A Concept of Behavioral Reputation System in Wireless Networks

Matej Kacic, Petr Hanacek, Martin Henzl and Ivan Homoliak
Faculty of Information Technology
Brno University of Technology
Bozetechova 2
612 66 Brno, Czech Rebuplic
Email: ikacic@fit.vutbr.cz

***Abstract*—Nowadays wireless networks are becoming important in personal and public communication. Most of them are secured by 802.11i standard with strong AES cipher - WPA2. In many cases an attacker has the ability to listen to all encrypted network traffic, which may become a potential intrusion. Each client in wireless network is vulnerable to a variety of threats and attacks. Many attacks, especially in corporate networks, are realized from internal environment. Identity theft is another serious problem of wireless networks. We present a concept of reputation system based on user behavior. Our goal is to precisely identify every entity in wireless network, and then determine malicious behavior of these entities.**

I. Introduction

A lot of research in this area usually focus on explicit identifiers such as MAC address which can be changed easily. Thus, it is challenging to track users and their behavior with always changing identifiers. In this paper we first analyze security issues of the newest standard (WPA2) in detail, and then we propose a concept of reputation system in 802.11i networks which can be used to achieve a correct identification of wireless entities and detection of malicious behavior of these entities.

The basement for reputation system is a creation of behavior model for each entity in the system (all devices and access points). This model is created by an algorithm which selects right attributes (signal strength, MAC address, FromDS, destination IP address, etc.) contained in an on-the-fly frame. These attributes are used in number of metrics which are able to detect or describe entity behavior. Our approach works across network layers; we take some attributes from radiotap header, all attributes from 802.11 frame, and many attributes from network, transport and application layer. The algorithm takes advantage from combination of radio-fingerprinting, link layer and all possible upper layers.

Obtaining data from upper layers than the link layer is very complicated because they are encrypted. We developed a way to gain data from upper layers by an extraction of cryptographic keys from access point and then using these keys to real-time decryption of 802.11 frames captured by wireless probe. Created model provides a behavior pattern of each entity in wireless system which is an important step for identification of entity. An artificial intelligence can take this model to detect potential malicious behavior and then raise or lower the value of reputation of entity. Entities with lower reputation than defined threshold are marked as intruders.

II. Security issues of 802.11i

Wireless networks since its creation have passed several phases of development with eliminating security vulnerabilities compromising wireless network. The 802.11i standard, finalized in 2004, is a security standard that can apply to other 802.11 standards. It involves many changes, including the addition of advance encryption cipher AES and better key management functionality. Today we use the 802.11i standard [1], known as WPA2, based on strong AES encryption algorithm. Standard WPA2 is considered safe for now, but we have to follow certain security rules in the design network and its configuration.

The wireless standard 802.11i is vulnerable to an attacker transmitting unauthorized management frames to force a client or access point to disconnect. By pretending to be a client or access point, an attacker can send either a de-authentication message to the other party to exit the authenticated state or a disassociation message to exit the association state. There are also hidden vulnerabilities based on the unprotected control packets CTS (clear to send) and RTS (request to send) allowing denial of service type of attack [2].

Conference Defcon 18 in summer 2010 brings new vulnerability of WPA2 enterprise encryption called ”hole 196” [3]. It allows a malicious insider (authorized user) to spoof the MAC address of the access point and to inject a GTK encrypted packet with broadcast destination address. The insider is able to launch several attacks such as ARP poisoning, DNS manipulation, Port scanning and DoS attack without detection. We have published [4] the way to inject malware to specific wireless client with purpose, for example, buffer overflow insertion attack to specific network application. We know that buffer overflow conditions exist when program allows to put into a buffer more data than it can hold. An attacker can use this vulnerability and insert malicious code into the memory of process and start execution. This can lead to gaining control of privileged application.

