*This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE CCNC 2006 proceedings.*

# Towards Adaptive Anomaly Detection in Cellular Mobile Networks

Bo Sun, Zhi Chen, Ruhai Wang
Dept. of Computer Science
Lamar University
Beaumont, TX 77710
{bsun, zhic}@cs.lamar.edu, wang@ee.lamar.edu

Fei Yu, Victor C.M. Leung
Dept. of Electrical and Computer Engineering
University of British Columbia
BC, Canada V6T 1Z4
{feiy, vleung}@ece.ubc.ca

***Abstract*— Location information is an important feature of users' mobility profile in cellular mobile networks. In this paper, continuing our existing work on constructing a mobility-based anomaly detection scheme, we further address a challenging problem - how to *adaptively* adjust the detection threshold of Intrusion Detection Systems (IDSs) in the context of cellular mobile networks. This is especially critical when we consider the different mobility patterns demonstrated by the mobile users. Utilizing a high order Markov model, we apply a weighted blending scheme to compute the entropy of our Exponentially Weighted Moving Average (EWMA) based mobility trie. This reflection of the *uncertainness* of the users' normal profile could help us *adaptively* adjust the detection threshold of our anomaly detection algorithm. Simulation results show that our proposed adaptive mechanisms can further reduce the false positive rate without decreasing the detection rate. Detailed analysis of the simulation results is also provided.**

## I. Introduction

Cellular-based wireless networks have become very popular as more and more users not only communicate with others using cellular phones but also perform important and sensitive tasks such as E-Shopping and E-Banking. On one hand, the advance of the wireless technology makes life easier. On the other hand, it introduces serious risks. *Prevention*-based techniques, like authentication and encryption, can effectively reduce attacks by keeping illegitimate users from entering the system. However, mobile devices are prone to being stolen and physically insecure due to their portability. This low physical security can make all secrets of the device open to malicious attackers and render *prevention* based approaches useless. At this time, Intrusion Detection (ID) approaches, utilizing different techniques to model the users' normal behavior and system vulnerabilities, come into place to help identify malicious activities.

In our previous work [2], we proposed an Exponentially Weighted Moving Average (EWMA) [6] based approach to construct an anomaly-based Intrusion Detection System (IDS) for Cellular Mobile Networks. We parse the normal users' daily movement activities by applying a Markov model at different orders, and then use the Lempel-Ziv 78 ($LZ78$) [5] algorithm to construct a mobility trie to store the route related information. The EWMA technique is applied to the mobility trie to keep it up-to-date and accurately reflect the user's movement pattern. A *threshold*-based scheme is then used to determine whether the mobile device is potentially compromised or not.

One more challenging problem is how to *adaptively* adjust the detection threshold of IDSs in the context of cellular mobile networks. In reality, it is highly possible that a single user will demonstrate different mobility behaviors. For example, even if the user demonstrates the same mobility level, a user will have a set of mobility patterns during weekdays, while demonstrating a different set of mobility patterns during weekends. Therefore, it is desirable to not only change the normal profile adaptively, but also adjust the threshold of the IDS automatically. An accurate threshold plays an important role in determining the performance of the IDSs.

Intrusion detection problems for users demonstrating different sets of mobility patterns are not easy. It is difficult to identify a metric, adjust the threshold appropriately and integrate the mobility impact into the construction of IDSs. The main contributions of this paper include proposing an effective measurement to capture the *uncertainness* of the users' normal profile and integrating the above measurement into cellular IDSs adaptively. To the best of our knowledge, there is no previous work that has been contributed in this area.

## II. Motivations

The complex cellular mobile network system could incur software errors and design errors. This could make many attacks possible. One example is the cell phone cloning: the mobile phone card of an authenticate user A is cloned by some attacker B, which enables B to use the cloned phone card to make fraudulent telephone calls. If cell phone cloning happens, the bills for the calls will go to the legitimate subscriber. Also, the masquerader could fake the International Mobile Equipment Identifier (IMEI) and the SIM (Subscriber Identity Module) card in order to get the service illegally. Subscription fraud could also enable the intruder to subscribe to the service using the authentic user's name. All these enable the necessity of a fraud detection system that can complement existing intrusion prevention system for cellular mobile networks. By

1-4244-0086-4/05/$20.00 ©2005 IEEE.

comparing the different behaviors demonstrated by the authentic user and the adversary, the system can detect the potential misbehavior. In [1], Lin *et al.* also discussed the potential fraudulent usage in mobile telecommunications networks.

