed denial of service attacks," in *Proceedings of the IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom '11)*, Changsha, China, November 2011.

[28] X. Bi, W. Tan, and R. Xiao, "A DDoS-oriented Distributed defense framework based on edge router feedbacks in Autonomous Systems," in *Proceedings of the 3rd International Multi-Symposiums on Computer and Computational Sciences (IMSCCS '08)*, pp. 132–135, October 2008.

[29] H. Yim, T. Kim, and J. Jung, "Probabilistic route selection algorithm to trace DDoS attack traffic source," in *Proceedings of the International Conference on Information Science and Applications (ICISA '11)*, April 2011.

[30] R. Yogesh Patil and L. Ragha, "A rate limiting mechanism for defending against flooding based distributed denial of service attack," in *Proceedings of the World Congress on Information and Communication Technologies (WICT '11)*, pp. 182–186, Mumbai, India, December 2011.

[31] H. R. Zeidanloo, M. J. Zadeh, P. V. Amoli, M. Safari, and M. Zamani, "A taxonomy of Botnet detection techniques," in *Proceedings of the 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT '10)*, pp. 158–162, July 2010.

[32] R. Vaidyanathan, A. Ghosh, Y.-H. Cheng, A. Yamada, and Y. Miyake, "On the use of BGP AS numbers to detect spoofing," in *Proceedings of the IEEE Globecom Workshops (GC '10)*, pp. 1606–1610, Miami, Fla, USA, December 2010.

[33] W. Feng, E. Kaiser, and A. Luu, "Design and implementation of network puzzles," in *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '05)*, 2005.

[34] T. J. McNevin, P. Jung-Min, and R. Marchany, "Chained puzzles: a novel framework for IP-layer client puzzles," in *Proceedings of the International Conference on Wireless Networks, Communications and Mobile Computing*, vol. 1, pp. 298–303, IEEE, June 2005.

[35] N. A. Fraser, D. J. Kelly, R. A. Raines et al., "Using client puzzles to mitigate distributed denial of service attacks in the tor anonymous routing environment," in *Proceedings of the IEEE International Conference on Communications (ICC '07)*, 2007.

[36] A. Shevtekar and N. Ansari, "A proactive test based differentiation technique to mitigate low rate DoS attacks," in *Proceedings of the 16th International Conference on Computer Communications and Networks (ICCCN '07)*, pp. 639–644, August 2007.

[37] K. S. Hin and A. Nakao, "Overfort: combating DDoS with peer-to-peer DDoS puzzle," in *Proceedings of the IEEE International Symposium on Parallel and Distributed Processing (IPDPS '08)*, pp. 1–8, Miami, Fla, USA, April 2008.

[38] L. Liu, X. Zhang, and S. Chen, "Botnet with browser extensions," in *Proceedings of the IEEE International Conference on Privacy,*

Security, Risk and Trust (PASSAT '11) and IEEE International Conference on Social Computing (SocialCom '11), pp. 1089–1094, Boston, Mass, USA, October 2011.

[39] K. Yamauchi, Y. Hori, and K. Sakurai, "Detecting HTTP-based botnet based on characteristic of the C&C session using by SVM," in *Proceedings of the 8th Asia Joint Conference on Information Security*, pp. 63–68, IEEE, Seoul, Republic of korea, July 2013.

[40] J.-S. Lee, H. C. Jeong, J.-H. Park, M. Kim, and B.-N. Noh, "The activity analysis of malicious http-based botnets using degree of periodic repeatability," in *Proceedings of the International Conference on Security Technology (SecTech '08)*, pp. 83–86, December 2008.

[41] N. H. Vo and J. Pieprzyk, "Protecting web 2.0 services from botnet exploitations," in *Proceedings of the 2nd Cybercrime and Trustworthy Computing Workshop (CTC '10)*, pp. 18–28, Victoria, Canada, July 2010.

[42] T. Cai and F. Zou, "Detecting HTTP botnet with clustering network traffic," in *Proceedings of the 8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM '12)*, pp. 1–7, September 2012.

[43] B. Wang, Z. Li, D. Li, F. Liu, and H. Chen, "Modeling connections behavior for web-based bots detection," in *Proceedings of the 2nd International Conference on e-Business and Information System Security (EBISS '10)*, pp. 141–144, May 2010.

[44] N.-Y. Lee and H.-J. Chiang, "The research of botnet detection and prevention," in *Proceedings of the International Computer Symposium (ICS '10)*, pp. 119–124, December 2010.

[45] G. Oikonomou and J. Mirkovic, "Modeling human behavior for defense against flash-crowd attacks," in *Proceedings of the IEEE International Conference on Communications (ICC '09)*, June 2009.

[46] A. Michalas, N. Komninos, N. R. Prasad, and V. A. Oleshchuk, "New client puzzle approach for DoS resistance in ad hoc networks," in *Proceedings of the IEEE International Conference on Information Theory and Information Security (ICITIS '10)*, pp. 568–573, Beijing, China, December 2010.

[47] S. H. Khor and A. Nakao, "DaaS: DDoS mitigation-as-a-service," in *Proceedings of the 11th IEEE/IPSJ International Symposium on Applications and the Internet (SAINT '11)*, pp. 160–171, Munich, Germany, July 2011.

[48] A. Lukyanenko, A. Gurtov, and A. Yla-Jaaski, "DISPUTE: Distributed puzzle tussle," in *Proceedings of the 16th IEEE Symposium on Computers and Communications (ISCC '11)*, pp. 775–778, July 2011.

[49] J. Chen, "Broadcast authentication protocol scheme based on DBP-MSP and safe routing in WSN against DDoS attacks," in *Proceedings of the 2nd International Conference on Networking and Distributed Computing (ICNDC '11)*, pp. 170–174, September 2011.

[50] G. No and I. Ra, "Adaptive DDoS detector design using fast entropy computation method," in *Proceedings of the 5th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS '11)*, pp. 86–93, July 2011.

[51] G. No and I. Ra, "An efficient and reliable DDoS attack detection using a fast entropy computation method," in *Proceedings of the 9th International Symposium on Communications and Information Technology (ISCIT '09)*, pp. 1223–1228, September 2009.

[52] Y. Liu, J. Cheng, J. Yin, and B. Zhang, "Detecting DDoS attacks using conditional entropy," in *Proceedings of the International Conference on Computer Application and System Modeling (ICCASM '10)*, pp. V13278–V13282, Taiyuan, China, October 2010.

[53] S. Yu, W. Zhou, R. Doss, and W. Jia, "Traceback of DDoS attacks using entropy variations," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 3, pp. 412–425, 2012.

[54] K. Kumar, R. C. Joshi, and K. Singh, "A distributed approach using entropy to detect DDoS attacks in ISP domain," in *Proceedings of the International Conference on Signal Processing, Communications and Networking (ICSCN '07)*, pp. 331–337, February 2007.

[55] A. T. Lawniczak, H. Wu, and B. N. Di Stefano, "Detection of anomalous packet traffic via entropy," in *Proceedings of the Canadian Conference on Electrical and Computer Engineering (CCECE '09)*, pp. 137–141, St. John's, Canada, May 2009.

[56] S. Schweizer, "mod_qos," 2014, http://opensource.adnovum.ch/mod_qos/index.html.

[57] C. Ciufo, "mod_evasive on Apache," 2011, https://library.linode.com/web-servers/apache/mod-evasive.

[58] K. Monshouwer, "Mod_antiloris for 2.4," 2012, http://www.apachelounge.com/viewtopic.php?p=20543.

[59] SpiderLabs/ModSecurity, ModSecurity Reference Manual. 2014, https://github.com/SpiderLabs/ModSecurity/wiki/Reference-Manual.

[60] T. Thapngam, S. Yu, and W. Zhou, "DDoS discrimination by Linear Discriminant Analysis (LDA)," in *Proceedings of the International Conference on Computing, Networking and Communications (ICNC '12)*, pp. 532–536, February 2012.

[61] H. Liu, Y. Sun, and M. S. Kim, "Fine-grained DDoS detection scheme based on bidirectional count sketch," in *Proceedings of the 20th International Conference on Computer Communications and Networks (ICCCN '11)*, pp. 1–6, Lahaina, Hawaii, USA, August 2011.

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2015, Article ID 703713, 5 pages
http://dx.doi.org/10.1155/2015/703713

Research Article

Using a Prediction Model to Manage Cyber Security Threats

Venkatesh Jaganathan, Priyesh Cherurveettil, and Premapriya Muthu Sivashanmugam

Department of Management Studies, Anna University Regional Centre Coimbatore, Jothipuram Post, Coimbatore, Tamilnadu 641 047, India

Correspondence should be addressed to Priyesh Cherurveettil; priyesh.cherurveettil@gmail.com

Received 12 September 2014; Revised 14 December 2014; Accepted 25 December 2014

Academic Editor: Aneel Rahim

Cyber-attacks are an important issue faced by all organizations. Securing information systems is critical. Organizations should be able to understand the ecosystem and predict attacks. Predicting attacks quantitatively should be part of risk management. The cost impact due to worms, viruses, or other malicious software is significant. This paper proposes a mathematical model to predict the impact of an attack based on significant factors that influence cyber security. This model also considers the environmental information required. It is generalized and can be customized to the needs of the individual organization.

1. Introduction

The digital assets of an organization are prone to attack any time. With threats gathering new dimensions, organizations should be able to objectively evaluate the risks of existing and new software applications. Based on this risk evaluation, sufficient resources can be allocated to mitigate cyber security risks. Quantitatively predicting proneness to attack can help organizations counter attack occurrences. The Common Vulnerability Scoring System (CVSS) is a standard framework used by many organizations. It communicates the characteristics and impacts of IT vulnerabilities. This framework has three groups, namely, Base, Temporal, and Environmental. The base group highlights the qualities of vulnerability that are unchanged over time and user. The temporal group covers the characteristics of vulnerability over time and the environmental group highlights the specific user environment. The CVSS helps establish a common language in the IT community. This paper proposes a mathematical model for predicting the impact of an attack based on the significant factors that influence cyber security. These factors are arrived at by considering several historical data points and mathematically verifying their significance to the impact and characteristics of attacks.

2. Related Work

Sheyner et al., Wang et al., and Kuhl et al. [1–3] highlighted that security analysis is based on attack graph generation and simulation. The focus of their work was on attack generation tools. This paper proposes establishing the quantitative relationship between the attack impact and the attack parameters. Tittel [4] explains about Unified Threat Management and why it is important to address it in his paper. Wu et al. [5] established a prediction model based on integrating environmental factors and attack graphs in a Bayesian network. They found that environmental information is important for accurate safety evaluations. This paper proposes mathematically proving that environmental information influences the characteristics and impacts of an attack. Anusha et al. [6] studied various models like unimodel and multimodel for enhanced security. They discussed about authentication at the beginning of the exam and the user system checks. Axelrad et al. [7] introduced a Bayesian network model for the motivation and psychology of the malicious insider. Khan and Hussain [8] established relationships between attack probability and vulnerability. However, the collective influence of attack environment factors on the attack was not revealed. Moore et al. [9] described a modeling and simulation foundation, based on the system dynamics methodology to test the efficacy of insider threat detection controls. The paper discusses risk management and early detection of risks based on insider threat.

The three impact metrics in the CVSS measure how vulnerability is assessed and how it impacts on an IT asset. These three metrics are Access Vector, Access Complexity,

and Authentication. It is also important to understand how vulnerability affects the integrity, confidentiality, and availability of these parameters. Cyber security metrics can be broadly classified into two based on the source of measurement. Several measurements are possible from the malicious user/group end. Some measurements are also possible from the host/victim end. From the side of a malicious user [10], the Intensity, Stealth and Time of attack are possible measurements. Further, the technical and cyber knowledge of personnel are relative measurements from the malicious user/group end. Measurements that are possible from the host end include the Vulnerabilities present in the tool, which are detectable through tools such as Nessus and X-Force, the Traffic to a particular application over a period of time, the Power of the protection tools installed at the target system, and the Value of the assets present in the network. From these two classifications, to build a prediction model, the measurement taken from the host end is used, whereas none of the malicious user/group end measurements is considered. With the increasing dimensions of attack nature, it is not possible to accurately calculate these measurements. Moreover, such measurements from the malicious user end are of little use if the attack possibilities need to be controlled by organizations hosting the target applications.

3. Prediction Model

Prediction models can be developed to predict different project outcomes and interim outcomes by using statistical techniques. A process performance model adopts the concepts of probability. This can also be explored further by building simulations. Output can be studied as a range. Depending on the predictions, midcourse corrections can be recommended. The model can be simulated to predict final outcomes based on the corrections suggested. It is thus a proactive model that helps the technical analyst to analyze the data and predict outcomes. Analysts can change the data and perform what-if analyses. They can then record these instances and decide on the best option. The model helps analysts decide which lever to adjust to meet the final project goal.

In the current system, the focus of fixing vulnerabilities is not based on the potential impact of the vulnerability. Further, more than adequate importance is given to fixing all the vulnerabilities or too little importance due to time and cost constraints. A proactive risk assessment prior to the release of the IT application is not available. The impact of any vulnerability is identified by using the CVSS calculator, only after analyzing how the attack took place.

In the proposed model, which was also piloted in a sample project organization, the potential impact of the vulnerability is predicted well before the IT application is released for usage. With this impact information, adequate cost and resources can be allocated to resolving vulnerabilities, thereby reducing the impact.

The multiple regression method is chosen to predict the impact of attacks in this proposal. Multiple regressions have certain underlying assumptions such as linearity, the nonexistence of multicollinearity, homoscedasticity, and normality. Each of these assumptions is validated for our attempt to establish the relationship between the impact of attack and the influencing factors. Based on the research above, factors that can influence the impact of an attack were identified, namely, Level of security protection on the target system, Usage or traffic in the identified network, Vulnerabilities present in the target system, and Value of assets present in the network. Oluwatosin and Samson [11] highlighted the challenges of existing computer applications that need to be considered from a security perspective.

Shar et al. [12] predicted vulnerabilities with features related to dataflow. Khan and Hussain [8] stated that it is safe to assume that all these factors have individual linear relationships with the probability of attack. A similar relationship is seen from the scatterplot of the impact of an attack and the environmental factors considered. This satisfies the first condition for using the multiple regression technique for our prediction model.

Multicollinearity and homoscedasticity are verified through the variance inflation factor (VIF). Normality is tested by performing the Anderson–Darling Test. In addition, a hypothesis test is run individually for each of the attack factors to ensure that the probability that the factor does not influence the independent factor is kept to a maximum of only 0.05. The coefficient of correlation (R^2) and adjusted coefficient of correlation (Adj. R^2) are maintained close to each other. This is sufficient to prove that among the many factors that can cause an attack, the selected factors predict the impact of an attack with the best possible accuracy. That is, the number of factors, beyond a certain point, becomes immaterial.

The operational definition of cyber security metrics considered is mentioned below.

(1) Y is the Overall CVSS Score, the dependent factor. The CVSS [13, 14] is predicted based on the environment and system characteristics of the target application.

(2) $X1$ is the number of vulnerabilities, namely, the total number of vulnerabilities detected by the static and dynamic vulnerability detection tools for the target application. The tools installed and run against the target application can identify several vulnerabilities based on algorithms such as but not limited to improved tainted algorithms or penetration testing. In a given application, the vulnerabilities reported by the tools can be broadly classified into 23 categories: API Abuse, Authentication Vulnerability, Authorization Vulnerability, Availability Vulnerability, Code Permission Vulnerability, Code Quality Vulnerability [15], Configuration Vulnerability, Cryptographic Vulnerability, Encoding Vulnerability, Environmental Vulnerability, Error Handling Vulnerability, General Logic Error Vulnerability, Input Validation Vulnerability, Logging and Auditing Vulnerability, Password Management Vulnerability, Path Vulnerability, Protocol Errors, Range and Type Error Vulnerability, Sensitive Data Protection Vulnerability [16], Session Management Vulnerability, Synchronization and

Table 1: Project data points.

Y CVSS Score	$X1$ Vulnerability	$X2$ Network Traffic
2.1	20	324
5.3	53	623
1.0	15	235
8.0	85	932
2.9	28	438
3.0	25	498
3.8	38	391
1.0	18	132
1.2	16	177
5.9	63	823
4.3	39	579
2.8	30	455
1.1	14	231
4.2	35	725
5.4	51	740
1.9	21	345
2.0	25	432
4.1	37	467
6.2	58	845
1.1	15	111
2.3	22	191
1.2	16	182
2.8	30	292
6.9	68	952
4.8	55	600

Timing Vulnerability, Unsafe Mobile Code and Use of Dangerous API [17].

(3) $X2$ is the Average Input Network Traffic recorded for the application during the week of the attack in KBPS.

Table 1 highlights the data points for each metric during all instances of an attack. In total, 25 such attack history data from a project are shown in Table 1. The data points from the CVSS calculator were recorded for every attack encountered. Output from the vulnerability tool was recorded for the target application. For the specified week range of the attack, network traffic was also recorded. The data from these three sources were tabulated every time an attack was encountered, as shown in Table 1. We also ensured that respective data points were taken from the same sample. The sample definitions were defined by the technical analyst. In this regression model, CVSS score (Y) was predicted by using the two X variables, vulnerability and network traffic.

The null hypothesis considered is that $X1$ and $X2$ have no influence over Y. In other words, vulnerability and network traffic have no influence over CVSS score. No mirror pattern can be found in the residual plot in Figure 1 and hence no heteroscedasticity is found. The normal probability plot shown in Figure 2 is approximately linear. From the figure, it is clear that the normality assumption for the errors has not been violated. With regard to the P value, since it is 0.02

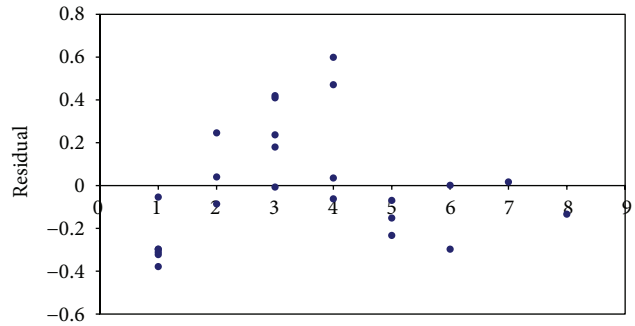


Figure 1: Residual plot.

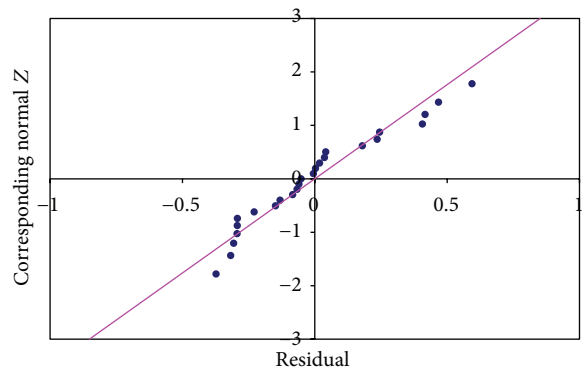


Figure 2: Normal probability plot.

Table 2: Regression equation.

Intercept	Vulnerability	Network Traffic
−0.2983	0.07174	0.0025

(<0.05), the null hypothesis is not valid, which means the variables selected have an impact on CVSS score.

As shown in Table 2, vulnerability has a positive influence on CVSS score. As vulnerability increases, CVSS score increases and hence the impact on IT assets is high. The influence of network traffic on CVSS is positive. This means that when network traffic is high, the impact of vulnerabilities is high and CVSS score is high. Thus, CVSS score is impacted positively both by vulnerability and by network traffic:

$$\begin{aligned}&\text{Predicted CVSS Score}\\&\quad = -0.2893 + 0.07174 * \text{Number of vulnerabilities}\\&\qquad \text{on the IT application reported by tools}\\&\qquad + 0.0025 * \text{Proposed average input network traffic}\\&\qquad \text{for the application for a week measured in KBPS.}\end{aligned} \tag{1}$$

As shown in Table 3, b is the coefficient that gives the least squares estimates, while $s(b)$ gives the standard errors of the least squares estimates for the x variables and t gives the computed t-statistic. This is the coefficient divided by the standard

Table 3: Multiple regression results.

	Intercept	Vulnerability	Network Traffic
b	−0.296	0.0706	0.002
$s(b)$	0.121	0.007	0.0005
t	−2.442	9.359	4.545
P	0.0231	0.0000	0.0002

Table 4: ANOVA table.

Source	SS	Df	MS	F	P
Regression.	96.22	2	48.15	597	0.000
Error	1.77	22	0.080		
Total	98	24			

error. The P value gives the P value for the hypothesis test. The VIF quantifies the severity of multicollinearity in an ordinary least squares regression analysis. The VIF for the given data is 6.41.

As shown in Table 4, SS is the sum of squares due to the regression. This measure of total variation in Y can be explained by the regression with the X variable. Df is the degrees of freedom. MS is the mean square, which is a measure of the sum of squares divided by the degrees of freedom. Mean square regression (MSR) and mean square error (MSE) are the two variables that define F: F = MSR/MSE. The F-statistic is used to test whether the Y and X variables are related.

For the given data, MSR is 48 and MSE is 0.08. The F-statistic determines that the P value is zero. This confirms the existence of a linear relationship between CVSS and the two variables, network traffic and vulnerabilities. R^2 provides information about the goodness of fit of a model. In the regression equation, the R^2 coefficient of determination determines how well the regression line approximates the real data points. Adjusted R^2 is a modified version that adjusts the number of predictors in the model. For the given data, R^2 is 0.9819 and adjusted R^2 is 0.9803.

The data prove that overall CVSS score is influenced by vulnerabilities in the network and network traffic. The infrastructure team in the organization shares the baseline data for these variables on a regular basis with the quality team. For each network process, based on the network type and applications hosted, a logical grouping can be considered and organization values can be baselined. Technical analysts can then refer to these baseline organizational data when they start the network design process. As part of the process, they can also use these reference values to determine the upper and lower specification limits. These values will be available for each of the subprocess parameters. The technical analyst can then determine and analyze which vulnerabilities need to be controlled and select threshold values based on that.

Based on the selected threshold values, what-if analysis is performed. Going by the different scenarios, vulnerability and network traffic values are assumed and provided as inputs to the model. The predicted outcome is then compared with the thresholds. It is important to note that while changing the parameters, technical analysts should understand the practical implications of the project. It is not only about the mathematical model, but about how it can be put into practice. For example, if CVSS score is high, how can it be reduced? How can vulnerabilities be reduced during design? Cost implications need to be considered. Then, the technical analyst has to look at the environmental constraints. As the prediction model considers the key influencing factors to predict the CVSS, the influencing factor values might affect the project schedule and project cost, which need to be analyzed as well. These forecasts serve as alerts that it should take action to mitigate the threat of cyber-attacks.

Predicting CVSS scores helps prioritize vulnerabilities and remediate those with high risks. CVSS scores are shared by software application vendors with their customers. This helps customers understand the severity of vulnerabilities and allows them to effectively manage their risks. Vulnerability bulletins are shared by few organizations. These bulletins share the date of attack, systems affected, and patches performed. Thus, the CVSS prediction model is vital and should be used extensively. Technical analysts should be comfortable using the prediction model extensively. For every scenario, the analyst should document the assumptions and associated risks. A detailed attack prevention plan should be in place. At every step, the attack, its type, cause, and preventive action should be documented. Different root-cause analysis techniques such as 5-why can be used to pinpoint the root cause. After identifying the root cause, the next steps in terms of corrective and preventive actions should also be thought through. Technical experts should review these plans so that they can bring in their experience and highlight any improvements.

Prediction models should not be a one-time activity. Technical analysts should use the model on an ongoing basis and also suggest shortcomings. Based on the scores, decisions need to be taken considering impact on cost and security. Prediction models are statistical and simulative in nature. These models should help simulating scenarios as well as determining outcomes. They can also model different variation factors and help the analyst with the predicted range or the variation of its outcomes.

4. Conclusion

Cyber-attack is an attempt to exploit computer systems and networks. Cyber-attacks use malicious codes to alter algorithms, logic, or data. Securing information systems is thus critical. Multiple countermeasures need to be built. The CVSS is an industry framework that helps quantify the vulnerability impact. This paper demonstrated a mathematical model to predict the impact of an attack based on significant factors that influence cyber security. Vulnerability and network traffic were selected as the influencing factors to predict CVSS score. Based on the score, the technical analyst can analyze the impact and take necessary preventive actions. This model also considers the environmental information required. It is thus generalized and can be customized to the needs of the individual organization.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] O. Sheyner, J. Haines, S. Jha, R. Lippmann, and J. M. Wing, "Automated generation and analysis of attack graphs," in *Proceedings of the IEEE Symposium on Security and Privacy*, pp. 273–284, May 2002.

[2] L. Wang, A. Singhal, and S. Jajodia, "Toward measuring network security using attack graphs," in *Proceedings of the ACM Workshop on Quality of Protection*, pp. 49–54, October 2007.

[3] M. E. Kuhl, J. Kistner, K. Costantini, and M. Sudit, "Cyber attack modeling and simulation for network security analysis," in *Proceedings of the Winter Simulation Conference (WSC '07)*, pp. 1180–1188, Washington, DC, USA, December 2007.

[4] E. Tittel, *Preventing and Avoiding Network Security Threats and Vulnerabilities*, 2013, http://www.tomsitpro.com/articles/threat_management-it_security-firewall-it_certification-network_security,2-477.html.

[5] J. Wu, L. Yin, and Y. Guo, "Cyber attacks prediction model based on Bayesian network," in *Proceedings of the 18th IEEE International Conference on Parallel and Distributed Systems (ICPADS '12)*, pp. 730–731, Singapore, December 2012.

[6] N. S. Anusha, T. S. Soujanya, and D. S. Vasavi, "Study on techniques for providing enhanced security during online exams," *International Journal of Engineering Inventions*, vol. 1, no. 1, pp. 32–37, 2012.

[7] E. T. Axelrad, P. J. Sticha, O. Brdiczka, and J. Shen, "A Bayesian network model for predicting insider threats," in *Proceedings of the 2nd IEEE Security and Privacy Workshops (SPW '13)*, pp. 82–89, May 2013.

[8] M. A. Khan and M. Hussain, "Cyber security quantification model," *Bahria University Journal of Information and Communication Technology*, vol. 3, no. 1, pp. 23–27, 2010.

[9] A. P. Moore, D. A. Mundie, and M. L. Collins, "A system dynamics model for investigating early detection of insider threat risk," Tech. Rep. DM-0000143, Program Software Engineering Institute, Carnegie Mellon University, 2013.

[10] M. Mateski, C. M. Trevino, C. K. Veitch et al., "Cyber threat metrics," Sandia Report SAND2012-2427, 2012.

[11] O. T. Oluwatosin and D. D. Samson, "Computer-based test security and result integrity," *International Journal of Computer and Information Technology*, vol. 2, no. 2, pp. 324–329, 2013.

[12] L. K. Shar, H. Beng Kuan Tan, and L. C. Briand, "Mining SQL injection and cross site scripting vulnerabilities using hybrid program analysis," in *Proceedings of the 35th International Conference on Software Engineering (ICSE '13)*, pp. 642–651, IEEE Press, May 2013.

[13] P. Mell, K. Scarfone, and S. Romanosky, *A Complete Guide to the Common Vulnerability Scoring System Version 2.0*, National Institute of Standards and Technology, Carnegie Mellon University, 2007.

[14] P. Mell, K. Scarfone, and S. Romanosky, *A Complete Guide to the Common Vulnerability Scoring System, Version 2.0*, 2007.

[15] Redhat, *Bash Code Injection Vulnerability via Specially Crafted Environment Variables (CVE-2014-6271, CVE-2014-7169)*, Redhat, 2014, https://access.redhat.com/articles/1200223.

[16] Sensitive Data Protection Vulnerability, OWASP, https://www.owasp.org/index.php/Category:Sensitive_Data_Protection_Vulnerability.

[17] https://www.owasp.org/index.php/Category:Vulnerability.

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2014, Article ID 612787, 11 pages
http://dx.doi.org/10.1155/2014/612787

Research Article

Intelligent Bar Chart Plagiarism Detection in Documents

Mohammed Mumtaz Al-Dabbagh,[1,2] Naomie Salim,[1] Amjad Rehman,[3] Mohammed Hazim Alkawaz,[1,2] Tanzila Saba,[4] Mznah Al-Rodhaan,[5] and Abdullah Al-Dhelaan[5]

[1] *Faculty of Computing, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia*
[2] *Faculty of Computer Sciences and Mathematics, University of Mosul, Mosul, Iraq*
[3] *MIS Department, CBA, Salman Bin Abdulaziz University, Alkharj, Saudi Arabia*
[4] *College of Computer and Information Sciences (CCIS), Prince Sultan University, Riyadh, Saudi Arabia*
[5] *Computer Science Department, College of Computer & Information Sciences, King Saud University, Riyadh, Saudi Arabia*

Correspondence should be addressed to Amjad Rehman; ar.khan@sau.edu.sa

Received 30 March 2014; Revised 21 June 2014; Accepted 7 July 2014; Published 17 September 2014

Academic Editor: Iftikhar Ahmad

This paper presents a novel features mining approach from documents that could not be mined via optical character recognition (OCR). By identifying the intimate relationship between the text and graphical components, the proposed technique pulls out the Start, End, and Exact values for each bar. Furthermore, the word 2-gram and Euclidean distance methods are used to accurately detect and determine plagiarism in bar charts.

1. Introduction

Detection, determination, and rectification of plagiarism are outstanding quests in every sphere of documentation and copyright. Lately, the significant advancement in information technology represented by digital libraries and World Wide Web is regarded as one of the main reasons for exponential growth in plagiarism appearance. It has become effortless for the plagiarist to utilize or copy the work of others without acknowledging or citing them due to the easy availability of most resources in digital format. Thus, plagiarism is regarded as one of the electronic crimes and intellectual thefts from others documents [1–3]. In academia, plagiarism posed a severe educational challenge which is acutely faced by research institutions, universities, and even schools. Several efforts are dedicated to detecting different types of plagiarism via programming code and text. Plagiarism detection began in the 1970s, where the identification of rate of plagiarism in programming code written by some computer languages such as C and Pascal was introduced [4]. Digital documents being the major carriers of information require extreme authentication in terms of their origins and trustfulness. The quest for achieving an accurate and efficient image forgery detection method in digital documentation is never ending. Developing a robust plagiarism detector by overcoming the limitations associated with human intervention is the key issue [5].

Recently, several researchers developed the algorithmic approach using computer codes to detect plagiarism in the homework of students [6]. Based on levels of plagiarism patterns some studies introduced plagiarism detection methods which are implemented in the algorithms and program codes [7]. Generally, the computerized or statistical approaches are exploited to detect plagiarism in natural language since the 1990s. The techniques used for natural language are based on various factors such as grammar, semantic, and grammar-semantics hybridizations [1, 4]. However, the grammar-based method is one of the restrictive ones to detect plagiarism. This type of method analyzes the sentences based on grammatical structure, which can be efficiently used to detect the Exact

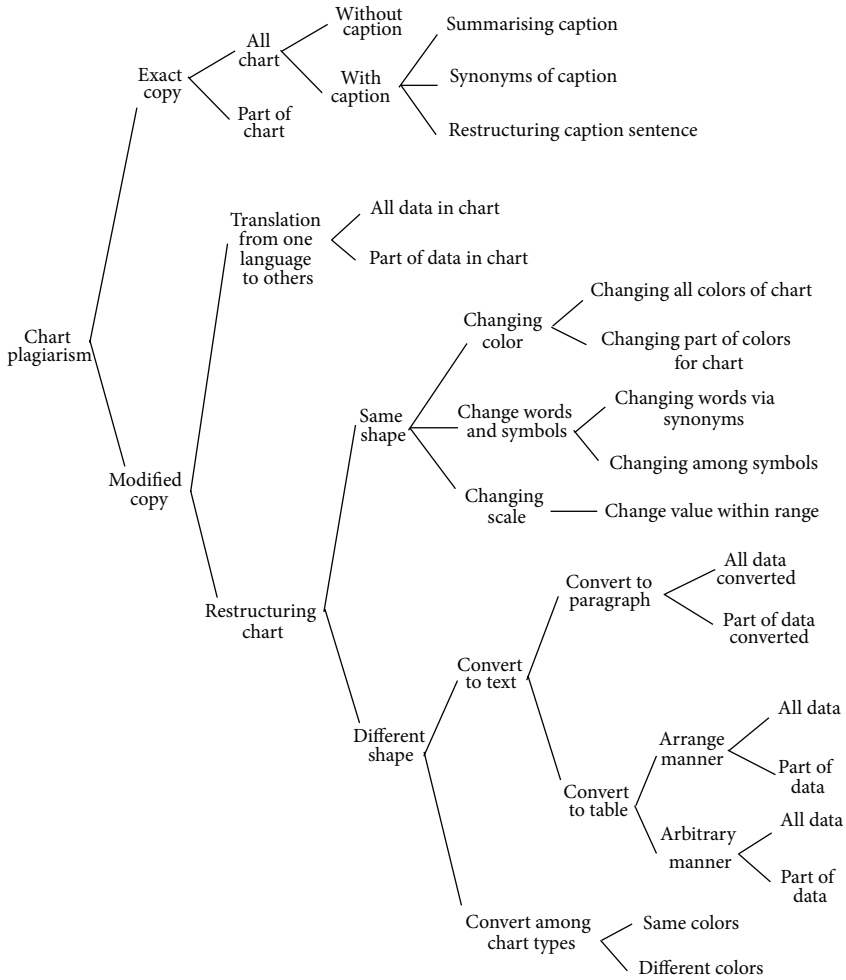


Figure 1: Taxonomy of chart plagiarism.

Copy of text. While semantic-based method utilizes vector space model to calculate the similarities among the texts. Undoubtedly, the grammar-semantics hybrid approach overcomes all disadvantages of the other methods. This is considered as one of the most versatile techniques to detect text plagiarism [1, 8, 9].

A new taxonomy is introduced to explain the concepts for various types and patterns of text plagiarism [4]. Plagiarism is divided into two main parts including literal and intelligent one. Each part consists of several subparts which cover all possibilities of text plagiarism. Generally, the representations of quantitative information are formulated via infographic form by using figures, charts, and tables. The information that is displayed in charts, figures, and tables includes results of experiments, framework, and statistical facts. These data and information in homogenous form can be formulated by using various shapes such as pie chart, bar chart, and 2D and 3D plots [10–12].

We report a new type of plagiarism detection method by highlighting the types of information that can be stolen from others work without referencing. Firstly, different types of forged information are organized into taxonomy of chart, figure, and table to highlight varieties of plagiarism patterns such as Exact and Modified Copy. Secondly, plagiarism detection in bar chart image is performed depending on ten features in images. Some of the features are extracted by OCR tool while others are acquired from the relationship of text and graphic components [13, 14]. Finally, the proposed technique is used to extract the features of bar chart images which cannot be extracted by OCR to detect plagiarism. The paper is organized as follows: Section 2 describes various existing techniques for extracting data from bar chart images [15]. The taxonomy of chart, figure, and table related to plagiarism is presented in Section 3. Section 4 discusses the methodology and Section 5 includes the experimental results of bar chart plagiarism detection. The discussions are elucidated in Section 6. Section 7 concludes the paper.

2. Related Work

Categorizations of bar chart images refer to their labeling into one of the predefined geometrical or nongeometrical classes. Though the classification is apparently manageable,

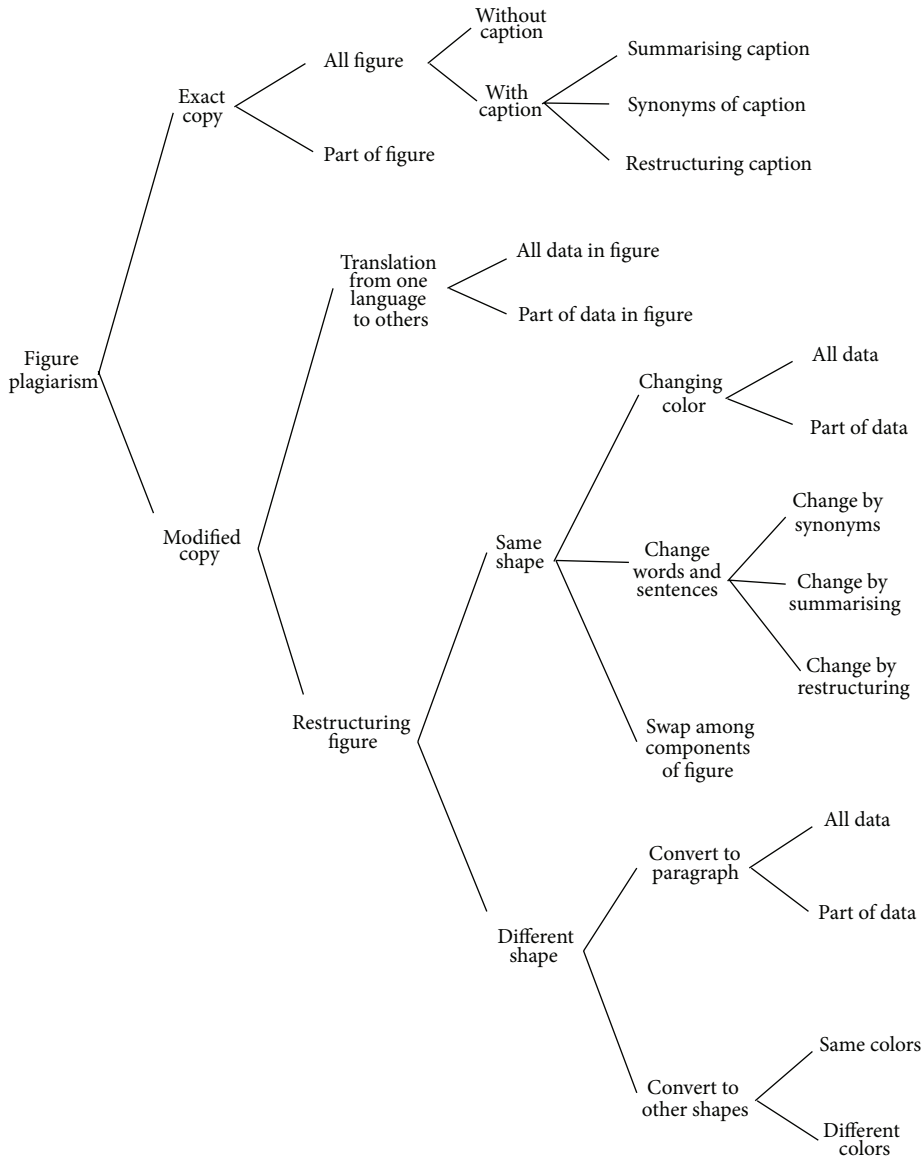


Figure 2: Taxonomy of figure plagiarism.

it is proven to be an extremely difficult problem in computer programming. Hence, there is an intense attention in developing automatic tools to categorize, describe, or retrieve images based on their contents.

Consequently, researchers attempted to extract the features and data from chart images. For automatic images categorization and description computational model is successfully introduced [16]. The analysis of local and global image characteristics by using text and image features is used in the model. The model is capable of differentiating geometrical and ordinary images. The computational model is comprised of classifier stage which is trained by the associated text features using advanced concepts and similarity matching stage.

Classification methods based on multiple-instance learning for chart images are also developed [10]. A re-revision system consisting of three concatenated major stages such as classification, extraction, and redesigned chart images is employed [17]. In the extraction stage, two types of charts (pie and bar) are focused on. Some techniques are presented in extracting data and graphical marks from chart images. Truly, the understanding and recognition of chart images require the preprocessing and extraction of data and information. Primarily, two types of available methods that deal with chart images are either to consider electronic chart directly [18–20] or to obtain them after converting into raster images [21–24]. Mishchenko and Vassilieva [25] introduced a model-based method for the classification of chart images which involved two main stages: firstly, predicting the location and the size of chart depending on the color distribution of chart image and secondly the extraction and matching of chart image edges to achieve the best match between query and database images.

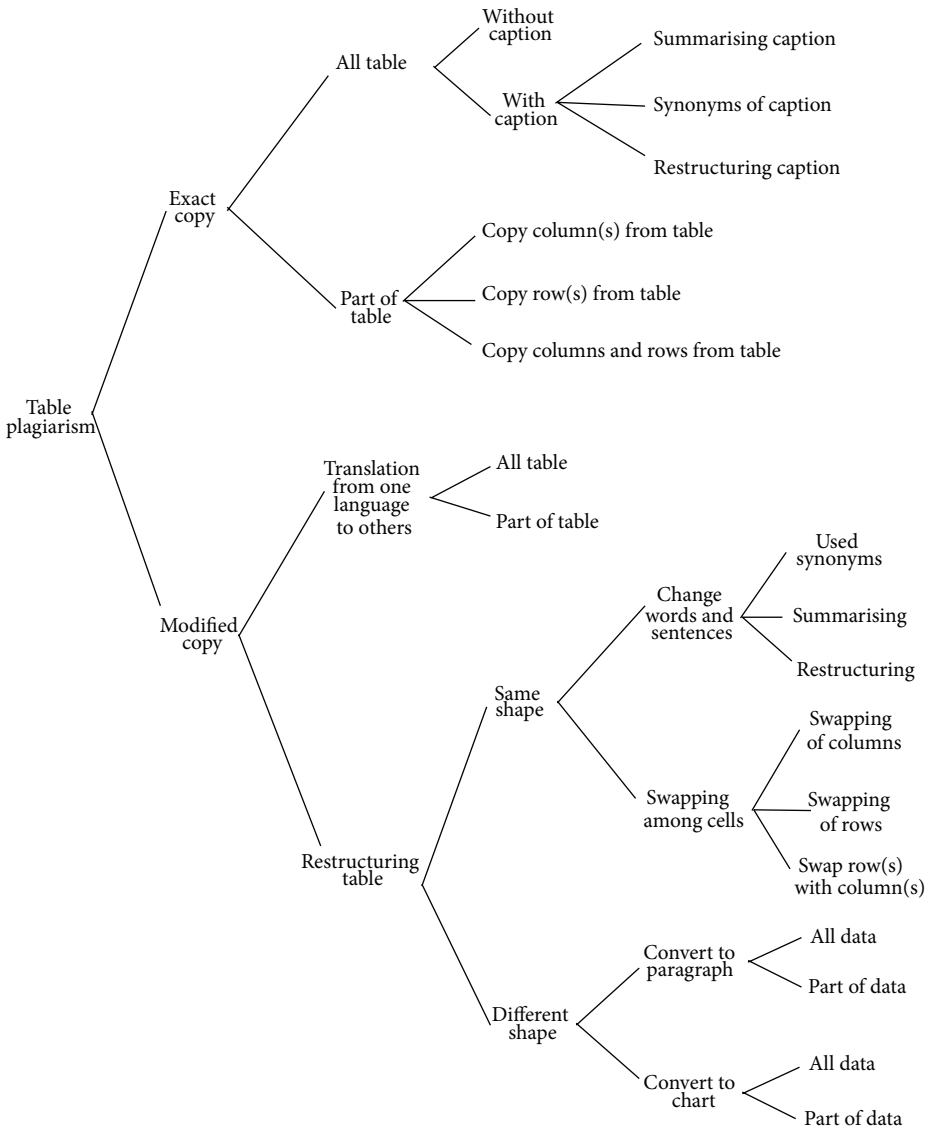


Figure 3: Taxonomy of table plagiarism.

The techniques for features extraction of image depend on the type of images such as chart or medical representation. Some techniques are applicable on two-dimensional plot of chart images while others work well for bar chart images. Hough transform technique is introduced as an approach to extract the features of bar chart images [23]. Some investigations are based on the edges of bars to extract the features [24, 26]. The learning-based method is established to recognize the chart images [22]. The features of bar chart images can be extracted by describing the height and width of each bar, which is applied on statistical images to determine the similarities [27]. Meanwhile, other techniques focused on geometric features rather than data and information of scientific bar chart images [28].

Currently, several techniques are developed to extract the features from medical images. The texture is one of the visual contents of a medical image used in content-based image retrieval (CBIR) to represent the image effectively for searching and recovering similar areas [29]. Gray-level statistical matrix technique is applied to extract the texture information for the content-based retrieval of mammograms from the MIAS database [30]. 3D texture features technique based on the cooccurrence matrixes of the gray-level, gradient, and curvature information regarding the nodule volume data for classifying the malignancy from benign is introduced [31, 32].

3. Plagiarism Taxonomy and Patterns

Three types of graphic plagiarism such as figure, chart, and table are important to emphasize. Each type highlights different levels of plagiarism. The patterns and types of plagiarism for figures, charts, and tables are presented as

taxonomy. Some kinds of text plagiarism are also evaluated [33]. The methods for detecting passages of text plagiarism for documents without appropriate citations are also suggested. Taxonomy is further extended to cover other types of plagiarism [4]. The taxonomy presented in various studies majorly demonstrates literal and intelligent plagiarism, where each kind includes many patterns of plagiarism. However, we are interested in detecting plagiarism of charts, as well as their taxonomy, figures, and tables. Alternatively, charts (pie, bar, and line) can be considered as one of the methods for representing the data and information of experimental results or comparing among techniques which are copied from other references without citation. Therefore, plagiarism of charts can be formulated in several forms to manifest the same information in various shapes. Taxonomy of chart plagiarism demonstrates many patterns and models which may be used to plagiarize the data of chart image.

Plagiarism patterns of chart, figure, and table are divided into Exact Copy and Modified Copy prototypes. The Exact Copy patterns of plagiarism are defined as the direct quote of data from other works without referencing, where copy and paste of the whole or part of the information image is performed. Simplicity is one of the important attributes of this type of plagiarism. Besides, this type of plagiarism does not require much time to hide the academic crime. The other type of graphic plagiarism is the Modified Copy for information of chart, figure, and table. This is more intelligently performed than the previous one because the same data can be formulated in many ways to exhibit the work in a different style than the original one. The goal of these intelligent means is that the plagiarist attempts to deceive the readers by doing some changes, such as translation from other languages or generating another shape for the same data.

