



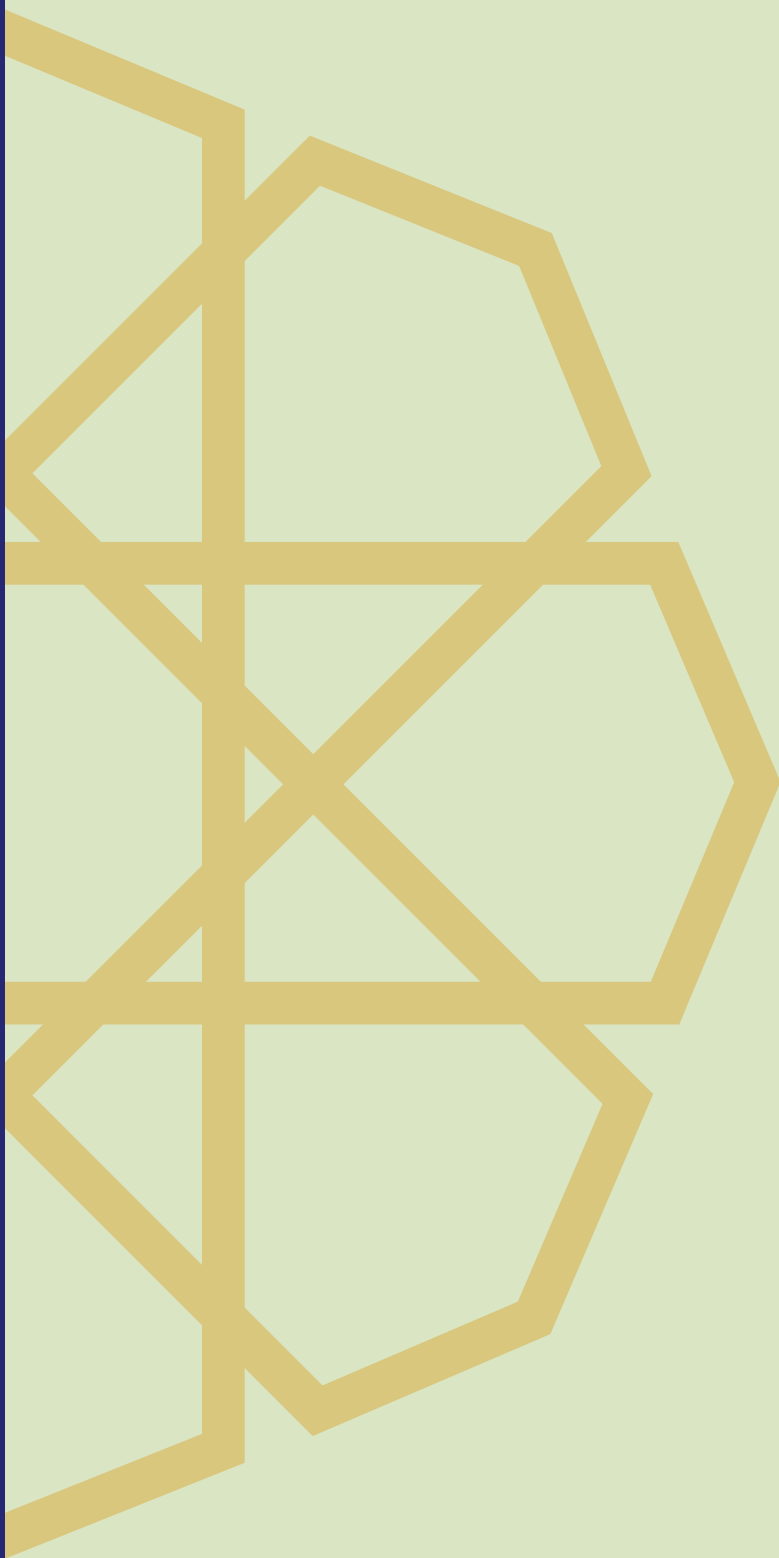
ISSN 2636-9680

OIC-CERT Journal of Cyber Security

Volume 1, Issue 1
January - December 2018

The Organisation of the Islamic Cooperation –
Computer Emergency Response Team
www.oic-cert.org

*Enhancing the knowledge on
cyber security among the
OIC member countries*



Editorial Panel

International Advisory Board

- Dato' Ts. Dr. Haji Amirudin Abdul Wahab, *CyberSecurity Malaysia (Malaysia)*
- Professor Datuk Ts. Dr. Shahrin Sahib@Sahibuddin, *Universiti Teknikal Malaysia Melaka (Malaysia)*
- Engr. Badar Al-Salehi, *Oman National CERT (Oman)*
- Dr. Rudi Lumanto, *Indonesia Security Incident Response Team on Internet Infrastructure / Coordination Center (Indonesia)*
- Abdul Hakeem Ajijola, *Consultancy Support Services Ltd (Nigeria)*
- Shamsul Bahri Kamis, *Brunei Computer Emergency Response Team (Brunei)*
- Professor Dr. Mohsen Kahani, *Ferdowsi University of Mashhad (Iran)*
- Professor Dr. Keith Martin, *Royal Holloway, University of London (United Kingdom)*
- Professor Xinyi Huang, *Fujian Normal University (China)*
- Professor Dr. Mohd Aizaini Maarof, *Universiti Teknologi Malaysia (Malaysia)*
- Professor Dr. Nathan Clarke, *University of Plymouth (United Kingdom)*
- Professor Dr. Mohammad Hossein Sheikhi, *Shiraz University (Iran)*

Editor-in-Chief

- Ts. Dr. Zahri Yunus, *CyberSecurity Malaysia (Malaysia)*
- Professor Ts. Dr. Rabiah Ahmad, *Universiti Teknikal Malaysia Melaka (Malaysia)*

Associate Editors-in Chief

- Mohd Shamir Hashim, *CyberSecurity Malaysia (Malaysia)*
- Dr. Shekh Faisal Abdul Latip, *Universiti Teknikal Malaysia Melaka (Malaysia)*

Editorial Board

- Ts. Dr. Solahuddin Shamsuddin, *CyberSecurity Malaysia (Malaysia)*
- Ts. Dr. Aswami Fadillah Mohd Arifin, *CyberSecurity Malaysia (Malaysia)*
- Professor Dr. Zulkalnain Mohd Yusoff, *Universiti Teknikal Malaysia Melaka (Malaysia)*
- Associate Professor Hatim Mohamad Tahir, *Universiti Utara Malaysia (Malaysia)*
- Associate Professor Dr. Noor Azurati Ahmad@Salleh, *Universiti Teknologi Malaysia (Malaysia)*
- Dr. S.M. Warusia Mohamed S.M.M Yassin, *Universiti Teknikal Malaysia Melaka (Malaysia)*
- Ts. Dr. Mohd Fairuz Iskandar Othman, *Universiti Teknikal Malaysia Melaka (Malaysia)*
- Dr. Muhammad Reza Z'aba, *University of Malaya (Malaysia)*
- Dr. Sofia Najwa Ramli, *Universiti Tun Hussein Onn Malaysia (Malaysia)*
- Dr. Azni Haslizan Ab Halim, *Universiti Sains Islam Malaysia (Malaysia)*

Technical Editorial Committee

- Noraini Abdul Rahman, *OIC-CERT Permanent Secretariat & CyberSecurity Malaysia (Malaysia)*
- Zaleha Abdul Rahim, *CyberSecurity Malaysia (Malaysia)*
- Ahmad Nasir Udin Mohd Din, *OIC-CERT Permanent Secretariat & CyberSecurity Malaysia (Malaysia)*
- Ts. Dr. Aslinda Hassan, *Universiti Teknikal Malaysia Melaka (Malaysia)*
- Dr. Raihana Syahirah Abdullah, *Universiti Teknikal Malaysia Melaka (Malaysia)*
- Dr. Nur Fadzilah Othman, *Universiti Teknikal Malaysia Melaka (Malaysia)*
- Ts. Dr. Zaki Mas'ud, *Universiti Teknikal Malaysia Melaka (Malaysia)*

Forward by the Editors-In-Chief

We would like to welcome everyone to the inaugural issue of the **OIC-CERT Journal of Cyber Security**, a peer-reviewed journal that aims to produce quality papers in the vast field of cyber security utilising a ready pool of cyber security professionals either from the industry or the academia from the OIC-CERT and the OIC member countries. The journal aspires to provide a platform for the academia and practitioners in cyber security to share experience and knowledge through research and publication thus contributing to the body of knowledge in cyber security.

The inaugural issue of the journal, an initiative by the Organization of the Islamic Cooperation – Computer Emergency Response Team (OIC-CERT), published seven papers that were reviewed and presented during the OIC-CERT Academic Colloquium 2018. The colloquium was held on 29 November in Shiraz, Iran in conjunction with the OIC-CERT Annual Conference and General Meeting 2018.

We are sincerely and deeply grateful to all authors, the editorial board, the technical committee and reviewers for their remarkable contributions and support. We invite submission of manuscripts for the next editions of the journal from cyber security professionals, scholars and practitioners involved in cyber security domains.

Ts. Dr. Zahri Yunos
CyberSecurity Malaysia

Professor Ts. Dr. Rabiah Ahmad
Universiti Teknikal Malaysia Melaka

Published by CyberSecurity Malaysia as the Permanent Secretariat to the OIC-CERT.
Level 5, Sapura@Mines, 7, Jalan Tasik, The Mines Resort City, 43300 Seri Kembangan,
Selangor Darul Ehsan, Malaysia.
Copyright © 2018 CyberSecurity Malaysia.
All rights reserved.

No part of this publication may be reproduced or distributed in any form or by means, or
stored in a database or retrieval system, without the prior written consent of CyberSecurity
Malaysia, including, but not limited to, in any network or other electronic storage or
transmission, or broadcast for distance learning.

Content

| | |
|--|----|
| SBPP: Statistical-Based Privacy-Preserving Approach for Data Gathering in Smart Grid <i>A. Ahadipour, M. Mohammadi, A. Keshavarz-Haddad</i> | 1 |
| A Hybrid Approach to Trust Inference in Social Networks <i>Maryam Fayyaz, Hamed Vahdat-Nejad, Mahdi Kherad</i> | 10 |
| Vulnerability Assessment and Penetration Testing of Virtualization <i>Ramin Vakili, Hamid Reza Hamidi</i> | 14 |
| Safeguarding Malaysia's Cyberspace against Cyber Threats: Contributions by CyberSecurity Malaysia <i>Fazlan Abdullah, Nadia Salwa Mohamad, Zahri Yunos</i> | 22 |
| Developing a Competency Framework for Building Cybersecurity Professionals <i>Ruhama Mohammed Zain, Zahri Yunos, Mustaffa Ahmad, Lee Hwee Hsiung, Jeffrey Bannister</i> | 32 |
| Preventing Reflective DLL Injection on UWP Apps <i>Mojtaba Zaheri, Salman Niksefat, Babak Sadeghiyan</i> | 41 |
| Crawler and Spiderin usage in Cyber-Physical Systems Forensics <i>M. Abedi, Sh. Sedaghat</i> | 53 |

SBPP: Statistical-Based Privacy-Preserving Approach for Data Gathering in Smart Grid

A. Ahadipour¹, M. Mohammadi², and A. Keshavarz-Haddad³

^{1,2} *PhD Candidate of Electrical Engineering*

³ *Faculty Member of Electrical Engineering*

School of Electrical and Computer Engineering

Shiraz University, Shiraz, Iran

ahadipour.alireza@shirazu.ac.ir, mojtaba.mohammadi@shirazu.ac.ir, keshavarz@shirazu.ac.ir

Abstract - As smart grids are getting popular and being employed widely, the privacy of users in such networks is getting more and more substantial. Decision making in smart grids depends on the information gathered from the users periodically. However, having access to the data relevant to the electricity consumption of users is inconsistent with their privacy. On the other hand, it is not sensible to entrust the responsibility of billing to consumers themselves. In this paper, we propose a statistical-based method for data gathering and billing in which the privacy of users is preserved, and at the same time, malicious consumers who try to send erroneous data would be detected.

KEYWORDS - Data Aggregator, Correlation Coefficient, Privacy, Smart Grid, Supplier, Statistical Method

I. INTRODUCTION

Recently, traditional grids underwent an alteration to smart grids which leads to many benefits including enhanced reliability and resilience, higher intelligence and optimized control, decentralized operation, higher operational efficiency, more efficient demand management, better power quality, and fraud detection [1]. Indeed, consumers minimize their expenses while providers maximize their revenue so that, a win-win partnership can be achieved.

The smart grid is envisaged to be the next generation of traditional grid. In contrast to the traditional grids, there is a bidirectional information flow between suppliers and consumers in smart. To provide this two-way communication, consumers should be equipped with smart meters by which they can measure their usage and send and receive their messages over various communication technologies such as power line communication, cable communication, and wireless communication.

Bidirectional information flows the supplier to generate the electricity based on the demands at any given time period; and at the same time, the supplier can define dynamic billing tariff, and regard to these tariffs that are sent to user periodically (e.g. every 15 minutes). Then, each user can decide whether to decrease its power consumption or not. Thus, electricity is

consumed in a more efficient way. On the other hand, in traditional grids, each user sends its electricity usage (by means of a third party) in fixed intervals (e.g. monthly) and its bill is calculated based on their whole usage; no matter their power consumption was in the pick hours or not. However, in smart grid, in the other direction of information flow, the users can declare their need for electricity; indeed, the users send their momentary electricity usage to the suppliers. As a result, unlike traditional grids, in smart grids suppliers provide electricity based on the need of consumers. Hence, ideally, no resource is wasted in the network [2].

In smart grids, one scenario for billing is that users send their electricity usage to local servers – which are responsible for gathering data – periodically by means of smart meters and then, local servers send the gathered data from users to local or central database. Then the server calculates the price of consumed electricity of each user based on the received data of that user. Criticism to this scenario is that the privacy would not be preserved in this method. As all consumers send their usage data to the server and these data are stored in a database, the pattern of each user's power consumption can be obtained by supplier; for instance, inhabitant's personal schedules, habits, religion, and so on.

Another scenario is that the supplier sends the time-varying tariffs periodically to the consumers and consumers compute their

electricity consumption price in the defined period (e.g. one month) based on the received tariffs. Eventually, at the end of each period, every user just sends its total billing amount to the supplier. In this case, the privacy of each consumer would be preserved. It is assumed that based on the existing information archived in databases regard to the power consumption of each user, the database can distinguish whether users are presenting correct billings or not. Consequently, one disadvantage of this scenario is that not only the supplier cannot find the malicious users, but also it would consider the honest ones guilty. For instance, if the power consumption pattern of a user alters over time, this user would be considered as a consumer who is declaring incorrect information; on the other hand, if there is a malicious user who ever sends artificial data, the database cannot notice this fact at all.

According to the afore mentioned scenarios, the main challenge in communications between consumers and suppliers is preserving the privacy of consumers and finding the malicious users simultaneously. To aim this goal, we propose a new statistical-based method for preserving privacy in data gathering of smart grids and at same time detecting the malicious users which manipulate their metering.

The remainder of this paper is organized as follows: In Related Works section, we briefly discuss related works. In System Model section, we introduce our system model. In Proposed Scheme section, we describe our proposed statistical-based scheme for data gathering in smart grid. In Simulation Results section, the simulation results of our scheme are presented. Finally, we conclude the paper in the last section.

II. RELATED WORK

Several algorithms for data gathering in smart grids have been studied in literature. In this section, we briefly review various privacy-preserving schemes for data gathering in smart grids.

In [3], an algorithm of data collection with self-awareness protection is proposed. They considered data collectors and respondents in their scheme and expressed that some of the respondents may not participate in contributing their personal data or submit erroneous data. To overcome this issue a self-awareness protocol was studied to enhance

trust of the respondents when sending their personal data to the data collector. All respondents collaborate with each other to preserve their privacy. The authors hired an idea, which allows respondents to know protection level before the data submission process is initiated. The paper is motivated by [4] and [5]. In [4], co-privacy (co-operative privacy) is introduced. Co-privacy claims that best solution to achieve privacy is to help other parties to achieve their privacy. More of co-privacy can be found in [4].

Many researchers focused on self-oriented privacy protection. One of the most interesting ones is [6] which proposes self-enforcing privacy (SEP) for e-polling. In this scheme, pollster must allow the respondents to track their submitted data in order to protect their privacy. In this case, respondents can accuse the pollster based on data they gathered during the collection process. Following this idea, a fair approach for accusation is presented in [7]. In [8], a respondent-defined privacy protection (RDPP) is introduced. It means that respondents are allowed to determine their required privacy protection level before delivering data to data collector. The main difference of this method is that unlike other methods, which data collector decides about the privacy protection level, respondents can freely define the privacy protection level.

To obtain privacy of residential users, a scheme named APED is proposed in [9]. It employs a pairwise private stream aggregation. They have shown that their scheme achieves privacy preserving aggregation and also executes error detection when some nodes fail to function normally. DG-APED is an improved form of APED, suggested in [10]. DG-APED propounds diverse grouping-based protocol with error detection. This research added differential privacy technique to APED. Moreover, DG-APED has an advantage of being efficient in term of communication and computation overhead compared to APED.

Authors in [11] first presented a new kind of attack, which adversary extracts information about the presence or absence of a specific person to access the smart meter information. They named this type of attack, human-factor-aware differential aggregation (HDA) attack and claimed that other proposed protocols cannot handle it. To solve this issue, they introduced two privacy-preserving protocols, a basic one and an advanced one.

They corroborated that their research can stand out against HDA attack by transmitting encrypted measurements to an aggregator in a way that aggregator cannot steal any information of human activities. By some implementations, it is demonstrated that the proposed method in [11] can guarantee privacy.

PDA is a scheme presented in [12]. It is a privacy-preserving dual-functional aggregation technique for smart grids in which, every user disseminates only one data and then data and control centre computes two statistical averages (mean and variance) of all users. Their simulations show that PDA is efficient concerning computational and communication overheads. The authors of [12], continued their researches leading to a privacy-preserving data aggregation with fault-tolerance called PDAFT [13]. In this work, a strong adversary is not able to gain any information, even in the case of compromising a few servers at the control centre (CC). Like PDA, PDAFT has a good communication overhead and is tenacious against many security threats. In a condition, which some users or servers fail, PDAFT can still work and this is the reason why they claimed that their proposed method has the fault-tolerance feature. DPAFT [14] is another privacy-preserving data collection scheme which supports both differential privacy and fault tolerance at the same time. It is claimed that, DPAFT surpass other schemes in many aspects, such as storage cost, computation complexity, utility of differential privacy, robustness of fault tolerance, and the efficiency of user addition or removal [14]. A new malfunctioning data aggregation scheme, named MuDA, is introduced in [15]. It is resistant to differential attacks and keeps users' information secret with an acceptable noise rate. PDAFT [15], DPAFT [14], and MuDA [15], shows nearly same characteristics. Their difference is in the cryptographic methods they use [16]. PDAFT employs homomorphic Paillier cryptosystem [17], while DPAFT and MUDA use Boneh-Goh-Nissim cryptosystem [18].

The paper [19] presents a secure power usage data aggregation for smart grid. By this method, supplier understands usage of each neighbourhood and makes decision about energy distribution, while it has no idea of the individual electricity consumption of each user. This scheme is designed to barricade

internal attacks and provide batch verification. Authors of [20] found out that [19] has the weakness of key leakage and the imposter can obtain the private key of user easily. It is proved that by using the protocol in [20], key leakage problem is solved and a better performance in term of computational cost is achieved. Neglecting energy cost is the disadvantage of this method.

Some other researches are also investigated in the field of privacy-preserving data collection. For example, in [21], authors designed a balanced anonymity and traceability for outsourcing small-scale linear data aggregation (called BAT-LA) in smart grid. They designed their protocol with the concern of providing both anonymity and traceability. Anonymity means that users' identity should be kept secret and traceability means that imposter users should be traced. Another challenge is that many devices are not capable of handling required complicated computations. Hence, they hired the idea of outsourcing computations with the help of public cloud. Authors of [21] utilized elliptic curve cryptography and proxy re-encryption to make BAT-LA secure. BAT-LA is evaluated by comparing it to two other schemes, RVK [22], and LMO [23] and it is shown that BAT-LA is more efficient in terms of confidentiality compared to the other two schemes [16].

The manuscript [24], a privacy-preserving protocol for smart grid is designed, which outsources computations to cloud servers completely. In this protocol, the data is encrypted before outsourcing and consequently cloud can perform any computations without decrypting data. It is claimed that their work became secure and efficient by using a multi-server framework. The paper [25] adopts perturbation techniques to preserve privacy and uses perturbation techniques and cryptosystems at the same time. This is designed in a way to be suitable for hardware-limited devices. Evaluations show that [25] is resilient to two types of attack, filtering attack, and true value attack. Authors of [26] divided their contribution to two parts. First it is described how an individual meter shares its readings to multiple users, and then the second part, where a user receives meter readings from multiple meters. Finally, they proposed a polynomial-based protocol for pricing. TPS3 [27] is security protocol, which is got its idea from Temporal Perturbation and Shamir's

Secret Sharing (SSS). Using both of these schemes simultaneously, makes it harder for adversary to obtain critical data of users. TPS3 guarantees privacy and reliability of users' data and begets a trade-off between communication cost and security. In [28], data collector tries to preserve privacy by adding some random noise to its computation result. To overcome the problem of computation accuracy reduction, an approximation method is proposed in [28] which leads to obtain a closed form of collector's decision problem.

In [29], a slightly different scenario is considered which data collector collects data from data providers and then spread it to data miner. The goal is to preserve providers' data privacy. Anonymization might be an answer, but it has its own challenges. To achieve a trade-off between privacy protection and data utility, interactions among three elements of scenario (data providers, data collector, and data miner) is modelled as a game and the Nash equilibria of the game is found. Simulations prove that the founded trade-off made an improvement to previous researches.

Some of the reviewed researches, such as [21] and [24] focused on outsourcing to clouds or distributed systems and prior to this, an encryption improves the security significantly. Based on which encryption method we use, it is important to use a secure key management scheme. The cryptographic technique ensures that no privacy sensitive information would be revealed. But, there is still the challenge of how to efficiently query encrypted multidimensional metering data stored in an untrusted heterogeneous distributed system environment [30]. The paper focused on this challenge and introduced a high performance and privacy-preserving query (P2Q) scheme and shows that it brings confidentiality and privacy in a semi-trusted environment.

III. SYSTEM MODEL

In this section, we present our system model. The essential elements of our SPBB approach include:

- i. *Consumer*: those who consume energy in a grid.
- ii. *Benign Consumer*: a consumer who reported its power consumption correctly.
- iii. *Malicious Consumer*: a consumer who reported its power consumption wrongly due to some purposes such as fraud or subversive goals.
- iv. *Supplier*: an entity whose responsibility is to provide energy for power consumers in a region.
- v. *Data Aggregator*: a local server whose liability is gathering the amount of power consumption information from consumers periodically and dispatching the gathered data to a supplier.
- vi. *Electricity Leakage*: the difference between the actual amount of consumed energy and the sum of quantity expressed by consumers as their power consumption.

Consider a grid consisting of M regions, each comprises one data aggregator and n_j consumers where j denotes the index of the region, that is $j \in \{1, \dots, M\}$. Consumers send their power consumption information measured by smart meters to the local aggregators. Data aggregators are responsible of gathering local data and sending it to the power supplier with a specific mechanism which will be presented in the subsequent section.

It is assumed that data aggregators are trusted. Indeed, no information leakage occurs at data aggregators, supposedly because after aggregation takes place, no raw information concerning power consumption of consumers would be at hand.

Besides, we assume that connections among above entities are secured by means of some cryptographic shared or public keys. Since smart meters on consumers' side cannot perform high computationally complex calculations, utilization of public key cryptography may not be sensible. Thus, employment of secret key cryptography would be a better option.

IV. PROPOSED SCHEME

In this paper we propose a method for data gathering with the purpose of informing the supplier of the instant power consumption. This algorithm provides the supplier with enough information about the demand for the power in the grid. Consequently, the power energy is produced based on the instantaneous

requirement and this would prevent waste of energy and supplies.

A. Data Gathering

Although the accuracy of smart grids' performance is engaged with the correctness of data gathered from consumers, this data gathering should not be in contrast with the privacy of consumers.

In this section we present a method for data gathering in smart grids which provide suppliers with data while keeping the users' power consumption information private and more importantly, find malicious consumers who try to send erroneous data to suppliers. We refer to this method as SBPP approach.

The proposed SBPP scheme for data gathering works as the following:

- i. Consumers send their power consumptions periodically to a local centre called data aggregator.
- ii. Each data aggregator selects one consumer randomly in each period.
- iii. It aggregates the power consumption of all consumers in that period except the randomly selected one.
- iv. Each data aggregator sends the aggregated amount of the previous step in accompany with the power

consumption of the randomly selected consumer to the supplier.

- v. The supplier provides energy based on the received power consumptions from data aggregators.

Figure 1 depicts how data gathering takes place. It is assumed that data aggregators are trusted, then power consumption information are not at hand any more after being aggregated by the data aggregators and being sent to the supplier. By this assumption, instead of having access to power consumption information of everyone at any period, a little portion of information is available about power consumption of each consumer. Suppose, for instance, there exist 100 consumers in a region with one data aggregator and let the period of data gathering be every 15 minutes. Without any data gathering algorithm, consumers would send their power consumption information to the supplier 2880 times ($30 \times 24 \times 60 / 15$) in a month, instead, by utilization of the above algorithm for data gathering, we have access to 0.01 of information corresponding to power consumption of users, that is, at most 29 times (0.01×2880) in a month.

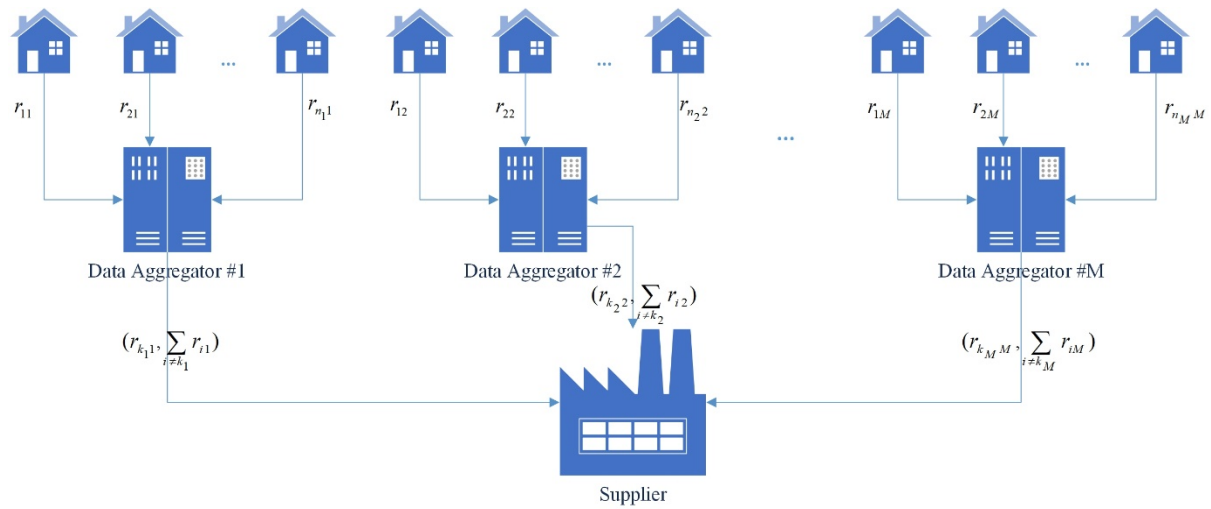


Figure 1: How power consumption information is sent to the supplier by data aggregators. Let P_{ij} be the power consumed by consumer i in region j and let k_j denotes the index of randomly chosen consumer in region j .

