Privacy Protection of Node Location and Data in Wireless Sensor Networks

https://doi.org/10.3991/ijoe.v12i11.6235

He Huan
Chongqing College of Electronic Engineering, Chongqing, China

***Abstract*—In order to solve the problem of low safety and low efficiency in distributed data protection in wireless sensor networks, a new method for privacy protection of node location and data in wireless sensor networks is researched in this paper. The author develops a novel architectural framework that observes user-specific patterns to distinguish between legitimate users and illegitimate user's thus enhancing and strengthening user authentication. An engine that calculates the authentication score (the similarity of features of the past and recent behaviors) for a given user is proposed. Simulation results show that the proposed method can improve overall system performance substantially.**

***Index Terms*—privacy protection, user authentication, authentication score, wireless sensor networks**

I. INTRODUCTION

New technology is constantly emerging in information age, and profoundly is changing the way of our work and study, like as cloud computing, big data, wireless sensor etc[1-3], however, data security becomes a problem. In general, security is the condition of being protected against danger or loss. Security is a concept similar to safety. The word "security" is synonymous with "safety", but as a technical term "security" means that is not only secure but that also has been secured. Therefore, security is defined as "the quality or state of being secured to be free from danger"[4]. In the computer technology area, the types of security are often addressed as following: Computer security, Data security, Application security, Information security and Network security. A successful network should have the following multiple layers of security in place to protect operations according to [5]:

Physical security-addresses the issue necessary to protect the physical items, objects, or areas of network from unauthorized access and misuse. Personal security- involves the protection of the individual or group of individuals who are authorized to access the network and its operations.

Operation security- focuses on the protection of the details of a particular operation or series of activities. Communication security-encompasses the protection of a network communications media, technology and its content.

Network security and information security is important. Security professionals try to protect their environment as effectively as possible using several actions. These actions can also be described as protecting confidentiality, integrity, and availability (CIA), or maintaining CIA. According to [6], CIA stands for:

Confidentiality-Ensures that no data is disclosed intentionally or unintentionally: Integrity-Ensures that no data is modified by unauthorized personnel, that no unauthorized changes are made by authorized personnel, and that the data remains consistent, both internally and externally.

Availability-although a secure computer must restrict access attempts by unauthorized users, it must still make the data availability to allow authorized users immediate access.

Authentication is one of the largest and fastest-growing segments of computer security. In this case, security refers to a collection of safeguards that the confidentiality of the information, protects the integrity of information, accounts for use of the system and protects the systems) availability of information, networks) used to process information [6]. Authentication is the process of establishing whether is who s/he declares to be or the process of verifying the claimed identity of a user. The network elements (NE) must offer features to verify the claimed identity of a user before giving that user operations access. Depending on the NE and the applications, there could be different kinds of identification and authentications. To illustrate this, consider the following example [6]:

a) The user can be associated with confidential information that only he or she is supposed to process such as: password, private key, or randomly time-varying PIN (such as those provided by single-use password token).

b) The user can be associated with a distinctive physical or logical address (e.g., user's authorized directory number, network address)

c) The user can be authenticated by certain unique attributes such as: voice or speech pattern, hard writing styles; palm, or retina scan.

The ways in which someone may be authenticated fall into three categories based on what are known as the factors of authentication: something you know, something you have, or something you are. Each authentication factor covers a range of elements used to authenticate or verify a person's identity prior to being granted access, approving a transaction request, signing a document or other work product, granting authority to others, and establishing a chain of authority. The three factors (classes) and some of elements of each factor are:

a) the ownership factors: Something the user has (e.g., wrist band, ID card, security token, software token, phone, or cell phone) the knowledge factors: Something the user knows (e.g., a password, pass phrase, or personal identification number (PIN), challenge response (the user must answer a question)).

b) the inherence factors: Something the user is or does (e.g., fingerprint, retinal pattern, DNA sequence (there are assorted definitions of what is sufficient), signature, face, voice, unique bio-electric signals, or other biometric identifier).

c) An important point is which type of users makes use of authentication systems, or most explicitly to whom authentication is targeted? An end-user is the final or ultimate user of a computer system, or an individual who is giving his physical characters (for example fingerprints; to a biometric authentication system [7]. The word "end-user" usually implies an individual with a moderately low level of computer expertise. This is a key area of this research.