The Modified Copy plagiarisms are primarily divided into translation and restructuring. In this research, new types of copying are organized by taxonomy which explains various patterns of graphic plagiarism. Furthermore, the primary focus of the bar chart image is to detect the proportion of plagarism. Figures 1, 2, and 3 depict the taxonomy of chart, figure, and table plagiarism, respectively.

4. Methodology

The methodology of bar chart plagiarism detection as shown in Figure 4 consists of three main stages, namely, planning and collection, feature extraction, and development with system evaluation. In the planning and collection stage, various patterns of graphical plagiarism via taxonomy of chart, figures, and table are presented (Figures 1, 2, and 3). The taxonomy of chart plagiarism explains different formulations which plagiarize the data of bar chart images. Therefore, varieties of bar chart images are collected and data sets are considered. The gatherings of the data sets consist of 100 bar chart images for storing in databases and twenty images for query including all possibilities of plagiarism for bar chart images. These data sets are collected from different resources such as thesis, which represented various types of bar chart images in 2D and 3D. Besides, vertical and horizontal bar chart images as shown in Figure 5 are taken into account.

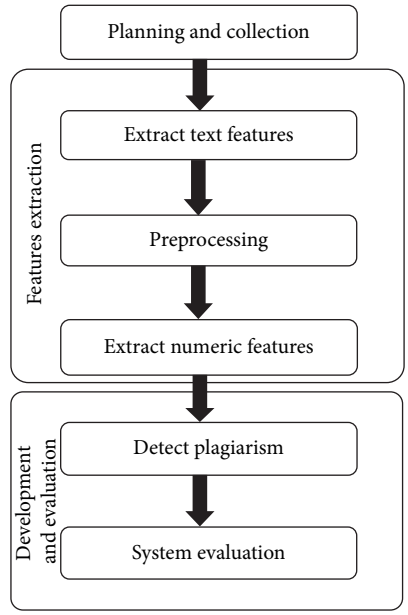
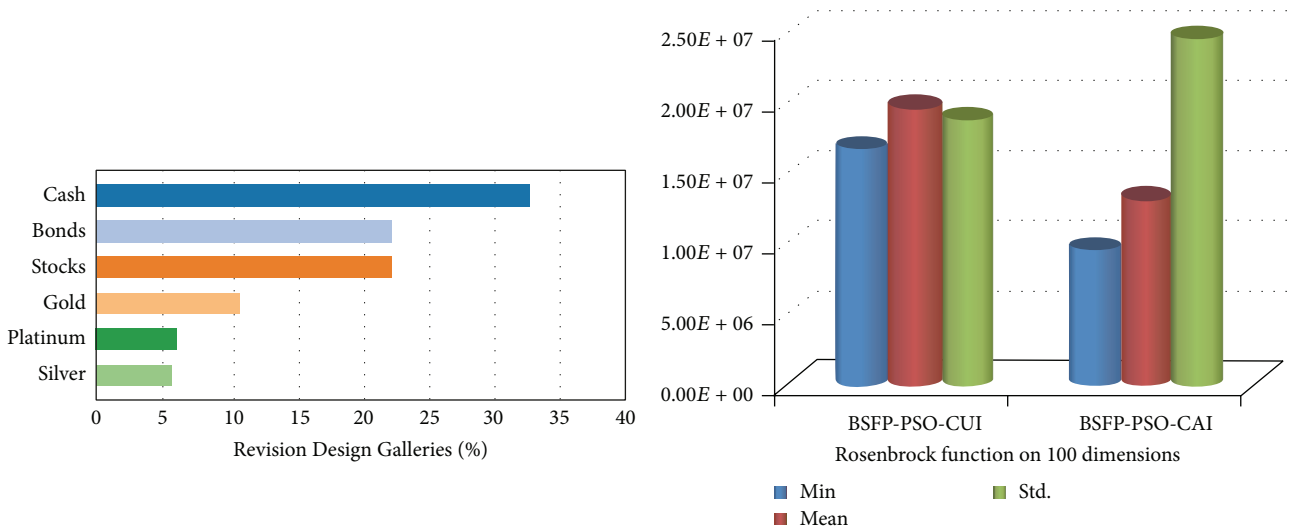


Figure 4: Flowchart of methodology.

In the feature extraction stage, the features of bar chart images are used to detect plagiarism. Various types of bar chart images are analyzed to detect the features of image. The bar chart images inferred to acquire maximum of ten features representing the information and the data of image. These features are common in different types of bar chart images, for instance, in 2D and 3D images. However, the number of uses for these features may different from each other.

The features extraction is an essential process to get the data from images which can be utilized to detect the rate of plagiarized data. Therefore, these ten features are categorized into low-level and high-level features. The low-level features refer to the text features of bar chart. The text features are the text which can be used in image to represent the information and data such as caption of image, label of each bar, label of coordinates, and values on coordinates. Generally, they are extracted from bar chart images using OCR tool. Conversely, the high-level features referring to numeric features cannot be extracted using OCR tool. The extraction of numeric features requires a relationship between the text and graphic components. The numeric features include values of bars in image. Each bar in an image has three numeric features which can be extracted by the proposed technique depending on Start, End, and Exact values. The Start and End values represent the first and last values while the Exact one corresponds to the real value of the bar. For instance, the text features for image in Figure 5 are *"Figure 10: Revision Design Galleries," "Silver, Platinum, ..., Cash,"* and *"0%, 5%, ..., 40%"* which represent caption of image and label of coordinates X and Y, respectively. Meanwhile, the numeric features represent the value of each bar, which can be detected by Start, End, and Exact features. For example, the Start, End, and Exact features of bar Silver are 5%, 10%, and 6%, respectively.

These features are used to detect the proportion of plagiarism for bar chart image. The extraction of Start, End, and



(a) (b)

Figure 5: Samples of data set [17, 36].

Exact values necessitates preprocessing of bar chart image to the adjacent coordinates of the image. Image scanning is then performed to detect the length of each bar in order to find the numeric features for each bar. Storing of the features in databases depends on the type of features whether numeric or text. The features which are extracted by the proposed technique are represented as vectors, while the text features that are extracted by OCR are characterized as string.

The detection methods for text plagiarism are mainly categorized based on character, semantics, structure, citation, cluster, cross language, and syntax. Comparatively, the smaller number of textual components than normal paragraph text allows us to use character-based methods to detect plagiarism of bar chart images. The character-based methods depend on character matching approaches to exactly or partially detect the identical string for features of bar chart images. Various algorithms of plagiarism are adopted in the text as character n-gram to identify the similarity between two strings based on the number of identical characters of features. Some researchers use 8-gram and 5-gram techniques [34, 35] for matching strings to detect plagiarism. We used 2-gram technique to detect plagiarism of bar chart images. This technique is used to represent the text features of bar chart images. Different similarity measures can be used to obtain the similarity for numeric features such as Euclidean distance, Jaccard, or cosine coefficient. The Euclidean distance is calculated by the following:

$$\mathrm{Ec}(x, y) = \sqrt{\sum_i |x_i - y_i|^2}. \quad (1)$$

Once the detection and storing of the proportion of plagiarism are completed then the performance of the system is evaluated. The performance is evaluated by overlapping of features using the relation of Precision and Recall given by the following:

$$\begin{aligned} \text{Recall} &= \frac{\text{Relavent Documents Retreived}}{\text{All Relavent Documents}}, \\ \text{Precision} &= \frac{\text{Relavent Documents Retreived}}{\text{All Documents Retrived}}. \end{aligned} \quad (2)$$

5. Experimental Results

The bar chart plagiarism detections are carried out in four main stages such as submission of query images, feature extraction of bar chart image, plagiarism detection, and highlighting results. The first stage is to submit various types of query images covering different kinds of possible plagiarism to detect and judge plagiarism of bar chart images, while the features of query are extracted in the second stage. The third stage includes detection of plagiarism by using word 2-gram and Euclidean distance techniques. Finally, the features of query bar chart image that are plagiarized from others are highlighted and the proportion of the similarity is displayed.

The bar chart plagiarism is further divided into *Exact Copy* and *Modified Copy* as explained in taxonomy. Figure 6(a) shows the query image while Figure 6(b) depicts the plagiarized images detected by the system. The first plagiarized image is similar to the whole data in the image while the second plagiarized image contains the same data that was plagiarized but presented as a horizontal bar chart image. The system extracts the features of query image and detects the proportion of plagiarism depending on Start, End, and Exact values for each bar as well as the label of each bar. The system highlights the data and information that are plagiarized and provides the proportion of plagiarism.

One of the patterns of plagiarism derived from *Modified Copy* is the stealing by changing scales. Each bar is modified

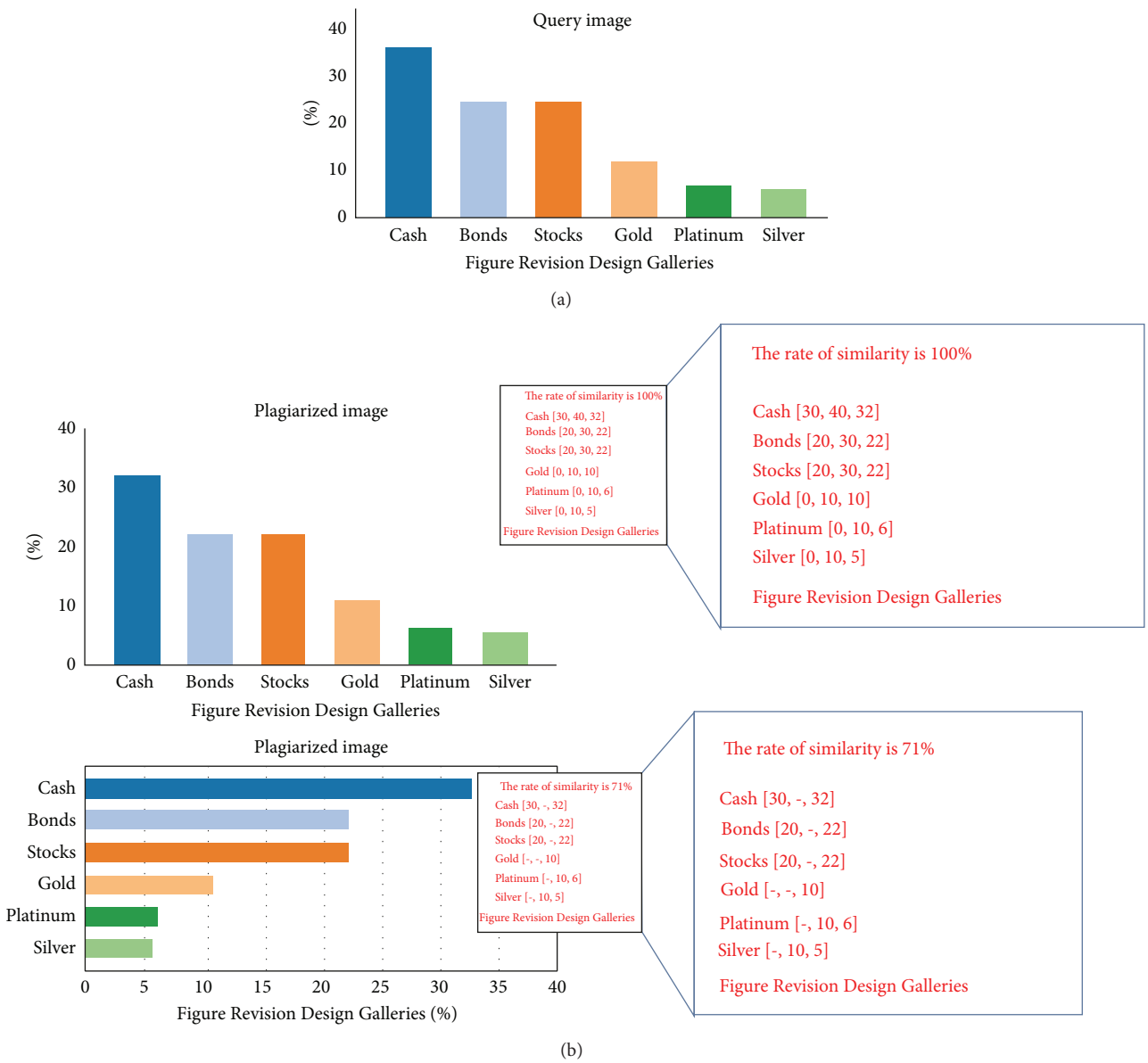


Figure 6: Plagiarism detection for one of Exact Copy patterns as plagiarism of the whole data of image.

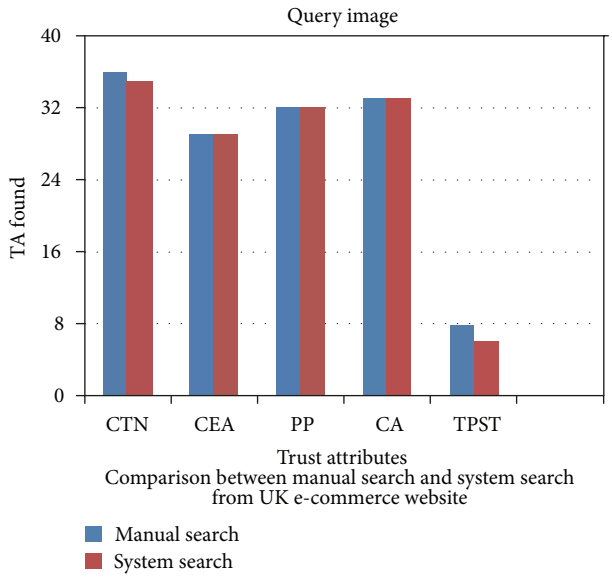
by plagiarists by changing Start and End values to be different from the original image. Figure 7 illustrates the significant role played by the Exact values to detect this type of plagiarism.

The plagiarists may use integration among patterns of possible bar chart plagiarism to present a more complex image which has the same data quoting from other works. Figure 8 displays the query image which is modified by changing colours and scales of bars as well as changing their location via swapping. The proposed system is capable of detecting this type of plagiarism and identifies the proportion of similarity.

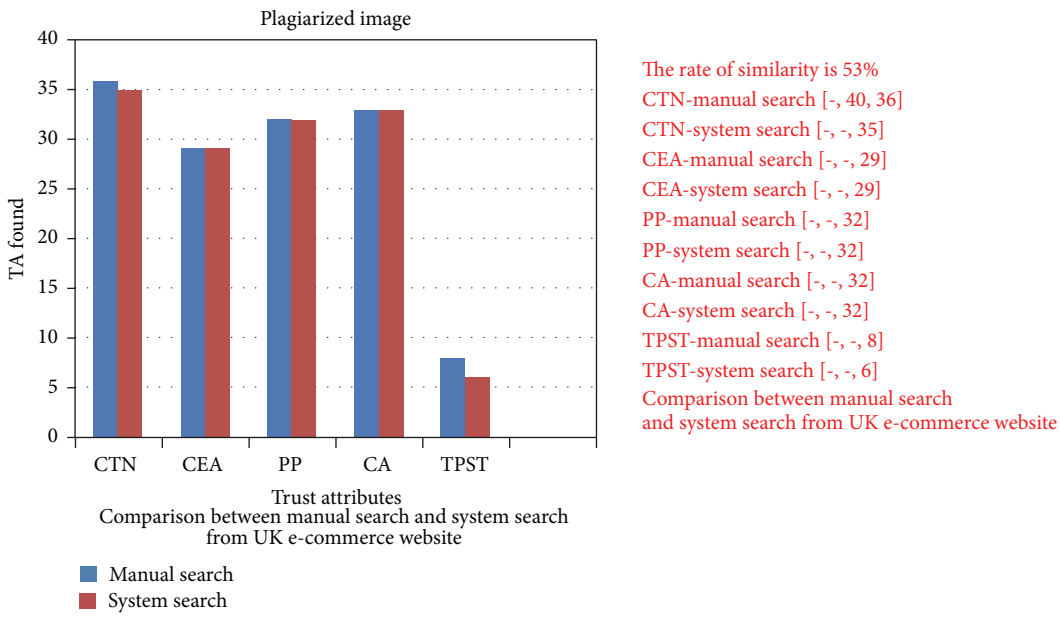
Figure 9 illustrates the performance of the system for plagiarism detection of *Exact Copy* and *Modified Copy* patterns, respectively.

6. Discussion

The state-of-the-art graphical plagiarism techniques and patterns are presented. The graphical plagiarism is considered one of the electronic crimes and thefts and the concepts of such stealing are newly viewed. Various important information and data can be represented as graphical forms such as results or frameworks for academic and business aspects. However, many systems of text plagiarism methods such as Turnitin are incapable of detecting plagiarism of images. In spite of the different styles of bar chart images, the extraction of features of image plays an important role in detecting plagiarism. Our proposed technique which is used to extract the numeric features played an essential role for bar chart plagiarism detection. The patterns of *Exact Copy* of bar chart



(a)



(b)

Figure 7: Plagiarism detection for one of the Modified Copy patterns which is the stealing by changing scales of image.

plagiarism detection including direct copy of all data or part of data are underscored. Plagiarism which is carried out by modifying caption of images via restructuring or summarizing for label sentences is emphasized. Alternatively, the patterns of *Modified Copy* which are regarded as more complicated than *Exact Copy* patterns are also analyzed. The difficulty of these patterns is the changing on image which appears as the same data and information in different forms. The restructuring of information for image within the same shape is also covered. The edition of bar chart images including the change of image bar colors or changing the bar locations either by swapping or via generating horizontal bars from vertical bars and vice versa is discussed in detail. Besides, more professional modification such as changing of scales on coordinates which is completely different from original one can be detected by the proposed method. In this case, the *Start* and *End* features of bars are completely different. Consequently, the *Exact* features as well as other attributes play significant role in detecting plagiarism in bar chart image.

7. Conclusion

We demonstrate the precise recognition of different plagiarized patterns in business documents using an intelligent bar chart detection system. The types and patterns of plagiarism

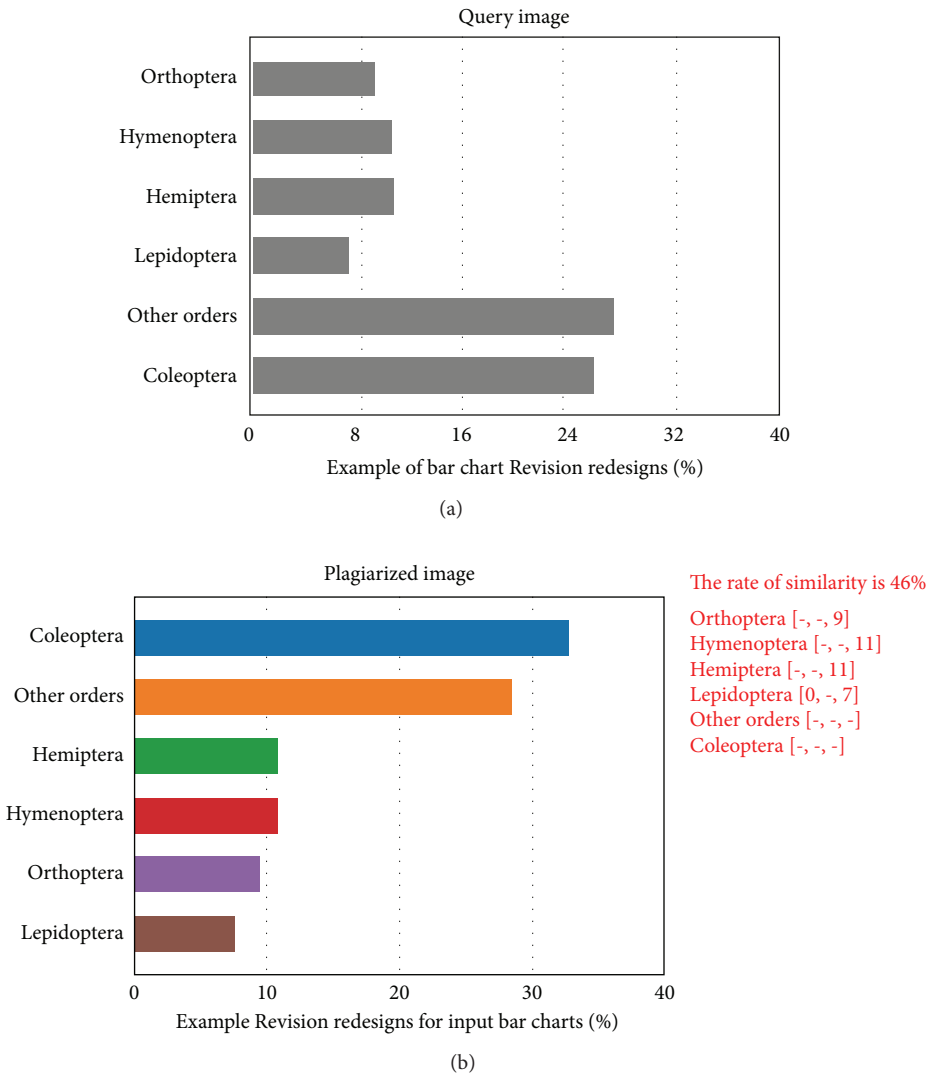


Figure 8: Plagiarism detection for integration among possible bar chart plagiarism.

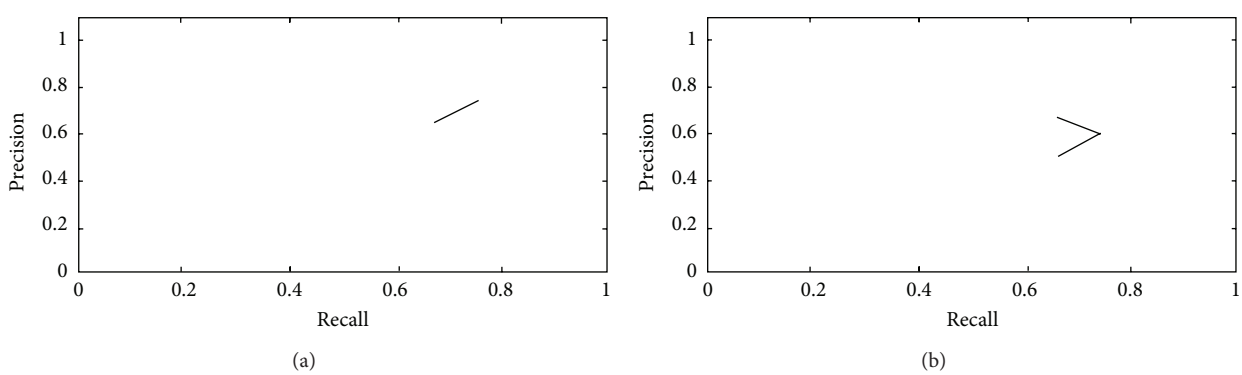


Figure 9: The evaluation of performance (a) for *Exact Copy* patterns and (b) for *Modified Copy* patterns.

are presented via taxonomy of figure, chart, and table. Various kinds of possible plagiarism are highlighted using taxonomy. Plagiarism of bar chart image as type of chart is detected by the newly proposed technique. It is established that the present technique is capable of extracting the features from a bar chart image which cannot be pulled out using OCR tool. Our technique first recognizes the connection between the text and graphical components to extract the Start, End, and Exact value for each bar. Using word 2-gram and Euclidean distance methods the accurate detection of plagiarism is performed. The detection of plagiarism is based on ten striking features. The system is capable of detecting different levels of plagiarism not only copy and paste of bar chart image but also modification on images such as changing color or scales. The present system efficiently and accurately distinguishes other possible alteration administered on these images such as swapping among bars location and even changes on caption via summarizing and restructuring. The proposed technique may be useful for intelligent plagiarism detection in business and academic documents.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this work through Research Group no. RGP-264. The authors are also thankful to Ministry of Science and Technology Innovation (MOSTI), Malaysia, and Research Management Center (RMC), Universiti Teknologi Malaysia (UTM), Johor, Malaysia, for their technical support and expertise in conducting this research.

References

[1] A. M. El Tahir Ali, H. M. D. Abdulla, and V. Snasel, "Survey of plagiarism detection methods," in *Proceedings of the 5th Asia Modelling Symposium (AMS '05)*, pp. 39–42, May 2011.

[2] A. Rehman and T. Saba, "Features extraction for soccer video semantic analysis: current achievements and remaining issues," *Artificial Intelligence Review*, vol. 41, no. 3, pp. 451–461, 2014.

[3] A. Rehman and T. Saba, "Evaluation of artificial intelligent techniques to secure information in enterprises," *Artificial Intelligence Review*, pp. 1–16, 2012.

[4] S. M. Alzahrani, N. Salim, and A. Abraham, "Understanding plagiarism linguistic patterns, textual features, and detection methods," *IEEE Transactions on Systems, Man and Cybernetics C: Applications and Reviews*, vol. 42, no. 2, pp. 133–149, 2012.

[5] T. Saba and A. Altameem, "Analysis of vision based systems to detect real time goal events in soccer videos," *Applied Artificial Intelligence*, vol. 27, no. 7, pp. 656–667, 2013.

[6] K. J. Ottenstein, "An algorithmic approach to the detection and prevention of plagiarism," *SIGCSE Bulletin*, vol. 8, pp. 30–41, 1976.

[7] A. Parker and J. O. Hamblen, "Computer algorithms for plagiarism detection," *IEEE Transactions on Education*, vol. 32, pp. 94–99, 1989.

[8] J. P. Bao, J. Y. Shen, X. D. Liu, and Q. B. Song, "Survey on natural language text copy detection," *Journal of Software*, vol. 14, no. 10, pp. 1753–1760, 2003.

[9] T. Saba and A. Rehman, *Machine Learning and Script Recognition*, Lambert Academic Publisher, 2012.

[10] W. Huang, S. Zong, and C. L. Tan, "Chart image classification using multiple-instance learning," in *Proceedings of the IEEE Workshop on Applications of Computer Vision (WACV '07)*, p. 27, Austin, Tex, USA, February 2007.

[11] T. Saba, A. Rehman, and M. Elarbi-Boudihir, "Methods and strategies on off-line cursive touched characters segmentation: a directional review," *Artificial Intelligence Review*, 2011.

[12] A. Rehman, D. Mohammad, G. Sulong, and T. Saba, "Simple and effective techniques for core-region detection and slant correction in offline script recognition," in *Proceedings of the IEEE International Conference on Signal and Image Processing Applications (ICSIPA '09)*, pp. 15–20, Kuala Lumpur, Malaysia, November 2009.

[13] A. Rehman and T. Saba, "Document skew estimation and correction: analysis of techniques, common problems and possible solutions," *Applied Artificial Intelligence*, vol. 25, no. 9, pp. 769–787, 2011.

[14] A. Rehman and T. Saba, "Off-line cursive script recognition: current advances, comparisons and remaining problems," *Artificial Intelligence Review*, vol. 37, no. 4, pp. 261–288, 2012.

[15] T. Saba, A. Rehman, and G. Sulong, "Improved statistical features for cursive character recognition," *International Journal of Innovative Computing, Information and Control*, vol. 7, no. 9, pp. 5211–5224, 2011.

[16] T. Helmy, "A computational model for context-based image categorization and description," *International Journal of Image and Graphics*, vol. 12, no. 1, Article ID 1250001, 2012.

[17] M. Savva, N. Kong, A. Chhajta, F.-F. Li, M. Agrawala, and J. Heer, "ReVision: automated classification, analysis and redesign of chart images," in *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*, pp. 393–402, Santa Barbara, Calif, USA, October 2011.

[18] S. Elzer, S. Carberry, I. Zukerman, D. Chester, N. Green, and S. Demir, "A probabilistic framework for recognizing intention in information graphics," in *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI '05)*, pp. 1042–1047, Edinburgh, UK, August 2005.

[19] S. Carberry, S. Elzer, N. Green, K. McCoy, and D. Chester, "Understanding information graphics: a discourse-level problem," in *Proceedings of the 4th SIGdial Workshop of Discourse and Dialogue (SIGDIAL '03)*, pp. 1–12, Sapporo, Japan, July 2003.

[20] M. S. M. Rahim, A. Rehman, S. Ni'matus, F. Kurniawan, and T. Saba, "Region-based features extraction in ear biometrics," *International Journal of Academic Research*, vol. 4, no. 1, pp. 37–42, 2012.

[21] N. Yokokura and T. Watanabe, "Layout-based approach for extracting constructive elements of bar-charts," in *Graphics Recognition Algorithms and Systems*, K. Tombre and A. Chhabra, Eds., vol. 1389, pp. 163–174, Springer, Berlin, Germany, 1998.

[22] Y. Zhou and C. L. Tan, "Learning-based scientific chart recognition," in *Proceedings of the 4th IAPR International Workshop on Graphics Recognition (GREC '01)*, pp. 482–492, 2001.

[23] Y. P. Zhou and C. L. Tan, "Hough technique for bar charts detection and recognition in document images," in *Proceeding of the International Conference on Image Processing (ICIP '00)*, vol. 2, pp. 605–608, Vancouver, Canada, September 2000.

[24] W. Huang, C. L. Tan, and W. K. Leow, "Model-based chart image recognition," in *Graphics Recognition. Recent Advances and Perspectives*, J. Lladós and Y.-B. Kwon, Eds., vol. 3088, pp. 87–99, Springer, Berlin, Germany, 2004.

[25] A. Mishchenko and N. Vassilieva, "Model-based chart image classification," in *Advances in Visual Computing*, G. Bebis, R. Boyle, B. Parvin et al., Eds., vol. 6939 of *Lecture Notes in Computer Science*, pp. 476–485, Springer, Berlin, Germany, 2011.

[26] W. Huang and C. L. Tan, "A system for understanding imaged infographics and its applications," in *Proceedings of the ACM Symposium on Document Engineering*, pp. 9–18, Manitoba, Canada, August 2007.

[27] M. M. Hassan and W. Al Khatib, "Similarity searching in statistical figures based on extracted meta data," in *Proceedings of the Computer Graphics, Imaging and Visualisation (CGIV '07)*, pp. 329–334, Bangkok, Thailand, August 2007.

[28] L. Yang, W. Huang, and C. L. Tan, "Semi-automatic ground truth generation for chart image recognition," in *Proceedings of the 7th international conference on Document Analysis Systems*, Nelson, New Zealand, 2006.

[29] G. D. Tourassi, "Journey toward computer-aided diagnosis: role of image texture analysis," *Radiology*, vol. 213, no. 2, pp. 317–320, 1999.

[30] D. A. Chandy, J. S. Johnson, and S. E. Selvan, "Texture feature extraction using gray level statistical matrix for content-based mammogram retrieval," *Multimedia Tools and Applications*, vol. 72, no. 2, pp. 2011–2024, 2014.

[31] F. Han, H. Wang, B. Song et al., "A new 3D texture feature based computer-aided diagnosis approach to differentiate pulmonary nodules," in *Medical Imaging: Computer-Aided Diagnosis*, Proceedings of SPIE, San Diego, Calif , USA, 2014.

[32] F. Han, H. Wang, B. Song et al., "Efficient 3D texture feature extraction from CT images for computer-aided diagnosis of pulmonary nodules," in *Proceedings of the SPIE, Medical Imaging: Computer-Aided Diagnosis*, vol. 9035, pp. 1–7, 2014.

[33] S. M. Eissen, B. Stein, and M. Kulig, "Plagiarism detection without reference collections," in *Advances in Data Analysis*, R. Decker and H.-J. Lenz, Eds., pp. 359–366, Springer, Berlin, Germany, 2007.

[34] J. Kasprzak, M. Brandejs, and M. Křipač, "Finding plagiarism by evaluating document similarities," in *Proceedings of the 3rd Workshop on Uncovering Plagiarism, Authorship and Social Software Misuse (PAN '09)*, pp. 24–28, Donostia, Spain, September 2009.

[35] C. Basile, D. Benedetto, E. Caglioti, G. Cristadoro, and M. D. Esposti, "A plagiarism detection procedure in three steps: selection, matches and squares," Donostia, Spain, 2009.

[36] Z. A. Hamed and S. Z. Mohd Hashim, *Hybrid particle swarm optimization and black stork foraging for functional neural fuzzy network learning enhancement [UTM Thesis]*, 2012.

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2014, Article ID 606570, 10 pages
http://dx.doi.org/10.1155/2014/606570

Research Article

Improving RLRN Image Splicing Detection with the Use of PCA and Kernel PCA

Zahra Moghaddasi, Hamid A. Jalab, Rafidah Md Noor, and Saeed Aghabozorgi

Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

Correspondence should be addressed to Hamid A. Jalab; hamidjalab@um.edu.my

Received 2 June 2014; Revised 26 July 2014; Accepted 13 August 2014; Published 14 September 2014

Academic Editor: Iftikhar Ahmad

Digital image forgery is becoming easier to perform because of the rapid development of various manipulation tools. Image splicing is one of the most prevalent techniques. Digital images had lost their trustability, and researches have exerted considerable effort to regain such trustability by focusing mostly on algorithms. However, most of the proposed algorithms are incapable of handling high dimensionality and redundancy in the extracted features. Moreover, existing algorithms are limited by high computational time. This study focuses on improving one of the image splicing detection algorithms, that is, the run length run number algorithm (RLRN), by applying two dimension reduction methods, namely, principal component analysis (PCA) and kernel PCA. Support vector machine is used to distinguish between authentic and spliced images. Results show that kernel PCA is a nonlinear dimension reduction method that has the best effect on R, G, B, and Y channels and gray-scale images.

1. Introduction

For many decades, photography has served a vital function in the lives of people and has been considered as one of the most important revolutions in recording moments. However, photographs have recently lost their trust because of the increasing use of manipulation tools, which have made photographs less trustable. The history of image tampering (i.e., creating photographs that never really happened in real life) is as old as the art of photography itself. Images have been applied for malicious purposes in many instances. In 2003, the Los Angeles Times printed on its front cover an image from photojournalist, Brian Walski, which showed a British soldier in Iraq trying to control a crowd of civilians in a passionate manner. However, the depicted moment never happened. The photograph was a composite of two different photographs merged to create an appealing image. The image tampering was discovered, and Walski was fired [1]. In 2009, the Brazilian newspaper Folha de Sao Paulo published an article that contained a spliced image created from different gray-scale images to show how the Brazilian Chief of Staff actively participated in the resistance against the military regime through such actions as the planning of and preparation for robberies and kidnappings [1].

An innovative research area called digital image forensics emerged from the necessity to regain the trustability of photographs. Digital image forensics is generally categorized into two basic groups: active methods and passive methods [2]. Active methods borrow additional information to insert into the original image, whereas passive methods require no prior information. Digital signatures and watermarking are the most common active detection methods. Passive methods focus on statistical modifications made on image content. Such modifications include splicing, retouching, healing, copy-move, and blurring. Image splicing is best defined as a combination of two or more parts of different photographs to create a new image, which is known as a spliced image. Image splicing deceives the human visual system to achieve its malicious purposes.

Studies were conducted on this topic, and several image splicing detection methods were proposed and developed. Farid [3] considered speech signal splicing as a highly nonlinear process and applied higher order spectral analysis to solve this problem. Ng et al. [4] extended the previously mentioned scheme into an image splicing detection method with the assumption that the image splicing procedure is nonlinear and that the image involved is nonstationary. Their expanded

method achieved an unsatisfactory detection performance of 72%.

Fu et al. [5] proposed an image splicing detection approach using the Hilbert-Huang transform (HHT). They considered the high nonlinearity and nonstationary nature of the image splicing operation and merged this technique with a statistical natural image model on the basis of moments of characteristic functions with wavelet decomposition. Their method obtained an accuracy rate of 80.15%. Li et al. [6] combined two methods, namely, moment features based on first-order histogram of the image discrete wavelet transform (DWT) coefficients and HHT, to extract features for image splicing detection. Their method obtained an average accuracy rate of 85.87%.

Shi et al. [7] proposed a blind, passive, and natural image splicing detection model based on statistical feature extraction methods. Their model includes the Markov transition probabilities and moments of characteristic functions of wavelet subbands applied to different 2D arrays and to the 2D arrays of multisize block discrete cosine transform. Their model achieved a detection accuracy rate of 91.8%, which is a promising improvement in the image splicing detection area.

Zhang et al. [8] applied an idea in image steganalysis [9] by merging the Markov features and discrete cosine transform (DCT) features. They achieved a detection rate of 91.5%. He et al. [10] expanded the natural Markov-based model applied in [7] by capturing the interblock correlation between the block DCT coefficients proposed in [11] and merging it with the features generated from the DWT domain. They also reduced their features by applying a feature selection method called support vector machine recursive feature elimination (SVM-RFE), which achieved a detection rate of 93.55% on a digital video multimedia (DVMM) image dataset [12].

Dong et al. [13] proposed an approach based on statistical features obtained from the run length method and on image edge statistics from the blind image splicing detection method. He et al. [14] improved the method proposed in [13] and developed a detection algorithm based on approximate run length. Their results showed a moderate detection accuracy rate (75% versus 69%) but with a lesser amount of time than in the original algorithm (6D versus 12D). Zhao et al. [15] proposed applying the run length run number (RLRN) vectors in four directions of the chroma spaces because detecting spliced images in one color space is difficult. Their results showed that the features extracted from the chroma channels were more accurate than those extracted from the R, G, B, and Y channels.

The existing methods concentrate only on the actual image splicing detection techniques, and handling the extracted high dimensional and redundant features can be a difficult and time consuming process. Therefore, dimensionality reduction methods are applied to remove redundant information from the extracted features and to obtain the most discriminative information with less dimensionality.

In this paper, we compared the two dimension reduction methods (PCA versus kernel PCA) that were developed. This paper aims to evaluate the detection accuracy and computational time of PCA and kernel PCA on RLRN [15] in dimensionalities of 10 and 50. The rest of this paper is organized as follows. The proposed approach is described in the second section, the experimental results are discussed in the third section, and the conclusion and future works are presented in Section 4.

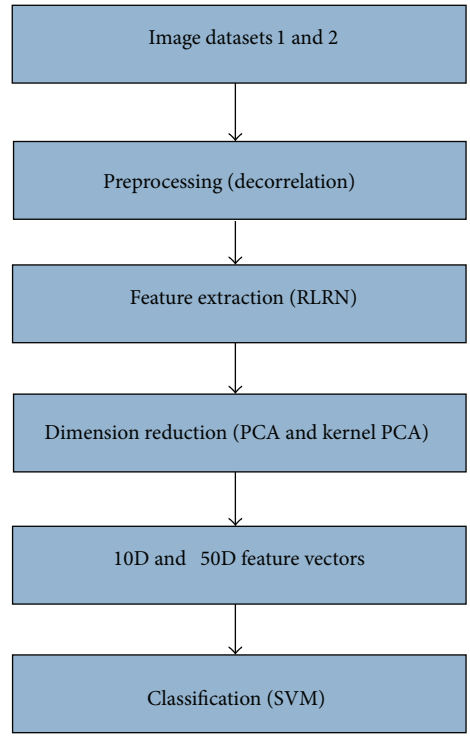


Figure 1: Proposed algorithm framework.

2. The Proposed Approach

Figure 1 demonstrates the approach proposed in this paper. The approach includes four phases: preprocessing, feature extraction, dimension reduction, and classification. The given image from image dataset 1 [12] or image dataset 2 [16] is initially preprocessed for extraction. The features are then extracted from the preprocessed image using RLRN [15]. The dimensionality is then reduced to 10 and 50 dimensions by applying PCA and kernel PCA. Finally, the n dimensional (nD) feature vector is used as input for the support vector machine (SVM) classifier to calculate the detection accuracy and computational time.

2.1. Preprocessing. Preprocessing improves image content by reducing undesired distortions and/or enhancing image features relevant to further processing. A difference between 2D arrays is used to reduce the correlation between image pixels/coefficients and the effect caused by the diversity in image content [15, 17]. The difference between 2D arrays is computed as follows [15, 17]:

$$E_h(i, j) = x(i, j) - x(i + 1, j)$$

$$E_v(i, j) = x(i, j) - x(i, j + 1)$$

$$E_d(i,j) = x(i,j) - x(i+1,j+1)$$
$$E_m(i,j) = x(i+1,j) - x(i,j+1), \quad (1)$$

where $x(i, j)$ represents the ith and jth element of the gray value of the images in the image matrix and $E_h(i, j)$, $E_v(i, j)$, $E_d(i, j)$, and $E_m(i, j)$ represent the elements along the horizontal, vertical, diagonal, and minor-diagonal directions, respectively.

2.2. Feature Extraction Methods. Image features include the global and local properties of an image such as average gray levels, intensity histogram shapes, circles, lines, texels, and contour shapes. Different methods have been developed to extract these features, and these methods have been applied in various image processing fields. Given that image splicing detection is a two-class problem (i.e., one class for authentic images and another class for spliced images), features extracted from images serve an important function in the detection and classification process, which aims to distinguish authentic images from spliced ones. In this paper, a recently used feature extraction method is presented to examine the effect of dimension reduction techniques on detection performance.

2.2.1. RLRN. RLRN is not extensively applied as a feature extraction method. However, the results which have been obtained in [15] show that this method can be used as an image splicing detection approach. The method was initially used in [18] in texture analysis to classify a set of terrain samples. A new run length algorithm was developed in [19] to extract texture features based on a multilevel dominant eigenvector estimation method, which improves classification accuracy.

This paper applied RLRN and its definition and mathematical equations are presented in this section. Zhao et al. [15] suggested using this method. A run in an image is defined as the number of pixels with the same gray level value in a specific direction. For a given image, a run length matrix $p_\theta(i, j)$ is defined as the number of runs with gray level i and run length j along a specific direction. Hence, the run length vector is defined as follows [15]:

$$p_{\theta r}(j) = \sum_{i=1}^{M} p_\theta(i,j) \quad 1 \le j \le N, \quad (2)$$

where M represents the number of gray levels and N is the maximum value of run lengths. Vector $p_{\theta r}(j)$ demonstrates the sum distribution of runs with length j in a given image. In the equation, run length represents the spread of the image structure and texture. The image with a long run length is smoother than that with a short run length because the latter has different regions with different structures. The gray level run length pixel number matrix is used to equally emphasize all run lengths in this paper. The matrix is defined as follows [15, 19]:

$$p_{\theta pr}(j) = \sum_{i=1}^{M} p_{\theta p}(i,j) \quad 1 \le j \le N, \quad (3)$$

where

$$p_{\theta p}(i,j) = p_\theta(i,j) \cdot j, \quad (4)$$

where $p_{\theta pr}(j)$ is the feature vector applied and referred to as RLRN. Four RLRN vectors were captured in four directions (0°, 45°, 90°, and 135°) from every channel (R, G, B, Y, Cb, and Cr) in image dataset 2 and from gray-scale images in image dataset 1 to distinguish the spliced images from the authentic images. The RLRN feature vector is shown as a 100D feature vector by obtaining the first 25 features of every vector in each orientation ($25 \times 4 = 100$).

The run length of an image represents its structure and texture, and the splicing procedure modifies the pixel correlations and image structure. Therefore, the RLRN feature extraction method can represent discontinuities and nonconformity and can be efficiently used as an image splicing detection method [13, 15].

2.3. Dimension Reduction Methods. Dimension reduction methods are applied to reduce feature dimensionality by eliminating redundant features and keeping important dimensions in the feature vector. Humans and machine learning methods find it difficult to interpret high dimensional data. Given that a feature matrix has rows that each represents a specific instance of an object and a large number of features exponentially increase the computational time. Thus, decreasing information into smaller sizes enhances method analysis and improves the training and testing phases during classification [20]. Several experiments were conducted to test and analyze this idea.

Different approaches such as finding the linear or nonlinear manifold that lies within the high dimensional data space can simplify interpretation. In this section, the PCA and kernel PCA are presented to improve the features extracted by the RLRN feature extraction technique discussed in the previous section. Moreover, linearity (PCA) and nonlinearity (kernel PCA) were considered in selecting the dimension reduction techniques to investigate their effects on the extracted features.

2.3.1. PCA. PCA is the most common and popular linear dimension reduction approach [21–23]. It has been used for years because of its conceptual simplicity and computation efficiency. The approach is applied in many areas such as noise reduction, pattern recognition, regression estimation, and image indexing [24]. It maps a dataset of n dimensions to a linear subspace with d dimensions, where $d < n$, and attempts to maintain most of the variability in the mapped dimensions. PCA is considered a second-order approach depending on the covariance matrix of the variables. The approach has different names in different fields such as singular value decomposition, Karhunen-Love transform, Hotelling transform, and empirical orthogonal function method [25].