On the other hand, by utilization of the SBPP algorithm for data gathering, only 29 information regarding the power consumption of each consumer is available at the supplier in an analogous period. Although it may seem that having access to power consumption information of consumers is in

contradiction with their privacy, availability of these information 29 times a month would not reveal any data concerning their life style compared with approachability of these information 2880 times within a month.

B. Finding Malicious Consumers

Malicious consumers pursue two distinct aims by sending erroneous data to suppliers. Either they declare their amount of power consumption lesser than their real consumed power to pay lower fee; or, they express their power consumption quantity much more so as to impose more expenditure to the supplier.

In this paper, we get use of correlation coefficient of power consumption of consumers to find malicious consumers in each region who try to send erroneous data to the supplier.

Correlation coefficient illustrates the statistical relationship between two variables and it is defined as follows:

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{cov}(X, X)\text{cov}(Y, Y)}} \quad (1)$$

where *corr* is a widely used alternative notation for the correlation coefficient and *cov* means covariance. Correlation coefficient possesses values in the range of -1 to +1, where -1 and +1 indicate the strongest possible agreement and disagreement respectively.

In order to find malicious consumers, it is assumed that data aggregators are aware of the total amount of power consumed in each region. By comparing this amount with the aggregated quantity declared by consumers, the shortage amount can be determined.

Having access to merely one quantity of power consumption information corresponding to a consumer does not suffice to distinguish if that consumer is benign or malicious. In other words, the more information we have regarding power consumption of each consumer, the better decision we can make about the sabotage of consumers. Thus, the algorithm for finding malicious consumers takes place at the end of each month.

So as to detect malicious consumers, each data aggregator stores the identity (ID) of the randomly selected consumer, its declared power consumption, and the leakage amount of power consumed in that region at every period. At the end of each month, for each consumer, the data aggregator computes the correlation coefficient of its reported consumed energy and the leakage amounts of power consumption. Henceforth, we define the leakage quantity as:

$$\text{leakage} = \text{actual amount} - \text{reported amount} \quad (2)$$

If the correlation coefficient turns to +1 for a consumer (according to (2)), it means that consumer had expressed its power consumption less than its actual used power. On the other hand, if the correlation coefficient for a user turns to -1, it means that consumer is declaring its power consumption more than its usage due to some subversive goals. Thus, the proposed scheme is capable of not only detecting malicious users, but also comprehending if that user is declaring its amount of power consumption less or more than its actual quantity.

Furthermore, it is possible that there exists more than one malicious user in a region. In this case, although the correlation coefficient corresponding to these users would not be equal to ± 1 , their correlation coefficient quantity will be maximum (or minimum) amongst other consumers. As a result, it is needed that a threshold (*th*) be defined where the absolute value of correlation coefficients fewer or more than the threshold indicate benign or malicious users respectively, as:

$$\begin{cases} \text{malicious user,} & -1 \leq \text{corr} \leq -th \\ \text{benign user,} & -th \leq \text{corr} \leq th \\ \text{malicious user,} & th \leq \text{corr} \leq 1 \end{cases} \quad (3)$$

It is apparent that the more the threshold is, the less malicious consumers are detected and on the other hand, the less the threshold is, the more benign users are considered malicious. Thus, a question that arises here is that *how should this threshold be determined?* The analysis concerning the detection of several malicious users in a region is left for future works, however, we briefly discuss the problem in the following. In this paper, according to the setting of the problem, we set the threshold to a fixed value namely 0.5.

As the proposed scheme is a statistical one, it is probable that the correlation coefficient of a benign user lies out of its defined region depicted in (3), or vice versa, that is, the correlation coefficient corresponding to a malicious consumer lies in the region belonging to benign ones.

C. Billing

In this section, we propose an algorithm for billing. As discussed in the preceding section, malicious consumers can be

distinguished by computing correlation coefficient of all consumers in a region. Malicious consumers' being determined, sent data corresponding to other consumers are considered trustworthy and error free. By this assumption, the liability for billing can be assigned to data aggregators. In every period, consumers send their amount of consumed energy to data aggregators. Based on the received data from consumers and the received tariffs from the supplier, data aggregators compute the cost of consumed power for each consumer before data aggregation takes place. In each period, data aggregators calculate the cost of consumed power for each consumer and add the cost to the previously calculated cost for that consumer and by the end of month, a bill will be issued and sent to each consumer.

Not only this algorithm decreases the signalling overhead, but also the privacy of consumers would be protected. It is merely required that suppliers send tariffs periodically to data aggregators and consumers simultaneously. Data aggregators compute the cost of consuming energy for every consumer and smart meters on the consumers' side adjust the power consumption based on the received tariffs, i.e., if tariff increases, smart meters force dispensable devices to be turned off. In this case, no information leakage and thus no privacy invasion would occur.

Besides, by finding malicious consumers in each region and by comparing the amount of power consumed by other consumers and the total amount of produced energy, the power consumption quantity of malicious consumers would be determined. However, that how the bill of these malicious consumers should be calculated and what penalties should be intended for these consumers are not considered in this paper.

V. SIMULATION RESULTS

In this section, we present the results of simulations for the proposed SBPP approach. We would show that our proposed scheme can detect malicious users who send bogus information concerning their power consumption quantity in a smart grid.

Consider a region consisting of 100 consumers and one data aggregator where data aggregation takes place every 15 minutes and assume that consumer # 25 is a malicious

user. Two cases are studied; user # 25 in case (a) expresses one tenth of its power consumption and in case (b) it reports its power usage 10 times more than its actual consumption. Figure 2 (a) illustrates case (a) where the correlation coefficient of expressed consumed energy and the leakage amounts of power consumption turns to +1 and Figure 2 (b) depicts case (b) where the correlation coefficient turns to -1.

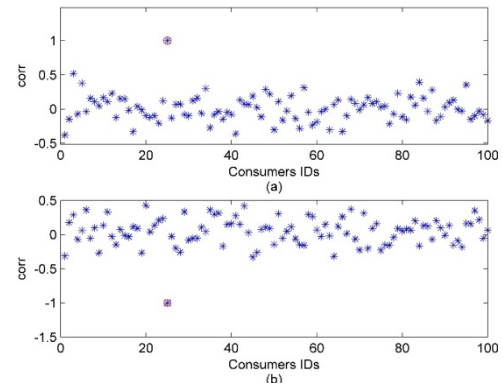
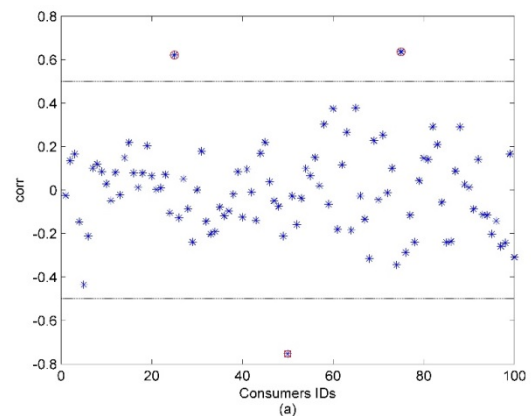


Figure 2: Correlation coefficient of reported energy consumption and the leakage amounts of power consumption for all users in the grid. (a) One malicious user declares its power consumption less than the actual quantity and (b) One malicious user declares its power consumption more than the actual quantity

Consider the previous assumptions except that there are three malicious consumers instead of one in that region with IDs 25, 50, and 75. Consumers with IDs 25 and 75 declare their power consumption less than their actual consumption and consumer # 50 expresses its power consumption more than its actual consumed energy. By setting the threshold to 0.5, consumers with absolute value of correlation coefficient greater than 0.5, that is, $|corr| \leq 0.5$, would be considered malicious, as depicted in Figure 3.



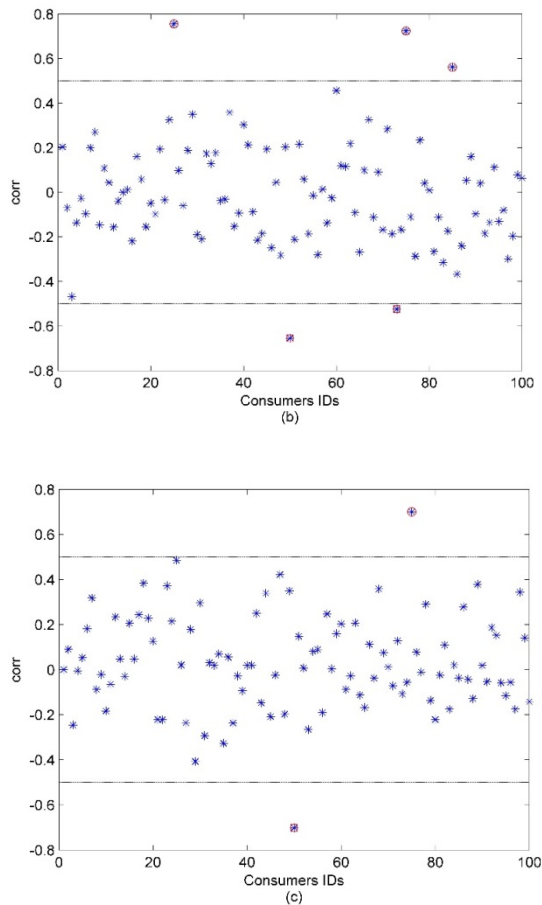


Figure 3: Detection of several malicious users (a) all malicious user are detected correctly, (b) in addition to malicious users, a number of benign users are found malicious, and (c) not all malicious users are detected.

As it can be seen from Figure 3, fixed threshold will result in 3 cases: 1) only malicious users been detected (Figure 3 (a)), 2) in addition to malicious users, some benign users found malicious (Figure 3 (b)), and 3) a subset of malicious users been detected (Figure 3 (c)).

VI. CONCLUSION

We presented a statistical-based approach for data gathering in smart grids which preserves the privacy of consumers. We investigated the capability of the proposed scheme in detecting malicious consumers who dispatch bogus data to service providers for a specific purpose such as abating their cost or imposing expenditure on suppliers (subversive goals). Furthermore, we showed that if there exists only one malicious user, it can definitely be detected if enough number of samples are gathered. When there are more malicious users, depending on the number of gathered samples, it is probable that all malicious consumers being detected, some

benign consumers found malicious, or a subset of malicious users being detected. We also presented a scheme for billing which concede the liability of billing to data aggregators in each region. By employing this scheme, not only the signalling overhead decreases significantly, but also billing occurs at a trusted entity where malicious consumers are distinguished from benign ones. Our simulation results verified these terms.

VII. REFERENCES

- [1] E. Fadel, V. C. Gungor, L. Nassef, N. Akkari, M. A. Malik, S. Almasri, and I. F. Akyildiz, "A survey on wireless sensor networks for smart grid," *Computer Communications*, vol. 71, pp. 22–33, 2015.
- [2] A. Alabdulatif, H. Kumarage, I. Khalil, M. Atiquzzaman, and X. Yi, "Privacy-preserving cloud-based billing with lightweight homomorphic encryption for sensor-enabled smart grid infrastructure," *IET Wireless Sensor Systems*, vol. 7, no. 6, pp. 182–190, 2017.
- [3] K.-S. Wong Wong and M. H. Kim, "Privacy-preserving data collection with self-awareness protection," in *Frontier and Innovation in Future Computing and Communications*. Springer, 2014, pp. 365–371.
- [4] J. Domingo-Ferrer, "Coprivacy: towards a theory of sustainable privacy," in *International Conference on Privacy in Statistical Databases*. Springer, 2010, pp. 258–268.
- [5] J. D. Ferrer, "Coprivacy: an introduction to the theory and applications of co-operative privacy," *SORT: statistics and operations research transactions*, pp. 0025–40, 2011.
- [6] P. Golle, F. McSherry, and I. Mironov, "Data collection with self-enforcing privacy," *ACM Transactions on Information and System Security (TISSEC)*, vol. 12, no. 2, p. 9, 2008.
- [7] M. Stegelmann, "Towards fair indictment for data collection with self-enforcing privacy," in *IFIP International Information Security Conference*. Springer, 2010, pp. 265–276.
- [8] R. Kumar, R. Gopal, and R. Garfinkel, "Freedom of privacy: anonymous data collection with respondent-defined privacy protection," *INFORMS Journal on Computing*, vol. 22, no. 3, pp. 471–481, 2010.
- [9] R. Sun, Z. Shi, R. Lu, M. Lu, and X. Shen, "Aped: An efficient aggregation protocol with error detection for smart grid communications," in *Global Communications Conference (GLOBECOM)*, 2013 IEEE. IEEE, 2013, pp. 432–437.
- [10] Z. Shi, R. Sun, R. Lu, L. Chen, J. Chen, and X. S. Shen, "Diverse grouping-based aggregation protocol with error detection for

- smart gri grouping-based aggregation protocol with error detection for smart grid communications,” *IEEE Transactions on Smart Grid*, vol. 6, no. 6, pp. 2856–2868, 2015.
- [11] W. Jia, H. Zhu, Z. Cao, X. Dong, and C. Xiao, “Human-factor-aware privacy-preserving aggregation in smart grid,” *IEEE Systems Journal*, vol. 8, no. 2, pp. 598–607, 2014.
- [12] C. Li, R. Lu, H. Li, L. Chen, and J. Chen, “Pda: a privacy-preserving dual-functional aggregation scheme for smart grid communications,” *Security and Communication Networks*, vol. 8, no. 15, pp. 2494–2506, 2015.
- [13] L. Chen, R. Lu, and Z. Cao, “Pdaft: A privacy-preserving data aggregation scheme with fault tolerance for smart grid communications,” *Peer-to-peer networking and applications*, vol. 8, no. 6, pp. 1122–1132, 2015.
- [14] H. Bao and R. Lu, “A new differentially private data aggregation with fault tolerance for smart grid communications,” *IEEE Internet of Things Journal*, vol. 2, no. 3, pp. 248–258, 2015.
- [15] L. Chen, R. Lu, Z. Cao, K. AlHarbi, and X. Lin, “Muda: Multifunctional data aggregation in privacy-preserving smart grid communications,” *Peer-to-peer networking and applications*, vol. 8, no. 5, pp. 777–792, 2015.
- [16] M. A. Ferrag, L. A. Maglaras, H. Janicke, and J. Jiang, “A survey on privacy-preserving schemes for smart grid communications (2016),” *arXiv preprint arXiv:1611.07722*, 2016.
- [17] P. Paillier, “Public-key cryptosystems based on composite degree residuosity classes,” in *International Conference on the Theory and Applications of Cryptographic Techniques*. Springer, 1999, pp. 223–238.
- [18] D. Boneh, E.-J. Goh, and K. Nissim, “Evaluating 2-dnf formulas on ciphertexts,” in *Theory of Cryptography Conference*. Springer, 2005, pp. 325–341.
- [19] C.-I. Fan, S.-Y. Huang, and Y.-L. Lai, “Privacy-enhanced data aggregation scheme against internal attackers in smart grid,” *IEEE Transactions on Industrial informatics*, vol. 10, no. 1, pp. 666–675, 2014.
- [20] D. He, N. Kumar, and J.-H. Lee, “Privacy-preserving data aggregation scheme against internal attackers in smart grids,” *Wireless Networks*, vol. 22, no. 2, pp. 491–502, 2016.
- [21] H. Wang, D. He, and S. Zhang, “Balanced anonymity and traceability for outsourcing small-scale data linear aggregation in the smart grid,” *IET Information Security*, vol. 11, no. 3, pp. 131–138, 2016.
- [22] H. Wang, B. Qin, Q. Wu, L. Xu, and J. Domingo-Ferrer, “Tpp: Traceable privacy-preserving communication and precise reward for vehicle-to-grid networks in smart grids,” *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 11, pp. 2340–2351, 2015.
- [23] C. Rottondi, G. Verticale, and C. Krauss, “Distributed privacy-preserving aggregation of metering data in smart grids,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1342–1354, 2013.
- [24] H. Chun, K. Ren, and W. Jiang, “Privacy-preserving power usage and supply control in smart grid,” *Computers & Security*, 2018.
- [25] U. B. BALOGLU and Y. DEMIR, “Lightweight privacy-preserving data aggregation scheme for smart grid metering infrastructure protection,” *International Journal of Critical Infrastructure Protection*, 2018.
- [26] A. Rial, G. Danezis, and M. Kohlweiss, “Privacy-preserving smart metering revisited,” *International Journal of Information Security*, vol. 17, no. 1, pp. 1–31, 2018.
- [27] M. U. Simsek, F. Yildirim Okay, D. Mert, and S. Ozdemir, “Tps3: A privacy preserving data collection protocol for smart grids,” *Information Security Journal: A Global Perspective*, vol. 27, no. 2, pp. 102–118, 2018.
- [28] G. Liao, X. Chen, and J. Huang, “Optimal privacy-preserving data collection: A prospect theory perspective,” in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.
- [29] L. Xu, C. Jiang, Y. Qian, Y. Ren, L. Xu, C. Jiang, Y. Qian, and Y. Ren, “Privacy-preserving data collecting: A simple game theoretic approach,” *Data Privacy Games*, pp. 45–57, 2018.
- [30] R. Jiang, R. Lu, and K.-K. R. Choo, “Achieving high performance and privacy-preserving query over encrypted multidimensional big metering data,” *Future Generation Computer Systems*, vol. 78, pp. 392–401, 2018.

A Hybrid Approach to Trust Inference in Social Networks

Maryam Fayyaz¹, Hamed Vahdat-Nejad², and Mahdi Kherad³

¹ Department of and Computer Engineering, Islamic Azad University of Birjand, Birjand, Iran

^{2,3} Faculty of Electrical and Computer Engineering, University of Birjand, Birjand, Iran

maryam_fayaz71@yahoo.com, vahdatnejad@birjand.ac.ir, m.kherad@birjand.ac.ir

Abstract - The trust inference issue in a social network is defined as anticipating the trust level which a user can have to another user who is not directly connected to him in the trust network. This paper proposes a method for trust inference using soft computing. To our best knowledge, it is the first time that soft computing is used to solve the trust inference issue. One of the main advantages of the proposed method is that, unlike previous methods, it is not limited to one type of trust network, and it can also be used for trust networks with different trust values. The proposed method is applied on the standard trust network and is compared to other similar methods. Experimental results show that it is able to produce more accurate results in comparison with previous methods.

KEYWORDS - Trust Inference, Social Network, Soft Computing

I. INTRODUCTION

Trust plays an important role in the formation of the relations between users. In fact, users share their information according to their trust on other users or make decision based on provided information by other users. We deal with a graph in social networks which its vertices are users and edges are relations between them. The main issue is how to inference trust between people who are not connected directly.

Social network is a term used for the first time in 1954 by [1] who was active in the field of Social studies [2]. He studied a research about social groups in Norway and used 'social network' term in that research to describe the relationship between humans and analyse communication mechanisms. A social network is a graph $G = (V, E)$ in which $V = \{v_1, v_2, v_3, \dots\}$ is set of vertices and $E = \{e_1, e_2, e_3, \dots\}$ is set of edges and each edge interconnects a pair of vertices together.

Any computational model, which is proposed for trust inference up to now, suggests a particular representation method. [3] and [4] consider a discrete set of values and a continuous numerical range to show trust, respectively. [5] and [6] select the continuous range of $[0,1]$ as the set of allowed values to show trust. [7] considers both continuous range of $[0,10]$ and discrete binary values of 0 and 1.

Models that utilize social network structure are specially based on trust concepts of web or friend of friend [8]. [5] presents an algorithm to traverse trust graph and infer trust. TidalTrust model [7] reviews the value of trust using numbers at the range of 0 to 10. This

model is simple and its low complexity leads to high scalability. In the current research, trust values are considered through paths, as a result only the shortest path from source to destination is checked.

Although soft computing is a powerful tool for solving similar problems, it has not been used in previous trust inference methods. One of the most important advantages of the proposed method is that unlike previous methods, it is not limited to one type of trust network, but applicable to different trust networks with various trust values.

This research aims to infer trust in a social network based on social behavior. In fact, the aim is predicting the trust that a user can have to another user who is not connected directly to. The genetic algorithm and neural network are used in the proposed method. Neural network has not been used in any of previous trust inference methods. In the proposed method, three features of the social network are exploited, which represent different aspects of trust. Therefore, a model based on neural network predicts trust values regarding these features. Finally, genetic algorithm is utilized to set the weights and balance the neural network. The experimental results show higher precision for the proposed method in comparison to BBK [9], Simple average [1], TidalTrust [7], TISoN [10] and κ -FuzzyTrust [11] methods in estimating the amount of trust.

After this introduction, the proposed method is presented in section 3. In section 4, the experimental results are discussed. Last of all, the final section deals with the conclusion and future research.

II. THE PROPOSED METHOD

We face with two problems when working with neural networks: choosing the right architecture, and choosing the right training algorithm. The architecture of neural network includes number of hidden layers, number of neurons in hidden layers and the stimulation function. Each of these parameters affects the performance of neural network, directly and significantly [12].

The most common neural training algorithm is Back propagation algorithm [13]. The problem of Back propagation algorithm is late convergence and also stopping in local optimized points. One approach in training neural networks is using innovative algorithms such as genetic that in fact, is considered as a part of soft computing [14]. Genetic algorithms are from a family of computational models inspired from Evolution theorem. They indicate a possible solution for specified problems using the data structure of chromosome and apply combined operations on this data structure to protect vital information [15]. The genetic algorithm is an optimization mechanism according to the process of selecting the best in the nature [16]. In a genetic neural system, every chromosome indicates weight values and biases. To determine the fitness value of each chromosome, neural network runs with weight and bias values of the chromosome and neural network error is calculated as the fitness function of the Genetic algorithm [14].

The main steps of the proposed method are as follows:

- i. Loading network information: At first, adjacent matrix of trust social network graph is loaded.
- ii. Feature extraction: In this stage, for each direct link in the network graph, four characteristics are calculated and a sample data is added to the training data. The output class corresponding to each of these data samples is the value of link or the trust between two.
- iii. Setting up the neural network: In this step, the proposed.
- iv. Setting up the genetic algorithm: In this stage, the genetic system is created for adjusting the parameters of the neural network. The length of a chromosome is equal to the number of weights and biases of the neural.
- v. Finalization of neural network: At the end, the best obtained chromosome

determines the best weights for neural network.

In each iteration, one link (u,v) of the trust graph is eliminated temporarily, and the features of the link are computed. The process is iterated for all links. These features contain following items, which are considered as input for neural network:

Mean trust of source node u (MST): This feature indicates the average of trust values that the source node u has to its neighbouring nodes.

$$MST_u = \frac{\sum_{j \in adj^+(u)} t_{uj}}{|adj^+(u)|} \quad (1)$$

Where $adj^+(u)$ is the set of neighboring nodes of u, that exists a link from u to them and t_{uj} is the trust value of node u to the node j.

Mean trust of destination node v (MDT): This feature shows the average of trust values that neighbouring nodes u have to node v.

$$MDT_v = \frac{\sum_{j \in adj^-(v)} t_{jv}}{|adj^-(v)|} \quad (2)$$

Where $adj^-(v)$ is the set of neighbors of v that there exist links from them to v and t_{jv} is the trust value of node j to node v.

Distance: This feature points to the value of the shortest path between a pair of source and destination nodes. The greater the distance between the two nodes of source and destination, the less influenced is the relation between source and destination user. In fact, the estimated trust value of source user to the destination user is influenced by the distance between them.

Multilayer perceptron neural network (MLP) is used for predicting trust. Since three features of MST, MDT and Distance are considered, the number of inputs of neural network is three. The proposed neural network consists of ten outputs, which are the estimated trust value. The number of neurons of the input layer with the number of features of input data and the number of neurons of output layer with the number of outputs are equal, respectively. The number of hidden layers is three, because a neural network with more than two layers is able to solve any kind of problem. Figure 1 shows the proposed neural network architecture.

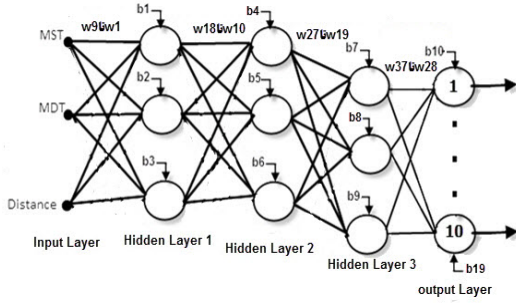


Figure 1: Architecture of the proposed neural network.

As it can be seen in Figure 1, the total number of neurons is equal to 19. Hence, the number of biases and weights to train the neural network is 19 and 37, respectively (one bias is considered for each neuron). The aim of the genetic algorithm is to determine the biases and the optimized weights of the neural network for the estimation of trust. In a chromosome, the genes of 1 to 37 indicate the weights from w_1 to w_{37} of the neural network and the genes from 38 to 56 indicate values of neurons' biases (b_1 to b_{19}). Therefore, each chromosome has 56 genes that are able to take a value in the range of -1 to 1. Figure 2, shows the structure of a chromosome for training the neural network.



Figure 2: The structure of a chromosome for training the proposed neural network.

Weights and biases are set using the genetic algorithm so that output trust has minimum error and maximum precision. The fitness function is given in formula 3.

$$f(x) = \sum_{i=1}^n |t_{ri} - t_{xi}| \quad (3)$$

Where $f(x)$ is the fitness function of the chromosome x , n is the number of training data elements, t_{ri} is the value of real trust for i th data element, and t_{xi} is the value of output trust of neural network generated by weights of chromosome x .

III. EXPERIMENTS

The social network used in this research is a part of trust project of mindswap [17] and FilmTrust [18]. Mindswap is created of obtained data from semantic web. In this network, users give the rank of trust between 1 (minimum trust) to 10 (maximum trust). Mindswap consists of about 2000 members with more than 2500 relations. FilmTrust is a

dataset of a website, in which people comment their opinions about different movies and also give a trust value between one to ten to others' opinions. This collection consists of about 900 users and 1067 links (direct trust) between them.

Matlab software is used for implementing the proposed method. 70 percent of data is considered for training, 15 percent as the test data, and 15 percent as validation data for neural network. In the genetic algorithm, initial population is 100, number of iterations is 1000, crossover rate is 0.8 and mutation rate is 0.2.

The proposed method is compared with five other methods of trust inference including BBK [9], simple average [1], TidalTrust [7], TISoN [10] and κ -FuzzyTrust [11]. These methods take two trust nodes in a trust network and calculate how much trust one node has to the other node. To determine the precision, Δ is calculated, which is the difference between actual value of trust between two nodes and the trust value inferred using the algorithm. In Table I, the average value of Δ is given for each of the methods over the dataset.