II. OVERVIEW

In computer security, strong authentication refers to systems that require rigorous user identity verification, which is realized through multiple factors for authentication and use advanced technology. It allows us to irrevocably determine the user's identity or the integrity of a specific data. Strong authentications also pre-suppose that access to a network is tremendously hard to break, therefore creating a secure network. The goal of a strong authentication is to reinforce the security by replacing the classic authentication methods of password for software-only authentication with dynamic password generators, or software-hardware authentication solutions like tokens and smartcards [8]. Classic authentication assumes we know something: the user and the password. On the other hand, strong authentication assumes we know something (for example password), but also employs something that will generate the password (like OTP token) combination of two or three authentication factors is a good method to employ a strong authentication. The need for strong authentication is greater than ever, the cost of solutions such as single sign on and strong two factor authentication has come down, and such solutions are now easier to use. It is time for companies to look at improving their authentication procedures, if they want to remain secure and avoid potential business disruption, financial loss and damage to reputation.

There are four authentication factors that can be used in authentication mechanism [9]:

a) Something you know, which usually refers to password and PINs. The simplest implementation of password and personal identification numbers (P1Ns) yield the simplest of all authentication mechanisms.

b) Something you have, which usually refers to cards or tokens. Physical authentication devices, such as smart cards and password tokens, were developed to eliminate certain weakness associated with the passwords. A major benefit of cards and tokens is that they can't be shared with the same freedom as sharing passwords.

c) Something you are, which refers to biometrics-the measurement of physical characteristic or personal traits. Common biometrics verification techniques try to match measurements from one user's fingerprint, hand, face, or voice to measurements that were previously collected from him/her.

d) Something only the user conveys, like a Microprocessor chip implanted under human skin that authenticates user's access to a device. An authentication factor is authentication information that is (information used to set up the validity of a claimed identity) used to check an identity demanded by or for a user. Consider the following scenario: Before a reliable security system (RSS) gives a legitimate user access to a computer system, network, or secure resource, the RSS must determine who he is, if he belongs to this system, if he has the right to access this system, and if he is the person he says he is.

Actually, the RSS has demanded a legitimate user the three distinct elements such as identification, authentication and authorization, which all together comprise the so-called access control. However how does the RSS confirm that the user (legitimate user) is who he says he is? For example, entering the password does not prove it is who he says s/he is. Hence, the RSS needs the authentication information to authorize access to legitimate user. The authentication information may be gathered from one of the authentication factors [10,11]. In that case, using any authentication factor alone provides single-factor authentication. Any two authentication factors may be combined to provide two-factor authentication which is more secure.

While it may appear that any of these means could provide strong authentication, there are problems associated with each. If people wanted to pretend to be someone else on a co-system, they can guess or learn that individual's password; they can also steal or fabricate tokens.

Each method also has drawbacks for legitimate users and system administrators: users forget passwords and may lose tokens, and administrative overhead for keeping track of I&A data and tokens can be substantial. Biometric systems have significant technical, user acceptance, and cost problems as well.

It covers most of authentication mechanisms that directly authenticate users. Most of explicit authentication methods require you to specify who or what you are and to relay appropriate credentials to prove that you are who you say you are. These credentials generally take the form of something you know, something you have, or something you are. What you know may be a password. What you have could be a token. What you are is the biometric in which sophisticated equipment is used to scan a person in order to provide authentication. The important element is used to recognize that different mechanisms provide authentication services with varying degrees of certainty. Choosing the proper authentication technology largely depends on the location where the user is being authenticated and the factors used in the authentication.

Passwords are the most prevalent form of authentication, but are only one of many technological methods available to secure systems from unauthorized access. Passwords provide "security at minimal inconvenience", offering adequate and inexpensive security in the early days of non-networked computers. There are numerous limitations to the password approach. Forgotten passwords are perhaps the most obvious problem, causing frustration and delay for users. A typical Internet user today has multiple passwords to memorize and recall on demand. This memory burden leads to types of behavior that can compromise security, e.g., writing passwords down or frequently reusing them to alleviate memory limitations. The biggest problem of a good password is that users can forget them easily. "Asking a user to recall a single user ID and password for one system may seem reasonable, but with proliferation of passwords, users are increasingly unable to cope with the many username and passwords.

