airiti Press

GPN 2007800019
ISSN 1018-1393
DOI: 10.6131/MISR

MIS REVIEW

Volume 21 Numbers 1/2 September 2015/March 2016

# MIS REVIEW
## An International Journal

Vol.21 Nos.1/2 September 2015/March 2016

*MIS Review* **Vol. 21, Nos. 1/2** Published in September 2015/March 2016

Publisher: Eldon Y. Li

Published by: National Chengchi University, Department of Management Information Systems & Airiti Press Inc.

Editor-in-Chief: Eldon Y. Li

Executive Editor: Shari S.C. Shang

Assistant Editor: Laurence Fang-Kai Chang

Printed by: Sincere Digital Printing, Co.

Cover Designer: Thomes Chen

Typesetting: Hsiao-Hsuan Wang

Publication Office:

National Chengchi University, Department of Management Information Systems
No. 64, Sec. 2, ZhiNan Rd., Wenshan District, Taipei City 116, Taiwan

Airiti Press Inc.
18F., No. 80, Sec. 1, Chenggong Rd., Yonghe District, New Taipei City 23452, Taiwan

Order Information:

Airiti Press Inc.
18F., No. 80, Sec. 1, Chenggong Rd., Yonghe District, New Taipei City 23452, Taiwan
Tel: +886-2-29266006
Fax: +886-2-29235151
E-mail: press@airiti.com

Price: NT$ 400

e-Journal: http://www.airitilibary.com

DOI: 10.6131/MISR

ISSN: 1018-1393

GPN: 2007800019

Printed in Taiwan

## Editorial Board

# Editor's Introduction

In this MISR issue, we are delighted to present four research papers. The summary of the four papers is as follows.

S. Raj Anand and E. Kannan in their paper "Probabilistic Approach of Improved Binary PSO Algorithm Using Mobile Sink Nodes" show that in Wireless Sensor Network (WSN) applications for efficient data accumulation, the use of mobile sinks plays a very important part. In sensor networks that make use of existing key pre distribution schemes of pairwise key establishment and authentication between sensor nodes and mobile sinks, the use of mobile sinks of data collection elevates a new security challenge. Improved Binary Particle Swarm Optimization algorithm (IBPSO) has been used to find the exact location of a three-way process such as sink, distribution of frequency, and localization. The Orthogonal Frequency Division Multiple Access (OFDMA) technique is used to identify the frequency in the communication channel for finding the exact frequency. The existing multiple access techniques have not been used to combine the three-way process such as sink the node, frequency, and positions for utilizing the efficiency of energy in the particular positions to transfer a packet. The proposed research is used to implement the IBPSO algorithm with OFDMA techniques for utilizing exact bandwidth to perform the energy level at the scheduled time. The experimental results have been implemented in the mathematical approach of Polynomial pool based scheme for finding the regions. In the region, the normal distribution procedure has measured optimally to produce the Quality of Service (QoS) for accessing the better outcome of bandwidth and it provides an easy way to access mechanism with higher energy efficiency.

Prasad Deshpande and Mark Stamp in their paper "Metamorphic Malware Detection Using Function Call Graph Analysis" state that previous work has shown that well-designed metamorphic malware can evade many commonly-used malware detection techniques, including signature scanning. In the paper, they consider a previously developed score which is based on function call graph analysis. They test the score on challenging classes of metamorphic malware and they show that the resulting detection rates yield an improvement over other comparable techniques. These results indicate that the function call graph score is among the stronger malware scores developed to date.

Rajeev Singh and Teek P. Sharma in their paper "Secure WLAN Handoff Scheme with Continuous Authentication" explore that Handoffs are essential for providing continuous mobility to a wireless Station (STA) in an Enterprise LAN. An important requirement of the handoff is to establish connection of the roaming STA with a new Access Point (AP) securely and quickly such that the undergoing communication remains unaffected. They propose a novel handoff scheme for enhancing the handoff performance and security. The scheme is a lightweight and reactive method for transferring the keying material i.e., STA context to new AP. Scheme utilizes Key Hiding Communication (KHC) scheme for ongoing data communication between STA and AP. It provides continuous authentication between STA and APs. Computation and communication cost for the handoff process are calculated and security analysis is done. A comparison with other handoff schemes is also provided.

Steven John Simon and Carol J. Cagle in their paper "An Analysis of Trust, Distrust, and Their Antecedents: Development of a Comprehensive Model of Consumer Intentions in Technology-Driven Transactions" argue that data breaches -- security incidents -- have become an everyday occurrence with hundreds of millions consumers having their lost personal identification information (PII), had their credit and debit card numbers stolen, and their credit compromised. Despite the risk, consumers continuously swipe their cards and share their personal information regularly. The study examines the impacts of trust and distrust on consumer intentions in the environment. More than 1,700 consumers involved in technology-driven transactions comprise the data sample. Trust, distrust, and their antecedents are investigated to determine (1) if trust and distrust are truly two distinct constructs, (2) if the two constructs have unique antecedents, and (3) their impacts on consumer intentions toward transactions. The study expands the literature treating trust and distrust as distinct yet inter-related constructs and by introducing new antecedents. Their findings suggest that trust and distrust are not the same construct and impact consumer intentions to transact.

As the final note, we would like to thank all the authors and reviewers for their collaborative efforts to make this issue possible. It is our sincere wish that this journal become an attractive knowledge exchange platform among information systems researchers. Last but not least, to our loyal readers around the world, we hope you find the contents of the papers useful to your work or research.

Dr. Eldon Y. Li
Editor-in-Chief and University Chair Professor

Department of Management Information Systems
College of Commerce
National Chengchi University
Taipei, Taiwan
Fall 2016

# *MIS Review*

September 2015/March 2016 Vol.21 Nos. 1/2

## Contents

### Research Articles

*MIS Review* Vol. 21, Nos. 1/2, September(2015)/March (2016), pp. 1-13
DOI: 10.6131/MISR.2015.2101.01


# Probabilistic Approach of Improved Binary PSO Algorithm Using Mobile Sink Nodes

*S. Raj Anand[1], E. Kannan[2]*
*[1]Department of Computer Science and Engineering, Vandayar Engineering College, India*
*[2]School of Computing and Information Technology, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, India*

ABSTRACT: *In Wireless Sensor Network (WSN) applications for efficient data accumulation, the use of mobile sinks plays a very important part. In sensor networks that make use of existing key pre distribution schemes of pairwise key establishment and authentication between sensor nodes and mobile sink, the use of mobile sinks of data collection elevates a new security challenge. Improved Binary Particle Swarm Optimization algorithm (IBPSO) has been used to find the exact location of a three-way process such as sink, distribution of frequency, and localization. The Orthogonal Frequency Division Multiple Access (OFDMA) technique is used to identify the frequency in the communication channel for finding the exact frequency. The existing multiple access techniques have not been used to combine the three-way process such as sink the node, frequency, and positions for utilizing the efficiency of energy in the particular positions to transfer a packet. The proposed research is used to implement the IBPSO algorithm with OFDMA techniques for utilizing exact bandwidth to perform the energy level at the scheduled time. The experimental results have been implemented in the mathematical approach of Polynomial pool based scheme for finding the regions. In this region, the normal distribution procedure has measured optimally to produce the Quality of Service (QoS) for accessing the better outcome of bandwidth and it provides an easy way to access mechanism with higher energy efficiency.*

KEYWORDS: *Mobile Sink, Improved Binary Particle Swarm Optimization, OFDMA, Wireless Sensor Network, Quality of Service*

## 1. Introduction

The Mobile Sink is used in many applications for producing accurate and timely phenomena detection is required. OFDMA based MAC protocol for underwater acoustic Wireless Sensor Networks described that is configurable to suit the operating requirements of the underwater sensor network. The protocol has three modes of operation, namely random, equal opportunity and energy-conscious modes of operation. The MAC design approach exploits the multi-path characteristics of a fading acoustic channel to convert

it into parallel independent acoustic sub-channels that undergo flat fading (Khalil et al., 2012). Delay tolerant data gathering in energy harvesting sensor networks with a Mobile Sink has specified that the energy conversion efficiently investigated. The impact of different parameters on the performance this is the first kind of work of data collection for energy harvesting sensor networks with mobile sinks (Ren & Liang, 2012). Particle swarm optimization for time difference-of-arrival based localization shown the PSO approach provided accurate source location estimation for both known and unknown propagation speed, and also gives an efficient speed estimate in the latter case (Lui et al., 2007). Reliable and load balanced multi-path routing for multiple sinks in wireless sensor networks specified that an efficient load balanced multipath routing for multiple sinks is obtained and the fault detection and recovery can be made effectively (Mahadevaswamy & Shanmukhaswamya, 2012). Architecture of wireless sensor networks with mobile sinks exploited the tradeoff between the successful information retrieval probability and the nodes energy consumption, a number of multiple node transmission scheduling algorithms are proposed (Song & Hatzinakos, 2007).

Energy-aware data aggregation for grid based Wireless Sensor Networks with a Mobile Sink proposed that each sensor node with location information and limited energy is considered. This approach utilized the location information and selects a special gateway in each area of a grid responsible for forwarding messages (Wang et al., 2007). Band based go casting for mobile sink groups in the wireless sensor Networks delivery-guaranteed and effective data dissemination for mobile sink groups in wireless sensor networks. A mobile sink group denotes a set of tightly coupled mobile sinks for team collaborations such as a team of firefighters and a group of soldiers (Park et al., 2013). Energy balanced routing algorithm based on Mobile Sink for Wireless Sensor Networks. The algorithm defined the transmitting coordinate (TC) of by mobile sink and the sensor nodes, which TC is the same formed a chain cluster using a greedy approach. It also defined collecting row (CR) and paralleling column (PC), and each PC transmits information to CR synchronously. Finally, information is transmitted to mobile sink by LEADER node which residual energy is the most (Guan et al., 2012). SRP-MS: a new routing protocol for delay tolerant wireless sensor network lifetime maximization of delay tolerant wireless sensor networks (WSNs) through the manipulation of Mobile Sink (MS) on different trajectories (Javaid et al., 2013; Paulus et al., 2012). Reducing delay data dissemination using mobile sink in Wireless Sensor Networks for reducing drastically for data packet collection from the networks and save energy consumption, congestion and average end to end delay problem for the collection of data packets in the network (Waghole & Deshpande, 2013). Zhao et al. (2014) described the positioning accuracy, they put forward an improved weighted centroid algorithm, then self-corrected. The distributed multi-cell beamforming algorithm converges to an NE point in just a few iterations with

low information exchange overhead. Moreover, it provides significant performance gains, especially under the strong interference scenario, in comparison with several existing multi-cell interference mitigation schemes, such as the distributed interference alignment method (Xu & Wang, 2012). The information related to the residual battery energy of sensor nodes to adaptively adjust the transmission range of sensor nodes and the relocating scheme for the sink (Wang et al., 2013). The key concept in virtual backbone scheduling is to minimize the energy consumption and more throughputs concentrated by Umesh et al. (2013). Huang et al. (2014) presented optimization of routing protocol in wireless sensor networks based on improved ant colony and particle swarm algorithm. The shown the results verify the effectiveness of the improved algorithm, and improve the search for optimal routing.

The objective of our research work is to find the appropriate position using IBPSO algorithm in all locations to acquire accurate bandwidth which minimizing the error through mobile sink. In order to communicate in scheduled time, the exact bandwidth can be used by OFDMA technique.

## 2. Architectural models

Figure 1 shows the architectural model of Mobile sink with IBPSO for finding the appropriate position in WSN. It shows the how sensor networks are distributed the data to various networks. The mobile sink has used to sink all the sensor nodes. Each sensor is used to find the exact position to transmit the data with IBPSO. The swarming of all the sensors has been assigning the data in the particular location. The IBPSO algorithm is used with time scheduled for distributing with higher bandwidth with respect to OFDMA/ TDMA for defining the appropriate periods.

## 3. Mobile sink with IBPSO

A mobile sink sends data request message to sensor nodes via a stationary access node. These mobile sink request messages will initiate the stationary access node to trigger sensor nodes for transmitting their data to the requested mobile sink. The main advantage of mobile sink is used to secure way of communication has to be established in the sensor networks. It is enabled to collect all information from a remote machine and the information has been timely improved with energy. The sink nodes find their geographic locations by the configuration of localization techniques. The OFDMA techniques also implemented in mobile sink for controlling the information without having any error. In WSN, the acknowledgement cannot be predicted to all transmissions. So, OFDMA

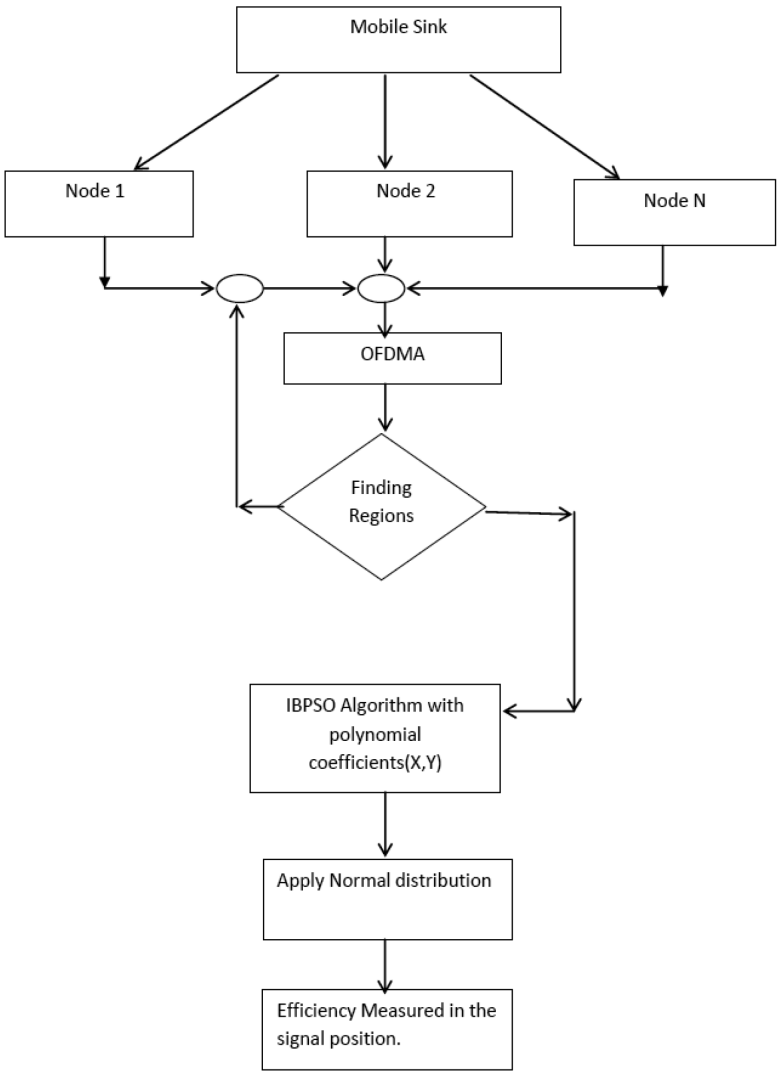



**Figure 1** Interconnection of Mobile Sink with IBPSO for Finding the Position

is scheduled the time for every transmission in mobile sink and finding the appropriate position through IBPSO.

Figure 2 depicts the mobile sinks which identifies all the sink nodes and neglect the unauthorized nodes. Then the data has been transferred to the particular location using an IBPSO algorithm for finding the exact location and also improving the performance.

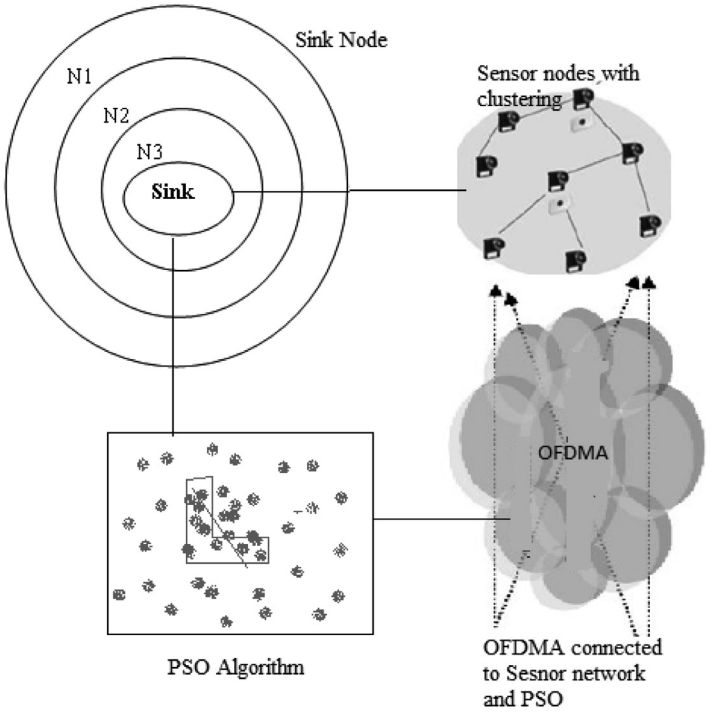



**Figure 2** Clustering of Sensor Node Sink with Position of Binary PSO in OFDMA

## 4. Improved Binary PSO algorithm

IPSO is a heuristic global optimization procedure, which is based on swarm Intelligence. IPSO is an algorithm for optimizing a non-linear and multidimensional problem which usually reaches good solutions efficiently, while requiring minimal parameterization. The overall function of this problem is optimized and also it finds the exact solution for an appropriate location. The function can be defined as $f(x_i)$ and the nodes are $f(x_{i,0})$, ..., $f(x_{i,d})$ and also it represents how the particle's position in the multidimensional space is relative to the desired goal. With this problem, the IPSO algorithm shows that how the frequencies are calculated at the rate of weights and also binds the "d" dimensions to be optimized for given a problem modeled as an optimization one of dimensions d. So, it has positioned as $x_i$ ($x_1$, $x_2$, ..., $x_n$) and velocity $v_i$ ($v_1$, $v_2$, ..., $v_n$).

The Improved Binary PSO (IBPSO) algorithm is called by IPSO, the particle position is not a real value, but either the binary 0 or 1. So that velocity $v_{i,d}$ mapped into interval [0,1]. In this case velocity has been updated in Equation (1) and positions have been updated in Equation (3). In this algorithm, swarm means all the particles share information and it says that best position in every visited and also by any particle in the swarm.Each particle has a position can be defined in the Equations (1) and (2).

$$x_{i,d}(it+1) = x_{i,d}(it) + v_{i,d}(it+1) \quad (1)$$

and

$$v_{i,d}(it+1) = v_{i,d}(it) + C_1 * Rnd(0,1) * \left[ p^{b_{i,d}}(it) - x_{i,d}(it) \right] + C_2 * Rnd(0,1) * \left[ g^{b_d}(it) - x_{i,d}(it) \right] \quad (2)$$

$$x_{i,d}(it+1) = \begin{cases} 1 & if\ \sigma < \frac{1}{1+e^{-v_{i,d}}} \\ 0 & otherwise \end{cases} \quad (3)$$

where

i, particle's index, used as a particle identifier

d dimension being considered, each particle has a position and a velocity for each dimension

it iteration number, the algorithm is iterative

$x_{i,d}$ position of particle i in dimension d

$v_{i,d}$ velocity of particle i in dimension d

$C_1$ acceleration constant for the cognitive component;

Rnd stochastic component of the algorithm, a random value between 0 and 1

$P^b_{i,d}$ the location in dimension d with the best fitness of all the visited locations in that dimension of particle i and gb is the global position

$C_2$ acceleration constant for the social component

σ Random factor in the [0,1] interval

## 5. Integrated polynomial pool based scheme

This scheme is used to define the combination of polynomial pool based scheme as well as the three tier security scheme. This scheme is defined position in the particular regions using IBPSO. Based on the region the packets are maintained efficiently with a short spot of time. The sink node is used to identify the unintended nodes, which are sent from the source place and also protect the unauthorized data along with two keys named as Sending key (S) and Receiving key (R). Consider the four regions in the particular

positions. The regions are located in the position with small particle $R_1$ = (-4, 0), $R_2$ = (-1, 0), $R_3$ = (1, 0) and $R_4$ = (3, 0). Among the four regions, the exact location can be identified by the following procedure and the polynomials of coordinating points are described in Figure 3 with sine wave.

$$P(x) > 0 \text{ for } x \in (-\infty, -4) \cup (3, \infty) \quad (4)$$

$$P(x) < 0 \text{ for } x \in (-4, -1) \cup (1, 3) \quad (5)$$

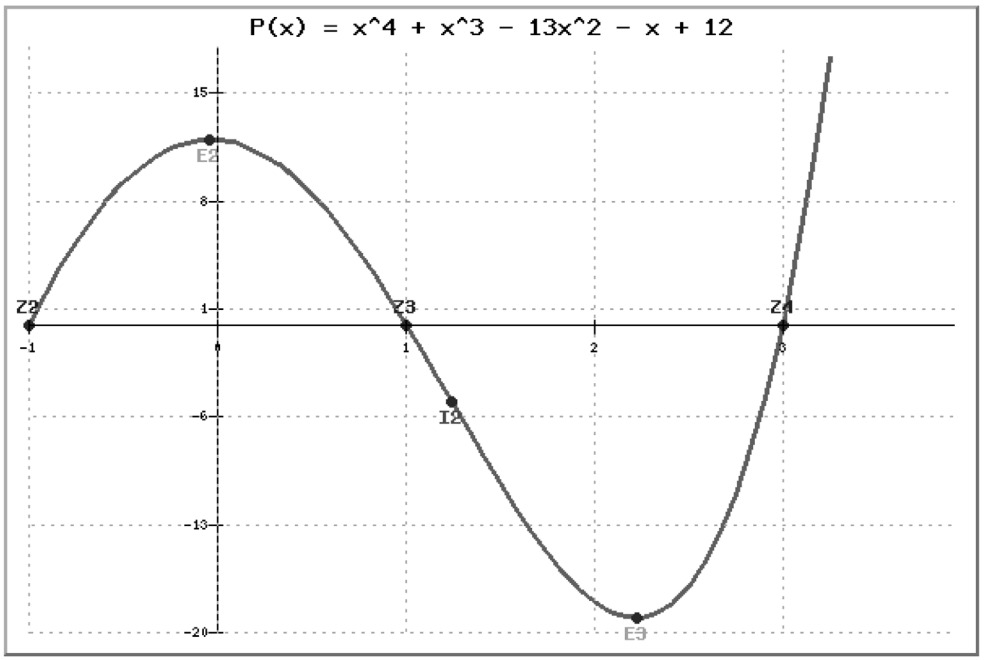



**Figure 3** Polynomial of Coordinating Points in the Particular Regions

IBPSO algorithm:

Step 1: Searching the nodes $N_1, N_2, ..., N_i$ for sending the data.

Step 2: Checking by mobile sink nodes, whether $N_1, N_2, ..., N_i$ are malicious or original data.

Step 3: Generating the keys $S_1, S_2, ..., S_i$ for the given data and encrypting the keys along with the data.

Step 4: For each particle i in s do

Step 5: For each dimension d in D d

$$x_{i,d} = Rnd(x_{\min}, x_{\max})$$

$u_{i,d} = Rnd(-v_{max}/3, v_{max}/3)$
End

Step 6: Evaluate the polynomial for the particular particle using the Equations (4) and (5). For each region in Swarm
(i) if (the position value of $X_i > P_i$) then $P_i = X_i$
(ii) if (the position value of $X_i > P_g$) then $P_g = X_i$
for each dimension of particle
(i) Update each dimension velocity using the Equation (2)
(ii) Update velocity if ($V_{id}(t + 1) > V_{max}$) then $V_{id}(t + 1) = V_{max}$
(iii) From Equation (3) $S = \frac{1}{1+e^{-v_{i,d}}} + 1$
(iv) S > rand(0,1) then Xid(t + 1) = 1
else $X_{id}(t + 1) = 0$

Step 7: Find the frequency with each dimension of the following

Frequency = $\frac{v_{i,d}}{S}$ and global position = Frequency

Step 8: $T_i$(Time) = global position + (Hours * 60) + Minutes
if($T_i$ == maxHours),
Message "Error: Request time is more"
Else Print the total time with frequency

Step 9: Print the position with appropriate frequency level.