Table 1: The Average of accuracy for different methods of trust inferecing

| Results on mindswap dataset | | | | | |
|------------------------------|-----------------------|-------|----------------|------|-------------|
| Proposed method | κ -Fuzzy Trust | TISoN | Simple Average | BBK | Tidal Trust |
| 1.07 | 1.33 | 1.24 | 1.43 | 1.59 | 1.09 |
| Results on FilmTrust dataset | | | | | |
| Proposed method | κ -Fuzzy Trust | TISoN | Simple Average | BBK | Tidal Trust |
| 1.41 | 1.52 | 1.49 | 1.72 | 1.64 | 2.38 |

As Table 1 shows, the proposed method achieves more accurate results in comparison with previous methods.

IV. CONCLUSION

In this paper, a hybrid model for trust inference in social networks using genetic algorithm and neural network has been proposed. In fact, the proposed neural network system is constituted based on the genetic algorithm. To evaluate the proposed method, the model has been coded in Matlab and implemented on validated social networks. Due to the obtained results, the proposed algorithm is an appropriate method in solving trust inference. The results confirm that this

method is able to produce trust values close to the actual ones.

V. REFERENCES

- [1] J. Golbeck, "Trust on the World Wide Web: A survey," *Found. Trends Web Sci.*, vol. 1, no. 2, pp. 131–197, 2006.
- [2] J. Scott, *Social network analysis*. Sage, 2017.
- [3] E. Elsalamouny, V. Sassone, and M. Nielsen, "HMM-based trust model," in *6th International Workshop on Formal Aspects on Security and Trust (FAST)* vol. 5983, pp. 21–35, 2010.
- [4] R. Xiang, J. Neville, and M. Rogati, "Modeling relationship strength in online social networks," in *19th International Conference on World Wide Web (WWW'10)*, New York, 2010, pp. 981–990: ACM Press.
- [5] A. Josang, "Probabilistic logic under uncertainty," in *the thirteenth Australasian symposium on Theory of computing*, Darlinghurst, Australia, 2007, vol. 65, pp. 101–110: Computer Society.
- [6] J. Tang, Y. Chang, C. Aggarwal, and H. Liu, "A survey of signed network mining in social media," *ACM Computing Surveys (CSUR)*, vol. 49, no. 3, p. 42, 2016.
- [7] J. A. Golbeck, "computing and a applying trust in web-based social networks," PhD thesis, Department of Computer Science, University of Maryland, Maryland, College Park, MD, USA, 2005.
- [8] W. Sherchan, S. Nepal, and C. Paris, "A Survey of Trust in Social Networks," *ACM Computing Surveys*, vol. 45, no. 4, 2013.
- [9] Y. Wang, Z. Cai, G. Yin, Y. Gao, and Q. Pan, "A trust measurement in social networks based on game theory," in *International Conference on Computational Social Networks*, 2015, pp. 236–247: Springer.
- [10] S. Hamdi, A. L. Gancarski, A. Bouzeghoub, and S. B. Yahia, "Tison: Trust inference in trust-oriented social networks," *ACM Transactions on Information Systems (TOIS)*, vol. 34, no. 3, p. 17, 2016.
- [11] S. Chen, G. Wang, and W. Jia, " κ -FuzzyTrust: efficient trust computation for large-scale mobile social networks using a fuzzy implicit social graph," *Information Sciences*, vol. 318, pp. 123–143, 2015.
- [12] B. D. Ripley, *Pattern recognition and neural networks*. Cambridge university press, 2007.
- [13] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.
- [14] X.-S. Yang and M. Karamanoglu, "Swarm intelligence and bio-inspired computation: an overview," in *Swarm Intelligence and Bio-Inspired Computation*: Elsevier, 2013, pp. 3–23.
- [15] S. Karakatič and V. Podgorelec, "A survey of genetic algorithms for solving multi depot vehicle routing problem," *Applied Soft Computing*, vol. 27, pp. 519–532, 2015.
- [16] K. Sastry, D. E. Goldberg, and G. Kendall, "Genetic algorithms," in *Search methodologies*: Springer, pp. 93–117, 2014,.
- [17] 2017, *Trust Project Network*. Available: <http://trust.mindswap.org>.
- [18] 2018, *FilmTrust*. Available: <https://www.librec.net/datasets/filmtrust.zip>.

Vulnerability Assessment and Penetration Testing of Virtualization

Ramin Vakili¹ and Hamid Reza Hamidi²

^{1,2}CERT Laboratory, Faculty of Engineering,
Imam-Khomeini International University, Qazvin, Iran
ramin.vakili@edu.ikiu.ac.ir, hamidreza.hamidi@eng.ikiu.ac.ir

Abstract - Virtualization brings us lots of significant usages and is a useful technology in data centres and cloud computing. Using virtualization could either reduce security issues or bring new ones. In this research we have tried to review security advantages and disadvantages of virtualization technology. Security specialists assess the security of a system using automatic tools for penetration testing and vulnerability assessment. In this paper, we also review some of the tools that can be used in security assessment of virtualization.

KEYWORDS - Virtualization, Cloud Computing, Penetration Testing, Vulnerability Assessment

I. INTRODUCTION

Virtualization is a platform which allows us to partition the computer system resources into multiple execution environments. Virtualization increases the utilization of systems and makes the managing of organizations infrastructure easier. This is one of the main reasons that has increased its popularity. Using virtualization would bring some security benefits and it also might cause new security issues [1].

Penetration testing and vulnerability assessment are a set of practical methods which is done by security specialist using tools to assess the security of systems. The goal of these methods is to find the vulnerable parts of a system and to confirm whether the current security measures are effective or not [2]. In this paper we first look at some benefits of security in virtualization and review the main security issues and what causes them. Then we introduce some security tools in the area of virtualization.

II. VIRTUALIZATION SECURITY BENEFITS

One of the main features of virtualization is the isolation between Virtual Machines and their execution environments. This feature makes it possible to have multiple guest operating systems in one host machine and each operating system (OS) runs its own programs in an isolated environment, thus the weaknesses of the programs in one guest OS will not harm the others. Virtualization also

has capabilities of recovering the systems to a normal state after any attacks.

The followings are some of virtualization security advantages [3][4]:

- **Better and faster recovery after attacks**

In case of attacks a compromised machine can be immediately restored to a good snapshot which this process is faster and easier than a physical server. Furthermore a copy of a compromised machine can be cloned for later analysis [4].

- **Patching safer and more effective**

Virtualization makes it possible to revert to a previous state if a patch is unsuccessful, making it more likely to install security patches. You can also make a clone of a running server and test the security patches on it [4].

- **Cost effective security devices**

Some security mechanisms and tools like intrusion detection and prevention systems and other security related appliance can be used more cost effective, because we can put them into a Virtual Machine (VM) instead of a physical server [4].

- **External monitoring**

Since VMs run on shared hardware resources, it allows detecting malicious activities and programs outside the VM, unlike the physical installation of OS on a host, which requires an antivirus. The

Hypervisor can monitor VMs and detects anomalies [5].

- **A safe place for testing malware**

A virtual machine can be suitable environment to test and evaluate malwares. Since VMs can be easily cloned, we can merely get a copy of a VM and test the malware. Although there are some malwares that are able to hide and disable some of their functionalities when they run on a virtual environment [5].

III. VIRTUALIZATION SECURITY CHALLENGES

We divide virtualization security issues into four categories, based on where does that particular vulnerability originate. Whether that vulnerability is from guest VM, host machine and VM Monitor (VMM) the security issues is an attack from outside of the virtualization environment or basically the challenge is a management problem [6].

A. Guest VM Security Challenges

In a virtualized environment multiple guest VMs can reside in a single host machine. Thus, these VMs actually run on a shared physical system, which causes some issues. The followings are security vulnerabilities related to the guest machine.

- **VM Hopping**

It happens when an attacker from one guest virtual machine gains access to another virtual machine within the same virtualized environment. Typically, after a successful attack, the attacker is able to monitor the resource usage info, modify configuration, delete data and cause confidentiality issues. Upon this attack happens, the Confidentiality, Integrity and Availability triangle is violated. Since in this scenario, the attacker migrate from one guest VM to another, it is also called guest-to-guest or cross-VM attack [6].

- **VM Escape**

All the allocations of the resources and system assets is monitored by VMM. In other word, guest VMs are never allowed to access the host machine without VMM interfering them. But some flaws and weaknesses may

cause a guest OS pass the VMM layer and access to the host machine [1].

If the attacker gains access to the host, consequently he has access to all the host resources including all other guest VMs. There are some types of VM Escape attacks like path traversal which uses command line syntax. VM-chat, VM-cat, VM-ftp and VM Drag-N-Sploit are some tools for communicating between the guest VM and host machine. These tool prove that the isolation between VMs can be violated in some situations [7].

- **Side Channel Attacks**

In side channel attacks, the physical characteristics of hardware like CPU, memory usage and other resources are exploited by the attacker. Because VMs in a virtual environment run on the same hardware, this attack is possible among VMs with shared hardware. These type of attacks requires direct access to the host, therefore they are hard to implement [1]. There are several types of side channel attacks in virtualization like timing attacks, power and electromagnetic analysis attacks, and fault induction attacks. [8].

- **VM Alteration**

Applications that run on a VM depend on infrastructure of virtual machine environment. Therefore these VMs which are running on applications must be trusted and any alteration on the VM will be a threat for the applications [1]. One way to protect VMs against this threat is using digital signature for validating virtual machine files. The signing key should never be placed anywhere it can be compromised and after making any external patches the VM should be resigned [9].

- **VM System Restore**

In the case of attacks or system crashes, system administrators usually restore the VM to the last good configuration. Due to simplicity and quickness, administrators prefer to roll back the system instead of installing new software. But rolling back may cause some security problems and make the system vulnerable. It may re-enable previous users and passwords or reveal the ciphers that were used for data encryption [6].

B. Host Machine or VMM Security Challenges

The followings are security vulnerabilities in the machine that is hosting the virtualized environment can be threatening for all the VMs running on the host machine.

- **Hypervisor Hyper-jacking**

Hypervisor poses some priorities which normal applications don't. In one type of attacks, the attacker tries to take the control of VMM which is running on the host machine. Typically the target of this attack is gaining access to the host machine [6].

- **Unsecure VM Migration**

One of the useful features of virtualization which is widely being used in cloud computing is live migration of VMs between two hypervisors. Even though in some virtualization technologies, VMs are encrypted for migration but most of the time the content of the VMs are not protected well enough. Some vulnerabilities have been seen on Xen and VMWare products [6].

In a project, they have managed to modify the memory of a VM during live migration [10]. They have developed a tool named Xensploit that is able to perform *a man in the middle* (MiTM) attack in live migration. To mitigate the probability of this attack, performing mutual authentication between the source and destination VMM can be done. Also using virtual network or a separate and secure physical network can be helpful [11]. An improved version of virtual Trusted Platform Modules (vTPM) protocol has been proposed for secure migration of VMs [12].

- **Resource Allocation**

As we mentioned earlier, the VMM is responsible for allocating system resources among the VMs and any resource usages must be intercepted by VMM. If an attacker takes control over the resource allocation, he can take most of the resources for one VM causing the entire virtual environment goes out of service and some type of *denial of service* attack happens [11].

C. External Security Challenges

In previous cases, malicious activities originated within the virtual environment,

either guest or host machine or VMM. But a virtual environment is also vulnerable to external threats. In this section we look at vulnerabilities that can be used by remote attackers.

- **Rootkit Attacks**

Rootkits are malware that are able to be present in a computer system without being detected and be hidden to the main parts of the system. Rootkits can be used by a remote attacker in different layers of virtualization [1]. For example, Blue-pill is an x86 architecture based virtualization rootkit that targets Microsoft Windows Vista. This rootkit is able to run inside an operating system in a virtual machine and take control the computer and act as a hypervisor and be an access point for other malwares [13].

- **Malicious Code Injections**

There are different types of vulnerabilities in software that might cause a malicious code injection be possible. For code injection, buffer overflow and accepting command line inputs are common. In these attacks, attacker tries to penetrate to VM and inject a malware code in different levels of virtualization [1].

D. Management Security Challenges

Cloud computing with demand on different types of services like Software as a Service (SaaS) and Infrastructure as a Service (IaaS), makes the management of virtualization environment and virtual machines very challenging and cause some security problems such as the followings.

- **VM Mobility**

VM mobility in cloud computing lets users importing a customized VM image into the infrastructure service. Since the content of VM can be transferred, this may lead to spreading the miss configurations and make sensitive data vulnerable. As mentioned previously in unsecure VM migration, this can cause a man in the middle attack [6], [14].

- **VM Sprawl**

Because creating new VMs can be easily done in couple of minutes, after a while there will be a lot of VMs with different types without proper IT management. VM Sprawl

is one of the biggest issues that data centres are facing. As the number of VMs increases, it makes the defining of rules and access permissions more complex and some rules might be overlooked. In these situations, service providers must ensure security of the services and the users keep their VMs secure and up to date [6], [15].

A management system has been proposed for managing virtual machines that allows to control the access to the versions of VMs and filtering and checking the integrity of VM file [16].

IV. VULNERABILITY ASSESSMENT AND PENETRATION TESTING

Both penetration testing and vulnerability assessment are for testing the security and identify the weak parts of a system, but there is a difference between these two. During vulnerability assessment usually, the computer systems are scanned by some tools to detect the vulnerable areas of that systems while penetration testing goes deeper and during its process they actually perform a real attack to see how the system work under a real attack and a report is created that specifies whether the attack was successful or not and it may contain details about the attack.

There are different types of penetration testing and we can categorize them based on their scope (attack by an insider or an external source) or what an organization wants to test. Generally, there are two approaches in penetration testing, Black-box and White-box. The difference of these two is the amount of information that the tester knows about the system [2].

A. Black-box testing

In this type of penetration testing which is also called “external testing” or “remote testing”, the tester has no prior knowledge about the infrastructure by deploying the number of real-world attack techniques. For example the tester will be provided with only the website or network IP address of organization [2].

B. White-box testing

In White-box testing, the tester has prior knowledge of some components of system like details of operating system, network IP address scheme, application code, and

sometimes even the passwords. The main goal of this testing is to verify the integrity of organization network and reduce the risk from internal attacks [2].

C. Virtualization Security Assessment Tools

There are many tools and software for security assessment and penetration testing which we can use for virtualization and other environments. We can consider a hypervisor like an operating system with some services and open ports running on a network, in this case there lots of tools which can be used to assess the hypervisor. Some tools are needed to run from a guest VM in a hypervisor.

• V.A.S.T.O and Metasploit

Metasploit is not just a vulnerability assessment tool but also a penetration testing framework for exploring vulnerabilities and exploiting them. Metasploit contains lots of modules for security assessments and attack simulations. Performing real attacks typically includes discovering vulnerabilities by some scan tools and finding appropriate attack tools for them which can be complex for many testers whose do not have enough experience in this field. The goal of Metasploit is to facilitate this process [17].

V.A.S.T.O is a penetration testing tool specific to virtualization, it has a set of modules that can be added to Metasploit framework. V.A.S.T.Os modules are mostly for VMWare and Xen products. Each module is for performing an assessment scan or an attack. The followings are some of the important modules of V.A.S.T.O [18]:

1. Abiquo_guest_stealer: Performing path traversal attack to escape to the host machine in Abiquo.
2. Abiquo_poison: Sniffing and performing MiTM attack in Abiquo communications.
3. Vmware_guest_stealer: Path traversal attack in VMWare.
4. Vmware_login: Performing brute-force attack to login to a VMWare server.
5. Vmware_lurker: Code execution during a MiTM attack in VMWare.
6. Vmware_version: For fingerprinting and extracting the details of any VMWare server.

For some attacks, multiple modules from V.A.S.T.O or Metasploit's own modules may be needed.

- **VM-Informer**

Unlike V.A.S.T.O which lets the tester to select the penetration test type, VM-Informer assess the security of virtual environment based on security policies and is not developed as an intruder's point of view. Policies are basically security benchmarks which can be modelled according to the requirements. After scanning the environment, it provides a report that identifies the security and insecurity of the environment. VM-Informer audits the following vulnerabilities [18]:

1. Miss configuration
2. Lack of security patches
3. Improper network scheme
4. Weakness in management layers

- **Nessus**

Nessus is one of the vulnerability assessment tools which is able to scan multiple host at the same time and evaluate the scan result with known dynamic vulnerability databases. According to Nessus developer, its aim is to be a "free, powerful, up- to-date and easy to use remote security scanner". The main part of Nessus is its plugins, written in either C language or NASAL (a script language specific to Nessus). Nessus can automatically scan the hosts and thus it is a useful tool when there are lot of servers and hosts. Some of the Nessus plugins not only detect the system vulnerabilities, they also provide some instruction for remediation. Nessus also let its users to add their own plugin which are written in NASAL.

Nessus can be used to discover vulnerabilities like DoS, code execution, buffer overflow, VM escape [19], [20]. For vulnerability assessment of VMWare with Nessus there is a capability that let you login with SOAP API which gives the tester more information about the virtualization environment and its vulnerabilities [19], [20].

- **Ettercap**

Ettercap is a multipurpose network sniffer/interceptor/Logger for LAN networks. When it lands on a network switch, it is able

to see all the communications are being passed by the switch and exploits them. Ettercap can be used for multiple types of man middle attack. It has some features that can be used during the attack [21]:

1. Character injection
2. Packet filtering
3. Automatic password collection for many common network protocols
4. SSH1 support
5. HTTPS support
6. PPTP suite
7. Kill any connection

In virtualization assessment this tool can be used to sniff and manipulate the messages sent between management client and hypervisor management API [20].

- **Hydra**

Hydra is a tool for password cracking using brute-force attack. A brute-force attack consists of an attacker trying many passwords with the hope of eventually guessing it correctly. Hydra supports many online services like POP3, HTTP, IMap and etc. In Virtualization Hydra can be used to test the brute force attack on the password authentication by examining whether there is any prevention mechanism in place [22], [20].

- **NMap**

NMap is a tool for scanning a range of IP addresses, identify active systems, discovering the open ports and what operating systems are running on those systems. Like other scanning tools NMap can be used by network administrators to find the vulnerabilities in the network or by an attacker for malicious activities. Typically, in security assessment of an environment first of all we need to gather information about the system we are trying to examine. We need to know what services are running on the system or the hypervisor and in what version in order to find proper vulnerabilities and methods to exploit them. NMap is of the best tools that can be used for information gathering of penetration testing [23].

- **TCP-Replay**

In the *man in the middle* attacks, captured packets can be used for a replay attack. A replay attack consists of sniffing a

communication between two parties and after capturing sensitive packets like password or password hashes it uses this packet to authenticate to the system later. In virtualization environment if a deletion of VMs are allowed, this environment probably is vulnerable to replay attack [20], [24].

- **Cain&Able**

Cain&Able is a tool for performing ARP-Spoofing which is also a MiTM attack. The aim of this attack is monitoring the packets that are sent to a machine or sent out by the machine. This tool can redirect the communication between two machines to be passed from the attacker's machine first and then goes to its destination. In virtualization this tool can be used to assess the security of communication between management client and hypervisor management API [20].

From these tools, some of them like V.A.S.T.O or VM-Informer are specific to virtualization, but most of them are general tools which do have applications for virtualization environments as well. Another difference is that for example V.A.S.T.O and Metasploit are penetration testing tools which are able to perform actual attacks and some manual steps are need using them while Nessus is a scan tool that detects vulnerabilities of a system. Regardless of what is the type of the tool and how can it assess a particular vulnerability we just consider a tool is able to assess a vulnerability, whether it can just detect the vulnerability or it is able to exploit them too. Table 1 shows what tools related to what vulnerabilities.

Table 1: Virtualization Security Assessment Tools and Vulnerabilities

| | V.A.S.T.O | NESSUS | CAIN&ABLE | TCP-REPLAY | HYDRA | ETTERCAP |
|--------------------|-----------|--------|-----------|------------|-------|----------|
| VM Escape | * | * | | | | |
| VM Hopping | | * | | | | |
| MiTM | * | * | * | * | | * |
| Denial of Service | * | * | | | | |
| Code Execution | * | * | | | | |
| Unauthorized login | * | * | | * | * | |
| Rootkits | * | | | | | |

Table 2: V.A.S.T.O for Penetration Testing of Virtualization

| | VMWARE | XEN | ABQUO | ORACLE-VM |
|--------------------|--------|-----|-------|-----------|
| VM Escape | * | | * | |
| VM Hopping | | | | |
| MiTM | * | | * | |
| Denial of Service | * | | | |
| Code Execution | * | * | | * |
| Unauthorized login | * | * | | |
| Rootkits | * | | | |

Table 3: Nessus for assessment of virtualization products

| | VMWARE | XEN | K.V.M |
|--------------------|--------|-----|-------|
| VM Escape | * | * | * |
| VM Hopping | * | | |
| MiTM | * | * | |
| Denial of Service | * | * | * |
| Code Execution | * | | * |
| Unauthorized login | * | | |
| Rootkits | | | |

Table 2 and 3 show the relation of V.A.S.T.O and Nessus for assessment of vulnerabilities based on virtualization products. The rest of the tools are kind of used for assessment of networks or can be used in combination to perform penetration testing.

V. DISCUSSION

To make the systems and environments secure for small companies that do not want to spend too much for security, the security assessment could be only exploring vulnerabilities, take a report and try to fix the issues based on their priorities. Tools like Nessus would be helpful for such purposes, because it is easy to use, and you can check your systems periodically, and it also provides useful information for remediation of the issues. VM-Informer is also can be used in these situations. But in companies that security has a big role they may want to go even deeper and find out too much about their systems, how their systems can be a target of attacks, how they react to that attacks and how much faster they can recover after. For this job, someone that has enough experience to

perform the penetration testing is needed and the process needs knowledge about the system and tools.

Some of the tests are tricky and most of the time the tester need to use a bunch of tools in combination. For penetration testing a tool like NMap can be used to scan the services and ports, the operating systems version and other information at information gathering phase. Beside Metasploit sniffing modules Ettercap or Wireshark are useful tools for sniffing and checking the hypervisor's network connections. Metasploit has also some modules that can be used for password cracking as well as Hydra itself.

In conclusion, as shown in Table 4, for a simple assessment Nessus or VM-Informer can be run to check the virtualization environment to find out what vulnerabilities are present, this scan can be used as a first step of a penetration testing operation too. We can use the information of the vulnerabilities to search and find appropriate tools to perform penetration testing.

Table 4: Tools for Vulnerability Assessment and Penetration Testing of Virtualization

| RECOMENDED TOOLS | |
|--------------------------|--|
| Vulnerability Assessment | Nessus, VM-Informer |
| Penetration Testing | Nessus, VM-Informer, Metasploit, Ettercap, Hydra, Cain&Able, ... |

VI. CONCLUSION

Although the virtualization is very practical in data centres and cloud computing, but it is necessary to assess its impacts on security components. In this paper we have tried to evaluate virtualization technology with security perspectives. Table 5 presents our reviewed virtualization security benefits and vulnerabilities and some recommended tools which can be used in security assessment of virtualization.

Table 5: Summary of virtualization security benefits, challenges and tools

| | | | |
|----------------------------------|---|----------|--|
| Virtualization Security Benefits | <ul style="list-style-type: none"> Better and faster recovery after attack Patching safer and more effective Cost effective security devices e.g. virtual IDS External monitoring by VMM VM is a safe place for testing malwares | | |
| Virtualization | <table> <tr> <td>Guest VM</td><td> <ul style="list-style-type: none"> VM Hopping </td></tr> </table> | Guest VM | <ul style="list-style-type: none"> VM Hopping |
| Guest VM | <ul style="list-style-type: none"> VM Hopping | | |

| | | |
|--|---|---|
| Security Challenges | Challenges | <ul style="list-style-type: none"> VM Escape Side Channel Attacks VM Alteration VM System Restore |
| | Host VM and VMM Challenges | <ul style="list-style-type: none"> Hypervisor Hyper-jacking Unsecure VM Migration Resource Allocation |
| | External Challenges | <ul style="list-style-type: none"> Rootkit Attacks Malicious Code Injections |
| | Management Challenges | <ul style="list-style-type: none"> VM Mobility VM Sprawl |
| Virtualization Security Assessment tools | <ul style="list-style-type: none"> V.A.S.T.O and Metasploit VM-Informer Nessus Ettercap Hydra NMap TCP-Replay Cain&Able | |

VII. REFERENCES

- [1] K. Pooja, R. Nagpal, and T. P. Singh, A Survey on Virtualization Service Providers , Security Issues , Tools and Future Trends, *Int. J. Comput. Appl.*, vol. 69, no. 24, pp. 36–42, 2013.
- [2] N. Shrestha, Security Assessment via Penetration Testing: A Network and System Administrator's Approach, Master's thesis, Univ. OSLO, 2012.
- [3] E. R. Rasmussen, Reducing IT Costs and Increasing IT Efficiency by Integrating Platform-Virtualization in the Enterprise, *Univ. Oregon.*, vol. 1277, no. February, 2009.
- [4] R. Randell, Virtualization Security and Best Practices, *RSA Secur. Conf.*, 2006.
- [5] G. Obasuyi and A. Sari, Security Challenges of Virtualization Hypervisors in Virtualized Hardware Environment, *J. Commun. Netw. Syst.*, no. July, pp. 260–273, 2015.
- [6] A. Mahjani, Security Issues of Virtualization in Cloud Computing Environments, Master's thesis, Luleå Univ. Technol., 2015.
- [7] S. Zahedi, Virtualization Security Threat Forensic and Environment Safeguarding, *Linnéus Univ.*, Degree project, 2014.
- [8] A. Yu and D. Brée, Side channel Attack-Survey Joy, *Inf. Technol. Coding*, vol. 1, no. 4, pp. 54–57, 2004.