For most mobile users in cellular networks, movement patterns can be captured and modeled. For example, public transportation drivers demonstrate a high regularity in their daily mobility patterns. The mobility history of an authentic user can be learned and compared to the current movement pattern in order to identify intruders. It is obvious that there are a certain number of users such as taxi drivers who do not exhibit regular movement patterns. We do not expect our detection based on mobility patterns to be accurate for all users in all situations.

In [2], we proposed an EWMA-based anomaly detection algorithm. Similar strategies proposed in [2] has been used in credit card companies. For example, a customer will be called if the abnormal usage of his/her credit card is detected, such as the card being used at another country that is not the owners residence and the owner frequently visits. In this paper, we aim at providing a better optional service to end users as well as a useful administration tool to service providers.

The approach proposed in [2] only considers the same set of a user's movement patterns. It is also possible that for a single user who goes to work every weekday and vacation every weekend, the user exhibits the different location *uncertainty*, which should be measured and reflected into our anomaly detection algorithm. This should impact the selection of the detection threshold. That is, when the set of movement patterns changes, the corresponding detection threshold should also change. We need a mechanism to help us *adaptively* adjust the detection threshold in order to make the IDS more practical.

Our proposed approach requires the tracking of people's locations. It is a location tracking service that is based on the system tracking users locations. This will give rise to user's location privacy issues. Therefore, our system provides the user an option to turn off this service. Privacy concerns must be properly addressed before we can deploy this kind of service. It is worth noticing that location privacy issues have attracted much attention from the research community. Therefore, it is promising to integrate our proposed service with other existing location privacy protection schemes.

In a cellular mobile network, location updates and registrations usually happen when the user enters or leaves a *location area*. This is true no matter whether a user is making a phone call or not. Considering that a user demonstrates regular movement patterns, a user will report its location at the intersection of the routes and the perimeter of the location area. When a user is making phone calls while traversing cells, cell IDs are available to the system. Our proposed scheme can utilize this feature to construct the IDSs for cellular mobility networks. In the following, we use cell IDs as an example feature to illustrate our adaptive algorithm.

## III. Background

### A. Assumptions

We assume that there is a mobility database for each mobile user that describes his normal activities. The mobility database could be stored in the Home Location Register (HLR) and cannot be hacked. We assume mobile devices can be compromised and all secrets associated with the compromised devices are open to attackers. In this way, we do not need to assume tamper resistant technologies. We further assume most users have favorite or regular itineraries. This makes us viable to construct their normal profiles.

### B. LZ-based Anomaly Detection in Cellular Mobile Networks

The general strategy is illustrated in Fig. 1.

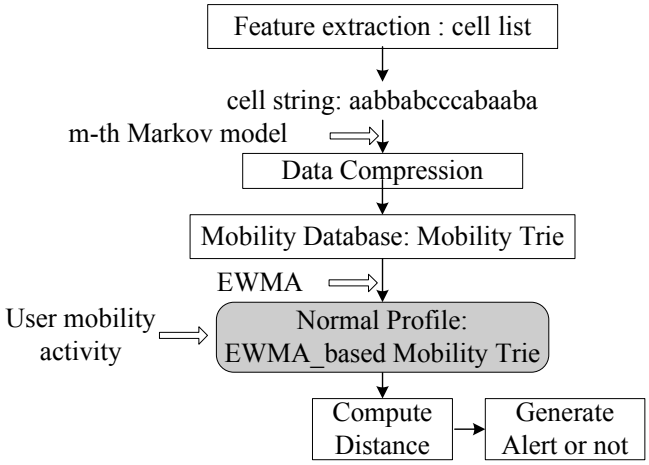


Fig. 1. Integrating EWMA into Mobility Trie Construction.

Basically, under the assumption that the user will have his own favorite itineraries, cell IDs traversed by each user is extracted to reflect the user's movement pattern. This makes the user mobility stable and results in a small alphabet. We apply a $m$-th Markov model to the extracted cell list and construct a mobility trie from the accumulative history of the user's movement pattern. A trie is a multiway tree with a path from the root to a unique node for each string represented in the tree. It could effectively store the parsed cell string.