Step 10: return $T_{i\_Time}$

## 6. Simulation

Figure 4 shows the parameter position finding without -1 value in the localization. In the particular location, nodes are having the appropriate position to find the signal strength. If the values are -1, returns the negative position and also not finding the localization.

Figure 5 shows how localization can be found in every location. When the position is identified, every node has connected to its neighboring node. The graph shows the square mark in every link to find its positions. The data are traversed through this node and finally reach to its destination. The efficiency of the sensor nodes are defined with normal distribution for providing the optimal solution in the particular location.

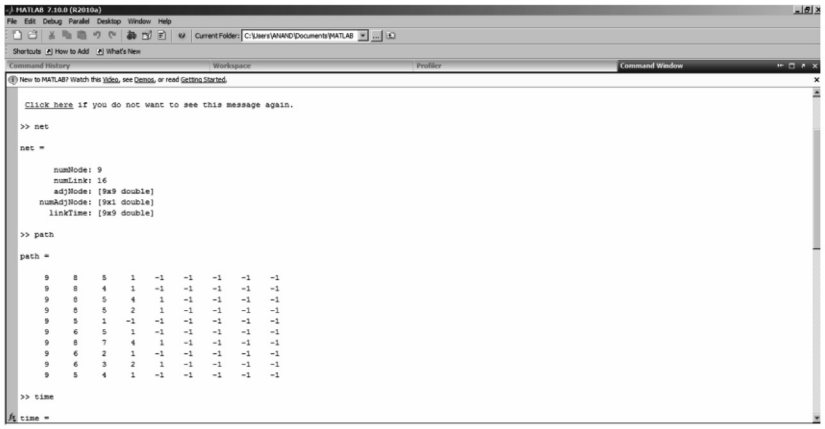



**Figure 4** Implementation of Node Position Finding in the Localization in Matlab

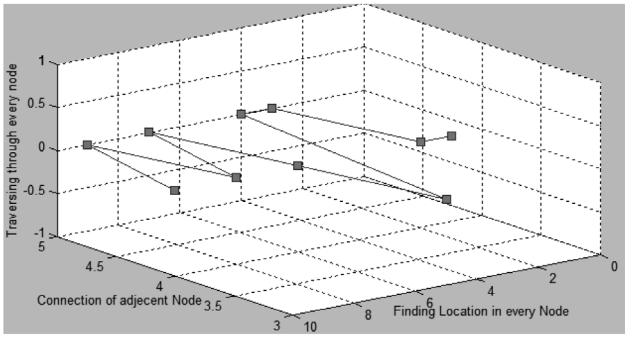



**Figure 5** Traversing through Every Node for Finding in the Localization

The solution has been optimally selected in the upper area of the normal curve with intersect point. Consider the 25 nodes have been participated in sending the data through the signal position of the 4 regions. In such a case all the nodes are intersected with the optimal solution with the efficiency of the time 0.5 seconds in the normal curve. The following steps are taken into the consideration with respect to finding optimal solutions.

Step 1: Since $\mu = 25$ and $\sigma = 4$ we have:

$$P(0 < X < 25) = P(0 - 25 < X - \mu < 25 - 25) = P\left(\frac{0 - 25}{4} < X - \mu\sigma < \frac{25 - 25}{4}\right)$$

Since $Z = \dfrac{x - \mu}{\sigma}$, $\dfrac{0 - 25}{4} = -6.25$ and $\dfrac{25 - 25}{4} = 0$ we have:

$$P(0 < X < 25) = P(-6.25 < Z < 0)$$

Step 2: Use the standard normal table to conclude that:

$$P(-6.25 < Z < 0) = 0.5$$

Such that $P(0 < X < 25) = 0.5$

Figure 6 shows the appropriate position finding with the various coordination. The graph shown the various coordination for finding the position and the average delay times are calculated in every position. This is normally represented for data distribution in the particular location, for how much of time can be utilized for transmitting data in the particular location. Figure 7 shows that the reliability of IBPSO algorithm. The reliability has measured in percentages and the IPSO algorithm has found the position with same reliability measurement. In this scenario, the time has been measured for finding the position which is incorporated in the particular regions.

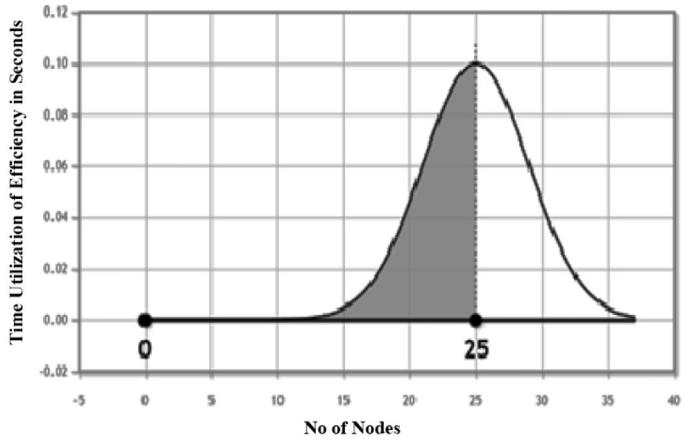



**Figure 6** The Node Position Finding in the Localization

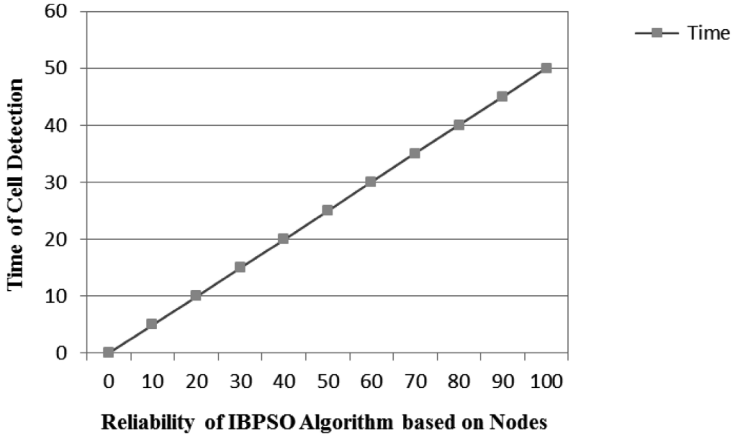



**Figure 7** imes for Reliability of IBPSO in Cell Detection

In every shot of signals, the average value has been calculated based on the number of nodes. If 100 nodes are transmitting the data, the overall reliability of data transfer in time is 0.5 seconds with percentage.

## 7. Conclusions

The main proposed system of the study described that three way distribution processes of improved particle swarm optimization frequencies for mobile sink in localization of wireless sensor network. In this study the IBPSO algorithm used for finding the location and to identify the localized environment. The mobile sink has been allocated the sink node for distribution of the data between one sink nodes to another sink node. During the distribution the OFDMA used to define the frequency with constant level. In the frequency channel all the data have distributed securely with the polynomial pool based method applied for protecting from unauthorized data. Finally, all the techniques are accumulated together in a channel without having any errors and also the efficiency have been utilized in the particular location to send the packets.

## References

Guan, J., Sun, D., Wang, A. and Liu, Y. (2012), 'Energy balanced routing algorithm based on mobile sink for wireless sensor networks', *Journal of Computational Information Systems*,Vol. 8, No. 2, pp. 603-613.

Huang, T., Li, X., Zhang, Z. and Lian, H. (2014), 'Optimization of routing protocol in wireless sensor networks by improved ant colony and particle swarm algorithm', *TELKOMNIKA Indonesian Journal of Electrical Engineering*, Vol. 12, No. 10, pp. 7486-7494.

Javaid, N., Khan, A.A., Akbar, M., Khan, Z.A. and Qasim, U. (2013), 'SRP-MS: a new routing protocol for delay tolerant wireless sensor networks', *Proceedings of 26th IEEE Canadian Conference on Electrical and Computer Engineering*, Regina, Canada, pp. 1-4.

Khalil, I.M., Gadallah, Y., Hayajneh M. and Kreishah, A. (2012), 'An adaptive OFDMA-based MAC protocol for underwater acoustic wireless sensor networks', *Sensors*, Vol. 12, pp. 8782-8805.

Lui, K.W.K., Zheng, J. and So, H.C. (2007), 'Particle swarm optimizatin for time-difference of arrival based localization', *Proceedings of 2007 15th European Signal Processing Conference*, Poznan, Poland, pp. 414-417.

Mahadevaswamy, U.B. and Shanmukhaswamya, M.N. (2012), 'Reliable and load balanced multi-path routing for multiple sinks in wireless sensor networks', *International Journal of Computer Applications*, Vol. 50, No. 12, pp. 14-21.

Park, S., Oh, S., Kim, J., Lee, E. and Kim, S.H. (2013), 'Band-based geocasting for mobile sink groups in wireless sensor networks', *Wireless Network*, Vol. 19, No. 6, pp. 1285-1298.

Paulus, R., Singh, G. and Tripathi, R. (2012), 'Energy efficient data transmission through relay nodes in wireless sensor networks', *ACEEE International Journal on Network Security*, Vol. 3, No. 1, pp. 40-45.

Ren, X. and Liang, W. (2012), 'Delay-tolerant data gathering energy harvestng sensor networks with a mobile sink', *Proceedings of IEEE Global Communications Conference*, Anaheim, CA, pp. 93-99.

Song, L. and Hatzinakos, D. (2007), 'Architecture of wireless sensor networks with mobile sinks: multiple access case', *International Journal of Distributed Sensor Networks*, Vol. 3, No. 3, pp. 289-310.

Umesh, B.N., Vasanth, G. and Siddaraju. (2013), 'Energy efficient routing of wireless sensor networks using virtual backbone and life time maximization of nodes", *International Journal of Wireless & Mobile Networks*, Vol. 5, No. 1, pp. 107-118.

Waghole, D.S. and Deshpande, V.S. (2013), 'Reducing delay data dissemination using mobile sink in wireless sensor networks', *International Journal of Soft Computing and Engineering*, Vol. 3, No. 1, pp. 305-308.

Wang, C.F., Shih, J.D., Pan, B.H. and Wu, T.Y. (2013), 'A network lifetime enhancement method for sink relocation and its analysis in wireless sensor networks', *IEEE Sensors Journal*, Vol. 14, No. 6, pp. 1932-1943.

Wang, N.C., Huang, Y.F., Chen, J.S. and Yeh, P.C. (2007), 'Energy-aware data aggregation for grid-based wireless sensor networks with a mobile sink', *Wireless Personal Communications*, Vol. 43, No. 4, pp. 1539-1551.

Xu, W. and Wang, X. (2012), 'Pricing-based distributed downlink beamforming in multi-cell OFDMA networks', *IEEE Journal on Selected Areas in Communications*, Vol. 30, No. 9, pp. 1605-1613.

Zhao, J.M., W.X. An, Li, D.A. and Zhao, D.D. (2014), 'Effective algorithms for WSN with weight principle in web of things', *IEEE Sensors Journal*, Vol. 14, No. 1, pp. 228-233.

## About the authors

**S. Raj Anand** received M.C.A. Degree from Bharathidasan University, Trichy. He has done M.E. Computer Science & Engineering from Anna University, Chennai. He is doing research work in Computer Science & Engineering at Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology (Vel Tech Dr. RR & Dr. SR Technical University). His research area in Wireless Sensor Network and Network Security. He has published 12 papers in National/International Journals. He has presented 12 papers in National and International conferences.

Corresponding author. Research Scholar, Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology Chennai, Tamilnadu, India 600062. Tel: +91-9443399113. E-mail address: rajsra_mca@rediffmail.com

**E. Kannan**, received M.Sc, Degree from Bharathidasan University, Trichy and M.E., in Computer Science and Engineering from Sathyabama University, Chennai. He has done his Doctorate in Computer Science from National Institute of Technology, Trichy. He is currently working as Dean, School of computing and Information Technology at Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology (Vel Tech Dr. RR & Dr.SR Technical University). His research interest includes Parallel Computing, Adhoc Network, Network Security and Natural Language Processing. He has published 58 papers in International Journals. He is a Member of Indian Society for Technical Education (ISTE) and Annual Member of IEEE. He is a recognized Supervisor for various Universities for guiding Ph.D. programs. E-mail address: ek081966@gmail.com

*MIS Review* Vol. 21, Nos. 1/2, September(2015)/March (2016), pp. 15-34
DOI: 10.6131/MISR.2015.2101.02


# Metamorphic Malware Detection Using Function Call Graph Analysis

*Prasad Deshpande, Mark Stamp*
*San Jose State University, San Jose, CA*

ABSTRACT: *Previous work has shown that well-designed metamorphicmalware can evade many commonly-used malware detection techniques, including signature scanning. In this paper, we consider a previously developed score which is based on function call graph analysis. We test this score on challenging classes of metamorphic malware and we show that the resulting detection rates yield an improvement over other comparable techniques. These results indicate that the function call graph score is among the stronger malware scores developed to date.*

KEYWORDS: *Malware, Function Call Graph, Metamorphic Software*

## 1. Introduction

Malware is a software that is designed to perform malicious activity (Panda Security, 2011). Examples of such malicious activity range from crashing a system to collecting and infiltrating sensitive data. There are many different categories of malware, including virus, worm, trojan horse, logic bomb, back door, and spyware (Aycock, 2006). In this paper, we use the term virus generically to refer to any type of malware.

According to Symantec (2011), the number of unique malware variants increased from about 286 million to more than 403 million between 2010 and 2011. Also, in 2011, Symantec claimed to have blocked more than 5.5 billion attacks (Symantec, 2011). These numbers give some indication of the scope and prevalence of the malware threat--a massive threat that shows no sign of abating anytime soon.

Code obfuscation is used to obscure the characteristics of code (Xu et al., 2013). Virus writers have developed a variety of code obfuscation techniques, many of which are designed to evade signature detection. Arguably, the most potent such technique is metamorphism, that is, code morphing that changes the internal structure with each infection, while maintaining the essentials of its original function (Shang et al., 2010). Metamorphic generators are readily available, so that even a novice attacker can easily take advantage of this powerful technique. Examples of notable metamorphic generators

include

- NGVCK (Next Generation Virus Creation Kit) (Snakebyte, 2000)
- MPCGEN (Mass Code Generator) (Tips Trik Dan Berbagi Informasi, n.d.)
- G2 (Second Generation Virus Generator) (VX Heavens, n.d.)
- VCL32 (Virus Creation Lab for Win32) (Attaluri et al., 2009)
- MetaPHOR (The Mental Driller, 2002)
- NRLG (NuKE's Random Life Generator) (Symantec, n.d.)
- NEG (NoMercy Excel Generator) (Symantec, n.d.)

Function call graphs have been previously applied to the malware detection problem. For example, Bilar (2007) propose and analyze a mechanism to generate call graphs for malware detection. The paper Shang et al. (2010) proposes an algorithm to determine similarity between function call graphs, while Karnik et al. (2007) uses a cosine similarity metric to measure the overall similarity between code samples, based on call graphs. In Christodorescu et al. (2007), a data mining algorithm is used to construct call graphs via dynamic analysis.

In this paper, we apply a call graph-based score to several challenging classes of metamorphic malware. We compare the results obtained using this call graph approach to previous results obtained using hidden Markov model (HMM) analysis (Wong & Stamp, 2006). These HMM results have previously served as a benchmark for comparing the effectiveness of a wide variety of detection techniques (Attaluri et al., 2009; Kazi & Stamp, 2013; Lin & Stamp, 2011; Runwal et al., 2012; Shanmugam et al., 2013; Sridhara & Stamp, 2012; Tamboli et al., 2014). We show that call graph analysis can yield improved results over many of these previous techniques in these particularly challenging cases.

This paper is the first to test call graph based scoring on such challenging classes of malware. Our results indicate that function call graphs are a powerful technique for scoring malware, and such scores are relatively immune to many common obfuscation techniques.

This paper is organized as follows. Section 2 provides background information on malware and detection techniques, including a discussion of Hidden Markov Models. We also briefly discuss various metamorphic techniques. In Section 3 we discuss call graph analysis and its application to malware detection and, of course, we emphasize the specific implementation that we have chosen. Section 4 contains our experimental results. Finally, Section 5 has our conclusion and suggestions for possible future work.

## 2. Background

In this section, we first discuss metamorphic malware and various code morphing techniques. Then we briefly discuss Hidden Markov Models (HMMs) and their use in malware detection. HMMs will serve as a benchmark for comparing the call graph scores analyzed in this paper.

### *2.1 Metamorphic techniques*

A metamorphic generator can produce a large number of different generations of code, where the functionality remains the same, but the internal structure differs. Such code obfuscation can alter instructions as well as program data and control flow (Borello & Mé, 2008; You & Kim, 2010). These techniques can be used to evade signature detection, as well as to evade statistical analysis. Next, we briefly consider some code morphing techniques.

#### *2.1.1 Register swap*

Register swapping is one of the easiest metamorphic techniques to implement, but it is also one of the least effective. RegSwap, which was arguably the first metamorphic viruses, used this technique exclusively (Szor, 2005). Table 1 shows code fragment from different generation of W95/RegSwap virus.

**Table 1** Two Generations of RegSwap

| pop edx | pop edx |
|---|---|
| mov edi, 0004h | mov ebx, 0004h |
| mov esi, ebp | mov edx, ebp |
| mov eax, 000Ch | mov edi, 000Ch |
| add edx, 0088h | add eax, 0088h |
| mov ebx, [edx] | mov esi, [eax] |
| mov [esi+eax*4+00001118], ebx | mov [edx+edi*4+00001118], esi |

Source: Szor, 2005

#### *2.1.2 Transposition*

Subroutine permutation is another elementary code morphing technique. If there are $n$ subroutines, then it is trivial to generate $n!$ different metamorphic copies by simply permuting the order of the subroutines. BadBoy and W32/Ghost are two viruses that employ subroutine permutation (Szor, 2005). BadBoy has 8 subroutines, so it can generate 8! = 40320 different variants.

More generally, if two instructions (or groups of instructions) are independent of each other then their order can be changed. Even more general transposition can be used, provided jump instructions are inserted to preserve the order of code execution.

### *2.1.3 Dead code insertion*

Dead code insertion can be a highly effective morphing strategy. Dead code may or may not be executed; if such code is executed, care must be taken so that it has no effect on the functioning of the program. Examples of dead code insertions are given Table 2 Note that none of the instructions in Table 2 change the value of the register.

**Table 2** Example of Dead Code

| Instruction | Description |
|---|---|
| add Reg,0 | Add value 0 to register |
| mov Reg,Reg | Transfer register value to itself |
| or Reg, 0 | Logical OR operation of register with 0 |
| nop | No operation |

Source: Szor and Ferrie (2001)

Dead code insertion is useful for evading signature detection and can also aid in evading statistical-based detection. Dead code insertion is used, for example, in Win95/Zperm (Szor, 2005) and also in the experimental metamorphic worm MWOR, which is analyzed in Sridhara and Stamp (2012).

### *2.1.4 Instruction substitution*

An instruction or group of instructions can be substituted for another equivalent instruction or group of instructions. For example, the instruction xor eax, eax can be replaced by sub eax, eax. Instruction substitution can be highly effective, but is relatively difficult to implement. Instruction substitution is used extensively in W32/MetaPHOR (Szor, 2005) and also to some extent in the MWOR worm Sridhara and Stamp (2012).

### *2.1.5 Formal grammar mutation*

A code morphing engine can be viewed as nondeterministic automata, where transitions are possible from every symbol to every other symbol (Zbitskiy, 2009). Here, the set of symbols consists of the set of possible instructions. By formalizing mutation techniques in this way, we can apply formal grammar rules and create malicious copies with large variation; see Zbitskiy (2009) for an example.

### *2.1.6 Host code mutation*

Some viruses mutate the code of the host along with their own code (Konstantinou & Wolthusen, 2008). Win95/Bistro is an example of malware that uses this concept of host code mutation Szor (2000).

### *2.1.7 Code integration*

Win95/Zmist implements a "code integration" technique. Specifically, Zmist decompiles a portable executable (PE) file, inserts itself into the code of the file, regenerates the code and data references, and recompiles the executable (Szor & Ferrie, 2001).

### *2.2 Hidden Markov model based detection*

Hidden Markov Model (HMM) analysis has proven useful in a wide array of fields, ranging from speech recognition (Rabiner, 1989) to software piracy detection (Kazi & Stamp, 2013). Previous research has shown that HMMs can be a highly effective tool for detecting metamorphic malware (Attaluri et al., 2009; Lin & Stamp, 2011; Wong & Stamp, 2006). Since HMMs have been widely studied, we use an HMM-based score as the benchmark for comparison with the call graph score considered in this paper.

An HMM includes a "hidden" Markov process, and a sequence of observations that are probabilistically related to this hidden process. We can train an HMM for a given sequence of observations. Then we can score a sequence against this trained model to determine how closely it matches the training data. The relevant notation commonly used in HMMs appears in Table 3.

A generic HMM is illustrated in Figure 1, where $X_t$ and $\mathcal{O}_t$ represent the (hidden) state sequence and the observation sequence, respectively. The underlying Markov process is driven by the $A$ matrix. The observations $\mathcal{O}_t$ are related to the current state of the Markov

**Table 3** HMM Notation

| Symbol | Description |
|---|---|
| $T$ | length of the observed sequence |
| $N$ | number of (hidden) states in the model |
| $M$ | number of distinct observation symbols |
| $\mathcal{O}$ | observation sequence ($\mathcal{O}_0, \mathcal{O}_1, \ldots, \mathcal{O}_{\mathrm{T}}-1$) |
| $A$ | $N \times N$ state transition probability matrix |
| $B$ | $N \times M$ observation probability matrix |
| $\pi$ | $1 \times N$ initial state distribution matrix |

Source: Stamp (2015).

process by probability distributions contained in the $B$ matrix. The matrices $A$, $B$, and $\pi$ are row stochastic, that is, the elements of each row satisfy the conditions of a probability distribution.

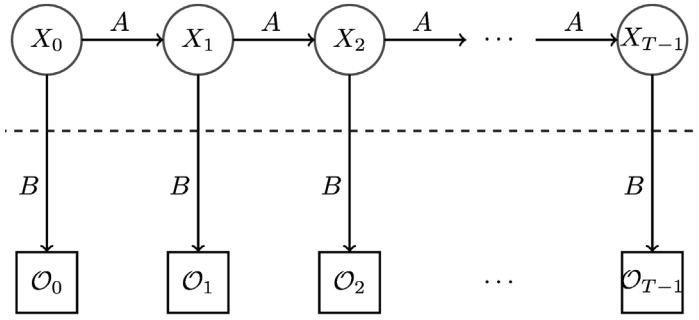



**Figure 1** Generic Hidden Markov Model (Stamp, 2015)

For the metamorphic malware detection problem considered in Wong and Stamp (2006), opcodes are extracted from several members of a given metamorphic family. These opcode sequences are concatenated to form a sequence $\mathcal{O}$, and an HMM is trained on $\mathcal{O}$. To score a given file, its opcode sequence is extracted and scored against the trained HMM. The results in Wong and Stamp (2006) indicate that this technique is highly effective at detecting hacker-produced metamorphic code.

These results have been confirmed and further analyzed in a substantial body of subsequent research, including Attaluri et al. (2009), Kazi and Stamp (2013), Lin and Stamp (2011), Runwal et al. (2012), Shanmugam et al. (2013), Sridhara and Stamp (2012) and Tamboli et al. (2014). Consequently, we use HMM scoring as a benchmark to measure the effectiveness of the call graph technique considered here.

## 3. Call Graph Analysis

In this section, we first discuss previous malware detection work based on using call graph analysis. Then we discuss function call graphs in general, and explain in detail the scoring algorithm used in this research.

### *3.1 Previous work*

Malware writers have developed a variety of techniques for evading signature detection (Aycock, 2006). In contrast to signature detection, functional call graph analysis relies on higher-level structure, that should be more difficult to obfuscate. The purpose

of this research is to determine the effectiveness of call graph-based techniques when confronted with advanced metamorphic malware. Such malware easily defeats signature scanning and, if properly constructed, can also evade statistical-based detection (Sridhara & Stamp, 2012).

A function call graph is created from the disassembled code of an executable as follows. Each function is represented by a vertex, with directed edges represent the caller-callee relationships between functions (Xu et al., 2013). In addition, edge "weights" can be considered, which can be based on opcode analysis and graph coloring techniques. Once such graphs have been constructed, determining the similarity between programs reduces to determining the similarity between their function call graphs.