- [9] J. Kirch, Virtual Machine Security Guidelines Version 1.0, *The Centre for Internet Security (CIS)*, 2007.
- [10] J. Oberheide, E. Cooke, and F. Jahanian, Empirical exploitation of live virtual machine migration, *Proc. BlackHat DC*, no. VMM, 2008.
- [11] A. Tayab et al., Virtualization and Information Security A Virtualized DMZ Design Consideration Using VMware ESXi 4.1, Unitec Institute of Tech, New Zealand, vol. 2, p. 89, 2012.
- [12] X. Wan, X. Zhang, L. Chen, and J. Zhu, An improved vTPM migration protocol based trusted channel, *Int. Conf. Syst. Informatics, ICSAI 2012*, no. Icsai, pp. 870–875, 2012.
- [13] U. Gurav and R. Shaikh, Virtualization – A key feature of cloud computing, *Int. Conf. Work. Emerg. Trends Technol.*, no. Icwet, pp. 227–229, 2010.
- [14] K. Benzidane, S. Khoudali, and A. Sekkaki, Secured architecture for inter-VM traffic in a Cloud environment, *2nd IEEE Lat. Am. Conf. Cloud Comput. Commun. LatinCloud 2013*, pp. 23–28, 2013.
- [15] H. Tsai, N. Chiao, R. Steinmetz, and T. U. Darmstadt, Threat as a Virtualization's Impact on Cloud Security, *28th IEEE Int. Conf. Data Eng.*, no. February, pp. 32–38, 2012.
- [16] J. Wei, X. Zhang, G. Ammons, V. Bala, and P. Ning, Managing security of virtual machine images in a cloud environment, *Proc. 2009 ACM Work. Cloud Comput. Secur. - CCSW '09*, no. Vm, p. 91, 2009.
- [17] B. Greenwood, An Introduction to Metasploit Project for the Penetration Tester, SANS Institute report, <https://cyber-defense.sans.org/resources/papers/gsec/introduction-metasploit-project-penetration-tester-107151> [Accessed: June 2018].
- [18] S. Chauhan, Hacking VMware with VASTO, Infosec Inst. report, <http://resources.infosecinstitute.com/virtualization-security/#gref>. [June 2018].
- [19] J. Mitchell, Proactive Vulnerability Assessments with Nessus, SANS Inst. report, <https://www.sans.org/reading-room/whitepapers/auditing/paper/78>. [Accessed: June 2018].
- [20] A. Thongthua and S. Ngamsuriyaroj, Assessment of hypervisor vulnerabilities, *Proc. - Int. Conf. Cloud Comput. Res. Innov. 2016*, pp. 71–77, 2016.
- [21] D. Norton, An Ettercap Primer, SANS Inst. report, <https://www.sans.org/reading-room/whitepapers/tools/paper/1406> [Accessed: June 2018].
- [22] C. Yiannis, Modern Password Cracking : A hands-on approach to creating an optimised and versatile attack, Inf. Secur. Group, R. Holloway, Univ. London, no. May, 2013.
- [23] T. Corcoran, *An Introduction to NMAP*, SANS Inst. report, <https://www.sans.org/readingroom/whitepapers/tools/paper/72> [Accessed: June 2018].
- [24] A. Hussain, Y. Pradkin, and J. Heidemann, Replay of malicious traffic in network testbeds, *IEEE Int. Conf. Technol. Homel. Secur. HST 2013*, pp. 322–327, 2013.

Safeguarding Malaysia's Cyberspace against Cyber Threats: Contributions by CyberSecurity Malaysia

Fazlan Abdullah¹, Nadia Salwa Mohamad², and Zahri Yunos³

^{1,2,3} CyberSecurity Malaysia, Seri Kembangan, Malaysia

fazlan@cybersecurity.my, nadia.salwa@cybersecurity.my, zahri@cybersecurity.my

Abstract - The world today is becoming dependent on Information and Communication Technology (ICT). Cyber threats on ICT infrastructures can lead to catastrophic damage and disruption, hence an effective information security policy framework is vital in securing the Critical National Information Infrastructure (CNII). Malaysia has implemented the National Cyber Security Policy (NCSP) to safeguard Malaysia's CNII against cyber threats. The implementation of NCSP initiatives requires the commitment and involvement of multiple stakeholders to ensure continuous momentum. Thanks to the implementation of the NCSP initiatives, Malaysia's commitment and effort in ensuring resilience against cyber threats has been recognized at the international level.

KEYWORDS - Critical National Information Infrastructure (CNII), Cyberattacks, Cyber Threat, Cyber Security, Cyber Security Policy

I. INTRODUCTION

The high dependency on the use of Information and Communication Technology (ICT) for social, political and economic activities makes many nations around the world vulnerable to the ever-increasing range of cyber threats. These threats can jeopardize every level of society and industry, from public users who use ICT equipment to the Critical National Information Infrastructure (CNII) which is dependent on the ICT systems for the operation of their infrastructure, for example in the banking, government, energy, water and telecommunications sector.

Interdependencies between these infrastructures have raised concerns that successful cyberattacks may have serious cascading effects on others, resulting in potentially disastrous impact. Therefore, it is necessary to have a strategy at the national level for protecting CNII from cyber threat activities.

II. RELATED WORK

A. Critical National Information Infrastructure (CNII)

Advancement in the use and dissemination of ICT are seen as closely connected to the notion of critical infrastructure protection. CNII are the foundation of a nation's economic, political, strategic and socio-economic activities [1][2][3]. In recent years,

CNII has become progressively more dependent on ICT, as there exist infrastructure interdependencies of CNII sectors [4]. In most cases, the ICT system forms the backbone of a nation's critical infrastructure (e.g. electrical grid), which means that a major security incident in a particular system could have significant impact on the reliability and safety of the operations of the physical systems dependent on it [5].

Interdependency is a bidirectional relationship between infrastructures, through which the state of each infrastructure is influenced by, or correlated to the state of the other. Many stakeholders are concerned with cyberattacks against interdependent critical infrastructures, such as telecommunications, power distribution, transportation, financial services and essential public utility services.

B. Theoretical Concept of Cyberattacks Targeting CNII

It is important to understand the infrastructure of computer networks that are at risk, especially those which support CNII operational functions [6]. Threats may be in the form of attacks launched using, or against, computer networks. Cyberattacks on CNII are possible, whereby the motives, resources and willingness to conduct operations of different kinds against specific targets are fundamental [7]. If perpetrators follow the lead of hackers, they theoretically have the capability to use ICT to conduct cyberattacks against specific targets. The cyber world, which encompasses

computer-related technologies such as the Internet and World Wide Web, gives perpetrators access and freedom over vast geographic areas. Among the most advanced countries, the US Department of Defense has placed cyber threats as the top national security threat to the United States.

There is a great deal of concern regarding serious attacks against CNII [8]. CNII is a complex, interconnected system with a vital role in underpinning our economy, security and way of life. CNII facilities pose high-value targets, which, if successfully attacked (physically or cyber-wise), have the potential to disrupt the normal rhythm of society, cause public fear and intimidation, and generate substantial publicity [9]. The CNII in a given country is often an attractive target for perpetrators owing to the large-scale economic and operational damage that can potentially occur with a major failure. In this case, the CNII's industrial control system is the potential target.

CNII organizations that provide critical services have long used a control system commonly known as Supervisory Control and Data Acquisition (SCADA) for gathering real-time data, controlling processes and monitoring equipment from remote locations [10]. SCADA serves to monitor and control the delivery of critical services, such as power, waste treatment, and nuclear, transport and water systems. These systems are frequently unmanned and accessed remotely by engineers via telecommunication links. Typically, SCADA systems are closed operating environments (or stand-alone systems). However, new research indicates a tendency for systems to move towards open standards (or networked architectures), such as Ethernet, TCP/IP and web technologies where vulnerabilities are more widely known [11].

A number of existing case studies represent the incidence of terrorist attack acts on CNII. One captured al-Qaeda computer reportedly contained engineering and structural features of a dam downloaded from the Internet [12]. In another case, it was found that al-Qaeda operators studied software and programming instructions for digital switches that run power, water and transportation grids. SCADA systems have also been accessed by terrorist and extremist groups to gather information on potential targets.

Therefore, it can be concluded that protecting CNII organizations against cyberattacks is deemed critical to a nation. The reason is that the destruction or disruption of ICT systems that provide critical services could significantly impact economic strength, image, defence and security, a government's functioning capabilities, and public health and safety. This observation is relevant, because CNII organizations are likely targets due to the high degree of interdependency between these critical sectors. Besides, the impact would be much greater and wider compared to non-CNII organizations. As a result of weaknesses or vulnerabilities in the SCADA system within CNII organizations, adversaries may conduct terrorist activities by utilizing the cyberspace to carry out cyberattacks on CNII facilities.

C. Cyberattacks on CNII: Case Studies

The Stuxnet attack against the Iranian Nuclear program demonstrates the impact that a sophisticated adversary with detailed knowledge of process control systems can have on critical infrastructure [13]. Stuxnet is believed to have destroyed 984 centrifuges at Iran's uranium enrichment facility in Natanz [14]. The attack alarmed the world towards vulnerabilities in the highly sophisticated facility and industrial control system.

Another cyberattack that has attracted the world's attention and raised concerns regarding e-banking systems is the Bangladesh Bank Heist that was reported in February 2016. The Bangladesh Bank was compromised through firewall exploitation, which facilitated a breach in the Society for Worldwide Interbank Financial Telecommunications' (SWIFT) Alliance Access Software for making payment instructions [15]. The US Central Bank approved five of the payment instructions and made the payments to accounts in Sri Lanka and Philippines – including \$81 million to four accounts in the names of individuals [16]. Investigation is ongoing and no arrests have been made despite the US Federal Bureau of Investigation, Interpol, Bangladesh police and authorities in the Philippines working on this case [17]. This cyberattack on the banking industry triggered cross-border action in safe audit procedures, security and architecture of the SWIFT network, as well as personnel negligence with e-banking systems and Standard Operating Procedures (SOP).

Global ransomware attacks are increasing as reported by Europol [18]. The most recent cyberattack, WannaCry, has affected hundreds of thousands of computers by exploiting vulnerabilities in Microsoft's Windows XP software and creating havoc around the world [19]. WannaCry is a dangerous combination of two malicious software components: a worm and a ransomware variant [20]. Hospitals, companies, universities and governments across at least 150 countries were hounded by a cyberattack that locked computers and demanded ransom [21]. CyberSecurity Malaysia's MyCERT department issued alerts and advisories on the WannaCry Ransomware threat [22] [23] [24].

The rise in planned cyberattacks by hacktivists on Malaysia with high damage potential for interdependent networks and information systems across the country has demanded high attention be paid to CNII protection initiatives. The most remembered cyber threat by hacktivists was the coordinated cyberattack called "Operation Malaysia" in 2011 by the Anonymous group, which conducted DDOS attacks on Malaysia's government websites in protest of Malaysia's blocking of certain websites [25] [26].

III. METHODOLOGY

The methodology used for this research is qualitative and the approach used is literature review from secondary sources. There will be no numeric data or quantitative data produced. Due to limited literature with regards to cyber incidents happening around the globe, the journal also looks at newspaper article for information and references.

IV. DISCUSSION

A. International Telecommunication Union (ITU) National Cyber Security Guideline

In this rapidly changing and sophisticated cyber-threat environment, all states and organizations need to have comprehensive, flexible and dynamic cybersecurity strategies. A national cybersecurity strategy is a plan of action to increase the security, resilience and self-reliance of national infrastructures in delivering services against cyber threats.

In 2011, the International Telecommunication Union (ITU) published

the ITU National Cybersecurity Strategy Guide as a reference model for national strategy elaboration. The ITU, a specialized agency of the United Nations (UN) for ICT, is an organization based on public-private partnership with current membership of 193 countries and 800 private sector entities and academic institutions.

Cyber security has been at the top of the UN agenda, for it is crucial to the socio-economy of the global community. UN has issued resolutions on five (5) cybersecurity matters: Combating Criminal Use of ICTs (A/RES/55/63 and A/RES/56/121), Culture of Cybersecurity (A/RES/57/239), Critical Infrastructure (A/RES/58/199) and Global Culture of Cybersecurity (A/RES/64/211) [27].

Based on the ITU National Cybersecurity Strategy Guideline, ten (10) elements are the main features of a holistic, multi-stakeholder and strategy-led cybersecurity program (Table 1).

Table 1: Elements of ITU National Cyber Security Guide

| No. | Element of ITU National Cyber Security Guide |
|-----|--|
| 1 | <u>Top Government Cybersecurity Accountability</u> Top government leaders are accountable for devising a national strategy and fostering local, national and global cross-sector cooperation |
| 2 | <u>National Cybersecurity Coordinator</u> An office or individual overseeing cybersecurity activities across the country |
| 3 | <u>National Cybersecurity Focal Point</u> A multi-agency body that serves as a focal point for all activities dealing with the protection of a nation's cyberspace against all types of cyber threats. |
| 4 | <u>Legal Measures</u> Typically, a country reviews and, if necessary, drafts new criminal laws, procedures, and policies to deter, respond to and prosecute cybercrime. |
| 5 | <u>National Cybersecurity Framework</u> Countries typically adopt such framework that defines minimum or mandatory security requirements on issues such as risk management and compliance. |
| 6 | <u>Computer Incident Response Team (CSIRT)</u> A strategy-led program that contains incident management capabilities with national responsibility. The role is to analyse cyber threat trends, coordinate responses and disseminate information to all relevant stakeholders. |
| 7 | <u>Cybersecurity Awareness and Education</u> A national program should exist to raise awareness about cyber threats. |
| 8 | <u>Public-Private Sector Cybersecurity Partnership</u> Governments ought to form meaningful partnerships with the private sectors |

| | |
|----|---|
| 9 | <u>Cybersecurity Skills and Training Program</u> A program that should help train cybersecurity professionals |
| 10 | <u>International Cooperation</u> Global cooperation is vital due to the transnational nature of cyber threats. |

The ITU National Cybersecurity Strategy Guideline is centred on matters that all countries should consider as part of the national cybersecurity strategy, such as national values, need and threat variance, national capabilities, culture and national interest. Being aware of the multi-stakeholder aspect of cybersecurity, ITU has thus developed the ITU Global Cybersecurity Agenda (GCA) -- a cross-border framework for international cooperation in cybersecurity.

GCA boosts cooperation between members and partners to prevent duplication in strategic initiative implementation. GCA recommends 5 pillars or areas in cybersecurity activities within ITU, as stated in Table 2.

Table 2: Global Cybersecurity Agenda (GCA) Pillars

| Pillar | Areas in Cybersecurity Activities |
|----------|-----------------------------------|
| Pillar 1 | Legal Measures |
| Pillar 2 | Technical and Procedural Measures |
| Pillar 3 | Organizational Structures |
| Pillar 4 | Capacity Building |
| Pillar 5 | International Cooperation |

B. Global Cybersecurity Index Framework by ITU

Cybersecurity ranges over a broad spectrum of fields across several industries and sectors. ITU, a specialized agency of the United Nations for ICTs, is committed to connecting nations, and protecting and supporting the fundamental rights of a person to communicate. The Global Cybersecurity Index (GCI) is a survey for measuring the commitment of Member States to cybersecurity. GCI is based on the ITU GCA, and is a framework for international cooperation to enhance confidence and security in the current information society. GCA is constructed upon the five (5) strategic areas of GCI: Legal Measures, Technical and Procedural Measures, Organizational Structures, Capacity Building and International Cooperation [27].

GCI is included under Resolution 130 (Rev. Busan, 2014), with the first survey held

in 2013-2014 in partnership with ABI Research. A new survey was carried out in 2017 using an enhanced reference model as a result of the extensive participation and collaboration of experts, industry stakeholders, contributing partners and GCI partners [28].

The objective of the GCI initiative is to assist member states identify areas for improvement in the field of cybersecurity, take constructive action for ranking as well as raise the countries' commitment to cybersecurity. Table 3 explains briefly the five (5) strategic pillars and sub-pillars of GCI [28].

Table 3: Strategic pillars and sub-pillars of GCI

| Strategic Pillars | Sub-Pillars |
|--|--|
| <u>Legal Measures</u> Existence of legal institutions and frameworks dealing with cybersecurity and cybercrime | <ul style="list-style-type: none"> • Cybercriminal legislation • Cybersecurity regulation • Cybersecurity training |
| <u>Technical and Procedural Measures</u> Existence of technical institutions and frameworks dealing with cybersecurity | <ul style="list-style-type: none"> • National CIRT • Government CIRT • Sectoral CIRT • Standards for organizations • Standards and certification for professionals • Child online protection |
| <u>Organizational Structures</u> Existence of policy coordination institutions and strategies for cybersecurity development at the national level | <ul style="list-style-type: none"> • Strategy • Responsible agency • Cybersecurity metrics |
| <u>Capacity Building</u> Existence of research and development, education and training programs; certified professionals and public sector agencies fostering capacity building | <ul style="list-style-type: none"> • Standardization bodies • Good practices • R&D programs • Public awareness campaigns • Professional training courses • National education programs and academic curriculums • Incentive mechanism • Homegrown cybersecurity industry |
| <u>International Cooperation</u> Existence of partnerships, cooperative frameworks and information sharing | <ul style="list-style-type: none"> • Inter-state cooperation • Multilateral agreements |

| | |
|----------|---|
| networks | <ul style="list-style-type: none"> • International forum participation • Public-private partnerships • Inter-agency partnerships |
|----------|---|

C. CNII Protection Framework in Malaysia

The revolution of information and interdependency of ICT infrastructures has increased the risk of various new vulnerabilities and dynamic threats to critical infrastructures. The Government of Malaysia is deliberately adopting ICT as a key enabler for socio-economic development. Thus, adopting an integrated and broad approach to protect critical infrastructure is necessary.

NCSP development started in 2005 and was accepted by the government for implementation in 2006. The NCSP aims to develop and establish a comprehensive program and framework to ensure the effectiveness of information security controls over critical assets and that the CNII is protected up to a level that is commensurate to the risks faced. Key areas considered during policy development are legislation, technology, institutional, public and private cooperation as well as international engagement.

The policy covers ten (10) CNII sectors identified and defined in the policy (Table 4).

Table 4: Ten (10) CNII Sectors Identified

| | |
|-----------------------------|--------------------|
| National Defence & Security | Water |
| Banking & Finance | Health Services |
| Information & Communication | Government |
| Energy | Emergency Services |
| Transportation | Food & Agriculture |

The NCSP has eight (8) Policy Thrusts (PT) covering the specific areas listed in Table 5.

Table 5: Elements of ITU National Cyber Security Guide

| Policy Thrust (PT) | Initiatives |
|----------------------------|--|
| PT 1: Effective Governance | <ul style="list-style-type: none"> • Centralize coordination of national cybersecurity initiatives. • Promote effective cooperation between public and private sectors. • Establish formal and encourage informal information exchange. |
| PT 2: | <ul style="list-style-type: none"> • Review and enhance Malaysia's |

| | |
|--|--|
| Legislative and Regulatory Framework | <p>cyber laws to address the dynamic nature of cybersecurity treats.</p> <ul style="list-style-type: none"> • Establish progressive capacity building programs for national law enforcement agencies. • Ensure that all applicable local legislation is complementary to and in harmony with international laws, treaties and conventions. |
| PT 3: Cybersecurity Technology Framework | <ul style="list-style-type: none"> • Develop a national cybersecurity technology framework that specifies cybersecurity requirement controls and baselines for CNII elements. • Implement an evaluation/certification program for cybersecurity products and systems. |
| PT 4: Culture of Security and Capacity Building | <ul style="list-style-type: none"> • Develop, foster and maintain a national culture of security. • Standardize and coordinate cybersecurity awareness and education programs across all CNII elements. • Establish an effective mechanism for cybersecurity knowledge dissemination at the national level. • Identify minimum requirements and qualifications for information security professionals. |
| PT 5: Research and Development Towards Self-Reliance | <ul style="list-style-type: none"> • Formalize the coordination and prioritization of cybersecurity research and development activities. • Enlarge and strengthen the cybersecurity research community. • Promote the development and commercialization of intellectual properties, technologies and innovations through focused research and development. • Nurture the growth of the cybersecurity industry. |
| PT 6: Compliance and Enforcement | <ul style="list-style-type: none"> • Standardize cybersecurity systems across all CNII elements. • Strengthen the monitoring and enforcement of standards. • Develop a standard cybersecurity risk assessment framework. |
| PT 7: Cybersecurity Emergency Readiness | <ul style="list-style-type: none"> • Strengthen the national computer emergency response teams (CERTs). • Develop an effective cybersecurity incident reporting mechanism. • Encourage all CNII elements to monitor cybersecurity events. • Develop a standard business continuity management framework. • Disseminate vulnerability advisories and threat warnings in a timely manner. • Encourage all CNII elements to perform periodic vulnerability assessment programs. |

| | |
|------------------------------------|---|
| PT 8: International Cooperation | <ul style="list-style-type: none"> Encourage active participation in all relevant international cybersecurity bodies, panels and multi-national agencies. Promote active participation in all relevant international cybersecurity events, conferences and forums. Enhance the strategic position of Malaysia in the field of cybersecurity by hosting an annual international cybersecurity conference. |
|------------------------------------|---|

D. Malaysia's NCSP Framework and ITU GCI

The elements used in the development of NCSP are similar to the elements recommended by ITU GCI for the development of a national cybersecurity policy. Table 6 compares the elements in both frameworks based on the people, technology and process components. GCI recommends five (5) key areas in the guideline, whilst NCSP identifies eight (8) key areas in policy development.

Table 6: Elements of NCSP and GCI

| Cyber security Policy | Strategic Areas / Pillars | Influencing Factor | Framework |
|-----------------------|---|--------------------|------------------------------------|
| Malaysia's NCSP | Culture of Security & Capacity Building | People | Awareness & Competency Development |
| GCI | Capacity Building | | |
| Malaysia's NCSP | R&D Towards Self Reliance. Cyber Security Emergency Readiness. Cyber Security Technology Framework. | Technology | Technology Development |
| GCI | Technical and Procedural Measures | | |
| Malaysia's NCSP | Legislative & Regulatory Framework | Process | Cyber Laws & Enforcement |
| GCI | Legal Measures | | |
| Malaysia's NCSP | Compliance & Enforcement (Standard) | Process | Security Management |
| GCI | Legal | | |

| | Measures | | |
|-----------------|---------------------------|---------|---------------------------|
| Malaysia's NCSP | International Cooperation | Process | International Cooperation |
| GCI | International Cooperation | | |

E. National Cybersecurity Policy Implementation Progress to Date in Malaysia

Since the policy was approved in 2006, multiple initiatives have been planned under each PT. Moreover, each PT's activities are driven by the respective ministries and government agencies as thrust drivers. The implementation approach of NCSP is to develop self-reliance in technology, develop human capital, monitor the compliance mechanism, evaluate and improve the mechanism, and create a cybersecurity culture. A brief description of the NCSP implementation is given as follows.

PT 1: Effective Governance

Initially, NCSP development and implementation was led by the Ministry of Science, Technology and Innovation Malaysia (MOSTI) with focus on establishing a governance structure and various committees. The committees cover each key aspect, such as policy, content, crisis management, legislation, acculturation and capacity building, and compliance and enforcement. To oversee the implementation of the NCSP thrusts and strategies, the National Cyber Security Coordination Committee (NC3) was formed in 2008.

In 2011, the stewardship of the NCSP was handed over to the National Security Council as the central coordinating body. Subsequently, the high-level e-Sovereignty Committee was established to oversee the overall cybersecurity governance in Malaysia, chaired by the Deputy Prime Minister of Malaysia.

On January 2017, the government of Malaysia established the National Cyber Security Agency (NACSA), which reflects the government's seriousness to address cybersecurity threats in a more coordinated manner.

PT 2: Legislative & Regulatory Framework

In boosting the legislative and regulatory aspects of cybersecurity, Malaysia adopted the

Information Security Legal and Regulatory Framework. A 'Study on the laws of Malaysia to accommodate legal challenges in the Cyber Environment' in 2009 and a 'Feasibility Study on the Cyber Security Standards Act' in 2015 were also conducted. As proposed by the adopted framework, the current legislation including the Computer Crime Act 1997, Communication and Multimedia Act, *Arahan Tetap Keselamatan Kerajaan*, and Evidence Act 1950 have been reviewed and are being amended.

In 2010, the Personal Data Protection Act 2010 and Department of Personal Data Protection were established for the protection and security of personal data. In supporting the law enforcement agencies and regulatory bodies in digital forensic investigation capabilities, CyberSecurity Malaysia's Digital Forensics Labs was established in year 2002. The capacity and capability of the lab was further enhanced with other expertise such as audio forensics, video forensics and closed-circuit television (CCTV) forensics.

PT 3: Cybersecurity Technology Framework

The framework was established for cybersecurity controls to be implemented and enforced based on recommended standards and guidelines. The security controls applied are commensurate with the potential organizational impact due to any security breaches caused by forfeiture of confidentiality, integrity or availability. The ISO 27001 Information Security Management Systems standard was identified as a baseline for compliance under PT 3. On 24th February 2010, the Malaysian Cabinet meeting had decided that CNII agencies shall implement MS ISO/IEC 27001 (Information Security Management System-ISMS) to safeguard and protect organizational data and information [29] [30].

Another initiative implemented under this framework is the Malaysia Common Criteria Certification (MyCC) Scheme, which is aimed to increase Malaysia's competitiveness in quality assurance of information security based on Common Criteria Standard ISO/IEC 15408. The scheme implements a security evaluation and certification program that will facilitate CNII to procure technology with documented assurance. The MyCC Scheme is operated by the Information Security Certification Body (ISCB), a department of Cybersecurity Malaysia, which manages

information security certification. Malaysia became a member of the Common Criteria Recognition Arrangement (CCRA) in 2007. The Government of Malaysia also agreed that the CC Certification would be one of the criteria in the procurement of information technology, especially local systems or products. To date, there are sixty-eight (68) products and systems have been certified under the MyCC Scheme. ITU has credited the establishment of CyberSecurity Malaysia's ISCB department and the establishment of the MyCC Scheme in the GCI 2017 survey report [28].

PT 4: Culture of Security & Capacity Building

The Government of Malaysia has been aware of the need for greater awareness and understanding of cybersecurity issues and for developing a positive cybersecurity culture. Hence, a study entitled National Strategy for Cyber Security Acculturation and Capacity Building was carried out in 2010 to evaluate current national and CNII awareness education programs and campaigns.

To ensure the success of the cybersecurity awareness, acculturation and education programs, coordinated initiatives and efforts have been driven by relevant organizations to increase the level of cybersecurity awareness, best practices and safe use of the Internet across all CNII as well as public elements.

One of the main initiatives is Cyber Security and Awareness for Everyone (CyberSAFE), which is a program that provides awareness for children, youth, parents and organizations. To date, more than 170,000 people have participated in the CyberSAFE Program. Another initiative is the development of the "Guideline to Determine Information Security Professional Requirements for CNII Agencies or Organizations." [31] This guides CNIIs with ensuring their organizations have sufficient trained professional to handle technical and non-technical cybersecurity issues within their organizations.

In addition, CyberSecurity Malaysia has collaborated with local universities in cybersecurity tertiary programs, such as Master of Cyber Security in collaboration with Universiti Kebangsaan Malaysia (UKM), Master of Protective Security Management with International Islamic University Malaysia (IIUM) and Degree of Cyber Security and Cyber Security Technology with

the National Defence University (Universiti Pertahanan Nasional Malaysia - UPNM).

In the ITU GCI 2017 survey, Malaysia was ranked second in the Asia Pacific region, scoring a perfect 100 on capacity building as a result of Malaysia's initiatives. The ITU GCI 2017 report also cited CyberSecurity Malaysia's professional training programs via higher education institutions in Malaysia as well as its CyberGuru website, dedicated to professional security training as contributing to the capacity building score in the survey. The professional training programs and the CyberGuru website are managed by CyberSecurity Malaysia's Cyber Security Professional Development (CSPD) department.

PT 5: Research & Development towards Self-Reliance

The NCSP implementation also focuses on Research & Development towards Self-Reliance through Policy Thrust 5. Led by MIMOS Berhad, an organization under MOSTI, MIMOS spearheaded the development of the National Cyber Security Research and Development Roadmap for Self-reliance in cybersecurity technologies.

The initiative of this thrust is to identify and monitor information security-related research and development projects. Among research projects and cooperation for supporting this thrust are CyberSecurity Malaysia's MyCERT National Malware Research Centre, CyberCSI, Cryptography Research, SCADA Research Lab collaboration between CyberSecurity Malaysia, and the cybersecurity industry.