The more recent the activity, the more weight it should be assigned when we construct the normal profile. Therefore, we apply the Exponential Weighted Moving Average (EWMA) [6] techniques to the mobility trie constructed. In this way, activities that happened long time ago are faded away exponentially. This modified mobility trie will serve as the normal profile of the user in the recent past. It reflects the *stationary* part of the users regular mobility pattern.

We then deploy a Prediction by Partial Matching (PPM) [5] scheme to calculate the probability predicted by the constructed EWMA-based normal profile. A context model, a *$m$-th Markov model*, is used to predict the next cell based on the previous *m* consecutive cells. The usage of PPM could take into consideration the trade-off between a too small *m* (results

in a poor prediction in the long run) and a too large *m*(results in the zero-frequency problem [5]).

In the detection phase, we extract the current activity of the mobile user and compute its probability predicted by the EWMA-based mobility trie. When the probability is less than a threshold $P_{thr}$ (a design parameter), the current activity is identified as anomalous. $P_{thr}$ is tuned in order to get a better tradeoff between the false positive rate and the detection rate.

## IV. Adaptive Anomaly Detection

EWMA-based mobility trie itself facilitates the differentiation between weekday and weekend routes because when the user changes its mobility patterns, say from weekday to weekend routes, the more recent the activities, the more weight they should have in the normal profile. The smoothing constant [6] in EWMA techniques plays an important role in determining how much weight the more recent activities should have. Basically the larger the smoothing constant is, the more weight they should have. Therefore, intuitively, the shorter the recent activities last, the larger the smoothing constant should be.

The EWMA-based approach only partially addresses the adaptation of normal profiles. In the followng, we detail our approach of how to tune the threshold for different users and different mobility levels.

### A. Feedback-based Approach

One simple approach to adjust the threshold is to apply the feedback principle. That is, based on the output of the detection algorithm (for example, in terms of detection rate and false positive rate), the system administrator can adaptively adjust the detection threshold in order to achieve the required performance. If the false positive rate is a more important metric, for example, when the system has been detected raising too many false alarms, the system administrator could lower the detection threshold correspondingly. However, in this approach, the decrease of the false positive rate is achieved at the risk of a decreased detection rate.

### B. Entropy-based Approach

We propose to use Shannon's entropy measure to identify the uncertainness of the up-to-date normal profile. Based on this, we could adjust the detection threshold correspondingly.

*1) Metric Selection:* The first step we need is to identify a metric that can effectively reflect the location *uncertainty*. In our case, it is the EWMA-based mobility trie. Shannon's *entropy* measure [4] is an ideal candidate for quantifying this uncertainty. Our previous work showed that for the non-adaptive mechanism, given a mobility level, the more varied the mobility pattern, the more dynamic the mobility trie. This motivates us to use *entropy* as a measure to reflect the dynamic level of the normal profile. The lower the uncertainty under the movement pattern, the richer the movement pattern is.

**Definition 1** Entropy: Suppose $X$ is a dataset, $C_x = \{C_x[1], C_x[2], \ldots, C_x[m]\}$ is a class set. Each data item of $X$ belongs to a class $x \in C_x[i]$. Then the entropy of $X$ related to this $|C_x|$-wise classification is defined as $H(X) = \sum_{i=1}^{m} -P_i lg P_i$, where $P_i$ is the probability of $x$ belonging to class $C_x[i]$.

*Entropy* can be interpreted as the number of bits required to encode the classification of a data item. It measures the uncertainty of a collection of data items. The lower the entropy, the more uniform the class distribution. If all data items belong to one class, then its *entropy* is 0, which means that no bits needs to be transmitted because the receiver knows that there is one class. The more varied the class distribution, the larger the *entropy*. When all of the data items are equally distributed over the $m$ classes, its *entropy* is $\lg m$. In the context of anomaly detection, entropy is a measure of the regularity of audit data.

**Definition 2** Conditional entropy: Suppose $X$ and $Y$ are two datasets, $C_x = \{C_x[1], C_x[2], \ldots, C_x[m]\}$ and $C_y = \{C_y[1], C_y[2], \ldots, C_y[n]\}$ are two class sets. Each data item of $X$ belongs to a class $x \in C_x[i]$ and each data item of $Y$ belongs to a class $y \in C_y[i]$. Then given $Y$ and $C_y$, the entropy of $X$ related to $C_x$ is defined as $H(X|Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} P_{ij} \lg \frac{1}{P_{i|j}}$, where $P_{ij}$ is the probability of $x \in C_x[i]$ and $y \in C_y[j]$, $P_{i|j}$ is the probability of $x \in C_x[i]$ given $y \in C_y[j]$.