We implemented the technique given in (Xu et al., 2013). We apply this technique to several advanced metamorphic generators. Our test results are given in Section 4. But first we discuss the process used to construct function call graphs and to measure the similarity of such graphs.

### *3.2 Function Call Graph Construction*

An graph can be represented as $G = (V, E)$, where $V$ is the set of vertices and $E$ is the set of edges. For a function call graph, the vertices represent functions while the edges represent caller-callee relationships between functions (Xu et al., 2013). The functions in $V$ are classified as one of two types, namely, local functions or external functions. Local functions are contained within the executable, while external functions are system or library functions. After disassembling an exe, functions begin with sub_xxxxxx and end with sub_xxxxxx where "xxxxxx" represents the name of the function. For local functions, the name of the called function is also found within the executable, while external functions names are not.

Given an executable, we first disassemble it using IDA Pro (Hex-Ray, n.d.). From the resulting assembly code, we search for functions and extract relevant information for each. Once the relevant information has been extracted for all functions, the function call graph is constructed. Figure 2 shows part of the function call graph for the virus Win32. Bolzano. As can be seen in Figure 2, the graph consists of local functions (those with names of the form sub_xxxxxx) and external functions, such as GetVersion. Note that local functions can call external functions, but an external function cannot call a local function.

As in Shang et al. (2010), we use a breadth first search (BFS) to determine caller-callee relationships between functions. In a BFS, we start from a root node and process successive levels. For our experiments, the entry point function serves as the root node and the the algorithm used is a straightforward BFS; for additional details, see Shang et al. (2010) or Deshpande (2013).

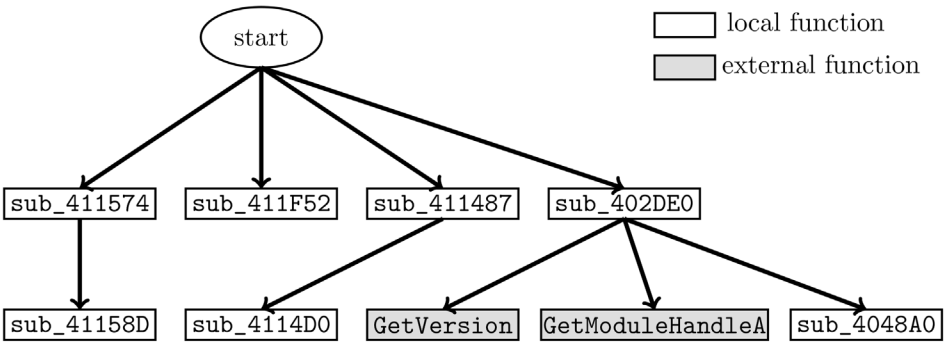



**Figure 2** Function Call Graph in Win32.Bolzano (Xu et al., 2013).

### *3.3 Function call graph similarity*

Once function call graphs have been constructed, we then determine the simi-larity of the corresponding programs by measuring the similarity of their graphs. External functions are matched using their symbolic names, since these will be the same across different programs. However, the symbolic names used for local functions need not be the same across metamorphic code variants. Consequently, we must analyze local functions to determine their similarity across different programs. We use three different techniques to match local functions.

#### *3.3.1 Matching external functions*

External functions have the same name across all executables and make no further calls within the call graph (Carrera & Erdelyi, 2004). Hence, in terms of the call graph, external functions have in-degree 1 and out-degree 0, and these functions can be matched based simply on their symbolic names. For example, the GetVersion function in one function call graph can be matched with same function in any other call graph.

Given function call graphs $G_1$ and $G_2$, we extract the external functions from each. As noted above, these functions are easily determined. Then both sets of external functions are compared, and for any common symbolic names, the corresponding vertex is saved to a common external vertex set, which will be used for scoring. Details of the scoring process are given in Section 3.3, below.

#### *3.3.2 Local function similarity based on external functions*

The first method that we use to find matching local functions consists of match-ing common external function calls. All local functions in the graphs $G_1$ and $G_2$ are compared and we simply tabulate matches in the external functions called. If the number of such matches is two or greater, the corresponding local functions are considered to match by this criteria, so they are saved to a common local set.

### *3.3.3 Local function similarity based on opcode sequences*

Local functions that do not match based on external functions are compared based on their opcode sequences. There are many different opcode-based similarity techniques (Attaluri et al., 2009; Runwal et al., 2012; Shanmugam et al., 2013; Wong & Stamp., 2006). So that our results will be consistent with previous work on call graph similarity, here we use the opcode similarity technique in Xu et al. (2013), which we now describe in detail.

Each vertex in the call graph is "colored" depending on the instructions used. Functions are considered to match provided their "colors" match. In case of matches, the score is computed using cosine similarity.

To make this score more robust against morphing techniques such as code substitution, we classify all X86 instructions into one of 15 categories, according to their function (Xu et al., 2013). These categories are listed in Table 4.

A 15-bit color variable is associated with each vertex corresponding to a local function in the graph. If an opcode of type *Ci* appears in the function, bit *i* of the color variable is set -- if no such opcode exists in the function, color bit *i* is 0. There is also a corresponding vector that holds the count of the number of instructions in each class. For example, the first column in Table 5 contains a function from Win32.DarkMoon. The

**Table 4** x86 Instruction Classification

| Class | Type | Description |
|---|---|---|
| $C_1$ | Data | data transfer such as mov |
| $C_2$ | Stack | stack operation |
| $C_3$ | Port | in and out |
| $C_4$ | Lea | destination address transmit |
| $C_5$ | Flag | flag transmit |
| $C_6$ | Arithmetic | shift, rotate, etc. |
| $C_7$ | Logic | bitbyte operation |
| $C_8$ | String | string operation |
| $C_9$ | Jump | unconditional transfer |
| $C_{10}$ | Branch | conditional transfer |
| $C_{11}$ | Loop | loop control |
| $C_{12}$ | Halt | stop instruction |
| $C_{13}$ | Bit | bit test and bit scan |
| $C_{14}$ | Processor | processor control |
| $C_{15}$ | Float | floating point operation |

Source: Xu et al. (2013).

second column in Table 5 lists the opcode, while the third column is the category of the opcode, as found in Table 4.

The color variable and vector of counts from Table 5 appear in Table 6. These are used for computing a score based on cosine similarity, which we now discuss.

**Table 5** Local Function from Win32.DarkMoon

| Assembly code | Opcode | Category |
|---|---|---|
| sub 4059DC proc near | — | — |
| push ebx | push | $C_2$ |
| mov ebx,eax | mov | $C_1$ |
| cmp ds:byte 4146C1, 0 | cmp | $C_6$ |
| jz short loc 405A04 | jz | $C_{10}$ |
| push 0 | push | $C_2$ |
| call SwapMouseButton | call | $C_9$ |
| mov ds:byte 4146C1, 0 | mov | $C_1$ |
| mov eax,ebx | mov | $C_1$ |
| mov edx,offset dword 405A28 | mov | $C_1$ |
| call sub 403DEC | call | $C_9$ |
| pop ebx | pop | $C_2$ |
| retn | retn | $C_9$ |
| push 0FFFFFFFFh | push | $C_2$ |
| call SwapMouseButton | call | $C_9$ |
| mov ds:byte 4146C1, 1 | mov | $C_1$ |
| mov eax,ebx | mov | $C_1$ |
| mov edx,offset dword 405A28 | mov | $C_1$ |
| call sub 403DEC | call | $C_9$ |
| pop ebx | pop | $C_2$ |
| retn | retn | $C_9$ |
| sub 4059DC endp | — | — |

Source: Xu et al. (2013).

**Table 6** Color Vector of Win32.DarkMoon

| | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | $C_6$ | $C_7$ | $C_8$ | $C_9$ | $C_{10}$ | $C_{11}$ | $C_{12}$ | $C_{13}$ | $C_{14}$ | $C_{15}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| color | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| count | 7 | 5 | 0 | 0 | 0 | 1 | 0 | 0 | 6 | 1 | 0 | 0 | 0 | 0 | 0 |

Source: Xu et al. (2013).

Let $X = (x_1, x_2, \ldots, x_{15})$ and $Y = (y_1, y_2, \ldots, y_{15})$ be the vectors of counts from two functions. Then the cosine similarity between these two vectors is computed as

$$\text{sim}(X, Y) = \frac{X \cdot Y}{||X||\,||Y||} = \frac{x_1 y_1 + x_2 y_2 + \cdots + x_{15} y_{15}}{\sqrt{x_1^2 + x_2^2 + \cdots + x_{15}^2}\,\sqrt{y_1^2 + y_2^2 + \cdots + y_{15}^2}} \tag{1}$$

If the color vectors match exactly, and the cosine similarity between the corresponding count vectors is greater than a predetermined threshold, then the functions are considered a match. In our experiments, we use the same parameters for scoring as in Xu et al. (2013).

### *3.3.4 Local Function Similarity Based on Matched Neighbors*

If two functions match, then it is more likely that functions corresponding to neighboring vertices in the function call graphs should match. For example, suppose that in Figure 3, vertex *A* has been matched with vertex *B*. Then there is a higher likelihood that one or more of vertices *U*, *V* will match with one or more of the vertices *X*, *Y*, *Z*.

Because of this higher likelihood of a match, we alter the score computation for such neighboring vertices. Successors and predecessors of previously matched functions are scored using a slightly relaxed version of the color-based score discussed in the previous section. The difference here is that there is no requirement that the color vectors match, that is, we compute the cosine similarity score in (1), regardless of the color vectors. In addition, we use a slightly different threshold for the similarity score.

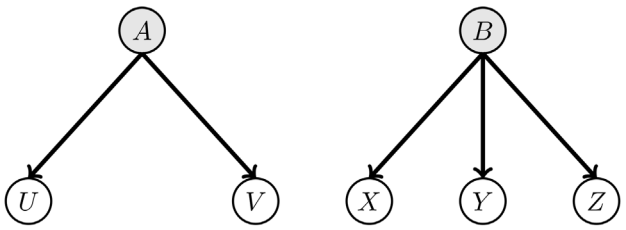



**Figure 3** Successors Functions of Matched *A* and *B*

### *3.3.5 Similarity Score*

Given two function call graphs $G_1$ and $G_2$, we determine all common vertices using the function matching algorithms outlined in Sections 3.3 through 3.3, above. Once we have found all common vertices, we determine the common edges. Suppose vertices *A* and *B* from $G_1$ have been matched to vertices *C* and *D* from $G_2$, respec-tively. If there is

an edge between $A$ and $B$ in $G_1$ and an edge between $C$ and $D$ in $G_2$, then $G_1$ and $G_2$ are said to have a common edge. Let common edge ($G_1$, $G_2$) be the set of such common edges. Then the similarity between two function call graphs is calculated as Xu et al. (2013).

$$\text{sim}(G_1, G_2) = 100 \cdot \frac{2\,|\text{common\_edge}(G_1, G_2)|}{|E(G_1)| + |E(G_2)|} \tag{2}$$

where $E(Gi)$ is the edge set of the graph $Gi$.

## 4. Experiments

In this section, we analyze the performance of the similarity scoring algorithm discussed in Section 3. We test the technique on two families of metamorphic malware, namely, the Next Generation Virus Generation Kit (NGVCK) Snake-byte (2000) and the experimental MWOR worms developed and analyzed in Sridhara and Stamp (2012). The NGVCK viruses have previously been shown to be highly metamorphic, but detectable using statistical-based techniques (Runwal et al., 2012; Shanmugam et al., 2013; Toderici & Stamp, 2013; Wong & Stamp, 2006). The MWOR worms were designed to be highly metamorphic and to evade statistical-based detection -- and they do successfully evade such detection (Sridhara & Stamp, 2012). Both of these metamorphic families have been used in studies of several other malware scoring techniques (Attaluri et al., 2009; Baysa et al., 2013; Lin & Stamp, 2011; Runwal et al., 2012; Shanmugam et al., 2013; Toderici & Stamp, 2013; Wong & Stamp, 2006), and hence they provide a basis for judging the effectiveness of the call graph similarity score considered here.

### *4.1 Test Data*

Our test data consists of 50 NGVCK virus files and a total of 120 MWOR files. The MWOR worms have an adjustable “padding ratio” parameter that specifies the fraction of dead code to worm code. For example, a padding ratio of 2.0 means that each worm has twice as much dead code as actual functioning worm code. The dead code is selected from benign files and, at higher ratios; it serves to effectively defeat statistical-based detection techniques (Sridhara & Stamp, 2012). We consider distinct sets of MWOR worms with padding ratios of 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0. For the NGVCK viruses we use 50 Cygwin utility files as representative examples of benign files. Since MWOR is a Linux worm, we use a set of 20 Linux library files for the representative benign set for the MWOR experiments. These data sets are consistent with those used in previous related research (Baysa et al., 2013; Runwal et al., 2012; Shanmugam et al., 2013; Wong & Stamp, 2006). In all experiments, we score all pairs of malware samples with each other, and we score all pairs consisting of one malware sample and one benign file.

### *4.2 Evaluation Criteria*

To evaluate our results, we use Receiver Operating Characteristic (ROC) curves. To construct an ROC curve, we plot the fraction of true positives versus the fraction of false positives as the threshold varies through the range of scores. The area under the curve (AUC) provides a single measure that enables us to directly compare experimental results. An AUC of 1.0 indicates ideal separation, that is, we can set a threshold for which no false positives of false negatives occur. At the other extreme, an AUC of 0.5 indicates that the binary classifier performs no better than flipping a coin.

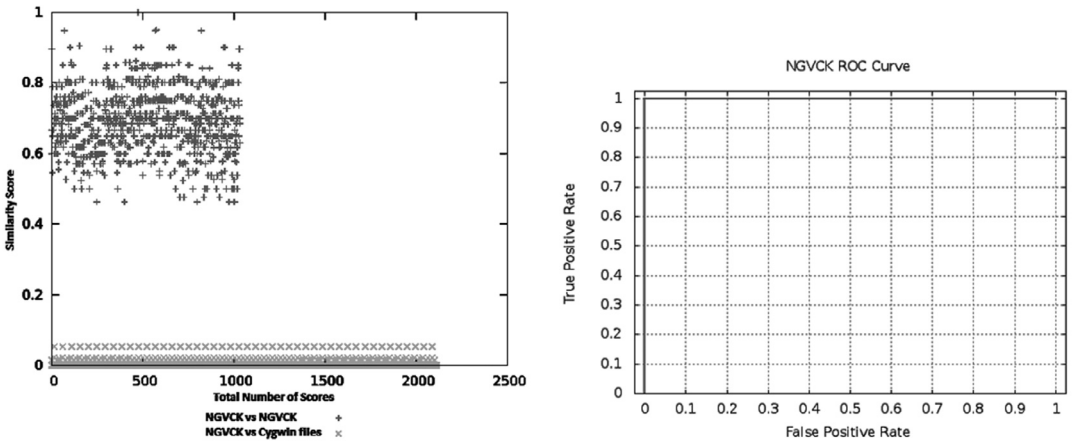
### *4.3 Test Results*

#### *4.3.1 NGVCK*

First, we tested the call graph based scoring technique on NGVCK viruses. A scatterplot of the resulting scores is given in Figure 4 (a), and the corresponding ROC curve appears in Figure 4 (b). In this case, the AUC is clearly 1.0, as we have ideal detection.

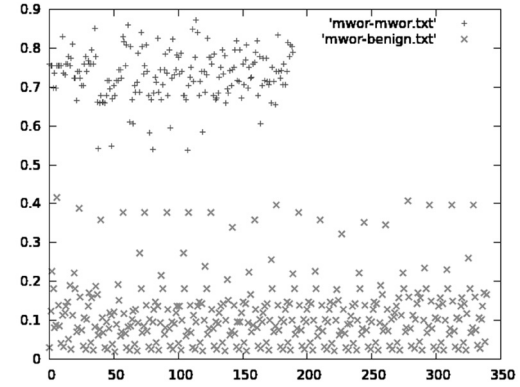
#### *4.3.2 MWOR*

Next, we tested the call graph score on MWOR worms, using a wide range of padding ratios. Recall that the MWOR padding ratio is the fraction of dead code to functional worm code. Figure 5 shows the similarity scores for MWOR worms, where the padding ratio ranges from 0.5 to 3.0. These results show that for padding ratios of 2.0 or less, we obtain ideal classification in each case. However, for padding ratios of 2.5 and above, there will be some misclassifications, regardless of the threshold.

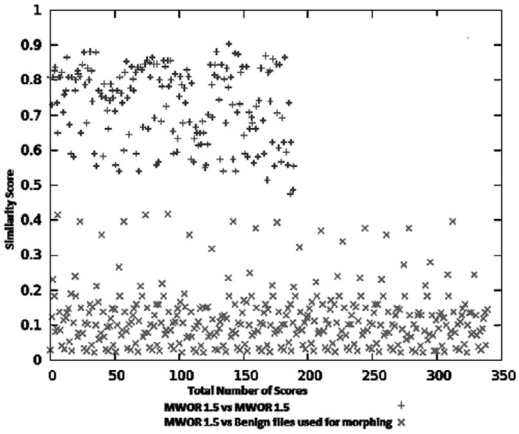



(a) Similarity scores (b) ROC curve

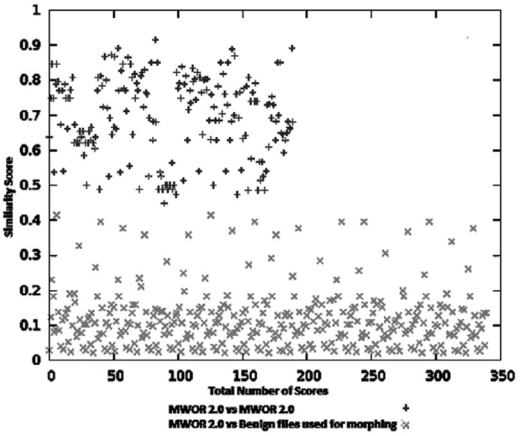
**Figure 4** Call Graph Similarity for NGVCK Virus Family



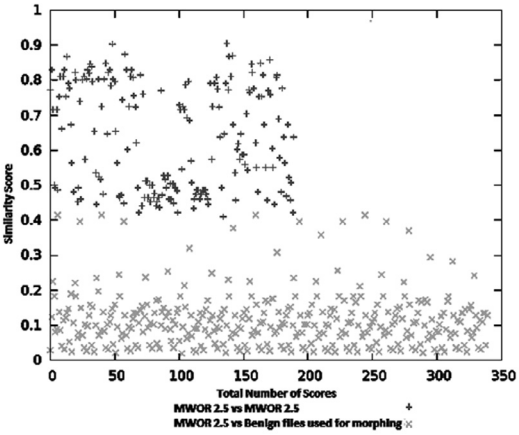

(a) Padding ratio 1.0



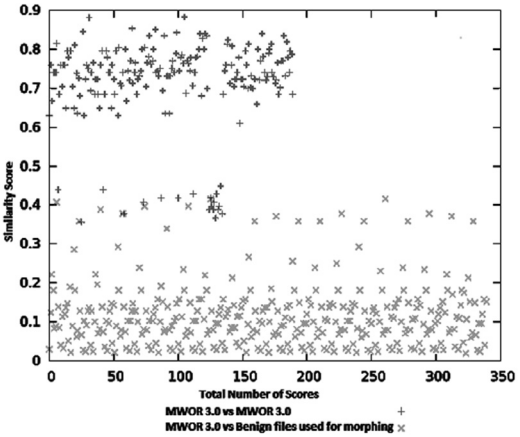

(b) Padding ratio 1.5



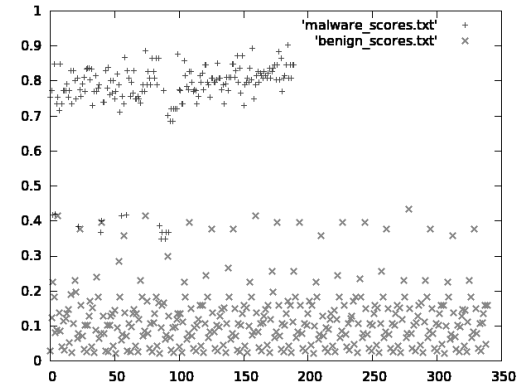

(c) Padding ratio 2.0




(d) Padding ratio 2.5




(e) Padding ratio 3.0




(f) Padding ratio 4.0

**Figure 5** MWOR Similarity Scores

ROC curves corresponding to the scores in Figure 5 were constructed, and the AUC for each computed. The first column of Table 7 contains the AUC statistic for each of the resulting ROC curves.

### *4.3.3 Comparison with Previous Work*

Next, we compare the results obtained using the call graph score to an HMM-based score. As previously mentioned, this HMM score has served as a benchmark for several previous studies on malware detection and hence provides a useful measure of the success of the call graph score, relative to previous work.

For the MWOR worms, a direct comparison (in terms of the AUC statistic) is provided in Table 7, where the HMM results are taken from Sridhara and Stamp (2012) and the "simple substitution" results are from Shanmugam et al. (2013) (which itself improved on the HMM score for the MWOR family). From these results, we see that the call graph technique is superior to both of these other techniques for padding ratios of 1.5 or greater. In Figure 6, we have plotted the results from Table 7 in the form of a bar graph, which clearly shows the robustness of the call graph score with respect to common morphing techniques.

## 5. Conclusion and future work

We implemented a function call graph technique and applied it to the malware detection problem. Opcode analysis and graph coloring techniques are employed to compute this score.

We tested this similarity score on two challenging metamorphic malware families. The results show that the function call graph score outperforms a straightforward HMM-

**Table 7** MWOR AUC Comparison

| Padding ratio | Area under the ROC curve (AUC) | | |
|---|---|---|---|
| | call graph | HMM | simple substitution |
| 0.5 | 1.0000 | 1.0000 | 1.0000 |
| 1.0 | 1.0000 | 0.9900 | 1.0000 |
| 1.5 | 1.0000 | 0.9625 | 0.9980 |
| 2.0 | 1.0000 | 0.9725 | 0.9985 |
| 2.5 | 0.9999 | 0.8325 | 0.9859 |
| 3.0 | 0.9989 | 0.8575 | 0.9725 |
| 4.0 | 0.9979 | 0.8225 | 0.9565 |

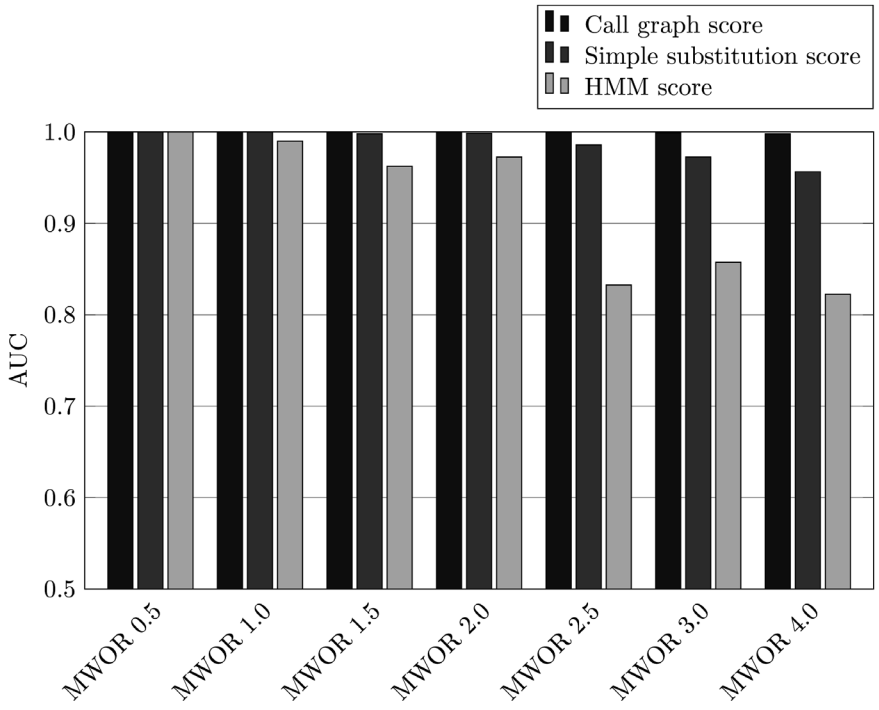



**Figure 6** AUC for Function Call Graph, Simple Substitution, and HMM Scores

based score and a "simple substitution" score. This is impressive, since the HMM score has served as a benchmark in several previous studies, and it has proven difficult to significantly improve on the HMM results.