Users create easily identifiable passwords and passphrases which are easily attacked by guessing and social engineering. As long as a user can enter the correct username and password, the computer and the system will assume that the operator who is using the computer is the legal user or original user. No continuous authentication throughout the sessions after the first username and password entry. If users select their own passwords, they tend to make them easy to remember. That often makes them easy to guess. The names of people's children, pets, or favorite sports teams are common examples. On the other hand, assigned passwords may be difficult to remember, so users are more likely to write them down. Many computer systems are shipped with administrative accounts that have preset passwords. Because these passwords are standard, they are easily "guessed." Although security practitioners have been warning about this problem for years, many system administrators still do not change default passwords. Another method of learning passwords is to observe someone entering a password or PIN. The observation can be done by someone in the same room or by someone some distance away using binoculars. This is often referred to as shoulder surfing.

III. METHOD AND ALGORITHM

With mobile commerce and more applications being accessed on mobile devices, services on the web and more data being stored, it shows that users are now carrying mobile/using devices that require greater level of protection. However, mobile devices are frequently used in insecure locations with little or no protection mechanisms. Because mobile devices are more susceptible of being theft, there is need to safeguard them in case they are stolen. In the recent years, mobile devices have become very powerful in our day to day life. There has been rapid growth of mobile commerce meaning an increase in the number of applications hosted. With our approach, we offer a stronger authentication method that is aims to identify stolen devices in the hands of illegitimate users.

The current authentication methods are vulnerable to threats and significantly more frustrating and difficult to perform on these devices, leading users using for example passwords to create and reuse shorter passwords and pins, or no authentication at all. Due to their popularity, mobile devices are frequently used in insecure locations with little or no protection mechanisms making them vulnerable to theft or accessed illegally.

To enhance the existing authentication methods is by combining multifactor authentication that is implementing two or more classes of human authentications factors. In our case we will combine something known to only the user that is "knowledge based" that is password and pass phrase together with "something inherent" to only the user that is user behavior and in other words we call it implicit authentication. The combination of the two "known" and "inherent" factors makes up the multifactor authentication methods and significantly improves the authentication strength, as it curtails the threat of stolen digital identities. Implicit authentication acts as a second factor and supplements passwords for higher assurance authentication.

To overcome the weaknesses already mentioned in previous sections, the proposed approach is based on user behavior patterns. The approach employs user behavior patterns to enhance and strengthen authentication mechanisms. In this work, the proposed approach strengthens the authentication mechanisms using the user patterns in order to identify and come up with a unique pattern for each user. Each user has a behavior pattern that differs from other users. Thus, the propose approach uses the user behavior to differentiate between the legitimate and the illegitimate users.

In my approach, I introduce a new user credential i.e. the user behavior. User Behaviors is used to enhance the existing authentication methods. As it has been explained earlier on, each user has its own unique characteristics by establishing a baseline or a profile defining each user based on user behavior. This is done by collecting the usage pattern of mobile device of users and compare whether the user's current behavior conforms to the previous behavior. This approach will establish a baseline or a profile defining each user based on user behavior. If the similarity is within an acceptable range then the user is likely to be legitimate user but if the similarity is not within an acceptable ranges, then the user is likely to be not who he claim to be.

This could be due to the following reasons: Co-worker trying to access the device; possible stealing of the device and Malware is installed in the device to collect password of sensitive websites.