Another well known threat is Rouge AP which is an AP installed to network without authorization and does not follow security policy or an AP that has setup based on the malicious intention to compromise the information system of organizations i.e. data sniffing[5]. There are four types of rouge AP. First, Employees rouge access point, which is installed on the organizations LAN without authorization. Next type is Attackers external AP which is setup outside the company

and it does not connect to LAN. Third type, Attackers internal AP, creates a backdoor to companies LAN. And the last type is Neighborhood rogue AP where this AP is setup by other company.

The security issues of the newest standard described in this section show us how important security research is in this area of interest. The rest of the paper describes one of the solutions for detection of these issues by exploring, comparing and evaluating behavioral pattern of all devices i.e. users in wireless network.

III. Related works

Many researchers have performed a number of approaches of user behavioral modelling - analysis of user or network activity. They are based on the analysis and founding of patterns and common actions in user behavior for prediction, anomaly detection, identification, etc. Creation of user behavior model involves the following steps:

1) Data collection - collecting useful information about user activity;data that are relevant or defining user behavior.
2) Feature extraction - preprocessing the collected data with different approaches such as data-mining methods and machine-learning methods.
3) Dimension reduction - reducing the size of the data
4) Behavioral pattern extraction - application of different approaches to obtain specific characteristics of user behavior.
5) Interpretation of the results

First approach, The complex neural network model of user behavior in distributed systems [6], describes different features of user's behavior and their model consists of three components: on-line model considers dynamics of user behavior by predicting user's action, off-line model is based on the analysis of statistical parameters and change detection module that is intended for detection of trends in user's activity. They use a neural network to predict of the next user's action.

Another approach, Characterizing user behavior and network performance in public wireless LAN [7],has analyzed user's behavior compared to network performance. They characterize user's behavior in terms of connection session length, user's distribution across APs, mobility and bandwidth requirements. They use a trace recorded at the ACM SIGCOMM'01 conference, capturing the workload of 300,000 flows from 195 users consuming 4.6GB of bandwidth.

Lackner et al. [8] method applied the concept of Activation patterns to email data in order to extract information related to chosen set of features. The generated patterns and the information extracted by various analysis method is used for creating behavioral fingerprints. For analysis they have extracted 8 features which can be integrated into the Activation patterns. For the creation of fingerprints they have applied neural network to Activation patters.

Many examples refer to Intrusion Detection systems - anomaly detection in computer systems. They are applied different artificial technique such as statistical methods [9], expert systems, neural networks [10], agent-based systems [11], rule-based networks, etc. [12]

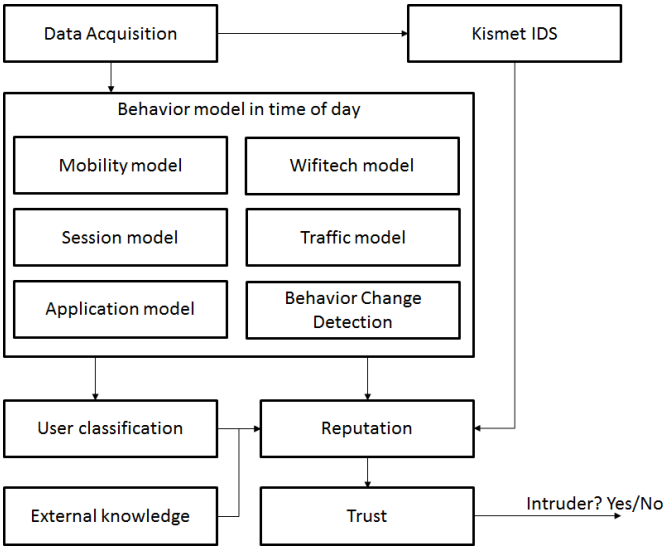

Fig. 1. Structure of behavioral reputation system

IV. Concept of the reputation system

This part describes the concept of reputation system in wireless network. First, the top level of the system is described and then each part is defined in detail. Figure 1 illustrates the main scheme of this system.

