A Rough Set Based Anomaly Detection Scheme Considering the Age of User Profiles

Ihn-Han Bae

School of Computer and Information Communication Eng.,
Catholic University of Daegu, GyeongSan 712-702, Korea
ihbae@cu.ac.kr

Abstract. This paper presents an efficient rough set based anomaly detection method that can effectively identify a group of especially harmful internal attackers – masqueraders in cellular mobile networks. Our scheme uses the trace data of wireless application layer by a user as feature value. Based on this, the use pattern of a mobile's user can be captured by rough sets, and the abnormal behavior of the mobile can be also detected effectively by applying a roughness membership function with the age of the user profile.

Keywords: Anomaly Detection, Rough Set, User Profile, Age.

1 Introduction

The nature of mobile computing environment makes it very vulnerable to an adversary's malicious attacks. The use of wireless links renders the network susceptible to attacks ranging from passive eavesdropping to active interfering. Unlike wired networks where an adversary must gain physical access to the network wires or pass through several lines of defenses at firewalls and gateways, attacks on a wireless network can come from all directions and target at any node [1]. In this paper, we propose the intrusion detection technique that the normal profile of each mobile is constructed by using rough sets and anomaly activities are effectively detected by using a rough membership function with the age of user profile. When an intrusion occurs, the attacker masquerading the legitimate user trends to have a different use pattern. Therefore, we can detect anomaly by comparing the use patterns. Therefore, we can detect anomaly by comparing the use pattern.

2 Related Work

The number of intrusions into computer system is growing because new automated intrusion tools are appearing every day. Although there are many authentication protocols in cellular mobile networks, security is still a very challenging task due to the open radio transmission environment and the physical vulnerability of mobile devices. Zhang, Y. proposed new model for Intrusion Detection System (IDS) and response in mobile, ad-hoc wireless networks. Sun, B. proposed a mobility-based anomaly detection scheme in cellular mobile networks. The scheme uses cell IDs traversed by a

user as the feature value [2]. In [3], based on fuzzy view of rough sets instead of exact rules, a soft computing approach that uses fuzzy rules for anomaly detection is proposed. Bae, I. H. proposed a rough set based anomaly detection scheme without considering the age of user profiles [4].

3 Rough Set

Let U be a finite set of objects called Universe, and $R \subseteq U \times U$ be an equivalence relation on U . The pair $A = (U, R)$ is called approximation space, and equivalence classes of the relation R are called elementary sets in A .

For $x \in U$, let $[x]_R$ denote the equivalence class of R , containing x . For each $X \subseteq U$, X is characterized in A by a pair of sets – its lower and upper approximation in A , defined as:

$$\begin{aligned}\underline{A}X &= \{x \in U \mid [x]_R \subseteq X\} \\ \overline{A}X &= \{x \in U \mid [x]_R \cap X \neq \emptyset\}\end{aligned}$$

The objects in $\underline{A}X$ can be with certainty classified as members of X on the basis of knowledge in R , while the objects in $\overline{A}X$ can be only classified as possible members of X on the basis of knowledge in A . The set $BN_A X = \overline{A}X - \underline{A}X$ is called the A -boundary region of X , and thus consists of those objects that we cannot decisively classify into X on the basis of knowledge in A .

Rough set can be also characterized numerically by the following coefficient called the accuracy of approximation, where $Card$ denotes the cardinality [5].

$$\alpha_A(X) = \frac{Card \underline{A}X}{Card \overline{A}X}$$

4 Anomaly Detection Considering the Age of User Profiles

In this section, we introduce the construction of a rough set based anomaly detection scheme that considers the age of user profiles. For each mobile user, the features of the user activities from the wireless application layer are captured. We use the traveled cell ID, the duration of the service and the service class of a user as the feature value. The feature values are stored in the user's feature profile database.

The equivalence classes from the user's feature profile database are computed by using rough sets. For a mobile user, based on both the user activity information and the equivalence class information, a deviation number is computed by a roughness membership function with the age of the user profile, where the deviation number represents the degree that the user behavior is deviated from the normal behavior. When a user activity occurs, if the deviation number is greater than the deviation threshold that is a system parameter, an alert message occurs, otherwise the user activity identified as normal. The whole scheme is illustrated in Fig. 1.