The general architecture covering all aspects of the proposed approach (i.e. how the past and recent behaviors) is shown in Figure 1. The architecture shows various sub components and how they relate to each other. For instance, it shows how the recent user behavior is gathered

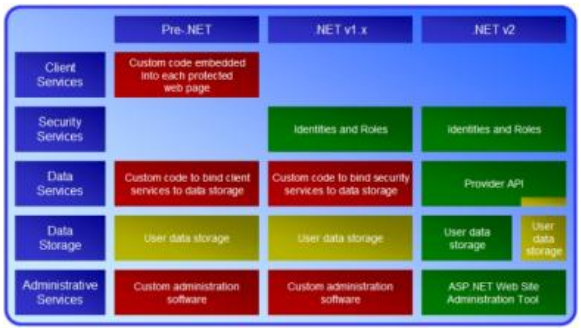

Figure 1. The general architecture

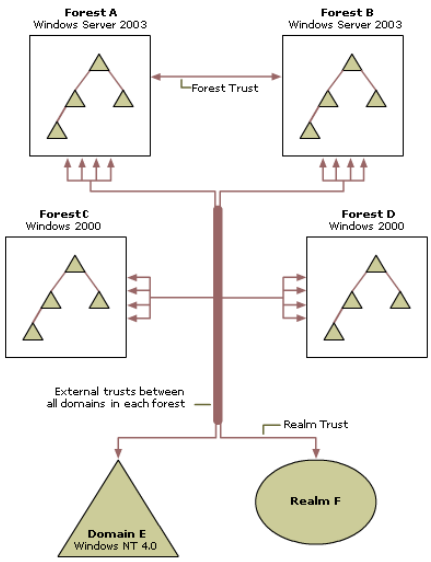

Figure 2. A framework of security transitions mechanism in wireless sensor network.

right from the point user details are submitted for verification by the authentication server. Logs manager help to track the details of how the user is interacting with network resources (such as email service, web service). There are numerous free source tools available on the Internet such as Snare and PAL that can help to generate the logs of different users and present them in a better format that is suitable for our proposed approach. In addition, there are several mobile devices tracking software available on the Internet such as Buddy Way and Mologogo. After the logs (data) are generated regarding a specific user, the same data is forwarded to the user behavior analyzer. The figure 2 shows the framework of security transitions mechanism in wireless sensor network. Figure 3 shows the diagram of security model.

Authentication servers are used to perform authorization or authentication operation of users. During operation, a user sends request to access a network resource. Authentication server collects contextual data about the user from users' devices. Logs manager that collects data on how users interact with network resources to facilitate authentication of the user. The authentication server periodically updates log manager. Devices such as laptops, PDA, smart mobile phones. Resources accessed such as file services, applications, email services, web services etc. The user behavior analyzer performs three tasks:

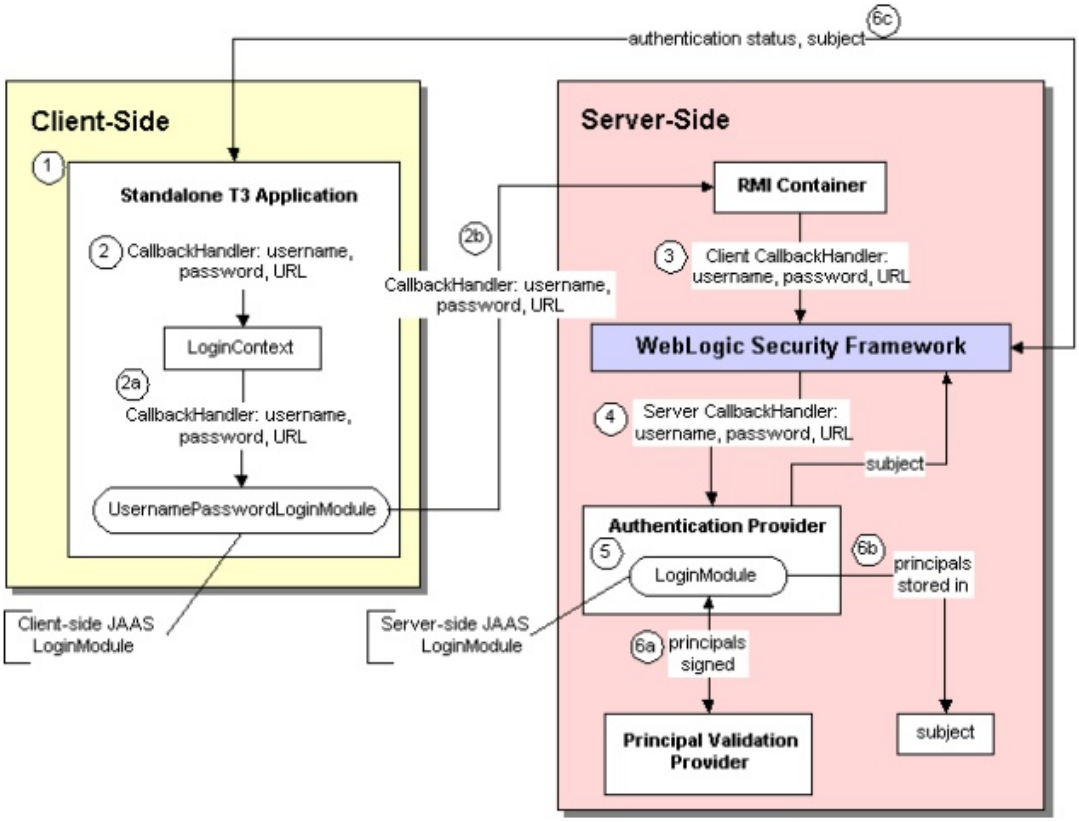

Figure 3. The diagram of security model.

To establish the baseline behavior of a given user, this baseline acts as the genuine behavior of a given user. This behavior is also referred to as past behavior.

-To establish a recent behavior of a given user.

-To compare the past behavior and the recent behavior of a given user.

S_{HM} solution vectors are randomly generated according to the variable range for each variable. Namely,

$$HMV = \left[\begin{array}{c|c} \boldsymbol{x}^1 & f(\boldsymbol{x}^1) \\ \boldsymbol{x}^2 & f(\boldsymbol{x}^2) \\ \mathrm{M} & \mathrm{M} \\ \boldsymbol{x}^{S_{HM}} & f(\boldsymbol{x}^{S_{HM}}) \end{array}\right] = \left[\begin{array}{cccc|c} x_1^1 & x_2^1 & \mathrm{L} & x_N^1 & f(\boldsymbol{x}^1) \\ x_1^2 & x_2^2 & \mathrm{L} & x_N^2 & f(\boldsymbol{x}^2) \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} \\ x_1^{S_{HM}} & x_2^{S_{HM}} & \mathrm{L} & x_N^{S_{HM}} & f(\boldsymbol{x}^{S_{HM}}) \end{array}\right] \quad (1)$$

Hence, we have equation (2):

$$x_i^{new} = \begin{cases} x_i^j,\ j \in \{1,2,...,S_{HM}\},\ if\ \ rand < HMCR \\ \overline{x}_i \in X_i,\ otherwise \end{cases} \quad (2)$$

As for x_i^{new} from harmony memory, we have equation (3):

$$x_i^{new} = \begin{cases} x_i^{new} + rand * BW, & if\ \ rand < PAR(\text{Continuous}) \\ (k+\lambda) * x_i^{new}, \lambda \in [-1,1], & if\ \ rand < PAR(\text{Discrete}) \\ x_i^{new}, & otherwise \end{cases} \quad (3)$$

The worst harmony is replaced with the new harmony, i.e.,

$$\boldsymbol{x}^{worst} = \boldsymbol{x}^{new},\ if\ \ f\left(\boldsymbol{x}^{new}\right) < f\left(\boldsymbol{x}^{worst}\right) \quad (4)$$

According to the analysis and description of routing problem in express delivery, the constraints can be expressed as:

(1) Distribution route length does not exceed the maximum value, i.e.,

$$\sum_{k=1}^{n} d\left(k-1,k\right) + d\left(n,0\right) \le MD \quad (5)$$

(2) The mathematical model of route programming can be defined as

$$L = \min\left(\sum_{k=1}^{n} d\left(k-1,k\right) + d\left(n,0\right)\right) \quad (6)$$