Through research and development efforts, CyberSecurity Malaysia has successfully developed services such as Cyber999 for handling cybersecurity incidents and the MyCyberSecurity Clinic for data recovery and sanitation.

PT 6: Compliance & Enforcement

On 24 February 2010, the government of Malaysia agreed for all CNIIs to implement and undergo certification based on MS ISO/IEC 27001 Information Security Management System (ISMS) standards within 3 years. A task force led by the National Security Council and comprising regulators and government bodies overseeing the CNIIs, was formed to ensure compliance to this directive. To date, more than one hundred

thirty-eight (138) CNIIs have been ISMS-certified [29].

PT 7: Cyber Security Emergency Readiness

The establishment of the Computer Emergency Response Teams (CERT) is one of the initiatives to reduce and mitigate cyber threats. Malaysian CERT (MyCERT) was formed on 13 January 1997 to facilitate and handle computer security incident responses to emergencies.

In 2008, the National Security Council developed the National Cyber Crisis Management Plan (NCCMP) in order to manage cyber emergencies. NCCMP was later further developed into the National Security Directive No. 24: National Cyber Crisis Management Policy and Mechanism, which was launched in 2013. This directive aims to ensure a high level of preparedness in the face of threats and cyberattacks at the national level.

The National Security Council, with CyberSecurity Malaysia as the technical expert agency, have co-organised a periodic national cyber crisis drill entitled X-Maya since 2008. The main objective of the drill is to exercise the workability of the National Cyber Security Response, Communication & Coordination Procedure and to raise awareness of the national security impact associated with the significant cyber incidents among CNII. To date, X-Maya has been held 6 times, with the latest drill held on 7th March 2017.

PT 8: International Corporation

This thrust is essential as cybersecurity threats are not affected by physical countries' boundaries and borders. One of the main objectives identified by this thrust is to increase Malaysia's involvement and participation at the international level in key international cyber security organizations and platforms to mitigate cyber threats from information sharing and to overcome cybersecurity challenges among member countries. Malaysia is a member of the Forum of Incident Response and Security Teams (FIRST) and the Regional Asia Information Security Exchange Forum Meeting (RAISE) -- a cooperative platform for information sharing, communications and promoting best practices.

Among key initiatives under this thrust, CyberSecurity Malaysia became the co-founder, first chair and permanent secretariat of the Organization of Islamic Cooperation – Computer Emergency Response Team (OIC-CERT). CyberSecurity Malaysia is also a co-founding member and current deputy-chair of the Asia Pacific Computer Response Team (APCERT).

V. CONCLUSION

Cyber threats are problems of today and the future. While developments in the area of ICT allow for enormous gains in efficiency, productivity and communications, they also create opportunities for those with devious ambitions to cause harm. We have to be prepared for the worst, especially to protect our critical national information infrastructure.

Securing CNII against cyber threat activities requires the efforts of the entire nation. The government alone cannot sufficiently secure CNII. It calls for public-private-community cooperation in addressing the matter. The government can take the lead in many of these efforts, provided it is supported by the private and community sectors. Thus, a comprehensive master plan to create a secure and sustainable CNII for Malaysia against cyber threats must be formulated and developed.

As a result of the successful implementation of the NCSP Thrusts and initiatives, Malaysia has managed to attain 3rd place among 193 countries worldwide in the ITU GCI 2014 survey and maintain its position in the subsequent ITU GCI 2017 survey. In securing CNII, Malaysia is recognized as a champion by the World Summit Information Society (WSIS) Prizes 2016 and 2017 for international collaboration.

Securing CNII is a continuous effort and policy reviews are crucial to ensure it is abreast with the latest, dynamic and complex technologies. Research in this area, especially policy updates and reviews, can possibly be further conducted to lead to the development of a better strategy and policy framework to counter cyber threats.

VI. REFERENCES

- [1] Ministry of Science, Technology and Innovation Malaysia, "National Cyber Security Policy." 2006.
- [2] US Department of Homeland Security, "Blueprint for a Secure Cyber Future - The Cybersecurity Strategy for the Homeland Security Enterprise," 2011.
- [3] J. Russell and R. Cohn, *Critical Infrastructure Protection*, Bookvika Publishing, 2012.
- [4] T. G. Lewis, T. J. Mackin, and R. Darken, "Critical Infrastructure as Complex Emergent Systems," *Int. J. Cyber Warf. Terror.*, vol. 1, no. 1, pp. 1–12, 2011.
- [5] C.-W. Ten, G. Manimaran, and C.-C. Liu, "Cybersecurity for Critical Infrastructures: Attack and Defense Modelling," *IEEE Trans. Syst. Man Cybern.*, vol. 40, no. 4, pp. 853–865, 2010.
- [6] H.-C. Chu, D.-J. Deng, and H.-C. Chao, "Potential Cyberterrorism via a Multimedia Smart Phone Based on a Web 2.0 Application via Ubiquitous Wi-Fi Access Points and the Corresponding Digital Forensics," *Multimed. Syst.*, vol. 17, no. 4, pp. 341–349, Nov. 2011.
- [7] R. Heickero, "Terrorism Online and the Change of Modus Operandi," *Swedish Def. Res. Agency, Stock. Sweden*, pp. 1–13, 2007.
- [8] I. Bernik and K. Prislan, "Cyber Terrorism in Slovenia - Fact of Fiction," in *The 3rd International Multi-Conference on Complexity, Information and Cybernetics*, 2012.
- [9] J. Jarmon, "Cyber-terrorism," *J. Terror. Secur. Anal.*, pp. 102–117, 2011.
- [10] S. W. Beildleman, "Defining and Deterring Cyber War," *Mil. Technol.*, pp. 57–62, 2011.
- [11] R. Lemos, "SCADA system makers pushed toward security," *Security Focus*, 2006.
- [12] The Lipman Report Editors, "Cyberterrorism: The Invisible Threat Stealth Cyber Predators in a Climate of Escalating Risk," *Guardsmark, LLC, Memphis, Tennessee, USA*. 2010.
- [13] B. Kesler, "The Vulnerabilities of Nuclear Facilities to Cyber Attacks," *Strategic Insights*, vol. 11, pp. 15–25, 2011.
- [14] W. J. Broad, J. Markoff, and D. E. Sanger, "Israeli Test on Worm Called Crucial in Iranian Nuclear Delay," *New York Times*, Jan 15, 2011.
- [15] S. Quadir, "Bangladesh Bank exposed to hackers by cheap switches, no firewall - Police," *Reuters*, 2016.
- [16] K. N. Das and J. Spicer, "How the New York Fed fumbled over the Bangladesh Bank Cyber-Heist," *Reuters*, 2016.

- [17] K. Lema, "Philippines Urges Bangladesh to Share Results of Heist Investigation," *Reuters*, 2016.
- [18] M. Hayden, "A Timeline of the WannaCry Cyber-Attack," *ABC News*, 2017.
- [19] AFP, "Global ransomware attacks on the rise: Europol," *The Star Online*, 2017.
- [20] "WannaCry Ransomware," *Europol*, 2017.
- [21] J. Wattles, "Who Got Hurt by the Ransomware Attack," *CNNMoney*, 2017.
- [22] The Sun, "WannaCry ransomware attack in Malaysia confirmed," May 16, 2017.
- [23] MA-661.052017: MyCERT Alert – WannaCry Ransomware," 2017. [Online]. Available:
<https://www.mycert.org.my/en/services/advories/mycert/2017/main/detail/1263/index.html>.
- [24] W. Z. A. Zakaria, M. F. Abdollah, O. Mohd, and A. F. M. Ariffin, "The Rise of Ransomware," *Proceedings of the 2017 International Conference on Software and e-Business*, [Online] pp. 66–70, 2017. Available:
http://delivery.acm.org/10.1145/3180000/3178224/p66-Zakaria.pdf?ip=175.139.192.49&id=3178224&acc=ACTIVE%20SERVICE&key=69AF3716A20387ED%2E624C05D357EE4F12%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1542614296_0b300d3e7203ebede3b3a3994bc7e32
- [25] N. Koswanage, "Malaysia tries to stop threatened cyber attack," *Reuters*, 2011.
- [26] C. Fuchs and D. Trottier, ed., *Social Media, Politics and the State*, Protests, revolutions, riots, crime and policing in the age of Facebook, Twitter and YouTube, New York, Routledge, 2014
- [27] D. F. Wamala, "ITU National Cybersecurity Strategy Guide." International Communication Union (ITU), 2011.
- [28] "Global Cybersecurity Index (GCI) 2017." International Communication Union (ITU), 2017.
- [29] Jabatan Perdana Menteri Malaysia, "Pelaksanaan Pensijilan MS ISO/IEC 27001:2007 Dalam Sektor Awam." 2010.
- [30] S. N. Hamdan, S. Ismail, and M. A. Khalid, "Preparation towards ISMS Certification 27001: An Experience in Malaysian Nuclear Agency," vol. 44, no. 49, 2011.
- [31] CyberSecurity Malaysia, "Guideline to Determine Information Security Professionals Requirements for the CNII Agencies / Organisations." 2013.

Developing a Competency Framework for Building Cybersecurity Professionals

Ruhama Mohammed Zain¹, Zahri Yunos², Mustaffa Ahmad³, Lee Hwee Hsiung⁴, and Jeffrey Bannister⁵

^{1,2,3,4} CyberSecurity Malaysia

⁵Orbitage Sdn Bhd, Malaysia

ruhama@cybersecurity.my, zahri@cybersecurity.my, mus@cybersecurity.my,
hh.lee@cybersecurity.my, jbannister@orbitage.com

Abstract - The provision of secure networks and services is becoming more critical with the continuing growth of online services and prevalent hacks against systems. In particular, at the national level, countries must protect their critical infrastructure from malicious attacks. Central to this is the requirement to have an adequate pool of industry professionals who are well-versed in cybersecurity. These skillsets must be built and maintained in a structured manner and have a roadmap of lifelong learning for sustainability. A wide range of cybersecurity certification schemes are available; however, many are either prohibitively expensive to build large pools of professionals or have assessment mechanisms that do not measure individual abilities practically. This paper presents an approach to define a structured framework for building core critical skills in cybersecurity that is in line with industry requirements, provides a lifelong learning roadmap, incorporates professionalism and has a practical, competency-based assessment mechanism.

KEYWORDS - Competency Framework, Cybersecurity Professional, Cybersecurity Education, Knowledge, Skill, Attitude, KSA

I. INTRODUCTION

According to a recent article in Forbes magazine [1] that cites figures from the Information Systems Audit and Control Association (ISACA), an information security advocacy group, a global shortage of two million cybersecurity professionals is predicted by 2019. In the U.S., employers are currently struggling to fill cybersecurity positions, with many job ads going unanswered. Cisco's 2017 security survey found that certification and talents are the third and fourth barriers respectively, to effective security implementation.

In addition to vendor specific certifications, there is a growing number of vendor-neutral certifications. In the cybersecurity domain, several well-respected certifications are in existence. Whilst some of these are specific to particular equipment or processes, many are not and the coverage is extensive. For instance, Law Enforcement Agencies are seeking forensics to capture criminals, "C" level addresses risk, governance and business continuity, and Government Armed Services are looking for ways to defend a country.

Numerous "generic" national, regional and international standards, recommendations and guidelines have been developed and can be

referenced by program developers in creating learning programs [2][3]. However, an assessment mechanism, particularly at the entry level, focuses on online assessments. In addition, many dominant assessment mechanisms are exorbitantly expensive for organisations to build large numbers of certified personnel.

II. METHODOLOGY

The Global Accredited Cybersecurity Education Scheme (Global ACE Scheme) introduced by CyberSecurity Malaysia, an agency under the Ministry of Communication and Multimedia, Malaysia, is a holistic cybersecurity professional certification framework. It outlines the overall approach, independent assessment requirements, examination impartiality, trainer competences, cybersecurity domain identification and classification, professional membership requirements and professional development action plans. This scheme, similar to cybersecurity itself, is applicable and relevant across all Critical National Information Infrastructure (CNII) sectors, including national defence and security, banking & finance, information & communications,

energy, transportation, water, health services, government, emergency services and food & agriculture, as they all rely on secure IT systems. The Global ACE Scheme was developed in line with international standards ISO/IEC 9000 series [4] on processes, ISO/IEC 17024 [5] on people certification and ISO/IEC 27001 [6] on security management.

Contributions of this paper are in describing the key features of the Global ACE Scheme framework and highlighting the principal benefits of the scheme, which centres on competency-based assessment and affordability. This article also explains the structure and elements of the Knowledge, Skills and Attitudes (KSA) descriptors and how KSA links to training and assessment.

III. DISCUSSION

A. The Need For Competence-Based Assessment

It is essential today to have controls, policies and processes in place to ensure business continuity. Every day major issues arise with online systems, such as large amounts of personal details, medical records, credit card and other sensitive information being stolen or locked and encrypted by ransomware, or systems/mechanisms being compromised to steal data. This is not only happening to industry organisations but also to governments [7].

In today's environment, security awareness, knowledge and skills need to be central rather than peripheral. This requires an adequate pool of industry professionals who are well-versed in cybersecurity. The skillsets must be built and maintained in a structured manner and have a roadmap of lifelong learning for sustainability.

Many recent cases of massive security breaches have made headlines, indicating that despite technical advances, systems are still vulnerable, while lack of skills and awareness in the cybersecurity area is a key contributing factor [8]. As an example, the recent 'WannaCry' ransomware attacks affected systems that were not patched and updated – a crucial area that should be addressed by a proper security policy implemented in an organisation [9].

Countries are now adding cybersecurity skills as part of the national agenda, right through the learning life cycle from promoting

cybersecurity as a career choice all the way through to reskilling and continual professional development. For instance, a UK government "National Cyber Security Strategy 2016-2021" report [10] stated the following in its opening lines, and committed £1.9b to the strategy over the next 5 years:

"The challenge of our generation is to build a flourishing digital society that is both resilient to cyber threats, and equipped with the knowledge and capabilities required to maximise opportunities and manage risks" [10]

In the 1990s the Internet took hold and began growing at a tremendous rate. This meant huge volumes of equipment to be sold and maintained. As such, a "quick" method of certifying personnel who could perform "configuration" correctly needed to be rolled out globally. This gauntlet was taken up by Information Technology (IT) vendors who quickly realised that the more people were certified, the more equipment they could sell. Many of these programs were very well-designed in terms of content; however, to scale up and reach the masses, a simple assessment method was required, consisting of sets of online multiple choice questions offered through "prometric" testing centres [11]. It should be noted that some vendors had structured pathways to advanced levels that incorporate "practical, hands-on" assessment. Although this met "quick-fix" needs in the 1990s, in today's world it is viewed as sorely lacking [12]. Two main concerns arising from these types of assessment that significantly reduce their effectiveness for employers are:

- i. They mainly measure knowledge and memory capacity and have limited effectiveness in measuring critical thinking skills;
- ii. A question bank is often available and training programs on passing exams are offered.

Technical personnel are now not only expected to configure but also to have an end-to-end view of a complete system, understand "why" a configuration is done in a particular way, and be able to configure various equipment from different vendors securely by having a transferrable skill set. All of this needs to be captured in the assessment mechanism, so that employers can be confident in somebody's ability rather than

their skills in memorising multiple choice questions [13].

It should be noted that DoD Directive 8570.01-M [14] requires personnel with privileged access to DoD systems to have recognised certification. CompTIA Advanced Security Practitioner (CASP) [15] currently meets this requirement via only 80 multiple choice questions. Clearly, there is a requirement for a better means of assessing whether the certified person can actually perform the tasks required of a given job role.

The Global ACE Scheme is designed to enhance both the knowledge and skill sets of cybersecurity professionals with current and state-of-the-art techniques for strategizing, mitigating, developing and providing cybersecurity services. This ensures optimal application of cybersecurity knowledge and skills in the wider community.

B. The Challenge For Human Resource Departments

In most organisations, it falls on Human Resources (HR) to manage staff development and up-skilling. It has been observed, particularly in large technical organisations, that there is often a disconnection between HR and technical managers in terms of training development. Since technical managers do not generally see development as their job, they may provide HR with limited feedback. Consequently, because HR personnel are not generally technical, they source the same programs and certifications used previously, as they might not be aware of alternatives or able to interpret the technical requirements adequately.

At this time, organisations need to be more agile to meet market requirements. Hence, HR is expected to provide more such as consider strategic plans for organisational competency development, whereby skills are developed in a structured manner [16]. In many cases, HR does not have in-house capabilities to identify critical security competences and thus needs to work with external consulting organisations that have the necessary track record and expertise in the area. In the context of cybersecurity, such framework provides HR with a ready-made solution for developing skills. The framework thus has already identified the skills required by the industry, has a roadmap from foundation through to specialization, and offers a practical, hands-on certification process that validates individual ability to apply their skills.

The Global ACE Scheme is designed to measure an individual's ability to "do" a given task and understand "why" it is done by taking context into consideration rather than relying solely on knowledge-based assessments. It consists of 3 levels: foundation, practitioner and specialist, as highlighted in Figure 1.

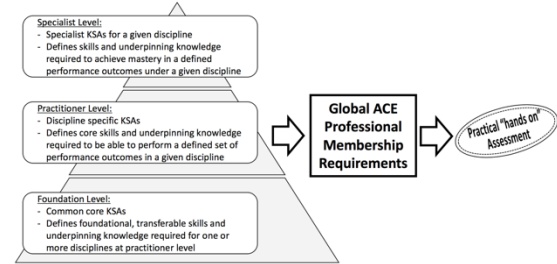


Figure 1: Competency framework

Each level consists of a number of competency modules referred to as KSA Descriptors (Knowledge, Skills, Attitudes) that prescribe a particular set of skills. For the purposes of this scheme, competency is defined as a skill plus the underpinning knowledge associated with that skill. At lower framework levels, these KSA Descriptors are written so as to enable the "transferability of skills" between job functions. Thus, a flexible, lifelong learning roadmap is possible with multiple career changes in the cybersecurity field. The framework is extendable in terms of the number of Descriptors based on industry requirements as identified via industry focus group workshops. Bloom's taxonomy [17] serves to ensure that the levelling complies with international norms and that there is consistency at a given level across descriptors. Further details on alignment with other reputable systems and how assessment reliability, validity and verification are ensured are given below.

C. Building A Structure For Identifying Competencies: The Ksa Descriptor (Knowledge, Skills, Attitudes)

Before it is possible to identify, develop, measure and maintain the "competencies" that the industry requires, a structured template is needed first, which can frame the requirements. This template provides a model to maintain consistency across each distinct area defined. For the purpose of this professional cybersecurity certification scheme, the template is referred to as a KSA Descriptor, the structure of which is the work product of a set of workshops conducted with a broad representation of industry players, cybersecurity experts, government

representatives and cybersecurity professionals. The KSA Descriptor's key purpose is to act as a reference guide, identifying the skills, underpinning knowledge and attitudes that professionals in the cybersecurity area require. The core functions of the KSA Descriptor are to act as:

- i. A reference for training providers to facilitate the development of suitable training courses relevant to the identified roles and functions;
- ii. A reference for developing examination questions to effectively assess the identified job roles and functions;
- iii. A reference for developing professional trainers able to effectively deliver training in line with the requirements of the identified job roles and functions.

One of the first questions to be addressed when developing the template is whether it should be framed from the perspective of a set of job functions or a set of learning outcomes. Since the main goal of the scheme is to develop cybersecurity professionals, we decided that it should lean towards training/development while keeping in mind that it should closely follow the performance requirements for a job. Therefore, a central part of the KSA Descriptor is to identify a set of performance outcomes for each given area; in other schemes, these are often referred to as 'tasks' [3].

The KSA Descriptor defines a benchmark of Knowledge, Skills and Attitudes onto which both training and assessment are mapped. Critical to success is for the certification to maintain quality throughout all processes to ensure that credibility is maintained. Therefore, in addition to the details of the KSA elements, a set of processes is also necessary to ensure quality and consistency are maintained throughout, as discussed in this paper under the heading "An ecosystem for skill development and assessment".

Another question arising during the definition phase is regarding the "A" in KSA. A survey of existing KSA type structures indicates that "A" referring to Ability or Attitude tends to occur in equal measures. However, it does become apparent that when referring to ability, it is challenging to discern the differences between a "skill" and an "ability" and there seems to be no consensus regarding this [18]. Using "attitude" fits in well with the overall philosophy of

professional certification in cybersecurity, since attitude is an important attribute of a professional, particularly when related to security matters. We found, for example, that "ethics" features extensively in matters related to security and should be blended into the fabric of skill development in this area.

The proposed framework shall address the three research areas and will not only focus on specific problems in isolation, for example, it assesses security in a SCADA network or makes a threat assessment of the latest zero-day vulnerability affecting a SCADA vendor [19]. The idea is to look at an overall research framework with the aim of increasing the dependability, resiliency and robustness of the SCADA network to support its critical processes.

The KSA Descriptor structure is split into five main sections, which are described in Table 1.

Table 1: Explanation of the Main KSA Descriptor Sections

| Section | Explanation |
|--------------------|---|
| Summary | Provides an overall summary of the scope and performance outcomes of the KSA descriptor, including pathway, document ID, version & date and an overview of the recommended training & assessment delivery mechanisms. |
| Knowledge (K) | Provides a set of Knowledge elements for the competency area. This is what one should "know." |
| Skills (S) | Provides a set of Skills elements for the competency area. This is what one should be "able to do." |
| Assessment Methods | Provides a legend to explain the different possible assessment methods for the K & S elements |
| Attitudes (A) | Provides a set of Attitudes elements for the competency area. This is what traits one should exhibit. Unlike the K & S elements, it is not expected that an assessment method should explicitly measure these, but rather that a training program should blend them into the learning fabric. This must be evaluated when the training program is submitted for evaluation. |

The major information elements of the summary section are explained in Table 2.

Table 2: KSA Descriptor - Summary Section

| Section | Explanation |
|----------|--|
| Synopsis | Provides an overview of the KSA descriptor scope. This is useful for HR personnel to get a summary of the KSAs and assist with mapping the competency area to the relevant job roles in an |

| | |
|------------------------------------|--|
| | organisation. |
| Performance Outcomes | Provides a set of outcomes that a successful individual should be able to demonstrate if they possess all KSA elements – these could also be termed “tasks”. |
| Learning Pathway | Identifies where this fits in the overall development roadmap. |
| Recommended learning time | Provides a minimum time benchmark for the duration of a course of building these KSAs in numbers of hours. |
| Training Strategy | Provides a summary of the type of learning environment to which a training program is expected to align. |
| Required Experience/Qualifications | Identifies pre-requisites expected before one would approach this set of KSAs. This is described in general terms and, if available, a KSA that identifies the pre-requisites. |

The Knowledge elements are explained below in Table 3.

Table 3: KSA Descriptor - Knowledge Section

| Section | Explanation |
|-------------------|---|
| Knowledge Element | Each knowledge element breaks the competency area down into the required knowledge at sufficient granularity at which it can be assessed. Training providers use this to ensure the knowledge element is covered sufficiently in training; exam question authors use this to ensure the element is assessed effectively. Both will utilize the Indicator for further scope clarification. |
| Indicator | The indicator provides further clarification on the knowledge element scope. It provides the information to allow both training organisations and examiners to build content & assessments to ensure the topic is addressed. |
| Weightage | Provides an indication of the amount of coverage there should be in the overall course/examination, e.g. 5% would indicate that in a 40-hour course, 2 hours should be spent on this Knowledge element. |
| Assessment Method | For element assessment, the method provides an indicator of the recommended way in which it should be assessed. A letter code is given to identify the method (e.g. PA – practical assessment, etc.) as shown in the legend below the elements (see Table 5). Appropriate learning & assessment |

| | |
|--|---|
| | techniques and educational best practices should be used in assessment development. |
|--|---|

The skills elements are explained as follows in Table 4.

Table 4: KSA Descriptor - Skills Section

| Section | Explanation |
|-------------------|---|
| Skills Element | Each skills element breaks down the competency area into the required skills at sufficient granularity for assessment. Training providers use this to ensure the skills element is covered sufficiently in training; exam question authors use this to ensure the element is assessed effectively. Both utilize the indicator for further scope clarification. |
| Indicator | The indicator provides further clarification on the skills element scope. It provides the information to allow both training organisations and examiners to build content & assessments to ensure the topic is addressed. |
| Weightage | This provides an indication of the amount of coverage there should be in the overall course/examination, e.g. 10% would indicate that in a 40-hour course, 4 hours should be spent on this skills element, i.e. practical activities |
| Assessment Method | For the assessment of this element, the method provides an indicator of the recommended way in which it should be assessed. A letter code is given to identify the method (e.g. PA – practical assessment, etc.) as shown in the legend below the elements (see Table 5). Appropriate learning & assessment techniques and educational best practices should be used in the development of assessments. |

Finally, each attitudes element breaks down the behaviours that should be developed and exhibited after training. Training providers use this to ensure the attitudes element is covered sufficiently in training; exam question authors do not need to use this, as the attitudes are not assessed separately but rather should be blended into the fabric of knowledge and skills development.

D. Identifying And Defining Key Industry Skill Requirements In The Cybersecurity Space

The approach adopted to identify and define industry requirements is to assemble a cross section of industry players for whom cybersecurity is critical, as well as academic representatives. This is done for two main reasons:

- i. Placing the two groups to work together means that skill requirements can be identified to meet industry requirements while also being structured in a way suitable for developing learning programs and assessment mechanisms.
- ii. Industry and academia are able to share their individual perspectives and appreciate each other's roles and viewpoints.

A number of workshops took place to identify the areas with wide appeal across industries as the core, in-demand skillsets, and subsequently build the KSA Descriptors for each.

One of the key outcomes is that to build skills in cybersecurity, technical practitioners need a solid foundation that addresses two fundamental areas: computer networks and operating systems. It was found that before an individual may consider security, they need to understand how services are offered and how traffic flows to and from these services. Thus, descriptors were built to identify these core skills and to act as pre-requisites for security-specific disciplines.

The descriptors were consolidated and circulated to produce a finalised set. The KSA Descriptors developed in this first phase are as follows:

- i. Cybersecurity Core/Foundations:
 - a. Computer Networking (security)
 - b. Operating Systems (security)
- ii. Cybersecurity-specific:
 - a. Business Continuity
 - b. Intrusion Detection, Monitoring & Prevention
 - c. Penetration Testing
 - d. Secure Application Development
 - e. Digital Forensics
 - f. Internet of Things (IoT) – security

E. An Ecosystem For Skill Development And Assessment

The KSA Descriptor forms a common benchmark for each defined area that specifies what the training and assessment outcomes should be. Figure 2 below shows the relationship between training, assessment and the KSA Descriptor.