The user feature profile database used in our scheme is shown in Table 1, where REQ#, CELL, DUR, CLASS and AGE represent the service request number, the traveled cell ID, the requested service duration, the requested service class and the age

of the data, respectively. As the age of profile data is smaller, the profile data is newer. The profile data are continuously come and faded (aging) away.

Fig. 1. The structure of rough set based anomaly detection scheme

Table 1. User feature profile information system

REQ#	CELL	DUR	CLASS	AGE
1	a	1	α	3
2	a	1	α	3
3	a	2	α	3
4	b	2	β	3
5	b	2	β	3
6	b	2	β	2
7	b	2	β	2
8	b	3	γ	2
9	c	3	γ	2
10	c	3	γ	2
11	a	1	α	1
12	a	1	α	1
13	a	2	α	1
14	b	2	β	1
15	b	3	β	1

The fuzzy inclusions are represented by the inequalities of membership functions. Further, we will allow certain errors as long as they are within the radius of tolerance. The fuzzy inclusion is computed by Eq. (1), roughness membership function.

$$\mu(X, Y) = \frac{\sum_{age=1}^g [Card((\bar{R}X \cup \bar{R}Y) - (\bar{R}X \cap \bar{R}Y)) \times (1 - w_{age})]}{Card(\bar{R}X \cup \bar{R}Y)} \quad (1)$$

where g represents the number of age grades, w_{age} represents the weighted value of the age, and X and Y represent condition attribute and decision attribute, respectively.

We assume that the deviation threshold (ϵ) is 0.58. In case that a user activity ($c, 2, \alpha$) is occurred, let $X=\{c, 2\}$ and $Y=\{\alpha\}$, $\bar{R}X=\{3,4,5,6,8,9,10,14,15\}$ and $\bar{R}Y=\{1,2,3,11,12,13\}$, so that $\mu(X, Y) = 8.4/14 = 0.6$. Accordingly, the user activity is identified as anomalous because that $\mu(X, Y) > \epsilon$, and an alert message is generated.

5 Performance Evaluation

We use the following two metrics: detection rate, false alarm rate to evaluate the performance of our proposed anomaly detection scheme.

We present and analyze the simulation results at different deviation threshold. We assume as follows. In cases of age 1 and age 2, if a user activity had two records those the two values of the user activity attribute values were matched with the feature values of the user profile data, the user activity is normal. Also, if a user activity had one record that two values of the user activity attribute values were matched with the feature values of age 1 within the user profile data and one or more records those two values of the user activity attributes values were matched with the feature values of age 2 or age 3 within the user profile data, the user activity is normal. All user activities of other cases are abnormal. Table 2 shows the parameter values for the simulation.

The performance of the proposed scheme is compared with that of a rough set based anomaly detection scheme without aging [4]. Simulation results of the detection rate and the false alarm rate over deviation threshold are illustrated in Fig. 2. The detection rate of our scheme (*Detection (age)*) is better than that of the rough set based anomaly detection scheme without aging (*Detection (no age)*) regardless of deviation thresholds.

Table 2. Simulation parameters

Parameter	Value
The number of user activities	1,000
The traveled cell IDs	Random (1,4)
The type of service durations	Random (1,4)
The type of service classes	Random (1,3)
The weight of ages, (w_1, w_2, w_3)	(0.6, 0.3, 0.1)

Fig. 2. Detection rate and false alarm rate at different deviation threshold

6 Conclusions

In this paper, we propose the intrusion detection technique that the normal profile of each mobile is constructed by using rough sets and anomaly activities are effectively detected by using a rough membership function with the age of user profile. From the results, we know that the performance of our scheme is better than that of the rough set based anomaly detection scheme without aging regardless of deviation thresholds.

References

1. Zhang, Y., Lee, W., Huang, Y-A.: Intrusion Detection Techniques for Mobile Wireless Networks, *ACM Wireless Networks Journal*, 9(5) (2003): 545-556
2. Sun, B., Yu, F., Wu, K., Leung V. C. M.: Mobility-Based Anomaly Detection in Cellular Mobile Networks, *Workshop on Wireless Security*, (2004): 61-69
3. Lin, T. Y.: Anomaly Detection - A Soft Computing Approach, *Proceedings in the ACM SIGSAC New Security Paradigm Workshop*, (1994): 44-53
4. Bae, I. H.: Design and Evaluation of a Rough Set Based Anomaly Detection Scheme for Mobile Networks, *advances in Natural Computation and Data Mining*, (2006): 262-268
5. Pawlak, Z.: *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic Pub., (1991)