The system consists from seven main module:

1) Data Acquisition - is responsible for monitoring, capturing and preprocessing the data from wifi communication. Preprocessed data is sent to the Behavioral model and captured frames are sent to Kismet IDS module.
2) Kismet IDS - Kismet is an 802.11 layer2 wireless network detector, sniffer, and intrusion detection system. Kismet identifies networks by passively collecting packets and detecting standard named networks, detecting (and given time, decloaking) hidden networks, and infering the presence of nonbeaconing networks via data traffic [13]. Kismet provides some additional information to compute the reputation.
3) Behavior model - is the place where behavioral pattern of each entity is created. It is divided into several submodule characterizing individual behavioral aspects of entities in wireless space.
4) User classification - this module classifies users based on their behavior and reputation into several categories like admin, guest, employee, access point, intruder, etc.
5) External knowledge - provides an additional information from external sources like network IDS systems, radius server, etc.
6) Reputation and trust - these two models are compute reputation and trust of each entity in wireless area.

A. Data Acquisition

Obtaining data from upper layers than the link layer is very complicated, because they are encrypted. We developed a way to gain data from upper layers by extraction of cryptographic

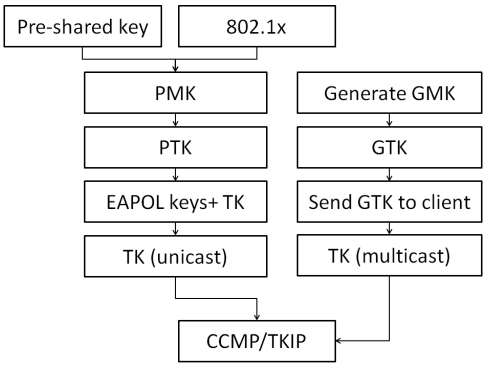

Fig. 2. 802.11i keys hierarchy

keys from access point and then to use these keys to real-time decryption of 802.11 frames captured by wireless probe. Wireless networks secured by 802.11i standard use several levels of cryptographic keys for different type of frames[14]. There are two possible top level keys that are used to generate the rest of the keys in the hierarchy. Those keys depend on chosen type of authentication of wifi network. The first key, pre-shared key, is used in home networks or in small business networks. On the other hand, huge corporate network use 802.1x[15], [16] authentication. Both keys serve to derive other key called Pairwise Master Key (PMK), which is then used for derivation of another key, Pairwise Transient Key (PTK). This key is unique for every connected client of wireless network and access point use this key for encrypting of communication between access point and wireless client.

The problem with maintaining an individual key with each client becomes apparent when dealing with multicast and broadcast traffic. If N clients are associated, the AP will have to retransmit the frame N times, encrypting it with a different key each time. To avoid this, the AP generates a random group master key (GMK). Every time a client associates or disassociates, the AP derives a new group transient key (GTK) from the GMK. This GTK is delivered to each of the clients to be used to encrypt and they decrypt multicast and broadcast traffic. Figure 2 illustrates full scheme of keys hierarchy of 802.11i standard.

The key extraction is not an easy party, because we need the PTK key for each connected station and we also need the GTK key and then use these keys to real-time decryption of 802.11 frames captured by wireless probe. It is obvious that key extraction procedure is one of possible compromising the privacy of connected users, but on the other hand the communications of users connected in wired network are also visible for network devices such as intrusion detection systems.

B. Model of User Behavior

The basement for reputation system is a creation of behavior model for each entity in the system (all devices and access points). This model is created by the algorithm which selects right attributes (signal strength, MAC address, FromDS, destination IP address, etc.) contained in an on-the-fly frame. These attributes are used in number of metrics which are able to detect or describe entity behavior. Our approach works across network layers; we take some attributes from radiotap header, all attributes from 802.11 frame, and many attributes from network, transport and application layer. The algorithm takes advantage from combination of radio-fingerprinting, link layer and all possible upper layers. Created model provides a behavior pattern of each entity in wireless system, which is an important step for identification of entity.

The model consist of five module: Mobility model, Traffic model, Wireless technology model, Session and Application model. In our modules we consider dynamical features of user's behavior and also many statistical properties.