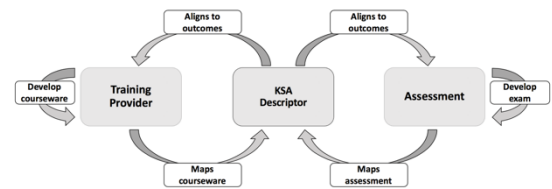


Figure 2: Relationship between training, assessment and the KSA Descriptor

To succeed, mechanisms and processes need to be in place to evaluate and validate training and assessment to ensure the following outcomes:

- i. Training and assessments align with the KSA descriptor
- ii. There is adequate and balanced coverage of each descriptor element based on the defined weightage
- iii. The training and assessment delivery mechanisms are consistent and meet the quality requirements set by Global ACE

For example, in training course development, the course developer must ensure that in developing the training materials:

- i. Each Knowledge element is covered in the training materials, e.g. slides and notes
- ii. Each Skills element is covered in the practical exercises
- iii. For each, the indicators are used to clarify the scope of coverage
- iv. The correct weightage is achieved for each element
- v. There is a strategy to develop and reinforce the Attitude elements throughout the training

Upon submitting course materials to an evaluation panel, the training organisation must adhere to the evaluation requirements. This includes marking all training materials to validate that all KSA elements are covered, for example:

- i. Provide highlighted slides, workbooks, notes, etc. to identify that each Knowledge & Skills element is addressed;
- ii. Provide a schedule to indicate the coverage of each element with the correct weightage
- iii. Provide a description of the training philosophy & mechanisms used to

build the Attitude elements through the Knowledge & Skills elements

For assessment delivery, the exam system must ensure that the appropriate assessment technique is used to assess each Knowledge & Skills element, e.g. if the descriptor indicates that “PA” practical assessment should be used, then the exam system must assess this in a practical context. It should be noted that does not preclude the use of a computer-based examination system; however, it must demonstrate how the system can emulate a live environment/scenario. The assessment must also ensure there is sufficient coverage of each Knowledge & Skills element in accordance with the weightage guidelines provided in the descriptor, e.g. if the Knowledge element indicates “MC” is the assessment method and 5% is the weightage and if the exam has 40 multiple choice questions, at least two should cover the element. The overall weightage in the exam must be maintained, e.g. if there is a set of short answer/written questions in addition to multiple choice questions, this should not dilute the weightage of the topic.

F. Assessment: The Importance Of Measuring Skills Practically

As mentioned earlier, effective assessment is a central requirement for structured skill development. The closer the assessment methods and criteria are to a real-world situation, the more successfully an organization can identify that an individual is competent [1][11].

For this reason, central to the KSA framework is that the assessment should cover both the Knowledge and Skills elements determined based on what the industry requires individuals to do as part of their jobs. The assessment methods are defined in the KSA Descriptor as follows:

Table 5: Assessment methods

| KSA | Associated Assessment Methods | When Assessed |
|-----------|---|----------------------|
| Knowledge | Continual assessment (CA) | During training |
| | Multiple Choice (MC) | Post training |
| | Theory/underpinning knowledge assessment (UK) | Post training |
| | Assignments (AS) | During/post training |
| | Case Studies (CS) | During/post training |

| | | |
|--------|---------------------------|----------------------|
| Skills | Continual assessment (CA) | During training |
| | Practical assessment (PA) | Post training |
| | Assignments (AS) | During/post training |
| | Case Studies (CS) | During/post training |

G. Managing & Tracking Professional Development

Managing and tracking certified professionals are two key activities to attract and retain scheme members. One vital mechanism to achieve this is to require that certified professionals maintain Continuing Professional Development (CPD) points in order to renew their membership status. It is a requirement under the scheme that certified members are constantly up-to-date with state-of-the-art developments in the field and technological changes. This will prevent the certifications from becoming outdated too quickly due to the fast-changing nature of cybersecurity. The Global ACE Scheme facilitates and enables opportunities for certified professionals to earn CPD points by organizing educational and professional events and publishing a list of recognized external events and activities. This fully supports the Malaysia Board of Technologists (MBOT) [20] function to promote education and training such that registered professionals may further enhance their knowledge related to their professions. Members will also benefit by having access to other experts in the course of attending the programs while at the same time enhancing their knowledge and skills.

H. Alignment With National Higher Education Ministries And Government Training Agencies

In Malaysia, the Ministry of Higher Education (MoHE), Malaysian Qualifications Agency (MQA) & Ministry of Human Resources/Department of Skills Development (JPK) are well-established and are the key organizations covering the spectrum of post school qualifications. MoHE and MQA govern both public and private universities and colleges, with JPK in charge of skills development with all three using the Malaysian Qualifications Framework (MQF) [21]. These organizations have a wealth of knowledge and processes in place to ensure quality mechanisms throughout the whole

value chain to ensure credibility, review of processes and sustainability [22][23].

The Global ACE scheme does not intend to reinvent the wheel in terms of certification, but recognizes that there are many Cybersecurity Professional Certifications on the market. Mechanisms will be put in place to determine how persons with such certifications can have a route to specialist certification if they so desire. The relevant committees will evaluate reputable certifications on the market and look at how to map them to the KSA Framework levels and standards [24].

I. Validation By Experts

The Global ACE Scheme framework has been validated by experts from industry, academia and the Malaysian government. The validation mechanism was a series of meetings and workshops during which all aspects of the framework were proposed, deliberated, revised based on feedback received and presented again for final acceptance by the relevant committees. Table 6 summarizes some of the meetings and workshops conducted to validate the scheme. The nature of engagement with experts from academia, government and industry is described along with the number of workshops held and the total number of attendees.

Table 6: Meetings and workshops conducted

| Sector | Nature of engagement | Number of workshops | Number of attendees |
|------------|---|---------------------|---------------------|
| Academia | <ul style="list-style-type: none"> • Scheme framework development • KSA descriptor development • Assessment questions development • Board of governance | 16 | 63 |
| Government | <ul style="list-style-type: none"> • Scheme framework development • KSA descriptor development • Scheme risk management • Board of governance | 16 | 157 |
| Industry | <ul style="list-style-type: none"> • Scheme | 15 | 95 |

| | | | |
|--|---|--|--|
| | framework development <ul style="list-style-type: none"> • KSA descriptor development • Assessment questions development • Board of governance • Training content mapping & alignment | | |
|--|---|--|--|

IV. LIMITATION

It is acknowledged that this is a preliminary study that seeks to identify and build the necessary components for a competency-based framework for developing cybersecurity professionals. In order to improve this framework further, an in-depth study of existing training and certification frameworks will have to be undertaken for the purpose of comparison and ensuring its continued relevance and currency. This is reserved as a future work.

V. CONCLUSION

The Global ACE scheme takes a competency-based approach that focuses on building and assessing both knowledge and skills in a practical context across key domains within the cybersecurity landscape. This approach was chosen to address the critically growing global shortage of talent in the cybersecurity field. The emphasis is on assessments that measure practical competence rather than purely theoretical and/or multiple-choice question assessments alone. In short, the scheme aims to produce cyber-security professionals with the necessary critical thinking skills, confidence and true ability to complete tasks. The scheme also outlines a structured roadmap to build and maintain professionals across the cybersecurity domain.

For future work, a detailed study to compare this scheme framework to other training and certification scheme frameworks is proposed. It would also be fruitful to research the outcome of implementing this scheme in terms of the number and quality of cybersecurity professionals produced.

VI. REFERENCES

- [1] J. Kauflin, "The Fast-Growing Job with a Huge Skills Gap: Cyber Security," *Forbes*, Mar-2017.
- [2] SFIA framework — SFIA," SFIA Foundation, 2015. [Online]. Available: <https://www.sfia-online.org/en/sfia-6>. [Accessed: 03-Jan-2018].
- [3] W. Newhouse, S. Keith, B. Scribner, and G. Witte, "National Initiative for Cybersecurity Education (NICE) Cybersecurity Workforce Framework," NIST Spec. Publ., pp. 800–181.
- [4] "ISO 9001:2015 Quality Management Systems." International Organization for Standardization, Geneva, Switzerland, 2015.
- [5] "ISO/IEC 17024:2012 Conformity assessment -- General requirements for bodies operating certification of persons." International Organization for Standardization, Geneva, Switzerland, 2012.
- [6] "ISO/IEC 27001:2013(en) Information technology — Security techniques — Information security management systems — Requirements." International Organization for Standardization, Geneva, Switzerland, 2013.
- [7] "Cisco 2017 Annual Cybersecurity Report," San Jose, California, 2017.
- [8] "Mitigating the Cybersecurity Skills Shortage Top Insights and Actions from Cisco Security Advisory Services," 2015.
- [9] S. Gibbs, "WannaCry: hackers withdraw £108,000 of bitcoin ransom | Technology | The Guardian," *The Guardian*, 2017. [Online]. Available: <https://www.theguardian.com/technology/2017/aug/03/wannacry-hackers-withdraw-108000-pounds-bitcoin-ransom>. [Accessed: 03-Jan-2018].
- [10] UK Government, "National Cyber Security Strategy 2016-2021," 2016.
- [11] Prometric, "Overview," 2017. [Online]. Available: <https://www.prometric.com/en-us/about-prometric/pages/prometric-advantage-overview.aspx>. [Accessed: 03-Jan-2018].
- [12] J. Richard, "Forensication Education: Towards a Digital Forensics Instructional Framework Forensication Education: Towards a Digital Forensics Instructional Framework GIAC (GCFE) Gold Certification Forensication Education 2," SANS Institute, InfoSec Read. Room, 2017.
- [13] H. Bound, A. Chia, and S. Yang, "Assessment for the changing nature of work," *Inst. Adult Learn.*, 2016.
- [14] "Information Assurance Workforce Improvement Program," DoD 8570.01-M, 2015.
- [15] "(CASP) Advanced Security Practitioner Certification | CompTIA IT Certifications," [certification.comptia.org](https://certification.comptia.org/certifications/comptia-advanced-security-practitioner), 2017. [Online]. Available: <https://certification.comptia.org/certifications/comptia-advanced-security-practitioner>. [Accessed: 03-Jan-2018].
- [16] J. Gothelf, "How HR Can Become Agile (and Why It Needs To)," *Harvard Business Review*, 2017. [Online]. Available: <https://hbr.org/2017/06/how-hr-can-become-agile-and-why-it-needs-to>. [Accessed: 03-Jan-2018].
- [17] D. R. Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. New York: David McKay Company. Inc., 1956.
- [18] D. H. P. R. G. & Collier, *Motor Learning and Development*. Human Kinetics, 2011.
- [19] E. Byres, D. Leversage, and N. Kube, *Security incidents and trends in SCADA and process industries. The industrial ethernet book*, 2007.
- [20] "Malaysia Board of Technologists," 2017. [Online]. Available: <http://www.mbot.org.my>. [Accessed: 08-Dec-2017].
- [21] Malaysia Qualifications Agency, "Malaysian Qualifications Framework Point of Reference and Joint Understanding of Higher Education Qualifications in Malaysia." 2016.
- [22] Jabatan Pembangunan Kemahiran, "Jabatan Pembangunan Kemahiran - Home," 2017. [Online]. Available: <http://www.dsd.gov.my/index.php/en/>. [Accessed: 08-Dec-2017].
- [23] Kementerian Pendidikan Tinggi, "KPT - Utama," 2017. [Online]. Available: <http://mohe.gov.my/>. [Accessed: 04-Jan-2018].
- [24] "Cyber Security Certifications | Explore Your Options," *Cyber Degrees*, 2017. [Online]. Available: <http://www.cyberdegrees.org/resources/certifications/>. [Accessed: 03-Jan-2018]

Preventing Reflective DLL Injection on UWP Apps

Mojtaba Zaheri¹, Salman Niksefat², and Babak Sadeghiyan³
^{1,2,3}APA Research Center, Amirkabir University of Technology

Abstract - Universal Windows Platform (UWP) is the Microsoft's recent platform-homogeneous application architecture. It allows a code to run on variety of devices including PC, mobile devices, etc., without needing to be rewritten or recompiled. UWP apps are becoming more and more popular and consequently this new application platform is becoming the next attack target for hackers and malware developers. In this paper, we first study the issue of host-based code injection attacks (HBCIA) in UWP apps. We show that despite the embedded mechanisms in UWP to maintain code integrity and to only allow legitimate DLLs to be loaded in memory, it is still possible to circumvent the defensive mechanisms and launch a variant of HBCIA called Reflective DLL Injection on UWP apps. We then propose a novel defence mechanism against reflective DLL injection attacks on UWP apps. Our proposed method can detect malicious/benign injection attempts on UWP apps and prevents malicious injections while allowing the benign injections to proceed as normal. Our experiments show that the proposed defence has less than 1% impact on system's overall performance and can be used inside anti-virus (AV) products to strengthen their protection capabilities.

KEYWORDS – DLL Injection, Universal Windows Platform, UWP

I. INTRODUCTION

Universal Windows Platform (UWP), first introduced in Windows 10, is the Microsoft's platform homogeneous application architecture. Its purpose is to allow development of universal applications that run on a variety of platforms including PC, mobile devices, and IoT devices. This relieves the code from the need to be rewritten or recompiled for each platform. Similar to Android and IOS, this platform has its own proprietary software store through which Microsoft can have more control over the distributed UWP applications. Since its release, Microsoft has encouraged the software developers to write code in UWP and the company itself included some UWP applications in Windows 10, including Microsoft Edge browser and Microsoft Groove Music.

With rapid popularity of UWP applications among software developers and considering the strong support of Microsoft, UWP apps are becoming more and more popular among end users and consequently this new platform has become the next attack target for hackers and malware developers. One important category of intra-host attacks that can potentially target UWP applications is Host Based Code Injection Attacks (HBCIA). HBCIA is defined as locally copying a code from a malicious source process into the address space of a target process and

executing the code [1]. A recent research in [1] shows that near 64% of the total of 162850 sample malware use HBCIA as part of their malicious behaviour.

One strong motivation for using HBCIA by malware is to evade detection and bypass host-based firewalls: Malware usually connect to their C&C¹ servers for sending information and receiving new commands. Thus host-based firewalls are generally sensitive to outgoing connections of locally running applications and they have rules to prevent unknown applications from accessing the network. To prevent being caught by the firewalls, new malware generally uses smart techniques for connecting to the Internet. One such technique is taking advantage of HBCIAs, i.e., injecting a software module to another legitimate running process such as Mozilla Firefox, Internet Explorer or Google Chrome and communicating using the injected module. Among these browsers, Microsoft Internet Explorer has been more promising for hackers as it is generally available in Windows family of operating systems by default. Moreover, Microsoft-Edge, the Microsoft's new UWP-based browser introduced in Windows 10, can be the next injection target for malware that are willing to launch a code-injection attack.

In this paper, we demonstrate that it is still possible to launch successful DLL injection attacks by a technique called reflective DLL injection [2][3] despite the new security

mechanisms embedded in UWP framework to maintain code integrity and prevent unsigned/malicious DLL injections. Then, we propose a defence mechanism against such attacks on UWP apps.

Some currently published methods such as [4][5] try to parse the victim process memory and find if a malicious DLL is loaded into the process memory. Then, they try to remove it and clean the memory. However, it's not a sound and complete countermeasure, as the malware is already loaded in the memory and can do its malicious activities before being removed from the memory. In contrast, in our proposed mitigation, we try to prevent the malware to load the malicious DLL from the very beginning. Another challenge in countering such attacks is that not all code injections are malicious. The operating system may inject some legitimate DLLs into processes. Moreover, processes may inject code into their own address space for purposes like loading plug-ins, etc. Therefore, we need a method to distinguish between malicious and benign injections. Our proposed defence mechanism does this with high precision. In case of a malicious injection, it successfully prevents the DLL to be written into the target process and raises an alarm. On the other hand, in case of a benign injection, the injection proceeds as normal. Finally, by taking advantage of PCMark benchmarking tool, we show that our proposed technique imposes a little overhead on operating system.

To summarize, our contribution in this paper is a mitigation technique against reflective DLL injection on UWP apps that provides the following original advantages:

- i. It entirely prevents a malware from loading its malicious module into the target process memory.
- ii. The proposed mechanism is very efficient as it only monitors and modifies the behaviour of one API (NtWriteVirtualMemory), which leads to a very low overhead on the system performance.
- iii. It doesn't have any effects on normal DLL injections, as it's possible to load legitimate/signed DLLs into target UWP apps through calling the LoadLibrary API.

This paper is organized as follows: In section II, we re-view related work including HBCIA methods and the existing

countermeasures. In section III, we review the security mechanisms embedded in UWP apps that are related to HBCIA attacks. In section IV, we demonstrate the methods that can circumvent the integrity mechanisms of UWP and perform the reflective DLL injection attack in UWP framework. In section V, we present our defensive mechanism to reflective DLL injection attacks. In section VI, we present the results of the evaluation of the proposed system. Finally, section VII concludes the paper.

II. RELATED WORK

The works in host-based code injection attacks can be classified into methods for performing such attacks, and mechanisms for detection and prevention. In this section, we review these works and considering the detection and prevention mechanisms, we claim that none of them is suitable for defending against reflective DLL injection on UWP apps.

A. Performing HBCIA

Since these methods have rather a technical nature, the concept has received much more attention in the technical forums rather than the research papers.

In [1], the authors have presented a semi-formal definition for host-based code injection attacks that we cited in the introduction. The paper has presented the basic idea of the technique in three main steps including I) Victim process selection, II) Code copying, and III) Code execution. This paper also mentions several motivations behind using HBCIAs including interception of critical information, privilege escalation, and detection avoidance.

In [6], a classification on various DLL Injection techniques is presented. This paper classifies these techniques as follows: CreateRemoteThread [7], Creating a Proxy DLL [8], Modification of Windows Registry [9], Windows Hooks [10][11], Using a Debugger [12], Patching the IAT [13] and Reflective Injection [2][14].

Most of the above techniques can't be used to inject into UWP apps because the LoadLibrary API has been limited by UWP

framework code integrity mechanism. However, a specific type of DLL Injection called Reflective Injection which was introduced in [2] can be used to circumvent this mechanism. This method can load a DLL on UWP apps through the concept of reflective programming without directly using LoadLibrary API. In section 4, further details about this technique is presented.

B. Detecting and Preventing HBCIA

Since the HBCIAs need to have local access to the target system, these types of attacks had not been considered very hazardous in the past. However, the advances in HBCIA techniques and ever-increasing number of malwares in recent years have motivated the security researchers to work on mitigation mechanisms for these attacks. In the following, we review some of these methods.

In [1], a mechanism named BeeMaster is proposed to prevent host-based code injections through using honeypot paradigm. In this mechanism, a master bee and multiple worker bees are used. The master bee creates and instruments the workers to find if a code injection is occurred. If so, the master bee creates a memory dump and terminates the worker bee. The downside of this mechanism is that the detection only works on the processes that are created by the master bee, and therefore it cannot detect the targeted injections that occur on other processes of the system.

[15] aims to detect malicious DLL injections by evaluating the injected DLLs through the information provided by the process snapshots. For this purpose, it checks some common malicious DLL characteristics in the loaded DLLs to find a match. Nevertheless, one of the drawbacks of this technique is that it cannot detect the attack before the injection, so it cannot prevent the malicious DLL from being loaded. In [16] a similar technique is opted for to detect malicious DLLs through their characteristics by using machine learning methods and has the similar defects of [17].

Some of the code injection methods mentioned in section 2.1 are useful for detection and prevention purposes. For instance, in [18] a mechanism called DLL Preemptive Injection is used that whenever the system is loading the UrlMon DLL to a

process, it interrupts the process and loads a monitoring module that later checks the API call patterns in the target process to see if its behavior is malicious. However, the proposed method is only effective against Trojan downloaders.

Also, Detecting the Code Injection Engine (DCIE) [19] tries to reject all the suspicious thread creating calls by hooking APIs and tracing three main steps of code injection attacks: allocating memory, writing to the memory, and creating the thread. Although this method prevents the injection attacks, it has two major weaknesses;

- i. It rules out injection of legitimate and signed DLLs, and
- ii. It hooks three APIs, which decreases system's performance.

In case of reflective injection, the articles [20][21] propose ideas to check the memory of running processes periodically and search to find if there is any malicious content, and then they try to delete the infected memory pages, change their permissions, or even kill the infected process. However, during the time span between the two checks the malware can harm the system.

In comparison with previous methods, in this paper we propose a countermeasure against reflective DLL Injection on UWP apps, which is very effective, hooks only one API so it does not depend on API succession and has a very low impact on the system's performance. Moreover, through its combination with UWP Binary mitigation mechanism, it still lets legitimate DLLs to be loaded without any limitation. In section V5, our proposed countermeasure is presented.

III. UWP SECURITY MECHANISMS

Before addressing the issue of code injection attacks on UWP apps, we should first review several security mechanisms embedded in UWP framework to prevent the classic injection attacks to happen. Microsoft's attitude toward UWP is not only a better user experience but a more secure environment for application development that makes it harder for malware to penetrate UWP-based devices. Two important security mechanisms in UWP are "App Container" and "Code Integrity Enforcement" which are

directly related to HBCIA attacks. We review these mechanisms in this section.

A. App Container

UWP framework is equipped with a new security sandbox called App Container which provides more fine-grained per-mission assignments and limits unauthorized read and write operations throughout the system. App Container helps to make sure that an UWP app is only restricted to its defined security permissions. In the following, we review a number of App Container capabilities.

Limit access to files and peripherals. UWP apps are restricted to access directly to only two directories: the app's WindowsApps directory in Program Files, and the app's package directory located in AppData. The full path to the WindowsApps is [Win_Drive]:\Program Files\WindowsApps.

All files stored by apps in WindowsApps have to be static files that don't change through the app's lifetime. To enforce this rule, files stored by applications in this directory go through integrity checks before the app is launched. If a file in this directory is modified, the app will fail those integrity checks and refuse to launch. Also, the app's local AppData directory is located in [Win_Drive]:\Users\[UserName]\AppData\Local\Packages.

This directory is meant to be a place for apps to store dynamic files that can change over the time. As such, files in this directory don't go through integrity checks because it is meant to be a place for apps to store cache files, settings files, save files, and more.

Integrity Levels. App Container is implemented using the concept of Integrity Levels. Considering the definition in Microsoft's MSDN (Microsoft, n.d.-c), the Level has one of labels as System, High, Medium, Low, Untrusted.

This notion has been introduced in Windows Vista and is attributed to processes and objects. This mechanism prevents low level processes from reading or modifying high level processes and objects.

In Windows 8, Integrity Levels have been combined with the App Container, and limit processes to only read and write in their restricted area. This concept helps to ensure that the program does not have any access to the areas that are out of its range, unless the access is explicitly granted. To address this issue, every app container is assigned with a

SID², and like users, the programs that are running in app containers.

Security Identifier can be part of Built-In groups, and consequently, have access to specific resources on the system. The associated name for these App Container Built-In groups is "Capabilities".

Specifically, in case of DLL loading, it's worth mentioning that all DLLs must have the read/execute per-missions of SID "S-1-15-2-1" which is equivalent ID for ALL_APPLICATION_PACKAGES, in DLL's Access Control List (ACL) (VoxelBlock, 2016).

B. Code Integrity Enforcement

Another important security mechanism in UWP apps is the Code Integrity Enforcement [22]. This mechanism is applicable in both process and kernel levels. The process-level enforcement is useful until the time the process is not compromised because the code integrity check can be disabled in a hacked process by the malware. Therefore, Microsoft has implemented the enforcement in the kernel-level to strengthen it against hacked processes and to prevent mal-ware from disabling this mechanism.

This mechanism activates during the LoadLibrary() API call. When a binary is going to be loaded, the kernel calls NtCreateSection() and then MiCreateSection() APIs. This last API finally invokes MiValidateSectionCreate() API which uses ci.dll (Code Integrity) to check the file signatures. If the verification does not match the defined policy, the kernel won't create the section and will return an error. The mitigation is performed by the kernel, so to turn off the mitigation, the intruder must have the kernel-level (ring 0) privilege [23].

The integrity check policies are defined in a structure called Process Signature Policy in "WinNT.h" (Microsoft, n.d.-a). Using the latest Windows SDK, one can see this structure as shown below:

```
typedef struct
_PROCESS_MITIGATION_BINARY_SIGNATURE_POLICY
{
    union {
        DWORD Flags;
        struct {
            DWORD MicrosoftSignedOnly : 1;
            DWORD StoreSignedOnly : 1;
            DWORD MitigationOptIn : 1;
            DWORD ReservedFlags : 29;
        };
    };
};
```

```

    }
    DUMMYSTRU
    CTNAME;    }
    DUMMYUNIO
    NNAME;
}
PROCESS_MITIGATION_BINARY_SIGNATURE_POLICY,
*PPROCESS_MITIGATION_BINARY_SIGNATURE_POLICY;

```

The flags specified in the structure enforce integrity restrictions. `MicrosoftSignedOnly` can be set to prevent the process from loading images that are not signed by Microsoft. `StoreSignedOnly` can be set to prevent the process from loading images that are not signed by the Windows Store and finally `MitigationOptIn` can be set to prevent the process from loading images that are not signed by Microsoft, the Windows Store and the Windows Hardware Quality Labs (WHQL).

All in all, the above integrity mechanism makes loading an unsigned DLL using `LoadLibrary` API impossible. Nevertheless, in the next section we review a number of recent techniques that allow intruders to circumvent this mitigation and load arbitrary DLLs into the memory of UWP apps even in the presence of an anti-virus.

IV. HBCIAS ON UWP APPS

One way to perform a host-based code injection attack is to put the code inside a DLL file and inject the DLL to the target process. This is called DLL injection. A classic DLL injection attack in Windows operating system is usually carried out by the following steps [7]:

- i. Obtaining a handle to the victim process through calling `OpenProcess` API by setting the process's ID as the input parameter of this API.
- ii. Allocating space inside the target process, by invoking `VirtualAllocEx` API.
- iii. Writing malicious DLL's path into the allocated memory space, by using `WriteProcessMemory` API.
- iv. Obtaining a handle of `Kernel32.dll` module by calling `GetModuleHandle` API.

- v. Obtaining the address of `LoadLibrary` API through using `GetProcAddress` API, with `Kernel32.dll`'s handle and `LoadLibrary`'s name as the input parameters.
- vi. Calling `LoadLibrary` API by one of thread creating APIs like `CreateRemoteThread`, `RtlCreateUserThread`, and `NtCreateThreadEx`, by using handle of the target process, address of `LoadLibrary` API, and written memory address of the DLL path as input parameters to accomplish the attack.

Due to the new code integrity security mechanism available for UWP apps, it is possible to only allow signed DLLs to be loaded this way [22]. Thus, attackers must not be able to inject arbitrary DLLs on a target process that is being protected by the code integrity mechanism.

However, Microsoft's code integrity mechanism only triggers on the `LoadLibrary` API call, it is still possible to inject binary shellcodes into the target process as stated in [23]. However, working with shellcodes is very difficult and the attacker has to handle many complexities. Hence, attackers are still looking for methods that despite the existence of Microsoft's binary mitigation mechanism, inject their arbitrary DLLs to the memory of processes. A little surfing of the security and hacking technical forums reveals that it is possible to use a tiny bootstrap shellcode to perform a so-called Reflective DLL Injection [2][3] and load an arbitrary DLL into a target process without directly using the `LoadLibrary` API call. However, the reflective DLL injection technique has been proposed for classic Windows applications and their use against UWP apps is not yet documented in academic papers or technical forums. We confirmed that this technique works successfully against UWP apps too by injecting an arbitrary DLL into the Microsoft Edge browser's memory. The details for the reflective DLL injection attack elaborate in the next section.