Future work could focus on possible improvements to call graph score technique considered in this paper. Specifically, the step where we match similar functions is worth reconsidering. Possible alternatives to the graph coloring approach used here include any of a variety of additional statistical techniques, such as HMM analysis (Wong & Stamp, 2006), chi-squared statistics (Toderici & Stamp, 2013), and the "simple substitution" distance in Shanmugam et al. (2013). In addition, structural techniques such as the entropy-based score in Baysa et al. (2013) and Sorokin (2011) or the compression-based score in Lee et al. (2015) could prove more robust than scores that rely directly on opcode-based analysis.

## References

Attaluri, S., McGhee, S. and Stamp, M. (2009), 'Profile hidden Markov models and metamorphic virus detection', *Journal in Computer Virology*, Vol. 5, No. 2, pp. 151-169.

Aycock, J.D. (2006), *Computer Viruses and Malware*, Springer-Verlag, New York, NY.

Baysa, D., Low, R.M. and Stamp, M. (2013), 'Structural entropy and metamorphic malware', *Journal of Computer Virology and Hacking Techniques*, Vol. 9, No. 4, pp. 79-192.

Bilar, D. (2007), 'On callgraphs and generative mechanisms', *Journal in Computer Virology*, Vol. 3, No. 4, pp. 285-297.

Borello, J. and Mé, L. (2008), 'Code obfuscation techniques for metamorphic viruses', *Journal in Computer Virology*, Vol. 4, No. 3, pp. 211-220.

Carrera, E. and Erdelyi, G. (2004), 'Digital genome mapping-advanced binary malware analysis', *Proceeding Virus Bulletin Conference*, City, State, pp. 187-197.

Christodorescu, M., Jha, S. and Kruegel, C. (2007), 'Mining specifications of malicious behavior', *Proceedings of the 6th joint meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering*, Dubrovnik, Croatia, pp. 5-14.

Deshpande, P. (2013), 'Metamorphic detection using function call graph analysis', Unpublished master thesis, San Jose State University, San Jose, CA.

Hex-Ray. (n.d.), 'IDA: about', available at: http://www.hex-rays.com/products/ida/index.shtml (accessed on February 14 2017).

Karnik, A., Goswami, S. and Guha, R. (2007), 'Detecting obfuscated viruses using cosine similarity analysis', *Proceedings of the First Asia International Conference on Modelling Simulation*, Phuket, Thailand, pp. 165-170.

Kazi, S. and Stamp, M. (2013), 'Hidden Markov models for software piracy detection', *Information Security Journal: A Global Perspective*, Vol. 22, No. 3, pp. 140-149.

Konstantinou, E. and Wolthusen, S. (2008), 'Metamorphic virus: analysis and detection', Technical Report RHUL-MA-2008-02, Department of Mathematics, Royal Holloway, University of London , Egham, UK.

Lee, J., Austin, T.H. and Stamp, M. (2015), 'Compression-based analysis of metamor-phic malware', *International Journal of Security and Networks*, Vol. 10, No. 2, pp. 124-136.

Lin, D. and Stamp M. (2011), 'Hunting for undetectable metamorphic viruses', *Journal in Computer Virology*, Vol. 7, No. 3, pp. 201-214.

Panda Security (2011) 'Virus, worms, trojans and backdoors: other harmful relatives of viruses', available at: http://www.pandasecurity.com/homeusers-cms3/security-info/about-malware/generalconcepts/concept-2.html (accessed on February 14 2017).

Rabiner, L.R. (1989), 'A tutorial on hidden Markov models and selected applications in speech recognition', *Proceeding of the IEEE*, Vol. 77, No. 2, pp. 257-286.

Runwal, N., Low, R. and Stamp, M. (2012), 'Opcode graph similarity and metamor-phic detection', *Journal in Computer Virology*, Vol. 8, No. 1-2, pp. 37-52.

Shang, S., Zheng, N., Xu, J., Xu, M. and Zhang, H. (2010), 'Detecting malware variants via function-call graph similarity', *5th International Conference Malicious and Unwanted Software*, Nancy, France, pp. 113-120.

Shanmugam, G., Low, R. and Stamp, M. (2013), 'Simple substitution distance and metamorphic detection', *Journal of Computer Virology and Hacking Techniques*, Vol. 9, No. 3, pp. 159-170.

Snakebyte (2000) 'Next generation virus construction kit (NGVCK)', available at: http://vxheaven.org/vx.php?id=tn02 (accessed on 14 February 2017).

Sorokin, I. (2011), 'Comparing files using structural entropy', *Journal in Computer Virology*, Vol. 7, NO. 4, pp. 259-265.

Sridhara, S. and Stamp, M. (2012), 'Metamorphic worm that carries its own morphing engine', *Journal of Computer Virology and Hacking Techniques*, Vol. 9, No. 2, pp. 49-58.

Stamp, M. (2015), 'A revealing introduction to hidden Markov models', available at: http://www.cs.sjsu.edu/~stamp/RUA/HMM.pdf (accessed on 14 February 2017).

Symantec. (2011) 'Internet security threat report, Vol. 17', available at: http://www.symantec.com/content/en/us/enterprise/other_resources/b-istr_main_report_2011_21239364.en-us.pdf (accessed on 14 February 2017).

Symantec. (n.d.), `Virus construction kit', available at: http://computervirus.uw.hu/ch07lev1sec7.html (accessed on 14 February 2017).

Szor, P. (2000), 'The new 32-bit medusa', available at: https://www.virusbulletin.com/uploads/pdf/magazine/2000/200012.pdf (accessed on 14 February 2017).

Szor, P. (2005), 'Advanced code evolution techniques and computer virus generator kits', available at: http://www.informit.com/articles/article.aspx?p=366890&seqNum=6 (accessed on 14 February 2017)

Szor, P. and Ferrie, P. (2001), 'Hunting for metamorphic', available at: https://www.symantec.com/avcenter/reference/hunting.for.metamorphic.pdf (accessed on 14 February 2017).

Tamboli, T., Austin, T. and Stamp, M. (2014), 'Metamorphic code generation from LLVM IR bytecode', *Journal of Computer Virology and Hacking Techniques*, Vol. 10, No. 3, pp. 177-187.

The Mental Driller. (2002) 'Metamorphism in practice or "How I made MetaPHOR and what I've learnt", available at: http://biblio.l0t3k.net/magazine/en/29a/ (accessed on 14 February 2017).

Tips Trik Dan Berbagi Informasi. (n.d.), 'Virus creation tools: VX heavens', available at: http://oktridarmadi.blogspot.com/2009/09/virus-creation-tools-vx-heavens.html (accessed on 14 February 2017).

Toderici, A.H. and Stamp, M. (2013), 'Chi-squared distance and metamorphic virus detection', *Journal of Computer Virology and Hacking Techniques*, Vol. 9, No. 1, pp. 1-14.

VX Heavens (n.d.), 'Access macro generator', available at: http://download.adamas.ai/dlbase/Stuff/VX%20Heavens%20Library/static/vdat/creatrs1.htm (accessed on 14 February 2017).

Wong, W. and Stamp, M. (2006), 'Hunting for metamorphic engines', *Journal in Computer Virology*, Vol. 2, No. 3, pp. 211-219.

Xu, M., Wu, L., Qi, S., Xu, J., Zhang, H., Ren, Y. and Zheng, N. (2013), 'A similarity metric method of obfuscated malware us-ing function-call graph', *Journal of Computer Virology and Hacking Techniques*, Vol. 9, No. 1, pp. 35-47.

You, I. and Yim, K. (2010), 'Malware obfuscation techniques: a brief survey', *International Conference on Broadband, Wireless Computing, Communication and Applications,* Fukuoka, Japan, pp. 297-300.

Zbitskiy, P. (2009), 'Code mutation techniques by means of formal grammars and automatons', *Journal in Computer Virology*, Vol. 5, No. 3, pp. 199-207.

## About the authors

**Prasad Deshpande** holds an undergraduate degree from Pune University in India, and he completed his Master's in Computer Science at San Jose State University in December 2013. Prasad currently works on disaster recovery and business continuity products at Symantec Corporation in Silicon Valley. E-mail address: prasad.0210@gmail.com

**Mark Stamp** can neither confirm nor deny that he spent more than seven years working as a cryptologic mathematician at the super-secret National Security Agency. However, he can confirm that he spent two years at a small Silicon Valley startup, developing a security-related product. For the past dozen years, Mark has been employed as a Professor of Computer Science at San Jose State University, where he teaches courses in information security, conducts research on topics in security and machine learning, writes security-related textbooks, and supervises ridiculously large numbers of Masters student projects. Perhaps not surprisingly, most of his students projects are security-related.

Corresponding Associate Professor, Department of Computer Science, San Jose State University, One Washington Square, San Jose, CA 95192. Tel: 408-492-5094. E-mail address: mark.stamp@sjsu.edu

*MIS Review* Vol. 21, No. 1, Nos. 1/2, September(2015)/March (2016), pp. 35-50
DOI: 10.6131/MISR.2015.2101.03


# Secure WLAN Handoff Scheme with Continuous Authentication

*Rajeev Singh[1], Teek P. Sharma[2]*
*[1]G. B. Pant University of Agriculture and Technology, Uttarakhand, India;*
*[2]National Institute of Technology, Hamirpur (H.P.) India*

ABSTRACT: *Handoffs are essential for providing continuous mobility to a wireless Station (STA) in an Enterprise LAN. An important requirement of the handoff is to establish connection of the roaming STA with a new Access Point (AP) securely and quickly such that the undergoing communication remains unaffected. We propose a novel handoff scheme for enhancing the handoff performance and security. The scheme is a lightweight and reactive method for transferring the keying material i.e., STA context to new AP. Scheme utilizes Key Hiding Communication (KHC) scheme for ongoing data communication between STA and AP. It provides continuous authentication between STA and APs. Computation and communication cost for the handoff process are calculated and security analysis is done. A comparison with other handoff schemes is also provided.*

KEYWORDS: *WLAN Security, Key Based Communication, Handoff, Lightweight Authentication.*

## 1. Introduction

Wireless networks supporting real time applications like Voice over IP (VOIP), e-conference etc. requires instant availability and seamless secure roaming (Bojkovic et al., 2005). This is achieved by wireless Station (STA) handoff to the new Access Point (AP). The desired properties of a good handoff are: (1) handoff should be completed within time limits as suitable for the multimedia and real time applications and (2) the handoff should be performed securely.

The communication interruption time tolerable for multimedia and real time applications during handoff is approximately 50 ms (Lee, 2010; Lee & Hunt, 2010). This is the time when mobile STA cannot send or receive data packets from its correspondent nodes. This time should be minimized such that the STA communicates with its correspondent node continuously. The 802.11i WLAN security standard provides the secure STA authentication at AP by utilizing the Authentication Server (AS) services. The secure STA authentication (default Full EAP/TLS) evolves shared secret key between STA and AP. The entire process takes time of the order of 300 ms to 4 seconds (Martinovic et

al., 2006, 2007) which make it unsuitable for the handoffs. This handoff time should be reduced while maintaining the security properties. For reducing the handoff time, pre-authentication is used where full 802.1X authentication is performed before starting the handoff. Here, STA starts EAP/TLS authentication with the candidate AP (new AP with which STA handoff is performed) through old AP connection. Old AP forwards the authentication messages to the AS. This process is termed as pre authentication. It ends when STA and new AP receive new PMK. Later, when handoff is initiated only 4-way handshake is required to complete the authentication. Involvement of AS is not required during the handoff reestablishment of trust relationship. This results in overall reduction in handoff delay and packet loss (Compton, 2008; IEEE 802.11i, 2004; Kassab et al., 2005).

Predictive authentication and proactive key distribution are proposed by the researchers to reduce time in locating and predicting the candidate AP. This involves overheads and security issues. In proactive key distribution, a group of candidate access points are determined and new shared key (i.e., PMK) is distributed among them before the handoff (Hur et al., 2007; Ling et al., 2010; Mishra et al., 2004a, 2004b; Pack et al., 2005). This introduces extra communications with all candidate APs instead of communicating with one candidate AP. In prediction based handoff techniques, if the prediction misses, the complete authentication is required for communication with the candidate AP (Chien et al., 2008; Kassab et al., 2005; Pack & Choi, 2004). This affects smooth handoff process. 802.11i also does not define candidate AP prediction as an inaccurate prediction may lead to large resource wastage. Thus, proactive key distribution is suggestive as compared to predictive authentication provided the extra communication with the group of candidate access points is lightweight and efficient. Another handoff termed as reactive method is also proposed by researchers where STA authentication is executed after the candidate AP is selected. The candidate AP is usually selected by the STA and then the security context (i.e., keying material) is transferred to this AP. For transferring security context, STA requests to AS via old AP, then AS transfers security context to the candidate AP. Security of intermediate messages that are used for authentication and transferring security context is an issue here.

For providing fast and secure handoff for the mobile STA in WLANs, standard bodies IEEE and IETF have defined protocols like Control and Provisioning of Wireless Access Points (CAPWAP), Hand Over Keying (HOKEY) and IEEE 802.11r (Task group r) (Clancy, 2008). CAPWAP supports centralized management of APs. HOKEY extends the Authentication, Authorization and Accounting (AAA) architecture to support key deriving and distribution with involving full EAP authentication. 802.11r depends upon passing credentials directly between APs for handover. Though CAPWAP takes very less time, it is more or less reauthentication with centralized Access Controller (AC), followed by key transfer to new Wireless Termination Points (WTP). HOKEY is successful in

multidomains but it takes more communication time. Among these three (CAPWAP, HOKEY and 802.11r), 802.11r is more efficient in terms of communication overheads. It still has issues concerning the safe transfer of key between APs.

We propose a novel Secure WLAN Handoff Scheme that maintains security properties while evolving and transferring the security context (key and initial vector) to the candidate AP. The scheme is lightweight and uses reactive method for handoff. The proposed secure handoff scheme not only provides the handoff within desired time limits required by multimedia and real time data traffic but also maintains desired security using primitives like lightweight authentication, encryption and Message Integrity Code (MIC) to all the messages involved in the handshake process. Two kinds of APs are defined in the scheme: normal AP and Domain Controller AP (DCAP). STA request DCAP through normal AP by putting ID of the candidate AP. DCAP in turn distributes the STA context (key and initial vector) to the candidate AP. Thus, when STA roams into the area of candidate AP, less time is involved in the STA authentication at the candidate AP.

The rest of the paper is divided into 4 sections. Section 2 presents the related work done. Section 3 proposes the secure handoff scheme. Section 4 discusses the performance issues and comparison among the related handoff schemes. Section 5 provides security analysis while section 6 provides conclusion.

## 2. Related work

Several predictive and pre-authentication schemes are proposed for enhancing the handoff (Compton, 2008; IEEE 802.11i, 2004; Kassab et al., 2005). Kassab et al. (2005) proposed statistical methods for modeling the mobility pattern of the STA. As a result of the model, a set of access points are selected in the handoff region. STA can associate with any one of them and thus, STA needs to exchange fewer messages with the candidate AP. An interesting concept of neighbor graph has been introduced in (Mishra et al., 2004a, 2004b; Shin et al., 2004) that identifies the candidate access points, one of which would associate with STA. The key material is distributed to these candidate access points. New Pairwise Master Keys (PMKs) are generated using PMK trees. As the key material is received by APs before the handoff, this process is termed as proactive key distribution. Communication between AP and AS is reduced in the scheme. Still the process introduces communication overheads between the candidate access points and the AS. For further reducing this overhead, improvements like selective neighbor caching and proactive key distribution with anticipated 4-way handshake are suggested in (Hur et al., 2007; Ling et al., 2010; Pack et al., 2005). In former, the STA context is transferred to only selective neighbors. In case only one neighbor is selected, it becomes similar to reactive handoff

method. In latter, the 4-way handshake is not required at the start of handoff rather STA generates PTKs before handoff with the help of candidate AP list sent by the AS. The method is useful mainly for 802.1X based networks.

Fast AP Transition Protocol (FATP) scheme uses proactive key distribution technique to transfer existing security context to the candidate AP before handoff (Lee, 2010; Lee & Hunt, 2010). After transferring the security context, the roaming STA and candidate AP mutually verify each other's identity and derive new session keys. This does not require involvement of AS during the STA reassociasation with the candidate AP. The resulting trust relationship has same properties of full EAP/TLS authentication and has less cost in terms of latency, computational power and network traffic overhead. It implements authentication followed by reassociation. Authentication leads to establishment of trust relationships and reassociation leads to changing the AP attachment. The scheme also claims to work under DoS attacks. The issue of DoS attacks is not addressed by any other handoff solutions. Scheme defines two types of intra-domain handoff scenarios namely "R0 to R1" handoff and "R1 to R1" handoff. A secure and fast handoff technique is proposed at (Maccari et al., 2006). It is based upon the concept that when STA moves towards the candidate AP, then candidate AP request for PMK from the AS along with proving its request as legitimate. For this STA gives a token to candidate AP who forwards it to the AS, proving request as legitimate. Token generation is based upon hash calculation using PMK shared between STA and AS. The scheme works only for 802.1X based networks. Another fast authentication scheme for the wireless LAN is proposed at (Zhang et al., 2010, 2011). The scheme reduces the authentication latency during the handoff using a tunnel technique. The tunnel technique provides secure communication. Roaming STA selects the new (candidate) AP and starts the fast authentication process. The ID of candidate AP is transferred to old AP. The old AP uses MAC address of the candidate AP for generating the pair wise tunnel key. This key is then transferred to both mobile STA and the candidate AP. The roaming STA now tries to associate with the candidate AP using this temporal tunnel key. The STA packets are still transferred to old AP which then forwards them to the destination. Mean while the candidate AP starts the EAP/TLS process with STA to generate PMK and PTK. Once PTK is generated temporal tunnel key is obsoleted and the communication starts using the new PTK.

## 3. Proposed secure WLAN handoff scheme

There are two types of APs involved in the scheme: Domain Controller Access Point (DCAP) and normal Access Points (AP1, AP2, ..., APn). There is only one domain controller AP in a particular domain while there are several normal APs in the domain. It is assumed that domain controller AP has high computation capacity. The main functionality

of domain controller is to authenticate wireless stations and access points. Hence, this role can even be performed by the authentication server (AS) in a domain. Apart from this, DCAP not only evaluates fresh communication key for STA but also forwards refreshed key to the new AP during handoff. The STA performs handoff among normal APs. The proposed scheme has initialization and communication phases similar to Key Hiding Communication (KHC) scheme (Singh & Sharma, 2013a); in addition it has handoff phase.

### *3.1 Initialization phase*

Each wireless station and access point initially authenticates itself to the DCAP and evolves shared master key for communication. For this initial authentication, STA and AP utilize the initialization phase of the Key Hiding Communication (KHC) scheme proposed by Singh and Sharma (2013a). During KHC initialization phase, pair of communicating nodes evolve master key (MK) between them. Using the KHC process, STA is initially authenticated at DCAP and then a master key ($MK^{STA\text{-}DCAP}$) is evolved at the STA and DCAP. Similarly, normal APs evolve secret master key ($MK^{AP1\text{-}DCAP}$, $MK^{AP2\text{-}DCAP}$, ..., $MK^{APn\text{-}DCAP}$) with the domain controller AP. $MK^{STA\text{-}DCAP}$ is termed as MK of the STA. DCAP transfers STA MK securely by encrypting using $MK^{AP1\text{-}DCAP}$ to the current communicating AP (say AP1). Using MK of the STA, initial parameters i.e., $CD_0$, C0 and C1 are shared between STA and AP. Such initial parameters are also shared between normal APs and domain controller AP. Thus, after initialization each pair of devices has its own set of $K_0$, $IV_0$, C0 and C1 required for secure communication. The naming conventions used in the paper are mentioned in Table 1. The wireless handoff scenario along with keys and parameters evolved after the initialization process is shown in Figure 1. As a station may perform frequent handoff, extra memory and computation overheads are involved at nodes. Hence, we assume 128 bit shared master key, 64 bit K0, 64 bit IV0, 64 bit C0 and 64 bit C1 in the KHC scheme.

### *3.2 Communication phase*

After initialization, the communication between STA and AP1 is performed like KHC communication phase. In communication phase of KHC, key refreshing and hiding concept for sharing the symmetric secret key ($K_i$) and initial vector ($IV_i$) is introduced. Key and IV refreshing are done using MK. After refreshing, the secret encryption key and IV are protected by XORing with counters C0 and C1 respectively. The key and IV are then mixed with each other before transferring them to the receiver. For mixing, the new byte locations for placing the $K_i$ and $IV_i$ in the $CD_i$ are calculated with the help of existing $K_{i-1}$. Mixed key and IV is termed as Codeword ($CD_i$). This codeword is added to the transmitted frame and delivered to the recipient. Corresponding frame MIC is calculated using $K_{i-1}$. The recipient extracts the key from the codeword, compares it with its own evaluated key, thereby authenticating the sender. Key ($K_i$) along with $IV_i$, is then used to

encrypt the data frame to be transmitted next. The key verification at the receiver also provides authentication per frame. The authentication is lightweight as key verification involves operations like increment, XOR and modulus evaluations. MIC of only the authenticated frames is checked. The verified key is utilized to encrypt the data and evaluate MIC for the next frame.

**Table 1** Naming Convention

| Domain Controller AP: DCAP | *Codewords* |
|---|---|
| *Device Identifiers* | New codeword: $CD_i$ |
| STA identifier: IDSTA (64 bits) | STA-domain controller: $CD^{STA\text{-}DCAP}$ |
| APi identifier: IDAPi (64 bits) | APi-domain controller: $CD^{APi\text{-}DCAP}$ |
| Domain Controller identifier: IDDC (64 bits) | STA-APi: $CD^{STA\text{-}APi}$ |
| *Shared Keys* | *Initial Vector* |
| Master Key: MK | Previous IV: $IV_{i-1}$ |
| Previous Key: $K_{i-1}$ | New IV: $IV_i$ |
| New Key: $K_i$ | STA-domain controller: $IV^{STA\text{-}DCAP}$ |
| STA-domain controller: $K^{STA\text{-}DCAP}$ | APi-domain controller: $IV^{APi\text{-}DCAP}$ |
| APi-domain controller: $K^{APi\text{-}DCAP}$ | STA-APi: $IV^{STA\text{-}APi}$ |
| STA-APi: $K^{STA\text{-}APi}$ | |

Thus, for transferring data between STA and AP in the proposed scheme, first refreshing of key and IV is done then key and IV protection is done which is then followed by key and IV mixing i.e., codeword ($CD_i^{STA}$) formation. The $CD_i^{STA}$ is sent as extra bytes in the WLAN header. AP verifies codeword and hence authenticates the STA. The contents within the frame body are encrypted using $K_{i-1}$ and $IV_{i-1}$. Each frame is protected via MIC addition to frame. The receiver verifies the $K_i$ and $IV_i$ from the received codeword ($CD_i^{STA}$) using protection and mapping. This verification provides per frame authentication. $K_i$ and $IV_i$ are then used to encrypt next frame. Thus, encryption key for each successive data frame is refreshed in this process. Similar key refreshing and verification also takes place between other two pairs i.e., AP and DCAP; STA and DCAP.

Two kinds of frames are used in the proposed handoff scheme: communication frames and handoff frames. The frame types and their corresponding contents are shown in Figure 2. Communication frames are same as that of the KHC scheme with the exception that the codeword size is now 128 bits only. Three different handoff frames are required in the following: between STA and AP1 (current AP); between AP1 and DCAP; and between DCAP and AP2 (new/candidate AP). This implies that 2 bits are required in frame header for indicating the frame type. We consider the proposed implementation strategy by Ren et al. (2004) and use bits B3 and B4 of the data frame control field for frame identification.