Based on the gradient descent method, node center and base width parameter are:

$$w_j\left(k\right) = w_j\left(k-1\right) + \eta\left(y\left(k\right) - y_m\left(k\right)\right)h_j + \alpha\left(w_j\left(k-1\right) - w_j\left(k-2\right)\right) \quad (7)$$

$$\Delta b_j = \left(y\left(k\right) - y_m\left(k\right)\right)w_j h_j \left(\frac{\left\|X - C_j\right\|^2}{b_j^3}\right) \quad (8)$$

$$b_j\left(k\right) = b_j\left(k-1\right) + \eta\Delta b_j + \alpha\left(b_j\left(k-1\right) - b_j\left(k-2\right)\right) \quad (9)$$

$$\Delta c_{j,i} = \left(y\left(k\right) - y_m\left(k\right)\right)w_j \frac{x_j - c_{j,i}}{b_j^2} \quad (10)$$

$$c_{ij}\left(k\right) = c_{ij}\left(k-1\right) + \eta\Delta c_{ij} + \alpha\left(c_{ij}\left(k-1\right) - c_{ij}\left(k-2\right)\right) \quad (11)$$

$$b_j\left(k\right) = b_j\left(k-1\right) + \eta\Delta b_j + \alpha\left(b_j\left(k-1\right) - b_j\left(k-2\right)\right) \quad (12)$$

$$\Delta c_{j,i} = \left(y\left(k\right) - y_m\left(k\right)\right)w_j \frac{x_j - c_{j,i}}{b_j^2} \quad (13)$$

$$c_{ij}\left(k\right) = c_{ij}\left(k-1\right) + \eta\Delta c_{ij} + \alpha\left(c_{ij}\left(k-1\right) - c_{ij}\left(k-2\right)\right) \quad (14)$$

Jacobian matrix algorithm is shown as follows:

$$\frac{\partial y\left(k\right)}{\partial u\left(k\right)} \approx \frac{\partial y_m\left(k\right)}{\partial u\left(k\right)} = \sum_{j=1}^{m} w_j h_j \frac{c_{1j} - x_1}{b_j^2} \quad (15)$$

Where $x_1 = u\left(k\right)$.

IV. EXPERIMENT RESULT

The experimental prediction slope one-join05 will crash. For small data throughput, we to forecast data set predict-probe scaled back, from 46 million users selected 5000 user data predict-probe0 as the forecast data set, so the final were root mean square error in the calculation of sample size also reduced to 17. Table 1 shows the experimental data set. Table II shows the operation time.

In addition to the experimental results of the application layer, through the experiment can also see the underlying platform in large data processing performance. Follow the interception of the core module of the experimental prediction of slopeone-join05, the implementation process of network, memory and CPU performance indicators. Among them, the most can reflect the distributed characteristics of the network are calculated using volume reached a peak of about 650m; amount of memory used the highest value of breakthrough 500g, although cluster distance to the total amount of memory and the gap from the image, but considering the platform is the cluster as a whole part, only 600g of actual amount of available memory, memory usage reached quite a high proportion; CPU usage peak also reached 60%.

TABLE I.
EXPERIMENTAL DATA SET

Data set	Record number
Original	100480507
Probe	1408395
Target-probe	1408395
Training-probe	99072112
Training-result-probe	2213532384
Training-result-probe0	157521183
Predict-result-probe	96438515
Predict-result-probe0	1030747
Predict-result-probe	935794590
Predict-result-probe0	84286169
rmse	17

TABLE II.
EACH STEP OPERATION TIME

Operation	Time(s)
Slopeone-Pjoin01	1598
Slopeone-Pjoin02	1787
Slopeone-Pjoin03	1676
Slopeone-Pjoin04	17636
Slopeone-Pjoin05	6626
Slopeone-Pjoin06	2417

V. DISCUSSION

The testing phase consisted of two stages. The first stage consists of a series of predetermined tasks which the user was required to perform thereby capturing user behavior over a wide cross-section of scenarios. During testing phase, the system captures the user's behavior for a predetermined time and tries to correctly identify the user based on the observed data. The results are validated by the user's responses. The behavioral attribute profiles of each user are then modified suitably thus having an accurate database of user profiles that detail the behavioral attributes of the users. The observation time varies based on the required granularity of user profile. The longer the session, the more the data is collected hence a better-defined profile of the user. However, it must be noted that increasing the observation session time period also increases the total testing time per user.