*Conditional entropy* describes the uncertainness of $X$ given $Y$, i.e., it indicates the coefficiencies between $X$ and $Y$. The smaller the *conditional entropy*, the more correlated $X$ and $Y$. If $X$ can be determined by $Y$, $H(X|Y)$ is 0. In the context of anomaly detection, *conditional entropy* can be used to explore the temporal sequential characteristics of audit data due to the temporal nature of the system activities.

There have been some work that utilizes the entropy of the trie, and use it to improve the performance adaptively [3]. In [3], the authors use an information-theoretic framework and Shannon' entropy measure as the basis to compare user mobility models. In this way, they can track user mobility efficiently and reduce paging costs.

*2) Compute the Entropy of a Trie:* When we compute the entropy of the EWMA-based mobility trie, we apply a weighted scheme at different orders. Specifically, based on the order of different finite contexts of the mobility trie, we calculate conditional entropies respectively and assign them different weights. The larger the order, the larger the weight. The sum of these weighted entropies is used as the measurement for adjusting system detection threshold. Let's consider the string $aaababbbbbaabccddcbaaaaaacabbbabcbdcadbdbb$. By applying $LZ78$ algorithm [5], we obtain a trie as illustrated in Fig. 2.

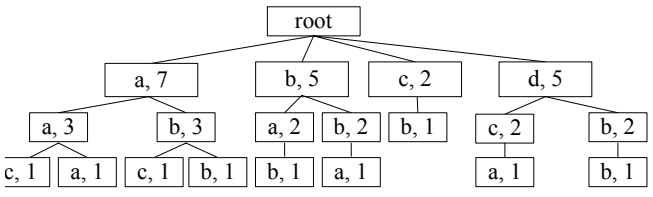


Fig. 2. An Example of EWMA-based Mobility Trie.

The maximum order $m$ and the corresponding weight $w_i$ are design parameters. In this example, let's assign 2 to $m$.

- Order-0 Model

$$\begin{aligned} & p(V1) \\ = \; & \frac{7}{19}lg\frac{19}{7} + \frac{5}{19}lg\frac{19}{5} + \frac{2}{19}lg\frac{19}{2} + \frac{5}{19}lg\frac{19}{5} \\ = \; & 1.88631. \end{aligned}$$

- Order-1 Model

$$\begin{aligned} & p(V2|V1) \\ = \; & \frac{7}{19}[(\frac{3}{6}lg\frac{6}{3}) \times 2] + \frac{5}{19}[(\frac{2}{4}lg\frac{4}{2}) \times 4] \\ = \; & 0.894737. \end{aligned}$$

- Order-2 Model

$$p(V3|V1V2) = \frac{3}{19}[(\frac{1}{2}lg\frac{2}{1}) \times 4] + 0 = 0.315789$$

When the context of a specific length is not found in the trie, we assign its conditional probability to 0. Further, we treat $0lg0$ as 0.

Generally, the larger the order, the larger the weight assigned to it, because context models with a larger order tends to be more accurate and should weight more in the current normal profile. If we assign 0.1, 0.2, and 0.7 to $w_1$, $w_2$, and $w_3$ respectively, the *weighted entropy* of the mobility trie in Fig. 2 can be calculated as:

$$\begin{aligned} & weighted\_entropy \\ = \; & w_1 \times H(V1) + w_2 \times H(V2|V1) + w_3 \times H(V3|V1V2) \\ = \; & 0.58863. \end{aligned}$$

*3) Adaptive Algorithm:* The proposed adaptive algorithm is illustrated in Fig. 3. After integrating a new string $s$ to a EWMA-based mobility trie, we compute its entropy. If the current entropy is larger than the previous one, we decrease the detection threshold by $\Delta$. Otherwise, we increase the detection threshold by $\Delta$. Here, $\Delta$ is a design parameter and is related to the mobility levels and mobility patterns.

After deciding the threshold, we can then apply it to the test trace and decide whether it is abnormal or not [2]. When the distance between the current trace and the EWMA-based mobility trie is less than the threshold, the current trace is normal. Otherwise, an alarm could be generated.