**For AP1 and DCAP communication**

Master Key (128 bit): $MK^{AP1\text{-}DCAP}$

Codeword (128 bit): $CD_0^{AP1\text{-}DCAP}$

Initial Key (64 bit): $K_0^{AP1\text{-}DCAP}$

Initial Vector (64 bit): $IV_0^{AP1\text{-}DCAP}$

Counters (64 bit each): $C0^{AP1\text{-}DCAP}$, $C1^{AP1\text{-}DCAP}$

**For STA and DCAP communication**

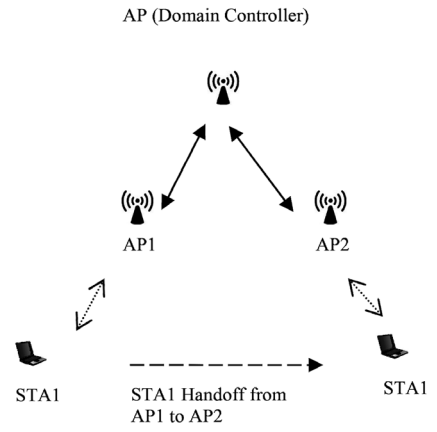
Master Key (128 bit): $MK^{STA\text{-}DCAP}$

Codeword (128 bit): $CD_0^{STA\text{-}DCAP}$

Initial Key (64 bit): $K_0^{STA\text{-}DCAP}$

Initial Vector (64 bit): $IV_0^{STA\text{-}DCAP}$

Counters (64 bit each): $C0^{STA\text{-}DCAP}$, $C1^{STA\text{-}DCAP}$




**For STA and AP1 communication**

Master Key (128 bit): $MK^{STA\text{-}AP1}$

Codeword (128 bit): $CD_0^{STA\text{-}AP1}$

Initial Key (64 bit): $K_0^{STA\text{-}AP1}$

Initial Vector (64 bit): $IV_0^{STA\text{-}AP1}$

Counters (64 bit each): $C0^{STA\text{-}AP1}$, $C1^{STA\text{-}AP1}$

**Figure 1** WLAN Handoff Scenario along with Keys and Parameters

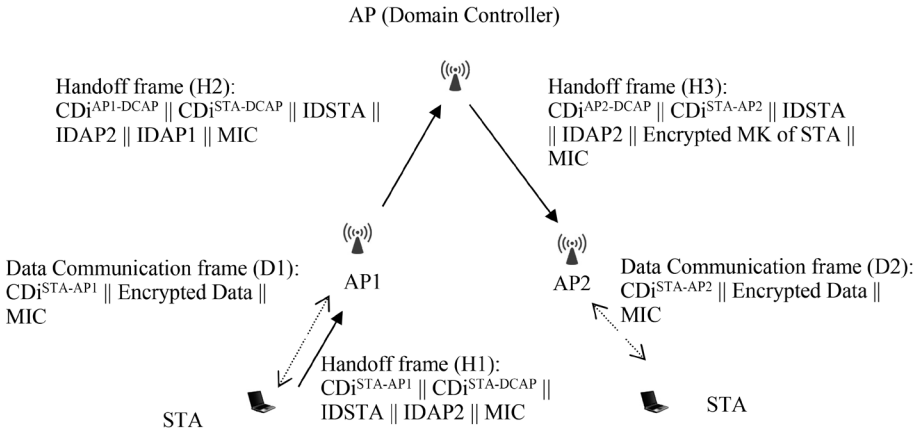
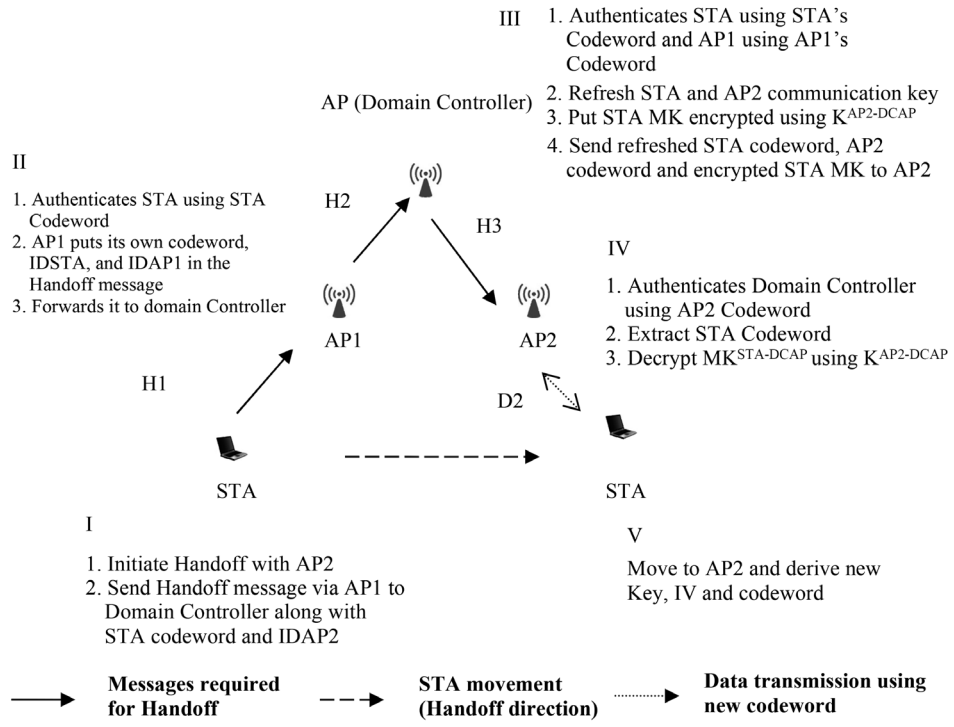



**Figure 2** Frame Types and Contents

Combination "00" indicates communication frame between STA and normal AP, "01" indicates handoff request from STA to AP, "10" indicates handoff request by old AP (AP1) to domain controller AP, "11" indicates handoff response from domain controller AP to new AP (AP2).

### 3.3 Handoff phase

STA which is currently under AP1 (old AP), sends handoff request to the AP1whenever handoff with AP2 (candidate AP) is required. STA sets the handoff bits in the frame header as "01" and puts its own ID as well as ID of AP2 in the frame body. New codewords ($CD_i^{STA\text{-}AP1}$ and $CD_i^{STA\text{-}DCAP}$) and MIC are appended to it. On receiving this handoff request, AP1 removes




Node Processing: I (At STA), II (At AP1), III (At Domain Controller), IV (At AP2), V (At roaming STA)

Handoff Messages: H1, H2, H3, H4

All the messages sent by any of the node are protected using the MIC evaluated using its own Key

**Figure 3** STA Handoff with AP2

$CD_i^{STA\text{-}AP1}$ and verifies authenticity of the STA through STA codeword. AP1 then appends IDSTA, its own IDAP1 and codeword ($CD_i^{AP1\text{-}DCAP}$) in the frame body along with

MIC. This request is forwarded to the domain controller. Domain controller authenticates STA and AP1 by verifying their codewords ($CD_i^{STA\text{-}DCAP}$ and $CD_i^{AP1\text{-}DCAP}$). It generates new codewords for AP2 ($CD_i^{AP2\text{-}DCAP}$) and STA ($CD_i^{STA\text{-}AP2}$), puts them in response frame to AP2. Domain controller also puts the encrypted MK for the current STA session. On receipt of response frame, AP2 extracts its own codeword, authenticates domain controller and extracts the STA codeword ($CD_i^{STA\text{-}AP2}$). The roaming STA request is verified using this extracted codeword. AP2 also extracts the MK of the STA session by decryption using $K_{i-1}^{AP1\text{-}DCAP}$. With the help of MK, AP2 further performs key refreshing and safe key transfer with STA. This accomplishes STA handoff with AP2. The entire handoff process the shown in Figure 3.

The computation and communication costs are involved in the proposed handoff scheme Three handoff messages i.e., H1, H2, H3 are exchanged among STA, AP1, DCAP and AP2. Scheme is reactive and therefore context/keying material is not supplied to all APs as done in proactive schemes, rather only one candidate AP is given STA communication key. The scheme provides secure communication as all the 3 handoff messages are protected by MIC and mutual authentication exists among all parties i.e., STA, AP1, DCAP and AP2 via codeword verification. The scheme also provides protection to handoff against DoS attacks at AP2. Before the handoff message H3 is received at AP2 all the DoS attack packets are dropped. Once STA's communication message i.e., D2 is received at the AP2, the entire process is nothing but the KHC communication and is safe under DoS attack.

## 4. Performance evaluation

For the proposed scheme, we calculate communication cost, network overload and computation cost required for performing the handoff and compare them with CAPWAP, HOKEY, IEEE 802.11r and FATP.

### *4.1 Communication cost and network overload*

STA requires a total network overload of four messages i.e. H1, H2, H3 and D2 to perform the handoff successfully with AP2. For simplification, we assume that the transmission latency between STA and AP is same as that of between AP and Domain Controller. The transmission time between two nodes using an IP socket (using UDP datagram) between two systems averages to 1.9796 milliseconds (Singh and Sharma, 2013b). In proposed handoff scheme, four such communications are required: between STA and AP1 (old AP); between AP1 and domain controller; between domain controller and AP2 (candidate AP); between STA and AP2.

Therefore, communication cost of our proposed handoff scheme is equal to 4 × 1.9796 ms = 7.9184 ms.

**Table 2** Computation Cost of the Proposed Handoff Scheme

| Device | Processing/Computations | Number |
|---|---|---|
| STA | Key and IV refresh | 03 |
| | Key protection and Key mapping | 03 |
| | Encryption | 01 |
| AP1 | Key and IV refresh | 02 |
| | Key protection and Key mapping | 02 |
| Domain Controller AP | Key and IV refresh | 04 |
| | Key protection and Key mapping | 04 |
| | Encryption | 01 |
| AP2 | Decryption | 02 |

### *4.2 Computation cost*

The proposed scheme involves computations at the STA, old AP (AP1), domain controller AP (DCAP) and candidate AP (AP2). The computations involved are listed in Table 2. Key protection and key mapping is done using XOR and modulus operations, respectively while Key and IV are refreshed using hash calculations. Both XOR and modulus are mathematical operations and takes negligible time as compared to cryptographic primitives. Therefore, we can ignore them from calculations. We consider the average time taken for hash calculation as 0.1256 ms (Singh & Sharma, 2013b). Total number of key and IV refreshing required are 09 while number of encryptions and decryptions required are 02 respectively. As key refreshing and IV refreshing both require hash calculation, 18 hash calculations are required for handoff. For maintaining integrity, MIC computation and verification is required for each of the 4 frames. Thus, computation time for the handoff process is:

$18 \times 0.1256 + 2 \times 0.1223 + 2 \times 0.0533 + 4 \times 0.193 = 3.3834$ ms

Total time required for handoff = communication time + computation time = 7.9184 ms + 3.3834 = 11.3018 ms (<< 50 ms)

Hence, the proposed scheme is well suited for multimedia and real time applications.

### *4.3 Comparison with other secure handoff schemes*

We compare proposed scheme with the existing handoff schemes and standards like CAPWAP, HOKEY, IEEE 802.11r and Fast AP Transition Protocol (FATP) in Table 3.

**Table 3** Comparison of Our Scheme with Other Secure Handoff Solutions

| Secure Handoff Solutions | Change Required in 802.11 | Fresh Session Key | Fresh Traffic key | Communication overhead | Issues |
|---|---|---|---|---|---|
| CAPWAP | No change in 802.11 but requires operations at IP layer | Not derived | Yes, AC executes 4-way handshake with STA and delivers new traffic keys for WTP | $4(T_w + T_c) + T_c = 85$ µsec | Fresh session keys are not evolved in the handover |
| HOKEY | No change in 802.11 but requires operations at IP layer | Yes | Yes | $2(T_w + T_a) + 4T_w =$ 130 µsec | Communication overhead is a concern |
| IEEE 802.11r | Yes, change in 802.11 protocol itself | Yes, initial AP (R0KH) generates session keys for other APs as PMK-R1, PMK-R2 etc. | Yes, Derived using PMK-R1, PMK-R2 etc. | $2T_w + 2T_c + 2T_w =$ 70 µsec | Caching of PMK by First Authenticator (PMK-R0), Key management involving key transfer between R0KH and R1KH will require $O(n^2)$ keys to be managed between n APs |
| FATP | Yes, termed as extended authentication frame format | Yes, using PMK-R0 | Yes, Derived using PMK-R1 | $2T_w + 4T_c^* = 50$ µsec | Caching of PMK is required. Scheme still suffer under DoS attacks. |
| Proposed scheme | Yes, change in frame formats required | Yes, By Domain controller | Yes, per frame by calculating hash | $2T_w + 2T_c = 40$ µsec | Scheme requires modifications in 802.11 frame formats |

Note. * considering secure inter Access Point communication requires only two messages and the handoff takes place from R0 to R1 AP

CAPWAP and HOKEY do not change the existing 802.11 frame structure. All except CAPWAP scheme generates fresh session keys. Fresh traffic keys are generated by all the schemes. Communication overhead in our scheme is less as compared to any other scheme. For calculating communication overhead, we assume in a typical network: transmission latency ($T_w$) between STA and AP is equal to 15 μsec, latency ($T_c$) between any two relative close devices including AP to AP and WTP to AC is equal to 5 μsec and latency ($T_a$) between infrastructure components and local AAA server is equal to 20 μsec (Clancy, 2008).

## 5. Security analysis

The proposed handoff scheme shortens the handoff latency by initiating a key transfer process prior to moving to the new AP and performing handoff. The security properties of the scheme are analysed in this section.

### *5.1 Protects STAs from re-associating to malicious APs*

As all the packets bear the codeword for authenticating a frame, no malicious STA is able to associate with the normal AP. AP1 authenticates STA by verifying its codeword ($CD_i^{STA\text{-}AP1}$). Domain controller authenticates STA and AP1 by verifying their codewords ($CD_i^{STA\text{-}DCAP}$ and $CD_i^{AP1\text{-}DCAP}$). New AP (AP2) authenticates domain controller by verifying AP2's codeword ($CD_i^{AP2\text{-}DCAP}$). On receipt of communication frame from STA, the codeword of the STA is also verified. Thus, all the frames used in the handoff are authenticated. This protect STAs from re-associating to Malicious APs.

### *5.2 Evolves fresh keys even during handshake*

At old AP (AP1), STA communicates using KHC and hence fresh key and IV are evolved per frame. On performing the handoff, STA refreshes its key and IV. Using key and IV, STA derives fresh codeword for communication with the new AP (AP2). Once codeword is verified, STA's communication with AP2 proceeds further with evolving fresh key and IV.

### *5.3 Continuous authentication is provided*

At old AP (AP1) STA communicates using KHC and hence enjoys continuous authentication. On performing the handoff, STA refreshes its key and IV. Using key and IV it derives a fresh codeword for communication with the new AP (AP2). For each frame, STA authentication process is continued with AP2. Hence, STA enjoys continuous authentication.

### 5.4 Protection against DoS attacks

In KHC scheme the computational DoS attack has less impact on AP1. AP1 protected by KHC scheme is able to maintain its communication under the computational DoS attack by verifying the codeword followed by MIC verification. This method of verifying codeword before MIC verification helps in protection against computational DoS attack. We realized that AP2 behavior under DoS attack while performing the handoff is same as that of KHC behavior. In handoff situation, STA1 moves to AP2 while communicating with STA2. During this period AP2 is under DoS attack and the attacker's objective is to hamper the handoff process at AP2. None of the attack packets are considered for processing till AP2 gets message H3 for STA handoff. This means all attack packets are dropped till H3 is received. After AP2 gets message H3 for handoff of STA, AP2 starts accepting the packets of the attacker node and STA1. Once the STA packet i.e., D2 is authenticated, the remaining process is same as that for an access point protected by KHC scheme under DoS attack. Thus, the proposed scheme enjoys enough security during the handoff.

## 6. Conclusion

In this paper, we propose a reactive handoff scheme. As the scheme is reactive, the security context (key and initial vector) is not supplied to all APs rather only one candidate AP is given STA communication key and initial vector. Thus, when STA roams into the area of candidate AP, less time is involved in the STA authentication at the candidate AP. The proposed scheme maintains security properties while evolving and transferring the security context to the candidate AP. The scheme is lightweight and provides continuous per frame authentication. All the handoff messages used in the scheme are protected. As frames are protected using MIC, frame modification is not possible. The proposed handoff scheme has low computation and communication cost (<50ms). This makes it suitable for real time scenarios with frequent handoffs. As compared to other secure handoff schemes, the proposed handoff scheme requires fewer messages, has less communication cost and is secure.

## References

Bojkovic, Z., Turán, J. and Ovseník, L. (2005), 'Towards to multimedia across wireless', *Journal of Electrical Engineering*, Vol. 56, No. 1-2, pp. 9-14.

Chien, H.Y., Hsu, T.H. and Tang, Y.L. (2008), 'Fast pre-authentication with minimized overhead and high security for WLAN handoff', *WSEAS Transaction on Computers*, Vol. 7, No. 2, pp. 46-51.

Clancy, T.C. (2008), 'Secure handover in enterprise WLANs: CAPWAP, HOKEY, and IEEE 802.11r', *IEEE Wireless Communications*, Vol. 15, No. 5, pp. 80-85.

Compton, S (2008), '802.11 denial of service attacks and mitigation', available at: http://www.sans.org/reading_room/whitepapers/wireless/80211-denial-service-attacks-mitigation_2108 (accessed 26 November 2014).

Hur, J., Park, C., Shin, Y. and Yoon, H. (2007), 'An efficient proactive key distribution scheme for fast handoff in IEEE 802.11 wireless networks', *Proceedings of the International Conference on Information Networking*, Estoril, Portugal, pp 629-638.

IEEE 802.11i (2004), 'IEEE standard for information technology-telecommunications and information exchange between systems-local and metropolitan area networks-specific requirements-part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications: amendment 6: medium access control (MAC) security enhancements', available at: https://standards.ieee.org/findstds/standard/802.11i-2004.html (accessed on 20 November 2014).

Kassab, M., Belghith, A., Bonnin, J.-M. and Sassi, S. (2005), 'Fast pre-authentication based on proactive key distribution for 802.11 infrastructure networks', *Proceedings of the Wireless Multimedia Networking and Performance Modeling*, Quebec, Canada, pp. 46-53.

Lee I. (2010), 'A novel design and implementation of Dos-resistant authentication and seamless handoff scheme for enterprise WLANs', Unpublished master thesis, University of Canterbury, Christchurch, New Zealand.

Lee, I. and Hunt, R. (2010), 'A novel design and implementation of Dos-resistant authentication and seamless handoff scheme for enterprise WLANs', *Proceedings of the 8th Australian Information Security Management Conference*, Perth, Australia, pp. 49-61.

Ling, T. C., Lee, J. F. and Hoh, K. P. (2010), 'Reducing handoff delay in WLAN using selective proactive context caching', *Malaysian Journal of Computer Science*, Vol. 23, No. 1, pp 49-59.

Maccari, L., Fantacci, R., Pecorella, T. and Frosali, F. (2006), 'Secure, fast handoff techniques for 802.1X based wireless network, Communications', *Proceedings of the IEEE International Conference on Communications*, Istanbul, Turkey, pp. 3917 - 3922.

Martinovic, I., Zdarsky, F.A., Bachorek, A. and Schmitt, J.B. (2007), 'Measurement and analysis of handover latencies in IEEE 802.11i secured networks', *Proceedings of the European Wireless Conference* (EW2007), Paris, France, pp.1-7.

Martinovic, I., Zdarsky, F. A., Bachorek, A. and Schmitt, J.B. (2006), 'Intro. of IEEE 802.11i and measuring its Sec. vs. performance tradeoff', Technical Report 351/06, Distributed Computer Systems Lab, University of Kaiserslautern, Kaiserslautern, Germany.

Mishra, A., Shin, M.H. and Arbaugh, W.A. (2004a), 'Context caching using neighbor graphs for fast handoffs in a wireless network', *Proceedings of the IEEE conference on Computer Communications*, Hong Kong, China, pp. 351-361.

Mishra, A., Shin, M.H., Petroni, N.L., Clancy, T.C. and Arbaugh, W.A. (2004b), 'Proactive key distribution using neighbor graphs', *IEEE Wireless Communications*, Vol. 11, No. 1, pp. 26-36.

Pack, S. and Choi, Y. (2004), 'Fast handoff scheme based on mobility prediction in public wireless LAN systems', *IEE Proceedings Communications*, Vol. 151, No. 5, pp. 489-495.

Pack, S., Jung, H., Kwon, T. and Choi, Y. (2005), 'A selective neighbor caching scheme for fast handoff in IEEE 802.11 wireless networks', *Proceedings of the International Conference on Communications*, Seoul, Korea, pp. 3599-3603.

Ren, K., Lee, H., Han, K., Park, J. and Kim, K. (2004), 'An enhanced lightweight authentication protocol for access control in wireless LANs', *Proceedings of the 12th IEEE International Conference on Networks*, Singapore, pp. 444-450.

Shin, M., Mishra, A. and Arbaugh, W. (2004), 'Improving the Latency of 802.11 hand-offs using neighbor graphs', *Proceedings of the 2nd International Conference on Mobile Systems, Applications and Services*, Boston, MA, pp. 70-83.

Singh, R. and Sharma, T.P. (2013a), 'A key hiding communication scheme for enhancing the wireless LAN security', *Wireless Personal Communications*, Vol. 77, No. 2, pp. 1145-1165.

Singh, R. and Sharma, T.P. (2013b), 'A secure WLAN Authentication Scheme', *IEEK Transactions on Smart Processing and Computing*, Vol. 2, No. 3, pp. 176-187.

Zhang, Z., Boukerche, A., Hussam M. and Ramadan, S. (2011), 'TEASE: a novel tunnel-based secure authentication scheme to support smooth handoff in IEEE 802.11 wireless networks', *Journal Parallel Distributed Computing*, Vol. 71, No. 7, pp. 897-905.

Zhang, Z., Pazzi, R.W. and Boukerche, A. (2010), 'Design and evaluation of a fast authentication scheme for WiFi-based wireless networks', *Proceedings of the IEEE International Symposium on World of Wireless Mobile and Multimedia Networks*, Quebec, Canada, pp. 1-6.

## About the authors

**Rajeev Singh** received his M. Tech. degree in computer science & engineering from Indian Institute of Technology, Roorkee, India in 2008 and his PhD degree from National Institute of Technology, Hamirpur, India in 2014. Currently, he is working as an assistant professor with the Department of Computer Engineering, Govind Ballabh Pant University of Agriculture & Technology, Uttarakhand, India. His research interest includes computer networks and network security.

Corresponding author. assistant professor, Department of Computer Engineering, Govind Ballabh Pant University of Agriculture & Technology, Uttarakhand, India 263145. Tel: +91-5944233338. E-mail address: rajeevpec@gmail.com

**Teek Parval Sharma** received his PhD degree from Indian Institute of Technology, Roorkee, India in 2009 in the area of wireless sensor networks. He is an associate professor at National Institute of Technology, Hamirpur, India. He has published numerous high quality research papers in international/ national journals and conferences, and has also contributed in various books of standard international publishers. His research interest includes distributed systems, wireless sensor networks, mobile Ad hoc networks, and wireless networks. E-mail address: teekparval@gmail.com

*MIS Review* Vol. 21, Nos. 1/2, September(2015)/March (2016), pp. 51-78
DOI: 10.6131/MISR.2015.2101.04


# An Analysis of Trust, Distrust, and Their Antecedents: Development of a Comprehensive Model of Consumer Intentions in Technology-Driven Transactions

*Steven John Simon, Carol J. Cagle*
*Stetson School of Business and Economics, Mercer University, Atlanta, GA, USA*

ABSTRACT: *Data breaches -- security incidents -- have become an everyday occurrence with hundreds of millions of consumers having their lost personal identification information (PII), had their credit and debit card numbers stolen, and their credit compromised. Despite the risk, consumers continuously swipe their cards and share their personal information regularly. This study examines the impacts of trust and distrust on consumer intentions in this environment. More than 1,700 consumers involved in technology-driven transactions comprise the data sample. Trust, distrust, and their antecedents are investigated to determine (1) if trust and distrust are truly two distinct constructs, (2) if the two constructs have unique antecedents, and (3) their impacts on consumer intentions toward transactions. The study expands the literature treating trust and distrust as distinct yet inter-related constructs and by introducing new antecedents. Our findings suggest that trust and distrust are not the same construct and impact consumer intentions to transact.*

KEYWORDS: *Trust, Distrust, Disposition, Data Breaches, Consumer Intentions*

## 1. Introduction

In 2014, a total of 79,790 cyber security incidents were reported in 61 countries (Brumfield, 2015). In the United States 43% of companies *reported* a data breach. An estimated $93 Billion was stolen from U.S. consumers between 2000 and 2014 (Ozawa, 2015).The widely publicized system breach at Home Depot exposed over 56 million customer records. Anthem's security lapse cost 80 million of its consumers and employees their personal information with similar losses at Ebay (145 million), JPMorgan (76 million), Court Ventures (200 million), Sony (77 million), AOL (92 million), and TJ Maxx (94 million). These attacks are not restricted to business as illustrated with the US government's Office of Personnel Management's (OPM) breach placing the security clearance and background information of over 22 million individuals at risk and the US military's loss of 70 million veteran's records. This phenomenon is not exclusive to any one country, continent, socio-economic class, or ethnic group. In South Korea, nearly 20 million people, almost 40% of the population had their personal data stolen and their credit

cards compromised (Thornhill, 2014). The magnitude of this crisis can be measured in media coverage. "Data breach" has become a term of everyday vernacular with *The New York* Times publishing more than 700 related articles in 2014, a 560% increase over 2013 (Brumfield, 2015). Despite the litany of successful cyber attacks and media coverage, individuals are continuously swiping their credit and debit cards, engaging in e-commerce transactions (many on unsecured connections), and providing a wealth of personal information to business, government, doctors, insurers, and almost anyone who asks for it. This extraordinary behavior in light of the obvious lack of security has prompted the authors to investigate technology-driven consumer behavior and transaction[1] intention as it relates to trust and distrust and their antecedents.