A. Reflective DLL Injection

Assuming the attacker has code execution capability in the target process and the whole content of the library (s)he wishes to inject has been written into an arbitrary location of

memory in the target process, Reflective DLL Injection [2][3] works as follows:

- i. Execution is passed via a tiny bootstrap shellcode to the library's ReflectiveLoader function which is an exported function found in the library's export table.
- ii. Since the library's image currently exists in an arbitrary location in memory, the ReflectiveLoader first calculates its own image's current location in memory so as to be able to parse its own headers for use later on.
- iii. The ReflectiveLoader will next parse the processes kernel32.dll export table in order to calculate the addresses of three functions required by the loader, namely LoadLibraryA, GetProcAddress and VirtualAlloc.
- iv. The ReflectiveLoader will then allocate a continuous region of memory into which it will proceed to load its own image. The location is not important as the loader will correctly relocate the image later on.
- v. The library's headers and sections are loaded into their new locations in memory.
- vi. The ReflectiveLoader will then process the newly loaded copy of its image's import table, loading any additional library's and resolving their respective imported function addresses.
- vii. The ReflectiveLoader will then process the newly loaded copy of its image's relocation table.
- viii. The ReflectiveLoader will then call its newly loaded image's entry point function, DllMain with DLL_PROCESS_ATTACH. The library has now been successfully loaded into memory.
- ix. Finally, the ReflectiveLoader will return execution to the initial bootstrap shellcode which called it.

Since the technique doesn't need a direct call to LoadLibrary, the security mechanism embedded in UWP apps is not able to prevent loading of the DLL. In the next section, we propose our mitigation mechanism to prevent this type of attack.

V. THE PROPOSED DEFENSE

In section 4, we discussed that despite the embedded mechanism in UWP framework against code injection attacks [22], it is still possible to bypass protection and inject arbitrary DLLs in UWP apps. We explained that the reflective DLL injection can be used to inject a DLL into UWP apps (e.g. Microsoft Edge browser) without direct call to the LoadLibrary API. In this section we propose a technique for defending against code injection attacks in UWP apps. The general idea for the defence is to precisely monitor an API call that is commonly used in reflective DLL injection attacks. More specifically, our idea is to monitor the input parameters to NtWriteVirtualMemory() API, which is used to write into the memory of a target process, and only allow valid parameters to get into.

To implement this, we use a hooking library to build a hooking DLL that hooks into all user-mode processes by means of a system-wide Kernel-mode injection driver. Since Microsoft strictly forbids patching or hooking in the driver land, we implemented the hooking in user-level, and made it system-wide by a driver that does the DLL injection in the kernel-level.

A. Preliminaries

Before presenting the proposed defence mechanism, we should first discuss some preliminaries about the underlying Windows internals that are used to build our mitigation engine.

User-Mode API Hooking is a technique by which developers can instrument and modify the behavior of API calls, for different purposes like monitoring programs' behavior, forcing them to function in a different way, etc. Hooks are widely used by anti-viruses, security applications, system utilities, programming tools, and so on. There are multiple hooking libraries such as Microsoft Detours [24], Mhook [25], Deviare [26], EasyHook [27], and others that can provide the user mode hooking capabilities. Their typical function is as follows:

- i. Storing beginning bytes of the original code of the target function somewhere else. It is needed for the correct behavior of the hooked function.

- ii. Overwriting the beginning bytes of the target function with a custom code (called trampoline). So, when the function executes, it jumps to the hook handler.
- iii. If needed, calling the stored original target function, at the end of the hook handler.

In this paper, we use Mhook [25] which is an open-source library and supports API hooking in both 32- and 64-bit programs. Microsoft also has introduced kernel-mode callbacks with Windows Vista. These callbacks are registered in kernel mode and provide notifications to the registrar upon a certain event (e.g. if you register a callback for a specific activity then you can have your callback function invoked before/after the action has occurred on the system). Three important callbacks for AV products are triggered for Create Process, Create Thread, and Load Image events. These callbacks are registered by invoking:

- i. PsSetCreateProcessNotifyRoutine
- ii. PsSetCreateThreadNotifyRoutine, and
- iii. PsSetLoadImageNotifyRoutine.

Our proposed mitigation mechanism which is written in a hooking DLL is deployed system-wide using the kernel-level injection in a LoadImage callback routine.

B. The Mitigation Engine

In this section we explain the proposed technique that mitigates the code injection attacks by monitoring the calls to NtWriteVirtualMemory API. Figures 1 and 2 depict reflective versus normal DLL injections while our proposed defence mechanism in action. Our proposed mechanism consists three main steps:

Determining if the Binary Mitigation is enforced in the target process: The proposed countermeasure aims to prevent the malware from circumventing the binary mitigation mechanism. In fact, we want to tighten up the mitigation currently enforced in UWP apps, and consequently the proposed mechanism should only be activated for UWP binaries that are already protected by Windows mitigation policy. In other words, if the binary mitigation is not active, the attacker

can use the LoadLibrary API directly to load its malicious DLL in to the target process.

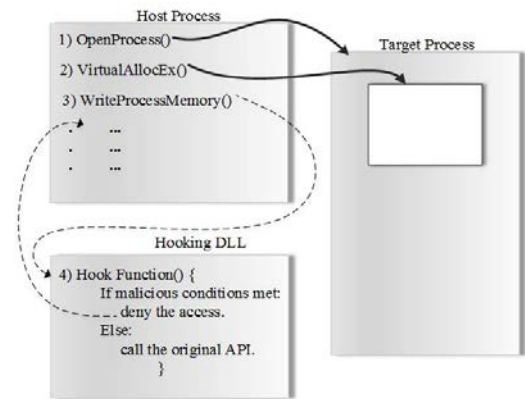


Figure 1: Proposed defence in action while a malicious reflective DLL Injection is being launched

For this purpose, we check the Signed Only flags in `PROCESS_MITIGATION_BINARY_SIGNATURE_POLICY` structure to find if the target app is forced to load only signed DLLs. This information is provided by "WinNT.h" in the Windows SDK version 10.0.14393.0, and can be accessed by calling `GetProcessMitigationPolicy()` API with `ProcessSignaturePolicy` type, and `ProcessHandle` structure passed to `NtWriteVirtualMemory` API, as inputs. This way, the mechanism neglects the injections to other windows applications, just like the way the Windows 10 itself does.

Detecting inter-process writes. It's possible for applications to write into their own address space using `NtWriteVirtualMemory` API call, which is apparently a non-malicious act. Therefore, we consider these intra-process writes as safe injections and continue to check whether we detect a `NtWriteVirtualMemory` call in which the process IDs of the caller and the target process are different. Since the `NtWriteVirtualMemory` API is invoked in the source process, the hook function is also executed in the con-text of this process and we can get the process ID of this process by calling `GetCurrentProcessId` API. The process ID of the target process can also be obtained from the `ProcessHandle` structure passed into the `NtWriteVirtualMemory` API. The `GetProcessId(ProcessHandle)` can obtain this data for us. If these two process IDs are equal, we consider it as a legitimate intra-process injection, and call the original `NtWrite-VirtualMemory` API without any modification. Otherwise we go to the next step for further checking.

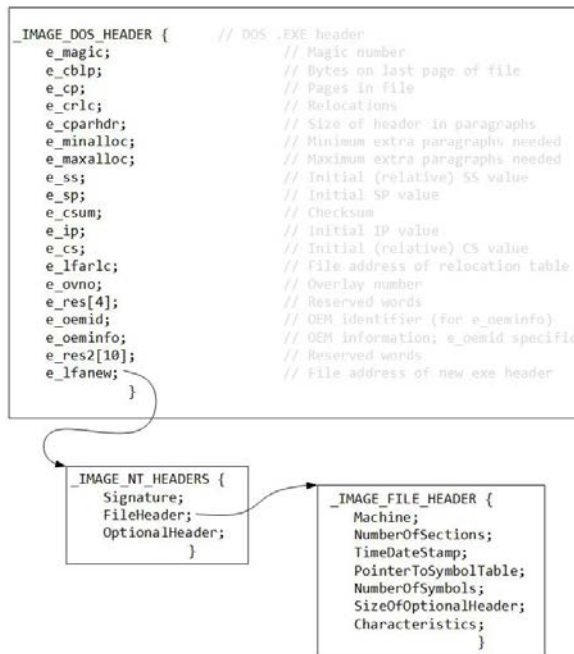


Figure 2: Proposed defence in action while a benign normal DLL Injection is being launched

Preventing the call if the input includes a DLL. The main difference between the reflective injection and normal DLL injection is that instead of writing the path of the desired DLL, it directly writes the DLL content into the target process memory, and consequently makes it possible to circumvent the Microsoft Mitigation Policy. So, we can utilize this fact, and prevent the write operation if the writing content contains a DLL. To check this please note that all Windows executables begin with a MS-DOS executable stub. So, we first check if a MS-DOS program header exists at the beginning of the injected data. We then check for markers for a Windows executable. If we learned that the writing content is a Windows executable, we look for information that determines whether the file is an application or is a DLL. So, we check the following conditions respectively:

- i. We check the first bytes of data for a valid DOS header. To do this we check the DOS header size field which should be 64 bytes at minimum.
- ii. All DOS program files (and therefore Windows executables) begin with a magic number; the word value \$5A4D ("MZ" in ASCII). So, we check if e_magic field of DOS header is equal to \$5A4D.
- iii. The Windows NT header begins with a magic number word whose value

indicates whether this is a NE³ format or PE⁴ format executable or a virtual device driver with LE⁵ format. The word is \$454E ("NE" in ASCII), \$4550 ("PE") or \$454C ("LE"). So, we check if the Signature field of NT header is equal to \$4550.

- iv. Windows executables have a file header immediately following the \$4550 magic number. This header structure has a Characteristics field which is a bit mask. If the bit mask contains the flag IMAGE_FILE_DLL then the file is a DLL, otherwise it is a program file.

Figure 3 illustrates the important structures in "WinNT.h" header file of Windows Kits version 10, considered in the pro-proposed mitigation. If all the conditions are met, the mitigation engine considers the API call as malicious, aborts the call, and raises an alarm.

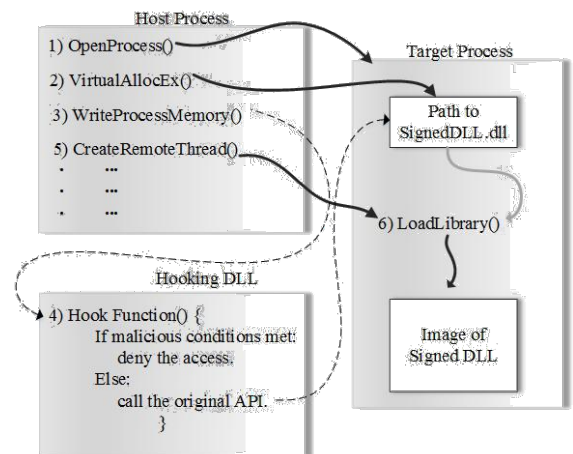


Figure 3: Structures in a Windows executable file

Since benign injections can be done in the normal way by writing only the path of the DLL into target process, there is no need to write the executable content directly, so the mitigation has no side effects on these benign injections.

C. System-Wide DLL Injection

We need a mechanism to load our mitigation engine DLL into all running processes upon their execution. To do this we have taken advantage of a system-wide DLL loading technique. A common method for system-wide DLL loading is the AppInit_DLLs infrastructure [9]. This mechanism loads an arbitrary list of DLLs in user-mode processes immediately after loading User32 DLL. However, it is not

enough as it does not load the DLLs in processes that don't load User32.dll. Like modern anti-virus products, we have written a kernel driver to implement an AppInit_DLLs-like infrastructure that loads our mitigation DLL immediately after loading Ntdll module instead of User32. This way, we will be ensured that the DLL is loaded in all windows processes, and the Mitigation is enforced system-wide. As mentioned earlier, the PsSetLoadImageNotifyRoutine is used to register a callback for Image Load events. This routine has the following signature:

```
NTSTATUS PsSetLoadImageNotifyRoutine(
    _In_ PLOAD_IMAGE_NOTIFY_ROUTINE
    NotifyRoutine
);
```

After setting this routine, whenever an Image Load event occurs our defined NotifyRoutine will be run with PUNICODE_STRING FullImageName, HANDLE ProcessId, PIMAGE_INFO ImageInfo, and BOOLEAN Create as input parameters. Our NotifyRoutine does the system-wide DLL injection in five steps:

- i. Check if the loading image is Ntdll. Ntdll is the first DLL that will be automatically loaded for every process on the system, and also contains the target API for hooking in our Mitigation DLL, the NtWriteVirtualMemory API.
- ii. Find the address of LdrLoadDll. Another reason to wait for Ntdll to be loaded is because we can parse the PE headers and find out the user mode address of LdrLoadDll. As explained in section 4, in user mode DLL injection, the LoadLibrary API is used for DLL loading, which is part of Kernel32 DLL. This API finally calls LdrLoadDll after some initializations. Thus, as we want to load our Mitigation DLL before loading Kernel32 DLL, we need to do the initialization in the callback routine and call the LdrLoadDll directly from Ntdll.
- iii. Prepare an assembly code to load the Mitigation DLL through LdrLoadDll call into target process. Since we are working on x64 Windows, we need to write two different x64 and x86 assembly codes, to call the LdrLoadDll with the name of the proper version of Mitigation DLL as input, into target 64- and 32-bit processes. Also, two distinct

versions of Mitigation DLL are placed in following directories:

64 bit : [Win_Drive]\Windows\System32

32 bit : [Win_Drive]\Windows\SysWOW64

- iv. Allocate memory into the target process and write the assembly code there. Since the callback is called in the context of the target process, we can simply use NtCurrentProcess() to specify what process the memory will be allocated and written into.
- v. Prepare an APC⁶ to call the assembly code. APCs allow user programs and system components to execute code in the context of a particular thread and, therefore, within the address space of a particular process. One advantage of APC is that it runs the code in the context of an existing thread and does not need to create a new thread for its operations, so makes it suitable for the case of system wide injection, as we need to load our DLL into all processes with no impact on performance. Following steps are required to add the code in the Thread APC Queue:

Find a thread in the target process

KeInitializeApc

KeInsertQueueApc

Then, the Mitigation DLL will be loaded into the process when the APC runs the assembly code. Finally, we have a mechanism like the AppInit_DLLs infrastructure that can load our Mitigation DLL in all processes immediately after loading the Ntdll. Our implementation codes for Mitigation Engine and System Wide Injection Driver are available in Github [28].

VI. EVALUATION

To evaluate the proposed mitigation and assess its efficiency, we first used PCMark benchmarking tool [29] to measure the impact of the new technique on the overall performance of the system. PCMark is one of a series of Windows performance testing tools that are provided by the Futuremark. It includes a variety of bench-mark tests reflecting the different ways people use their computers. Each benchmark produces detailed results for gaining a deep understanding of performance during each

individual workload. The technical guide in [29] explains specific tests the tool conducts on systems, and the formulas it uses to produce the scores.

We used a virtual machine with the following specifications in our experiments which is Windows 10 x64 Enterprise Build 14393 as Operating System, Intel Xeon X5670 @ 2.93 GHz @ 2933 MHz as CPU, 1 Core(s), 1 Logical Processor(s) and 8.00 GB Memory.

We selected 5 common benchmarks and measured the system performance while our mitigation engine is on or off. Based on the results provided in Table 1, the overall performance degradation is at most 0.59 percent which is very small and negligible. In fact, the mitigation DLL does not have significant influence on typical user activities like web browsing, text writing, video chat and others. Since the technique only checks the NtWriteVirtualMemory API calls for inter-process writes into UWP apps and this event is not very common in ordinary usages of the system, it doesn't have tangible impact on the system's usual functionalities and performance.

Next, we assess the proposed countermeasure's impact on NtWriteVirtualMemory which the specific API is involved in the mitigation. To do so, we called the API to write a 100 KB memory block into a target process for 10000 times and calculated the average time. The detailed results are provided in Table 2. Whenever a DLL is being written into an UWP process memory, the write operation will be aborted, and the user will be informed about the malicious activity, so the first row of the table is not a usual write operation and its overhead doesn't have any impact on the system's performance. If the target of the write operation is a non-UWP process, the mitigation will be stopped in the first step, and based on the results of the second and fourth rows of the table, its overhead impact is around 4.6 percent. However, if the writing content is not a DLL, and the target process is UWP, the mitigation mechanism will be stopped in the third step and will have an overhead around 6.5 %. However, since it doesn't occur commonly in the system, it doesn't have a tangible impact on the system overall performance, as shown in Table 1.

Finally, to assess the number of NtWriteVirtualMemory API calls during execution of common Windows programs, we

took advantage of API Monitor program [30] to illustrate the fact that the NtWriteVirtualMemory API call is not frequently used in prevalent Windows programs. API Monitor is a free monitoring tool that lets us monitor and control API calls made by applications and services. We selected a set of Windows programs, and ran each program for five minutes, to check call frequency of NtWriteVirtualMemory API. As illustrated in Table 3, call frequency of the API is at most 0.0002% in Google Chrome application.

Table 1: Overall Performance Impact on System (Time-Based).

| Benchmark | Normal | Hooked | Overhead % |
|--------------------------------------|----------|----------|------------|
| Web Browsing - JunglePin | 0.373 s | 0.375 s | 0.54 |
| Web Browsing - Amazonia | 0.141 s | 0.141 s | 0.0 |
| Writing | 6.31 s | 6.31 s | 0.0 |
| Phone Editing v2 | 1.867 s | 1.878 s | 0.59 |
| Video Chat v2/Video Chat Encoding v2 | 704.7 ms | 706.5 ms | 0.26 |

Table 2: Performance Impact on NtWriteVirtualMemory API.

| Content is DLL | Target is UWP | Normal ms | Hooked ms | Overhead % |
|----------------|---------------|-----------|-----------|------------|
| ✓ | ✓ | 0.0481 | 1.0270 | 2035.14 |
| ✓ | ✗ | 0.0482 | 0.0504 | 4.56 |
| ✗ | ✓ | 0.0480 | 0.0511 | 6.46 |
| ✗ | ✗ | 0.0481 | 0.0503 | 4.57 |

Table 3: NtWriteVirtualMemory Call Frequency in Windows Programs

| Program | NtWriteVirtualMemory Call | Total Number of Call |
|----------------------------|---------------------------|----------------------|
| Vmware Workstation | 1 | 2330721 |
| Telegram | 0 | 2495273 |
| Twitter | 0 | 1658813 |
| Spark Instant Messenger | 0 | 5255403 |
| Notepad++ | 0 | 1317342 |
| Windows Media Player | 1 | 12404209 |
| VLC Media Player | 0 | 11506828 |
| TeamViewer | 0 | 6176240 |
| Mozilla Firefox | 0 | 15501030 |
| Google Chrome | 23 | 10272584 |
| Microsoft Edge | 0 | 2464095 |
| Wireshark Network Analyzer | 5 | 2843753 |
| Internet Download Manager | 0 | 4944558 |

VII. CONCLUSION

In this paper, we studied the issue of reflective DLL injection attacks on UWP apps and proposed a defence mechanism to counter such attacks. We discovered that despite the embedded security mechanism in UWP framework, it is still possible to inject malicious/unsigned DLLs into UWP apps even in the presence of an antivirus software. To defend against these attacks, we proposed a mechanism that monitors the input parameters to `NtWriteVirtualMemory()` API and aborts malicious DLL injection attacks. We implemented the proposed idea by leveraging the hooking libraries and Windows kernel callbacks. This allows us to monitor the processes and prevent malicious injections into UWP apps while allowing the benign injections to proceed as normal.

VIII. REFERENCES

- [1] Barabosch, T., Eschweiler, S., & Gerhards-Padilla, E. (2014). Bee master: Detecting host-based code injection attacks [Conference Proceedings]. In International conference on detection of intrusions and malware, and vulnerability assessment (p. 235-254). Springer.
- [2] Fewer, S. (2008). Reflective DLL injection [Journal Article]. Harmony Security, Version, 1.
- [3] Staples, D. (2015). Improved reflective DLLinjection [Web Page]. <https://github.com/dismantl/ImprovedReflectiveDLLInjection>.
- [4] Mertsarica. (2010). Antimeter tool [Web Page]. <https://www.mertsarica.com/antimeter-tool/>.
- [5] King, A. (2012). Detecting reflective injection [Web Page]. <https://www.defcon.org/html/defcon-20/dc-20-speakers.html#King>. DEF CON R 20 Hacking Conference.
- [6] Berdajs, J., & Bosnic, Z. (2010). Extending applications using an advanced approach to DLL injection and API hooking [Journal Article]. Software: Practice and Experience, 40(7), 567-584.
- [7] Richter, J. (1994). Load your 32 bit DLL into another process's address space using injlib [Journal Article]. Microsoft Systems Journal-US Edition, 13-40.
- [8] Lam, L.-c., Yu, Y., & Chiueh, T.-c. (2006). Secure mobile code execution service. In Proceedings of the 20th conference on large installation system administration (pp. 5-5).
- [9] Help, M., & Support. (2010). Working with the appinit DLLs registry value [Web Page]. <https://support.microsoft.com/en-us/help/197571/working-with-the-appinit-dlls-registry-value>.
- [10] Kuster, R. (2003). Three ways to inject your code into another process [Web Page]. <https://www.codeproject.com/Articles/4610/Three-Ways-to-Inject-Your-Code-into-Another-Process>
- [11] Newcomer, J. M. (2001). Hooks and DLLs [Web Page]. <https://www.codeproject.com/Articles/1037/Hooks-and-DLLs>.
- [12] Shewmaker, J. (2010). Analyzing DLL injection [Web Page]. <http://www.bluenotch.com/>.
- [13] NTCORE. (2012). Explorer suite [Web Page]. www.ntcore.com/exsuite.php.
- [14] Barabosch, T., & Gerhards-Padilla, E. (2014). Host-based code injection attacks: A popular technique used by malware [Conference Proceedings]. In Malicious and unwanted software: The americas (malware), 2014 9th international conference on (p. 8-17). IEEE.
- [15] Jang, M., Kim, H., & Yun, Y. (2007). Detection of DLL inserted by windows malicious code [Conference Proceedings]. In Convergence information technology, 2007. international conference on (p. 1059-1064). IEEE.
- [16] Glendowne, D., Miller, C., McGrew, W., & Dampier, D. (2015). Characteristics of malicious DLLs in windows memory [Conference Proceedings]. In Ifip international conference on digital forensics (p. 149-161). Springer.
- [17] VoxelBlock. (2016). Basic and intermediate techniques of uwp app modding [Web Page]. <https://www.unknowncheats.me/forum/general-programming-and-reversing/177183-basic-intermediate-techniques-uwp-app-modding.html>.
- [18] Yucheng, G., Peng, W., Juwei, L., & Qingping, G. (2011). A way to detect computer trojan based on DLL preemptive injection [Conference Proceedings]. In Distributed computing and applications to business, engineering and science (dcabes), 2011
- [19] Sun, H.-M., Tseng, Y.-T., Lin, Y.-H., & Chiang, T. (2006). Detecting the code injection by hooking system calls in windows kernel mode [Conference Proceedings]. In 2006 international computer symposium, ics.
- [20] DLL [Web Page]. <http://www.codeguru.com/cpp/g-m/directx/directx8/article.php/c11453/>

- Intercept-Calls-to-DirectX-with-a-Proxy-DLL.htm.
- [21] Microsoft. (n.d.-a). Process mitigation binary signature policy structure [Web Page]. [https://msdn.microsoft.com/en-us/library/windows/desktop/mt706242\(v=vs.85\).aspx](https://msdn.microsoft.com/en-us/library/windows/desktop/mt706242(v=vs.85).aspx)
 - [22] Cowan, C. (2015). Protecting microsoft edge against binary injection [Web Page]. <https://blogs.windows.com/msedgedev/2015/11/17/microsoft-edge-module-code-integrity/>.
 - [23] Rascagneres, P. (2016). Microsoft edge binary injection mitigation overview [Web Page]. <http://www.sekoia.fr/blog/microsoft-edge-binary-injection-mitigation-overview/>.
 - [24] Microsoft. (2002). Detours [Web Page]. <https://www.microsoft.com/enus/research/project/detours/>.
 - [25] Mhook, an API hooking library [Web Page]. (2014). <https://github.com/martona/mhook>.
 - [26] Deviare API hook [Web Page]. (2017). <http://www.nekra.com/products/deviare-api-hook-windows/>.
 - [27] Easyhook the reinvention of windows API hooking [Web Page]. (2017). <https://easyhook.github.io/>.
 - [28] Zaheri, M., & Niksefat, S. (2017). Github project for preventing reflective DLL injection on UWP apps [Web Page]. <https://github.com/m0jt4b4/UWPHardening>.
 - [29] FutureMark. (2016a). Pcmark 8: The complete benchmark for windows [Web Page]. <http://www.futuremark.com/benchmarks/pcmark>.
 - [30] rohitab.com. (2017). APIMonitor: Spy on API calls and COM inter-faces [Web Page]. <http://www.rohitab.com/apimonitor>.

Crawler and Spiderin usage in Cyber-Physical Systems Forensics

M. Abedi¹ and Sh. Sedaghat²

¹Jahrom University APA CENTER, Jahrom, Iran

²Faculty of Information Technology Engineering Department, Jahrom State University, Jahrom, Iran
clvmoein@gmail.com, shsedaghat@jahromu.ac.ir

Abstract - As a featured subset of cyber-physical-systems, Mobile cyber-physical-systems can make use of Mobile devices, such as smartphones, which serve as a convenient and economical platform for Mobile applications in all places between humans and the geographic world around it. Today, cyber physical systems are popular in power grids, healthcare devices, transportation networks, industrial processes and infrastructure. Cyber- physical systems (CPS) are used more widely, the security of physical cyber systems in the design, implementation, and research of the system is very important. Various types of attacks in the cyber-physical-system (e.g. Stuxnet worms) cause severe casualties and potentially serious security risks. Over the past few years, researchers have focused on aspects of the security of cyber-physical systems. In this paper, after analysing CPS security objectives and CPS security approaches, we propose a security technique to provide security and improve intrusion detection methods for cyber-physical systems, which is used to improve CPS immunization. Mobile CPS that has expanded the benefits and scope of CPS applications in recent years has become increasingly popular. For example, mobile CPS can be a kind of basic techniques to support the development of transport network systems, thus protecting the privacy and security of users in the dynamic transport environments Improves. In this article, we first recognize the Mobile CPS of the traditional CPS. Then, we recommend a solution using the Crawling and Spidering techniques used in search engines to detect and cope with the influence of information security systems

KEYWORDS - Cyber-Physical System Security, Intrusion Detection, Information Security, Crawling, Spidering

I. INTRODUCTION

Cyber-Physical System or CPS combining the physical world with cyber-components is a key research field for more than a decade [1]. Traditional CPS is effective in many engineering projects such as intelligent power grids, manufacturing systems, aerospace systems and defence systems [2]. Today, with the development of inclusive Mobile devices, Mobile CPS has attracted more attention. Compared to the traditional CPS, which rely on fixed machines or massive sensors and emphasizes the use of cyberparks to dominate the physical world, the Mobile CPS focuses on its mobility, which can be integrated seamlessly and everywhere. Everyday life gets people. Therefore, Mobile CPS can easily be used in each person's life and be deployed in a wider range of physical worlds.