PCA is based on finding the d orthogonal linear vectors, known as principal components, of n dimensions with maximum variance. Therefore, the number of reduced dimensions is not more than n. The approach works well if the most significant modes of variability are almost linear. Hence,

high dimensional samples are best remade from their low dimensional linear projections. Otherwise, PCA becomes ineffective if the most vital significant modes of variability are nonlinear [21]. In mathematical terms, PCA finds Y as the new feature vector set with d dimension ($d \leq D$), in which X is the original feature vector set with D dimension [26]:

$$Y = XM. \tag{5}$$

To find linear mapping M, PCA attempts to maximize the following function:

$$M^T \operatorname{cov}(X) M, \tag{6}$$

where $\operatorname{cov}(X)$ is the covariance of the original feature vector set X. However, M consists of d principal eigenvectors of the sample covariance matrix of the zero-mean data [26]. Therefore, the following eigenproblem must be solved for the d principal components λ:

$$\operatorname{cov}(X) M = \lambda M. \tag{7}$$

2.3.2. Kernel PCA. PCA is a linear dimension reduction method. Some datasets have a nonlinear nature, and PCA cannot reduce the dimensions of these datasets efficiently. Thus, kernel PCA was designed to address this problem. Kernel PCA was applied in some pattern recognition experiments [24] and exhibited better recognition rates than linear PCA. Kernel PCA is a nonlinear form of PCA that attempts to identify complicated correlations between given features. It computes principal components in the original dataset through nonlinear mapping. It also discovers major components that are nonlinear in relation to the input space by running, which results from nonlinear mapping in which the low dimensional hidden structures are likely to be simple [21]. Kernel PCA locates the principal eigenvectors of the kernel matrix instead of the covariance matrix [26]. Thus, the computational complexity of kernel PCA is independent of the dimensionality of the feature set, which allows it to work on feature sets with different possible dimensionalities. Kernel PCA does not require any nonlinear optimization; it only needs to solve an eigenvalue problem as in the case of standard PCA. Thus, kernel PCA is free of local minima trap during training. The original feature set must be mapped to a higher dimensional feature set to calculate kernel PCA [24]:

$$\Phi : R^N \longrightarrow F, \quad x \longrightarrow X. \tag{8}$$

Then, the covariance matrix of data is calculated to obtain the principal components by solving the eigenvalue problem using the following equations:

$$C_F = \frac{1}{N}\sum_{1}^{N} \Phi(x_i)\, \Phi(x_i)^T$$
$$C_F v = \lambda v. \tag{9}$$

Subsequently, the eigenvector can be expressed as a linear combination of features:

$$v = \sum_{1}^{N} \alpha_i \Phi(x_i)$$
$$\alpha_i = \frac{1}{\lambda N} v. \tag{10}$$

Therefore, the kernel matrix is defined as follows:

$$k_{ij} = \mathscr{K}(x_i, x_j) = \left(\Phi(x_i) \cdot \Phi(x_j)\right) = \Phi(x_i)^T \Phi(x_j), \tag{11}$$

where k_{ij} represents the elements of kernel matrix K, x is the feature set, and $\mathscr{K}$ is the kernel function with conditions that result in a positive semidefinite kernel K [26]. $\Phi(x_i)$ may not be zero-mean such that the features must be centered. The corresponding kernel is obtained using the following equation:

$$k_{ij}^c = k_{ij} - \frac{1}{N}\sum_{i=1}^{N} k_{ik} - \frac{1}{N}\sum_{j=1}^{N} k_{jk} + \frac{1}{N^2}\sum_{i,k}^{N} k_{ik}. \tag{12}$$

Consequently, the following equation represents the low dimensional feature set y_i:

$$y_i = \sum_{i=1}^{N} \alpha_{ji} \mathscr{K}(x, x_i), \tag{13}$$

where α_{ji} represents the jth value in the vector α_i.

High correlations are generally found among the extracted features using RLRN. PCA and kernel PCA are applied to reduce the correlations by eliminating the information redundancies from the features. Figure 2 shows the standard deviation distribution of the features extracted from gray-scale images (R and Cb channels of the colored images) before and after applying PCA and kernel PCA, respectively. The standard deviation measures and shows how data are spread out from the mean. In this case, a high standard deviation implies a high correlation between the features. Figure 2 shows that the original features are highly correlated and their standard deviations are spread over a wide range in 100D. After applying PCA, the standard deviations mostly concentrate on the first few features and decrease as dimensionality increases. However, the standard deviations are still high and considerable. In contrast, the standard deviations greatly reduced after applying kernel PCA on the original features. The features after applying kernel PCA were obviously highly uncorrelated.

2.4. Classifier. SVM is one of the most popular supervised machine learning algorithms applied in pattern recognition. The Matlab codes for this classifier are available in [27]. In this work, LIBSVM was specifically used as a classifier, while radial basis function was used as a kernel function. A grid search method was applied to obtain the best value for the c and g parameters. All authentic images were labeled as −1 and all spliced images were labeled as +1 during the classification process.

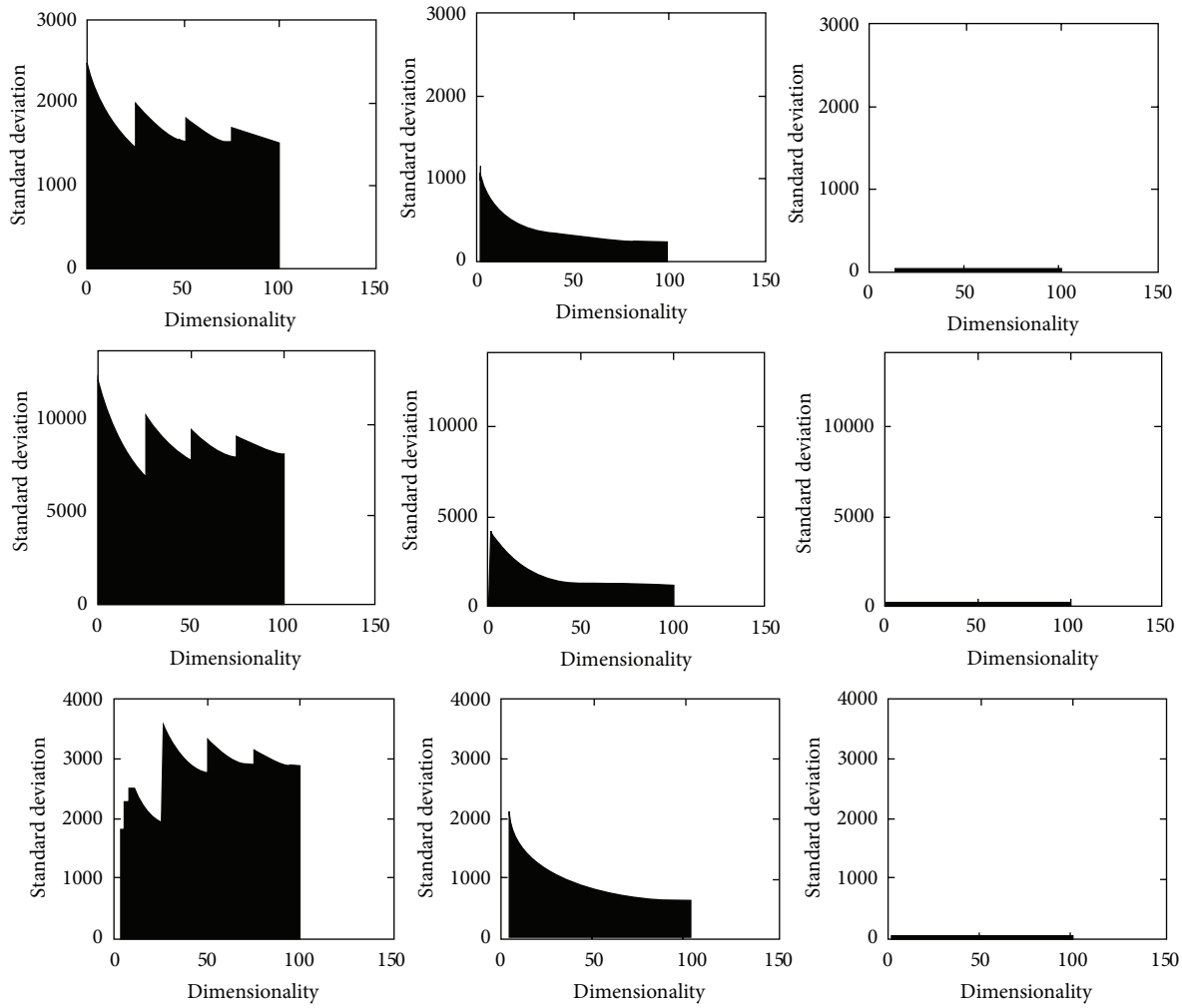


Figure 2: Standard deviation distributions of the extracted features. The rows indicate the standard deviation distributions of the features extracted from gray-scale images, red channel, and Cb channel of the colored images, respectively. The first column indicates the original features. The second column shows the features after applying PCA, while the third column shows the features after applying kernel PCA.

3. Experimental Results

A set of experiments that demonstrate the effectiveness of the proposed approach are described. Our classification system was implemented using MATLAB R2012a on a 2.40 GHz Intel (R) Core i5 processor with 4 GB RAM on a Windows 7 platform.

3.1. Image Dataset. Two image datasets (gray and colored) were applied to evaluate the proposed method. The first image dataset was the Columbia Image Splicing Detection Evaluation Dataset provided by the Digital Video MultiMedia (DVMM) Laboratory, Columbia University (2007) [12]. This dataset contained 1845 gray-scale images (933 authentic images and 912 spliced images) in BMP format. DVMM was the only gray-scale image dataset designed for image splicing detection evaluation. Almost all papers applied DVMM, so it was also used in this work for better comparison with other methods. Another image dataset designed by the Chinese Academy of Sciences, Institute of Automation (CASIA), with high resolution images was also applied.

The CASIA tampered image detection evaluation database [16] is another image dataset designed to evaluate image splicing detection methods. Version 1.0 of this dataset included 1721 color images (800 authentic images and 921 spliced images) with 384 pixels × 256 pixels in JPEG format and was used in our experiments to evaluate the proposed approach. Examples of both image datasets are presented in Figure 3 (DVMM image dataset) and Figure 4 (CASIA image dataset). The first row consists of authentic images, while the second row consists of spliced images [12, 16]. Each image dataset is divided into two groups during the classification process: training (5/6 of the images) and testing (1/6 of the images). These groups were randomly selected to reduce the nondeterministic properties of the classifier.

3.2. Classification. For satisfactory results, PCA and kernel PCA were applied on RLRN feature extraction method with different dimensionalities (10D and 50D) to evaluate the detection accuracy and computational time. The results were presented in three tables. The true positive (TP) and true negative (TN) represent the detection rate of the authentic

Figure 3: Examples of DVMM image dataset.

Figure 4: Examples of CASIA image dataset.

Table 1: Detection accuracy and computational time for RLRN with 100D.

	100D			
	TP (%)	TN (%)	Acc. (%)	Time (s)
R	63.91	74.68	69.30	388.43
G	57.89	71.43	64.66	402.83
B	55.64	75.32	65.48	426.27
Gray	70.51	75.00	72.75	489.72
Y	58.65	67.53	63.09	408.01
Cb	90.23	88.96	89.59	258.73
Cr	94.74	92.86	93.80	252.32

and spliced images, respectively. Accuracy (Acc.) represents the average detection rate. Computational time is represented in seconds (s) in each table.

Table 1 illustrates the results from the original dimension of the RLRN method with 100D obtained from different channels (RGB, gray, and YCbCr). The results show that the detection accuracies in the R, G, and B channels varied in a close range (64.66% to 69.30%), which verifies the strong correlation among the three channels because of the color filter array interpolation process. The luma channel (Y) that is correlated with the RGB channel (i.e., Y is a linear combination of R, G, and B channels) [15] also shows similar results (63.09%) with RGB. The results obtained from the Cb and Cr channels exhibited the best detection accuracies (89.59% and 93.80%) among the channels, since RLRN is more sensitive to the chroma channels (Cb and Cr) than the luma one [15]. However, the computational times for the Cb and Cr channels were the least among the channels.

Table 2 presents the detection accuracy and computational time results from applying PCA on RLRN with 10 and

Table 2: Detection accuracy and computational time for RLRN with PCA in 10D and 50D.

PCA	10D				50D			
	TP (%)	TN (%)	Acc. (%)	Time (s)	TP (%)	TN (%)	Acc. (%)	Time (s)
R	60.90	73.38	67.14	447.86	66.17	74.68	70.43	302.57
G	55.63	72.08	63.86	457.46	61.65	77.92	69.79	270.28
B	67.67	69.48	68.57	504.21	64.66	72.73	68.69	293.34
Gray	67.94	75.00	71.43	462.49	65.38	73.68	69.53	297.29
Y	54.89	72.73	63.81	449.55	59.40	70.78	65.09	275.09
Cb	81.20	87.01	84.11	203.25	87.97	93.51	90.74	182.66
Cr	83.46	92.21	87.83	187.71	90.98	87.66	89.32	187.93

Table 3: Detection accuracy and computational time for RLRN with kernel PCA in 10D and 50D.

Kernel PCA	10D				50D			
	TP (%)	TN (%)	Acc. (%)	Time (s)	TP (%)	TN (%)	Acc. (%)	Time (s)
R	84.96	83.12	84.04	122.07	85.71	90.91	88.31	183.91
G	84.96	83.12	84.04	114.4	87.22	90.26	88.74	193.37
B	84.96	83.12	84.04	125.67	85.71	90.91	88.31	223.77
Gray	100	0	50.00	63.87	90.38	86.18	**88.28**	236.49
Y	82.71	86.36	84.54	112.78	87.22	90.26	88.74	188.15
Cb	84.96	83.12	84.04	111.67	87.81	90.91	89.36	186.7
Cr	84.96	83.12	84.04	111.58	85.71	90.91	88.31	186.62

50 dimensions. The results indicate the same behavior among different channels with the results obtained from 100D. A reduction in detection accuracy was observed for 10D compared with 100D. All detection accuracies slightly increased for 50D, except for the gray and Cr channels. Computational time in 50D also reduced compared with 100D. The results presented in Table 2 indicate a high correlation between the features after PCA application, which was proven in Figure 3.

Table 3 demonstrates the detection accuracy and computational time that resulted from combining the RLRN method with kernel PCA. These results show that gray-scale images with 50D exhibited a considerable increase in detection accuracy compared with those obtained from the original and the combined methods (88.28% versus 72.73% and 69.53%), respectively. The R, G, B, and Y channels also exhibited a substantial growth of 25% in detection accuracy. These results were anticipated from Figure 3, which indicated a low correlation among features after kernel PCA application.

However, the Cb and Cr channels did not follow the same trend and exhibited a slight decrease in 50D. The other dimensionalities were also tested for the Cb and Cr channels because of this decrease, and the results indicated that the optimal dimensionality for these two channels was in 95D with detection accuracy of 92.68%. Thus, the Cb and Cr channels contain important information focused on the first 95 dimensions. All computational time generally reduced for 10D and 50D in comparison with 100D.

Combining the RLRN method with kernel PCA in 50D generally exhibited the best results among the methods, which verifies the nonlinear nature of RLRN. Figures 5, 6, and 7 demonstrate the receiver operating characteristic (ROC) curves for gray-scale images and R and Cr channels, respectively. A comparison of the features extracted from the original feature extraction method in 100D was made in each figure with the extracted features from the merged ones with PCA and kernel PCA dimension reduction methods in 50D. The results obtained from 10D were ignored because of the high amount of results presented here. Figure 5 indicates the best effect of kernel PCA on RLRN, while PCA has almost the same results with the original features.

The R, G, B, and Y channels have similar behaviors; the R channel was selected to represent the ROC in Figure 6. According to this ROC, the R, G, B, and Y channels performed better when the dimensionality of the features was reduced to 50D by applying kernel PCA.

Figure 7 shows an inverse effect of kernel PCA on the obtained features from the Cb and Cr channels. The results show that the original features performed the best, and the dimension reduction methods did not improve accuracy rate.

3.3. Comparison with Other Methods. A comparison of some of state-of-the-art image splicing detection methods was conducted for a comprehensive evaluation of the entire system. Table 4 indicates the comparison between different methods and the proposed framework for gray-scale images in image dataset 1.

Table 4 shows that the accuracy rates exhibited different trends. Computational time could not be compared because it was not provided in the other methods. The best results were observed for the expanded DCT Markov + DWT Markov and expanded DCT Markov [10], which reduced to 100D by applying the SVM-RFE method (93.55% and 90.07%). The next best accuracy rate was observed for our proposed method (88.28%) with 50D. The results show that

Table 4: Comparison between proposed approaches and other methods.

Feature extraction methods	Dimensionality	TP (%)	TN (%)	Acc. (%)
Expanded DCT Markov [10]	100	89.92	90.21	90.07
DWT Markov [10]	100	87.58	85.39	86.50
Expanded DCT Markov + DWT Markov [10]	100	93.28	93.83	93.55
HHT + moments of characteristic functionswith wavelet decomposition [5]	110	80.25	80.03	80.15
Run length and edge statistics based model [13]	163	83.23	85.53	84.36
RLRN + kernel PCA (proposed)	**50**	**90.38**	**86.18**	**88.28**

Table 5: Comparison between the original RLRN [15] and the proposed method (RLRN + kernel PCA).

	RLRN [15]			RLRN + kernel PCA		
	TP (%)	TN (%)	Acc. (%)	TP (%)	TN (%)	Acc. (%)
R	56.30	83.70	70.9	85.71	90.91	88.31
G	51.80	83.20	68.50	87.22	90.26	88.74
B	57.20	83.70	71.30	85.71	90.91	88.31
Gray	60.40	81.70	71.80	86.47	90.91	88.69
Y	53.30	83.10	69.20	87.22	90.26	88.74
Cb	91.70	96.50	94.30	87.81	90.91	89.36
Cr	91.80	97.10	94.70	85.71	90.91	88.31

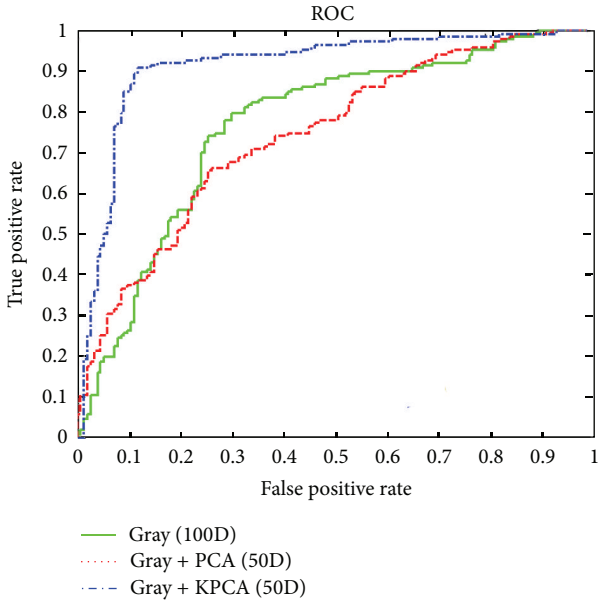


Figure 5: Comparison between gray-scale images in 100D, with PCA in 50D and kernel PCA in 50D.

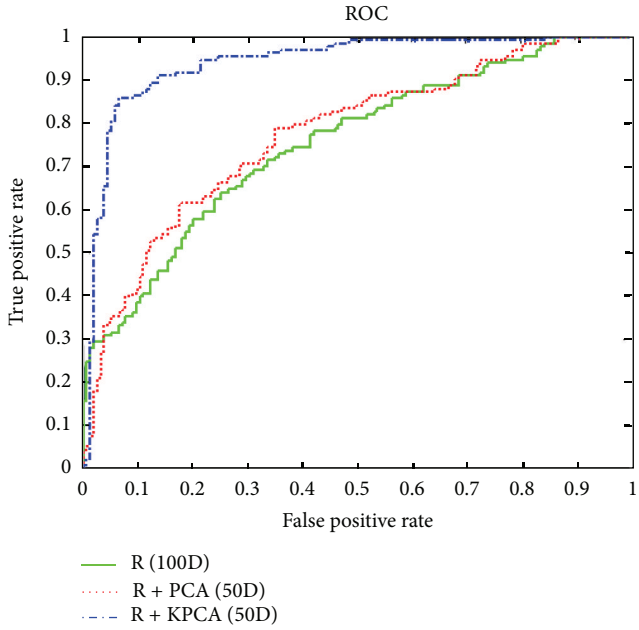


Figure 6: Comparison between gray-scale images in 100D, with PCA in 50D and kernel PCA in 50D.

our proposed method performed the best compared with similar methods [13] with less dimensions (163D versus 50D).

Table 5 shows another comparison between the original RLRN [15] in 60D and the proposed method (RLRN + kernel PCA) in 50D on image dataset 2 (CASIA). The Gray indicated in the table was obtained by converting the colored images in image dataset 2 to gray-scale images. As previously discussed, a huge increase in detection accuracies in the R, G, B, and Y channels, as well as the gray one, was observed.

4. Conclusion

The literature review presented in this work showed that several image splicing detection methods were proposed. Unfortunately, many of them are unable to handle the extracted high dimensional and redundant features well. In addition, processing these features is time consuming. Therefore, this paper focused on evaluating the effectiveness of dimension reduction methods on image splicing detection

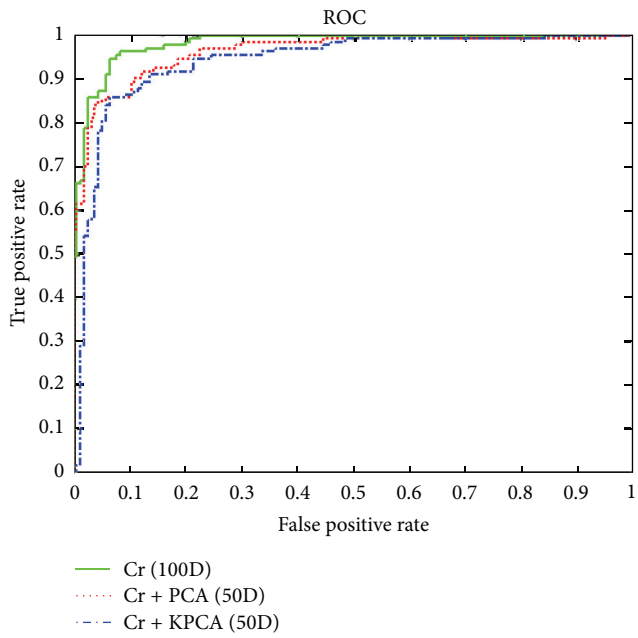


Figure 7: Comparison between gray-scale images in 100D, with PCA in 50D and kernel PCA in 50D.

methods to reduce dimensionality and computational time and remove redundant information from the features. Two dimension reduction methods (PCA: linear and kernel PCA: nonlinear) were selected and applied on an instance image splicing method (RLRN). A set of experiments were designed and tested to determine the effectiveness of these methods on RLRN. Each dimension reduction method was applied on RLRN to reduce dimensionality to 10D and 50D. The results showed that the R, G, B, and Y channels and gray-scale images performed best when merged with kernel PCA and in 50D, which verifies the nonlinear nature of the RLRN features. However, the results also demonstrate that this area of study requires further research. Other dimension reduction methods and their effects on other image splicing detection methods must be investigated, and the optimal dimension number for every image splicing detection approach must be identified.

Abbreviations

*n*D: *n* dimensional
RLRN: Run length run number
PCA: Principle component analysis
SVM: Support vector machine
HHT: Hilbert-Huang transform
DWT: Discrete wavelet transform
DCT: Discrete cosine transform
SVM-RFE: Support vector machine recursive feature elimination
DVMM: Digital video multimedia
CASIA: Chinese Academy of Sciences, Institute of Automation
TP: True positive
TN: True negative
Acc.: Accuracy
ROC: Receiver operating characteristics.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank both reviewers for their valuable comments which helped to improve the paper. This research is supported by University of Malaya Research Grant (UMRG) RG312-14AFR.

References

[1] A. Rocha, W. Scheirer, T. Boult, and S. Goldenstein, "Vision of the unseen: current trends and challenges in digital image and video forensics," *ACM Computing Surveys*, vol. 43, no. 4, article no. 26, 2011.

[2] W. Wang, J. Dong, and T. Tan, "A survey of passive image tampering detection," in *Digital Watermarking*, pp. 308–322, Springer, 2009.

[3] H. Farid, *Detecting Digital Forgeries Using Bispectral Analysis*, 1999.

[4] T.-T. Ng, S.-F. Chang, and Q. Sun, "Blind detection of photomontage using higher order statistics," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS '04)*, pp. V-688–V-691, IEEE, May 2004.

[5] D. Fu, Y. Q. Shi, and W. Su, "Detection of image splicing based on hilbert-huang transform and moments of characteristic functions with wavelet decomposition," in *Digital Watermarking*, pp. 177–187, Springer, New York, NY, USA, 2006.

[6] X. Li, T. Jing, and X. Li, "Image splicing detection based on moment features and Hilbert-Huang Transform," in *Proceedings of the IEEE International Conference on Information Theory and Information Security (ICITIS '10)*, pp. 1127–1130, December 2010.

[7] Y. Q. Shi, C. Chen, and W. Chen, "A natural image model approach to splicing detection," in *Proceedings of the 9th workshop on Multimedia & security*, pp. 51–62, ACM, September 2007.

[8] J. Zhang, Y. Zhao, and Y. Su, "A new approach merging Markov and DCT features for image splicing detection," in *Proceedings of the IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS '09)*, pp. 390–394, November 2009.

[9] Y. Q. Shi, C. Chen, and W. Chen, "A Markov process based approach to effective attacking JPEG steganography," in *Information Hiding*, Springer, 2007.

[10] Z. He, W. Lu, W. Sun, and J. Huang, "Digital image splicing detection based on Markov features in DCT and DWT domain," *Pattern Recognition*, vol. 45, no. 12, pp. 4292–4299, 2012.

[11] C. Chen and Y. Q. Shi, "JPEG image steganalysis utilizing both intrablock and interblock correlations," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS '08)*, pp. 3029–3032, Seattle, Wash, USA, May 2008.

[12] T.-T. Ng, S.-F. Chang, and Q. Sun, "A data set of authentic and spliced image blocks," ADVENT Technical Report 203-2004, Columbia University, New York City, NY, USA, 2004, http://www.ee.columbia.edu/ln/dvmm/.

[13] J. Dong, W. Wang, T. Tan, and Y. Q. Shi, "Run-length and edge statistics based approach for image splicing detection," in *Digital Watermarking*, pp. 76–87, Springer, 2009.

[14] Z. He, W. Lu, and W. Sun, "Improved run length based detection of digital image splicing," in *Digital Forensics and Watermarking*, pp. 349–360, Springer, 2012.

[15] X. Zhao, J. Li, S. Li, and S. Wang, "Detecting digital image splicing in chroma spaces," in *Digital Watermarking*, vol. 6526 of *Lecture Notes in Computer Science*, pp. 12–22, Springer, Berlin, Germany, 2011.

[16] "CASIA tampered image detection evaluation database," http://forensics.idealtest.org/.

[17] D. Zou, Y. Q. Shi, W. Su, and G. Xuan, "Steganalysis based on Markov model of thresholded prediction-error image," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '06)*, pp. 1365–1368, IEEE, July 2006.

[18] M. M. Galloway, "Texture analysis using gray level run lengths," *Computer Graphics and Image Processing*, vol. 4, no. 2, pp. 172–179, 1975.

[19] X. Tang, "Texture information in run-length matrices," *IEEE Transactions on Image Processing*, vol. 7, no. 11, pp. 1602–1609, 1998.

[20] K. Anusudha, S. A. Koshie, S. S. Ganesh, and K. Mohanaprasad, "Image splicing detection involving moment-based feature extraction and classification using artificial neural networks," *ACEEE International Journal on Signal & Image Processing*, vol. 1, no. 3, p. 9, 2010.

[21] A. Ghodsi, *Dimensionality Reduction a Short Tutorial*, Department of Statistics and Actuarial Science, University of Waterloo, Ontario, Canada, 2006.

[22] J. E. Jackson, *A User's Guide to Principal Components*, John Wiley & Sons, New York, NY, USA, 1991.

[23] I. T. Jolliffe, *Principal Component Analysis*, vol. 487, Springer, New York, NY, USA, 1986.

[24] B. Schölkopf, A. Smola, and K.-R. Müller, "Kernel principal component analysis," in *Artificial Neural Networks—ICANN '97*, vol. 1327 of *Lecture Notes in Computer Science*, pp. 583–588, Springer, Berlin, Germany, 1997.

[25] I. K. Fodor, "A survey of dimension reduction techniques," Tech. Rep. UCRL-ID-148494, Lawrence Livermore National Laboratory, 2002.

[26] L. J. E. O. Postma, H. J. van den Herik, and L. J. van der Maaten, "Dimensionality reduction: a comparative review," *Journal of Machine Learning Research*, vol. 10, pp. 66–71, 2009.

[27] C.-C. Chang and C.-J. Lin, "LIBSVM: a Library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, article 27, 2011.

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2014, Article ID 562787, 7 pages
http://dx.doi.org/10.1155/2014/562787

Research Article

Security Considerations and Recommendations in Computer-Based Testing

Saleh M. Al-Saleem[1,2] and Hanif Ullah[2]

[1] *Department of Information System, College of Computer and Information Sciences, King Saud University, P.O. Box 51178, Riyadh 11543, Saudi Arabia*
[2] *Department of Computerized Based Testing, National Center for Assessment in Higher Education, P.O. Box 68566, Riyadh 11537, Saudi Arabia*

Correspondence should be addressed to Hanif Ullah; h.ullah@qiyas.org

Received 29 April 2014; Revised 5 July 2014; Accepted 14 August 2014; Published 1 September 2014

Academic Editor: Adeel Javed

Many organizations and institutions around the globe are moving or planning to move their paper-and-pencil based testing to computer-based testing (CBT). However, this conversion will not be the best option for all kinds of exams and it will require significant resources. These resources may include the preparation of item banks, methods for test delivery, procedures for test administration, and last but not least test security. Security aspects may include but are not limited to the identification and authentication of examinee, the risks that are associated with cheating on the exam, and the procedures related to test delivery to the examinee. This paper will mainly investigate the security considerations associated with CBT and will provide some recommendations for the security of these kinds of tests. We will also propose a palm-based biometric authentication system incorporated with basic authentication system (username/password) in order to check the identity and authenticity of the examinee.

1. Introduction

During the last few decades and especially from 1990 onwards computer-based testing (CBT) has become one of the most conspicuous ways of organizing and delivering the tests. The reason behind this prominence of CBT over the paper-and-pencil based testing is its ease of administration, immediate display of results, improved item development, enhanced identification and authentication, and so forth [1]. However, most of the organizations and institutions are still relying on the examinations where the examinees have face to face exams in an identified place under an administered situation. This may help the organizations to check the authenticity of the examinee by checking his identity using ID card and also ensuring that no cheating is going on during the exam [2]. But due to the enormous advantages of computer-based testing over the traditional paper-and-pencil based testing, most of the organizations are now moving towards CBT.

Due to the hasty and considerable development in the provision of detached and internet-delivered computer-based testing, a number of issues in the administration standards, security, and control over the testing processes are raised. The most important one is that of authentication of examinee and examiner. Both the examinee and examiner should be authenticated in order to secure the computer-based testing. Different techniques have been proposed for the authentication purpose. A theoretical approach was proposed by [3] that incorporates the fingerprint biometrics with e-learning environments in order to restrain the unethical conduct during the exams. The proposed solution would enhance the current authentication scheme by adding the fingerprint biometrics. The technique is randomly exercised during the exam with a very short fingerprint scanning response time in order to get additional security. The approach could discourage the learners and examinees from having someone else taking the exam for them.

Similarly a multimodal biometric approach was proposed by Asha and Cellappan in [4] in order to authenticate the examinees in the test. The proposed technique was a combination of fingerprint and mouse dynamics. The basic aim of the approach was to reduce the cheating during the exam.

Some other techniques have been used to authenticate the examiner or invigilator. The authenticity of the examiner is also more important because the examiner has access to the examinee registration data as well as the examinee test data. Different techniques have been proposed for the authenticity of the examiner. Moreover an e-monitoring based scheme has also been proposed to monitor the examinee during the exam and to rectify or detect the problem of cheating during the exam.

The remainder of this paper is organized as follows. Section 2 describes the related work relevant to our paper. Section 3 investigates privacy and security issues/considerations among the available computer-based testing standards. Section 4 provides recommendations for secure computer-based testing environment and elaborates the proposed palm-based authentication technique. Section 5 offers conclusions.

2. Related Work

Deepthi et al. proposed a novel approach to enhance the security for online exams by introducing the idea of group cryptography with an e-monitoring scheme. Through the e-monitoring system the examinees can be monitored similar to the offline exams by webcams [5]. They also proposed different methods to detect and prevent ongoing exam cheating like the identities of the entities on the system that are verified by webcams and these reference photos taken during the verification process are stored for further authentication during the exam. Similarly the examinees monitoring data are recorded and stored during the exam. Also screen shots are taken during the exam so that the proctor can better determine the status of the examinee during the exam [5].

Castella-Roca et al. proposed a secure e-exam management system where all data and information should be in digital format. Also they proposed a cryptographic scheme that should be executed at every stage of exam in order to get the maximum security [2]. Their system is based on different cryptographic protocols offering high level of security during the entire exam. The authors identified different stages of exam, for example, setting up an exam, beginning, holding and submitting, grading, obtaining, and revising the exams. They also identified different security requirements like authenticity, privacy, correction, secrecy, receipt, copy detection, and so forth [2]. The authors also proposed that the exam should be in a supervised environment.

Oluwatosin and Samson examined some of the challenges of the existing computer-based testing systems and came up with a new system that could be deployed on either internet or intranet. The proposed scheme was designed by unified process using UML [6]. Security and result integrity features were integrated in the system. The examiner should answer some questions on the first login and should update his authentication details. All authentications are logged automatically by the system in order to determine any illegal access. Also an examinee could not have more than one active session running simultaneously.

Anusha et al. studied various authentication techniques like unimodel, multimodel, and data visualization and proposed a method called enhanced security using data visualization (ESDW) for online exams [7]. The method included the examinee authentication at the beginning of the exam and continuous monitoring through a webcam during the entire exam. The authentication process of the examinee was first carried out by preprocessing of the image through filtering, normalization, and segmentation. The extraction of feature was done based on the color, texture, and shape of the desired image.

El-Khatib and Korba examined the provisions and limitations for privacy and security by investigating some of the most popular e-learning standards. The capabilities of many existing privacy enhancing technologies including network privacy, policy-based privacy, and trust systems were reviewed and assessed [8]. The problems of privacy and security for distributed e-learning systems, where the learner can access the learning contents from anywhere, using any suitable device, that is, desktop computers, PDA, and so forth, were also investigated. Overall privacy requirements for e-learning systems based on "privacy principals" were highlighted.

Huszti and Pethő described a cryptographic scheme that retains the security requirements without the intervention of a trusted third party. Authenticity, anonymity, secrecy, robustness, and correctness are the main security features highlighted in this paper. These requirements were accomplished by applying cryptographic primitives [9].

Similarly Chang and Ansley [10] investigated and compared the properties of item exposure control methods in order to estimate the abilities of examinees in computer adaptive testing context. Different item pool sizes and different desired maximum exposure rates, effectiveness, and psychometric properties of the various control methods were evaluated in terms of test security, item overlap rate, and so forth; the objective of the study was to offer more information about the properties of the various exposure control strategies, to provide information about how the exposure control methods would be affected by using different sizes of item pools, and to provide guidelines for selecting a specific exposure control method.

Asha and Chellappan proposed a multimodel biometric technique in [4] for the authentication of e-learner in an e-learning environment. They classified the authentication system into three categories, that is, method based on human memory, method based on physical devices, and method based on biometrics. The proposed approach incorporated the fingerprint biometrics with the behavioral biometrics or mouse dynamics. These mouse dynamics were collected passively and verified throughout the session. The authors concluded that the proposed approach would enhance the authentication process.

Levy and Ramim presented a theoretical approach for biometric authentication of e-exams in [3]. The proposed approach used a fingerprint biometric solution in e-exams. During the exam, the learner's access is authenticated once at login for the whole duration. The authors proposed that the

current authentication system would be enhanced by using the fingerprint biometrics solution.

In [11–14] the authors proposed different palm-based authentication techniques. In [11] the authors proposed a scanner-based personal authentication system using the palmprint features. The proposed technique suits for many network based systems. Preprocessing, feature extraction, and modeling modules were used to generate the matching templates. Similarly in [12–14] the authors used wavelet-based, symbolic representation for palm-based authentication methods.

In paper [15–17] the authors proposed a genetic algorithm based palm recognition method for authentication process. The authors suggested that fingerprint and facial recognition systems are slow and required more expensive technical equipment, while palm recognition method did not require special equipment and could be used in systems where fast detection is required. Similarly a hand-based and palmprint and iris-based authentication system were proposed in [18–20]. Palmprint and iris-based authentication were used in order to overcome the problem of dictionary attacks.

3. Security Considerations

Many organizations and institutions around the world are moving towards computer-based testing from traditional paper-and-pencil based testing. However, this move will require substantial resources in terms of cost, management and administration, and security considerations. For a successful movement, it is recommended to prepare a plan and to highlight all of the above mentioned issues. In this paper we are going to highlight all the considerable issues related to the security of computer-based testing (CBT).

3.1. Required Resources for CBT. A number of resources are required for CBT administration. These resources include the cost associated with items such as computers, printers, and handheld devices although the costs related to traditional test or paper-and-pencil based tests are removed, but still the cost for CBT is much higher than the other. Moreover the cost associated with the online exam management systems and the licensing cost for different online/desktop item bank creation software's should also be considered in order to shift from paper-and-pencil based testing environment to CBT.

3.2. Item Bank Creation and Exposure Control. One of the main issues related to CBT is the creation and administration of item banks. CBT normally requires a huge number of item banks because of its frequent scheduling. Creation of item banks by subject-matter experts (SMEs) and their related security is one of the most important issues. The SMEs should be a trusted body and should not provide the detail about the item banks to a third party or to the examinee directly. The second main issue with item banks is its exposure control. The issue arises when some items appear more frequently within a short period of time or within a specific geographical area.

Some of the test-takers or examinees may have access to these frequent appearing test items before the test schedule and these test items could be compromised. This may result a significant concern for high stake test makers [10].

Several algorithms have been proposed in order to control item exposure rates. The algorithm in [21] employs an exposure control parameter for each item in the bank determined by a series of repetitive simulations. For items appearing rarely, the associated exposure control parameters could be as high as 1.0, meaning that once these items are selected, they are almost presented.

Similarly the algorithm proposed in [22, 23] for unconditional multinomial procedure was derived from the Stocking and Lewis approach. Rather than using the procedure of Stocking and Lewis, they used the idea of multinomial model for the selection of next item that has to be administered. They also presented an algorithm by using conditional multinomial procedure to control the item exposure to the examinees at similar level of proficiency [23].

3.3. Identification and Authentication of Examinee. Test security is one of the most important aspects of any exam, whether it is administered as paper-and-pencil based or CBT. The primary element of security in CBT is the procedure used to identify and authenticate the examinee. Different people used different ways to identify the examinee.

A range of levels of authentication could be used for the examinee [24]. The authentication could be exercised by username and password. Also the access could be limited to specific machines over the internet by allowing specific IP addresses. Both of these methods could also be merged in order to get high level of authenticity. Further enhancements could be made to the system by adding fingerprint and retinal eye-pattern recognition. Moreover high stake exams would require the presence of an administrator to check the identity of the examinees and to make sure that the exam is completed under good conditions.

A multimodal biometric approach was introduced in [4] for the authentication process of examinees. They used the combination of fingerprint and mouse dynamics. Mouse with a fingerprint scanner could be used in order to capture the fingerprint of the examinee along with the measurement of mouse dynamics. The examinee would be authenticated once by the server at login for the whole duration of the activity session. The proposed solution would enhance the authentication process and would reduce the chances of cheating during the exam.

Another biometric recognition system was applied to evaluate the basic knowledge in high school students in [25]. The assessment was carried out in order to answer the main problem, who is there? The authentication process was assessed by means of index fingerprint. During the enrollment process the proposed scheme saved the student fingerprint and indexed it in the features database. Student personnel ID was assigned in the features database in order to link the students personnel information with the fingerprint image.

A theoretical fingerprint biometric authentication system was presented in [3] for the delivery of e-learning courses. In their approach they incorporated the existing fingerprint

biometric authentication technologies with e-learning environments to restrain unethical behavior during exam taking. The approach suggested practical solutions to incorporate a random fingerprint biometrics user authentication during exams. The solution will enhance the current authentication system by adding the fingerprint biometric solution.

3.4. Invigilator/Proctor Authentication. Besides the examinee authentication, invigilator or proctor authentication is also an important aspect in online or computer-based tests, because the proctor has access to many aspects of the exam, including the examinee registration data, test data, or examinee test. In both of these cases either the examinee data or the examinee test data could be compromised. In order to rectify this problem, an invigilator code should be generated along with the examinee code, so that the invigilator would never have access to the examinee code and, therefore, have no access to the examinee test. Also the invigilator could be authenticated by providing secure connection to the invigilator computer during username/password authentication. A role-based access control could also be used in order to keep the access secure [26].

3.5. Cheating during Exam. Another important security issue in the computer-based testing is cheating during the ongoing exam. The examinee can cheat either by communicating with their other colleagues or by browsing over the internet. To overcome this problem, a continuous monitoring system should be implemented in order to monitor the examinee throughout the exam. Different people have used different techniques for the same problem. In [7] the authors proposed the use of webcams to monitor the examinees during the exam. Also this monitoring data should be stored on a monitor server for future use. The authors continuously monitor and save the recording of the entire exam in order to rectify the problem of cheating. Also screen shots of the examinee computer were also saved in parallel with video in order to better determine what actually the examinee was doing with his computer.

Another way to rectify this problem is to arrange the test in a web lock environment, where the test is displayed in a full screen option and the browser and other applications are locked during the exam. The examinee can only minimize or exit the full screen once the test is submitted for grading. Also print, print screen and capturing options, copy and paste, right-click menu options, browser menu and toolbar options, function keys, and so forth are all disabled during the exam.

3.6. E-Monitoring System. E-Monitoring system is a system where the examinee or the student is monitored by a webcam during the entire exam. This system is one of the best solutions to rectify the problem of cheating during the exam. The system could be implemented and centralized, where the examinees of different centers are monitored through a centralized system from a particular place, or decentralized where each and every test center has its own monitoring system inside the test center and the examinees are monitored from there. This video streaming of the entire exam is stored on different hard drives for future use.

4. Proposed Authentication Scheme

It is necessary to combine many biometric traits to secure computer-based tests. Many models and schemes have been discussed in the previous section for the same purpose. In this section we are going to propose a new authentication scheme for computer-based tests. Figure 1 summarizes the entire biometric authentication scheme. The proposed scheme incorporates the traditional username/password technique to the palm-based biometric authentication technique in order to get highest level of security during the exams. The examinee will first enter his/her login detail. These credentials will be evaluated and if it is correct, the examinee will be asked to put his palm on the device for the second phase of the authentication process; otherwise, a message will be displayed with incorrect username/password. After the palm detection process, some of the features of palm will be extracted in order to check it against the features in the database.

Once the features are extracted, they are checked for the verification of examinee. If the verification is successful, the examinee will be permitted to the exam; otherwise, the examinee will be instructed that he/she is not an authorized examinee. Also in order to get the optimum security during the exam, the examinee will be continuously monitored by a webcam to avoid any inconvenience during the exam. The webcam monitoring will detect the threat of cheating during the exam. Also the authentication process can be enhanced by the monitoring scheme. The username and password are basically an exam specific generated codes or credentials that are produced during the test creation process. For each and every examinee a separate exam ID or username and password are generated. When the examinee arrives at the specified center, these credentials are provided to the examinees. Also the basic identification process (checking of ID cards or other identification documents) is carried out at the reception of the exam center. Once all of these requirements are fulfilled, then the examinee is asked to put his palm on the device for the second phase of the authentication process.

For the entire process secure socket layer (SSL) will be used because the secure tunnels that can be created with SSL offer a means to exchange the authentication information or credentials without it being easily intercepted. The proposed method will handle the basic authentication problems like someone else is trying to attempt the exam instead of actual examinee.

The proposed technique is different from all other techniques because of its emergence. Palm-based authentication is latest trend instead of fingerprint and other authentication methods. Also the proposed technique will be implemented in our center in the near future and we have already the devices for palm-based authentication system. Moreover the retinal authentication is an expensive and difficult solution because the devices are costly and not easily available in the market. Also the entire authentication and verification

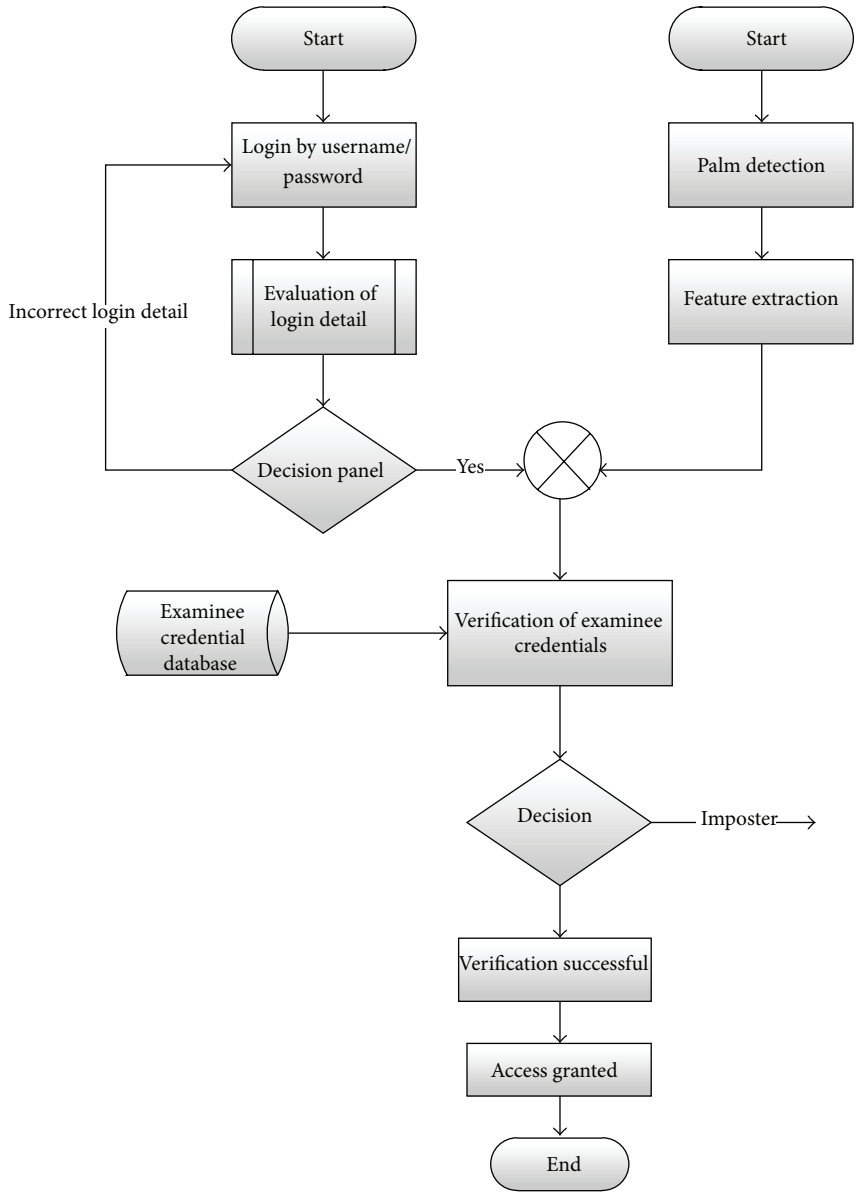


Figure 1: Proposed biometric authentication scheme.

process for the retinal system is much difficult than palm-based authentication and verification method.