Interestingly, of the 70 million individuals impacted by the 2013 Target breach only 35% indicated that their trust and behavior towards Target had changed (Silver, 2014). Trust is a widely researched construct that has been linked to all economic transactions and interpersonal exchanges (Alga 2014; Gambetta, 1988; Gefen, 2004; Mayer et al., 1995). In a related technology-driven area, the role of trust in e-commerce is highly correlated with consumer intentions and purchase behavior (Kim et al., 2004, Pavlou & Gefen, 2005; Wu & Tsang 2008). Less understood and researched is distrust -- the unwillingness to be vulnerable to others (Benamati et al., 2010). Some researchers believe that distrust is a distinct construct from trust and not just ends of the same continuum (Lewicki et al., 1998).

This study seeks to answer several research questions while contributing to the body of knowledge. First, since the literature has focused almost exclusively on trust, this research develops a comprehensive model of trust and distrust, positing that trust and distrust are two distinct yet interrelated constructs that influence intentions. Few prior studies (Benamati et al., 2010; Moody et al., 2014) have attempted this with Benamati's work focused on online banking, a more restrictive domain. Second, assuming that trust and distrust are distinct constructs, they should have antecedents that influence their development. Therefore, this study examines previously researched antecedents in the context of both trust and distrust. Third, we further develop a comprehensive model with factors validated through a large sample (n = 1,763) of real-worldconsumers. This work seeks insights into trust and distrust in our dynamic tech-driven, data theft environment while determining the impact of the model's constructs on an individual's intentions.

## 2. Trust

Trust can be thought of as the glue that holds society together. It is a defining

[1] For the purposes of this work, transactions refer to any monetary exchange or the providing of any personal identification information (PII) to include credit or debit card numbers, social security numbers, medical information, etc.

feature of most economic and social transactions in which uncertainty is present (Pavlou, 2003). Trust is commonly invoked by individuals, businesses and organizations, as well as governments. Trust is mentioned in mottos, slogans, pitches, and even appears on US currency. It is a pervasive concept that has been widely studied across disciplines and yet a common definition has eluded researchers and practitioners alike.

Gambetta (1988) stated that when a trust-related topic is discussed, trust is always considered a fundamental or crucial element -- one that we cannot do without in human interactions. Trust is generally crucial in business and social interactions that are characterized by dependence of one party on another. It is a common perception, that trust is one of the key and perhaps most important factors in completing a transaction and thus of economic trade (Alga, 2014). In these transactions, trust -- in part -- binds all parties together based on the expected utility or return from the interaction (Ganesan 1994; Mayer et al., 1995). Trust has been linked to consumer confidence with consumers willing to transact with organizations they trust more than those who they do not (Keen, 2000). Gefen (2004) suggests that trust caters to a basic need to predict, understand, and control the social environment while attempting to determine the behavior of others and foresee the outcomes of actions.

Trust is the foundation of commerce (Su & Han, 2003) and is important because it helps consumers overcome perceptions of risk (McKnight et al., 2002). When conducting commerce as in all social interactions, trust is a mechanism that is employed to reduce uncertainty (Su & Han, 2003) and complexity (Gefen et al., 2003; Luhmann, 1979).Trust in online merchants has been positively associated with their attitude towards the store and intent to conduct transactions (Jarvenpaa & Tractinsky, 1999; Kim et al., 2004; Macintosh & Lockshin, 1997). Ba and Pavlou (2002) argue that trust refers to the subjective assessment of one party that another party will perform a particular transaction according to his or her confident expectations, in an environment characterized by uncertainty.Lack of trust has a negative consequence for consumers both online (Wu & Tsang, 2008) and in the physical environments. Consumers whose trust is deficient will not engage in financial transactions (Hoffman et al., 1999). In e-commerce, a number of studies (Hoffman et al., 1999, Liu et al., 2005, Pavlou & Gefen, 2005) have shown that trust is a major barrier to acceptance and inhibits Internet transactions (Kim et al., 2004).

Trust has traditionally been difficult to define (Rousseau et al., 1998) and has been regarded as a *confusing pot-pourri* (Shapiro, 1987). McKnight et al. (2002) call for conceptual clarity and quote Keen et al. (1999), " ... the basic conclusion in all these fields [is] trust is becoming more and more important, but we really cannot say what it exactly is" (pp. 4-5). The reason for this confusion (McKnight & Chervany, 2001) is that researchers have conceptualized trust within a narrow perspective in their specific field.

Economists view trust from the reputation of the parties and the impact on transactions (Cave, 2005). The key to successful economic transactions is avoiding opportunistic behavior (Williamson, 1985). Managerial and marketing researchers focus on strategies for consumers and trust building (Fogg, 2002) using trust as a mediator of the influence of a company's actions on consumer behavior (Johnson, 2007). Mayer et al. (1995) define trust as the "willingness of one party to be vulnerable to the actions of another based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" (p. 712). Human computer interaction views the relationship between the user and system usability (Riegelsberger et al., 2005), while information systems researchers investigate system adoption, acceptance, and use[2]. Lee and Turban (2001) hold that trust is the willingness of a consumer to be vulnerable to the actions of an Internet merchant in an Internet shopping transaction, based on the expectation that the Internet merchant will behave in certain agreeable ways. Sociologists investigate trust from an interpersonal and group perspective (Salovery & Rothman, 2003) with Rotter (1971) defining trust as a generalized expectancy held by an individual that the word of another can be relied upon. Zucker (1986) suggests that trust is a set of expectations shared by all involved in an exchange, which encompass social rules.

For purposes of this work, trust is not a behavior or a choice, but a psychological condition and can be defined as the willingness to be vulnerable under conditions of risk and interdependence (Bhattacharya, 2002). McKnight et al. (2002) proposed three means to measure trust -- willingness to provide personal information, willingness to engage in a purchase, and willingness to act on provided information. This study focuses on willingness to provide personal information and willingness to engage in a purchase in trusting intentions. Further McKnight et al. (1998) indicate that trusting beliefs directly influence trust intentions. This relationship between trust and intentions/behaviors has been pervasive across the literature and disciplines (Ho & Chau, 2013; Jarvenpaa & Tractinsky, 1999; Kim et al., 2004; Liu et al., 2005; Macintosh & Lockshin, 1997; Pavlou and Gefen, 2005; Wu & Tsang, 2008).

Prior research has also found that trust is a predictor of consumer behavior and directly influences trust intentions (Bhattacherjee, 2002; McKnight et al., 2002). The higher the trusting beliefs, the more likely they (consumers) are to interact (Luhmann, 1979). Recently, there have been efforts to integrate the concept of trust in technology with the technology acceptance model (Ghazizadeh et al., 2012; Pavlou, 2003) to predict a user's intent. In a study involving expectation disconfirmation theory, Lankton et al. (2014) demonstrate that technologytrusting expectations influence intentions. Trust

[2] Several prior studies have provided a meta-analysis of the trust literature (see Gefen et al., 2003; Johnson, 2007, Kim & Tadisina, 2007; Rousseau et al., 1998.).

researchers have found a strong relationship between trusting beliefs and intentions (Vance et al., 2008).

Therefore,
Hypothesis 1: Trust positively influences intentions.

## 3. Distrust

Prior research has mainly focused on trust and largely ignored distrust, partly because of the assumption that trust and distrust are two ends of one continuum (Benamati et al., 2010; Chau et al., 2013; Komiak & Benbasat, 2008; Ou & Sia, 2010; Schoorman et al., 2007; Seckler et al., 2015). Early studies (Rotter, 1971) viewed distrust as the opposite of high trust. This conceptualization of distrust as a single dipolar construct has been questioned (Lewicki et al., 1998). They state that distrust is a distinct construct from trust and that low trust is not equivalent to high distrust with the two constructs actually coexisting. To support this assumption, they extend Luhmann (1979) and suggest that trust and distrust have different consequences and develop a model which helps explain how both trust and distrust reflect the complexities and risks associated with interpersonal and business relationships. Trust focuses on more positive emotional reactions (*hope, faith, confidence, assurance*) toward others, while distrust is based in negative overtones (*fear, skepticism, cynicism, vigilance*) (Benamati et al., 2010; Lewicki et al., 1998).

Although both trust and distrust serve as mechanisms to reduce complexity and uncertainty (Kramer, 1999), distrust may exert a more critical role for consumers (Ou & Sia, 2010). In most people the desire to avoid a negative outcome is stronger than seeking a potentially more positive one (Moody et al., 2014), suggesting that distrust should provide stronger motivations and behaviors (Chau et al., 2013). Trust in the workplace has been found to foster improved working relationships and performance (Dirks & Ferrin, 2001) while distrust of institutionalized roles and structures lead to greater negative consequences (Sitkin & Roth, 1993).

Akin to trust, distrust simplifies an individual's decision-making process by determining which high risk or undesirable outcomes should be avoided. Trust reduces complexity by enabling individuals to take actions that expose them to risk while distrust reduces complexity, encouraging individuals to take protective actions to reduce risk (Benamati et al., 2010). In other words, trust and distrust balance each other leading a decision-maker to a state of equilibrium and potential action. Trust without distrust might lead to a consumer who fails to take full account of the risks associated with a decision. In the context of this work, a consumer that freely provides personal identification information (PII) on all occasions might find their identity compromised. This is not to

suggest that trust is good and distrust is bad, a simplistic view that has been pervasive in organizational and social research (Lewicki et al., 1998). Distrust can be thought to represent caution before or after taking an action. A famous example is Ronald Reagan's quote "Trust but Verify," during arms negotiations. While distrust is not explicitly mentioned it is implied and in this context not viewed as negative or bad but merely responsible and a means to reduce uncertainty which in turn led to an agreement. It has been further suggested (Benamati et al., 2010) that trust and distrust need not vary simultaneously. Lewicki et al. (1998) state that it is possible to like and dislike, to love and hate, and may be possible to trust and distrust. In this case, one might trust a company and its products but have minimal trust in their IT support systems as a result of a data breach or failure to protect personal information.

If distrust is a distinct construct from trust, then focusing exclusively on trust may explain only part or provide a bias estimation of behavior (Benamati et al., 2010; McKnight et al., 2004; Ou & Sia, 2010). This study therefore assumes that trust and distrust exist as separate yet related constructs allowing the authors to study the constructs independently and interdependently as they relate to intentions and behavior. This study follows prior definitions of distrust, the negative expectations regarding an action, "the positive expectation of injurious action" (Luhmann, 1979, p. 72), or "fear of, a propensity to attribute sinister intentions to, and a desire to buffer oneself from the effects of another's conduct" (Lewicki et al., 1998, p. 439).

In the context of consumer transactions distrust could lead to higher uncertainty on the part of the consumer and thus negatively impacting future transaction.The distrust is one of the most frequently cited reasons for consumers not usingmobile banking (Lin, 2011). Although both trust and distrust serve as mechanisms to reduce complexity and uncertainty (Kramer, 1999), distrust may exert a more critical role for consumers (Ou & Sia, 2010). In most people the desire to avoid a negative outcome is stronger than seeking a potentially more positive one (Moody et al., 2014), suggesting that distrust should provide stronger motivations and behaviors (Chau et al., 2013).

Therefore:

Hypothesis 2: Distrust negatively influences intentions.

## 4. Antecedents to trust and distrust

### *4.1 Disposition to trust or distrust*

"Disposition to trust is a general not situation specific, inclination to display faith in humanity and to adopt a trusting stance toward others" (Gefen, 2000, p. 728). Disposition

to trust does not imply that others are trustworthy, only that they are more willing to depend on others (McKnight & Chervany, 2001). Conlon and Mayer (1994) found the willingness to trust was significantly related to behavior and performance.

Disposition to trust has an impact on the formation of trust, especially when consumers have insufficient information or are in unfamiliar or abnormal situations (Gefen, 2004; Zhou & Tian, 2010). The concepts have been linked to faith in humanity and the assumptions of people and organizations in general. In contrast, disposition to distrust is also a persistent view that a person holds across situations, irrespective of the others involved (McKnight & Chervany, 2001; Moody et al., 2014; Zhaou & Tian, 2010). Thisconcept implies a general unwillingness to depend on or become vulnerable to others (McKnight et al., 2004).

Both dispositional trust and distrust develop over a lifetime as a result of learned outcomes from varied experiences (McKnight et al., 2004; Merritt et al., 2013; Rotter, 1971). Furthermore, both constructs are thought to be relatively stable, although not static (Mayer et al., 1995; Merritt and Ilgen, 2008) and may change as individuals experience both positive and negative experiences. The degree or severity of an experience could yield a greater impact. Therefore, we expect that as with trust and distrust, disposition to trust and distrust may not vary simultaneously. A person with high disposition trust is more likely to trust others than a person with low dispositional trust, while an individual with high dispositional distrust is likely to be more distrustful.

Therefore:
Hypothesis 3: Disposition to trust positively influences trust.
Hypothesis 4: Disposition to distrust positively influences distrust.

### *4.2 Reputation*

Company reputation reflects the amount of regard that stakeholders, particularly customers, assign to the company (Fombrun & Rindova, 2000). The literature recognizes that reputation plays a critical and primary role in building productive customer relations (Abimbola & Vallester, 2007; Garbarino & Johnson, 1999). Fombrun (2005) noted that customers judge companies constantly, adding that reputation is widely seen as a powerful intangible corporate asset, which the leaders of well-respected companies actively measure and work to enhance.

Prior research on cognitive trust suggests that a trustor may categorize an unfamiliar trustee astrustworthy or untrustworthy based on the reputation of the trustee. The reputation categorization processinfers that a trustee with a good reputation is believed to be trustworthy (McKnight et al., 1998). When nodirect experiential information is available, the trustee's reputation may affect people's beliefs (Powell, 1996). Thus, reputation isconsidered an important moderator of trust.

Keh and Xie (2009) examined the relationship of corporate reputation to trust and purchase intent in Chinese companies. Their findings showed that reputation was very strongly related to trustworthinessand with a positive link to the customers' intent to purchase. The strength of the relationship suggests that it is a major component of overall corporate reputation. Consequently, higher levels of trust can exercise a significant and favorable impact on customer behavior. Reputation has been categorized as a factor through which individuals build cognitive trust (McKnight et al., 1998) and view an organization as trustworthy (Jarvenpaa et al., 2000; Kim, 2012). An established reputation has been linked to integrity and ability – an organization that is capable and insures its products and services (Liu et al., 2005; Mayer et al., 1995).

Therefore:
Hypothesis 5a: Reputation positively influences trust.
Hypothesis 5b: Reputation negatively influences distrust.

### *4.3 Knowledge*

Knowledge has been recognized as one of the most important cognitive factors influencing behavioral processes (Jeng & Fesenmaier, 2002; Vogt & Fesenmaier, 1998) and consumer behavior (Klink & Smith, 2001). Selnes and Howell (1999) found that decision-making behavior and information processing differ between customers with high and low levels of knowledge. Knowledge plays a central role in many theoretical models of attitude because it is hypothesized to influence behavior (Barber et al., 2009). Hadar et al. (2013) investigated the impact of subjective knowledge states on financial decision-making with findings showing that when customers felt less confident in their knowledge, they were less likely to make a risky investment. With high levels of knowledge, consumers are confident in their own ability to undertake information-searching tasks (Schmidt & Spreng, 1996) with results leading to higher levels of trust, which reinforce their purchase behavior (Alba & Hutchinson, 1987; Crowley & Mitchell 2003).

Knowledge is a multidimensional construct comprising three categories: (1) subjective knowledge -- familiarity, (2) objective knowledge -- expertise, and (3) product experience -- possession (Alba & Hutchinson, 1987; Bettman & Park 1980; Brucks, 1985; Johnson & Russo, 1984; Park & Moon, 2003; Ratchford, 2001; Sujan, 1985). Hence, subjective knowledge equates to the familiarly with the laws and structural procedures in place to safeguard their information, objective knowledge is expertise but focused on the understanding of the nature of the technology related to information transmission and protection, while experience relates to the conduct of the transaction gained through use. This study incorporates knowledge that would protect an individual from the dangers present in technology-driven transactions.

Therefore:
Hypothesis 6a: Knowledgenegatively influences trust.
Hypothesis 6b: Knowledgepositively influences distrust.

### *4.4 Technology trust*

It has been suggested (Benamati et al., 2010) that trust and distrust need not vary simultaneously. Lewicki et al. (1998) state that it is possible to like and dislike, to love and hate, and may be possible to trust and distrust. Therefore, one might trust a company and its products but distrust its IT support systems as a result of a data breach or failure to protect personal information. McKnight et al. (2011) explained that trust in technology relates to individuals depending on, or beingwilling to depend on the technology to accomplish a specific task because the technology has positivecharacteristics.

The relationship between trust in an organization's technology and the organization itself has been suggested by Evenstad (2016). An individual's trust in technology may also influence his/her trust in other elements of an organization such as institutional trust (Muir, 1994). Higher trust in technology leads to higher trust culture (in the organization) and greater adoption and use (Xu et al., 2014). In a medical setting, Montague et al. (2009) uncovered a link between trust in an organization's technology and social trust (trust in the organization itself). Distrust in technology prevents the user from utilizing systems to their full extent, and can lead to a decrease in productivity (Xu et al., 2014). The lack of trust is one of the most frequently cited reasons for consumers not using mobile banking (Lin, 2011; Masrek et al., 2014). Technology trust has been empirically supported as a moderator of trusting beliefs and intentions (Li et al., 2008; Vance et al., 2008).

Therefore:
Hypothesis 7a: Technology trust positively influences trust.
Hypothesis 7b: Technology trust negatively influences distrust.

## 5. Methodology

### *5.1 Instrument*

The questionnaire was derived from pervious literature (see Appendix) with a number of questions rewritten to make them more applicable to this study's intent. The initial instrument consisted of just under a hundred (7 point scale, Likert-type) items including demographic information. The instrument was initially pilot tested with a group of approximately 250 individuals drawn from a population representing the desired sample. As a result of the pilot test, the analysis of the data, and feedback from subjects, the instrument was reduced by greater than one half with redundant items eliminated and

the wording of several questions modified. The resulting instrument was then retested and after very minor changes, a final instrument was made available. The first page of the questionnaire explained to participants the intent and objectives of the study and provided them with some contextual background for the work, while asking them to answer the questions with regard to first hand interaction with a specific organization. It further explained that the research was for academic purposes only and that any information provided would be held in strictest confidence.

### *5.2 Participants*

The authors surveyed a large global sample (2,000+) of consumers to understand the impacts of data breaches. Participants were recruited into the study utilizing business and institutional contacts of the researchers and received no compensation. The only conditions for participation was that individuals must be over the age of 18, be personally responsible for a credit card or online pay account (e.g., PayPal or direct phone billing), and have either completed (1) an Internet transaction, (2) used a credit card or online pay account at any merchant -- on-line or otherwise, or (3) provided personal information to any organization (business or otherwise) in the preceding six months. Participants were screened to insure they met the previously stated requirements of the study. Participants were informed that when answering the study's questions they should consider their interaction with any organization through which they either conducted transactions or provided personal identification information regardless of if that organization had suffered a breach or not. Surveys were taken and collected via a controlled access website that randomized all sample questions. While not excluding students, this study focused on participants from the general population. The sample was not restricted to any group or country although the majority of responses (approximately 79%) came from the United States. For this sampling, 1,763 fully completed and useable surveys were obtained.

The sample was composed of 653 (37%) females and 1110 (63%) males with a mean age of 36.8 years (standard deviation 12.6). Participants had the following education make-up -- 1.2% no high school degree, 6.9% high school completion, 19.2% some college work, 12.5% associates degree, 39.2% bachelor degree, and 21% graduate degree. The questionnaire also collected income demographics with 43.9% having income of less than $75,000 per year, 38.4% between $75,000 & $150,000, 9.7% between $150,000 & $200,000, and 8.1% in excess of $200,000.

## 6. Research analysis

Analysis and data validation were conducted in phases following recommended research guidelines (Cenfetelli & Bassellier, 2009; Hair et al., 2013; Moody et al., 2014).

First, data were examined to determine the extent of multicollinearity, reliability, or common-method bias. The data were found to be free of missing data for all observed items with the exception of control variables gender, age, highest level of education, ethnicity, and income.

The second phase of the analysis proceeded by establishing factorial validity using exploratory factor analysis (EFA). EFA was conducted using Unweighted Least Squares (ULS) with oblimin rotation used to extract the maximum shared variance, leaving unique and error variance in the model (Osborne & Banjanovi, 2016). ULS was considered to be more robust in the presence of violations of multivariate normality. The assumption that the manifest variables were correlated suggested that an oblimin roation was appropriate. Initial analysis indicated an 8 factor model with 32 variables after examining the scree plot and data pattern matrices for the full sample.

The next phase of the analysis examined the model using path analysis. The path analysis was used to test the theoretical model to determine directional relationships and assess the overall viability of the model. After examining the final path analysis results, all paths were retained. Additionally, the Comparative Fit Index (CFI) (0.85), Standardized Root Mean Square Residual (SRMR) (0.06), Root Mean Square Error of Approximation (RMSEA) (0.167) and the full RMSEA 90% confidence interval were within acceptable limits, ($0.153 \leq \text{RMSEA} \leq 0.181$).

Continuing with the analysis, the theoretical model was examined using Confirmatory Factor Analysis (CFA). Using the approach suggested by Anderson and Gerbing (1988), CFA was used to provide evidence that the indicator variables effectively measured the underlying constructs of the measurement model. Initial goodness-of-fit statistics were less than favorable. After examining results and Lagrange Multipliers (LM), two manifest variables were removed, retaining 30 variables. The CFI (0.94), GFI (0.89), RMSEA (0.064), and RMSEA 90% confidence interval ($0.062 \leq \text{RMSEA} \leq 0.066$), results of the CFA indicated an acceptable fit.

The final step,assessing the structural model, indicated that the model identified five exogeneous variables, Reputation, Knowledge, Technology Trust, Disposition to Trust and Disposition to Distrust and two endogeneous variables that moderated levels of Trusting and Distrusting Beliefs affecting levels of Intent. All reported godness-of-fit values were withing acceptable ranges based on research (Hair et al., 2013). RMSEA (0.068), RMSEA 90% confidence interval ($0.066 \leq \text{RMSEA} \leq 0.07$), and SRMR (0.058) support the conclusion that the model did a reasonable job of accounting for the covariance in the data. Additionally, 27% of the variance in Intentions is explained by Trusting and Distrusting Beliefs, 63% of the variance in Trusting Beliefs is explained by Disposition to Trust, Reputation, Knowledge, and Technology Trust, and 48% of the variance in

Distrusting Beliefs is explained by Disposition to Distrust, Reputation, Knowledge, and Technology Trust.

### 6.1 Measurement model analysis

Common-method bias is a measurement error attributable to systematic error in the measurement method observed most often in research studies involving self-reported measures in measurement instruments. The Pearson correlation coefficient matrix for manifest variables was examined for mono-method bias and reliability (Podsakoff et al., 2003). All correlation coefficients were between -0.63 and 0.89 suggesting a lack of mono-method bias (See Table 1). Research suggests (Cenfetelli and Bassellier, 2009; Hair et al., 2013) that indicators with VIF values less than 10 are generally acceptable. The model's VIF values ranged from 1.919 to 6.720.Cronbach's Alphas ranged between 0.837 and 0.934 with only the indicators for Knowledge being below 0.9. The literature suggests that Cronbach's alphas greater than 0.70 are acceptable and values greater than 0.80 are ideal (Nunnally & Bernstein, 1994) (See Table 2). Statistical power is estimated at 0.99 for this model.