1) Mobility model: One of the important criteria for determining the wireless user's behavior is user's mobility. The mobility pattern can be different from day to day, for reason of our nature. Sometimes we sit all day at one place, another time we pass from office to office with our wireless device. There are also some device types with very high mobility - mobile phones and tablets with VoIP services. An example of this type of user might be a person using VoIP application while walking to the office.

We can monitor the user's mobility by different approaches:

1) Changing Access Points - user can move between access points in time i.e. we can trace this movement at during the day.
2) Location of wireless device - triangulation on signal strengths from multiple access points can be used to pinpoint location down to a few meters [17].

The measure of mobility of wireless device is defined in five levels: Stationary, Low, Medium, High, Very high. These levels of mobility of wireless device define how often the device changes its position at times of day.

2) Session model: Differences of human variability are part of human existence and characterize a person in society. Every human has different habits, or tastes and these activities are reflected into their behavior in wireless network. Very good example is smoking break, lunchtime, when somebody has this habit and they have to leave their working place at specific time of day, thus the interruption of application and network session occurs.

User's behavior can be modelled by monitoring these activities or habits on network layer and application layer separately on timeline. Session model uses various metrics such as time of connection and connection duration in correlation with some access point, on which the time metrics (time of day, average time per day, maximum time per day,etc.) are applicated. On application layer the session type (specific application, multicast session, udp session, etc.), session start and length are monitored with time metrics used in network layer.

The output of this model is the measure of similarity between the corresponding session model and actual user action in time. Very low level of similarity is considered as suspicious behavior and the reputation of corresponding user is reduced.

3) Traffic model: Traffic model is based on statistical parameters of user behavior i.e. network metrics. Wireless Network Metrics are the measurable parameters or features that can be used to model the various behaviors on the network.

This model uses a number of metrics working on different network layers. For example, we count a number of input and output MAC frames at various time intervals (time of day, day, week) and, on the other hand, we track the traffic of all kind of application in time. Table I shows much more metrics as example.

MAC address	ARP/IP pair changes
Change of MAC	Spoofed disassociate message
Adhoc Network	Seq. number of Client
802.11x in use	NIC vendor
Authentication attempts	Frames RTS/CTS/ACK
Timeout for RTS	Time between frames
Fragmented headers	Deauth messages
Spoofed messages	Connection time
...	

TABLE I. RELEVANT WIRELESS NETWORK METRICS

Every metric in our design is extended by the dimension of time and it is extended by basic statistical function such as modus, median, average, the standard deviation, etc. For now we use a native Bayes classifier due to its effectiveness in application traffic classification [18]. More sophisticated classifiers exist [19], which will be one of our next research in this area.

4) Wireless technology model: Like a human fingerprint, network traffic has unique characteristics that can be used to identify a sender device. The wireless technology model tries to identify users in wireless network. The MAC address embedded in every frame may partially identify a user because of the uniqueness of each MAC address. On the other hand users can change the MAC address of their wireless devices. Under this assumption, we expect that the user MAC address is changing variable in this model. MAC address is assumed as an explicit identifiers. Unlike explicit identifiers, such as the MAC address, implicit identifiers cannot be associated directly with senders but they may have unique characteristics to be distinguished from other traffic. The MAC address of devices are used only for purpose of testing correctness in identification process.

This model uses two approaches for fingerprinting wireless devices [20]: First, the Active fingerprinting, where specifically crafted 802.11 frames are sent to the device and precise timings of the responses are gathered. This approach allows for adaptive fingerprinting, where the types and contents of subsequent frames depend on the timings gathered so far. This is roughly similar to a chosen-plaintext attack against a cipher. This approach is potentially detectable by the node being fingerprinted. The second approach for fingerprinting wireless devices is the passive fingerprinting, where we make the fingerprint by listening transmitted 802.11 frames and then by measuring their timings. Inputs of this model are properties of radiotap headers and MAC headers for example signal strength, antenna type, wireless standard type (a/b/g/n), transfer rate, timings between ACK frames, number of fragmented frames, error rates, number of retransmissions, average payload size, etc.