5. Results and Discussion

The results of the proposed palm-based authentication scheme are described in this section. The block diagram of the enrollment and verification stage is also given in Figure 2. In the preprocessing stage of the enrollment and verification stage of the palm, region of interest (ROI) of the image is obtained. The images are taken from the IIT Delhi palmprint database and are given in Figure 3. After the preprocessing stage, the desired features of the given images are extracted and are stored in a database in order to XOR (XOR or Exclusive-OR is the process of combining the bits extracted from region of interest of the palm image and the bits obtained from the username and password in order to make the process secure and strong) with the username and password of the selected examinee. All these credentials are then stored in a database for the verification process. The same process is repeated for the verification process. Once the features are extracted, they are verified with the stored features in the database; if both the features are matched, the verification is successful; otherwise the verification process is denied.

The proposed approach is tested with different sets of input images and there results are stored in the template database for further processing. The detailed results and further discussion of the proposed scheme will be included

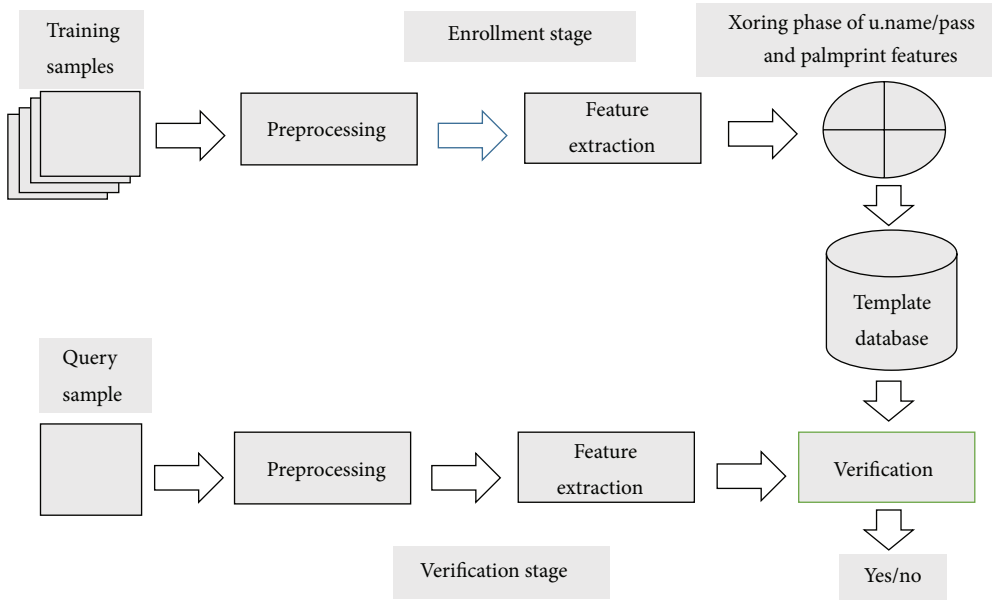


Figure 2: Enrollment and verification stage of the proposed system.

Figure 3: Images from the IIT Delhi database.

in the future work and will be a part of the next paper on the same topic.

6. Conclusion and Future Work

In this paper we have explored many security considerations in computer-based and online-based testing. Also we have explored many authentication schemes for the authentication of examinees. Moreover we have proposed a new authentication scheme that incorporates the username/password with the palm-based biometrics. The proposed scheme will enhance the authentication process up to an optimum level during the computer-based or online-based testing. The video capturing technique will also be merged with the existing system, in order to make it more secure in the future. Also CATSIM simulation tool will be used to simulate the entire authentication and verification process.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] M. Thurlow, S. S. Lazarus, D. Albus, and J. Hodgson, *Computer-Based Testing: Practices and Considerations*, University of Minnesota, National Center on Educational Outcomes, Minneapolis, Minn, USA, 2010.

[2] J. Castella-Roca, J. Herrera-Joancomarti, and A. Dorca-Josa, "A secure e-exam management system," in *Proceeding of the 1st International Conference on Availability, Reliability and Security (ARES '06)*, April 2006.

[3] Y. Levy and M. M. Ramim, "A theoretical approach for biometrics authentication of e-exams," in *Proceedings of the Chais Conference on Instructional Technologies Research*, pp. 93–101, The Open University of Israel, Raanana, Israel, 2007.

[4] S. Asha and C. Chellappan, "Authentication of e-learners using multimodal bometric technology," in *Proceedings of the IEEE-International Symposium on Biometrics and Security Technologies (ISBAST '08)*, April 2008.

[5] C. H. Deepthi, T. P. Shekhar, and N. Chiranji, "A novel approach to enhance security for online exams," *International Journal on Computer Science and Technology*, vol. 2, no. 3, pp. 85–90, 2011.

[6] O. T. Oluwatosin and D. D. Samson, "Computer-based test security and result integrity," *International Journal of Computer and Information Technology*, vol. 2, no. 2, pp. 324–329, 2013.

[7] N. S. Anusha, T. S. Soujanya, and D. S. Vasavi, "Study on techniques for providing enhanced security during online exams," *International Journal of Engineering Inventions*, vol. 1, no. 1, pp. 32–37, 2012.

[8] K. El-Khatib and L. Korba, "Privacy and security in E-learning," *International Journal of Distance Education*, vol. 1, no. 04, pp. 1–16, 2003.

[9] A. Huszti and A. Pethő, "A secure electronic exam system," *Journal of Publications Mathematician Debrecen*, vol. 77, no. 3-4, pp. 299–312, 2010.

[10] S. Chang and T. N. Ansley, "A comparative study of item exposure control methods in computerized adaptive testing," *Journal of Educational Measurement*, vol. 40, no. 1, pp. 71–103, 2003.

[11] C. Han, H. Cheng, C. Lin, and K. Fan, "Personal authentication using palm-print features," *Pattern Recognition*, vol. 36, no. 2, pp. 371–381, 2003.

[12] N. Swathi, S. Satish, V. S. Satyanarayana et al., "New palm print authentication system by using wavelet based method," *Signal & Image Processing*, vol. 2, no. 1, pp. 191–203, 2011.

[13] J. Chen, Y. Moon, M. Wong, and G. Su, "Palmprint authentication using a symbolic representation of images," *Image and Vision Computing*, vol. 28, no. 3, pp. 343–351, 2010.

[14] K. B. Ray and R. Misra, "Extracting region of interest for palm print authentication," *International Journal of Advanced Studies in Computer Science and Engineering*, vol. 2, no. 6, 2013.

[15] A. Cenys, D. Gibavicius, N. Goranin, and L. Marozas, "Genetic algorithm based palm recognition method for biometric authentication systems," *Elektronika ir Elektrotechnika*, vol. 19, no. 2, pp. 69–74, 2013.

[16] M. Hanmandlu, "A comprehensive study of palmprint based authentication," *International Journal of Computer Applications*, vol. 37, no. 2, pp. 17–23, 2012.

[17] S. R. Hiltz and M. Turoff, "Education goes digital: the evolution of online learning and the revolution in higher education," *Communications of the ACM*, vol. 48, no. 10, pp. 59–64, 2005.

[18] G. Amayeh, G. Bebis, A. Erol, and M. Nicolescu, *A New Approach to Hand-Based Authentication*, Computer Vision Laboratory, University of Nevada, Reno, Nev, USA.

[19] P. T. Selvi and N. Radha, "Palmprint and iris based authentication and secure key exchange against dictionary attacks," *International Journal of Computer Applications*, vol. 11, no. 11, pp. 7–12, 2011.

[20] V. K. Jain, "A technique to ROI of palmprint for palmline matching," *International Journal of Engineering Research and Applications*, vol. 2, no. 6, pp. 1007–1009, 2012.

[21] J. B. Sympson and R. D. Hetter, *Item Exposure Control in CAT-ASVAB*, Navey Personnel Research and Development Center, San Diego, Calif, USA, 1995.

[22] M. L. Stocking and C. Lewis, "A new method of controlling item exposure in computerized adaptive testing," Research Report 95-25, Educational Testing Service, Princeton, NJ, USA, 1995.

[23] M. L. Stocking and C. Lewis, "A new method of controlling item exposure in computerized adaptive testing," ETS Research Report RR-95-25, Educational Testing Service, Princeton, NJ, USA, 1995.

[24] D. Bartram, "Testing on the internet: issues, challenges and opportunities in the field of occupational assessment," in *Computer-Based Testing and the Internet: Issues and Advances*, John Wiley & Sons, 2006.

[25] J. A. Hernandez, "Biometrics in online assessments: a study case in high school students," in *Proceedings of the 18th International Conference on Electronics, Communications and Computers (CONIELECOMP '08)*, 2008.

[26] N. H. Lin, L. Korba, G. Yee, T. K. Shih, and H. W. Lin, "Security and privacy technologies for distance education applications," in *Proceedings of the18th International Conference on Advanced Information Networking and Applications (AINA '04)*, pp. 580–585, March 2004.

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2014, Article ID 753659, 13 pages
http://dx.doi.org/10.1155/2014/753659

Research Article

Network Anomaly Detection System with Optimized DS Evidence Theory

Yuan Liu,[1] Xiaofeng Wang,[2] and Kaiyu Liu[1]

[1] *School of Digital Media, Jiangnan University, Wuxi, Jiangsu 214122, China*
[2] *School of Internet of Things Engineering, Jiangnan University, Wuxi, Jiangsu 214122, China*

Correspondence should be addressed to Yuan Liu; lyuan1800@sina.com

Received 24 April 2014; Revised 28 July 2014; Accepted 12 August 2014; Published 31 August 2014

Academic Editor: Iftikhar Ahmad

Network anomaly detection has been focused on by more people with the fast development of computer network. Some researchers utilized fusion method and DS evidence theory to do network anomaly detection but with low performance, and they did not consider features of network—complicated and varied. To achieve high detection rate, we present a novel network anomaly detection system with optimized Dempster-Shafer evidence theory (ODS) and regression basic probability assignment (RBPA) function. In this model, we add weights for each senor to optimize DS evidence theory according to its previous predict accuracy. And RBPA employs sensor's regression ability to address complex network. By four kinds of experiments, we find that our novel network anomaly detection model has a better detection rate, and RBPA as well as ODS optimization methods can improve system performance significantly.

1. Introduction

With the development of computer network technology and the increasing of the networks scale, computer networks are under the threat of attack from hackers and other technologies, so the security status of the computer networks is becoming the focus of people's attention. Intrusion detection technology, protecting the network security behind the firewall, is becoming the research focus in the recent network security field. As the emphasis and difficulty of the network intrusion detection technology [1], network anomaly detection technology has the deficiency of the low detection rate, high false positive rate, and high false negative rate at present. So in this domain, many researchers proposed lots of useful algorithms [2–8], but these methods are so simple and single that they cannot be fully adapted to complicated and changeable network. Thus, a novel network anomaly detection mechanism is required to solve the above troubling problems.

Recently, some researches cope with network anomaly detection utilized by Dempster-Shafer (DS) evidence theory [9, 10] proposed by Dempster in 1976 and then improved by his student Shafer, which has been widely used in many fields of data fusion, such as expert advisory system, forecasting, image processing, artificial intelligence, and identifying classification. Intrusion detection is a problem of multi-classification essentially, which divides network data into normal data and various types of attacking data. Since simple detection algorithms always suffer from limitations such as low detection rate and high false alarm rate, many researchers apply DS evidence theory into intrusion detection systems. For example, some researchers divide the characteristics of network data into the basic feature set, the content feature set and the traffic feature set. Then they utilize detection algorithm to detect these three feature sets and fuse data through DS evident theory to get the final results. Though the IDS theory based on DS evidence theory has a good detection rate, most of these studies based on the classic DS evidence theory should assume that the intercepted data is independent of each other without confliction. However, conflicts between network data are inevitable, so those researches will lead to unreasonable fusion result, high false alarm rate and miss alarm rate.

To better solve this serious issue, we present a novel network anomaly detection mechanism based on optimized DS evidence theory (ODS) which also can achieve better

reasonable result in network conflict data, unlike conventional DS evidence theory. In this mechanism, we employ ODS to merge 3 classifiers, support vector machine classifier (SVM) [11, 12], biased minimax probability machine classifier (BMPM) [13], and back propagation network classifier (BP) [14]. Unlike the original fusion rule using the classification feature of those classifiers, the new one utilizes its regression feature, because regression feature can better reflect real-time network environment. Note that, since network environment is almost complicated and varied, in DS evidence theory, each sensor cannot be equally computed. So we assign different weights for each sensor, respectively, according to its previous prediction accuracy. In addition, based on different distance sizes between network connections, we present a new construction method for basic probability assignment function (BPA) based on regression ability, regression BPA (RBPA), unlike simple BPA (SBPA). Finally, through comparison of 4 fusion methods and 3 single methods, the experimental results with KDD99 [15] show that the ODS algorithm can overcome the conflict problem among the evidences, and the proposed module can improve the detection performance of the anomaly detection system.

The reminder of this paper is organized as follows. In Section 2, we will introduce class DS evidence theory, analyze its limitation, and present an optimized DS evidence theory with weights for each sensor. Then we propose a novel network anomaly detection model with ODS in Section 3. In Section 4, we introduce 4 key issues: how to combine ODS evidence theory with network anomaly detection, how to construct and decide BPA value in ODS evidence theory, how to decide weight w_i in fusion rules of ODS evidence theory, and how to train 6 classifiers. With data set KDD99, we evaluate our novel network anomaly detection model by 4 kinds of experiments in Section 5. Next, Section 6 introduces related work. Finally, we conclude our main work and propose future work in Section 7.

2. Optimized DS Evidence Theory

2.1. Class DS Evidence Theory. DS evidence theory [9, 10] is considered as a general extension of the traditional classical probabilistic inference theory in the finite field. Unlike the conventional Bayes inference method, DS evidence theory without a priori probability still can be used to deal with uncertainty and imprecision information. So we can see that DS evidence theory has greater flexibility.

DS evidence theory is considered as theory built on a nonempty finite field Θ called the recognition framework, which includes a limited number of independent system state $\{A_1, A_2, \ldots, A_n\}$. An element in $P(\Theta)$ as a power set of system state Θ is called a system state hypothesis H_i. Through the observation results $E_1, E_2, \ldots, E_m$ for system state by each sensor, DS evidence theory can merge these results and infer the former state of system. Here it mainly involves the following concepts.

Definition 1. Basic probability assignment function (BPA) is defined as a map from a power set of Θ to $[0, 1]$ interval. It is represented as $m : P(\Theta) \rightarrow [0, 1]$, $m(\Phi) = 0$, $\sum_{A \in P(\Theta)} m(A) = 1$, where $m(A)$ is called confidence value which means that current sensor decides hypothesis A the degree of confidence according to the observation results.

Definition 2. Belief function is defined as

$$\text{Belief}_i(A) = \sum_{E_k \subseteq A} m_i(E_k). \tag{1}$$

This function represents the degree of confidence for hypothesis A. And the function result is composed of basic confidence values of observation results E_k which supports hypothesis A.

Definition 3. Plausibility function is defined as

$$\text{Plausibility}_i(A) = 1 - \sum_{E_k \cap A = \Phi} m_i(E_k). \tag{2}$$

This function represents the degree of plausibility for hypothesis A. And the function result is composed of basic confidence values of observation results E_k which supports hypothesis A.

Definition 4. DS fusion rules, for any hypothesis A, defining m_i and m_j as the basic probability assignment function (BPA) of two evidences, respectively, state that one obtains basic belief assignment function of the combination evidence from two evidences above as follows:

$$\left(m_i \oplus m_j\right)(A) = \frac{\sum_{E_k \cap E_{k'} = A} m_i(E_k) m_j(E_{k'})}{1 - \sum_{E_k \cap E_{k'} = \emptyset} m_i(E_k) m_j(E_{k'})}. \tag{3}$$

Likewise, one can achieve DS general synthesis rules for the combination evidence from n evidences as follows:

$$m_{1\cdots n}(A) = \frac{\sum_{\cap_i E_i = A} m_1(E_1) m_2(E_2) \cdots m_n(E_n)}{\sum_{\cap_i E_i \neq A} m_1(E_1) m_2(E_2) \cdots m_n(E_n)}. \tag{4}$$

2.2. Drawback of Class DS Evidence Theory. The advantage of DS evidence theory mainly focuses on several parts as follows: it can satisfy axiom system that is weaker than the probability, distinguish unknown and uncertainty situation, and continuously shrink the hypothesis set in the light of the accumulation of evidences.

The disadvantage of DS evidence theory is that, when dealing with the issue with confidence degree tending to 0, the result computed by DS evidence theory will conflict with expectation result. That is to say, when confidence degree is too small or 0, the results achieved are very different. In the same way, BPA is required to give so many results that calculation is also more complicated. If the hypothesis set is too large, the calculation complexity of evidence theory will increase exponentially.

2.3. Optimized DS Evidence Theory. From formula (4), we can see that the confidence degree of each sensor is the same. That is, each sensor has the same accuracy. Obviously, it does not fit the facts, for example, with doctor for treatment.

A doctor considers that this patient may be suffering from X disease with 99%, or Y disease with 1%. However, B doctor considers that this patient may be suffering from Y disease with 1%, or Z disease with 99%. Then we can achieve this patient suffering from Y disease according to formula (4) merging these different two pieces evidence. But this result is fully fault and does not match reality. Therefore, the synthesis rule from formula (4) only applies to the case with the same precision in all sensors.

To solve this serious issue, we present a novel method to change with conventional DS evidence theory. That is, we combine weights with DS evidence theory. The detailed implementation is that we add weight value to each sensor according to its previous predict accuracy. So we define m_i as the basic confidence values obtained by sensor S_i, and w_i as the previous predict accuracy of sensor S_i; similarly, m_j as the basic confidence values obtained by sensor S_j, and w_j as the previous predict accuracy of sensor S_j. The DS evidence theory combination with weights as follows:

$$\left(m_i \oplus m_j\right)(A) = \frac{\sum_{E_k \cap E_{k'} = A} \left[w_i m_i\left(E_k\right) \cdot w_j m_j\left(E_{k'}\right)\right]}{1 - \sum_{E_k \cap E_{k'} = \emptyset} \left[w_i m_i\left(E_k\right) \cdot w_j m_j\left(E_{k'}\right)\right]}. \quad (5)$$

3. ODS Network Anomaly Detection Model Design

In this paper, we present a novel network anomaly detection module based on optimized DS evidence theory merging with several kinds of classifiers. In this module, we utilize BMPM, SVM, and BP network as classifiers. Unlike the original fusion rule using the classification feature of those classifiers, the new one utilizes its regression feature, because regression feature can better reflect real-time network environment. Then we consider the merged result as one parameter used to construct BPA of DS evidence theory. And then we will introduce this novel network anomaly detection model in detail, which is depicted in Figure 1.

As shown in Figure 1, this module mainly consists of five modules: network connection record module, feature extraction module, data preprocessing module, early detection module, and ODS fusion module, respectively.

Network connection record module utilizes some network sniffer tools, for example, Sniffer, to collect network packets in the network where network anomaly detection host is, and then stores it. That is, this module is used to collect network data.

Feature extraction module is used to extract some features impacting network anomaly detection, which are in the network packets stored by network connection record module. And then we record corresponding features into a feature vector, in order to preprocessing module to use it. Similarly, this module gets rid of unconcerned features for network anomaly detection. Essentially, this module is used to complete feature reduction.

Data preprocessing module is used to cope with feature vector after feature extraction. In addition, some futures in one feature vector are discrete type, such as protocol type, service type and logo, and others are continuous type, such as connection time type, the length of data sent, and the length of data received. Since discrete data needs inputting into detection module in early phase, in order to following work, continuous features need to be discretized. At the same time, these feature vectors also need to be standardized and normalized, in order for these vectors to be normally operated in BP network. In essence, this module is used to do data normalization, discretization, and standardization.

Early detection module is employed to detect the feature vectors that have processed by data preprocessing module and gives the corresponding detection results for DS fusion module later. It is composed of 3 sensors: SVM, BMPM, and BP. To fit with complicated network environment, we optimize sensors, that is, add weights into each sensor and construct 6 classifiers: BMPM_N, BMPM_A, SVM_N, SVM_A, BP_N, and BP_A (Section 4.3). Then we train these classifiers according to distance theory [16] (Section 4.4). Finally, we achieve several results when a network record is coming.

ODS fusion module will utilize ODS evidence theory to merge and analyze these detection results from early detection module. That is, according to regression ability of sensors, we fuse these results by (15) and give decision results, that is, whether the attack or not.

4. ODS Network Anomaly Detection Model Implementation

In ODS network anomaly detection model, we should solve several key issues: how to combine ODS evidence theory with network anomaly detection, how to construct and decide BPA value in ODS evidence theory, how to decide weight w_i in fusion rules of ODS evidence theory, and how to train 6 classifiers in detail, and so forth. Therefore, in this section, we will introduce the solutions for these serous issues mentioned above in detail.

4.1. Combining ODS Evidence Theory with Network Anomaly Detection. Since in ODS network anomaly detection model, system judges that whether current connection is unusual only according to network feature observed, we only define ODS evidence theory identification framework with two elements: normal status and abnormal status.

Therefore, according to DS evidence theory, we define ODS evidence theory identification framework as $\{N, A\}$, where N represents normal status and A represents abnormal status. We can see that status N and A are mutual exclusion, that is, $N \cap A = \emptyset$. Similarly, we can redefine BPA function as $m : P(\{N, A\}) \rightarrow [0, 1]$, $m(\emptyset) = 0$, $m(\{N, A\}) + m(N) + m(A) = 1$. In above formula, $m(N)$ represents the observation results of current feature by current sensor and considers that reliability of current status belongs to abnormal status. On the other hand, $m(\{N, A\})$ represents the observation results of current feature by current sensor and cannot decide reliability of current status belongs to normal or abnormal status. We will introduce detailed BPA function in next subsection.

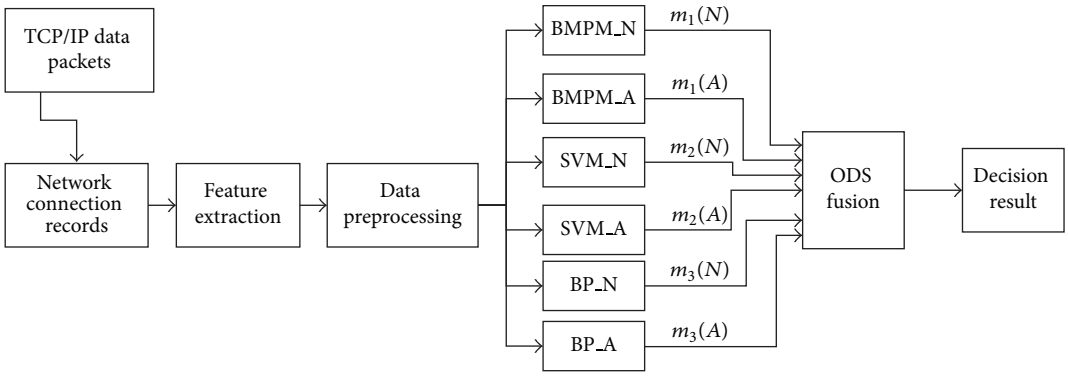


Figure 1: A novel network anomaly detection model with ODS evidence theory.

4.2. Regression BPA. In this subsection, we first give a hypothesis about network connection status for BPA value and then depict sensors' regression ability and how to compute RBPA value in detail.

4.2.1. Hypothesis and Discussion for Network Connection Status. Reference [16] said that the distance between abnormal network connection and normal network connection is larger than that between normal network connection and normal network connection. That is, for classifiers, the distance of different data is larger than that of same data. According to this rule, here we give a hypothesis: for a network connection to be seen (unknown), the prediction result will be $m(N)$ with N-classifier. That is to say, N-classifier considers all the network connection as normal network connection all the time, but only gives corresponding different support degrees according to difference of real network connection: high support degree for real normal network connection and low support degree for real abnormal connection. Through this hypothesis, we can see that for a real normal network connection, the prediction result $m(N)$ computed by N-classifier is larger than $m(A)$ computed by A-classifier and vice versa.

From Figure 1, three kinds of classifiers in early detection module, such as SVM, BMPM, and BP are also considered as three sensors. SVM_N and SVM_A are, respectively, represented support degree of normal and abnormal network connection from SVM sensor. Similarly, this rule is also suitable for BMPM and BP. So when we assign the same parameter for SVM_N and SVM_A, respectively, they can be considered as a whole sensor. And this whole sensor can give different support degrees to normal and abnormal network connection, respectively. Similarly, this is suitable for BMPM and BP sensor. Therefore in this novel model, the fusion part is considered as ODS evidence theory combining with three kinds of sensors, SVM, BMPM, and BP. If a normal network connection needs processing, no matter which senor (SVM, BMPM, and BP), the support degrees $m(N)$ and $m(A)$ for this network connection are, respectively, achieved by the chosen sensor. And these results satisfy the objective fact of this network connection. That is, in ODS evidence theory, if $m(N)$ is larger than $m(A)$, this network connection is considered as a normal one.

Based on this hypothesis above, if this novel model is required to automatically give current network connection support degree of normal and abnormal status by each sensor, and this result can also satisfy the objective fact of real network connection, we will utilize the features of sensors, such as study ability, regression ability, associative memory ability, and generalization ability. That is because sensors (e.g., BMPM, SVM, and BP) can achieve similar result after training, learning, and regression operations, which is almost equal to the actual result. So we see that sensors with their features can reflect the real network environment.

4.2.2. The Features of Sensors. In the above subsection, we said that some features of sensors will be selected to help with BPA function construction, so here we utilize regression ability, supervised learning ability of SVM, BMPM, and BP sensor. That can be explained in detail as follows: here we define one class of data as normal network connection data N and its corresponding training data NT, and define another class of data as abnormal network connection data A and its corresponding training data AT. Then these two kinds of data and their corresponding training data are used to train these classifiers depicted in Figure 1. As long as the two kinds of data distribution and training data have obvious difference, when a data record satisfies any kinds of data mentioned above, we can estimate the value of this data record (NT or AT corresponding with training data) utilizing the regression ability of these classifiers. In addition, the estimate values show obvious difference due to data record satisfying different kinds of data distribution.

4.2.3. BPA Based on Regression Ability. Since network connection status can be represented as normal or abnormal status by different sensors which give different support degrees for them, with this rule, we construct BPA function in ODS network anomaly detection model. When ODS evidence theory is combined with network anomaly detection, assuming that the current network connection is a normal one, corresponding BPA value can be different achieved by different sensors (various classifiers in fusion model). And the BPA values are corresponding with hypothesis N, A or $\{N, A\}$. Similarly, we also expect that the BPA value of normal

network connection assigned by hypothesis N is larger, but on the contrary, the BPA value assigned by hypothesis A (abnormal status) or hypothesis $\{N, A\}$ (unknown status) should be smaller.

After training N-classifier and A-classifier (training classifiers will be introduced in Section 4.4), we can compute BPA value in ODS evidence theory. Currently, SVM_N and SVM_A can be considered as a whole one, a SVM sensor. For a network connection record, the regression estimates value $m(N)$ computed by SVM_N and $m(A)$ computed by SVM_A. Due to associated ability of SVM classifier, if this record is a normal network connection record, $m(N)$ will be larger than $m(A)$, vice versa. In addition, this rule is also suitable for BMPM and BP. Therefore, these three sensors, SVM, BMPM, and BP, in system can assess current status for a coming network connection. That is, we can achieve support degrees $m(N)$ and $m(A)$ for normal status and abnormal status, respectively.

Noticeably, in this paper we present a novel method to deal with unknown network connection status as follows:

$$m(\{N, A\}) = \begin{cases} 1 - m(N) - m(A) & m(N) + m(A) < 1 \\ 0 & m(N) + m(A) \geq 1. \end{cases} \tag{6}$$

From formula (6), we can see that when $m(N) + m(A) < 1$, we define $m(\{N, A\})m(\{N, A\})$. Similarly, when $m(N) + m(A) \geq 1$, the value of $m(\{N, A\})$ is 0. $m(N)$ and $m(A)$ should be normalized at the same time. Thus, this is suitable for the requirement of ODS evidence theory $m(N) + m(A) + m(\{N, A\}) = 1$, and the computation of RBPA function is completed in ODS evidence theory.

4.3. Weights for Each Sensor. In the traditional network anomaly detection system, the performance is decided by an estimate parameter *F-Score*, which reflects that an intrusion detection system performance is good or bad. And the greater the *F-Score* indicates that the better performance of this system. So in this paper, we extend this important parameter *F-Score* and propose two new parameters, *F-Score-N* and *F-Score-A*. Since with ODS evidence theory, the w_i value depends on its previous accuracy in the process of sensor prediction, we utilize these new parameters to add weights for each senor. Then we will introduce these new parameters in detail.

Here, we define several parameters, respectively, as follows:

(i) TP: the number of abnormal connection detected by anomaly detection system (abnormal connection itself);

(ii) FN: the number of normal connection detected by anomaly detection system (abnormal connection itself);

(iii) FP: the number of abnormal connection detected by anomaly detection system (normal connection itself);

(iv) TN: the number of normal connection detected by anomaly detection system (normal connection itself);

(v) Precision: the proportion of true abnormal connections of abnormal connections detected by anomaly detection system;

(vi) Recall: the proportion of abnormal connections detected by anomaly detection system of true abnormal connections;

(vii) *F-Score*: a balance average parameter for Precision and Recall used to estimate a network anomaly detection system.

With these parameters mentioned above, the formulas are depicted as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad F\text{-}Score = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{7}$$

But these conventional formulas do not satisfy this novel network anomaly detection system presented in this paper, so we propose these new parameters as follows:

(i) Precision_A: The proportion of abnormal connections of abnormal connections detected by anomaly detection system (abnormal connection itself);

(ii) Recall_A: The proportion of abnormal connections detected by anomaly detection system of all true abnormal connections (abnormal connection itself);

(iii) Precision_N: The proportion of normal connections of abnormal connections detected by anomaly detection system (normal connection itself);

(iv) Recall_N: The proportion of normal connections detected by anomaly detection system of all true normal connections (normal connection itself);

(v) *F-Score-N*: The accuracy for normal connections estimated by sensors, that is, N-classifiers' weights;

(vi) *F-Score-A*: The accuracy for abnormal connections estimated by sensors, that is, A-classifiers' weights.

With these parameters mentioned above, the formulas are depicted as follows:

$$\text{Precision\_A} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{8}$$

$$\text{Precision\_N} = \frac{\text{TN}}{\text{TN} + \text{FN}}, \tag{9}$$

$$\text{Recall\_A} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{10}$$

$$\text{Recall\_N} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \tag{11}$$

$$F\text{-}Score\text{-}N = \frac{2 * \text{Precision\_N} * \text{Recall\_N}}{\text{Precision\_N} + \text{Recall\_N}}, \tag{12}$$

$$F\text{-}Score\text{-}A = \frac{2 * \text{Precision\_A} * \text{Recall\_A}}{\text{Precision\_A} + \text{Recall\_A}}. \tag{13}$$

4.4. Training Classifiers. In this subsection, we mainly introduce how to train 6 different classifiers including collecting training data set, preprocessing training data set, and training classifiers.

4.4.1. Collecting Training Data Set. First, collecting training data set is introduced in detail. Here we simulate network attack to attack this system and implement network connection record module at the same time. And then these records are stored. The detailed process is depicted as follows.

(1) We utilize corresponding attack software to simulate all kinds of attacks, for example, DOS attack which can be simulated by the combination DOS attack simulator with ping command, and others also can be achieved by this way.

(2) Before these abnormal connections attacking system, we should record IP address and attack types of attack hosts, respectively, IP address of destination hosts, and simultaneously implement network connection record module in network anomaly detection system to record network packets. Here the time window for attack time is 1 hour, and network connection record module also records all the network packets in this period.

(3) Based on these network packets recorded, we filter these network packets according to effective network attack packet standard corresponding with IP address and attack type of attack host recorded before attacking. Then these filtered network packets should be discretized, standardized, and normalized by feature extraction module and data preprocessing module, respectively, in Figure 1. Finally, we can obtain different kinds of network connection feature vectors, such as Normal, Dos, Probe, R2L, and U2R.

4.4.2. Preprocessing Training Data Set. According to the rule mentioned in [16], we utilize this different distance for same or different kinds of network connection to define BPA value in ODS evidence theory. Before train 6 classifiers, we must preprocess this training data set.

As shown in Figure 1, we can see that six basic classifiers in early detection module can be divided into two categories, namely, N-classifier and A-classifier. Before training N-classifier and A-classifier, we should preprocess training data set, and this will be introduced in detail. First, we analyze training N-classifier as follows.

(1) When a training data set includes normal connections, N-Train and abnormal connections, A-Train, it should be processed before training N-classifier. First, we compute clustering center, N-CORE of N-Train in training data set.

(2) Then we should compute the distance between this normal connection and N-CORE and define this distance value as a positive value, which also corresponds with this normal connection.

(3) Then we should compute the distance between this abnormal connection and N-CORE and define this distance value as a negative value, which also corresponds with this abnormal connection.

(4) Finally, the results from (2) and (3) are stored into N-Dist which corresponds with connection records. And this list N-Dist will be normalized from 0 to 1 and is considered as a training label to train N-classifier.

After processing above, distance corresponding with normal connection is larger than that with abnormal connection in training data set. If these distance values are used as supervised learning training label when training classifiers, these classifiers will learn this phenomenon through associated ability. So we can see that the regression value for normal network connection will be larger than that for abnormal network connection, when a normal network connection and an abnormal network connection need processing. Indeed, this rule mentioned above is also suitable for training A-classifier.

Then we will train 6 classifiers: N-classifiers utilizing training data set and corresponding N-Dist, A-classifiers utilizing training data set and corresponding A-Dist.

4.4.3. Training Classifiers. In this phase, we mainly train 6 classifiers in early detection module and compute some parameters in ODS fusion module. Here we divide network connection feature vectors (Normal, Dos, Probe, R2L, and U2R) into two parts according to attack type. Each part includes processed training data set and corresponding list (N-Dist, A-Dist) stored distance value.

(1) One part is used to train 6 classifiers, SVM_N, SVM_A, BMPM_N, BMPM_A, BP_N, and BP_A, and then these classifiers trained will be stored to do prediction in future.

(2) Another part is employed to predict these trained classifiers, and it should record these results including all kinds of attacks and normal connections, that is, these results for TP, TN, FP, and FN depicted in Section 4.3.

(3) Then we can get weights *F-Score-N* and *F-Score-A* of all sensors, SVM, BMPM, and BP computed by formulas (12) and (13). Finally, these values are stored into array *F-Score-N* and *F-Score-A,* respectively.

4.5. Execution Flow of ODS Network Anomaly Detection System. After training 6 classifiers introduced in Section 4.4, we can easily get weights for each sensor (in Section 4.3). As the same, the BPA values of each sensor, support degree $m(N)$, $m(A)$, and $m(\{N, A\})$, can be achieved easily introduced in Section 4.2. In this section, execution flow of ODS network anomaly detection system will be introduced in detail.

(1) First, we can get a network connection packet by network connection record module, and then a network connection feature vector can be obtained by feature extraction module and data preprocessing

module which process the network connection packet achieved one by one.

(2) Then this network connection feature vector will be processed to do regression estimate by 6 classifiers (3 sensors) in early detection module. So we can obtain support degree $m_1(N)$ and $m_1(A)$ for normal network connection status and abnormal network connection status after sensor SVM processing this network connection. Next computed by formula (6), support degree $m_1(\{N, A\})$ for unknown status is also achieved easily.

(3) In this way, we can be easy to obtain $m_2(N)$, $m_2(A)$, and $m_2(\{N, A\})$ corresponding with BMPM, and $m_3(N)$, $m_3(A)$, and $m_3(\{N, A\})$ corresponding with BP.

(4) Here we achieve an ODS evidence theory with weights for n-sensors inferred by formulas (4) and (5):

$$m_{1\cdots n}(A) = \frac{\sum_{\cap_i E_i = A} w_1 \cdot m_1(E_1) \cdot w_2 \cdot m_2 \cdot (E_2) \cdots \cdot w_n \cdot m_n(E_n)}{\sum_{\cap_i E_i \neq A} w_1 \cdot m_1(E_1) \cdot w_2 \cdot m_2 \cdot (E_2) \cdots \cdot w_n \cdot m_n(E_n)}. \quad (14)$$

In this novel model, we choose SVM, BMPM, and BP as sensors, so the parameter i is define from 1 to 3. By formula (14), we can obtain the support degree of this network connection, $m_{123}(N)$, $m_{123}(A)$ and $m_{123}(\{N, A\})$ through fusion 3 sensors. This process needs that we should bring support degree for N, A and $\{N, A\}$ computed by SVM, BMPM, and BP sensors, and weight vector *F-Score-N* and *F-Score-A* into formula (14).

(5) The final decision result by system is depicted in

$$\text{Decision}(x) = \begin{cases} \text{normal} & (\text{if } m_{123}(N) \\ & \quad = \max\{m_{123}(N), m_{123}(A), \\ & \qquad m_{123}(\{N, A\})\}) \\ \text{abnormal} & (\text{if } m_{123}(A) \\ & \quad = \max\{m_{123}(N), m_{123}(A), \\ & \qquad m_{123}(\{N, A\})\}) \\ \text{uncertain} & (\text{if } m_{123}(\{N, A\}) \\ & \quad = \max\{m_{123}(N), m_{123}(A), \\ & \qquad m_{123}(\{N, A\})\}). \end{cases} \quad (15)$$

The final decision result can be explained in detail: if $m_{123}(N)$ is larger than $m_{123}(A)$ and $m_{123}(\{N, A\})$, this system considers current network connection as a normal one; as the same, if $m_{123}(A)$ is larger than $m_{123}(N)$ and $m_{123}(\{N, A\})$, this system considers current network connection as an abnormal one; if $m_{123}(\{N, A\})$ is larger than $m_{123}(A)$ and $m_{123}(N)$, this system cannot judge current network connection as a normal or abnormal one.

5. Experiments and Analysis

In this section, we would verify the effectiveness of combining ODS evidence theory with SVM, BMPM, and BP sensors and prove that this novel ODS network anomaly detection model can get higher detection rate (DR) and lower false positive rate (FR) for not only traditional attacks but also new attacks.

5.1. Data Set

5.1.1. KDD99 Data Set. The KDD-Cup99 data set from UCI repository has been widely used as the benchmark data for network anomaly detection evaluation. It consists of several components depicted in Table 1. As in the case of the International Knowledge Discovery and Data Mining Tools Competition, only the "10% KDD" data is employed for the purposes of training. This contains 22 attack types and is essentially a more concise version of the "Whole KDD" data set. So in our experiments, we apply its 10% training data consisting of 494 021 connection records for training. Each connection record represents a sequence of packet transmission starting and ending at a time period and can be classified as normal traffic, or one of 22 different classes of attacks. All attacks fall into four main categories.

(i) Denial-of-service (Dos)—denial of the service that are accessed by legitimate users, for example, SYN flooding.

(ii) Remote-to-local (R2L)—unauthorized access from a remote machine, for example, password guessing.

(iii) User-to-root (U2R)—unauthorized access to gain local super-user (root) privileges, for example, buffer overflow attack.

(iv) Probing (Probe)—surveillance and probing for information gathering, for example, port scanning.

The test data set has not the same probability distribution as the training data set. There are 4 new U2R attack types in the test data set that are not presented in the training data set. These new attacks correspond to 92.90% (189/228) of the U2R class in the test data set. On the other hand, there are 7 new R2L attack types corresponding to 63% (10196/16189) of the R2L class in the data set. In addition there are only 104 (out of 1126) connection records presented in the training data set corresponding to the known R2L attacks presented simultaneously in the two data sets. However there are 4 new Dos attack types in the test data set corresponding to 2.85% (6555/229853) of the Dos class in the test data set and 2 new Probing attacks corresponding to 42.94% (1789/4166) of the Probing class in the test data set.

5.1.2. Data Set Preprocessing. Since a connection record in KDD 99 includes not only symbol feature but also continuous and discrete features, we must cope with these features before do experiments. Here Naïve algorithm in Rosetta software [17] is used to deal with continuous feature, and symbol feature can be discretized by general mapping method directly. Then in order to remove different features of various data and

Table 1: KDD data set.

Data set	Total	Normal (%)	DOS (%)	Probe (%)	U2R (%)	R2L (%)
10% KDD	494,020	19.79	79.2	0.8	0.01	0.2
Test KDD	311,029	19.58	73.9	1.3	0.02	5.2
Whole KDD	4,898,430	19.8	79.3	0.84	0.001	0.02

achieve general feature and same weights for discretized data, these data should be standardized, and these standardization formulas are introduced as follows:

$$\mathrm{I}_j^e = \frac{1}{n}\sum_{i=1}^{n}\mathrm{I}_{ij},$$
$$\mathrm{I}_j^\delta = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{I}_{ij} - \mathrm{I}_j^e\right)^2}, \quad (16)$$
$$\mathrm{I}_j' = \frac{\mathrm{I}_j - \mathrm{I}_j^e}{\mathrm{I}_j^\delta}.$$

5.2. Experimental Design. In order to prove that network anomaly detection system with ODS and RBPA has better performance, we design 4 kinds of experiments.

The first experiment is that we choose 3 single methods (SVM, BMPM, and BP) and 4 fusion methods (DS with SBPA, DS with RBPA, ODS with SBPA, and ODS with RBPA) to do detection in the same data set. In this data set, 4000 network connections of each connection type (Normal, Dos, Probe, and R2L) are selected and 249 network connections are chose from U2R type. These data chosen constructed a data set which is divided into 2 parts: training data set and test data set. This experiment is used to prove that the method we presented can detect various attacks and has higher DR and lower FR.

The second experiment is that we also choose these 7 network anomaly detection methods to do detection in R2L data set which has 4000 network connections. And the former 2000 network connections are normal connections and the later 2000 network connections are abnormal connections. This experiment is utilized to prove that the method with RBPA outperforms the method with ODS, and two optimization methods we presented can be used in network anomaly detection simultaneously with better performance.

The third experiment is that we also choose these 7 network anomaly detection methods to do detection in the same data set, like the first experiment. But we compare several parameters mentioned in Section 4.3, such as Precious, Recall, and *F-Score*. In addition, we utilize ROC curve which shows DR and FR of corresponding method, and AUC which represents the area under corresponding ROC curve to estimate the performance of network anomaly detection system.

The fourth experiment is that we choose 2 network anomaly detection methods (ODS with RBPA and DS with SBPA) to do detection. But here we choose 10% KDD99 data as training data set and test data set mentioned in Section 5.1.1 as test data set. This experiment is used to estimate the new model's detection ability for new attack type.

Table 2: The detection result of all the attacking types with SVM (FR = 1.91%).

Attack type	Attack number	Attack number detected	DR (%)
DOS	2000	1989	99.45
Probe	2000	1992	99.6
U2R	124	102	80.64
R2L	2000	1950	82.20
Total	6124	6033	98.51
	Variance		9.76

Table 3: The detection result of all the attacking types with BP (FR = 1.90%).

Attack type	Attack number	Attack number detected	DR (%)
DOS	2000	1971	98.55
Probe	2000	1986	99.30
U2R	124	95	76.61
R2L	2000	1974	98.70
Total	6124	6026	98.39
	Variance		9.90

5.2.1. Experiments with 4 Attack Types. From Tables 2, 3, 4, 5, 6, and 7, we can see that the false positive rate (FR) of single detection model, such as SVM, BMPM, and BP, is higher than that of fusion detection model. This reflects that fusion detection method can effectively reduce the FR in the anomaly detection system. From detection rate (DR) and attack number detected by anomaly detection methods, the DR of fusion detection method outperforms that of single method, and fusion method will bring lower FR. In addition, compared to fusion detection model, the variance of almost single detection model is larger, meaning that fusion detection model is not easy to shake, that is, relatively stable. Though the DR of some models for various attacks is high, its FR is still high, for example, BMPM model. Therefore, this novel model with ODS and regression BPA outperforms than others, and it has lower FR and better DR.

Here, we not only analyze the whole performance of this novel model, but also discuss ODS with weights and regression BPA performance. According to whether BPA and DS evidence theory redesigned, we can achieve 4 results for different combinations shown in Tables 5, 6, 7, and 8 respectively, including DS with Simple BPA (SBPA), ODS with SBPA, DS with regression BPA (RBPA) and ODS with RBPA.

Table 4: The detection result of all the attacking types with BMPM (FR = 1.32%).

Attack type	Attack number	Attack number detected	DR (%)
DOS	2000	1931	96.55
Probe	2000	1982	99.10
U2R	124	115	92.74
R2L	2000	1987	99.35
Total	6124	6015	98.22
	Variance		2.72

Table 5: The detection result of all the attacking types with classic DS evidence theory fusion of simple BPA method (FR = 0.62%).

Attack type	Attack number	Attack number detected	DR (%)
DOS	2000	1987	99.35
Probe	2000	1990	99.50
U2R	124	116	93.54
R2L	2000	1986	99.30
Total	6124	6079	99.26
	Variance		2.60

Table 6: The detection result of all the attacking types with ODS evidence theory fusion of Simple BPA method (FR = 0.59%).

Attack type	Attack number	Attack number detected	DR (%)
DOS	2000	1985	99.25
Probe	2000	1993	99.65
U2R	124	118	95.16
R2L	2000	1987	99.35
Total	6124	6083	99.33
	Variance		1.90

Table 7: The detection result of all the attacking types with classic DS evidence theory fusion of regression BPA method (FR = 0.32%).