Although some may believe that Mobile CPS is a subset of traditional CPS [3], this is not the case, because they have unique features that offer opportunities in many functional areas that traditional CPS cannot do it. Because Mobile devices are equipped with a variety of sensors, the Mobile CPS benefits

from the continued acquisition of information in the physical world. So, compared with traditional CPS, the Mobile CPS can have much more information resources and can analyse physical systems with more data. Additionally, Mobile CPS integrates traditional features of the CPS with the help of technology development, benefiting from their combination.

Therefore, the Mobile CPS is not a subgroup of the traditional CPS but overlapping it. Due to this characteristic, there is a common challenge for the traditional CPS and Mobile CPS, and some examples are shown as a subset between traditional CPS and Mobile CPS in Figure 1. Additionally, due to the fact that the traditional CPS and Mobile CPS share common challenges and some similarities in the architecture of the system, some traditional CPS solutions for Mobile CPS can also be used. However, as shown in Figure1, since Mobile CPSs are more than a subset of the CPS, they have particular challenges, including Mobile device power constraints, unstable Mobile networks, and very dynamic environments.

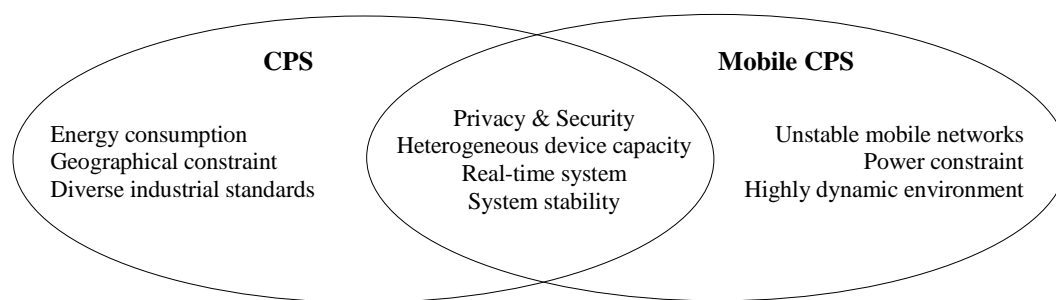


Figure 1: Relation between traditional CPS and Mobile CPS

CPS can be described as intelligent systems that comprise computing components (i.e., hardware and software) and physical components that act seamlessly and closely together to control the changing real-world situation. The prevalence and vulnerability of CPS has left researchers and influencers focused on these systems. In order to ensure the safety of Mobile cyber-physical security systems, there are several security goals to achieve, including six major security objectives: Confidentiality, integrity, availability, robustness, reliability and trustworthiness.

Compared to Internet attacks, it is more difficult to detect and prevent attacks on the CPS goal. To prevent intrusion detection, hackers may apply multiple steps and combine types of attacks to access a traditional or Mobile Cyber-physical system. The continuous integration of cloud technology in all aspects of our daily lives creates business opportunities, operational risks, and research challenges. But as companies continue to provide services and increase access to customers and employees, they continue to expand software access and create new supply chain management chains, the risk of cyber-physical attacking increases. Increasing the level of digital communication between physical devices (such as sensors and thyristors) and cyber-equipment (such as intelligent decision-making systems), CPS (such as power grids) has turned to large ecosystems that require a scalable and flexible infrastructure. Integrating cyber-physical-systems through a cloud computing infrastructure is a Cyber-Physical Cloud or CPC that not only potentially improves the interaction between cyber-physical devices, it also provides the ability to store and analyse large-scale data [4]. News organizations are increasingly highlighting the dangers of integrating this technology. For example, another article cited a cyber-physical attack report that had damaged an explosive furnace

in a steel plant in Germany. An excellent example of an attack on a cyber-physical system is the Stuxnet virus that targets Iran's nuclear power plant and reduces the efficiency of systems [5].

In fact, moving from a cyber-physical-network to the cloud can lead to various security issues. There are only a few cyber crime cases known in CPS, but a successful attack could have catastrophic consequences. A recent survey found that the role of digital forensic in managing CPC incidents was not well understood [6]. Although Digital forensic tools and techniques are unlikely to stop an attack in real-time, a forensic approach to design can help provide several methods. For example, this approach can help identify an incident by its source and determine its type, maintain and analyse critical vital data, rebuild parts of the data, and obtain results and speed. Microsoft proposes a "assume breach" approach to cloud security - an innovative design, engineering, and operational approach that predicts an attack has already occurred [7]. Ensuring the environment is like a castle because of the asymmetric nature of cyber space. For example, to protect an information space, Kaspersky should ensure that different security technologies are in place, all systems are installed in time, and so on.

However, an internet attacker should only have one or more vulnerabilities in the network to attack and exploit them.

In a security incident, referral plays an important role in research, such as tracking and identifying the source of the attack. This can be facilitated by a digital pharma. Researchers have highlighted potential issues in digital forensic research in cloud environments, such as the appropriateness of data recording techniques, tools, multiple sources of evidence, and qualification issues.

II. CYBER-PHYSICAL CLOUD SYSTEMS

The continued amalgamation of cloud technology into all aspects of our daily lives creates business opportunities, operational risks, and investigative challenges. But as businesses continue to offer customers and employees increased access, improved software functionality, and new supply chain management opportunities, the risk of cyber-physical attacks on CPCs grows. Increasing digital interconnectivity between devices at the physical (such as sensors and actuators) and cyber (such as intelligent decision systems) levels has transformed CPS (such as the electric power grid) into large ecosystems requiring a scalable and flexible infrastructure.

In reality, moving from an internal cyber-physical network to the cloud can lead to various security issues. There are only a few known cyber-attack incidents on CPS, but a successful attack can have real-world and catastrophic consequences. A recent survey suggested that the role of digital forensics in CPCs incident handling isn't widely understood.

As technology dependency and cloud integration continue to escalate, ensuring CPCs security becomes a critical factor in delivering trustworthy and robust services. The nature of Cyber-physical and cloud computing infrastructures, however, presents inherent challenges to ensuring data confidentiality, integrity, and availability.

A. Risk Management Principles and Practices

It would be unrealistic to expect any organization to have infinite resources to identify and act on all potential threats and risks. Therefore, based on the "assumed breach" approach[7], to achieve CPCs systemic resilience the system developer and forensic expert need to adopt risk management principles and practices to identify and prioritize current and emerging threats (for example, potential vulnerabilities in both cloud computing and CPS and how these vulnerabilities can be exploited), risk areas (including risks arising from unexpected and highly unpredictable causes, also known as the "black swan" problem), and potential evidence source and type (see the forensic readiness principles).

B. Incident-Handling Principles and Practices

Guiding principles and practical strategies can minimize the impact of loss after a

security incident and help prevent and mitigate future incidents. As earlier work noted, incident handling and digital forensic practices overlap, and both practices should be integrated into an incident-handling strategy [6]. For example, intrusion detection systems can help determine attack sources. In addition, having a forensic database (for pre-incident collection) would benefit incident responders during a preliminary incident response. In earlier work, Grispos and his colleagues note that organizations have opportunities to strengthen policies, standards, and procedures prior to migrating to cloud environments. Organizations need to investigate these opportunities from a CPCS perspective. Additional work by Grispos and his colleagues in the area of security incident response criteria demonstrate the type of industry practices that need to be identified and verified for CPCS incident handling. However, we need to ensure that activities undertaken during incident handling (for example, evidence collection) don't result in service disruption, and therefore system backup and redundancy must be carefully planned in incident handling.

C. Laws and Regulations

When designing forensic strategies, it's important to consider international and local legal and regulatory requirements, because different national laws and regulations might have different evidence requirements. A law designated for data protection might only be applicable to the country in which the data resides, for example. In some scenarios, cloud providers might be required to comply with a court order and surrender user data without notifying the data owner. Relevant standards and industry best practices should also be considered in the design and development phases. The Payment Card Industry-Data Security Standard (PCI-DSS), for instance, mandates regular monitoring of access to network resources, which would require the system to include an efficient logging capability for compliance purposes as well as the digital evidence source.

D. CPC Hardware and Software Requirements

The interdependencies between hardware and software within a CPCS complicate the identification and collection of evidential data. Potential evidence artefacts would exist across several CPC layers (for example, from field devices to cloud aggregators); thus, providing an embedded forensic agent is a potential

solution to remotely collecting the evidential data. Furthermore, specific communication protocols used in cyber-physical systems, such as ModBUS, to control field devices would require a customized forensic approach as compared to the common network protocol (for example, TCP/IP). Understanding hardware and software requirements are, therefore, critical in supporting the collection of forensically sound evidence.

E. Industry specific requirements

Because of the diversity of cyber-physical components (for example, sensor, controller, and networked systems) and data types (for example, sensor data from in-vehicle systems are quite different from sensor data from power grid systems), we must also consider industry-specific (for example, energy, automotive, and transportation) requirements. Therefore, identifying and collecting evidence data sources requires careful planning. Moreover, each industry has a different security risk profile, which would affect the choice of forensic strategies.

F. Validation and Verification

Once a prototype of the system has been designed and developed, it's important to validate and verify to ensure that the evidence collected is adequate and reliable, and that the forensic processes and functions used are sound (for example, there's no contamination of evidence). As Yinghua Guo and his colleagues discuss, "validation refers to the confirmation by examination and the provision of objective evidence that a tool, technique or procedure functions correctly and as intended" and "verification is the confirmation of a validation with laboratories tools, techniques and procedures." [8].

Ensuring reliable evidence data is an important aspect of producing digital evidence that's admissible in a court of law (that is, forensically sound). We can use Rodney McKemmish's criteria as guidelines to establish forensic soundness [9]:

- *Meaning.* Design digital forensic processes that won't change the data's meaning.
- *Error.* Design digital forensic processes that can avoid undetectable error. If an error is encountered when undertaking forensic processes, it must be identified and explained as evidence.
- *Transparency.* Verify evidence by documenting the chain of custody,

including identifying the forensic software and hardware used, detailing the analysis environment, and specifying any problems, errors, and inconsistencies throughout the forensic processes.

- *Experience.* Be sure to task an individual with sufficient and relevant expertise with finding digital evidence.

Assurance refers to the measurement of forensic processes and functions using relevant metrics, such as those involving security incidents, maturity level, and IT performance, and can include incident simulation or testing (for example, penetration testing) as input. The system designer can refine the CPCs based on the validation and verification results before finalizing. As part of the final check, the designer defines a set of actions that constitutes a strategy for incident handling and creates (or updates) digital forensic practices to manage incident occurrence in the product's post release phase.

Any problems resulting from the validation and verification will involve refining the related factors. The completed CPCs should be forensically ready in the aforementioned key areas. To sum up, defining and planning what evidence will be required ensures that better security mechanisms and architecture are in place, and that they can provide the evidence when it's required.

Internet search engines use two crawling and spidering capabilities to get information from web space. On these search engines like Google and Bing, the spider is responsible for loading the pages and the crawler plays the role of commander in the spider. In fact, the crawler decides which pages to load, and ultimately the spider is responsible for loading [10].

III. RELATED RESEARCHS ON CPS SECURITY TECHNIQUES

[11] provides an overview of smart grid operation, associated cyber infrastructure and power system controls that directly influence the quality and quantity of power delivered to the end user. The paper identifies the importance of combining both power application security and supporting infrastructure security into the risk assessment process and provides a methodology for impact evaluation. A smart grid control classification is introduced to clearly identify communication technologies and control messages required to support these control functions.

Table 1 summarizes the most and least studied IDS techniques in the literature grouped by the application type in the order of most to least.

We see that for all applications studied, the most commonly used configurations are behavior-based detection techniques and host-based auditing. Table I indicates that there is little research with regard to automotive applications, knowledge-based detection techniques and network-based auditing.

[12] developing mobile cyber-physical

and system-theory-based security are essential to securing cyber-physical systems.

Vita, a novel mobile CPS for crowdsensing, which leverages the advantages of social computing, service computing, cloud computing, and a number of open source techniques across mobile devices and cloud platform, to provide a systematic approach that supports both application developers and users for mobile crowdsensing applications have been presented in [15].

[16] introduces various research

Table 1: Most and Least Studied IDS Techniques, by Citations (some used more than one detection technique)

| CPS Application | Detection Technique | Audit Material | Unique CPS Aspects |
|--------------------|--|--------------------------|--|
| Smart utility (18) | Behavior (10) Behavior-Specification (6) Knowledge (3) | Host (11) Network (7) | Physical Process Monitoring (8) Closed Control Loops (2) Attack Sophistication (9) Legacy Technology (14) |
| SCADA (6) | Behavior (5) Behavior-Specification (1) Knowledge (1) | Network (5) Host (1) | Physical Process Monitoring (1) Closed Control Loops (0) Attack Sophistication (1) Legacy Technology (2) |
| Medical (3) | Behavior (2) Behavior-Specification (1) Knowledge (0) | Host (3) Network (0) | Physical Process Monitoring (1) Closed Control Loops (0) Attack Sophistication (1) Legacy Technology (2) |
| Aerospace (2) | Behavior (1) Behavior-Specification (1) Knowledge (0) | Host (2) Network (0) | Physical Process Monitoring (1) Closed Control Loops (0) Attack Sophistication (0) Legacy Technology (2) |
| Automotive (1) | Behavior (1) Behavior-Specification (0) Knowledge (0) | Host (1) Network (0) | Physical Process Monitoring (0) Closed Control Loops (0) Attack Sophistication (1) Legacy Technology (0) |

applications in the context of WreckWatch and related projects yielded some lessons, like: Many components of the solutions are highly related, Analysis of properties, such as safety, that span a combination of devices and services is difficult, Factoring social/human properties of systems into system analysis is not well understood, It is hard to integrate mobile Internet devices with conventional sensor networks, Individual mobile devices are prone to unexpected unavailability.

In [13], Researchers developed a mathematical model to analyse survivability of a mobile cyber physical system (MCPS) comprising sensor-carried mobile nodes with voting-based intrusion detection capabilities.

[14] shows that cyber-physical system security demands additional security requirements, such as continuity of power delivery and accuracy of dynamic pricing, introduced by the physical system. Such requirements are usually closely related to the models and states of the system, which are difficult to address by information security alone. Therefore, both information security

applications which required cyber-physical testbeds to provide representative environments to explore and validate potential solutions.

[17] explores the development of a probability model to analyse the reliability of a cyber physical system (CPS) containing malicious nodes exhibiting a range of attacker behaviours and an intrusion detection and response system (IDRS) for detecting and responding to malicious events at runtime.

The paper [18] gives a comprehensive review on CPS security following the security framework from diverse perspectives.

The forensic-by-design framework presented in [19] provides a starting point for conversations, research and solutions that could be used to address this issue.

[20] Authors have introduced the applications and key challenges and techniques of mobile CPS and distinguished them from the traditional CPS.

IV. A SOLUTION FOR INTRUSION DETECTION IN CYBER-PHYSICAL SYSTEMS

We will explain our method in 3 phases:

Phase 1- We clear how does our solution can be implemented in software and hardware and infrastructure.

Phase 2- We tested the solution for a case by using OPManager software and show the results.

Phase 3- We explain the role of mobile CPS and spidering and crawling techniques in our method.

A. Phase 1

As previously mentioned, the property of Identifying and collecting information around the whole surface web on the search engines is the responsibility of a technique called “spidering”, and after identifying web pages, the spider tells the crawler the necessary commands. Our proposed strategy, including the use of this Web space feature in intrusion detection systems, which, of course, requires cyber-physical cloud systems to manage it. Our proposed strategy includes the following steps:

- **Step 1:** Monitor the critical security and firewall systems throughout the network in the medium-term and long-term time periods (to defining a true network state pattern in traffic, active devices and so on). At this stage, first of all, we should provide an environment that includes samples of our real internal network of organisation components. This environment could be a virtual space or a real- local network in some place. The important issue about the real or virtual network environment is that it must include exact hardware instruments, software applications and cloud technology infrastructures. for virtualization such an environment, we could use different software such as VMware, GNS3, Cisco Packet tracer, etc. and if we want to have a real local network to find out the true state of network and monitor different components of network, like network traffic and users’ activities in network, there are useful software such as: OPManager, PRTG, SolarWinds, and etc. Our next action is to monitor all hardware, software and cyber-physical cloud systems infrastructure activities of the cyber-physical and network system before the

launch and introduction of the related system and infrastructure, and more critical and more important than previous actions is storing the information has been obtained in a secure and secret database and protect it from stealing or injecting information from the database. It should be noted that this stage is being implemented only by IT security professionals who are fully trusted by the organization, and no internal or external staff are aware of the implementation of this phase.

In fact, our goal to implementing this step, is to determine and store the normal and ideal functional conditional of our isolated (not connected to the internet) internal organisation’s network.

Now and after the implementation of the first step, we determined the normal state of the network and we know the whole information around the internal network when it does not have any malware, spyware and abnormal traffics.

- **Step 2:** Monitor all hardware and software parts and critical security components and systems traffic throughout the network. In the second step, after introducing and launching the system, we examine all of the network traffics and system activities in real-time (Current network status). Some software like OPManager, PRTG, SolarWinds can be useful for monitoring the whole CPS network properties such as bandwidth or memory usage.
- **Step 3:** Match and compare the information obtained in the first step with the data collected in the second stage. In the third step, you can compare the current status and performance of the network with the normal state of the CPS, and if you see the slightest change to the ideal function, check this change, identify the suspect and all the information and potential hazards around the change of the information and report them to the information security specialist.

B. Phase 2

For example; we have monitor memory usage in a practical CPS case by using OPManager and inserted the result in Figure 2. As shown in Figure 2, the amount of memory used in the network and server during the test (yellow lines) is greater than the normal amount of

memory consumption that should be taken in a natural and safe manner according to the pattern (green lines). The blue lines in the form show the maximum amount of memory usage tolerance by the infrastructure on the network. As long as the distance between the blue lines and the yellow lines is lower, the problem with the server and other components and network infrastructure is more likely to occur, and reporting and processing need to be done faster.

Now and after the implementation of our solution in first to phases, we should start phase 3 and detect the abnormal issues through the network and report them to the information security administrator of the organization.

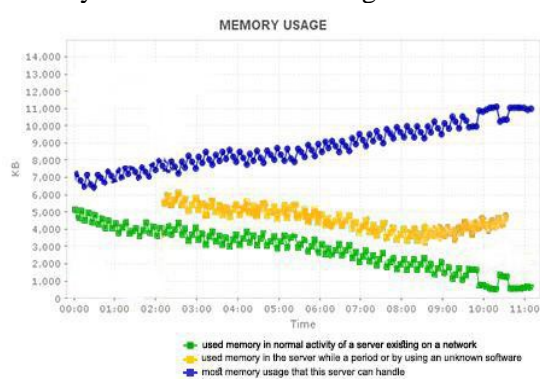


Figure 2: Memory usage in a practical CPS case compared with the normal memory usage

C. Phase 3

Informing the organization's security authorities can be done using the cloud computing, Fog computing and cyber-security tools. In this way, changes made after Real-Time analysis are reported to security administrators via Fog computing technology (which speeds up the operation of cloud computing), and they are also using Mobile cyber-physical devices that always have the ability to set Crawler in a way that disrupts the performance of a malicious or intruder after it is detected and prevents potential attackers from causing damage.

In recent years, the capabilities of Mobile devices have improved dramatically. These features, such as impressive computing resources, multiple radios, sensor modules and high-level programming languages enable Mobile devices to create a Mobile cyber-physical system in our everyday lives. Mobile CPS is the result of the integration of distributed sensors with computing and connectivity all over the internet. It also integrates Mobile CPS, computing, cyber and physical resources, and facilitates the interaction of the digital world with the physical world, and potentially enriches the

everyday life of citizens anytime and anywhere. Therefore, the Mobile CPS can be a convenient and affordable platform that facilitates complex and all-round intelligent applications between humans and the physical world around them.

Mobile CPS can be used in various fields including (1) Mobile smart robots and robotic systems, The use of multiple smart sensors, Mobile devices, Intelligent services, Cloud robots, and Improving the efficiency and scalability of complex work processing that is not feasible under the constraints of local resources in different application areas; (2) Intelligent transportation systems, for example, The ability to measure, calculate and communicate with control vehicles in the physical world; To deal with safe challenges (for example, reducing latency in response to traffic accidents), Efficient transportation Fashion and green; for example, Smart city, environmental monitoring, health systems and smart grids, which improves information, comfort, operational Safety and green energy of the human community. Solutions that are defined by software, Distributed systems, Cloud computing, social networking, Security and privacy, Human-centred computing, and other methods and technologies that can be used for moving CPSs are also welcome. The last phase of our proposed solution has two main steps:

- Step 1: Detection any malicious activity on the Network;
In this step; we need a special spider and crawler to constantly search the different parts of the network and compare current status of network with the normal state.
- Step 2: inform and alarm the information security staff personals and administrators throughout the Mobile CPS to make the network secure.

If our spider and crawler found detect any differences between current and normal states of network, then is time to use Mobile CPS technology to be useful for inform the intrusion detection to security managers and help to make the network more secure.

V. FUTURE WORKS

Several research fields that facilitate the deployment and securing use of Mobile CPS include:

- Architectural platform for distributed Mobile CPS
- Smart Mobile Robots and Robotic Systems
- Software Solutions for Mobile CPS
- Smart city and smart grid technology
- Man-centric calculations in mobile CPS
- Automobile networks and intelligent transportation systems
- Evaluations and security solutions, privacy, and issues related to the reliability of the Mobile CPS
- Distributed intelligent systems and applications
- Mobile social networks and inclusive apps
- Mobile cloud computing
- Mobile Service-centric and calculations
- Design and optimization of Mobile CPS
- Asymmetric networks in Mobile CPS
- Intelligence processing for Mobile CPS
- Big data analysis on Mobile CPS
- Data mining, machine learning, and sophisticated system design for Mobile CPS
- Scalable monitoring systems with Mobile wireless networks
- Resource Management in Mobile CPS
- Experience to deploy real-world Mobile CPS

VI. CONCLUSION

We first provided some explanations about CPS and named their variants, in terms of the differences and similarities between traditional CPS and Mobile CPS and the security objectives for CPS systems. By pointing out the features of CPS, we conclude that intrusion detection and the prevention of attack on these scalable systems are of great importance in the industry, security systems and even the lives of people every day. Then, using a common technique in internet search engines, such as Spidring and Crowling, have proposed a strategy and idea to detect malicious devices, hacker activities, and manipulate the information network by unauthorized persons. In our proposed approach, the security experts of any organization that needs to protect the information of their organization can remotely attack the attackers and those who intend to sabotage the organization's information space and neutralize their actions. In the end, we also looked at Mobile CPS, and several research areas were proposed to improve the security of the Mobile CPS forensics.

VII. REFERENCES

- [1] L. Sha, S. Gopalakrishnan, X. Liu, and Q. Wang, "Cyber-physical systems: A new frontier", in Proc. IEEE Int. Conf. Sensor Netw., Ubiquitous Trustworthy Comput. (SUTC), Jun. 2008, pp. 1–9.
- [2] R. R. Rajkumar, I. Lee, L. Sha, and J. Stankovic, "Cyber-physical systems: The next computing revolution", in Proc. 47th Design Autom. Conf. ACM, 2010, pp. 731–736.
- [3] T. Hanz and M. Guirguis, "An abstraction layer for controlling heterogeneous Mobile cyber-physical systems", in Proc. IEEE Int. Conf. Autom. Sci. Eng. (CASE), Aug. 2013, pp. 117–121.
- [4] S. Karnouskos, A.W. Colombo, and T. Bangemann, "Trends and Challenges for Cloud-Based Industrial Cyber-Physical System", *Industrial Cloud-Based Cyber-Physical Systems*, A.W. Colombo et al., eds. Springer Int'l Publishing, 2014, pp. 231–240.
- [5] R. Langner, "Dissecting a Cyberwarfare Weapon", *IEEE Security & Privacy*, vol. 9, no. 3, 2011, pp. 49–51.
- [6] N.H. Ab Rahman and K.-K.R. Choo, "A Survey of Information Security Incident Handling in the Cloud", *Computer Security*, vol. 49, Mar. 2015, pp. 45–69.
- [7] Microsoft, "Microsoft Enterprise Cloud Red Teaming", 2014; http://download.microsoft.com/download/C/1/9/C1990DBA-502F-4C2A-848D-392B93D9B9C3/Microsoft_Enterprise_Cloud_Red_Teaming.pdf.
- [8] Y. Guo, J. Slay, and J. Beckett, "Validation and Verification of Computer Forensic Software Tools—Searching Function", *Digital Investigations*, vol. 6, 2009, pp. 12–22.
- [9] R. Mckemmish, "When Is Digital Evidence Forensically Sound?", *Advances in Digital Forensics IV*, I. Ray and S. Sheno, eds., Springer, 2008, pp. 3–15.
- [10] N.H. Ab Rahman, W.B. Glisson, Y. Yang, K.-K.R. Choo, "Forensic by Design Framework for Cyber-Physical Cloud Systems", *IEEE Cloud Computing*, 2016.
- [11] Siddharth Sridhar, Adam Hahn, Manimaran Govindarasu, "Cyber-Physical System Security for the Electric Power Grid", *Proceedings of the IEEE*, 2012.
- [12] Jules White, Siobhan Clarke, Christin Groba, Brian Dougherty, Chris Thompson, Douglas C. Schmidt, "R&D Challenges and Solutions for Mobile Cyber-Physical

- Applications and Supporting Internet Services*”, Journal of Internet Services and Applications.
- [13] Robert Mitchell, Ing-Ray Chen, “*On Survivability of Mobile Cyber Physical Systems with Intrusion Detection*”, Springer Science and Business Media, 2012.
- [14] Yilin Mo, Tiffany Hyun-Jin Kim, Kenneth Brancik, Dona Dickinson, Heejo Lee, Adrian Perrig, and Bruno Sinopoli, “*Cyber-Physical Security of a Smart Grid Infrastructure*”, Proceedings of the IEEE, 2012.
- [15] Xiping Hu, Terry H. S. Chu, Henry C. B. Chan, Victor C. M. Leung, “*Vita: A Crowdsensing- Oriented Mobile Cyber-Physical System*”, IEEE Transactions on Emerging topics in Computing, 2013.
- [16] Adam Hahn, Aditya Ashok, Siddharth Sridhar, Manimaran Govindarasu, “*Cyber-Physical Security Testbeds: Architecture, Application, and Evaluation for Smart Grid*”, IEEE Transactions on smart grid, 2013.
- [17] Simrandeep Kaur chana, S. J. Karale, “*Analysis of Intrusion Detection Response System (IDRS) In Cyber Physical Systems (Cps) Using Regular Expression (Regexp)*”, IOSR Journal of Computer Engineering, 2014.



OIC-CERT Permanent Secretariat

CyberSecurity Malaysia, Level 5, Sapura@Mines, 7, Jalan Tasik
The Mines Resort City, 43300 Seri Kembangan
Selangor Darul Ehsan, Malaysia
secretariat@oic-cert.org