### 6.2 Convergent validity

Research suggests that convergent validity is established using three measures (Hair et al., 2013). The first indicator of convergent validity is the size of the factor loadings. All standardized factor loadings ranged between 0.5943 and 0.9306. With the exception of one variable, all loading exceeded 0.61. The second indicator of convergent validity is the average variance extracted (AVE) for each item. AVE ranged between 0.504 and 0.766 (values exceeding 0.50 suggest adequate convergence). The third indicator of convergence uses reliability estimates where values greater than 0.6 are acceptable, i.e. internal consistency. Reliability estimates range between 0.604 and 0.915 and consistently represent the same latent construct. Therefore, there is empirical evidence to suggest that the model achieves convergent validity (See Table 2).

### 6.3 Discriminant validity

Discriminant validity is achieved when the variables that should not be related are not. This is evidenced in exploratory factor analysis by variables *hanging* together without cross loading on other factors. The suggested approach for determining discriminant validity is to compare the AVE to the square of the correlation estimate for each latent construct (Hair et al., 2013). Empirically, AVE estimates suggest that the model violates this condition. However, the variables exhibit high loadings on each latent construct and exhibited no cross loading below the 0.55 level.

**Table 1** Pearson Correlation Coefficients for Manifest Variables

| | I1 | I2 | I3 | TR1 | TR2 | TR3 | TR4 | TR5 | DB1 | DB2 | DB3 | DB4 | DB5 | DT1 | DT2 | DT3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I1 | 1.00 | | | | | | | | | | | | | | | |
| I2 | 0.88 | 1.00 | | | | | | | | | | | | | | |
| I3 | 0.79 | 0.80 | 1.00 | | | | | | | | | | | | | |
| TR1 | 0.44 | 0.44 | 0.47 | 1.00 | | | | | | | | | | | | |
| TR2 | 0.41 | 0.43 | 0.42 | 0.81 | 1.00 | | | | | | | | | | | |
| TR3 | 0.42 | 0.42 | 0.46 | 0.81 | 0.79 | 1.00 | | | | | | | | | | |
| TR4 | 0.38 | 0.38 | 0.44 | 0.70 | 0.67 | 0.74 | 1.00 | | | | | | | | | |
| TR5 | 0.46 | 0.46 | 0.50 | 0.78 | 0.73 | 0.81 | 0.73 | 1.00 | | | | | | | | |
| DB1 | -0.35 | -0.37 | -0.35 | -0.53 | -0.49 | -0.54 | -0.41 | -0.55 | 1.00 | | | | | | | |
| DB2 | -0.33 | -0.37 | -0.33 | -0.50 | -0.47 | -0.52 | -0.43 | -0.54 | 0.90 | 1.00 | | | | | | |
| DB3 | -0.36 | -0.37 | -0.38 | -0.54 | -0.49 | -0.57 | -0.50 | -0.63 | 0.83 | 0.85 | 1.00 | | | | | |
| DB4 | -0.34 | -0.37 | -0.37 | -0.52 | -0.46 | -0.52 | -0.49 | -0.59 | 0.78 | 0.81 | 0.86 | 1.00 | | | | |
| DB5 | -0.36 | -0.39 | -0.37 | -0.57 | -0.51 | -0.58 | -0.46 | -0.63 | 0.82 | 0.82 | 0.84 | 0.83 | 1.00 | | | |
| DT1 | 0.17 | 0.17 | 0.18 | 0.35 | 0.29 | 0.33 | 0.32 | 0.33 | -0.18 | -0.18 | -0.16 | -0.21 | -0.21 | 1.00 | | |
| DT2 | 0.19 | 0.19 | 0.19 | 0.34 | 0.31 | 0.32 | 0.29 | 0.33 | -0.09 | -0.06 | -0.10 | -0.14 | -0.14 | 0.68 | 1.00 | |
| DT3 | 0.18 | 0.17 | 0.17 | 0.36 | 0.37 | 0.37 | 0.35 | 0.37 | -0.22 | -0.18 | -0.16 | -0.17 | -0.22 | 0.66 | 0.61 | 1.00 |

**Table 1** Pearson Correlation Coefficients for Manifest Variables (continued)

| | DDT1 | DDT2 | DDT3 | DDT4 | DDT5 | REP1 | REP2 | REP3 | REP4 | K1 | K2 | K3 | K4 | TT1 | TT2 | TT3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DDT1 | 1.00 | | | | | | | | | | | | | | | |
| DDT2 | 0.72 | 1.00 | | | | | | | | | | | | | | |
| DDT3 | 0.53 | 0.69 | 1.00 | | | | | | | | | | | | | |
| DDT4 | 0.61 | 0.69 | 0.69 | 1.00 | | | | | | | | | | | | |
| DDT5 | 0.58 | 0.60 | 0.55 | 0.72 | 1.00 | | | | | | | | | | | |
| REP1 | 0.01 | 0.00 | -0.01 | 0.00 | 0.00 | 1.00 | | | | | | | | | | |
| REP2 | 0.01 | -0.03 | -0.03 | -0.02 | -0.02 | 0.79 | 1.00 | | | | | | | | | |
| REP3 | 0.02 | 0.01 | -0.03 | 0.00 | 0.02 | 0.78 | 0.81 | 1.00 | | | | | | | | |
| REP4 | 0.02 | 0.01 | -0.01 | 0.00 | 0.04 | 0.71 | 0.76 | 0.81 | 1.00 | | | | | | | |
| K1 | 0.26 | 0.22 | 0.20 | 0.25 | 0.20 | 0.11 | 0.07 | 0.11 | 0.09 | 1.00 | | | | | | |
| K2 | 0.30 | 0.23 | 0.18 | 0.24 | 0.21 | 0.13 | 0.12 | 0.12 | 0.10 | 0.65 | 1.00 | | | | | |
| K3 | 0.31 | 0.27 | 0.27 | 0.30 | 0.23 | 0.16 | 0.15 | 0.15 | 0.15 | 0.54 | 0.57 | 1.00 | | | | |
| K4 | 0.36 | 0.29 | 0.23 | 0.31 | 0.32 | 0.22 | 0.18 | 0.21 | 0.19 | 0.52 | 0.53 | 0.56 | 1.00 | | | |
| TT1 | -0.05 | -0.05 | -0.06 | -0.05 | -0.08 | 0.27 | 0.25 | 0.28 | 0.30 | 0.10 | 0.07 | 0.08 | 0.06 | 1.00 | | |
| TT2 | -0.07 | -0.09 | -0.06 | -0.04 | -0.09 | 0.25 | 0.24 | 0.27 | 0.27 | 0.18 | 0.11 | 0.08 | 0.12 | 0.78 | 1.00 | |
| TT3 | -0.01 | -0.03 | -0.02 | 0.02 | -0.04 | 0.30 | 0.29 | 0.30 | 0.32 | 0.21 | 0.14 | 0.10 | 0.16 | 0.71 | 0.80 | 1.00 |

**Table 2** Factor Loadings, Average Variance Extracted (AVE), and Cronbach Alphas

| Constructs and Indicators | Standardized Loading $L_i$ | Indicator Reliability $L_i^2$ | Error Variance $1 - L_i^2$ | Composite Reliability | Variance Extracted (AVE) | Squared Multiple Correlations of the Variables with Each Factor | Cronbach Coefficient Alpha |
|---|---|---|---|---|---|---|---|
| Intentions (F1) | | | | | 0.7655 | 0.9463 | 0.9331 |
| V1 | 0.9069 | 0.8225 | 0.1775 | 0.9069 | | | |
| V2 | 0.9306 | 0.8660 | 0.1340 | 0.8476 | | | |
| V3 | 0.7798 | 0.6081 | 0.3919 | 0.6692 | | | |
| Trust_Beliefs (F2) | | | | | 0.5201 | 0.9302 | 0.9003 |
| V4 | 0.6347 | 0.4029 | 0.5971 | 0.7474 | | | |
| V5 | 0.7833 | 0.6136 | 0.3864 | 0.7056 | | | |
| V6 | 0.6918 | 0.4786 | 0.5214 | 0.6939 | | | |
| Distrust_Beliefs (F3) | | | | | 0.5748 | 0.8928 | 0.9622 |
| V7 | 0.7650 | 0.5853 | 0.4147 | 0.8263 | | | |
| V8 | 0.8220 | 0.6757 | 0.3243 | 0.8166 | | | |
| V9 | 0.7616 | 0.5800 | 0.4200 | 0.7771 | | | |
| V10 | 0.7338 | 0.5385 | 0.4615 | 0.6811 | | | |
| V11 | 0.7033 | 0.4947 | 0.5053 | 0.7427 | | | |
| Disp_Trust (F4) | | | | | 0.6166 | 0.9455 | 0.8472 |
| V12 | 0.8307 | 0.6901 | 0.3099 | 0.8278 | | | |
| V13 | 0.7970 | 0.6352 | 0.3648 | 0.7336 | | | |
| V14 | 0.7242 | 0.5244 | 0.4756 | 0.6076 | | | |
| Disp_Distrust (F5) | | | | | 0.5036 | 0.9279 | 0.8988 |
| V15 | 0.5942 | 0.3531 | 0.6469 | 0.7283 | | | |
| V16 | 0.7814 | 0.6106 | 0.3894 | 0.7901 | | | |
| V17 | 0.6792 | 0.4613 | 0.5387 | 0.7627 | | | |
| V18 | 0.7748 | 0.6002 | 0.3998 | 0.7062 | | | |
| V19 | 0.7019 | 0.4926 | 0.5074 | 0.7157 | | | |
| Reputation (F6) | | | | | 0.6703 | 0.8462 | 0.9337 |
| V20 | 0.7896 | 0.6235 | 0.3765 | 0.8744 | | | |
| V21 | 0.8577 | 0.7356 | 0.2644 | 0.8672 | | | |
| V22 | 0.8593 | 0.7383 | 0.2617 | 0.7954 | | | |
| V23 | 0.7641 | 0.5838 | 0.4162 | 0.7459 | | | |
| Knowledge (F7) | | | | | 0.5231 | 0.8565 | 0.8374 |
| V24 | 0.7782 | 0.6056 | 0.3944 | 0.8001 | | | |
| V25 | 0.7997 | 0.6395 | 0.3605 | 0.7444 | | | |
| V26 | 0.6874 | 0.4725 | 0.5275 | 0.5944 | | | |
| V27 | 0.6124 | 0.3750 | 0.6250 | 0.6348 | | | |
| Tech_Trust (F8) | | | | | 0.6977 | 0.8901 | 0.9071 |
| V28 | 0.7483 | 0.5600 | 0.4400 | 0.8728 | | | |
| V29 | 0.9357 | 0.8755 | 0.1245 | 0.8673 | | | |
| V30 | 0.8109 | 0.6576 | 0.3424 | 0.6576 | | | |

### 6.4 Structural model analysis

The theoretical model identifies three first order constructs, Reputation (F6), Knowledge (F7), and Tech Trust (F8) as well as two second order constructs, Disposition to Trust (F4) and Disposition to Distrust (F5). Three other constructs were identified, Trusting Beliefs (F2), Distrusting Beliefs (F3), and Intentions (F1). Intentions, Trusting Beliefs and Distrusting Beliefs are second order variables. Goodness-of-fit indices for the SEM are within acceptable limits. All paths linking the latent constructs were determined to be in the appropriate direction and significant at $p < 0.0001$. The standardized path coefficients from Trusting Beliefs (F2) to Intentions (F1) is 0.423, t = 18.58, $p < 0.0001$) from Distrusting Beliefs (F3) to Intentions (F1) is -0.182, t = -7.667, $p \leq 0.0001$.

## 7. Results

The theoretical model (see Figure 1) exhibits the constructs and their respective relationship magnitude and direction. The higher trusting beliefs, the more likely consumers' intentions will be manifested by their intentions. Prior research suggests that

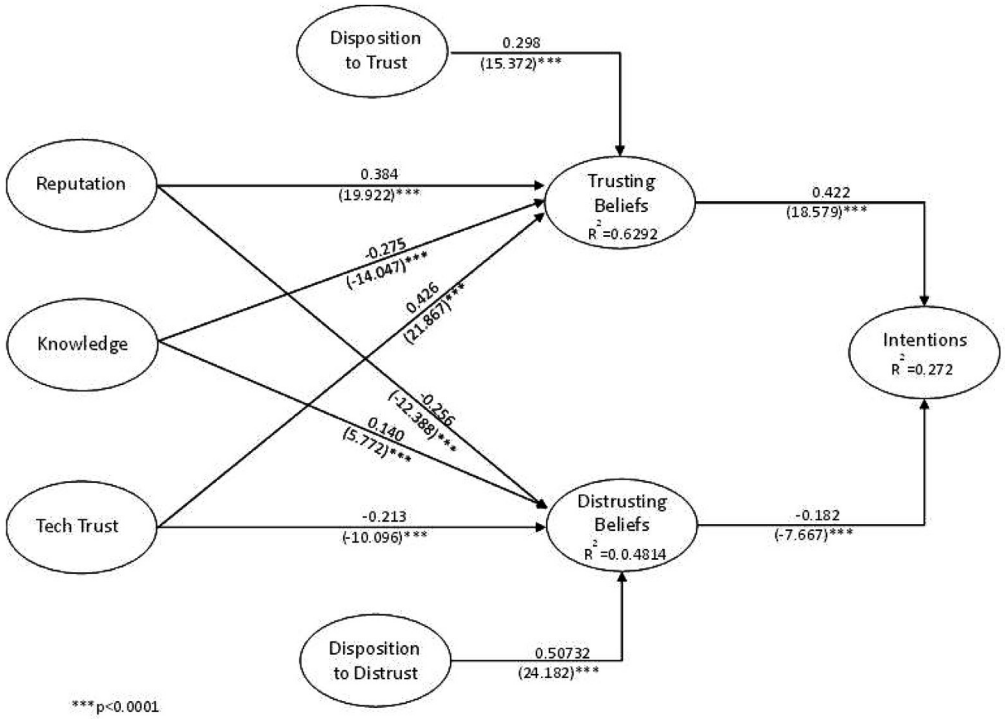



**Figure 1** Comprehensive Model

Note. $^{***}p < 0.0001$

there is a strong relationship between consumers' trusting beliefs and their intentions. The theoretical model portraying the relationship between trusting beliefs and intentions confirms this relationship. The research shows that there is a strong, positive (0.422, $p < 0.0001$) relationship between trusting beliefs and intentions. Therefore, Hypothesis 1 is supported. The relationship between distrusting beliefs and intentions is negatively related (-0.180, $p < 0.000$). This supports prior research findings that distrusting beliefs are a conceptually separate construct from trusting beliefs. Therefore, Hypothesis 2 is supported.

Hypotheses 3 and 4 for disposition to trust and disposition to distrust were both supported (0.290, $p < 0.0001$ and 0.510, $p < 0.0001$). As antecedents, disposition to distrust appears to have a stronger relationship to distrusting beliefs than disposition to trust with trusting beliefs. This finding is not surprising given that trusting beliefs and distrusting beliefs are conceptually different concepts. Reputation indicated a strong positive relationship (0.0.390, $p < 0.0001$) with trusting beliefs and and negative relationship with distrusting beliefs (0.140, $p < 0.0001$). Prior research indicates that reputation often plays a role in people's beliefs. Therefore, the theoretical model supports both hypotheses 5a and 5b.

Prior research suggests that knowledge is an important antecedent in many theoretical models of attitude due to its direct and indirect effects on behavior. Knowledge has a negative relationship (-0.280, $p < 0.0001$) with trusting beliefs and a somewhat weaker, positive relationship (0.140. $p < 0.0001$) with distrusting beliefs. Knowledge that would protect an individual from technology-driven transactions would strengthen this argument. The model supports hypotheses 6a and 6b.Technology trust exhibits a strong, positive relationship with trusting beliefs (0.430, $p < 0.0001$) and a negative relationship with distrusting beliefs (-0.220, $p < 0.0001$) suggesting that as a consumer becomes more dependent on technology to complete specific tasks or transactions, the stronger the trusting beliefs. On the other hand, a lack of trust in an organization's technology suggests that distrusting beliefs or avoidance of such technology. Hypotheses 7a and 7b are supported in the model.

## 8. Discussion and conclusion

This research provides interesting insights into an academic discussion -- the relationship of trust and distrust -- as well as understanding of consumer behavior in data theft environments, a new phenomenon. This study and was undertaken in part because both researchers had their information compromised during data breaches. Trust and

distrust were historically believed to be opposite ends of the same continuum. The belief that they are distinct constructs with unique antecedents was introduced by Lewiciki et al. (1998) and is still widely argued. The researchers recognize that one study neither confirms nor rejects theory but while using a large sample of real-world participants, as opposed to students, the empirical findings support the concept that trust and distrust are distinct constructs and those constructs directly impact the intentions of consumers.

While shedding light on a new phenomenon, the study created as many questions as it provided answers. Since this study was a snapshot in time a follow-up study of a longitudinal nature of trust/distrust would be very interesting to understand the lasting impacts of data breaches on trust/distrust as well as the impact on consumer intentions. Additionally, this study did not explore the degree of impact, only if the participants had been directly exposed to a data breach. Clearly the magnitude of the incident should create varied reactions by consumers, leading to stronger changes in intentions. A future study of the degree of loss and its impact would provide an interesting extension. Further, while not emphasized in this study, our sample was drawn from a global population (although predominately USA). We did not differentiate on national or cultural origin but it would be interesting to understand any cultural implications that future studies might derive.

The subject material of this research is of interest to both academics and practitioners. Future studies from an academic standpoint should extend this work while attempting to understand the relationships amongst trust and distrust and their antecedents. Those relationships would further clarify the nature of trust and distrust. For practitioners, future studies of the relationships -- amongst trust, distrust, their antecedents, and intentions -- and their strengths could assist in the understanding of how changes in one factor impact a consumer's behaviors and intentions. This in turn could provide organizations remedies that go beyond 'locking down' a system. For instance, strong relationships were discovered between reputation and trust and distrust. Assuming that these relationships are generalizable, organizations could enhance their reputation which in turn could provide benefits if they suffer as data breach. The same can be inferred for knowledge and technology trust.

All investigations and reports suggest that data breaches and loss of data will continue to increase and that the magnitude of the losses will expand correspondingly. This exploratory study examined the impact of data breaches/hacks on trust/distrust, their antecedents, and intentions. This research demonstrates that, within this context, trust and distrust are distinct constructs with distinct antecedents.

## References

Abimbola, T. and Vallaster, C. (2007), 'Brand, organization identity and reputation: SMEs as expressive organizations', *Qualitative Market Research: An International Journal*, Vol. 10, No. 4, pp. 416- 430.

Alba, J.W. and Hutchinson, J.W. (1987), 'Dimensions of consumer expertise', *Journal of Consumer Research*, Vol. 13, No. 4, pp. 411-454.

Anderson, J.C. and Gerbing, D.W. (1988), 'Structural equation modeling in practice: a review and recommended two-step approach', *Psychology Bulletin*, Vol. 103, No. 3, pp. 411-423.

Ba, S. and Pavlou, P.A. (2002), 'Evidence of the effect of trust building technology in electronic markets: price premiums and buyer behavior', *MIS Quarterly*, Vol. 26, No. 3, pp. 243-268.

Barber, N., Taylor, C. and Strick, S. (2009), 'Wine consumers' environmental knowledge and attitudes: influence on willingness to purchase', *International Journal of Wine Research*, Vol. 2009, No. 1, pp. 59-72.

Benamati, J., Serva, M.A. and Fuller, M.A. (2010), 'The productive tension of trust and distrust: the coexistence and relative role of trust and distrust in online banking', *Journal of Organizational Computing and Electronic Commerce*, Vol. 20, No. 4, pp. 328-346.

Bettman, J.R. and Park, C.W. (1980), 'Effects of prior knowledge and experience and phase of the choice process on consumer decision processes: a protocol analysis', *Journal of Consumer Research*, Vol. 7, No. 3, pp. 234-248.

Bhattacherjee, A. (2002), 'Individual trust in online firms: scale development and initial test,' *Journal of Management Information Systems*, Vol. 19, No. 1, pp. 211-241.

Brucks, M. (1985), 'The effects of product class knowledge on information search behavior', *Journal of Consumer Research*, Vol. 12, No. 1, pp. 1-16.

Brumfield, J. (2015), 'Verizon 2015 data breach investigations report', avalible at: http://www.verizon.com/about/news/2015-data-breach-report-info (accessed on Feb 23, 2016).

Cave, J. (2005), 'The economics of cyber trust between cyber partners', in Mansell, R. and Collins, B.S. (Eds.), *Trust and Crime in Information Societies*, Edward Elgar, Chelteham, UK, pp. 380-427.

Cenfetelli, R.T. and Bassellier, G. (2009), 'Interpretation of formative measurement in information systems research', *MIS Quarterly*, Vol. 33, No. 4, pp. 689-707.

Chau, P.Y.K., Ho, S.Y., Ho, K.K. and Yao, Y. (2013), 'Examining the effects of malfunctioning personalized services on online users', *Decision Support Systems*, Vol. 56, pp. 180-191.

Colesca, S.E. (2009), 'Understanding trust in e-government', *Economics of Engineering Decisions*, Vol. 3, pp. 7-15.

Conlon, E.J. and Mayer, R.C. (1994), 'The effect of trust on principal-agent dyads: an empirical investigation of stewardship and agency', *Academy of Management Meeting*, Dallas, TX.

Cowley, E. and Mitchell, A.A. (2003), 'The moderating effect of product knowledge on the learning and organization of product information', *Journal of Consumer Research*, Vol. 30, No. 3, pp. 443-454.

Dirks, K.T. and Ferrin, D.L. (2001), 'The role of trust in organizational settings', *Organizational Science*, Vol. 12, No. 4, pp. 450-467.

Doney, P.M. and Cannon, J.P. (1997), 'An examination of the nature of trust in buyer-seller relationships', *Journal of Marketing*, Vol. 61, No. 2, pp. 35-51.

Evenstad, L. (July 12-18, 2016), 'The importance of trust and ethics in a digital world', *Computer weekly*, pp. 23-28.

Fogg, B. (2002), *Persuasive Technology: Using Computers to Change What We Think and Do,* Morgan Kaufmann, Boston, MA.

Fombrun, C. and Rindova, V.P. (2000), 'The road to transparency: reputation management at royal Dutch/Shell', in Schultz, M., Hatch, M.J. and Larsen, M.H. (Eds.), *The Expressive Organization: Linking Identity, Reputation, and the Corporate Brand*, Oxford University Press, Oxford, UK, pp. 233-258.

Fombrun, C. (2005), 'The leadership challenge of building resilient corporate reputations', in Doh, J.P. and Stumpf, S.A. (Eds.), *Handbook on Responsible Leadership and Governance in Global Business*, Edward Elgar, Northampton, MA, pp. 54-70.

Gambetta, D.G. (1988), 'Can we trust trust?', in Gambetta, D (Ed.), *Trust: Making and Breaking Cooperative Relations*, Blackwell, New York, NY, pp. 213-237.

Ganesan, S. (1994), 'Determinants of long-term orientation in buyer-seller relationships', *Journal of Marketing*, Vol. 58, No. 2, pp. 1-19.

Garbarino, E. and Johnson, M.S. (1999), 'The different roles of satisfaction, trust and commitment in customer relationships', *Journal of Marketing*, Vol. 63, No. 2, pp. 70-87.

Gefen, D. (2000), 'E-commerce: the role of familiarity and trust', *Omega*, Vol. 28, No. 6, pp. 725-737.

Gefen, D., Karahanna, E., and Straub, D.W. (2003), 'Trust and TAM in online shopping: an integrated model', *MIS Quarterly*, Vol. 27, No. 1, pp. 51-90.

Gefen, D. (2004), 'What makes an ERP implementation relationship worthwhile: linking trust mechanisms and ERP usefulness', *Journal of Management Information Systems*, Vol. 21, No. 1, pp. 263-288.