Attack type	Attack number	Attack number detected	DR (%)
DOS	2000	1997	99.85
Probe	2000	1991	99.55
U2R	124	116	93.54
R2L	2000	1994	99.70
Total	6124	6098	99.57
	Variance		2.74

According to whether DS evidence theory redesigned (whether adding weights into DS), we can divide these models into two groups without considering BPA design: one is Tables 5 and 6, and another is Tables 7 and 8. In this way, we can compare the performance of ODS, DS with weights (*F-Score* value as weights) with that of DS. From these two groups, we can see that the FR of ODS is lower than that of DS, and the total DR of ODS is also lower than that of DS. Clearly, most of DR of various attack types with ODS outperform that with DS. Thus, ODS with weights is effective compared with DS.

Table 8: The detection result of all the attacking types with ODS evidence theory fusion of regression BPA method (false positive rate = 0.27%).

Attack type	Attack number	Attack number detected	DR (%)
DOS	2000	1997	99.85
Probe	2000	1997	99.85
U2R	124	119	95.97
R2L	2000	1992	99.60
Total	6124	6105	99.69
	Variance		1.69

Similarly, according to whether BPA redesigned (whether with sensors' regression ability), we can divide these models into two groups without considering DS design: one is Tables 5 and 7, and another is Tables 6 and 8. In this way, we can compare the performance of RBPA with that of SBPA. For FR and total DR, RBPA is better than SBPA significantly. So RBPA with sensors' regression ability is effective compared with SBPA.

5.2.2. Experiments with R2L Attack. In this subsection, we mainly focus on the novel model for single attack type according to formula (15). From Figures 2, 3, 4, and 5, we can see that they are achieved by different groups with redesigned or conventional BPA and DS. In these figures, corresponding with $m_{123}(N)$ in formula (15), parameter *MNP* represents the support degree of normal connection for current network connection after it is detected by SVM, BMPM, and BP sensors and merged by ODS. On the contrary, *MAP* corresponds with $m_{123}(A)$. By formula (15), if *MNP* is larger than *MAP*, current network connection is considered as a normal one, and vice versa. In this experiment, there are 4000 network connections in each figure, and the former 2000 network connections are normal connections and others are abnormal connections.

First, we analyze and compare Figures 4 and 5 in one group. In Figure 4, some normal connections of the former 2000 network connections overlap together for *MNP* and *MAP*. Significantly, some parts of *MAP* are above *MNP*, that is, this normal network connection is wrongly considered as an abnormal one, leading to a higher FR. On the contrary, the overlap in Figure 5 is less than that in Figure 4. In this way, Figure 3 outperforms Figure 2 with *MNP* and *MAP*. Without considering BPA design, ODS with weights is further effective.

Next, we analyze and compare Figures 2 and 4 in one group. In Figure 2, almost all the normal connections (the former 2000) overlap together for *MNP* and *MAP*. However, this overlap is further less in Figure 4. Clearly, this also occurs in the later 2000 connections, abnormal connections. In the same way, Figure 5 is better than Figure 3. In essence, this shows that RBPA method outperforms SBPA method, with lower FR and higher DR. This conclusion is consistent with the results from Tables 5 and 7 or Tables 5 and 8. Without considering DS design, RBPA with regression is further effective compared with SBPA.

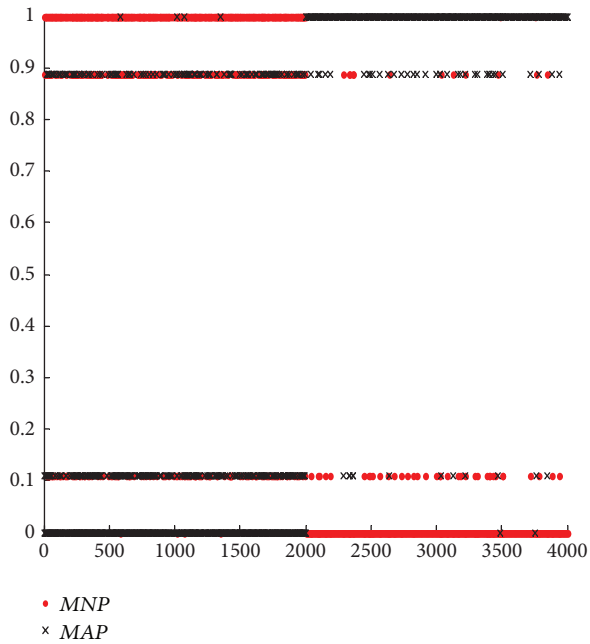


Figure 2: Results of classic DS evidence theory fusion of SBPA.

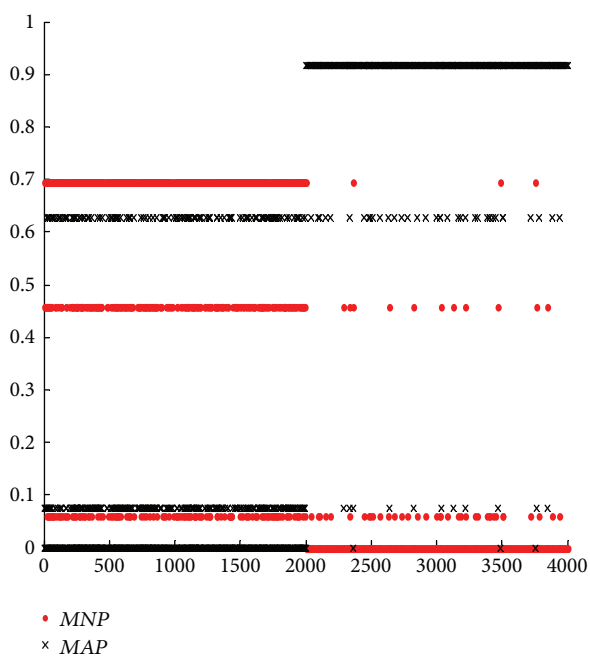


Figure 3: Results of ODS evidence theory fusion of SBPA.

Moreover, based on Figure 2, we compare Figure 3 with Figure 4. We can see that the results of *MNP* and *MAP* are distinguished easily and are suitable for real network better in Figure 4. But the opposite results are obtained in Figure 3, meaning that fuzzy and unseparated results. This leads a higher FR. Only verifying a condition, DS or BPA design, we can get Figure 3 with ODS optimization and Figure 4 with RBPA optimization. From Figures 3 and 4, we conclude

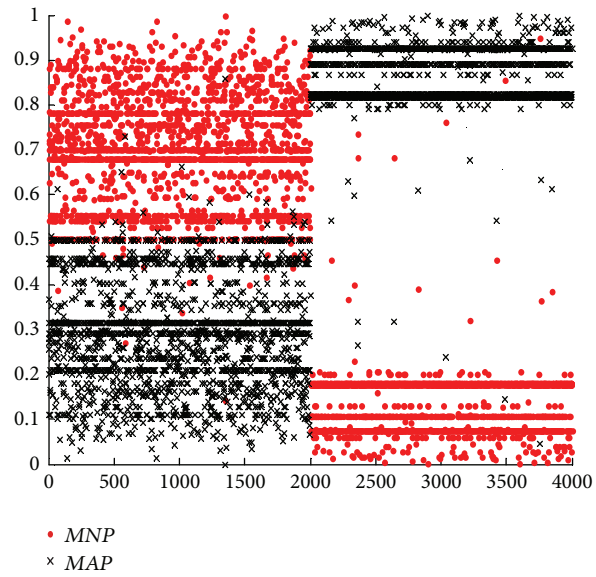


Figure 4: Results of classic DS evidence theory fusion of RBPA.

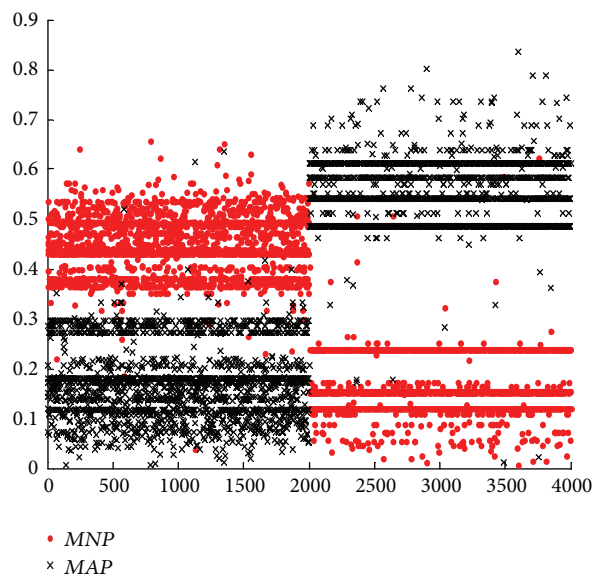


Figure 5: Results of ODS evidence theory fusion of RBPA.

that RBPA method is better than ODS method in network anomaly detection system.

Similarly, based on Figure 2, we compare Figures 3 and 4 with Figure 5. We can see that the result of Figure 5 is better than both Figures 3 and 4 significantly. Compared with Figure 2, Figure 5 is improved enormously. In a word, no matter which one system chooses, the performance of optimized network anomaly detection system will be improved clearly. Specially, these two optimization methods can be utilized by network anomaly detection system simultaneously, leading a better result than the one with either optimization method.

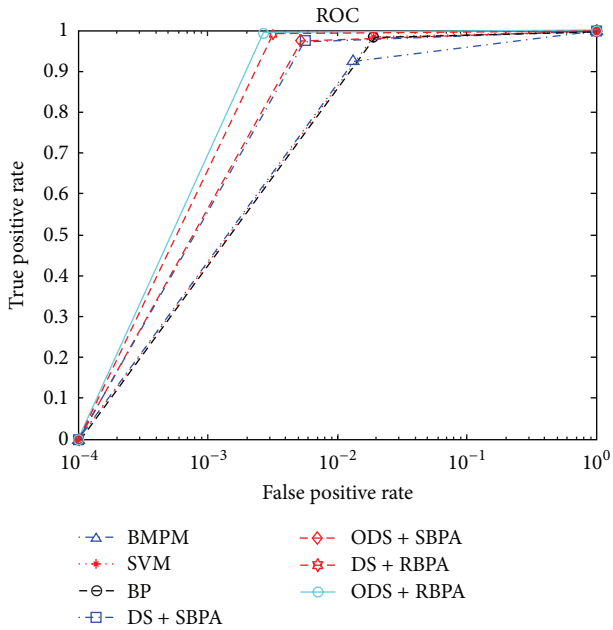


Figure 6: The ROC curve of BMPM, SVM, BP, DS with SBPA, ODS with SBPA, DS with RBPAS, and ODS with RBPA.

Table 9: Comparison of BMPM, SVM, BP, DS, ODS, SBPA, and RBPA.

Detection method	Precision	Recall	*F-Score*	AUC
BMPM	0.9835	0.9263	0.9540	0.9851
SVM	0.9752	0.9851	0.9801	0.9780
BP	0.9753	0.9839	0.9796	0.9781
DS with SBPA	0.9926	0.9748	0.9836	0.9934
ODS with SBPA	0.9933	0.9749	0.9840	0.9940
DS with RBPA	0.9957	0.9910	0.9934	0.9962
ODS with RBPA	0.9964	0.9930	0.9947	0.9968

5.2.3. Experiments Based on ROC and AUC. Table 9 shows that the results of all network normal and abnormal connections used by various anomaly detection methods. And the ROC curve of each method is depicted in Figure 6. In these two experiments, we employ ROC curve that shows the relationship of FR and DR, and AUC that represents the area under ROC curve. Here several parameters are utilized to estimate network anomaly detection system, such as Precious, Recall, and *F-Score* which are introduced in Section 4.3. Specially, the larger the values of parameters (*F-Score*, AUC) are, the better the performance of corresponding system is.

First, compared single detection methods, SVM, BMPM, and BP with fusion detection methods, we can see that single detection methods have smaller values of *F-Score* and AUC from Table 9; that is, the performance of single detection methods is lower than that of fusion methods. Ensuring an invariable condition in 4 fusion methods, we can analyze the effectiveness of RBPA and ODS. In this way, compared DS with SBPA and DS with RBPA, ODS with SBPA and ODS with RBPA, we can see that the methods with RBPA have higher *F-Score* and AUC values. As the same, the methods with ODS have higher *F-Score* and AUC values.

From Figure 6 that shows the ROC curve of 7 network anomaly detection methods, the network anomaly detection method merged with ODS and RBPA has the largest area under corresponding ROC curve (the largest AUC value in Table 9). When they have the same DR, FR of the network anomaly detection method merged with ODS and RBPA is the smallest one. Similarly, when they have the same FR, DR of the network anomaly detection method merged with ODS and RBPA is the highest one. So this fusion method is the best one in these 7 fusion methods.

5.2.4. Experiments with New Attacks. In this experiment, we utilize 3 network anomaly detection systems (BP, DS with SBPA and ODS with RBPA) to detect new attacks. Unlike experiments mentioned above, the data set used in this experiment is 10% KDD99 and test data set with 17 new attack types in Section 5.1.1.

From Table 10, we can see that the performance of single method is lower than that of fusion method. Most of new attack connections DR are higher than BP, but there still exist some abnormal DR, "sqlattack" for example. With ODS and RBPA optimization, this novel method we presented makes up this defect, which has a better new attack detection performance than others.

6. Related Work

The use of data fusion in the field of network anomaly detection is presented by Siaterlis and Maglaris [18]. The Dempster-Shafer theory of evidence is used as the mathematical foundation for the development of a novel anomaly detection engine. The detection engine is evaluated using the real network traffic. The superiority of data fusion technology applied to intrusion detection systems is presented in the work of Wang et al. [19]. This method used information collected from the network and host agents and application of Dempster-Shafer theory of evidence. Another work incorporating the Dempster-Shafer theory of evidence is by Hu et al. [20]. Wu et al. [21], proposed a framework of client-server architecture where the mobile agent continuously extracted various features and send to the server to detect anomaly using anomaly detectors. They used multiple distributed servers with different machine learning as a detector for analyzing the feature vector and D-S Evidence theory of information fusion is used to fuse the results of detectors, also proposed a cycle-based statistical approach to find anomaly activity. Zhouzhou et al. [22] presented a new algorithm based on D-S evidence theory to reduce energy consumption in wireless sensors network, which modifies D-S evidence theory and fuses it on cluster-head selection phase and adjusts operation period. The Dempster-Shafer theory of evidence in data fusion is observed to solve the problem of how to analyze the uncertainty in a quantitative way.

Table 10: Comparison of BP, DS with SBPA and ODS with RBPA for new attacks.

Attack name	Total connections	Detected connections			DR (%)		
		BP	DS + SBPA	ODS + RBPA	BP	DS + SBPA	ODS + RBPA
Apache2	794	792	794	794	99.75	100.00	100.00
httptunnel	158	155	157	158	98.10	99.37	100.00
mailbomb	5000	4893	5000	500	97.86	100.00	100.00
mscan	1053	1050	1050	1052	99.72	99.72	99.91
named	17	16	17	17	94.12	100.00	100.00
processtable	759	758	759	759	99.87	100.00	100.00
ps	16	16	16	16	100.00	100.00	100.00
saint	736	736	735	736	100.00	99.86	100.00
sendmail	17	14	16	17	82.35	94.12	100.00
snmpgetattack	7741	7704	7716	7739	99.52	99.68	99.97
snmpguess	2406	2404	2404	2406	99.92	99.92	100.00
sqlattack	2	2	1	2	100.00	50.00	100.00
udpstorm	2	0	2	2	0.00	100.00	100.00
worm	2	2	2	2	100.00	100.00	100.00
xlock	9	7	6	8	77.78	66.67	88.89
xsnoop	4	3	4	4	75.00	100.00	100.00
xterm	13	13	13	13	100.00	100.00	100.00
Average					89.65	94.67	99.34

Reference [11] presented a novel intrusion detection approach combining SVM and KPCA to enhance the detection precision for low-frequent attacks and detection stability. In order to shorten the training time and improve the performance of SVM classification model, an improved radial basis kernel function (N-RBF) based on Gaussian kernel function is developed, and GA is used to optimize the parameters of SVM. [14] proposed a flow-based anomaly detection system, which is trained with a flow-based data set. In this new system, multilayer Perceptron neural network with one hidden layer is used, which is added interconnection weights by a Gravitational Search Algorithm. Giacinto et al. [23] utilized general classifiers to divide various feature subspaces from the same data set and then merged voting, mean algorithm, Bayes, and decision module together. However, there exists less analysis about detection algorithm and fusion method. Another drawback of this model is higher false alarm rate. The formulation of the intrusion detection problem as a pattern recognition task using data fusion approach based on multiple classifiers is attempted by Didaci et al. [24]. The work confirms that the combination reduces the overall error rate, but may also reduce the generalization capabilities. Ambareen Siraj et al. [25] brought fuzzy cognitive map into fusion network anomaly detection and presented an intelligent network anomaly detection model. Thomas and Balakrishnan [26–28] selected artificial neural network as fusion algorithm and constructed fusion network anomaly detection model based on SNORT, PHAD, and ALAD that are open source detection systems. Although it was proved as an effective system, but its detection rate for some attacks was lower. In [29], performance of this fusion model is decided by diversity of various classifiers. [30] presented a novel network anomaly detection system with DS evidence theory and regression neural network, but its detection rate is lower.

7. Conclusion and Future Work

In this paper, we present a novel network anomaly detection model with ODS evidence theory and RBPA. When applying DS evidence theory on network anomaly detection model, we set weight for each sensor. And the weight value is from prior knowledge, *F-Score-N* and *F-Score-N,* which are extended by *F-Score*. Another key contribution is a new BPA function. We utilize regression ability of classifiers, SVM, BMPM, and BP, to compute various support degrees ($m(N)$, $m(A)$ and $m(\{N, A\})$) for each status (normal, abnormal, and unknown). Finally, we design 4 kinds of experiments to prove that this novel network anomaly detection model has higher DR and lower FR.

To further improve the performance of this new model, we will choose other sensors as classifiers to cope with complicated network records. And we will also utilize new attacks to evaluate this model. Finally, we will try to optimize DS evidence theory according to features of sensors.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by The National Natural Science Foundation of China (Grant no. 61103223) and the Key Program for Basic Research of Jiang Su (Grant no. BK2011003).

References

[1] A. G. Tartakovsky, A. S. Polunchenko, and G. Sokolov, "Efficient computer network anomaly detection by changepoint detection methods," *IEEE Journal on Selected Topics in Signal Processing*, vol. 7, no. 1, pp. 4–11, 2013.

[2] S. Staniford, J. A. Hoagland, and J. M. McAlerney, "Practical automated detection of stealthy portscans," *Journal of Computer Security*, vol. 10, no. 1-2, pp. 105–136, 2002.

[3] S. M. Bridges and R. B. Vaughn, "Fuzzy data mining and genetic algorithms applied to intrusion detection," in *Proceedings of 23rd National Information Systems Security Conference*, pp. 13–31, 2000.

[4] A. H. Sung and S. Mukkamala, "Identify important features for intrusion detection using support vector machines and neural networks," in *Proceedings of the Symposium on Application and the Internet*, pp. 209–216, 2003.

[5] G. Zhu and J. Liao, "Research of intrusion detection based on support vector machine," in *Proceedings of the International Conference on Advanced Computer Theory and Engineering (ICACTE '08)*, pp. 434–438, December 2008.

[6] W. Ren, J. Cao, and X. Wu, "Application of network intrusion detection based on Fuzzy C-means clustering algorithm," in *Proceedings of the 3rd International Symposium on Intelligent Information Technology Application (IITA '09)*, vol. 3, pp. 19–22, Nanchang, China, November 2009.

[7] T. Li and J. Wang, "Research on network intrusion detection system based on improved k-means clustering algorithm," in *Proceedings of the International Forum on Computer Science-Technology and Applications (IFCSTA '09)*, pp. 76–79, December 2009.

[8] C.-M. Bao, "Intrusion detection based on one-class SVM and SNMP MIB data," in *Proceeding of the 5th International Conference on Information Assurance and Security (IAS '09)*, vol. 2, pp. 346–349, Xia'n, China, September 2009.

[9] L. Lin, X. Xie, and S. Zhong, "A multiple classification method based on the DS evidence theory," in *Proceedings of the 9th International Symposium on Linear Drives for Industry Applications*, vol. 271 of *Lecture Notes in Computer Science*, pp. 587–596, 2014.

[10] W. Hu, J. Li, and Q. Gao, "Intrusion detection engine based on Dempster-Shafer's theory of evidence," in *Proceedings of the International Conference on Communications, Circuits and Systems (ICCCAS '06)*, pp. 1627–1631, Guilin, China, June 2006.

[11] F. Kuang, W. Xu, and Z. Siyang, "A novel hybrid KPCA and SVM with GA model for intrusion detection," *Applied Soft Computing*, vol. 18, pp. 178–184, 2014.

[12] C.-C. Chang and C.-J. Lin, *LIBSVM: A Library for Support Vector Machines*, 2010, http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[13] Y. Haiqin and H. Kaizhu, "MATLAB Toolbox for Biased Minimax Probability Machine," 2004, http://appsrv.cse.cuhk.edu.hk/~miplab/mempm_toolbox/index.htm.

[14] Z. Jadidi, V. Muthukkumarasamy, E. Sithirasenan, and M. Sheikhan, "Flow-based anomaly detection using neural network optimized with GSA algorithm," in *Proceedings of the IEEE 33rd International Conference on Distributed Computing Systems Workshops (ICDCSW '13)*, pp. 76–81, Philadelphia, Pa, USA, July 2013.

[15] "KDD Cup 1999 Data," 1999, http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

[16] J. Zhuge, D. Wang, Y. Chen, Z. Ye, and W. Zou, "Network anomaly detector based on the D-S evidence theory," *Journal of Software*, vol. 17, no. 3, pp. 463–471, 2006.

[17] The ROSETTA Homepage, 1998, http://www.lcb.uu.se/tools/rosetta/.

[18] C. Siaterlis and B. Maglaris, "Towards multisensor data fusion for DoS detection," in *Proceedings of the ACM Symposium on Applied Computing*, pp. 439–446, March 2004.

[19] Y. Wang, H. Yang, X. Wang, and R. Zhang, "Distributed intrusion detection system based on data fusion method," in *Proceddings of the 5th World Congress on Intelligent Control and Automation (WCICA '04)*, vol. 5, pp. 4331–4334, June 2004.

[20] W. Hu, J. Li, and Q. Gao, "Intrusion detection engine based on Dempster-Shafer's theory of evidence," in *Proceedings of the International Conference on Communications, Circuits and Systems (ICCCAS '06)*, pp. 1627–1631, June 2006.

[21] Z. Wu, X. Zhou, and J. Xu, "A result fusion based distributed anomaly detection system for android smartphones," *Journal of Networks*, vol. 8, no. 2, pp. 273–282, 2013.

[22] L. Zhouzhou, W. Fubao, and W. Wei, "A clustering algorithm based on D-S evidence theory for wireless sensor networks," *Information Technology Journal*, vol. 13, no. 13, pp. 2211–2217, 2014.

[23] G. Giacinto, F. Roli, and L. Didaci, "Fusion of multiple classifiers for intrusion detection in computer networks," *Pattern Recognition Letters*, vol. 24, no. 12, pp. 1795–1803, 2003.

[24] L. Didaci, G. Giacinto, and F. Roli, "Intrusion detection in computer networks by multiple classifiers systems," in *Proceedings of the International Conference on Pattern Recognition*, 2002.

[25] A. Siraj, R. B. Vaughn, and S. M. Bridges, "Intrusion sensor data fusion in an intelligent intrusion detection system architecture," in *Proceedings of the Hawaii International Conference on System Sciences*, pp. 4437–4446, January 2004.

[26] C. Thomas and N. Balakrishnan, "Advanced sensor fusion technique for enhanced intrusion detection," in *Proceedings of the IEEE International Conference on Intelligence and Security Informatics (ISI '08)*, pp. 173–178, Taipei, Taiwan, June 2008.

[27] C. Thomas and N. Balakrishnan, "Performance enhancement of intrusion detection systems using advances in sensor fusion," in *Proceedings of the 11th International Conference on Information Fusion (FUSION '08)*, pp. 1–7, July 2008.

[28] C. Thomas and N. Balakrishnan, "Improvement in intrusion detection with advances in sensor fusion," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 3, pp. 542–551, 2009.

[29] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–239, 1998.

[30] Y. Yuan, S. Shang, and L. Li, "Network intrusion detection using D-S evidence combination with generalized regression neural network," *Journal of Computational Information Systems*, vol. 7, no. 5, pp. 1802–1809, 2011.

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2014, Article ID 514652, 14 pages
http://dx.doi.org/10.1155/2014/514652

Research Article

A Hybrid Digital-Signature and Zero-Watermarking Approach for Authentication and Protection of Sensitive Electronic Documents

Omar Tayan,[1,2] Muhammad N. Kabir,[1,3] and Yasser M. Alginahi[1,4]

[1] *IT Research Center for the Holy Quran and Its Sciences (NOOR), Taibah University, Madinah 41411, Saudi Arabia*
[2] *College of Computer Science and Engineering (CCSE), Department of Computer Engineering, Taibah University, Madinah 41411, Saudi Arabia*
[3] *Department of Multimedia and Graphics, Faculty of Computer Systems & Software Engineering, Universiti Malaysia Pahang, Malaysia*
[4] *Academic Services, Department of Computer Science, Taibah University, Madinah, Saudi Arabia*

Correspondence should be addressed to Omar Tayan; omar_tayan@yahoo.co.uk

Received 12 April 2014; Revised 17 June 2014; Accepted 29 June 2014; Published 28 August 2014

Academic Editor: Iftikhar Ahmad

This paper addresses the problems and threats associated with verification of integrity, proof of authenticity, tamper detection, and copyright protection for digital-text content. Such issues were largely addressed in the literature for images, audio, and video, with only a few papers addressing the challenge of sensitive plain-text media under known constraints. Specifically, with text as the predominant online communication medium, it becomes crucial that techniques are deployed to protect such information. A number of digital-signature, hashing, and watermarking schemes have been proposed that essentially bind source data or embed invisible data in a cover media to achieve its goal. While many such complex schemes with resource redundancies are sufficient in offline and less-sensitive texts, this paper proposes a hybrid approach based on zero-watermarking and digital-signature-like manipulations for sensitive text documents in order to achieve content originality and integrity verification without physically modifying the cover text in anyway. The proposed algorithm was implemented and shown to be robust against undetected content modifications and is capable of confirming proof of originality whilst detecting and locating deliberate/nondeliberate tampering. Additionally, enhancements in resource utilisation and reduced redundancies were achieved in comparison to traditional encryption-based approaches. Finally, analysis and remarks are made about the current state of the art, and future research issues are discussed under the given constraints.

1. Introduction

Recent advancements in information and communication technologies combined with the widespread growth of the Internet have enabled the ease of digital content distribution, communication, and reproduction. Consequently, millions of users from the digital community are able to benefit from the advantages of the fast and simple digital information exchange. However, it is pointed out that such benefits come together in-hand with the problems and threats associated with ensuring digital copyright protection, preventing digital counterfeiting, proof of authenticity, and content-originality verification as an essential requirement largely for online disseminations of sensitive and specialized, formal, legal, financial, and religious content. Essentially, all such digital multimedia contents in the Internet can be classified into images, text, audio, and video, with the challenge being to ensure secure and reliable communications for each media type. This paper is primarily concerned with document integrity and source traceability with regard to widely disseminated digital text resources while reducing resource redundancies from traditional schemes when applied for our target domain. The problem of achieving authenticity and integrity verification for sensitive online text documents/media was presented in the literature as a challenging research problem in [1–15].

Most related studies on cryptography and copyright protection for authentication and integrity protection ignored the performance impact due to the high complexities and relatively large redundant implementation overheads used and particularly when applied for simpler applications that only require confirming authentication and integrity protection (rather than secrecy) of sensitive content [16]. In other cases, such schemes involved the overhead requirement for distributing algorithms and/or keys between communicating parties. For instance, well-known encryption-based digital-signature techniques had offered secrecy during data transmission, applied restrictions on data-access for copyright protection, and were able to detect unauthorized changes to the data. However, most of those schemes had involved large overheads in complex algorithmic computations and in the number of keys required, in addition to the distribution of those keys and algorithms between the communicating parties [16–18].

Other reasons also exit as to why encryption alone cannot provide a completely workable solution in particular applications. For instance, encryption carries overheads in resources and for some applications; it may be preferable to send data with no secrecy or such redundant overhead. Other cases include scenarios when some network management protocols separate confidentiality and integrity functions, rendering encryption alone as inappropriate.

A number of works based on hashing and message digests (MDs) were then proposed for achieving authentication and integrity with reduced overhead as a tradeoff for removing secrecy measures during transmission in order to achieve improved performance when applied in scenarios involving public-key algorithms [16, 19]. However, those works were primarily concerned with *accuracy* (e.g., only confirming authenticity and integrity) of the data rather than with the *performance* overhead incurred. Moreover, hashing approaches had involved the initial problem of exchanging "public-keys" between potentially many communicating parties. It is to the best of our knowledge that only few prior studies had focused on authentication and integrity schemes in the domain of both conflicting requirements (e.g., accuracy and enhanced performance) for those performance-dependent applications involving sensitive electronic documents.

More recently, steganography and watermarking techniques have been found in the literature for embedding hidden marker data in cover media without significantly degrading the quality of the media. Essentially, the hidden watermark serves to identify ownership and to verify its authenticity and integrity or otherwise to detect any modifications to the data. However, watermarking approaches are unable to control access to the data and hence are mainly ideal for applications that require integrity/authenticity verification rather than secrecy in the communications channel. In each of the above approaches (e.g., digital signatures, hashing, and watermarking), the primarily *accuracy*-based requirements were achieved by operating on any media type and sequence of bits (e.g., images, text, and audio bit patterns).

In this work, we focus on confirming authenticity and intact integrity of sensitive text content whose primary motive may compromise the need for secrecy in the communications channel during transmission. The motive here is that it may be required or even desirable that particular sensitive content should be freely propagated via multiple publishers/servers for wider outreach and dissemination. Hence, the well-understood relation between the client(s) and publisher/server now differs from the common one-to-one relation as in e-commerce transactions that had typically involved hashing or encryption algorithms being distributed between two or more known parties. Furthermore, the use of private keys for each (particular) client/receiver (as in public-key cryptosystems) is no longer required or applicable in our system, in which the goal of integrity robustness would require swiftly checking that sample documents from any client browser are authentic and untampered. This paper considers digital-signature and watermarking schemes for our target application domain and proposes a hybrid approach that employs concepts taken from digital-signature and watermarking schemes to achieve our goal. Our approach was evaluated through extensive experiments, with the results demonstrating that our scheme could be optimized for the target application domain of sensitive online texts that require authenticity and integrity verification with no secrecy in the communications channel. Significantly, results from our scheme had demonstrated that our goal could be achieved whilst avoiding the overhead of registering secret keys from all parties with a certification authority (e.g., as in symmetric-key signature schemes) as well as removing the need for separate public and private keys (the need for private keys was completely removed in our approach) for each communicating party (e.g., as in public-key signature schemes).

This paper is organized as follows: Section 2 provides the related work on digital-signatures and watermarking schemes, Section 3 explains the proposed hybrid digital-signature and zero-watermarking based framework, Section 4 discusses the analysis of the proposed framework, and finally Section 5 concludes the paper.

2. Related Work

2.1. Digital-Signature Schemes. Cryptography is used to protect information during the transmission process in applications that include emails, banking, sales, and corporate/private data. Cryptographic schemes are classified into symmetric-key systems and asymmetric-key systems [20]. Digital signature schemes are based on symmetric-key or asymmetric-key systems and offer effective mechanisms for facilitating content authenticity, integrity, and data-secrecy during transmission. The two most commonly used public-key digital-signature schemes are the Rivest-Shamir-Adleman (RSA) public-key encryption algorithm and the digital-signature algorithm (DSA) [21].

The work in [22] presents a theoretical performance analysis of DES and RSA with their working mechanisms. This study presents cases where public-keying schemes are preferred to secret-key systems. In [23], the comparison

between different symmetric cryptosystems concluded that DES is the most widely used encryption scheme, with 3DES being the slowest algorithm. In comparison, RC4 required the least memory space for implementation and had minimum simulation times. A summary of some traditional and commercial digital-signature techniques is classified as shown in Figure 1.

A number of works can be found in the literature with contributions mainly associated with limited improvements to the existing digital-signature techniques and algorithms. Examples of improvements developed in the literature include [17, 24–27]. In [24], the ElGamal digital-signature scheme was improved using a random number to increase the difficulty of a third-party obtaining the decipher key. Lui and Li [25] report on computation and communication improvements to a previously enhanced digital-signature scheme in the literature. Reference [27] discusses an efficiency enhancement to the RSA algorithm by speeding up certain array-based computations. Lin and Qiu [17] report on two improved digital-signature schemes based on a previous design of a directed signature scheme. Finally, a number of hybrid approaches had also reported some improvements to the existing and commercial techniques by combining digital signatures with either of watermarking, random numbers, and hash functions [18, 19, 24, 28, 29].

2.2. Steganography and Digital-Watermarking Schemes. In the literature, the techniques employed to provide the necessary copyright protection and integrity robustness for digital content are known as digital watermarking. A watermark is a signature or unique logo of an organization or individual who owns the rights to digital content [1] and typically contains information related to the copyrights, ownership, publisher, and document information [2]. Watermarking extends the information in the cover text and becomes an attribute of the watermarked document, in which "the object of communication is the packaging and the hidden message only references that packaging" [3]. Traditionally, digital-watermarking techniques are mainly used to embed identification data into the host cover document, in which the embedded data is a function of the host data/content bit sequences [4, 5, 30, 31]. Security issues of text-watermarking are the characteristic of its specific requirements and features and differ greatly from those of other multimedia watermarking schemes [6]. For example, it is relatively easy to insert watermark data into images as compared with plain text since the images contain plenty of redundant areas allowing the watermark data to be inserted whilst retaining perceptual similarity with the original file [2]. Plain text, on the other hand, has a clear structure and little/no redundant data (as found in the case of many languages including English), which negatively affects both the watermark capacity and security [7], therefore increasing the difficulty involved addressing this research problem.

Some of the objectives of the state of the art in digital text-watermarking can be classified into assuring authenticity and integrity of documents, identifying the origin or publisher/distributer of the contents, usage control, and general protection of documents [3]. Figure 2 outlines the important phases in the life cycle of a generic text-watermarking model.

A review of the literature evidences the maturity of watermarking and steganography based techniques in digital natural-language documents and digital text content in some languages including English, Persian, Turkish, and Chinese [4, 8–10], with only fewer techniques presented for the case of other semitic languages such as Arabic electronic texts [7, 8, 11]. Furthermore, watermarking of text documents has been classified into linguistic steganography and nonlinguistic steganography [12]. In the former, the techniques employed would typically manipulate the lexical, syntactic, and semantic properties while trying to preserve the meanings, whilst, in the latter approach, techniques are characterized by the file types and amendments are made to the text by using different text attributes to embed a message. Text-based watermarking has traditionally used shifting techniques or natural-language based watermarking [12]. Three types of text-watermarking shifting codes include line-shift coding, word-shift coding, and feature/character coding, whilst natural-language watermarking involves either of synonym substitutions or semantic transformation techniques which are very language-dependent [12]. On the other hand, the work on [13] classifies text-watermarking techniques into image-based techniques, syntactic-based manipulation, and semantic-based manipulation techniques which involve replacing the original text with alternative words in order to embed a hidden message whilst preserving the meanings as far as possible. Figure 3 summarizes some of the traditional watermarking techniques found in the literature for the different world languages.

In [6], Zhou et al. classified text-watermarking schemes into four categories of embedding modes: format watermarking, content watermarking, zero watermarking, and binary-image document watermarking [6]. The literature evidences, however, that text-watermarking is a relatively new field as compared with other forms of multimedia with slow development of techniques due to the simplicity and nonredundancy of the text [9]. Comparing fragile, semifragile, and robust watermarking, robust watermarking approaches have attracted attention of more researchers to date [9]. In either case, the designer's choice of watermarking approach should take into consideration the nature/characteristics of the target application since no single optimal scheme exists for all application types [6].

A key requirement for document protection arises with the need for users to confirm authenticity and integrity of the received text [14]. Many traditional text-watermarking techniques based on format-related embedding by modifying text layout and appearances have weak robustness [14]. Such approaches are vulnerable to the detection of the watermark data in the cover text and are more entitled to present themselves more for possible security attacks. Generally, text-watermarks can be attacked in a number of ways, which include inserting, deleting, and rearranging words and phrases [1]. Recently, however, zero-watermarking schemes have been proposed to overcome the problems of weak imperceptibility as well as the tradeoff that exists between robustness and imperceptibility [14, 15]. In such approaches,

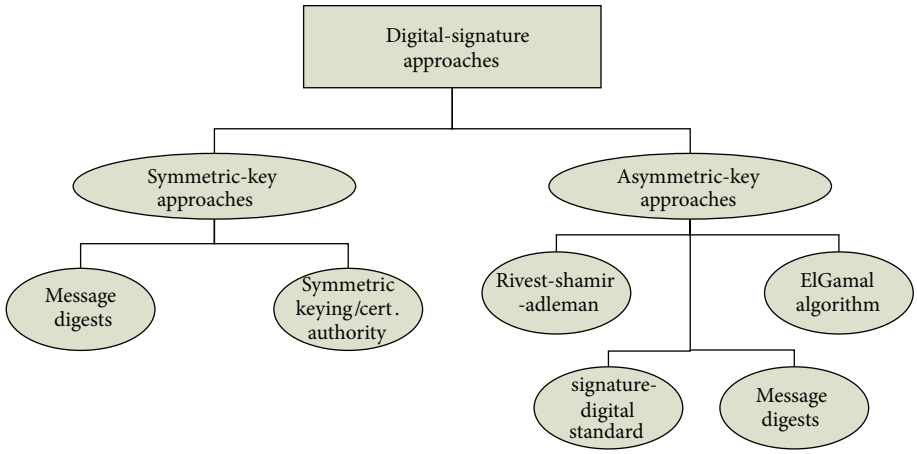


Figure 1: Classification of traditional digital-signature schemes.

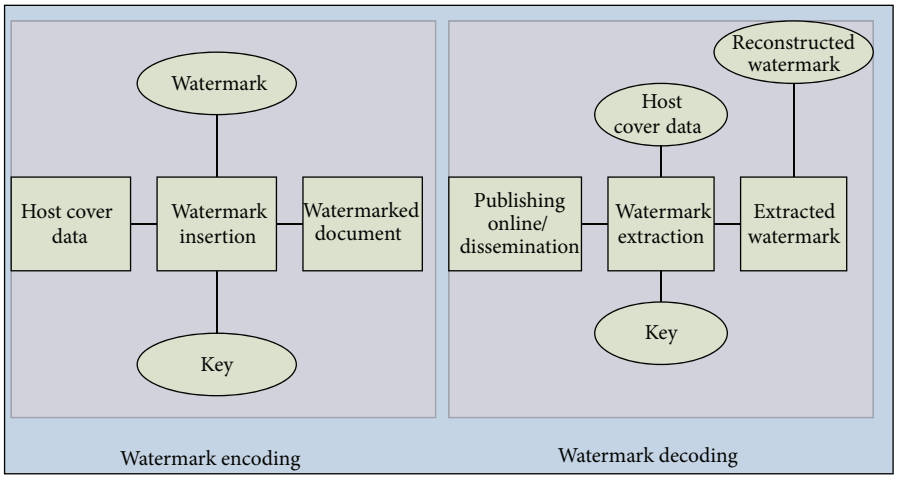


Figure 2: Phases in the watermarking life-cycle.

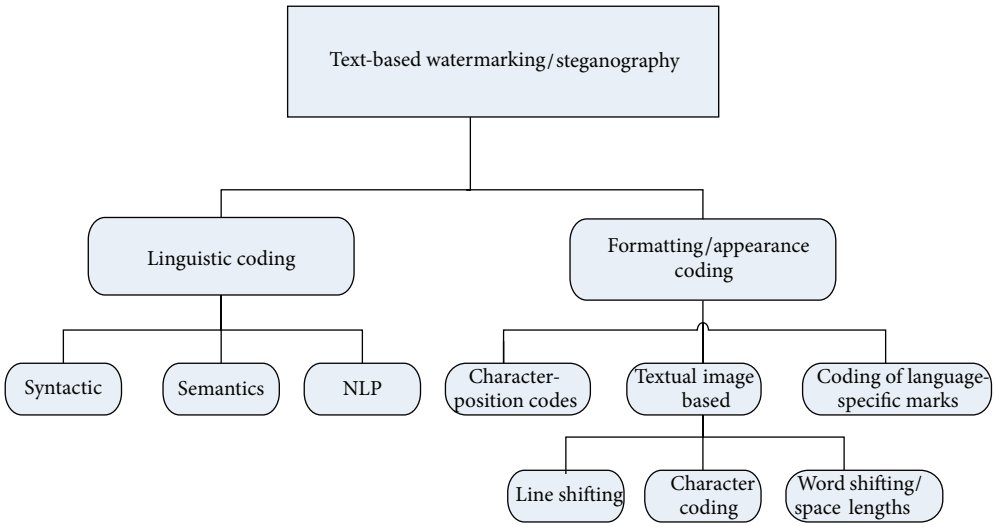


Figure 3: Digital-watermarking classification.

an attacker's examination of nonoriginal/unnormal formatting codes (causing distortion) in the cover text would be completely removed by eliminating the need for any physical embedding. Here, rather than physically inserting the watermark data, zero-watermarking schemes generate binary patterns during the encoding process by extracting essential characteristics from the host data which are then used in the detection process [14]. It is noted, however, that most of the existing zero-watermarking approaches are designed for image or audio media, with insufficient research conducted using such methods for text documents.

Furthermore, text-watermarking methods found in the literature are very limited and specific to few languages only, in addition to the lack in robustness, integrity, accuracy, and generality [13]. Hence, this work has been motivated by the need to address the deficiencies in text-watermarking, whilst addressing the challenges of generality, integrity, and robustness. In the proposed zero-watermarking approach presented here, no use of steganography is required, since no physical embedding of data is performed on the document. On the contrary, manipulations are performed on the document to determine whether or not the document has been modified and in order to verify the source. The next section describes our proposed hybrid scheme which addresses the above problems by ensuring language independency, invisibility, and robustness and preserves data integrity.

3. Proposed Hybrid Digital-Signature and Zero-Watermarking Approach

This paper introduces an implementation of a new design approach for integrity and authentication protection of plain text documents based on zero-watermarking with manipulations also related to digital-signature schemes. The proposed approach resembles digital-signature schemes through the manipulations required at the encoder and decoder as well as through the use of watermark keys/signatures used to verify source authenticity. On the other hand, our approach differs from traditional digital-signature schemes since in our scheme complex encryption operations and their associated overheads are not required during transmission. The goal here is to provide a mechanism for the secure dissemination of critical and sensitive documents in which any physical modification can render the document invalid for the user. Application examples of such requirements are numerous and include formal/official, financial, political, and religious text documents used to prove the original publisher in addition to assuring accuracy and integrity of the data. In the proposed approach, a novel hybrid framework related to digital signatures and zero-watermarking is described.

3.1. Description. The proposed algorithm performs a logical embedding of the watermark-data in the cover document. As such, the algorithm does not modify the text in the cover data to embed the watermark, but rather, watermark keys, W_{KG1} and W_{KG2}, are generated based on the *characteristics* of the text document. The Unicode standard is used in the encoder and decoder in order to encode all characters of the main worldwide languages and therefore provide support for worldwide language compatibility. Additionally, the objective of this paper is achieved using a blind watermark-extraction approach, since the original document is not required in the decoding phase and any detected change in the transmitted document/document-under-scrutiny is considered invalid for client use.

The embedding process (Figure 4) begins with an image logo, W_I, being converted into a character sequence, W_{CS}, and embedded in a copy of the cover document, T_C. The *image-to-text converter* block at the encoder can be generalised/replaced with other media converters and therefore made applicable to any multimedia input or digital information that converts the data into a binary string prior to the encoding process. Meanwhile, the original document, T_O, is unaltered and sent for online dissemination. The watermark logo, W_I, is the unique signature of an organization/publisher or individual that owns rights to the digital content/online document. The embedding phase is based on a spread-spectrum technique that inserts one-watermark character per set (insertions only into the first word of each set), with the set size, S, being set to two words, forming a word pair. The result of the embedding is then passed for processing within the document analyzer and classifier (DAC), which uses the Unicode standard to numerate the words into binary Unicode summations (sum_j for the first word and sum_{j+1} for the second word) for further processing. Next, we use a logical XOR operation/function of the kth bit-positions of both words in each word pair set to produce an (F_k) function code for each of the bit positions.

An example of generating a partial function code from W_I and T_C is illustrated in Figure 5 (example bit sequences shown may not be representative of actual words used).

The example in Figure 5 shows the publishers logo, W_I, in binary format before being converted into the corresponding character sequence, W_{CS}. Each of the embedding characters (e_i) is then embedded into the first words of each word set in T_C with e_0 being embedded into the first word of the first word set and e_1 being embedded into the first word of the second word set and so on.