Wireless technology model does not follow the behavior of users, but rather describes the characteristics of network devices on the lowest level i.e. specify hardware and operating system fingerprint, which is very important in device identification.

5) Application Model: User's behavior represents a complex non-linear process, but there are some regularities that can be revealed. We believe that the sequence of user's actions in time and appropriate lengths of these actions can be used for exploring regularities in user' behavior. Thus the user's action can be defined as three-tuple consisting of port number, ip address and direction, which are coded to numeric representation. Our idea is based on using the neural network to predict the probability of actions in specific time T. The time is sampled as time of day and appropriate weekday with hour respectively day precision. Neural network inputs are time of day-weekday, action and duration of action. The result of neural network will be likelihood of whether the action occurs and its delay from the time T. It is obvious that for each user a neural network has to be built and trained.

The output of application model is similarity between the neural network output of each actions and the real sequence of actions in given point in time correlated with the duration of each action. The value of similarity is an input of the computation of reputation.

6) Behavior change detection: Human or user's behavior is changing in the course of time, hence we consider it as a dynamic process. Changes in user's behavior are caused by different reasons, e.g. due to new tasks, software version changes, salary reductions, etc. Therefore, the behavior change detection is required.

The rate of the behavior change is reflected by the computed reputation. If user's behavior has not changed, then actual value of reputation would not differ significantly from previous reputation values. On the other hand, if anomaly occurrs, reputation will differ considerably from previous reputation values and it will be above some threshold. Natural changes in user behavior are related with small reputation difference i.e. difference under some threshold. The change of behavior model is occurrs when reputation is changed in long term and at the same term the entity is considered as trust entity. When model is updated the value of reputation is set to default value according to user classification.

C. Reputation and Trust

Reputation is one of the factors upon which trust is based. Trust is the expectation of one person about the actions of others that affects the first person's choice when an action must be taken before the actions of others are known [21]. Reputation is a basis for trust [22].

Reputation systems are primarily seen as tools for ubiquitous computing, where the use of classical security mechanism is very difficult. The result of reputation system is security control with possibility of finding some new potentially suspicious behavior. On the other hand, reputation system works with a certain degree of inaccuracy and time delay.

In previous sections the user behavior model is described. This model reflects the behavior of users in dimension of time. Our idea is to use outputs of this model and also previous computed reputation values for each user to compute new reputation value. The computing of reputation is based on behavior

deviation i.e. the fluctuations (variations) in the behavior. There is also a feedback from wireless IDS system which provides us information about well-known attacks. A trust update algorithm maintains current trust state and combines it with a new observation in user behavior. Reputation value of each entity is an input of trust module which determine whether we trust this entity.

V. Conclusion

Advanced attacks have become more and more sophisticated and their detection is becoming more complicated. Many corporates have a number of employees and in many cases we are not aware of importance of security from inside environment. On the other hand, the tracking of the behavior of a large number of people is not easy. It would be nice if we could determine the level of trust of the entities in wireless network at any time given.

This paper describes the concept of reputation system based on user respectively device behavior in time dimension. The purpose of this concept is identification of entities and their evaluation from the viewpoint of trust in wireless network system. This concept consists of many modules modelling the aspects of behavior from the lowest layer to application layer. Each module needs to be explored in detail and this will be our future research. We would like to note that it is not able to reproduce the behavior precisely.

In accordance with concept presented in this paper, the formal description of each part is necessary. We are also planning to do some testing data, which can be used by other research group in world.

Acknowledgment

This work was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070), by the project CEZ MSM0021630528 Security-Oriented Research in Information Technology and by project FIT-S-11-1 Advanced Secured, Reliable and Adaptive IT.

References

[1] *IEEE Std 802.11i-2004*. IEEE, 2004, iSBN 0-7381-4074-0.

[2] A. Rachedi and A. Benslimane, "Impacts and solutions of control packets vulnerabilities with ieee 802.11 mac," *Wireless communications and mobile computing*, vol. 9, no. 4, pp. 469–488, 2009.