## V. Simulation Study

A graph resembling the cellular mobile network is used in our simulation study. On average, each cell is surrounded by another six cells. Assumption is made that in reality, people show different itinerary patterns in weekdays and weekends. Therefore, for a mobile user, we designate five different routes for weekday activities and three routes for weekend activities. The probability distribution of the user taking the weekday routes is 0.6, 0.2, 0.1, 0.05 and 0.05 respectively.

```
Initialize mobility database := null
LOOP
        Based on LZ algorithm, wait for a sequence s
        IF (The mobility trie of the mobile exists)
          IF (a path p corresponding to s is found)
            Add s to the mobility trie
            Using EWMA to modify the frequencies of nodes
          ELSE
            Create new nodes, and initialize their frequencies to λ
        ELSE
            1) Create a mobility trie := single sequence s
            2) Initialize the frequencies for every node in sequence
               s to λ
        Compute the entropy e1 of the EWMA-based mobility trie
        If (e1 > e)
        /* e is the entropy of the previous EWMA-based mobility trie */
            Decrease the detection threshold by Δ
        Else
            Increase the detection threshold by Δ
        e = e1;
FOREVER
```

Fig. 3. Adaptive EWMA-based Mechanims.

The probability distribution of weekend routes is 0.5, 0.3, and 0.2 respectively. Five mobility levels are considered to study the performance of the adaptive mechanisms. Specifically, we set the mobile speed to 20, 30, 40, 50, and 60 miles/hour respectively. Call durations are the same for all calls and exponentially distributed with mean value of 3 minutes. Given fixed call duration, the higher the mobility level, the more the cells traversed.

We apply a blended Markov model with $m$ set to 2. $w_0$, $w_1$, and $w_2$ are set to 0.1, 0.2, and 0.7 respectively. These values are used to calculate both the probability of the current activity and the entropy of the EWMA-based mobility trie.

### A. Performance Metrics

- *False Alarm Rate*: It is measured over normal itineraries. Suppose $m$ normal itineraries are measured, and $n$ of them are identified as abnormal, *false alarm rate* is defined as $n/m$.
- *Detection Rate*: It is measured over abnormal itineraries. Suppose $m$ abnormal itineraries are measured, and $n$ of them are detected, *detection rate* is defined as $n/m$.

### B. Simulation Results

For the false alarm rate of the non-adaptive mechanism, we apply the detection threshold constructed using weekday data and apply them to weekend data. Note that the normal profiles of weekday data are adapted to those of weekend because of the usage of EWMA techniques. For the false alarm rate of the adaptive mechanism, we use the algorithm illustrated in Fig. 3 to adaptively adjust the threshold. We measure the detection rate using the same methodology.

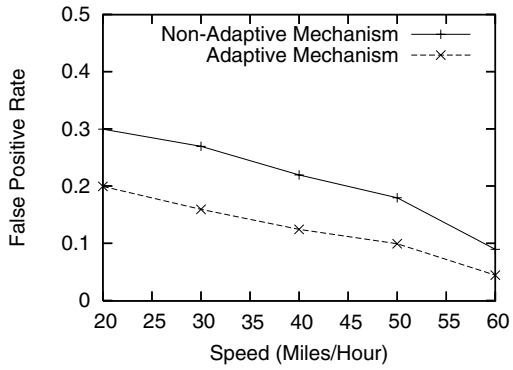


Fig. 4. False Alarm Rate at Different Mobility Levels.

*1) False Alarm Rate:* Simulation results of the false alarm rate are illustrated in Fig. 4. For both mechanisms, the false alarm rate decreases with the increase of the mobility. For a normal user who has traversed more cells with regular movement patterns, his itinerary will demonstrate more resemblance to his regular activities that is recorded in the mobility trie. Therefore, the probability normalized by the itinerary length is relatively stable. In the way, the false positives are reduced.

At the same mobility level, the false alarm rate of the adaptive mechanism is lower than that of the non-adaptive mechanism. Adaptive mechanisms take into consideration the changes of the user's mobility patterns and can adjust the detection threshold correspondingly. The online IDS can thus adapt itself to the environment better. False positives, which are the main concerns when we deploy any intrusion detection systems, can be reduced correspondingly.