One of the main components in the encoding process is the use of the DAC, which is comprised of a document-analysis phase and a bit-pair classifier. The DAC consists of two main components: the *analyzer* which converts each word into Unicode summations and a logical-XOR *classifier* of similar bit positions of adjacent words. The *document analyzer* is used for the conversion of each word in the cover text into a binary summation of its constituent characters, whereas the *classifier* passes through the document, sampling similar bit positions in adjacent words of each set and producing a one-bit result of the XOR operation, an (F_k) function code for each bit position operation between the two words. Similar function codes are then generated for the remaining bit-positions in the set. It is assumed that after all necessary summation operations, each word is represented using a 17-bit binary result. In this algorithm, two 16-bit Unicode values, as in the standard Unicode table [32], were added together, which produces a 17-bit result in the case of an overflow. Hence, each word set allocated 17-bit storage/memory to

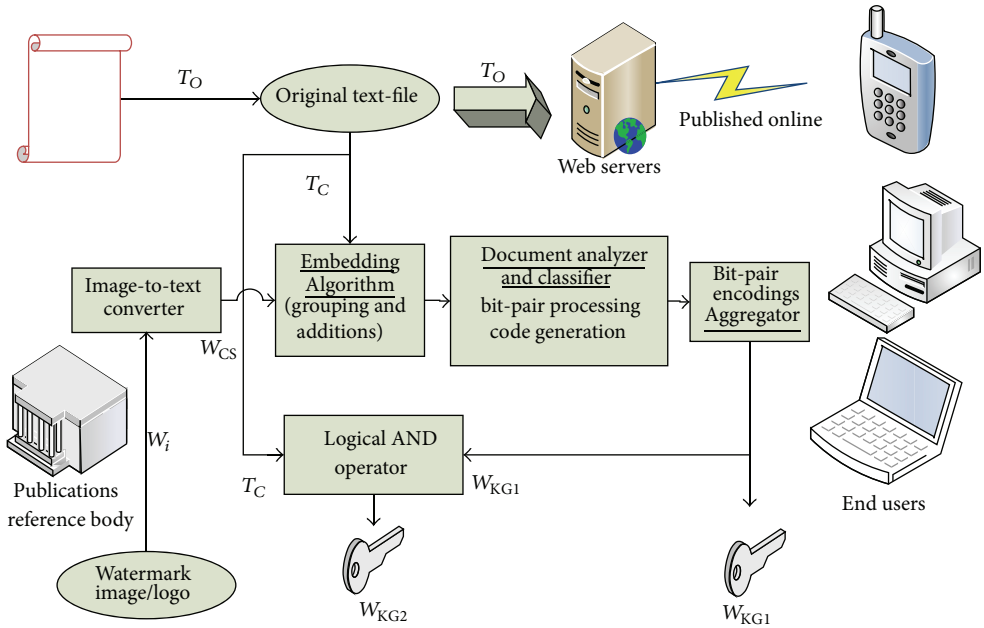


Figure 4: Watermark encoding process.

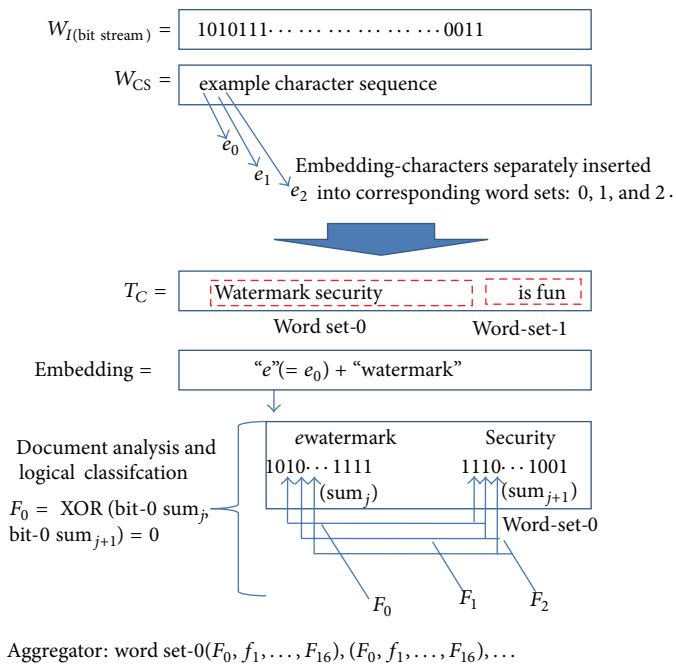


Figure 5: Partial function code generated in the encoder.

contain the result of the addition operation from the previous step. The DAC operation in the example of Figure 5 is shown to produce an F-result bit sequence through the logical operations on each bit position of the words in each set, before being output and aggregated.

The DAC processes each set in turn, with each word pair in a set being clearly separated by spaces, and words are assumed to begin with only nonspace characters. During the analysis phase, spaces are considered as part of the previous word encountered. By passing through the entire document, the *classifier* would be responsible for generating the individual function codes which are then used as input to the aggregator to produce a unique key W_{KG1}. In the algorithm proposed in this paper, all logical and comparison operations are performed on Unicode binary values to extend our approach to all Unicode-supported languages. Next, a logical AND operation using the W_{KG1} and T_C is used to generate a second unique key, W_{KG2}, as shown in Figure 4.

Notably, all input characters (from T_C and W_{CS}) were padded to 16-bit Unicode values, ensuring that logical operations (like ADD during the DAC stage or in the last/AND stage between W_{KG1} and T_C) result with no loss in the 17-bit results generated. Finally, the two keys generated, together with the original document and time-stamped logo, are registered with a certification authority (CA)—a trusted third-party intermediary body in the digital community.

Enquiries pertaining to document authenticity, source tracing, and tamper detection tests of an online document are addressed using the decoding process, whereby the document under scrutiny, T_S, is passed for processing, in which the *analyzer* algorithm converts each word into binary Unicode values. This proceeds with CA embedding the stored signature/key, W_{KG2}, into the output produced by the *analyzer* using a NAND operation, the result of which is passed to the *comparator*. Simultaneously, CA passes the unique key, W_{KG1}, into the comparator for equivalence testing (w) between the W_{KG1} and W_{KE}. If the document is valid, the decoder extracts the characters ($e_0 \cdots e_i$) of the embedded W_{CS} character stream and converts the embedded data into a watermark image, W_I, (using a text-to-image convertor) which thereby identifies the true owner. The details of the proposed decoder are shown in Figure 6.

In the proposed system, the document owner/publisher is responsible for generating the watermark keys/signatures (W_{KG1} and W_{KG2}) and registering the time-stamped key, logo, and algorithm with the CA, whilst the CA is responsible for decoding the digital content and examining the watermark during the verification/decoding process for purposes of authenticity and source verification upon client requests. Hence, the correct keys (to be stored at the CA) required for verification checking at the client side can only be generated from the known publisher given that the original document is used as input to the encoder. Furthermore, the algorithm is only required by the publisher (and not the CA or client side) and hence is not made public.

3.2. Encoder and Decoder Algorithm Design. The watermark encoding and decoding algorithms are presented in Algorithms 1 and 2.

3.3. Design Issues and Advantages. The approach proposed in this paper ensures that the hybrid logical-watermark concept remains intact and valid in the following scenarios:

(i) when the font style, size, colour, and so forth are modified;

(ii) when the whole document is copied (e.g., transported) onto another empty or nonempty document;

(iii) when document integrity remains robust in the face of OCR techniques and exact retyping with the support of the standard Unicode format;

(iv) when the detected watermark cannot be destroyed without distortion and therefore invalidating the document at the end user.

Furthermore, the logical watermark is characterized by the following.

(i) It cannot be detected, derived, or extracted from the host document, therefore achieving 100% imperceptibility.

(ii) There is no additional increase to the original file size.

(iii) A partial copy of the document does not allow the watermark to be detected.

(iv) Scrutinizing the authenticity of a document in question can be performed by extracting/detecting the watermark to prove the rightful author.

(v) During the detection process, tampered documents may be evaluated as traceable to an original source based on the "*closeness measure,*" which measures the degree of similarity (e.g., as in the ratio of similar bits) of the extracted/recovered watermark image with the closest CA-registered watermark image. This in turn may be used to identify the locations of the modified bits in the document.

(vi) Our encoding method supports circularly embedding of the watermark image in the document allowing for increased robustness and tamper detection abilities, since the watermark can be extracted from multiple segments of the document and compared for locating modified characters.

A drawback of this approach is evident in the space required at the CA's database for storing the keys W_{KG1} and W_{KG2}, generated at the encoder side. In this study, a set size of 2 was considered (as an inner parameter), which for a document of 20,000 words requires 10,000 word sets $*$ 17 bits per word set = 170,000 bits of storage at the CA. This problem of large storage requirements can be addressed since the encoder design enables the set size to be readjusted at the publisher side (since it is only a fixed-value input to the encoder algorithm) to accommodate the CA's space limitations when necessary.

4. Results Analysis/Summary of Results

This section provides results and analysis of the proposed logical-watermarking approach in terms of our computational cost and application-driven cost-function requirements: *imperceptibility* and *document-integrity robustness* for authenticity and tamper detection. The *imperceptibility* requirement is addressed given that no one other than the owner and CA can know about the existence of any watermark in the document since the original text, T_O, is unchanged after encoding. Consequently, unauthorized parties are not able to detect any existing watermarks, thereby reducing the probability of attacks or tampering via the communications channel. The *document-integrity robustness* requirement is essential for document authenticity and tamper detection and is addressed by detecting any change in the original document (e.g., at the comparator stage) as when the document has been subject to third-party modifications which would invalidate the document for end users. Notably, our design had enabled the retrieval of the original publisher logo, following the validity decision in

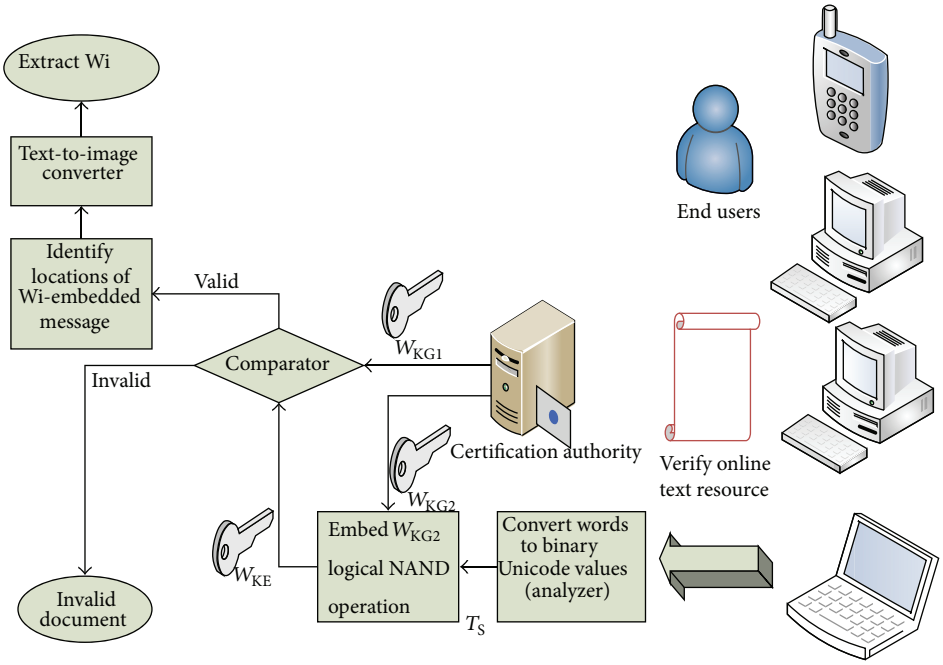


Figure 6: Watermark decoding process.

Input: original cover-document T_0, logo-watermark W_I, fixed set-size, L, with $L = 2$ (an inner-parameter).
Output: watermark key data-sequence, W_{KG1}, based on aggregation of (DAC) and second key W_{KG2} generated by a logical operation on T_0 and W_{KG1}
(1) Convert W_I to a character-stream, W_{CS}.
(2) Make a *copy-document* of the original *cover-document*, T_C.
(3) Divide T_C into n-word sets according to the set-size L.
(4) **for** $i = 1$ to n
    get i th word-set from T_C.
    add the Unicode values of all characters in each word of set-i to produce sum_j and sum_{j+1}
    get next embedding-character e from W_{CS}
    add the Unicode value of e to sum_j, and convert sum_j and sum_{j+1} to binary values.
    while bits left in sum_j to process **do**
        perform XOR-classification using logical-bitwise operations on the kth bit-position of sum_j and sum_{j+1} to generate an (F_k) result.
        aggregate all (F_k) results produced by passing through each of kth bits (of first and second words in current set) in turn.
    end while
    aggregate all F_k bit-sequences produced by combining result for each of the previous i word-sets, into F.
end for
(5) Obtain the first key-generated, W_{KG1}, as the result of aggregation of F produced for all word-sets.
(6) Perform logical AND-operation between the W_{KG1} and T_C to generate the second-key, W_{KG2}.
(7) Output W_{KG1} and W_{KG2} for the whole document.

Algorithm 1: Encoding algorithm.

Input: Document under scrutiny, T_S, watermarked keys; W_{KG1}, W_{KG2}, fixed set-size, $L = 2$.
Output: Validity decision, output image logo W_I that identifies the owner.

(1) Divide T_S into n-word sets according to the L.
(2) **for** $i = 1$ to n
    get i-th word-set from T_S.
    apply *analyzer* to add the Unicode values of all characters in each word of set-i to produce sum_j and sum_{j+1}
    convert sum_j and sum_{j+1} into binary values.
end for
(3) Embed W_{KG2} into T_S by performing a logical NAND-operation, to produce an extracted key, W_{KE}.
(2) Compare W_{KE} with the W_{KG1} to test for equivalence, w.
(3) **if** $(w == 1)$ **then**
  while bits left in F-result, in W_{KE} **do**
    Find the bit-pattern of the first-words of each set, using knowledge from the valid-document, T_S, of the binary values of the second words in each word-set together with generation of possible inputs of the logical-XOR operation (from encoder)—as in the F_k-results found in W_{KE}-to identify the unknown bits of the first-word within each word-set.
    Convert the derived first-word bit-results (r_1) and the first-word bit-results of T_S (r_2) into decimal.
    Perform $r_1 - r_2$ to obtain the Unicode-decimal values.
    Determine the embedded characters in each of the first-words of W_{KE} from the Unicode-values.
  end while
end if
(4) Convert extracted characters to produce an image W_I, using the Text-To-Image convertor.
(5) Output W_I logo or invalid-decision.

Algorithm 2: Decoding algorithm.

the decoder using a function code generation scheme that allows us to obtain the embedded W_I-bits at the encoding phase. Furthermore, as explained in the "*design issues*" discussion (Section 3.3) above, our encoder supports accumulative (circular) watermark embedding which had also resulted in increased robustness.

On the other hand, traditional text-watermarking involves embedding a watermark through the modification of document layout and appearance therefore possessing poor robustness since they cannot recover the watermark following simple formatting operations on the document [14]. The numerical results pertaining to real-sample tests are highlighted in Table 1, with the analysis and discussion of the benefits and features of those approaches being presented in Table 2. In Tables 1 and 2, the proposed method is compared with traditional format-encoding based watermarking and text-based zero-watermarking methods from the literature.

Our encoding and decoding algorithms were implemented in C++. The programs were compiled by a C++ compiler of version GCC 4.8 under the Linux operating system *Ubuntu 11.10*. Tests were run on a Pentium i3 processor of 1.7 GHz. The computational times were computed using the standard C function *clock()* which requires the header file *ctime*. The following C++ code fragment was used to calculate the execution time of an algorithm in seconds:

```
const clock_t beginTime = clock();
//encoding/decoding algorithm;
double computationalTime = double (clock () − beginTime)/CLOCKS_PER_SEC.
```

In Table 1, five sample text *files* were used in all computations of encoder and decoder times for the algorithms being compared; those had consisted of algorithms [33–36] and our new proposed algorithm in this paper. Each of the encoder and decoder computational times was then calculated for each of the sample text files arranged in increasing order of size. Additionally, an average computational time per character metric (ave./character) was calculated for each algorithm to provide an indication of the average delay performances over various input text sizes. Table 1 shows that our proposed approach is very comparable with the existing approaches in

Table 1: Computational-cost comparison between relevant approaches.

File name	No. of Chars	Computational time [encoder (ms)]						Computational time [decoder (ms)]					
		(Tayan et al., [33])	(Jalil et al., [34])	(Meng et al., [35])	Method A from [36]	Method B from [36]	Proposed approach	(Tayan et al., [33])	(Jalil et al., [34])	(Meng et al., [35])	Method A from [36]	Method B from [36]	Proposed approach
Text 1	28915	30	180	210	20	40	**20**	20	170	224	10	20	**20**
Text 2	47974	40	160	350	30	60	**40**	30	240	371	20	40	**40**
Text 3	54839	60	195	410	50	100	**60**	40	278	460	40	40	**70**
Text 4	116794	80	1714	1850	70	190	**100**	70	1684	1891	60	80	**110**
Text 5	166166	130	590	620	120	300	**160**	100	520	640	100	130	**170**
Ave./char		0.00089	0.0063	0.0083	0.00071	0.0016	**0.00089**	0.00065	0.0067	0.0088	0.00052	0.00074	**0.00095**

Table 2: Comparison of features and benefits between watermarking methods.

Concept/metric	Traditional watermark approaches	Zero-watermarking approach [1]	Zero-watermarking approach [35]	Zero-watermarking approach [34]	Zero-watermarking approach [33, 36]	Proposed watermarking method
Overhead (additional)	Proportional to embedded key size	None	None	None	None	None
Embedding mode	Format-encoding	Logical: zero-watermark encoding	Logical: zero-watermark encoding	Logical: zero-watermark encoding	Logical: zero-watermark encoding	Logical: zero-watermark encoding
Location of watermark message (W_M)	Embedded in T_O	Embedded in T_C	Embedded in T_C	Embedded in T_C	Embedded in T_C	Embedded in T_C
Processing and embedding decision	Based on searching through text for candidate words, lines, and spaces	Based on double-letter words in English language	Based on sentence entropy	Based on the first letter with specific word lengths	Based on comparing Unicode summations	Based on comparing Unicode summations and logical operations
Compatible with various formats?	Limited	Yes, only English character support needed	Supports Chinese language only	Supports English language only	Yes, host-document language character support/Unicode needed	**Yes, host-document language character support/Unicode needed**
Language-dependent	No	Yes	Yes	Yes	No	**No**
Document- authenticity verification	No	Yes	Yes	Yes	Yes	Yes
Tamper-detection and identification capabilities	No	No	No	Yes	No	**Yes**
Integrity robustness	Weak	Strong	Strong	Strong	Strong	Strong
Perceptibility performance	Low-medium	High	High	High	High	High
Capacity ratio performance	Inversely proportional to perceptual similarity	High	High	High	High	High
Capability to extract publishers watermark logo	No	Yes	No	No	No	**Yes**

terms of computational cost requirements. At the encoder stage, our proposed method had provided improvements over the works in [34, 35] and Method B from [36], whilst performing approximately equal to the approach in [33] and less than that of Method A [36]. Meanwhile at the decoder stage, the proposed method had produced results better than that of [34, 35] and longer times than those of [33, 36] due to the additional times required in identifying tampered locations and/or extracting the publisher's logo when needed.

Next, a comparison of the features and benefits between traditional watermarking methods, prior zero-watermarking methods [1, 33–36], and the new proposed watermarking method is now examined in detail in Table 2. A number of frequently recurring performance metrics and concepts from the digital-watermarking literature were considered as the basis for our comparisons. Those comparative metrics had included properties such as overhead complexity, embedding type, effect of watermark embedding on the original file, compatibility, and language dependencies. Moreover, a comparison of some powerful capabilities had included authenticity verification, tamper detection/identification, and the degree of robustness, imperceptibility, and capacity. Finally, each algorithm's ability to extract the owner's/publisher's watermark logo at the client side is presented in the final row of Table 2.

From Table 2, numerous improvements are identified by comparing the proposed approach in this paper (e.g., the rightmost column) with that of other traditional approaches from the literature. Specifically, it is shown that key advantages are gained over traditional approaches in terms of overhead, watermark transportability, compatibility, document authenticity verification, and tamper detection capabilities. Hence, the proposed approach had resulted in improvements in the key performance metrics of integrity robustness, imperceptibility, and capacity ratio (last three rows in Table 2). Furthermore, several outstanding improvements were also evident when comparing our proposed approach with existing zero-watermarking approaches from the literature. For instance, key enhancements were observed in format-compatibility, language independency, tamper detection, identification capabilities, and finally, through its ability to extract the publishers watermark logo. From Table 2, further benefits are also evident over traditional approaches, such as imperceptibility performance and integrity robustness.

From Table 2, the advantages of our approach are also highlighted in the final approach (shown in the rightmost column). In all approaches considered, the overhead parameter had only referred to the overhead at the publisher/encoder side. In Section 3, a discussion was given on the overhead of our approach on the CA. Notably, the work in [33, 36] was closely matched in most advantages except that our new approach here has the ability to exactly extract the publisher's watermark logo rather than simply determine whether the text-under-scrutiny (T_S—from the decoder-side) is valid or not. Significantly, our new approach was able to localise tamper attempts performed on T_S, an improvement not previously found in the related work [33, 35, 36]. These contrasting features that improve the work from [33] are observed in the last row and the fifth last row in Table 2.

The proposed technique is a new zero-watermarking approach which can deal with sensitive documents. From Tables 1 and 2, it is noted that the proposed method may not outperform other methods in terms of computational time for the encoding and decoding phases; however, compared to other methods the proposed method addresses some of the weaknesses found in the current available techniques; that is, it is not language-dependent; it has tamper detection and identification capabilities; it is robust and capable of extracting publishers watermark logo.

5. Discussion and Conclusion

The proliferation and expansion of the Internet suggest that more attention is required for the security and protection of online multimedia data and particularly for the dominant text medium. Many existing text-watermarking algorithms lack the requirements of robustness, imperceptibility, and document authenticity verification. This paper has proposed a novel hybrid approach involving concepts from digital signatures and logical text watermarking independent of the underlying language, given that it can be encoded in standard Unicode. The proposed algorithm can be used to protect electronic documents and digital textual content from tampering, forgery, and illegal content manipulation, whilst removing many implementation redundancies and complexities found in previous schemes. Additionally, the proposed approach can achieve effective protection and authenticity verification, while its computational costs and quality of results obtained are completely practical. The drawbacks that are being considered for improvement in future work involve reducing storage requirements at the CA and further enhancing computational times, both of which become more significant for very large text document samples. Significant contributions of this paper include introduction of a new design framework for a text-based logical watermarking scheme, a mechanism for adapting and optimizing the framework for specific target-applications, and finally demonstrating how such an approach can bypass the menace of most publishers' watermark targeted attacks by avoiding all such physical and vulnerable/suspicious modifications in the text due to the encoding process.

Future work and open research issues have emerged as a result of this work and primarily involve, first, testing our approach with larger varieties of sensitive online document samples and enhancing the proposed approach to become a commercially viable solution, to be developed as an essential tool for reference/certification bodies/organizations concerned with the dissemination of sensitive/critical text resources. Second, the planned next phase of this work considers evaluating language-specific embedding characters and their benefits on our performance metrics of interest and whether they can be used to enhance our cost parameters. Other opportunities for future work involve adapting our approach to the other major applications of text-watermarking, namely, copyright protection of text documents, by comparing the recovered/decoded watermark from

the illegally copied document with the watermarks stored at the CA, in terms of their *degrees-of-similarity*. Finally, it is also anticipated that this work will open new research directions aimed at developing and advancing the state of the art in multimedia-based logical watermarking in the two major application domains of copyright protection and authenticity verification/tamper detection.

Conflict of Interests

The authors declare that they have no conflict of interests regarding the publication of this paper.

Acknowledgment

The authors would like to thank and acknowledge the IT Research Centre for the Holy Quran and Its Sciences (NOOR) at Taibah University for their financial support during the academic year 2012/2013 under research Grant reference no. NRC1-126.

References

[1] Z. Jalil, A. M. Mirza, and T. Iqbal, "A zero-watermarking algorithm for text documents based on structural components," in *Proceedings of the International Conference on Information and Emerging Technologies (ICIET '10)*, pp. 1–5, Karachi, Pakistan, June 2010.

[2] M. A. Qadir and I. Ahmad, "Digital text watermarking: Secure content delivery and data hiding in digital documents," in *Proceedings of the 39th Annual 2005 International Carnahan Conference on Security Technology (CCST '05)*, October 2005.

[3] B. Barán, S. Gómez, and V. Bogarín, "Steganographic watermarking for documents," in *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*, Maui, Hawaii, USA, January 2001.

[4] Z. Yu and X. Liu, "A new digital watermarking scheme based on text," in *Proceeding of the 1st International Conference on Multimedia Information Networking and Security (MINES '09)*, vol. 2, pp. 138–140, Hubei, China, November 2009.

[5] C. Chen, S. Wang, and X. Zhang, "Information hiding in text using typesetting tools with stego-encoding," in *Proceedings of the 1st International Conference on Innovative Computing, Information and Control 2006 (ICICIC '06)*, pp. 459–462, September 2006.

[6] X. Zhou, W. Zhao, Z. Wang, and L. Pan, "Security theory and attack analysis for text watermarking," in *Proceedings of the International Conference on E-Business and Information System Security (EBISS '09)*, pp. 1–6, Wuhan, China, May 2009.

[7] F. Al-Haidari, A. Gutub, K. Al-Kahsah, and J. Hamodi, "Improving security and capacity for arabic text steganography using 'Kashida' extensions," in *Proceedings of the 7th IEEE/ACS International Conference on Computer Systems and Applications (AICCSA '09)*, pp. 396–399, May 2009.

[8] M. Shirali-Shahreza and S. Shirali-Shahreza, "Persian/arabic unicode text steganography," in *Proceedings of the 4th International Symposium on Information Assurance and Security (IAS '08)*, pp. 62–66, September 2008.

[9] Y. Zhang, H. Qin, and T. Kong, "A novel robust text watermarking for word document," in *Proceedings of the 3rd International Congress on Image and Signal Processing (CISP '10)*, pp. 38–42, October 2010.

[10] R. Davarzani and K. Yaghmaie, "Farsi text watermarking based on character coding," in *Proceedings of the International Conference on Signal Processing Systems (ICSPS '09)*, pp. 152–156, May 2009.

[11] M. A. Aabed, S. M. Awaideh, A. M. Elshafei, and A. A. Gutub, "Arabic diacritics based steganography," in *Proceedings of the 2007 IEEE International Conference on Signal Processing and Communications (ICSPC '07)*, pp. 756–759, November 2007.

[12] A. O. Adesina, H. O. Nyongesa, and K. K. Agbele, "Digital watermarking: a state-of-the-art review," in *Proceedings of the 5th IST-Africa Conference and Exhibition*, May 2010.

[13] Z. Jalil and A. M. Mirza, "A review of digital watermarking techniques for text documents," in *Proceeding of the International Conference on Information and Multimedia Technology (ICIMT '09)*, pp. 230–234, Jeju Island, Republic of Korea, December 2009.

[14] X. Zhou, Z. Wang, W. Zhao, S. Wang, and J. Yu, "Performance analysis and evaluation of text watermarking," in *Proceedings of the 1st International Symposium on Computer Network and Multimedia Technology (CNMT '09)*, December 2009.

[15] X. Zhou, W. Zhao, Z. Wang, and L. Pan, "Security theory and attack analysis for text watermarking," in *Proceedings of the International Conference on E-Business and Information System Security (EBISS '09)*, May 2009.

[16] A. Tanenbaum, *Computer Networks*, Prentice Hall, New York, NY, USA, 5th edition, 2010.

[17] B. Lin and H. Qiu, "Two improved digital signature schemes," *Journal of Systems Engineering and Electronics*, vol. 12, no. 1, pp. 78–81, 2001.

[18] L. Zhu, "Electronic signature based on digital signature and digital watermarking," in *Proceedings of the 2012 5th International Congress on Image and Signal Processing (CISP '12)*, pp. 1644–1647, October 2012.

[19] C. Zhou, G. Zhu, B. Zhao, and W. Wei, "Study of one-way hash function to digital signature technology," in *Proceedings of the International Conference on Computational Intelligence and Security (ICCIAS '06)*, pp. 1503–1506, October 2006.

[20] O. P. Verma, R. Agarwal, D. Dafouti, and S. Tyagi, *Performance Analysis of Data Encryption Algorithms*, IEEE Delhi Technological University, New Delhi, India, 2011.

[21] S. R. Subramanya and B. K. Yi, "Digital signatures," *IEEE Potentials*, vol. 25, no. 2, pp. 5–8, 2006.

[22] A. Kumar, S. Jakhar, and S. Makkar, "Distinction between secret key and public Key cryptography with existing glitches," *Indian Journal of Education and Information Management*, vol. 1, no. 9, pp. 392–395, 2012.

[23] H. Agrawal and M. Sharma, "Implementation and analysis of various symmetric cryptosystems," *Indian Journal of Science and Technology*, vol. 3, no. 12, pp. 1173–1176, 2010.

[24] L. Xiao-Fei, S. Xuan-Jing, and C. Hai-Peng, "An improved ElGamal digital signature algorithm based on adding a random number," in *Proceedings of the 2nd International Conference on Networks Security, Wireless Communications and Trusted Computing (NSWCTC '10)*, pp. 236–240, Wuhan, China, April 2010.

[25] J. Liu and J. Li, "Cryptanalysis and improvement on a digital signature scheme without using one-way hash and message redundancy," in *Proceedings of the 2nd International Conference on Information Security and Assurance (ISA '08)*, pp. 266–269, Busan, Republic of Korea, April 2008.

[26] H. Wang and S. Zhao, "Cryptanalysis and improvement of several digital signature schemes," in *Proceedings of the 2nd International Conference on Networks Security, Wireless Communications and Trusted Computing (NSWCTC '10)*, pp. 306–309, April 2010.

[27] C. Fu and Z. Zhu, "An efficient implementation of RSA digital signature algorithm," in *Proceedings of the International Conference on Wireless Communications, Networking and Mobile Computing (WICOM '08)*, October 2008.

[28] C. Hu and X. Wang, "Zero watermark protocol based on timestamp and digital signature," in *Proceedings of the International Forum on Information Technology and Applications (IFITA '09)*, pp. 193–196, May 2009.

[29] S. S. P. Shukla, S. P. Singh, K. Shah, and A. Kumar, "Enhancing security & integrity of data using watermarking & digital signature," in *Proceedings of the 2012 1st International Conference on Recent Advances in Information Technology (RAIT '12)*, pp. 28–32, March 2012.

[30] C. I. Podilchuk and E. J. Delp, "Digital watermarking: algorithm and application," *IEEE Signal Processing Magazine*, vol. 18, no. 4, pp. 33–46, 2001.

[31] M. D. Swanson, M. Kobayashi, and A. H. Tewfik, "Multimedia data-embedding and watermarking technologies," *Proceedings of the IEEE*, vol. 86, no. 6, pp. 1064–1087, 1998.

[32] Unicode Character Table, http://unicode-table.com/en/.

[33] O. Tayan, Y. Alginahi, and M. N. Kabir, "An adaptive zero-watermarking approach for authentication and protection of sensitive text documents," in *Proceedings of the International Conference on Advances in Computer and Information Technology (ACIT '13)*, May 2013.

[34] Z. Jalil, A. M. Mirza, and H. Jabeen, "Word length based zero-watermarking algorithm for tamper detection in text documents," in *Proceeding of the 2nd International Conference on Computer Engineering and Technology (ICCET '10)*, vol. 6, pp. V6-378–V6-382, Chengdu, China, April 2010.

[35] Y. Meng, T. Guo, Z. Guo, and L. Gao, "Chinese text zero-watermark based on sentence's entropy," in *Proceedings of the International Conference on Multimedia Technology (ICMT '10)*, October 2010.

[36] O. Tayan, Y. Alginahi, and M. N. Kabir, "Performance assessment of zero-watermarking techniques for online arabic textual-content," *Life Science Journal*, 2013.

Hindawi Publishing Corporation
The Scientific World Journal
Volume 2014, Article ID 536930, 21 pages
http://dx.doi.org/10.1155/2014/536930

Research Article

A Chaotic Cryptosystem for Images Based on Henon and Arnold Cat Map

Ali Soleymani,[1] Md Jan Nordin,[2] and Elankovan Sundararajan[1]

[1] *Software Technology and Management Center (Softam), Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia*
[2] *Center for Artificial Intelligence Technology (CAIT), Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia*

Correspondence should be addressed to Ali Soleymani; ali.soleymani@gmail.com

Received 30 April 2014; Accepted 15 July 2014; Published 28 August 2014

Academic Editor: Iftikhar Ahmad

The rapid evolution of imaging and communication technologies has transformed images into a widespread data type. Different types of data, such as personal medical information, official correspondence, or governmental and military documents, are saved and transmitted in the form of images over public networks. Hence, a fast and secure cryptosystem is needed for high-resolution images. In this paper, a novel encryption scheme is presented for securing images based on Arnold cat and Henon chaotic maps. The scheme uses Arnold cat map for bit- and pixel-level permutations on plain and secret images, while Henon map creates secret images and specific parameters for the permutations. Both the encryption and decryption processes are explained, formulated, and graphically presented. The results of security analysis of five different images demonstrate the strength of the proposed cryptosystem against statistical, brute force and differential attacks. The evaluated running time for both encryption and decryption processes guarantee that the cryptosystem can work effectively in real-time applications.

1. Introduction

Some researchers utilized conventional cryptosystems to directly encrypting images. But this is not advisable due to large data size and real-time constraints of image data. Conventional cryptosystems require a lot of time to directly encrypt thousands of image pixels value. On the other hand, unlike textual data, a decrypted image is usually acceptable even if it contains small levels of distortion. For all the above mentioned reasons, the algorithms that function well for textual data may not be suitable for multimedia data [1]. Many studies have been performed on the use of textual encryption algorithms for images by modifying the algorithms to adapt with image characteristics. One such option for encrypting an image is to consider a 2D array of image pixels value as a 1D data stream and to then encrypt this stream with any conventional cryptosystem [2, 3]. This would be considered a naïve approach and usually is suitable for text and occasionally for small images files that are to be transmitted over a fleet dedicated channel [4]. Subramanyan et al. [5] proposed an image encryption algorithm based on AES-128 in which the encryption process is a bitwise XOR operation on a set of image pixels. This method employs an initial 128-bit key and an AES key expansion process that changes the key for every set of pixels. The secret keys are generated independently at both the sender and the receiver sides based on the AES key expansion process. Therefore, the initial key alone is shared rather than the whole set of keys.

2. Chaos and Cryptography

Chaotic maps are simple functions and are iterated quickly. Chaos-based image encryption systems are therefore fast enough for real-time applications. Chaos is a natural phenomenon discovered by Edward Lorenz in 1963 while studying the butterfly effect in dynamical systems. The butterfly effect describes the sensitivity of a system to initial conditions as mentioned in Lorenz's paper titled "Does the Flap of a Butterfly's Wings in Brazil set off a Tornado in Texas?" [6].

The flapping wings represent a tiny variation in the initial conditions of the dynamic system that causes a chain of events leading to large-scale changes in the future. Had the butterfly not flapped its wings, the trajectory of the system might have been vastly different [7]. In general, this means that a small variance in the initial parameters (even in ten-millionth place value) could yield widely divergent results. Hence, for a chaotic system, rendering long-term prediction is impossible in general. This means that having initial conditions of these systems makes their future behavior predictable. This behavior, which derived from a natural phenomenon, is known as deterministic chaos or, simply, chaos and exhibits by chaotic maps. Such maps are classified as continuous maps and discrete maps.

In the 1990s, numerous researchers found that there are some relationships between properties that have counterparts in chaos and cryptography. A high sensitivity to initial conditions, with deterministic pseudorandom behavior, is an interesting similarity between chaotic maps and cryptographic algorithms. Furthermore, confusion and diffusion are two general principles in the design of cryptography algorithms that lead to the concealing of the statistical structure of pixels in a plain image and to a decrease in the statistical dependence of a plain image and the corresponding encrypted one. Applying a mixing property on chaos-based encryption algorithms will increase the complexity of the cipher image.

Chaotic maps are assigned to discrete and continuous time-domains. Discrete maps are usually in the form of iterated functions, which corresponded to rounds in cryptosystems. This similarity between cryptography and discrete chaotic dynamic systems is utilized to propose chaotic cryptosystems. Each map has some parameters that are equivalent to the encryption keys in cryptography. In stream cipher, a chaotic system is applied to generate a pseudorandom key stream but in block ciphers, the plaintext or the secret key(s) are used as the initial and control parameters. Finally, some iteration is applied on the chaotic systems to obtain the cipher-text. Security and complexity are significant concerns in cryptosystems. These should be considered when selecting a map and its parameters for use in cryptography [8].

2.1. Related Works. The first chaos-based cryptosystem was proposed by Matthews in 1989 [29]. Subsequently, the amount of research on chaotic cryptography increased rapidly, while trying to break (and find the weakness of) the proposed schemes in order to improve chaos-based cryptosystems.

The algorithm proposed by Wang et al. [30] for encrypting color images utilizing a logistic map was broken by Li et al. [31]. Another cryptosystem analyzed by Li et al. [32] is the recent work of Zhu [26]. Zhu applied hyperchaotic sequences to generate the key stream but Li in his work proved that the proposed algorithm was not sufficiently robust against a chosen plaintext attack. Another weak cryptosystem is the combination of the Lorenz map and perceptron model of the neural network proposed by Wang et al. [33]. This chaotic algorithm was cracked by Zhang et al. [34] after analyzing its security by simulated attacks. The experimental results show that the secret key can be reconstructed after one pair of known-plaintext/ciphertext attacks. Furthermore, the effect of changing one bit in the plain image is a change in only one bit at the same position in an encrypted image. This is another weakness of Wang's proposed algorithm.

Many similar works have failed in security analysis. Hence, when designing and implementing a chaos-based cryptographic system, some important requirements should be kept in mind. A common framework was proposed by Alvarez and Li [35] for chaos-based cryptosystem designers. Implementation rules, key management tips, and security analysis approaches are three main issues suggested in their work. Adhering to these basic guidelines guarantees an acceptable level of security with the chaos-based cryptosystem scheme. Moreover, Alvarez and Li in [36] established a practical security analysis of a cryptosystem based on the Baker map [37]. In addition to breaking this cryptosystem due to vulnerability of the key, some countermeasures are introduced for improving and enhancing the security of similar cryptosystems. Alvarez and Li in another cryptanalysis work [38] presented that the nonlinear chaotic algorithm by Gao et al. [39] is insecure according to failure in the plaintext attack and statistical and key space analysis.

Chaos-based encryption algorithms are based on diverse types of chaotic maps and also on discrete maps. Most of these are a combination of two or more chaotic maps to achieve a greater level of complexity, security, and expanded key space. A combination of the Arnold cat map and the Chen map was the work of Guan et al. [40]. The Arnold cat map was applied to clutter the position of the pixels followed by XOR with the discrete output signal of the Chen map to modify the gray value of the cluttered pixels. This was analyzed and improved by Xiao et al. [41]. They found the weakness of the proposed algorithm and overcame the flaws.

To overcome the disadvantages of permutation-only cryptosystems, Fu et al. [12] proposed a novel shuffling algorithm which performs an efficient bit-level permutation in two stages of chaotic sequence sorting and Arnold cat map. Their analysis results show that this scheme is more secure and has much lower computational complexity than previous similar works.

Xu et al. [10] analyzed the improved work of Xiang et al. [42] and found two drawbacks. In their proposed letter, iterating Chen chaotic system generates random number sequence, which is more random in comparison with the sequence that was generated by logistic map in [42]. The second drawback is overcome by setting the parameter of Chen map using the last one byte of encrypted plaintext after every iteration that leads to a higher sensitivity of encrypted image to the plain one. This scheme is fast and secure according to simulation results and large size of key space, respectively.

To overcome the drawback of time-consuming real number arithmetic calculations in chaos-based image encryption techniques, a block cipher cryptosystem was proposed by Fouda et al. [27]. This fast and secure chaotic scheme is based on sorting the integer coefficients of linear diophantine

equation (LDE), which is generated dynamically by only two rounds of any chaos map.

The scheme of Chen et al. [28] is another work that is proposed to enhance the efficiency of chaos-based encryption. They found that permutation-diffusion encryption approaches are produce with high computation of at least two chaotic maps and weak against known/chosen plaintext attacks. Hence, they proposed a dynamic mechanism to generate the state variables from the 3D or hyperchaotic maps for snake-like diffusion and pixel-swapping confusion. A tiny change (e.g., one pixel) will make a totally different key stream sequence at the first round of encryption.

Table 1 is a brief overview of some chaotic maps applied in image encryption. The Arnold cat map is the most commonly used map in chaos-based image encryption works with the main purpose of shuffling pixels of an image in a pseudorandom order.

2.2. Henon Map. Henon is a two-dimensional dynamic system proposed [43] to simplify the Lorenz map [44] with the same properties and is defined by (1). This might be easier to implement than the differential equations of the Lorenz system. Consider

$$\begin{aligned} x_{i+1} &= y_{i+1} + 1 - \alpha x_i^2, \\ y_{i+1} &= \beta x_i. \end{aligned} \tag{1}$$

The initial parameters are α, β and the initial point is (x_0, y_0). Each point (x_n, y_n) is mapped to a new point (x_{n+1}, y_{n+1}) through the Henon map. For $\alpha = 1.4$ and $\beta = 0.3$, the Henon function has chaotic behavior and the iterations have a boomerang-shaped chaotic attractor. Figure 1 is the outline on a two-dimensional plane for the Henon map obtained from a distinct number of iterations starting from the chosen initial point $(0.1, 0.1)$. Minute variations in the initial point will lead to major changes and different behavior.

2.3. Arnold Cat Map. ACM is a mixing discrete ergodic system that performs an area preserving stretch and fold mapping discovered by V. Arnold in 1968 using the image of a cat. This 2D transformation is based on a matrix with a determinant of 1 that makes this transformation reversible and described as

$$\Gamma : \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & P \\ Q & PQ+1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \bmod n. \tag{2}$$

Here, P and Q are integers and (x, y) is the original position that is mapped to the new position (x', y'). This transformation randomizes the original order of pixels or bits in an image. However, after sufficient iterations, the original image is reconstructed. Reverse mapping using (3) is a phase in decryption process to transform the shuffled image into the input image. The number of iterations in the permutation step must be equal to that of the reverse transformation. Consider

$$\Gamma' : \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} PQ+1 & -P \\ -Q & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \bmod n. \tag{3}$$

3. Proposed Cryptosystem Model

3.1. Initializing Prerequisite Values. In addition to α, β, and initial point (x_0, y_0) in (1), there are some other variables that must be initialized before running the algorithm. The proposed encryption architecture is shown in Figure 2. This scheme is based on two secret images and permutation steps in the bit level and pixel level. In the bitwise permutation, the pixel values are distorted but, in the pixel permutation, the pixels are shuffled without any alteration in value and histogram.

Creating the secret images and a set of parameters P and Q for the Arnold cat map are prerequisites for the encryption and decryption processes. Secret images have pseudo-random-like gray pixel distributions and are created using coordinates x and y generated by a Henon map. The secret images are the same as the plain image in height and width; therefore, the number of iterations for the Henon map depends on the total pixels in the plain image. In this work, experiments are performed on $m \times m$ gray-level images. Hence, the minimum iterations of Henon map should be m^2. The first few iterations seem fairly close together. Therefore, the total number of iterations is $m^2 + 100$, but the first 100 points are discarded to achieve higher randomness. Secret image pixels are generated using (4) and (6). The pix$X$ and pix$Y$ are sets of pseudorandom numbers ($0 \le \text{pix}X_i$, $\text{pix}Y_i \le 255$) created by x-coordinates and y-coordinates of the Henon map and are considered pixel values. To shape the one-dimensional pixel values into an image, (5) and (7) are applied to create the 2D *secImgX* and *secImgY* secret image. The final secret image is generated by a combination of *secImgX* and *secImgY* by performing the XOR operation on the corresponding pixels as described by (8). Consider

$$\text{pix}X_i = \text{abs}\left(\lfloor x_{100+i} \times \gamma \rfloor\right) \bmod 256, \quad i = 1, \ldots, m^2, \tag{4}$$

$$secImgX = \text{reshape}\left(\text{pix}X, m, m\right), \tag{5}$$

$$\text{pix}Y_j = \text{abs}\left(\lfloor x_{100+i} \times \lambda \rfloor\right) \bmod 256, \quad j = 1, \ldots, m^2, \tag{6}$$

$$secImgY = \text{reshape}\left(\text{pix}Y, m, m\right), \tag{7}$$

$$secImg = \text{xor}\left(secImgX, secImgY\right). \tag{8}$$

The permutation steps by the Arnold cat map are based on the parameters P and Q. The x-coordinates and y-coordinates that result from the iterations of the Henon map are applied to generate parameters for the ACM. The pixels or bits of an input image that are permuted by the Arnold cat map return to its preliminary position after finite iterations. Attackers may be able to restore the original image by using this periodicity. To avoid such reconstruction of the input image, iteration is repeated for q rounds with different values for the parameters P and Q in each round. Equations (9) and (10) generate parameter values for the Arnold cat map. The

Table 1: Applied chaos maps in some proposed image encryption techniques.