[3] M. S. Ahmad, "Wpa too!" *DEF CON*, vol. 18, 2010.

[4] M. Kacic, P. Hanacek, M. Henzl, and P. Jurnecka, "Malware injection in wireless networks," in *Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), 2013 IEEE 7th International Conference on*, 2013.

[5] S. Srilasak, K. Wongthavarawat, and A. Phonphoem, "Integrated wireless rogue access point detection and counterattack system," in *Information Security and Assurance, 2008. ISA 2008. International Conference on*, april 2008, pp. 326 –331.

[6] A. Shelestov, S. Skakun, and O. Kussul, "Complex neural network model of user behavior in distributed systems," in *Proc of XIII-th Int Conf Knowledge-Dialogue-Solutions, Varna, Bulgaria. Sofia, Bulgaria: FOI ITHEA*, 2007, pp. 42–9.

[7] A. Balachandran, G. M. Voelker, P. Bahl, and P. V. Rangan, "Characterizing user behavior and network performance in a public wireless lan," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 30, no. 1. ACM, 2002, pp. 195–205.

[8] G. Lackner, P. Teufl, and R. Weinberger, "User tracking based on behavioral fingerprints," in *Cryptology and Network Security*. Springer, 2010, pp. 76–95.

[9] H. Javitz and A. Valdes, "The sri ides statistical anomaly detector," in *Research in Security and Privacy, 1991. Proceedings., 1991 IEEE Computer Society Symposium on*, 1991, pp. 316–326.

[10] C. I. Ezeife and M. Z. Rahman, "Neudetect: a neural network data mining wireless network intrusion detection system," in *Proceedings of the Fourteenth International Database Engineering & Applications Symposium*, ser. IDEAS '10. New York, NY, USA: ACM, 2010, pp. 38–41. [Online]. Available: http://doi.acm.org/10.1145/1866480.1866487

[11] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. Suppl 3, pp. 7280–7287, 2002.

[12] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 15:1–15:58, Jul. 2009. [Online]. Available: http://doi.acm.org/10.1145/1541880.1541882

[13] "Kismet [online]," http://www.kismetwireless.net, 2010.

[14] Kevin Benton, *The Evolution of 802.11 Wireless Security*. UNLV Informatics, 2010-04-18.

[15] A. Earle, *Wireless security handbook*. Auerbach, 2005. [Online]. Available: http://books.google.cz/books?id=DojR6q1E5ZUC

[16] EAP Working Group, "Protected eap protocol (peap) version 2," http://tools.ietf.org/html/draft-josefsson-pppext-eap-tls-eap-10, 2004.

[17] R. Hansen, R. Wind, C. S. Jensen, and B. Thomsen, "Algorithmic strategies for adapting to environmental changes in 802.11 location fingerprinting," in *Indoor Positioning and Indoor Navigation (IPIN), 2010 International Conference on*. IEEE, 2010, pp. 1–10.

[18] A. W. Moore and D. Zuev, "Internet traffic classification using bayesian analysis techniques," *SIGMETRICS Perform. Eval. Rev.*, vol. 33, no. 1, pp. 50–60, Jun. 2005. [Online]. Available: http://doi.acm.org/10.1145/1071690.1064220

[19] S. Zander, T. Nguyen, and G. Armitage, "Automated traffic classification and application identification using machine learning," in *Local Computer Networks, 2005. 30th Anniversary. The IEEE Conference on*, 2005, pp. 250–257.

[20] B. Sieka, "Active fingerprinting of 802.11 devices by timing analysis," in *Consumer Communications and Networking Conference, 2006. CCNC 2006. 3rd IEEE*, vol. 1. IEEE, 2006, pp. 15–19.

[21] D. Gambetta, "Trust: Making and breaking cooperative relations," 1990.

[22] P. J. Windley, K. Tew, and D. Daley, "A framework for building reputation systems," *Www2007. Banff, Canada*, vol. 49, 2007.