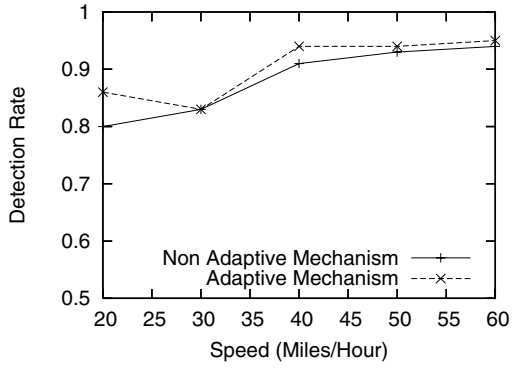


Fig. 5. Detection Rate at Different Mobility Levels.

*2) Detection rate:* Simulation results of the detection rate are illustrated Fig. 5. For both mechanisms, the detection rate increases with the increase of the mobility. With the increase of the mobility level, each user tends to have more cells traveled. Therefore, for a masquerader, his itinerary tends to deviate significantly from the normal profile. In this way, the detection rate is improved with the increase of mobility.

Detection rates with and without adaptive mechanisms do not show much difference. When the attacker demonstrates a significantly different route compared to normal profiles, the *adaptive* mechanism will not enable the abnormal changes to have a high probability predicted by the EWMA-based mobility trie. The same is true for non-adaptive mechanism. Therefore, for an abnormal route change, it will have a low probability under both the adaptive and non-adaptive mechanism. In this way, the detection rate cannot be improved.

To summarize, the main benefit of our adaptive mechanisms is to lower the false positive rate, while keeping roughly the same performance in terms of the detection rate.

## VI. Related Work

Two important intrusion detection techniques exist: *misuse detection* and *anomaly detection*. [7] presents a good taxonomy of existing technologies. Many research efforts have been devoted to different detection techniques. Most of them take into consideration domain specific knowledge.

Relatively few research efforts have been devoted to intrusion detection research of wireless networks. In [8], Samfat *et al.* proposed IDAMN (Intrusion Detection Architecture for Mobile Networks) that includes two algorithms to model the behavior of users in terms of both telephony activity and migration patterns. Buschkes *et al.* [9] presented a new model based on the Bayes decision rule and applied this rule to mobile user profiles.

## VII. Conclusions and Future Work

We investigate using entropy as a metric to integrate the *adaptive* threshold into the construction of mobility-based IDSs. Based on entropy, we can adaptively adjust the detection threshold. Simulation results demonstrated that our adaptive mechanisms can lower the false positive rate, while keeping roughly the same performance in terms of the detection rate.

We plan to further investigate better approaches to lower false positive rate. More features such as call activities will be accommodated into the system to make it suitable to all users.

## References

[1] Y.-B. Lin, M. Chen, and H. Rao, "Potential fraudulent usage in mobile telecommunications networks", *IEEE Transactions on Mobile Computing*, vol. 1, no. 2, 2002, pp. 123-131.

[2] B. Sun, F. Yu, K. Wu and VCM Leung, "Mobility-Based Anomaly Detection in Cellular Mobile Networks," *Proceedings of ACM Wireless Security (WiSe'04)* Philadelphia, PA, 2004, pp. 61-69.

[3] A. Bhattacharya, and S.K. Das, "LeZi-Update: An Information-Theoretic Framework for Personal Mobility Tracking in PCS Networks," ACM/Kluwer Journal on Wireless Networks, Vol. 8, No. 2-3, pp. 121-135, Mar.-May 2002.

[4] T. Cover and J. Thomas, *Elements of Information Theory,* John Wiley & Sons, 1991.

[5] T.C. Bell, J.G. Cleary, and I.H. Witten, *Text Compression*, Prentice-Hall Advanced Reference Series, Prentice-Hall, Englewood Cleffs, NJ 1990.

[6] R.A. Johnson and D.W. Wichern, *Applied Multivariate Statistical Analysis,* Upper Saddle River, NJ: Prentice Hall, 1998.

[7] H. Debar, M. Dacier, and A. Wespi, "A Revised Taxonomy for Intrusion-Detection Systems," Annales des T*é*l*é*com., vol. 55, 2000, pp. 361 - 378.

[8] D. Samfat, and R. Molva, "IDAMN: An Intrusion Detection Architecture for Mobile Networks," *IEEE Journal on Selected Areas in Communications,* vol. 15, no. 7 , pp. 1373-1380, Sept. 1997.

[9] R. Buschkes, D. Kesdogan, and P. Reichl, "How to Increase Security in Mobile Networks by Anomaly Detection," *1998 Computer Security Applications Conference,* Phoenix, AZ, USA, Dec. 1998, pp. 3-12.