Ghazizadeh, M., Lee, J.D. and Boyle, L.N. (2012), 'Extending the technology acceptance model to assess automation', *Cognition, Technology & Work*, Vol. 14, No. 1, pp. 39-49.

Grazioli, S and Jarvenpaa, S. (2000), 'Perils of internet fraud: an empirical investigation of deception and trust with experienced internet users', *IEEE Transactions on Systems, Man & Cybernetics, Part A : Systems and Humans*, Vol. 30, No. 4, pp. 395-412.

Hadar, L., Sood, S. and Fox, C. (2013), 'Subjective knowledge in consumer financial decisions', *Journal of Marketing Research*, Vol. 50, No. 3, pp. 303-316.

Hair, J.F., Black, W.C. Babin B.J. and Anderson R.E. (2013), *Multivariate Data Analysis*, Pearson Education, Harlow, UK.

Ho, S.Y. and Chau, P.Y. (2013), 'The effects of location personalization on integrity trust and integrity distrust in mobile merchants', *International Journal of Electronic Commerce*, Vol. 17, No. 4, pp. 39-71.

Hoffman, D.L., Novak, T.P. and Peralta, M. (1999), 'Building consumer trust online', *Communications of the ACM*, Vol. 42, No. 4, pp. 80-85.

Jarvenpaa, S.L. and Tractinsky, N. (1999), 'Consumer trust in an internet store: a cross-cultural validation', *Journal of Computer-Mediated Communication*, Vol. 5, No. 2, pp. 0.

Jarvenpaa, S.L., Tractinsky, N. and Vitale, M. (2000), 'Consumer trust in an internet store', *Information Technology and Management*, Vol. 1, No. 1, pp. 45-71.

Jeng, J. and Fesenmaier, D.R. (2002), 'Conceptualizing the travel decision-makinghierarchy: a review of recent developments', *Tourism Analysis*, Vol. 7, No. 1, pp. 15-32.

Johnson, D.S. (2007), 'Achieving customer value from electronic channels through identity commitment, and trust in technology', *Journal of Interactive Marketing*, Vol. 21, No. 4, pp. 2-22.

Johnson, E.J. and Russo, J.E. (1984), 'Product familiarity and learning new information', *Journal of Consumer Research*, Vol. 11, No. 1, pp. 542-550.

Keen, P, G. W., Balance, C., Chan, S. and Schrump, S. (1999), *Electronic Commerce Relationships: Trust by Design*, Prentice-Hall, Englewood Cliffs, NJ.

Keen, P.G. (2000), 'Ensuring e-trust', *Computerworld*, Vol. 34, No.11, pp. 46.

Khe, H.T. and Xie, Y. (2009), 'Corporate reputation and customer behavioral intentions. the roles of trust, identification and commitment', *Industrial Marketing Management*, Vol. 38, No. 7, pp. 732-742.

Kim, E. and Tadisna, S. (2007), 'A model of customers' trust in e-businesses: micro-level inter-party trust formation', *Journal of Computer Information Systems*, Vol. 48, No. 1, pp. 88-104.

Kim, H.W., Xu, Y. and Koh, J. (2004), 'A comparison of online trust building factors between potential customers and repeat customers', *Journal of the Association for Information Systems*, Vol. 5, No. 10, pp. 392-420.

Kim, J.B. (2012), "An empirical study on consumer first purchase intention in online shopping: integrating initial trust and TAM', *Electronic Consumer Research*, Vol. 12, No. 2, pp. 125-150.

Klink, R. R. and Smith, D.C. (2001), 'Threats to the external validity of brand extension research', *Journal of Marketing Research*, Vol. 38, No. 3, pp. 326-335.

Komiak, S.Y. and Benbassat, I. (2008), 'A two-process view of trust and distrust building in recommendation agents: a process-tracing study', *Journal of the Association for Information Systems*, Vol. 9, No. 12, pp. 727-747.

Kramer, R.M. (1999), 'Trust and distrust in organizations: emerging perspectives, enduring questions', *Annual Review of Psychology*, Vol. 50, No. 1, pp. 569-598.

Lankton, N., McKnight, D.H. and Thatcher, J.B. (2014), 'Incorporating trust-in-technology into expectation disconfirmation theory', *Journal of Strategic Information Systems*, Vol. 23, No. 2, pp. 128-145.

Lewicki, R.J. and Bunker, B.B. (1996), 'Developing and maintain trust in work relationships', in Kramer, R.M. and Tyler, T.M. (Eds.), *Trust in Organizations: Frontiers of Theory and Research*, Sage, Thousand Oaks, CA, pp. 114-139.

Lewicki, R.J., McAllister, D.J. and Bies, R.J. (1998), 'Trust and distrust: new relationships and realities', *The Academy of Management Review*, Vol. 23, No. 3, pp. 438-458.

Li, X., Hess, T. and Valacich, J.S. (2008), 'Why do we trust new technology? a study of initial trust formation with organizational information systems', *Journal of Strategic Information Systems*, Vol. 17, No. 1, pp. 39-71.

Lin, H.F. (2011), 'An empirical investigation of mobile banking adoption: the effect of innovation attributes and knowledge-based trust', *International Journal of Information Management*, Vol. 31, No. 3, pp. 252-260.

Liu, C.T., Marchewka, J. and Yu, C. (2005), 'Beyond concern: a privacy trust-behavioral intention model of electronic commerce', *Information and Management*, Vol. 42, No. 2, pp. 289-304.

Luhmann, N. (1979), *Trust and Power: Two Works*, Wiley, Chichester, UK.

Macintosh, G. and Lockshin, L.S. (1997), 'Retail relationships and store loyalty: a multi-level perspective', *International Journal of Research in Marketing*, Vol. 14, No. 5, pp. 478-497.

Masrek, M.N., Salwani, I., Daud, N.M. and Omar, N. (2014), 'Technology trust and mobile banking satisfaction: a case of Malaysian consumers', *Procedia -- Social and Behavioral Sciences,* Vol. 129, pp. 53-58.

Mayer, R.C., Davis, J.H. and Schoorman, F.D. (1995), 'An integrative model of organizational trust', *The Academy of Management Review*, Vol. 20, No. 3, pp. 709-734.

McKnight, D.H., Carter, M., Thatcher, J.B. and Clay, P.L. (2011), 'Trust in a specific technology: an investigation of its components and measures', *ACM Transactions on Management Information Systems*, Vol. 2, No. 2, doi: 10.1145/1985347.1985353.

McKnight, D.H., Cummings, L.L. and Chervany, N.L. (1998), 'Initial trust formation in new organizational relationships', *Academy of Management Review*, Vol. 23, No. 3, pp. 437-490.

McKnight, D.H. and Chervany, N.L. (2001), 'What trust means in e-commerce customer relationships: an interdisciplinary conceptual typology', *International Journal of Electronic Commerce*, Vol. 6, No. 2, pp. 35-59.

McKnight, D.H., Choudhury, V. and Kacmar, C. (2002), 'Developing and validating trust measures for e-commerce: an integrative typology', *Information Systems Research*, Vol. 13, No. 3, pp. 334-359.

McKnight, D.H., Kacmar, C. and Choudhury, V. (2004), 'Dispositional trust and distrust distinctions in predicting high- and low-risk internet expert advice site expectations', *E-Service Journal*, Vol. 3, No. 2, pp. 35-58.

Merritt, S.M. and Ilgen, D.R. (2008), 'Not all trust is created equal: dispositional and history-based trust in human-automation interactions', *Human Factors: The Journal of the Human Factors and Ergonomics Society*, Vol. 50, No. 2, pp. 194-210.

Merritt, S.M., Heimbaught, H., LaChapell, J. and Lee, D. (2013), 'I trust it, but I don't know why: effects of implicit attitudes toward automation on trust in an automated system', *Human Factors: The Journal of the Human Factors and Ergonomics Society*, Vol. 55, No. 3, pp. 520-534.

Montague, E.N., Kleiner, B.M. and Winchester, W.W. (2009), 'Empirically understanding trust in medical technology', *International Journal of Industrial Ergonomics*, Vol. 39, No. 4, pp. 628-634.

Moody, G.D., Galletta, D.F. and Lowry, P.B. (2014), 'When trust and distrust collide online: the engenderment and role of consumer ambivalence in online consumer behavior', *Electronic Commerce Research and Applications*, Vol. 13, No. 4, pp. 266-282.

Muir, B.M. (1994), 'Trust in automation: part I. theoretical issues in the study of trust and human intervention in automated systems', *Ergonomics*, Vol. 37, No. 11, pp. 1905-1922.

Nunnally, J.C. and Bernstein, I.H. (1994), *Psychometric Theory*, McGraw-Hill, New York, NY.

Osborne, J.W. and Banjanovic, E.S. (2016), *Exploratory Factor Analysis with SAS*, SAS Institute, Cary, NC.

Ou, C.X. and Sia, C.L. (2010), 'Consumer trust and distrust: an issue of website design', *International Journal of Human-Computer Studies*, Vol. 68, No. 12, pp. 913-934.

Ozawa, N. (2015), '$16 billion stolen from 12.7 million identity fraud victims in 2014', available at: https://www.javelinstrategy.com/press-release/16-billion-stolen-127-million-identity-fraud-victims-2014-according-javelin-strategy (accessed on July 22, 2015).

Park, C.W. and Moon, B.J. (2003), 'The relationship between product involvement and product knowledge: moderating roles of product type and product knowledge type', *Psychology and Marketing*, Vol. 20, No. 11, pp. 977-997.

Pavlou, P.A. (2003), 'Consumer acceptance of electronic commerce: integrating trust and risk with the technology acceptance model', *International Journal of Electronic Commerce*, Vol. 7, No. 3, pp. 101-134.

Pavlou, P.A. and Gefen, D. (2005), 'Psychological contract violation in online marketplaces: antecedents, consequences, and moderating role', *Information Systems Research*, Vol. 16, No. 4, pp. 372-399.

Podsakoff, P.M., MacKenzie, S.B., Jeong-Yeon, L. and Podsakoff, N.P. (2003), 'Common method biases in behavioral research: a critical review of the literature and recommended remedies', *Journal of Applied Psychology*, Vol. 88, No. 5, pp. 879-903.

Powell, W.W. (1996), 'Trust-based forms of governance', in Kramer, R.M. and Tyler, T.R. (Eds.), *Trust in Organizations: Frontiers of Theory and Research*, Sage, Thousand Oaks, CA, pp. 1-67.

Ratchford, B.T. (2001), 'The economics of consumer knowledge. *The Journal of Consumer Research*, Vol. 27, No. 4, pp. 397-411.

Riegelsberger, J., Sasse, M.A. and McCarthy, J.D. (2005), 'The mechanics of trust: a framework for research and design', *International Journal of Human-Computer Studies*, Vol. 62, No. 3, pp. 381-422.

Rotter, J.B. (1971), 'Generalized expectancies for interpersonal trust', *American Psychologist*, Vol. 26, No.5, pp. 443-452.

Rousseau, D.M., Sitkin, S.B., Burt, R.S. and Camerer, C. (1998), 'Not so different after all: a cross-discipline view of trust', *Academy of Management Review*, Vol. 23, No. 3, pp. 393-404.

Schmidt, J.B. and Spreng, R.A. (1966), 'A proposed model of external consumer information search', *Journal of the Academy of Marketing Science*, Vol. 24, No. 3, pp. 246-256.

Schoorman, F.D., Mayer, R.C. and Davis, J.H. (2007), 'An integrative model of organizational trust: past, present, and future', *Academy of Management Review*, Vol. 32, No. 2, pp. 344-354.

Seckler, M., Heinz, S., Forde, S., Tuch, A.N. and Opwis, K. (2015), 'Trust and distrust on the web: Uuser experiences and web site characteristics', *Computers in Human Behavior*, Vol. 45, pp. 39-50.

Selnes, F. and Howell, R. (1999), 'The effect of product expertise on decision making and search for written and sensory information', *Advances in Consumer Research*, Vol. 26, No. 1, pp. 80-89.

Silver, H. (2014) 'The impact of target's data breach on consumer trust', available at: http://connexity.com/blog/2014/05/the-impact-of-targets-data-breach-on-consumer-trust/ (accessed on July 22, 2015).

Shapiro, S.P. (1987), 'The social control of interpersonal trust', *American Journal of Sociology*, Vol. 93, No. 3, pp. 623-658.

Sitkin, S.B. and Roth, N. (1993), 'Explaining the limited effectiveness of legalistic "Remedies" for trust/distrust', *Organizational Science*, Vol. 4, No. 3, pp. 367-392.

Suh, B. and Han, I. (2003), 'The impact of consumer trust and perception of security control on the acceptance of electronic commerce', *International Journal of Electronic Commerce*, Vol. 7, No. 3, pp. 135-161.

Sujan, M. (1985), 'Consumer knowledge: effects on evaluation strategies mediating consumer judgments', *Journal of Consumer Research*, Vol. 12, No. 1, pp. 31-46.

Thornhill, T. (2014), 'Nearly half of South Koreans have their bank details stolen (including the President) as anti-fraud worker arrested', *Daily Mail*, available at http://www.dailymail.

co.uk/news/article-2543167/Data-100MILLION-South-Korean-credit-cards-stolen-scam-affecting-40-population-including-President-Park-Geun-hye.html (accessed July 22, 2015).

Vance, A., Elie-Dit-Cosaque, C. and Straub, D.W. (2008), 'Examining trust in information technology artifacts: the effects of system quality and culture', *Journal of Management Information Systems*, Vol. 24, No. 4, pp. 73-100.

Vogt, C.A., and Fesenmaier, D.R. (1998), 'Expanding the functional information search model', *Annals of Tourism Research*, Vol. 25, No. 3, pp. 551-578.

Williamson, O.E. (1985) *The Economic Institutions of Capitalism*. Free Press, New York, NY.

Wu, J.J. and Tsang, A.S. (2008), 'Factors affecting members' trust belief and behavior intention in virtual communities', *Behaviour & Information Technology*, Vol. 27, No. 2, pp. 115-125.

Xu, J., Kim, L., Dietermann, A. and Montague, E. (2014), 'How different types of users develop trust in technology: a qualitative analysis of the antecedents of active and passive user trust in a shared technology', *Applied Ergonomics*, Vol. 45, No. 6, pp. 1495-1503.

Zhou, M. and Tian, D. (2010), 'An integrated model of influential antecedents of online shopping initial trust: empirical evidence in a low-trust environment', *Journal of International Consumer Marketing*, Vol. 22, No. 2, pp. 147-167.

Zucker, L.G. (1986), 'Production of trust: institutional sources of economics structure, 1840-1920', *Research in Organizational Behavior*, Vol. 8, pp. 53-111.

## About the authors

**Steven John Simon** holds a PhD in Information Technology and International Business from the University of South Carolina and is currently an Associate Professor of Information Technology at Mercer University in Atlanta Georgia. He is a retired US Navy Captain having served 26 years. His assignments included Deputy CIO for the Office of Naval Research and US Sixth Fleet, CIO/J-6 for USSTRATCOM's WMD Center, Director of the Cyber Security Center at USNA, and Commanding Officer for both the Department of the Navy's Communication Security (COMSEC) System and Naval Information Operations Center – Georgia. He has published over 75 scholarly papers in journals including *ISR*, *EJIS*, *CACM*, and was editor-in-chief for *JIST*.

Corresponding Associate Professor, Stetson School of Business and Economics, Mercer University Dr., Atlanta, GA, 30341. Tel: +01-678-547-6118. E-mail address: Simon_sj@mercer.edu

**Carol J. Cagle** holds a Ph.D. in Business Administration and Management Science from the University of Texas at Arlington, M.S. in Management of Technology from Georgia Institute of Technology, and M.S. and B.S. in Computer Science from George Washington University and Naval Postgraduate School, respectively. Dr. Cagle teaches graduate and undergraduate operations and supply chain management and business analytics courses. Her current research interests include operations strategy and business analytics. Dr. Cagle has extensive experience with Fortune 500 companies managing technological developments. She is an active member of the Institute of Electronic and Electrical Engineers (IEEE), the Association of Computing Machinery (ACM), the American Society for Quality (ASQ), The Institute for Operations Research and Management Sciences (INFORMS), and the Council of Supply Chain Management Professionals (CSCMP). E-mail address: Cagle_cj@mercer.edu

## Appendix

### *Retained Survey Instrument Items*

(source in italics)

Have you, a family member, or close friend had your personal information compromised (information/files stolen or financial data, e.g. credit card or bank numbers) as a result of data theft or a hack with a retailer or organization you provided information?

**Knowledge** (Kim et al., 2004)

K1 I understand the safeguards in place to protect my personal information.

K2 I understand about encryption that can protect my information when stored and during transmission.

K3 I understand that there are laws in place to limit my liability if my personal information is compromised.

K4 I know that organizations are responsible for protecting my personal information.

**Disposition to Trust** (Benamati et al., 2010; Ho & Chau, 2013; McKnight et al., 2002)

DT1 I usually trust people until they give me a reason not to trust them.

DT2 I generally give people the benefit of the doubt when I first meet them.

DT3 My typical approach is to trust new acquaintances until they prove I should not trust them.

**Disposition to Distrust** (Ho & Chau, 2013)

DDT1 Most people are usually out for their own good.

DDT2 Most people pretend to care more about one another than they really do.

DDT3 Most people inwardly dislike putting themselves out to help other people.

DDT4 Most people would tell a lie if they could gain by it.

DDT5 Most people would cheat on their income tax if they thought they could get away with it.

**Technology Trust** (Cloesca, 2009)

TT1 I believe the technologies supporting ... are reliable.

TT2 I believe the technologies supporting ... are secure.

TT3 Overall, I have confidence in the technology used by....

**Reputation**

REP1 People say this organization has a good reputation. (Doney & Cannon, 1997)

REP2 In public opinion, this organization is favorably regarded. (Kim et al., 2004)

REP3 People say this organization has a good image. (Grazioli & Jarvenpaa, 2000)

REP4 This organization is well respected by people. (McKnight et al., 2002)

**Trust (in Organization)**

TR1 This organization is trustworthy. (Grazioli & Jarvenpaa, 2000)

TR3 This organization keeps customers' best interests in mind. (Grazioli & Jarvenpaa, 2000)

TR4 This organization would do the job right even if not monitored. (Suh & Han, 2003)

**Distrust (in Organization)** (Ou & Sia, 2010)

DB1 This organization looks suspicious.

DB2 This organization seems distrustful.

DB3 I feel skeptical (i.e., have doubts) about this organization.

DB4 I must be very watchful and wary when dealing with this organization.

DB5 I am fearful of dealing with this organization.

**Intentions** (Pavlou, 2003; Suh & Han, 2003)

I1 I intend to continue doing business with … in the future.

I2 I expect I will continue working with ... in the future.

I3 I will strongly recommend others to use....

# CALL FOR PAPER

## *MIS Review: An International Journal*

Published 2 Issues Annually by Airiti Press Inc.

*MIS Review* is a double-blind refereed academic journal published jointly by Airiti Press Inc. and Department of Management Information Systems, College of Commerce, National Chengchi University in Taiwan. The journal is published both in print and online. We welcome submissions of research papers/case studies in the areas including (but not limited to):

**1. MIS Roles, Trends, and Research Methods**

Roles, positioning and research methods of management information systems, and the impacts & development trends of information technology on organizations.

**2. Information Management**

Information infrastructure planning and implementation, information technology and organizational design, strategic applications of information systems, information system project management, knowledge management, electronic commerce, end-user computing, and service technology management.

**3. Information Technologies**

Database design and management, decision support systems, artificial intelligence applications (including expert systems and neural networks), software engineering, distribution systems, communication networks, multimedia systems, man-machine interface, knowledge acquisition & management, data mining, data warehouse, cooperative technology, and service science & engineering.

**4. Information Applications and Innovations**

The applications and innovations of business functional information systems (e.g., production, marketing, financial, human resources, and accounting information systems), enterprise resource planning, customer relationship management, supply chain management, intellectual capital, geographic information systems, and integrated information systems.

**5. Information Technology Education and Society**

Information education, e-learning, and information impacts on society.

**6. Others**

Other MIS-related topics.

## INSTRUCTIONS FOR SUBMISSION

1. Papers can be prepared in either Chinese or English. If your paper is written in Chinese, it will be translated into English once it is accepted for publication.
2. There is no submission deadline for MIS Review. All papers will be double-blind reviewed by at least two reviewers, who will be recommended by the Editorial Board. The processing time for the first-round formal reviews is about six weeks. Subsequently rounds of reviews tend to be faster.
3. To simplify file conversion effort, PDF or Microsoft Word 2000/2003 (for Windows) format is advised. Then, please submit your paper via the MIS Review website (URL: http://www.icebnet.org/misr/).
4. MIS Review is an academic journal. According to international practice, once an article is accepted and published, MIS Review will not give or take any payment for the publishing. An electronic copy of the paper will be sent to the article author(s) for non-profit usage.
5. The submitted and accepted paper should follow the author guidelines for paper submission format provided on the MIS Review website.

**The submitted paper should include the title page, abstract, key words, the paper body, references, and/or appendices. You must submit three files. The information of author(s) should not appear anywhere in the paper body file, including page header and footer.**

1. On a separate (cover letter) file, please follow the author guidelines provided on the MIS Review website to prepare the letter.
2. On a separate (title page) file, please note the title of the paper, names of authors, affiliations, addresses, phone numbers, fax numbers, and E-mail addresses.
3. On a separate (paper body) file, please include the paper title, an abstract, a list of keywords, the paper body, the references, and/or appendices. The abstract must contain the research questions, purposes, research methods, and research findings. The abstract should not exceed 500 words and the number of keywords must be 5-10 words.
4. The submitted and accepted paper should follow the author guidelines for paper submission format provided on the MIS Review website.

## CONTACT

Editorial Assistant
Department of Management Information Systems
College of Commerce, National Chengchi University
No. 64, Sec. 2, ZhiNan Road, Wenshan District,
Taipei 11605, Taiwan R.O.C.
Phone: +886-2-29393091 ext.89055
E-mail: misr@mis.nccu.edu.tw

# airiti Press Subscription Form

## MIS Review

You may subscribe to the journals by completing this form and sending it by fax or e-mail to

Address: 18F., No. 80, Sec. 1, Chenggong Rd., Yonghe District, New Taipei City 23452, Taiwan (R.O.C.)
Tel: +886-2-29266006 ext. 8695 Fax: +886-2-29235151 E-mail: press@airiti.com Website: http://www.airitipress.com

| PERSONAL | | | | LIBRARIES / INSTITUTIONS | | | |
|---|---|---|---|---|---|---|---|
| | Europe | US/CA | Asian/Pacific | | Europe | US/CA | Asian/Pacific |
| 1 Issue | € 34 | US$ 41 | US$ 38 | 1 Issue | € 53 | US$ 65 | US$ 62 |

| Vol. | | No. | | ~ | Vol. | | No. | | Copies | | US$ | | Total US$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

*All Price include postage

**PLEASE NOTE**

- Issues will be sent in two business days after receiving your payment.
- Please note that all orders must be confirmed by fax or email.
- Prices and proposed publication dates are subject to change without notice.
- Institutions include libraries, government offices, businesses, and for individuals where the company pays for the subscription.
- Personal rates are available only to single-user personal subscribers for personal and non-commercial purposes.
- Airiti Press reserves its right to take appropriate action to recover any losses arising fromany intended or unintended misrepresentation of the term "Personal Subscriber".

## BILLING INFORMATION

| | |
|---|---|
| Name | |
| Company | |
| Tel | Fax |
| E-mail | |
| Shipping Address | |

## INTERNATIONAL PAYMENTS

| **Pay by Credit Card** | |
|---|---|
| Card Type ☐JCB ☐MasterCard ☐Visa | |
| Card Name | |
| Card Number | |
| Expiry Date _____ / _____ | CVV number |
| Signature | |

| **Direct Bank Transfer** | |
|---|---|
| Beneficiary | AIRITI INC. |
| Address | 18F., No. 80, Sec. 1, Chenggong Rd., Yonghe District, New Taipei City 23452, Taiwan (R.O.C.) |
| Bank Name | E.Sun Commercial Bank, Ltd.Yong He Branch |
| Account No | 0107441863017 |
| Swift Code | ESUNTWTP |
| Bank Address | No.145, Zhongzheng Rd., Yonghe District, New Taipei City 23454, Taiwan (R.O.C.) |