	ACM	Logistic	Henon	Lorenz	Baker	Chen	Tent	CML	Standard map
Zhu et al. [9]	×	×							
Xu et al. [10]						×			
Zhang and Cao [11]	×			×					
Fu et al. [12]	×								
Zhang et al. [13]	×	×	×						
Ghebleh et al. [14]	×						×		
Elshamy et al. [15]					×				
Ye and Zhou [16]							×	×	
Ye and Zhou [17]							×	×	
Wang et al. [18]					×				
Al-Maadeed et al. [19]		×							
Patidar et al. [20]		×							×
Wong et al. [21]									×
Guanghuia et al. [22]		×							
Zhang et al. [23]		×							
Liu et al. [24]	×	×							

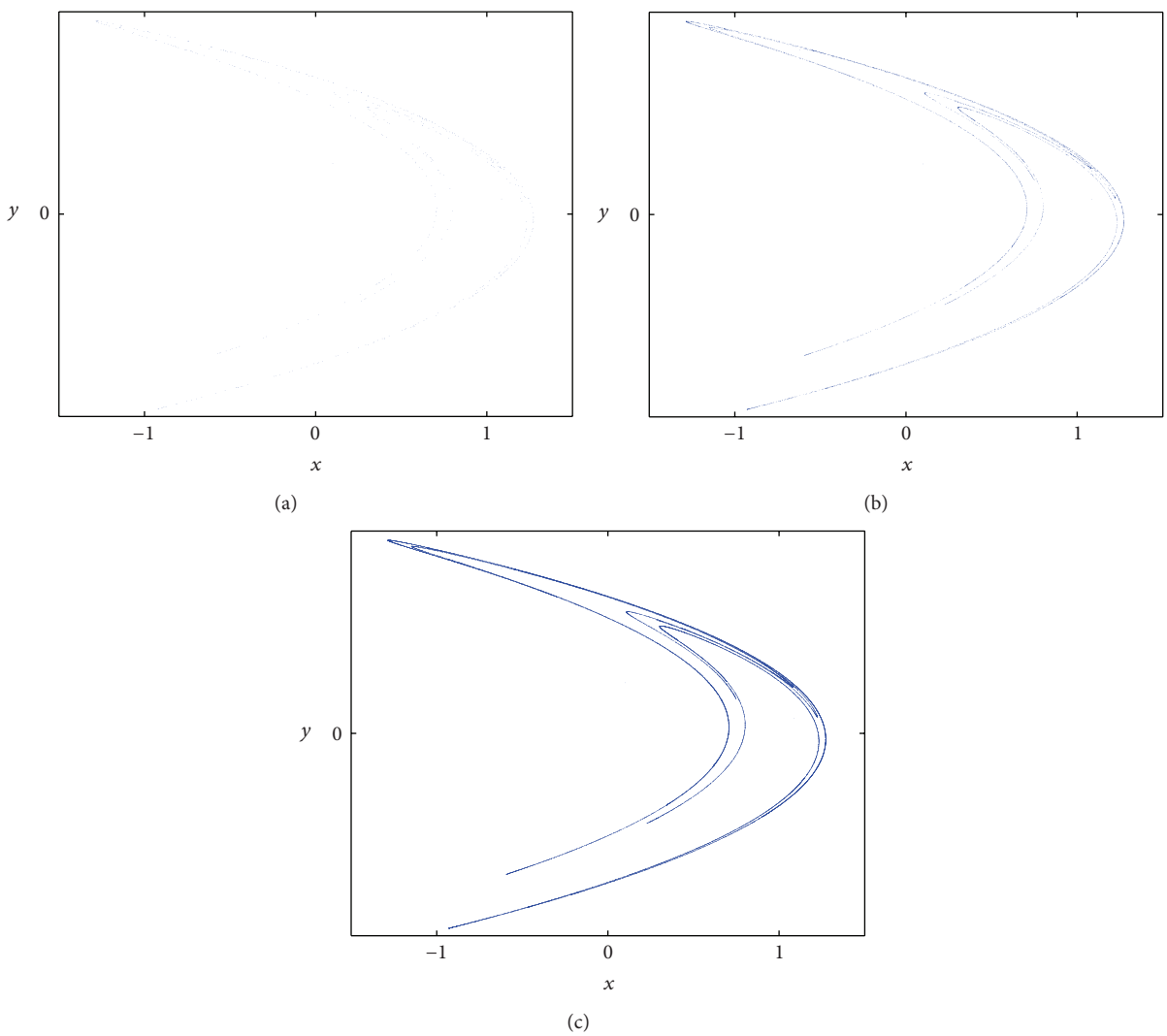


Figure 1: Henon map attractor after (a) 500, (b) 5000, and (c) 50000 iterations with initial parameters $\alpha = 1.4$ and $\beta = 0.3$ and initial point $(0.1, 0.1)$.

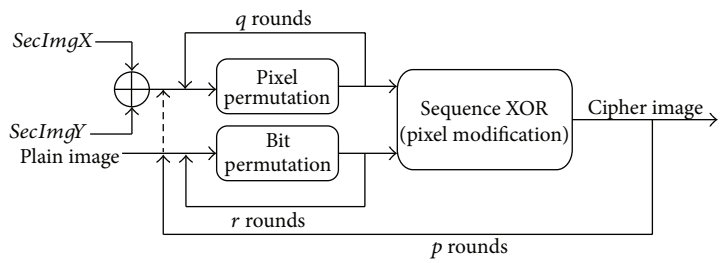


Figure 2: Proposed encryption scheme architecture.

number of generated parameters is equal to the total number of permutation rounds:

$$P_i = \text{abs}\left(\left\lfloor x_{100+i} \times 10^{14} \right\rfloor\right) \bmod \delta, \quad i = 1, \ldots, p * (q + r), \tag{9}$$

$$Q_i = \text{abs}\left(\left\lfloor y_{100+i} \times 10^{14} \right\rfloor\right) \bmod \vartheta, \quad i = 1, \ldots, p * (q + r). \tag{10}$$

3.2. Encryption Process. Figure 2 presents the architecture of the proposed encryption scheme. This scheme has three inputs and three main functions, and the final result is the encrypted image. The plain image, *secImgX,* and *secImgY* are the three main inputs for this model. The primary functions are bit permutation, pixel permutation, and pixel modification. As illustrated in Figure 2, at the first step, *secImgX* and *secImgY* are XORed pixel by pixel to generate *secImg*. Then, pixels of the secret image are permuted q rounds. A simultaneous step is r rounds of bit-level permutation of the plain image. The outputs of these two phases are applied to pixel modification, which is a sequence XOR of consecutive pixels. The result is fed back to the bit permutation function (instead of the plain image) for additional $p - 1$ rounds while the *secImg* is permuted with new parameters at each round. The functions details are described in the following sections.

3.2.1. Bit Permutation. For a gray-level image with a size of $m \times m$ pixels, the total bits are $m \times m \times 8$. Prior to bit permutation, the input image is divided into eight subimages. Each subimage is $m \times m/8$ pixels or $m \times m$ bits in height and width as shown in Figure 3. Matrix (11) shows how the kth subimage is created. Replacing the corresponding pixel values of the input image at the proper position of the matrix would create the subimage. A pixel in the position (i, j) is an 8-bit value in the form of (12), where $b(8)$ is the most significant bit (MSB) and $b(1)$ is the least significant bit (LSB) of the pixel value in binary form. Every pixel value of the subimage converts to its binary format and creates the bit-plane. The bit-plane is a matrix with m rows and m columns and each element is one bit 0 or 1. Matrix (13) shows how to create the matrix for the bit-plane. Each subimage is converted to the equivalent $m \times m$ bit-plane and each bit-plane is permuted separately and independently:

$$subImage_k = \begin{bmatrix} \text{img}\left(1, (k-1) * \frac{m}{8} + 1\right) & \cdots & \text{img}\left(1, k * \frac{m}{8}\right) \\ \text{img}\left(2, (k-1) * \frac{m}{8} + 1\right) & \cdots & \text{img}\left(2, k * \frac{m}{8}\right) \\ \vdots & & \\ \vdots & & \\ \text{img}\left(m, (k-1) * \frac{m}{8} + 1\right) & \cdots & \text{img}\left(m, k * \frac{m}{8}\right) \end{bmatrix}, \quad k = 1, \ldots, 8, \tag{11}$$

$$\text{img}(i, j) = b(8)\, b(7) \cdots b(1), \tag{12}$$

$$bitPlane_k = \begin{bmatrix} b_{1,(k-1)*(m/8+1)}(8) & b_{1,(k-1)*(m/8+1)}(7) & \cdots & b_{1,k*(m/8)}(1) \\ b_{2,(k-1)*(m/8+1)}(8) & b_{2,(k-1)*(m/8+1)}(7) & \cdots & b_{1,k*(m/8)}(1) \\ & \vdots & & \\ & \vdots & & \\ b_{m,(k-1)*(m/8+1)}(8) & b_{m,(k-1)*(m/8+1)}(7) & \cdots & b_{m,k*(m/8)}(1) \end{bmatrix}. \tag{13}$$

After creating the *bitPlane* matrices, the Arnold cat map was applied to these matrices to permute the bits. In the permutation phase, the new location of each bit is calculated by (14). The pair of (x', y') is the new position of (x, y). At the

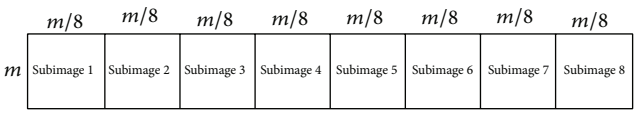


Figure 3: Splitting an image to 8 subimages.

first phase, the input image is permuted r times with different parameters P and Q where $i = 1, \ldots, r$. Consider

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & P_i \\ Q_i & P_i Q_i + 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \bmod m. \tag{14}$$

After finishing this phase, the *bitPlane* matrices are changed to decimal values to reconstruct the image pixels.

3.2.2. Pixel Permutation. Concurrent with bit permutation of the plain image, the secret image is permuted for q rounds at the pixel level to change the position of the pixels in a random manner. In contrast to bit permutation, pixel permutation does not affect the pixel values. Therefore, histograms of the plain image and the shuffled image are entirely the same. The permutation is performed q times with different parameters P and Q with $j = 1, \ldots, q$. Consider

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & P_j \\ Q_j & P_j Q_j + 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \bmod m. \tag{15}$$

3.2.3. Pixel Modification. The concluding phase is a sequence XOR to modify the pixel values. This step will cause extreme changes in the pixels of cipher image with even one bit change in a pixel in plain image. This step is based on the shuffled secret image and the bit-level permuted plain image. Equations (16) and (17) are used to change pixels consecutively. The output of this step is a modified image. For more confusion and modification, this step will repeat for p rounds. After each round, the result is replaced with a plain image, as input, and the secret image will permute q rounds. After completing p rounds, the final output is the cipher image. Consider

$$\mathrm{img}(1) = \mathrm{xor}\left(\mathrm{img}(1), \mathrm{secImg}(1)\right), \tag{16}$$

$$\mathrm{img}(i) = \mathrm{xor}\left(\mathrm{img}(i-1), \mathrm{img}(i), \mathrm{secImg}(i)\right). \tag{17}$$

3.3. Decryption Process. Figure 4 shows the decryption process. On the receiving side, the secret image is generated by XORing *secImgX* and *secImgY,* which are recreated using private parameters. Inverse pixel modification is performed on the cipher image and the secret image after $p \times q$ rounds of pixel permutations. The result of the secret image pixel permutation in the first step is saved for subsequent steps and at each round this is inverse permuted for q rounds and applied as an input to the inverse pixel modification function. The additional input is (the feedback of) the output of the inverse pixel modification function after r rounds of inverse bit permutation.

4. Experiments and Security Analysis

Experiments are carried out to analyze the proposed algorithm and evaluate their security and robustness. A vigorous encryption algorithm is resistant to attacks by an opponent or to unauthorized access. Due to the various types of attacks, a comprehensive security analysis is inevitable. This section analyzes the results of simulated attacks such as a statistical attack, a differential attack, a brute-force attack, and a known/chosen plaintext attack to demonstrate the strength of the proposed technique. Key space, encryption time, and decryption time are additional parameters that will affect the decision-making regarding the choice of applied cryptosystem. The selected images for the experiments are "Peppers," "Baboon," and "Fingerprint," which are 512 × 512 gray images. "Cameraman" and "Chess-plate" are two additional images of size 256 × 256. The proposed algorithm is implemented using the MATLAB programming language on a PC with a 64-bit OS, Core i5 CPU, and 8 GB installed RAM.

4.1. Initial Values. The Henon map is the main function in this cryptosystem and the generated chaotic sequence is employed to produce both the secret image and the parameters for the Arnold cat map. Its initial value is one of the secret keys in this scheme. The chosen point (1.210000001, 0.360000001) is the starting point for generating the Henon chaotic sequence. α and β have fixed values of 1.4 and 0.3, respectively. As mentioned above, the test images sizes are 512 × 512 and 256 × 256 pixels. The number of iterations for the Henon map is based on the image size. The larger image has a total of 262144 pixels. The Henon map (18) is iterated 262244 times, but the first 100 points are discarded. Consider

$$\begin{aligned} x_{i+1} &= y_{i+1} + 1 - 1.4x_i^2 \\ y_{i+1} &= 0.3x_i \\ &\left(x_0 = 1.21000001, y_0 = 0.36000001\right), \quad i = 0, \ldots, 262243. \end{aligned} \tag{18}$$

Having obtained the set of x and y, we can create pixels of *secImgX* and *secImgY* by setting the values $\gamma = 12345678$ and $\lambda = 87654321$ in (4) and (6) as shown in (19) and (21), respectively. Then, we reshape them to form an image with the same size as the plain image using (20) and (22). The final secret image is the result of (23). Consider

$$\mathrm{pix}X_i = \mathrm{abs}\left(\lfloor x_{100+i} \times 12345678 \rfloor\right) \bmod 256, \quad i = 1, \ldots, 262144, \tag{19}$$

$$secImgX = \mathrm{reshape}\left(\mathrm{pix}X, 512, 512\right), \tag{20}$$

$$\mathrm{pix}Y_j = \mathrm{abs}\left(\lfloor x_{100+i} \times 87654321 \rfloor\right) \bmod 256, \quad j = 1, \ldots, 262144, \tag{21}$$

$$secImgY = \mathrm{reshape}\left(\mathrm{pix}Y, 512, 512\right), \tag{22}$$

$$secImg = \mathrm{xor}\left(secImgX, secImgY\right). \tag{23}$$

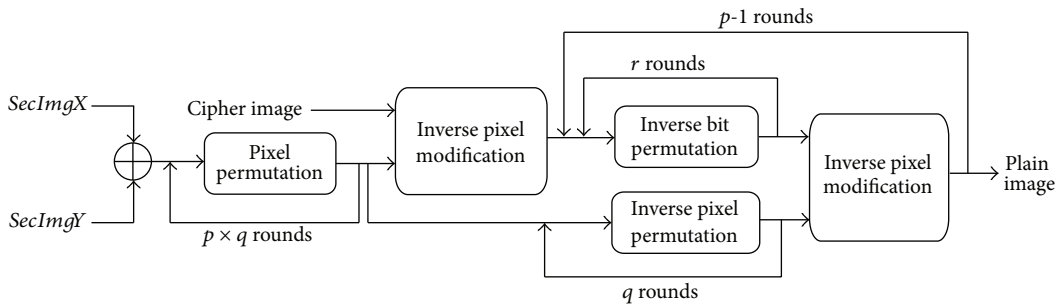


Figure 4: Decryption model architecture.

Table 2: Evaluated running time for main functions in proposed cryptosystem.

Variable	Function	Running time for one round (ms)
T_{BP}	Bit permutation running time	62
T_{PP}	Pixel permutation running time	13
T_{PM}	Pixel modification running time	4
T_{IBP}	Inverse bit permutation running time	75
T_{IPP}	Inverse pixel permutation running time	23
T_{IPM}	Inverse pixel modification running time	4

Table 3: Comparison of encryption time for proposed algorithm with recent similar works.

	Proposed scheme	[12]	[10]	[25]	[26]	[27]
Encryption time (ms)	19.75	23	22	20.79	32	52

The parameters for the Arnold cat map that perform permutations on the plain image and secret images are the set of P and Q. Each pair of P and Q is a modified value of a point on the Henon sequence. These coordinates are real numbers. They converted to integer numbers using multiplied, modular, and absolute operations as described in (8) and (9). By setting quantities $\delta = 12345$ and $\vartheta = 67890$, P and Q are in the form of

$$\begin{aligned} P_i &= \operatorname{abs}\left(\left\lfloor x_{100+i} \times 10^{14} \right\rfloor\right) \bmod 12345, \\ &\qquad i = 1, \ldots, p(q+r), \\ Q_i &= \operatorname{abs}\left(\left\lfloor y_{100+i} \times 10^{14} \right\rfloor\right) \bmod 67890, \\ &\qquad i = 1, \ldots, p(q+r). \end{aligned} \tag{24}$$

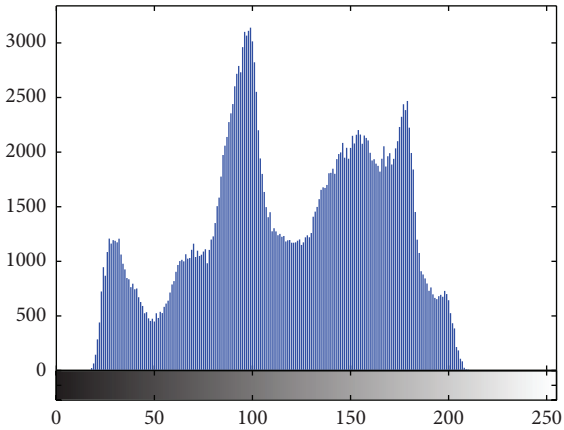
4.2. Running Time. Pixel permutation, bit permutation, and pixel modification are three key functions in this cryptosystem. The encryption time depends on the run time of each function and the number of rounds. Table 2 presents the average run time for one round of each function in this cryptosystem on a 512 × 512 image. Additional tests show a linear relation between running time and number of pixels. The results for 256 × 256 were almost a quarter of the given intervals in Table 2. The total encryption time is calculated by (25) and the decryption time is calculated by (26). Table 3 is the encryption time of proposed cryptosystem in comparison with recent similar works for a 256×256 gray image. Consider

$$T_E = p * \left(r * T_{BP} + q * T_{PP} + T_{PM}\right), \tag{25}$$

$$\begin{aligned} T_D &= p * q * T_{PP} + T_{IPM} + (p-1) \\ &\quad \times \left(r * T_{IBP} + q * T_{IPP} + T_{IPM}\right). \end{aligned} \tag{26}$$

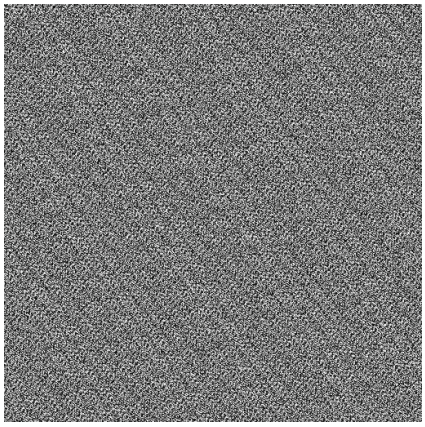
4.3. Encryption and Decryption Illustrations and Histogram. The image histogram is the graphical illustration of the pixel distribution at different gray levels. A great deal of statistical information regarding the image is extractable from its histogram [45]. The histogram of an encrypted image should have a uniform distribution and be completely different from that of the plain image. This prevents the leakage of any meaningful information from the plain image. In addition to the histogram, the encrypted image must be absolutely unique in appearance without any similar pattern to the original image. Basically, the histogram analysis results are demonstrated for a bit-permuted plain image based on the proposed technique to compare with the proposed scheme

(a)

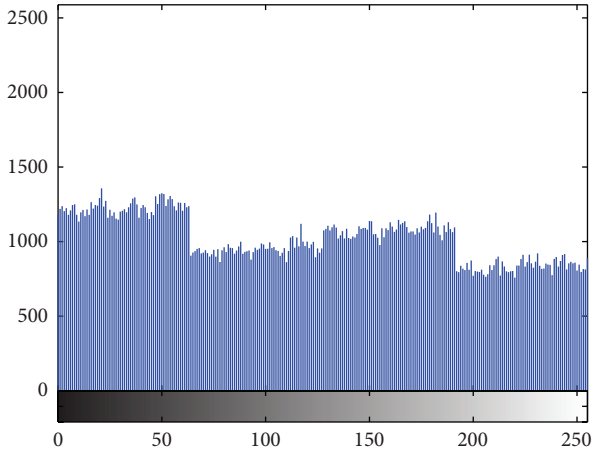


(b)

Figure 5: (a) Plain image and (b) its histogram.



(a)



(b)

Figure 6: (a) Plain image after four rounds of bit permutation based on Zhu's algorithm and (b) its histogram.

by Zhu et al. [9]. Despite the pixel value alteration by Zhu's model, the histogram of the permuted image is not uniform. This is due to the permutation of a group of bits in the same position. According to Shannon's theory, the 8th bit (MSB) pixel values carry almost 50% of the total information of the image. Permuting all of the bits at the same location will not vastly change the image pixels' value. Hence, the plain image, which is shown in Figure 5, is bit permuted for four rounds, and the result and its histogram are illustrated in Figure 6. It seems that there are patterns that still appear in the image. To overcome this vulnerability, a bit permutation scheme is proposed and shuffles the bits entirely and independently of its position. Figure 7 shows the subimages of the plain image that would be permuted independently. After four rounds of permuting each subimage with altered parameters, the result and its histogram are shown in Figure 7. Visual comparison of images and histograms in Figures 6 and 7 demonstrates the efficacy of the proposed bit-permutation model. Figures 8 and 9 are the images and the histograms after the encryption and decryption process, respectively. The decrypted image is the same as the original one and this technique is found to be lossless.

4.4. Key Analysis

4.4.1. Key Space Analysis. The total number of possible keys that an attacker must try to break a cryptosystem is called key space and it should be large enough to prevent brute-force attack. In the proposed cryptosystem, the initial point (x_0, y_0) of the Henon map was used as one of the secret keys. Other control parameters are γ, λ, δ, ϑ, p, q, and r. These parameters should be kept secret and be used as secret keys. Table 4 is the upper bound for each variable. A combination of these parameters will provide a large key space of approximately 2^{300} that is sufficient to make brute-force attack infeasible and very large rather than similar works that are compared in Table 5.

4.4.2. Key Sensitivity Analysis. In addition to a sufficiently large key space to protect an encrypted image from brute-force attacks, a strength algorithm should also be absolutely sensitive to both encryption and decryption keys. Changing even one bit in a secret key will cause a completely different result in either the encrypted image or the decrypted image. Key sensitivity is analyzed in both the encryption

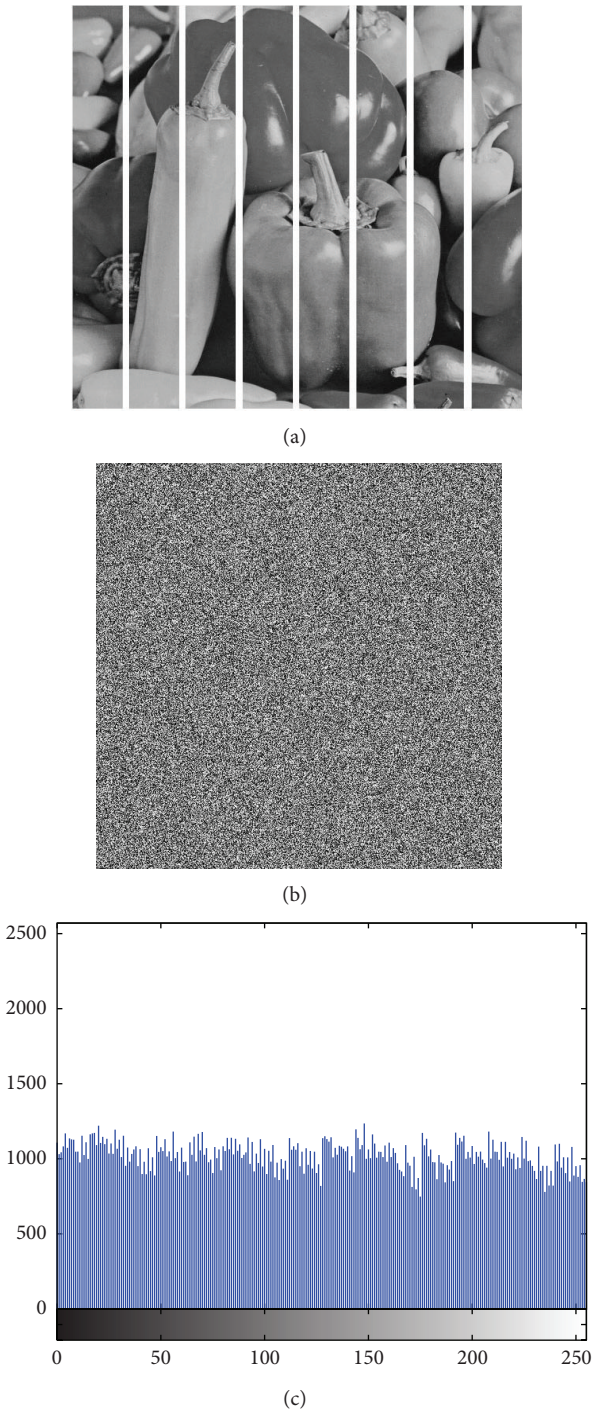


Figure 7: (a) Plain subimages, (b) plain image after four rounds of proposed bit permutation and (c) its histogram.

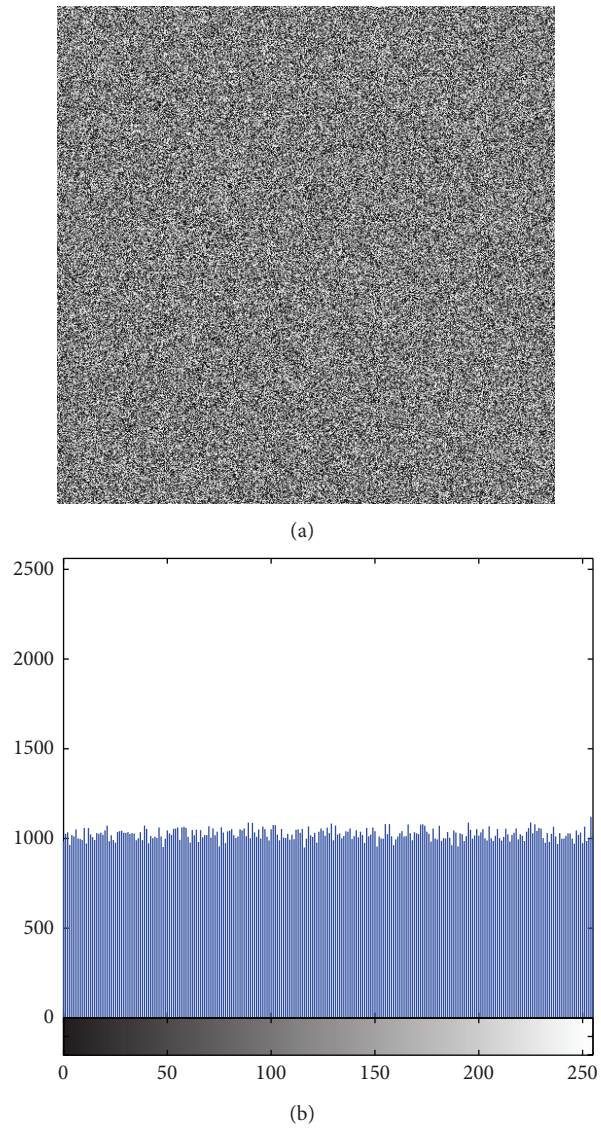


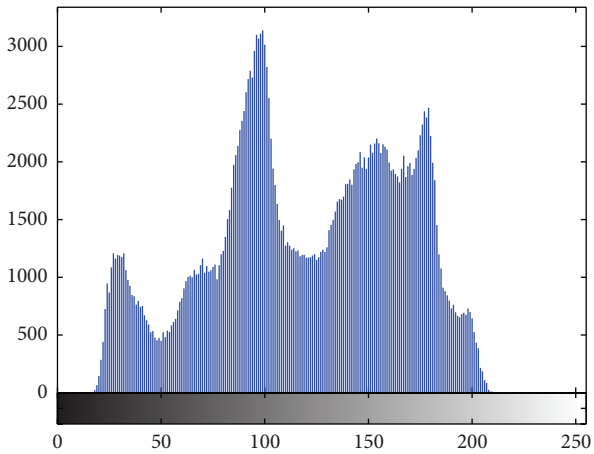
Figure 8: (a) Encrypted image and (b) its histogram.

and the decryption phase. In the encryption phase, the cipher image that results from changing even one bit in any one of the initial values is compared with the encrypted image that resulted before changing the key. The results are given in Table 6. Several experiments were performed and, in each experiment, only one parameter was manipulated while others were unchanged. The changed values and the difference rates for the produced images are listed in the table.

In the decryption phase, key sensitivity means that the encrypted image cannot be decrypted by slight variations in the secret key. Based on the results in Table 7, changing even one bit in the decryption key will result in a wholly different decrypted image.

4.5. Statistical Analysis. Statistical analysis can extract the relationships between the original and the encrypted image. Shannon in his theory of information and communication [45] proved that it is possible to break many types of cryptograms by statistical analysis. This can be thwarted by dissipating the redundancy in the structure of the message by diffusion or by increasing the complexity of the relationship between the encrypted message and the secret key by

(a)



(b)

Figure 9: (a) Decrypted image and (b) its histogram.

confusion. Either confusion or diffusion is presented in the proposed cryptosystem to frustrate statistical attacks.

4.5.1. Correlation Analysis. Two adjoining pixels in a regular image are strongly correlated in horizontal, vertical, and diagonal positions. Scatter plots in Figures 10 and 11 reveal the correlation of two adjacent pixels in horizontal, vertical, and diagonal distributions in the plain and the cipher image, respectively. Correlation coefficients are calculated for test images by (27) and the results for plain images and cipher images are listed in Table 14. For an ordinary image, the correlation coefficients are very close to 1, which is the highest possible value. The produced encrypted image is ideal and resists statistical attack if the correlation coefficients are very low and close to 0. Consider

$$r_{x,y} = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}}. \quad (27)$$

4.5.2. Entropy Analysis. Entropy is a statistical parameter that is defined to measure the uncertainly and randomness of a bundle of data. According to Shannon theory, image entropy

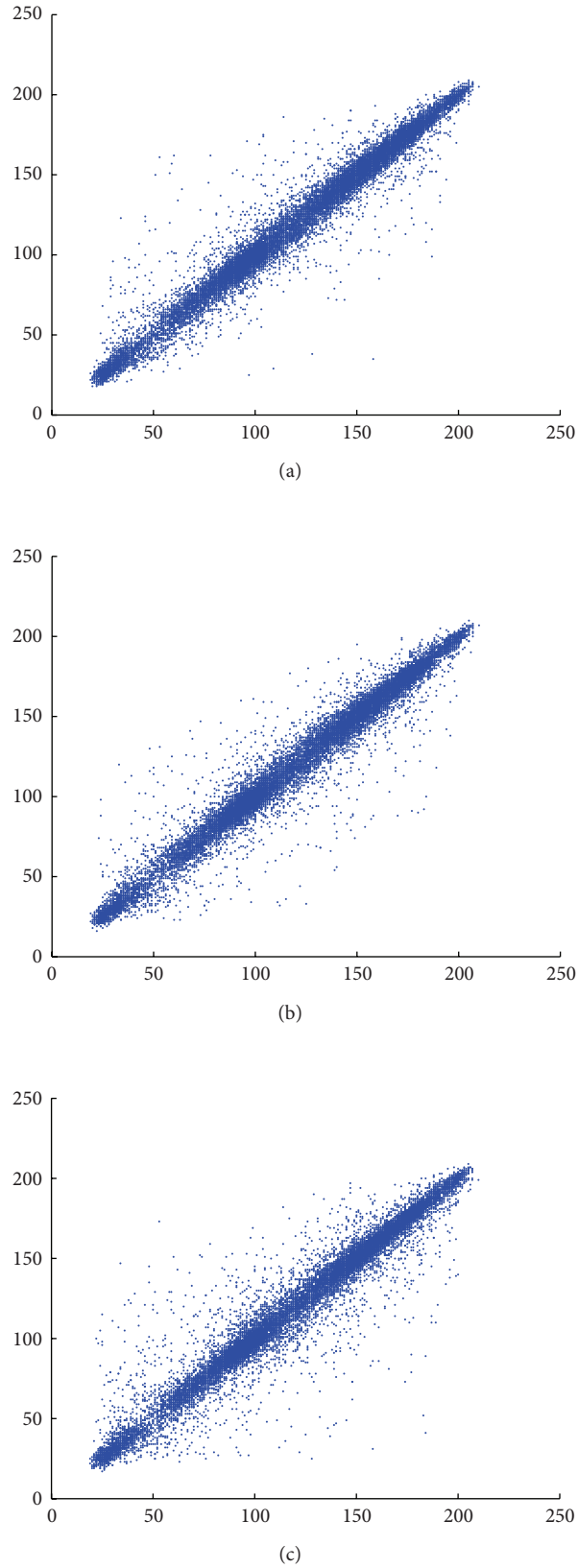


Figure 10: Correlation of plain image's pixels in (a) horizontal, (b) vertical, and (c) diagonal position.

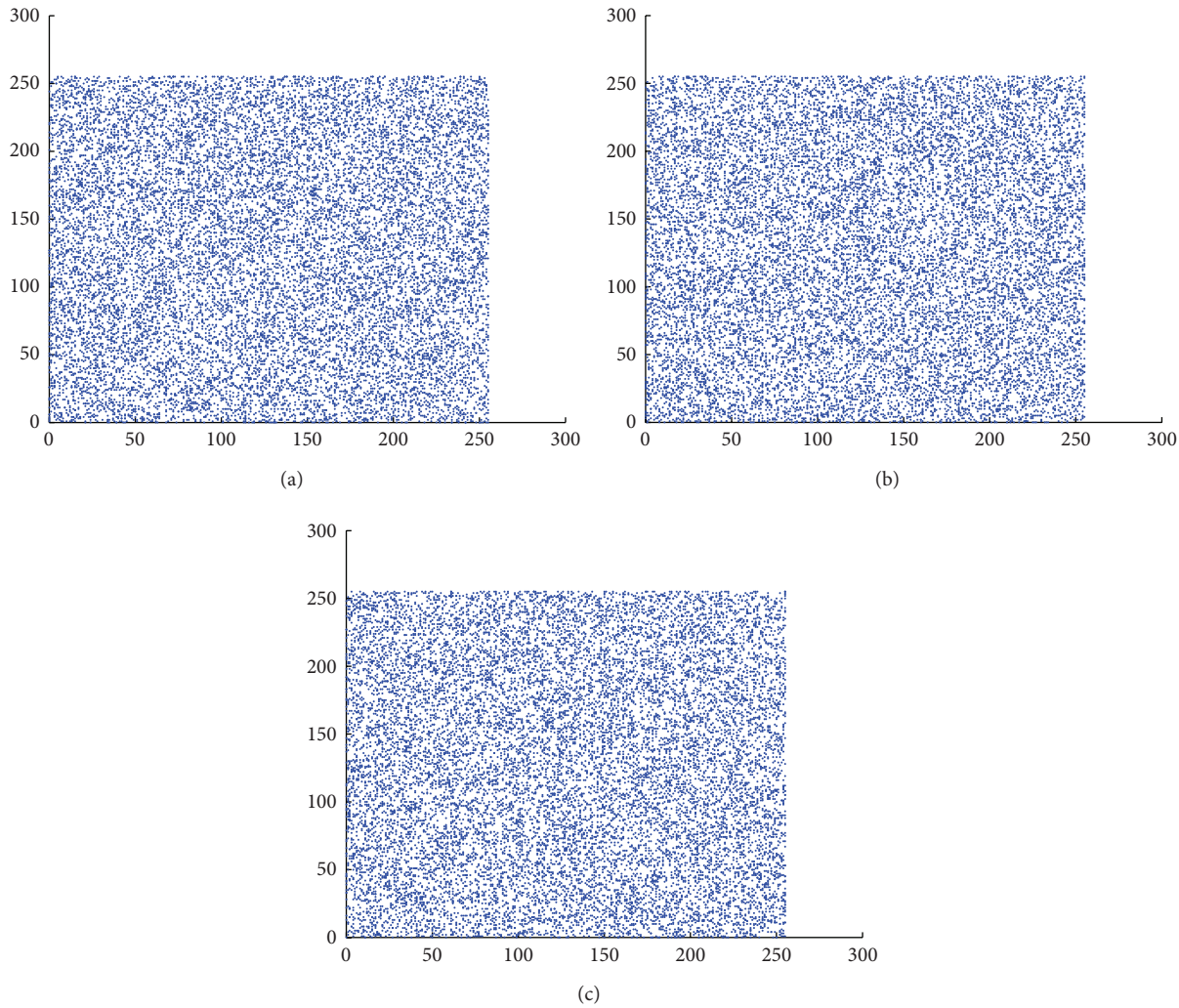


Figure 11: Correlation of cipher image's pixels in (a) horizontal, (b) vertical, and (c) diagonal position.

Table 4: Secret parameters length in bit.

Parameter	Length (bit)
p	10
q	10
r	10
γ	48
λ	48
δ	24
ϑ	24
x_0	64
y_0	64

is the number of bits that is necessary to encode every pixel of the image. The optimal value for entropy of an encrypted image is ~8. This quantity describes the random pattern and texture of pixels in an encrypted image and is calculated by

$$\text{entropy} = \sum_{i=0}^{n} P_i \log_2 P_i, \tag{28}$$

where n is the total number of gray levels (i.e., 256) and P_i is the probability of incidence of intensity i in the current image. P_i is the number of pixels with intensity i divided by the total number of pixels. The base-2 logarithm will present the calculated entropy in bits. The entropy values for plain images and encrypted images are given in Table 14.

Table 5: Comparison of encryption time for proposed algorithm with recent similar works.

	Proposed scheme	[12]	[10]	[25]	[26]	[28]	[27]
Key space	2^{300}	2^{153}	2^{120}	2^{128}	2^{186}	2^{199}	2^{256}

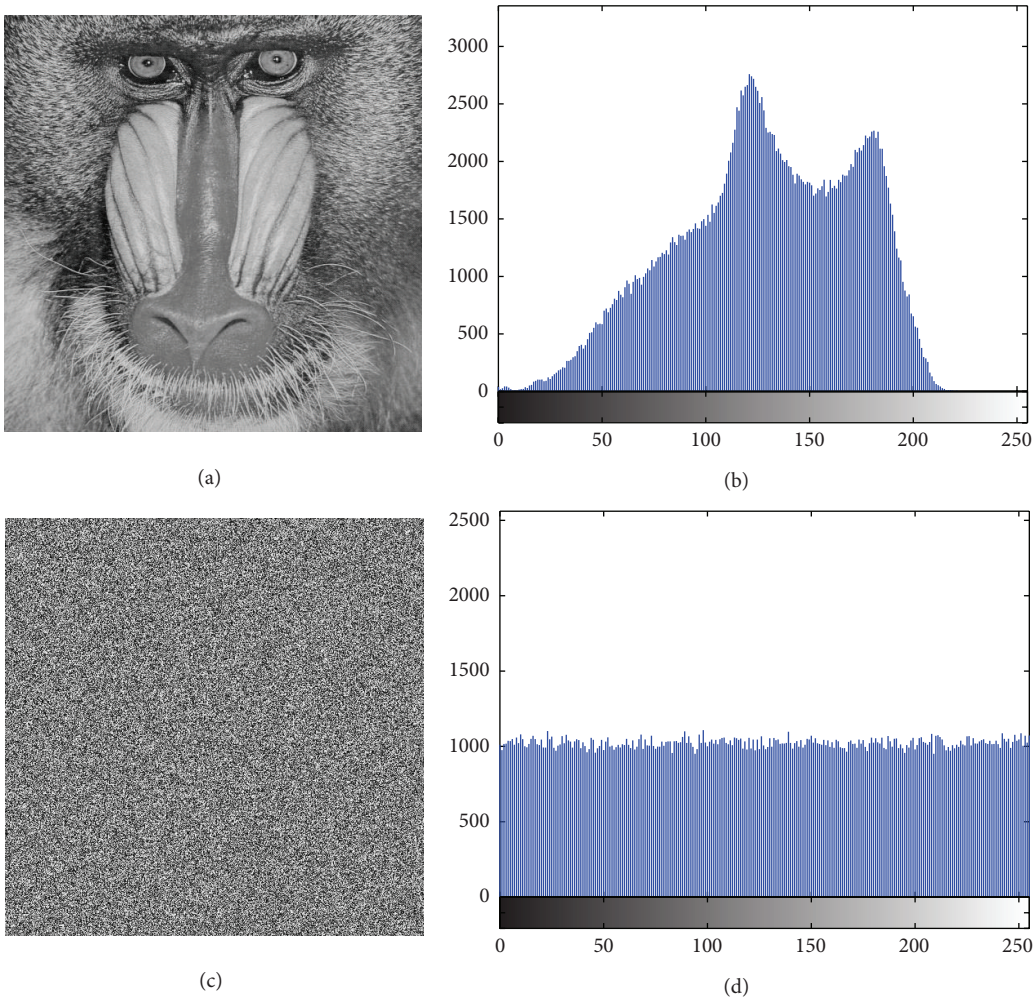


Figure 12: (a) Baboon image and (b) its histograms and (c) encrypted image and (d) its histogram.

Table 6: Difference rates of two encrypted images with slight change in a parameter.

Parameter	Initial value	Changed value	Encrypted images difference rate
p	6	7	99.59%
q	2	1	99.60%
r	2	1	99.62%
Γ	12345678	12345679	99.58%
λ	87654321	87654320	99.62%
δ	12345	12346	99.63%
ϑ	67890	67891	99.60%
x_0	1.21000001	1.21	99.59%
y_0	0.36000001	0.36	99.60%

4.6. Differential Analysis. For the purpose of differential attack, an attacker changes a specific pixel in the plain image and traces the differences in the analogous encrypted image to find a meaningful relation. This is also known as a chosen-plaintext attack. A robust encrypted image must be sensitive to minor changes and even changing one bit in the plain image should result in a wide range of changes in the cipher image.

The NPCR measures the number of pixels change rate in an encrypted image when 1 bit is changed in the plain

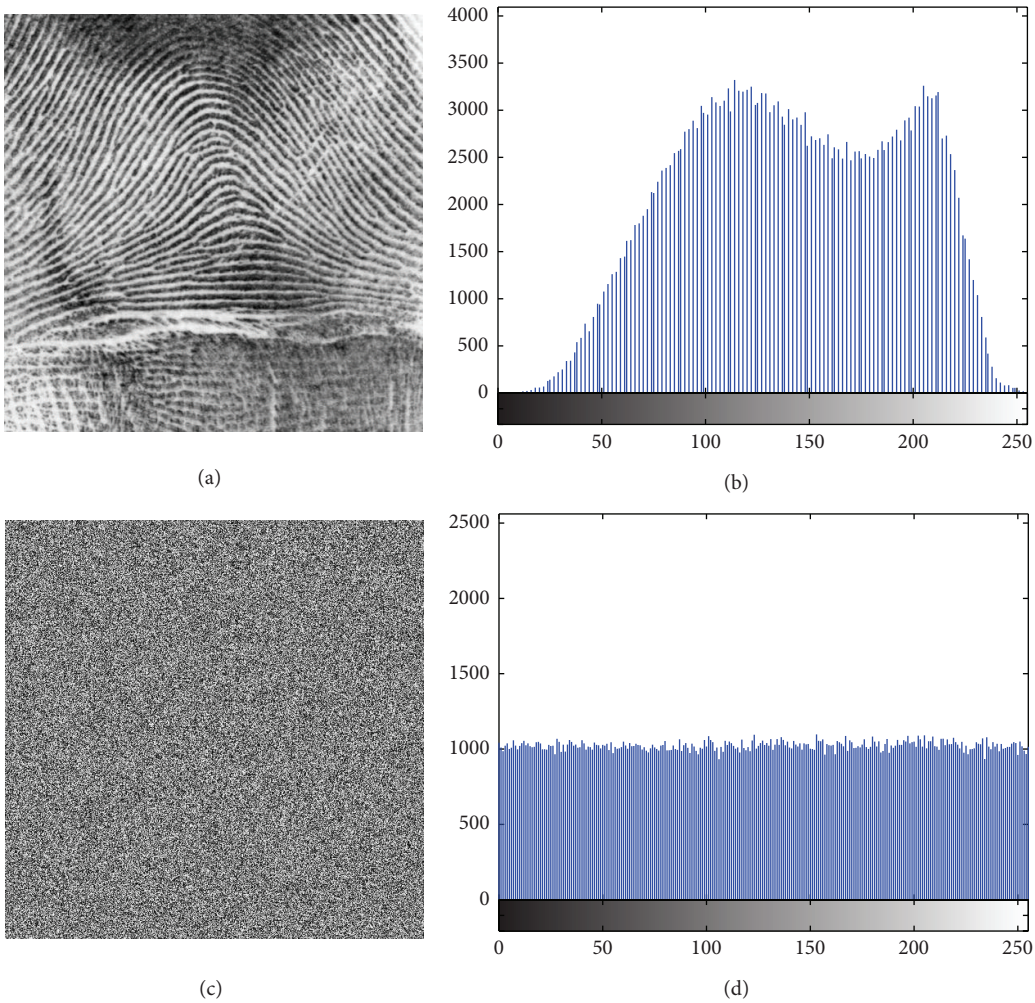


Figure 13: (a) Fingerprint image and (b) its histograms and (c) encrypted image and (d) its histogram.

image. This parameter is calculated by (29) and for an ideal encryption algorithm it is 1. Consider

$$\mathrm{NPCR} = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} f(i,j) \times 100,$$
$$f(i,j) = \begin{cases} 0, & \text{if } c_1(i,j) = c_2(i,j), \\ 1, & \text{if } c_1(i,j) \neq c_2(i,j), \end{cases} \tag{29}$$

where c_1 and c_2 are obtained by encrypting two $m \times n$ plain images and one random bit dissimilarity.