Once the testing phase is completed for all users, the system enters the verification phase where the user logs in to the system via a normal initial authentication way. Then, the users proceeds to perform different tasks on various applications. The proposed system is application independent and can collect information across applications from different devices thus ensuring that the user re-authentication is not limited to particular applications and the user is not constrained to use a particular application throughout the usage session. For every observation session, the system collects data from the user's actions. The behavioral attributes are extracted and passed to the verification (analysis) engine. The engine computes the authentication score by comparing the captured data and baseline data (past behavior) of that user and then recommends the best action for the user under consideration. It should be noted that repeated deviations (not within the acceptable range) from the past user's profile lead the authentication

system to suspect a security compromise. In this case, the system notifies the security analyst and then ends the current session by logging out. The user is required to re-login and begin a new session.

VI. CONCLUSION

Users are increasingly dependent on mobile devices. However, the current authentication methods are generally vulnerable to threats and significantly more frustrating and difficult to perform on these devices. This has contributed users, for example, to create and reuse shorter passwords and pins, or no authentication at all. Research in implicit authentication suggests the time period a specific trait is monitored is useful in identification of user behavior. Most people are creatures of habit a person goes to work in the morning, perhaps with a stop at the coffee shop, but almost always using the same route. Focusing on this aspect, this thesis has proposed an implicit user re-authentication approach that uses observations of users' behaviors for authentication. The proposed technique observes users- specific patterns in file system activity and network access to build models of normal behavior. These will help distinguish between normal use and anomalous use.

REFERENCES

[1] C. Guo, X. Liu, M. Jin, et al., "The research on optimization of auto supply chain network robust model under macroeconomic fluctuations," *Chaos, Solitons and Fractals,* September 2015.

[2] H. Jing, "Node deployment algorithm based on perception model of wireless sensor network," *International Journal of Automation Technology*, vol.9, no.3, pp. 210-215, April 2015. https://doi.org/10.20965/ijat.2015.p0210

[3] H. Jing, "Routing optimization algorithm based on nodes density and energy consumption of wireless sensor network," *Journal of Computational Information Systems*, vol. 11, no.14, pp. 5047-5054, July 2015.

[4] N. Aldin, P. Brehmer, and A. Johansson, "Business development with electronic commerce: refinement and repositioning," *Business Process Management Journal,* pp. 101-112, 2004. https://doi.org/10.1108/14637150410518329

[5] Y. Xu, X. Xie, and H. Zhang, "Modeling and Analysis of Electronic Commerce Protocols Using Colored Petri Nets". *Journal of Software,* pp. 1181-1187, 2011. https://doi.org/10.4304/jsw.6.7.1181-1187

[6] Y. Jiao, "Electronic Commerce Logistics Network Optimization Based on Swarm Intelligent Algorithm," *Journal of Networks,* pp. 89-98, 2013.

[7] Y. Zhang, and F. Han, "Embedded Spectrum Sensor Network Architecture and Transmission Medium Test Based on TCP/IP," *International Journal of Online Engineering,* vol. 12, no.5, pp. 38-42, May 2016. https://doi.org/10.3991/ijoe.v12i05.5734

[8] X. Zhang, et al., "Rotation-based privacy-preserving data aggregation in wireless sensor networks," *ICC 2014 - 2014 IEEE International Conference on Communications,* pp. 4184-4189, 2014. https://doi.org/10.1109/icc.2014.6883977

[9] J. Zheng, et al., "Auction-based adaptive sensor activation algorithm for target tracking in wireless sensor networks," *Future Generation Computer Systems,* vol. 39, no. 1, pp.88-99, 2014. https://doi.org/10.1016/j.future.2013.12.014

[10] R. Zhao, and C. Yue, "Toward a secure and usable cloud-based password manager for web browsers," *Computers & Security*, pp. 46-59, 2014. https://doi.org/10.1016/j.cose.2014.07.003

[11] B. Qin, H. Wang, Q. Wu, et al., "Simultaneous authentication and secrecy in identity-based data upload to cloud," *Cluster Computing,* pp. 164-171, 2013. https://doi.org/10.1007/s10586-013-0258-7

AUTHOR

He Huan is with the Chongqing college of Electronic Engineering, Chongqing, China (xhcqhs@sina.com).

Submitted 09 September 2016. Published as resubmitted by the author 16 October 2016.