The UACI in differential analysis is the unified average changing intensity between two encrypted images with a difference in only one bit in corresponding plain images. The UACI can be calculated by (30):

$$\mathrm{UACI} = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{|c_1(i,j) - c_2(i,j)|}{255} \times 100. \tag{30}$$

To evaluate the sensitivity of the proposed algorithm to differential attacks, a random bit is changed in the plain image. Encrypting two plain images with a difference in only one bit produces two encrypted images. The rates of pixel and intensity differences in the two encrypted images are calculated. Tables 8, 9, and 10 present calculated UACI and NPCR values for different combinations of p, q, and r to trade off the encryption speed and the overall rounds to find a threshold that achieves the highest rate. The following tables are related to the Peppers image. From Tables 8 and 9, it was concluded that increasing the value of the parameter q that is related to the pixel permutation rounds does not affect the NPCR and UACI values. These values depend only on p and r. However, to increase the level of confusion and increase the key space, pixel permutation is required. The results in Table 10 were used to find the minimum values for p and r that result in the ideal value for NPCR and UACI with the smallest run-time. It was concluded that at least three rounds of p were required to obtain the highest values for UACI and

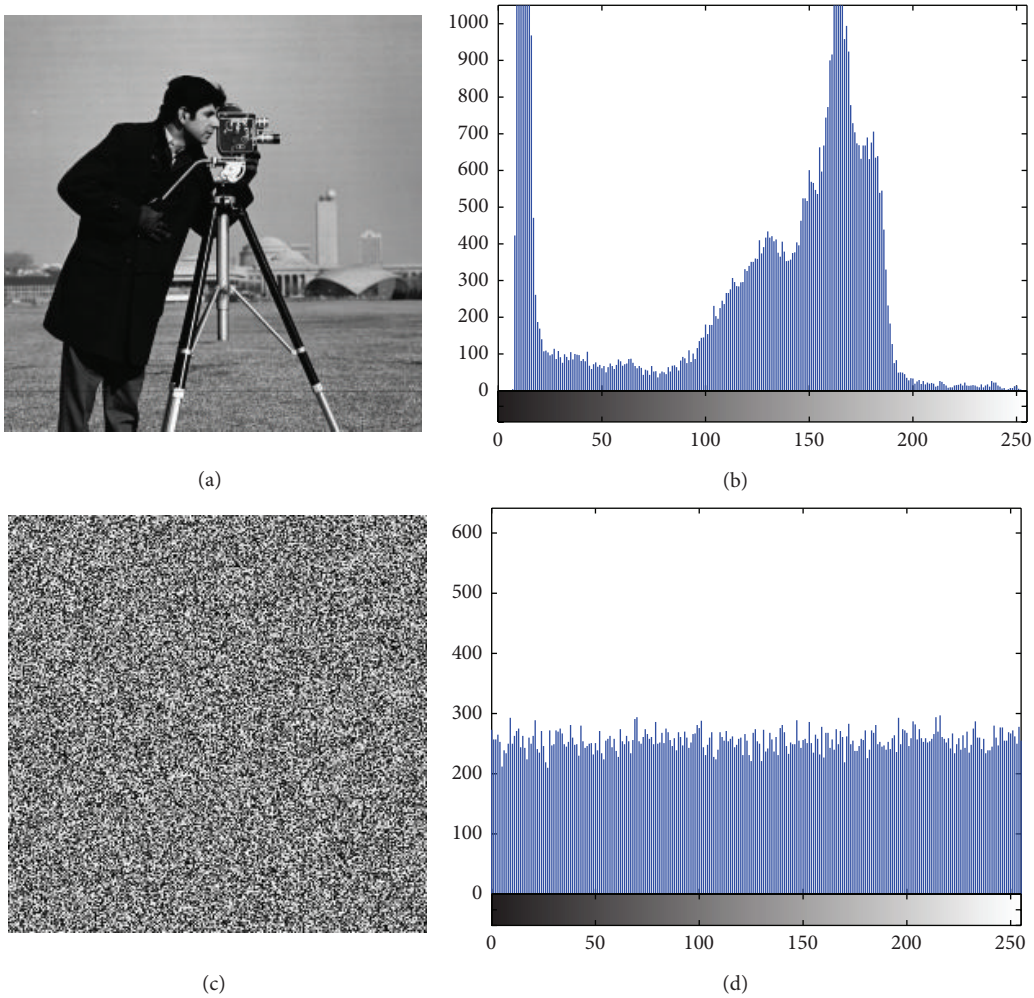


Figure 14: (a) Cameraman image and (b) its histograms and (c) encrypted image and (d) its histogram.

Table 7: Difference rate of two decrypted images with slight change in a parameter.

Parameter	Encryption parameters	Decryption parameters	Decrypted images difference rate
p	6	5	99.59%
q	2	1	99.62%
r	2	1	99.59%
γ	12345678	12345677	99.62%
λ	87654321	87654322	99.58%
δ	12345	12344	99.59%
ϑ	67890	67889	99.60%
x_0	1.21000001	1.21000002	99.60%
y_0	0.36000001	0.36000002	99.62%

NPCR. After the 3rd row, all of the combinations are ideal. Different combinations of p, q, and r are calculated with (25) and the combination of (p, q, r) that encrypts the image with the smallest run-time is (3, 1, 1).

In addition to Peppers, the same experiments were performed on Baboon, Figure 12, Fingerprint, Figure 13, Cameraman, Figure 14, and Chess-plate, Figure 15. The results of the experiments on these three images are investigated to determine the strength of the proposed cryptosystem. Tables 11, 12, 13, and 14 list the calculated values of UACI and NPCR for different combinations of p and r in Baboon, Fingerprint, Cameraman, and Chess-plate images, respectively. Because the results were similar to other combinations of the Pepper image, the other similar tables were discarded and only

Table 8: Calculated UACI and NPCR for different combinations of q and r while $p = 1$ in Peppers.

q/r		1	2	3	4	5	6	7	8	9	10
1	UACI	5.1163	2.7810	2.0331	1.1860	1.5810	6.1305	1.4349	3.7728	3.3215	0.0564
	NPCR	81.5414	44.3218	32.4032	18.9026	25.1965	97.7051	22.8695	60.1284	52.9366	0.8984
2	UACI	5.1163	2.7810	2.0331	1.1860	1.5810	6.1305	1.4349	3.7728	3.3215	0.0564
	NPCR	81.5414	44.3218	32.4032	18.9026	25.1965	97.7051	22.8695	60.1284	52.9366	0.8984
3	UACI	5.1163	2.7810	2.0331	1.1860	1.5810	6.1305	1.4349	3.7728	3.3215	0.0564
	NPCR	81.5414	44.3218	32.4032	18.9026	25.1965	97.7051	22.8695	60.1284	52.9366	0.8984
4	UACI	5.1163	2.7810	2.0331	1.1860	1.5810	6.1305	1.4349	3.7728	3.3215	0.0564
	NPCR	81.5414	44.3218	32.4032	18.9026	25.1965	97.7051	22.8695	60.1284	52.9366	0.8984
5	UACI	5.1163	2.7810	2.0331	1.1860	1.5810	6.1305	1.4349	3.7728	3.3215	0.0564
	NPCR	81.5414	44.3218	32.4032	18.9026	25.1965	97.7051	22.8695	60.1284	52.9366	0.8984
6	UACI	5.1163	2.7810	2.0331	1.1860	1.5810	6.1305	1.4349	3.7728	3.3215	0.0564
	NPCR	81.5414	44.3218	32.4032	18.9026	25.1965	97.7051	22.8695	60.1284	52.9366	0.8984
7	UACI	5.1163	2.7810	2.0331	1.1860	1.5810	6.1305	1.4349	3.7728	3.3215	0.0564
	NPCR	81.5414	44.3218	32.4032	18.9026	25.1965	97.7051	22.8695	60.1284	52.9366	0.8984
8	UACI	5.1163	2.7810	2.0331	1.1860	1.5810	6.1305	1.4349	3.7728	3.3215	0.0564
	NPCR	81.5414	44.3218	32.4032	18.9026	25.1965	97.7051	22.8695	60.1284	52.9366	0.8984
9	UACI	5.1163	2.7810	2.0331	1.1860	1.5810	6.1305	1.4349	3.7728	3.3215	0.0564
	NPCR	81.5414	44.3218	32.4032	18.9026	25.1965	97.7051	22.8695	60.1284	52.9366	0.8984
10	UACI	5.1163	2.7810	2.0331	1.1860	1.5810	6.1305	1.4349	3.7728	3.3215	0.0564
	NPCR	81.5414	44.3218	32.4032	18.9026	25.1965	97.7051	22.8695	60.1284	52.9366	0.8984

Table 9: Calculated UACI and NPCR for different combinations of q and p while $r = 1$ in Peppers.

q/p		1	2	3	4	5	6	7	8	9	10
1	UACI	5.1163	16.2147	**33.4573**	**33.4727**	**33.4075**	**33.4355**	**33.4484**	**33.5372**	**33.4775**	**33.4735**
	NPCR	81.5414	89.1617	**99.6342**	**99.6208**	**99.6059**	**99.6231**	**99.5930**	**99.6277**	**99.5930**	**99.6174**
2	UACI	5.1163	16.2173	**33.4317**	**33.4902**	**33.4462**	**33.4644**	**33.5382**	**33.4616**	**33.5023**	**33.4503**
	NPCR	81.5414	89.1617	**99.6342**	**99.6208**	**99.6059**	**99.6231**	**99.5930**	**99.6277**	**99.5930**	**99.6174**
3	UACI	5.1163	16.2191	**33.3994**	**33.5234**	**33.4093**	**33.4373**	**33.4497**	**33.4928**	**33.5363**	**33.5198**
	NPCR	81.5414	89.1617	**99.6342**	**99.6208**	**99.6059**	**99.6231**	**99.5930**	**99.6277**	**99.5930**	**99.6174**
4	UACI	5.1163	16.2003	**33.4263**	**33.4481**	**33.3943**	**33.4181**	**33.5062**	**33.4770**	**33.4784**	**33.5147**
	NPCR	81.5414	89.1617	**99.6342**	**99.6208**	**99.6059**	**99.6231**	**99.5930**	**99.6277**	**99.5930**	**99.6174**
5	UACI	5.1163	16.1772	**33.4520**	**33.4995**	**33.3855**	**33.4539**	**33.4610**	**33.5291**	**33.4940**	**33.4577**
	NPCR	81.5414	89.1617	**99.6342**	**99.6208**	**99.6059**	**99.6231**	**99.5930**	**99.6277**	**99.5930**	**99.6174**
6	UACI	5.1163	16.1845	**33.4262**	**33.5164**	**33.3721**	**33.4232**	**33.4606**	**33.5161**	**33.4626**	**33.4400**
	NPCR	81.5414	89.1617	**99.6342**	**99.6208**	**99.6059**	**99.6231**	**99.5930**	**99.6277**	**99.5930**	**99.6174**
7	UACI	5.1163	16.1885	**33.4141**	**33.4955**	**33.4241**	**33.4615**	**33.4464**	**33.5113**	**33.4486**	**33.4434**
	NPCR	81.5414	89.1617	**99.6342**	**99.6208**	**99.6059**	**99.6231**	**99.5930**	**99.6277**	**99.5930**	**99.6174**
8	UACI	5.1163	16.1882	**33.4694**	**33.4696**	**33.3941**	**33.4834**	**33.5019**	**33.4556**	**33.4589**	**33.5000**
	NPCR	81.5414	89.1617	**99.6342**	**99.6208**	**99.6059**	**99.6231**	**99.5930**	**99.6277**	**99.5930**	**99.6174**
9	UACI	5.1163	16.1954	**33.4277**	**33.4527**	**33.3872**	**33.5124**	**33.4510**	**33.4750**	**33.4249**	**33.5083**
	NPCR	81.5414	89.1617	**99.6342**	**99.6208**	**99.6059**	**99.6231**	**99.5930**	**99.6277**	**99.5930**	**99.6174**
10	UACI	5.1163	16.2147	**33.4573**	**33.4727**	**33.4075**	**33.4355**	**33.4484**	**33.5372**	**33.4775**	**33.4735**
	NPCR	81.5414	89.1617	**99.6342**	**99.6208**	**99.6059**	**99.6231**	**99.5930**	**99.6277**	**99.5930**	**99.6174**

Table 10: Calculated UACI and NPCR for different combinations of p and r while $q = 1$ in Peppers.

p/r		1	2	3	4	5	6	7	8	9	10
1	UACI	5.1163	2.7810	2.0331	1.1860	1.5810	6.1305	1.4349	3.7728	3.3215	0.0564
	NPCR	81.5414	44.3218	32.4032	18.9026	25.1965	97.7051	22.8695	60.1284	52.9366	0.8984
2	UACI	16.2119	2.0258	0.9736	0.4904	0.8663	32.6779	0.4922	3.8314	4.1271	0.1998
	NPCR	89.1617	85.5148	81.7390	75.0000	76.3233	94.8944	75.3403	86.6261	89.8193	50.9499
3	UACI	**33.3879**	**32.8273**	**34.7519**	**32.7804**	**33.5263**	**33.5459**	**33.3332**	**34.4429**	**33.5705**	**33.2778**
	NPCR	**99.6342**	**99.5850**	**99.6552**	**99.5987**	**99.6357**	**99.6017**	**99.6098**	**99.6044**	**99.6033**	**99.5075**
4	UACI	**33.4882**	**33.3894**	**33.4144**	**33.4373**	**33.5037**	**33.5170**	**33.4163**	**33.4185**	**33.4494**	**33.5380**
	NPCR	**99.6208**	**99.6124**	**99.5979**	**99.5918**	**99.6075**	**99.5892**	**99.6185**	**99.6075**	**99.5960**	**99.5995**
5	UACI	**33.3994**	**33.3844**	**33.5083**	**33.4886**	**33.4846**	**33.5354**	**33.4447**	**33.5059**	**33.3940**	**33.4656**
	NPCR	**99.6059**	**99.6124**	**99.6258**	**99.6159**	**99.6033**	**99.6067**	**99.6101**	**99.6166**	**99.6120**	**99.5953**
6	UACI	**33.4314**	**33.4144**	**33.4623**	**33.3797**	**33.4872**	**33.5663**	**33.4599**	**33.4731**	**33.4311**	**33.4141**
	NPCR	**99.6231**	**99.6296**	**99.6250**	**99.6006**	**99.6208**	**99.6265**	**99.6037**	**99.6040**	**99.5926**	**99.6120**
7	UACI	**33.4744**	**33.4282**	**33.4541**	**33.5446**	**33.4563**	**33.4700**	**33.4543**	**33.4240**	**33.4748**	**33.4333**
	NPCR	**99.5930**	**99.5968**	**99.6128**	**99.6017**	**99.6082**	**99.6021**	**99.6311**	**99.6132**	**99.5964**	**99.6185**
8	UACI	**33.4535**	**33.3790**	**33.5941**	**33.5118**	**33.5437**	**33.4958**	**33.5057**	**33.5276**	**33.4622**	**33.4553**
	NPCR	**99.6277**	**99.5972**	**99.5758**	**99.6044**	**99.6094**	**99.5983**	**99.6170**	**99.6136**	**99.5888**	**99.5983**
9	UACI	**33.4703**	**33.4465**	**33.4612**	**33.5285**	**33.5804**	**33.4783**	**33.4213**	**33.4463**	**33.4063**	**33.4409**
	NPCR	**99.5930**	**99.5975**	**99.6006**	**99.6094**	**99.6185**	**99.6262**	**99.5903**	**99.6052**	**99.6082**	**99.6071**
10	UACI	**33.4631**	**33.4266**	**33.3962**	**33.4113**	**33.4824**	**33.5212**	**33.5256**	**33.4979**	**33.3891**	**33.4482**
	NPCR	**99.6174**	**99.6109**	**99.6178**	**99.6105**	**99.6212**	**99.6071**	**99.5857**	**99.6014**	**99.6071**	**99.5960**

Table 11: Calculated UACI and NPCR for different combinations of p and r while $q = 1$ in Baboon.

p/r		1	2	3	4	5	6	7	8	9	10
1	UACI	34.5764	18.5622	1.9700	39.7193	43.0302	12.3125	4.4068	40.1488	48.6447	20.7776
	NPCR	68.8828	36.9793	3.9246	79.1283	85.7243	24.5289	8.7791	79.9839	96.9093	41.3929
2	UACI	7.9196	0.9566	0.1960	15.6506	16.3404	0.5035	0.1959	14.2109	32.8500	1.7190
	NPCR	87.5687	81.3938	49.9817	93.7302	93.2835	76.1761	49.9420	86.6169	93.3601	67.2352
3	UACI	**33.4245**	**32.9577**	**32.9665**	**33.6173**	**33.4770**	**33.5504**	**33.4176**	**32.8382**	**33.4229**	**32.5148**
	NPCR	**99.6040**	**99.6105**	**99.6094**	**99.6319**	**99.6090**	**99.6128**	**99.6162**	**99.6017**	**99.5922**	**99.5224**
4	UACI	**33.5433**	**33.4850**	**33.4328**	**33.4770**	**33.4972**	**33.4388**	**33.4096**	**33.4726**	**33.4217**	**33.4147**
	NPCR	**99.6140**	**99.6307**	**99.6166**	**99.6136**	**99.6037**	**99.6071**	**99.6059**	**99.5964**	**99.6056**	**99.6235**
5	UACI	**33.4336**	**33.4809**	**33.4824**	**33.4264**	**33.4568**	**33.5143**	**33.4470**	**33.3774**	**33.4297**	**33.3378**
	NPCR	**99.6151**	**99.5892**	**99.6128**	**99.6189**	**99.6273**	**99.5922**	**99.6407**	**99.6044**	**99.5819**	**99.6120**
6	UACI	**33.4377**	**33.4569**	**33.4803**	**33.4672**	**33.5039**	**33.4647**	**33.4143**	**33.5158**	**33.4548**	**33.4117**
	NPCR	**99.6021**	**99.5987**	**99.6178**	**99.5975**	**99.6174**	**99.6155**	**99.6090**	**99.6067**	**99.6140**	**99.6143**
7	UACI	**33.4281**	**33.4751**	**33.4826**	**33.4711**	**33.4626**	**33.4911**	**33.4976**	**33.4263**	**33.4891**	**33.5239**
	NPCR	**99.6052**	**99.6120**	**99.5922**	**99.5892**	**99.5998**	**99.6334**	**99.6040**	**99.6052**	**99.6258**	**99.6120**
8	UACI	**33.4534**	**33.5000**	**33.4770**	**33.4986**	**33.4616**	**33.4514**	**33.4376**	**33.4205**	**33.4672**	**33.5179**
	NPCR	**99.6265**	**99.5892**	**99.5953**	**99.6082**	**99.6101**	**99.6059**	**99.6014**	**99.6334**	**99.6258**	**99.6128**
9	UACI	**33.4524**	**33.5607**	**33.4740**	**33.5257**	**33.3783**	**33.5049**	**33.4948**	**33.4504**	**33.4197**	**33.4726**
	NPCR	**99.6223**	**99.6204**	**99.6002**	**99.5934**	**99.6124**	**99.5991**	**99.6025**	**99.5911**	**99.5953**	**99.6094**
10	UACI	**33.3259**	**33.4977**	**33.3791**	**33.5318**	**33.5857**	**33.4664**	**33.4249**	**33.4084**	**33.5136**	**33.5256**
	NPCR	**99.6155**	**99.6353**	**99.6536**	**99.6338**	**99.5850**	**99.5987**	**99.6536**	**99.6063**	**99.6155**	**99.6353**

Table 12: Calculated UACI and NPCR for different combinations of p and r while $q = 1$ in Fingerprint.

p/r		1	2	3	4	5	6	7	8	9	10
1	UACI	0.1102	0.2833	0.2429	0.3697	0.3291	0.4672	0.6168	0.1744	0.0714	0.1102
	NPCR	14.0533	36.1149	30.9715	47.1390	41.9651	59.5711	78.6449	22.2366	9.0992	14.0533
2	UACI	0.4539	1.0997	1.0502	2.0179	2.0101	4.0957	16.3335	0.4905	0.1923	0.4539
	NPCR	71.9505	85.4733	86.7638	85.9703	84.3555	88.2622	89.8018	75.0031	49.0311	71.9505
3	UACI	**32.7711**	**33.2829**	**23.0805**	**33.4200**	**33.3053**	**33.4104**	**33.9360**	**33.4891**	**34.1633**	**32.7711**
	NPCR	**99.6140**	**99.6277**	**99.3843**	**99.6178**	**99.5754**	**99.6014**	**99.6368**	**99.6098**	**99.7284**	**99.6140**
4	UACI	**33.5087**	**33.4496**	**33.4820**	**33.5291**	**33.3559**	**33.4629**	**33.5104**	**33.4655**	**33.4892**	**33.5087**
	NPCR	**99.6105**	**99.6220**	**99.6216**	**99.6178**	**99.6067**	**99.6181**	**99.6120**	**99.5991**	**99.6105**	**99.6105**
5	UACI	**33.5129**	**33.4573**	**33.4886**	**33.4295**	**33.5528**	**33.5534**	**33.4449**	**33.4749**	**33.5176**	**33.5129**
	NPCR	**99.6117**	**99.6181**	**99.5953**	**99.6269**	**99.5880**	**99.6021**	**99.6094**	**99.6166**	**99.5991**	**99.6117**
6	UACI	**33.4461**	**33.4562**	**33.5025**	**33.5266**	**33.4495**	**33.4262**	**33.4276**	**33.4117**	**33.4122**	**33.4461**
	NPCR	**99.5983**	**99.6094**	**99.6044**	**99.5956**	**99.5930**	**99.6132**	**99.6391**	**99.6204**	**99.5926**	**99.5983**
7	UACI	**33.3924**	**33.5282**	**33.4733**	**33.4923**	**33.5048**	**33.4279**	**33.4765**	**33.5338**	**33.4224**	**33.3924**
	NPCR	**99.6265**	**99.6075**	**99.6014**	**99.6273**	**99.6212**	**99.6075**	**99.5922**	**99.6101**	**99.6155**	**99.6265**
8	UACI	**33.4546**	**33.4874**	**33.5073**	**33.4015**	**33.4098**	**33.4831**	**33.3897**	**33.4127**	**33.5300**	**33.4546**
	NPCR	**99.6063**	**99.6025**	**99.6132**	**99.6159**	**99.6147**	**99.6292**	**99.5827**	**99.6315**	**99.5949**	**99.6063**
9	UACI	**33.5202**	**33.4528**	**33.5220**	**33.4781**	**33.4966**	**33.4986**	**33.4642**	**33.4896**	**33.5301**	**33.5202**
	NPCR	**99.6223**	**99.6201**	**99.5995**	**99.6304**	**99.5895**	**99.6002**	**99.6078**	**99.5880**	**99.5781**	**99.6223**
10	UACI	**33.3561**	**33.4246**	**33.4130**	**33.4110**	**33.3838**	**33.4363**	**33.5498**	**33.4906**	**33.4433**	**33.3561**
	NPCR	**99.6170**	**99.5872**	**99.5777**	**99.5884**	**99.6170**	**99.6048**	**99.5926**	**99.6140**	**99.5987**	**99.6170**

Table 13: Calculated UACI and NPCR for different combinations of p and r while $q = 1$ for Cameraman.

p/r		1	2	3	4	5	6	7	8	9	10
1	UACI	7.4904	2.6620	5.1333	11.4982	3.7695	5.6794	5.3902	11.9543	8.4184	7.4904
	NPCR	29.8447	10.6064	20.4529	45.8130	15.0192	22.6288	21.4767	47.6303	33.5419	29.8447
2	UACI	1.0214	0.1943	0.4910	1.9258	0.3592	0.5052	0.4911	1.8747	1.0055	1.0214
	NPCR	83.5632	49.5514	75.0793	86.7004	63.8062	76.2024	75.0595	84.4025	81.1432	83.5632
3	UACI	**33.7662**	**29.9298**	**32.1345**	**32.5679**	**33.5877**	**33.5277**	**33.6352**	**34.8690**	**33.6583**	**33.7662**
	NPCR	**99.6262**	**99.5071**	**99.5102**	**99.5285**	**99.5224**	**99.5850**	**99.6384**	**99.6552**	**99.6521**	**99.6262**
4	UACI	**33.4563**	**33.3303**	**33.4678**	**33.5979**	**33.4952**	**33.5201**	**33.4910**	**33.2905**	**33.3876**	**33.4563**
	NPCR	**99.6292**	**99.5453**	**99.5972**	**99.6582**	**99.5972**	**99.6170**	**99.5682**	**99.5743**	**99.6475**	**99.6292**
5	UACI	**33.2851**	**33.4706**	**33.3841**	**33.5766**	**33.3711**	**33.4932**	**33.3911**	**33.4998**	**33.4473**	**33.2851**
	NPCR	**99.6155**	**99.6262**	**99.5834**	**99.5972**	**99.6109**	**99.5956**	**99.5956**	**99.6368**	**99.6201**	**99.6155**
6	UACI	**33.4854**	**33.3612**	**33.6393**	**33.4269**	**33.3583**	**33.6404**	**33.4790**	**33.3590**	**33.2436**	**33.4854**
	NPCR	**99.6567**	**99.5911**	**99.5682**	**99.6078**	**99.6277**	**99.5880**	**99.6277**	**99.6124**	**99.6414**	**99.6567**
7	UACI	**33.4159**	**33.4013**	**33.5464**	**33.5039**	**33.4404**	**33.4768**	**33.4303**	**33.3951**	**33.2653**	**33.4159**
	NPCR	**99.6674**	**99.5850**	**99.6094**	**99.6262**	**99.5880**	**99.6140**	**99.5941**	**99.5850**	**99.6002**	**99.6674**
8	UACI	**33.4321**	**33.4036**	**33.6195**	**33.4753**	**33.3993**	**33.5394**	**33.4112**	**33.5314**	**33.4387**	**33.4321**
	NPCR	**99.5621**	**99.5926**	**99.6368**	**99.6277**	**99.5895**	**99.5941**	**99.5819**	**99.5987**	**99.6094**	**99.5621**
9	UACI	**33.4651**	**33.3438**	**33.4917**	**33.4837**	**33.4551**	**33.4780**	**33.4704**	**33.4806**	**33.4516**	**33.4651**
	NPCR	**99.6140**	**99.6323**	**99.5636**	**99.6216**	**99.6429**	**99.6109**	**99.6338**	**99.5758**	**99.6002**	**99.6140**
10	UACI	**33.6242**	**33.5821**	**33.2618**	**33.4312**	**33.5802**	**33.6133**	**33.3934**	**33.6385**	**33.5515**	**33.6242**
	NPCR	**99.5972**	**99.5895**	**99.6078**	**99.5773**	**99.6033**	**99.5911**	**99.6323**	**99.6109**	**99.5850**	**99.5972**

Table 14: Calculated UACI and NPCR for different combinations of p and r while $q = 1$ for Chess-plate.

p/r		1	2	3	4	5	6	7	8	9	10
1	UACI	49.7511	38.9683	0.6828	0.3194	17.8328	9.1234	1.8446	0.6682	24.8476	45.3508
	NPCR	99.1135	77.6321	2.7206	81.4575	71.0526	72.7020	14.6988	85.1974	99.0021	90.3473
2	UACI	30.8111	26.3993	34.1747	33.5065	33.3622	33.4134	33.0488	33.5821	33.5865	33.5562
	NPCR	72.5723	90.6128	99.3164	99.3210	99.5193	99.5987	99.4949	99.6674	99.5697	99.6399
3	UACI	**33.6698**	**33.2530**	**33.4078**	**33.4257**	**33.5033**	**33.4428**	**33.4594**	**33.5224**	**33.3517**	**33.4282**
	NPCR	**99.5483**	**99.6460**	**99.5850**	**99.5682**	**99.6063**	**99.6033**	**99.5819**	**99.6338**	**99.5972**	**99.6002**
4	UACI	**33.4568**	**33.4550**	**33.4542**	**33.4903**	**33.5600**	**33.3137**	**33.4928**	**33.5251**	**33.4225**	**33.3196**
	NPCR	**99.6185**	**99.6201**	**99.6262**	**99.5819**	**99.5850**	**99.5972**	**99.5895**	**99.6216**	**99.5956**	**99.6063**
5	UACI	**33.4745**	**33.5885**	**33.5112**	**33.5035**	**33.4013**	**33.4363**	**33.4302**	**33.5542**	**33.4494**	**33.4572**
	NPCR	**99.6155**	**99.6414**	**99.6292**	**99.6262**	**99.5850**	**99.5956**	**99.6063**	**99.6262**	**99.6109**	**99.6002**
6	UACI	**33.4729**	**33.4588**	**33.3505**	**33.5044**	**33.5310**	**33.5108**	**33.3838**	**33.6189**	**33.3733**	**33.4620**
	NPCR	**99.6277**	**99.5728**	**99.6292**	**99.5987**	**99.6048**	**99.6689**	**99.5865**	**99.6017**	**99.5941**	**99.6109**
7	UACI	**33.4701**	**33.4639**	**33.3825**	**33.5426**	**33.4590**	**33.4672**	**33.3788**	**33.4522**	**33.4046**	**33.4853**
	NPCR	**99.5926**	**99.6246**	**99.5758**	**99.6368**	**99.5941**	**99.6078**	**99.6201**	**99.6399**	**99.6048**	**99.5682**
8	UACI	**33.4681**	**33.4529**	**33.3931**	**33.3834**	**33.6822**	**33.4259**	**33.3729**	**33.3856**	**33.4834**	**33.4206**
	NPCR	**99.5926**	**99.6811**	**99.6063**	**99.5941**	**99.5972**	**99.6002**	**99.6643**	**99.6368**	**99.5987**	**99.6155**
9	UACI	**33.3928**	**33.5164**	**33.5825**	**33.4455**	**33.5033**	**33.4134**	**33.4258**	**33.5132**	**33.5198**	**33.6166**
	NPCR	**99.5956**	**99.6231**	**99.6078**	**99.6216**	**99.6262**	**99.6094**	**99.6033**	**99.6094**	**99.6216**	**99.6292**
10	UACI	**33.2732**	**33.5652**	**33.5405**	**33.6319**	**33.3404**	**33.4587**	**33.3828**	**33.4207**	**33.4494**	**33.5986**
	NPCR	**99.5911**	**99.5895**	**99.6185**	**99.6002**	**99.6078**	**99.6063**	**99.6231**	**99.6216**	**99.5590**	**99.6536**

Table 15: Results of security analysis.

Image name	p	q	r	Plain entropy	Cipher entropy	Plain image correlations			Cipher image correlations			UACI	NPCR
						HC	VC	DC	HC	VC	DC		
Cameraman 256 × 256	1	1	1	7.0097	7.9969	0.8390	0.7189	0.6973	0.0003	0.0012	0.0013	7.4904	29.8447
	3	1	1		7.9976				0.0057	−0.0049	0.0027	33.7662	99.6262
	1	3	1		7.9971				0.0013	0.0035	−0.0030	7.4904	29.8447
	1	1	3		7.9972				0.0011	0.0011	−0.0042	5.1333	20.4529
Chess-plate 256 × 256	1	1	1	1	7.9970	0.9775	0.9800	0.9637	−0.0096	−0.0056	0.0056	49.7511	99.1135
	3	1	1		7.9974				0.0193	−0.0231	0.0048	33.6698	99.5483
	1	3	1		7.9972				−0.0010	0.0102	0.0111	49.7511	99.1135
	1	1	3		7.9973				0.0123	−0.0053	0.0206	0.6828	2.7206
Baboon 512 × 512	1	1	1	7.3579	7.9993	0.8644	0.7587	0.7261	−0.0038	0.0033	0.0015	34.5764	68.8828
	3	1	1		7.9993				0.0015	−0.0004	0.0009	33.4245	99.6040
	1	3	1		7.9993				0.0012	−0.0004	−0.0007	34.5764	68.8828
	1	1	3		7.9993				0.0016	0.0018	−0.0024	1.9700	3.9246
Peppers 512 × 512	1	1	1	7.5714	7.9993	0.8642	0.7587	0.7261	−0.0030	0.0018	−0.0017	0.1867	47.6059
	3	1	1		7.9993				−0.0047	−0.0032	−0.0009	33.9480	99.3217
	1	3	1		7.9993				0.0005	0.0007	0.0012	0.1867	47.6059
	1	1	3		7.9993				−0.0008	−0.0025	−.0009	11.3506	90.4499
Fingerprint 512 × 512	1	1	1	6.7279	7.9993	0.8644	0.7587	0.7261	0.0040	−0.0010	0.0049	0.1102	14.0533
	3	1	1		7.9992				−0.0009	0.0009	−0.0032	32.7711	99.6140
	1	3	1		7.9994				−0.0014	−0.0002	−0.0013	0.1102	14.0533
	1	1	3		7.9993				0.0043	−0.0007	−0.0007	0.2429	30.9715

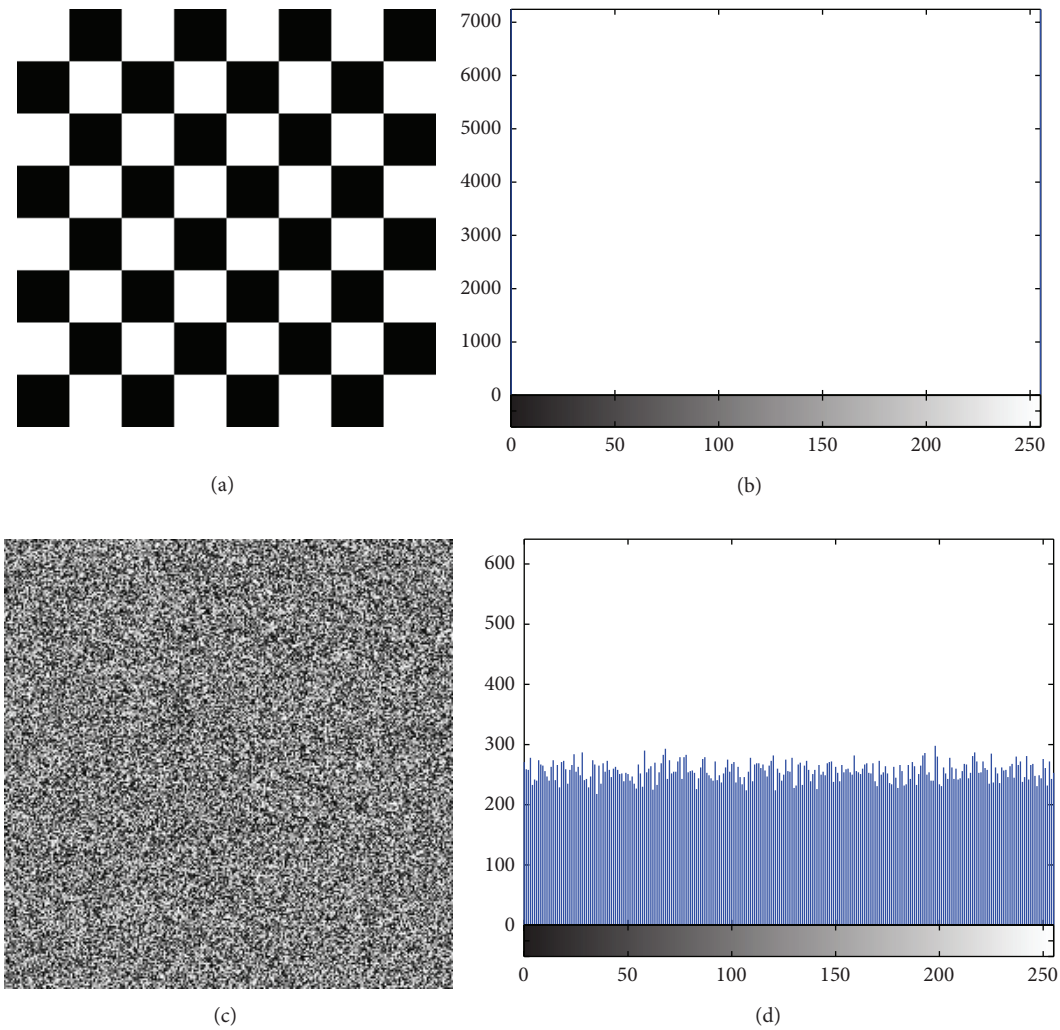


Figure 15: (a) Chess-plate image and (b) its histograms and (c) encrypted image and (d) its histogram.

the set of p and r was surveyed. The results for entropy, correlation, and brief of the UACI and NPCR are listed in Table 15.

5. Conclusion and Future Works

In this paper, a new chaos-based cryptosystem has been proposed for encrypting images. The Arnold cat map and the Henon map are two discrete chaotic maps that are used in this scheme. Bit shuffling and pixel shuffling are reversible transformations that are performed using the Arnold cat map with various secret parameters. Improving the randomness of transformation and the efficiency of bit permutation are two advantages of this cryptosystem that increases the strength of the ciphered image in comparison with previous works. Iterating the Arnold cat map with different parameters at each round prevents undesirable reconstruction of the input image. These parameters are generated by the Henon map with secret initial values. The points generated by the Henon map are also applied to create secret images for more confusion and diffusion and to increase the key space. Sequential XOR of the bit-permuted plain image and the pixel-permuted secret image is another phase of modifying the pixels values. This creates a slight distortion in the plain image to prevent successful differential attacks. The results of security analysis of five images demonstrate the resistance of the encrypted image to statistical attacks and to the chosen-plaintext attack. In addition, a sufficiently large key space makes a brute force attack impractical. As the future work, the proposed cryptosystem in this paper will combine with a public key technique such as ECC or RSA to propose a hybrid encryption method. This technique is a chaotic asymmetric cryptosystem.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors wish to thank Universiti Kebangsaan Malaysia (UKM) and Ministry of Higher Education Malaysia for supporting this work by research Grants FRGS/1/2012/SG05/UKM/02/1 and ERGS/1/2012/STG07/UKM/02/9.

References

[1] A. Soleymani, Z. Ali, and M. Nordin, "A survey on principal aspects of secure image transmission," in *Proceedings of World Academy of Science, Engineering and Technology*, pp. 247–254, 2012.

[2] P. P. Dang and P. M. Chau, "Image encryption for secure internet multimedia applications," *IEEE Transactions on Consumer Electronics*, vol. 46, no. 3, pp. 395–403, 2000.

[3] P. P. Dang and P. M. Chau, "Implementation IDEA algorithm for image encryption," in *Mathematics and Applications of Data/Image Coding, Compression, and Encryption III*, vol. 4122 of *Proceedings of SPIE*, pp. 1–9, August 2000.

[4] B. Furht, E. Muharemagic, and D. Socek, *Multimedia Encryption and Watermarking*, vol. 28 of *Multimedia Systems and Applications*, Springer, 2006.

[5] B. Subramanyan, V. M. Chhabria, and T. G. Sankar Babu, "Image encryption based on AES key expansion," in *Proceedings of the 2nd International Conference on Emerging Applications of Information Technology (EAIT '11)*, pp. 217–220, Kolkata, India, February 2011.

[6] E. Lorenz, *Predictability: Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in Texas*, American Association for the Advancement of Science, 1972.

[7] M. G. Signorini and M. Ferrario, "Nonlinear analysis of experimental time series," in *Advanced Methods of Biomedical Signal Processing*, pp. 347–378, John Wiley & Sons, New York, NY, USA, 2011.

[8] L. Shujun, M. Xuanqin, and C. Yuanlong, "Pseudo-random bit generator based on couple chaotic systems and its applications in stream-cipher cryptography," in *Progress in Cryptology — INDOCRYPT 2001*, vol. 2247 of *Lecture Notes in Computer Science*, pp. 316–329, 2001.

[9] Z. Zhu, W. Zhang, K. Wong, and H. Yu, "A chaos-based symmetric image encryption scheme using a bit-level permutation," *Information Sciences*, vol. 181, no. 6, pp. 1171–1186, 2011.

[10] S.-J. Xu, X.-B. Chen, R. Zhang, Y.-X. Yang, and Y.-C. Guo, "An improved chaotic cryptosystem based on circular bit shift and XOR operations," *Physics Letters A*, vol. 376, no. 10-11, pp. 1003–1010, 2012.

[11] Z. Zhang and T. Cao, "A chaos-based image encryption scheme with confusion- diffusion architecture," *Communications in Computer and Information Science*, vol. 152, no. 1, pp. 258–263, 2011.

[12] C. Fu, B. Lin, Y. Miao, X. Liu, and J. Chen, "A novel chaos-based bit-level permutation scheme for digital image encryption," *Optics Communications*, vol. 284, no. 23, pp. 5415–5423, 2011.

[13] Y. Zhang, P. Xu, and L. Xiang, "Research of image encryption algorithm based on chaotic magic square," *Advances in Intelligent and Soft Computing*, vol. 149, no. 2, pp. 103–109, 2012.

[14] M. Ghebleh, A. Kanso, and H. Noura, "An image encryption scheme based on irregularly decimated chaotic maps," *Signal Processing: Image Communication*, vol. 29, no. 5, pp. 618–627, 2014.

[15] A. M. Elshamy, A. N. Z. Rashed, A. E. A. Mohamed et al., "Optical image encryption based on chaotic baker map and double random phase encoding," *Journal of Lightwave Technology*, vol. 31, no. 15, pp. 2533–2539, 2013.

[16] R. Ye and W. Zhou, "An image encryption scheme based on 2D tent map and coupled map lattice," *International Journal of Information and Communication Technology Research*, vol. 1, pp. 344–348, 2011.

[17] R. Ye and W. Zhou, "A chaos-based image encryption scheme using 3D skew tent map and coupled map lattice," *International Journal of Computer Network and Information Security*, vol. 4, pp. 38–44, 2012.

[18] X. Wang, F. Chen, and T. Wang, "A new compound mode of confusion and diffusion for block encryption of image based on chaos," *Communications in Nonlinear Science and Numerical Simulation*, vol. 15, no. 9, pp. 2479–2485, 2010.

[19] S. Al-Maadeed, A. Al-Ali, and T. Abdalla, "A new chaos-based image-encryption and compression algorithm," *Journal of Electrical and Computer Engineering*, vol. 2012, Article ID 179693, 11 pages, 2012.

[20] V. Patidar, N. K. Pareek, and K. K. Sud, "A new substitution-diffusion based image cipher using chaotic standard and logistic maps," *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, no. 7, pp. 3056–3075, 2009.

[21] K. Wong, B. S. Kwok, and W. Law, "A fast image encryption scheme based on chaotic standard map," *Physics Letters, Section A: General, Atomic and Solid State Physics*, vol. 372, no. 15, pp. 2645–2652, 2008.

[22] C. Guanghuia, H. Kai, Z. Yizhi, Z. Jun, and Z. Xing, "Chaotic image encryption based on running-key related to plaintext," *The Scientific World Journal*, vol. 2014, Article ID 490179, 9 pages, 2014.

[23] Q. Zhang, X. Xue, and X. Wei, "A novel image encryption algorithm based on DNA subsequence operation," *The Scientific World Journal*, vol. 2012, Article ID 286741, 10 pages, 2012.

[24] H. Liu, Z. Zhu, H. Jiang, and B. Wang, "A novel image encryption algorithm based on improved 3D chaotic cat map," in *Proceedings of the 9th International Conference for Young Computer Scientists (ICYCS '08)*, pp. 3016–3021, November 2008.

[25] K. Wong, "Image Encryption Using Chaotic Maps," 2009.

[26] C. Zhu, "A novel image encryption scheme based on improved hyperchaotic sequences," *Optics Communications*, vol. 285, no. 1, pp. 29–37, 2012.

[27] J. S. A. Fouda, J. Y. Effa, S. L. Sabat, and M. Ali, "A fast chaotic block cipher for image encryption," *Communications in Nonlinear Science and Numerical Simulation*, vol. 19, no. 3, pp. 578–588, 2014.

[28] J. Chen, Z. Zhu, C. Fu, H. Yu, and L. Zhang, "A fast chaos-based image encryption scheme with a dynamic state variables selection mechanism," *Communications in Nonlinear Science and Numerical Simulation*, 2014.

[29] R. Matthews, "On the derivation of a "chaotic" encryption algorithm," *Cryptologia*, vol. 13, no. 1, pp. 29–42, 1989.

[30] X. Wang, L. Teng, and X. Qin, "A novel colour image encryption algorithm based on chaos," *Signal Processing*, vol. 92, no. 4, pp. 1101–1108, 2012.

[31] C. Li, L. Y. Zhang, R. Ou, K.-W. Wong, and S. Shu, "Breaking a novel colour image encryption algorithm based on chaos," *Nonlinear Dynamics. An International Journal of Nonlinear Dynamics and Chaos in Engineering Systems*, vol. 70, no. 4, pp. 2383–2388, 2012.

[32] C. Li, Y. Liu, T. Xie, and M. Z. Q. Chen, "Breaking a novel image encryption scheme based on improved hyperchaotic sequences," *Nonlinear Dynamics. An International Journal of Nonlinear Dynamics and Chaos in Engineering Systems*, vol. 73, no. 3, pp. 2083–2089, 2013.

[33] X. Wang, L. Yang, R. Liu, and A. Kadir, "A chaotic image encryption algorithm based on perceptron model," *Nonlinear Dynamics*, vol. 62, no. 3, pp. 615–621, 2010.

[34] Y. Zhang, C. Li, Q. Li, and D. Zhang, "Breaking a chaotic image encryption algorithm based on perceptron model," *Nonlinear Dynamics*, vol. 69, no. 3, pp. 1091–1096, 2012.

[35] G. Alvarez and S. Li, "Some basic cryptographic requirements for chaos-based cryptosystems," *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering*, vol. 16, no. 8, pp. 2129–2151, 2006.

[36] G. Alvarez and S. Li, "Breaking an encryption scheme based on chaotic baker map," *Physics Letters A*, vol. 352, no. 1-2, pp. 78–82, 2006.

[37] R. F. Machado, M. S. Baptista, and C. Grebogi, "Cryptography with chaos at the physical level," *Chaos, Solitons and Fractals*, vol. 21, no. 5, pp. 1265–1269, 2004.

[38] G. Alvarez and S. Li, "Cryptanalyzing a nonlinear chaotic algorithm (NCA) for image encryption," *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, no. 11, pp. 3743–3749, 2009.

[39] H. Gao, Y. Zhang, S. Liang, and D. Li, "A new chaotic algorithm for image encryption," *Chaos, Solitons and Fractals*, vol. 29, no. 2, pp. 393–399, 2006.

[40] Z. Guan, F. Huang, and W. Guan, "Chaos-based image encryption algorithm," *Physics Letters A: General, Atomic and Solid State Physics*, vol. 346, no. 1–3, pp. 153–157, 2005.

[41] D. Xiao, X. Liao, and P. Wei, "Analysis and improvement of a chaos-based image encryption algorithm," *Chaos, Solitons and Fractals*, vol. 40, no. 5, pp. 2191–2199, 2009.

[42] T. Xiang, X. Liao, G. Tang, Y. Chen, and K. Wong, "A novel block cryptosystem based on iterating a chaotic map," *Physics Letters A: General, Atomic and Solid State Physics*, vol. 349, no. 1–4, pp. 109–115, 2006.

[43] M. Hénon, "A two-dimensional mapping with a strange attractor," *Communications in Mathematical Physics*, vol. 50, no. 1, pp. 69–77, 1976.

[44] E. Lorenz, "Deterministic nonperiodic flow," *Journal of Atmospheric Sciences*, vol. 20, no. 2, pp. 130–141, 1963.

[45] C. Shannon, "Communication theory of secrecy systems," *The Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, 1949.