Substantiating Timing and Behavioral Anomalies in Wireless LANs Using GCL

Elankayer Sithirasenan and Vallipuram Muthukkumarasamy
School of Information and Communication Technology, Griffith University, Gold Coast, Australia
Email: {e.sithirasenan, v.muthu}@griffith.edu.au

***Abstract*—With the increasing dependence on wireless LANs (WLANs), businesses, educational institutions and other organizations are in need of a reliable security mechanism. The latest security protocol, the IEEE 802.11i assures rigid security for WLANs with the support of IEEE 802.1x protocol for authentication, authorization and key distribution. Nevertheless, fresh security threats are emerging often to oust these new defense mechanisms. Further, many organizations based on superficial vendor literature, believe their wireless security is sufficient enough to prevent any unauthorized access. Having wide ranging options for security configurations, users are camouflaged into deep uncertainty. This volatile state of affairs has prevented many organizations from fully deploying WLANs for their secure communication needs, though WLANs may be cost effective and flexible. In this paper, we present a novel mechanism to detect and substantiate anomalies caused by both known and unknown security threats in WLANs. We monitor the wireless environment for timing and/or behavior anomalies during the security association process and use outlier based data association approaches to substantiate their legitimacy. The proposed concept was tested on our experimental setup. The results obtained from wireless hosts configured for EAP-LEAP, PEAP and TLS security association show high confidence for EAP group events.**

***Index Terms*—anomaly detection, intrusion detection, wireless security, group outliers, IEEE 802.11, security threat validation**

I. Introduction

IEEE 802.11i [1] provides mutual authentication, key management, and data confidentiality protocols that may execute concurrently over a network in which other protocols are also used. On the assumption of upgrading the hardware, 802.11i defines CCMP that provides strong confidentiality, integrity, and replay protection. In addition, an authentication process, combining 802.1x port-based authentication and key management procedures, is performed to mutually authenticate the devices and generate a fresh session key for data transmission. Since 802.11i promises to be in the right direction for wireless security, it should be able to prevent an adversary from any attacks even if the adversary has powerful equipment and techniques for breaking into the system. In other words, an implementation of 802.11i protocol in a WLAN should provide sufficient data confidentiality, integrity, and mutual authentication.

Although this newly introduced security standard provides effective measures to protect the wireless networks from confidentiality and integrity threats, their reliance on authenticity and availability are still a major concern. Further, there are many wireless installations, where appropriate security mechanisms are neither used nor implemented effectively [2]. Hence in addition to improving the existing security mechanisms we also need other protection techniques to safeguard the wireless environment from new security threats. Effective use of the wireless network will only be possible if security threats are mitigated, preventing any potential catastrophe. In this respect, tracking the management frames during IEEE 802.11i security association and effectively analyzing them may reveal vital information about impending threats to the wireless environment.

In this regard our proposed system exploits the use of management frame timings and the behavior of wireless hosts, during the security association process. Unusual timing values exhibited by wireless stations may be the start of an intended security breach. Legitimate wireless hosts attempting to connect to an authorized authenticator usually demonstrates a particular practical behavior. Tracking and profiling such behavior of all stations in the wireless environment may present some useful information for detecting misbehaving stations. Therefore, tracking the management frames and effectively analyzing them for timing and/or behavior anomalies is expected to enhance the ability of our proposed system to discover security related issues in advance.

It is explicable that the behavior of a wireless access point varies from one vendor to another and depends on many factors. Hence, maintaining a behavior profile for every participating wireless host - access point duo will be more effective rather than having a generic profile for the environment. Throughput and resistance to one single node flooding depends on the actual implementation and operative system of the access point. A variation in timing during the association process for a given station depends on factors such as type of transversing traffic (number of packets and their sizes), UDP vs TCP traffic, number of traffic flows etc. Therefore, maintaining dynamic profiles

Manuscript received March 14, 2008; revised November 10, 2008. This work was supported in part by the Institute for Integrated and Intelligent Systems, Griffith University, Australia.

Elankayer Sithirasenan is with the Institute for Integrated and Intelligent Systems, Griffith University, Brisbane, Australia (phone: +61-7-555-28728; e-mail: E.Sithirasenan@griffith.edu. au).

Vallipuram Muthukkumarasamy is with the School of Information and Communication Technology, Faculty of Engineering, Griffith University, Gold Coast, Australia (phone: +61-7-555-28256; e-mail: v.muthu@griffith.edu.au).

will be advantageous than having static profiles. Hence, we employ an adaptive technique to update the timing and behavioral profiles depending on the nature of the operative environment.

Trying to model what is "normal" in a wireless environment is not an easy task, hidden nodes will trigger retransmissions and heavy real-time UDP traffic will have an impact on how packet queues are handled and hence timing of management frames. Therefore, detecting a security threat in the wireless environment, merely based on timing profiles can lead to large number of false positives. Hence, we explore the use of outlier based data association techniques to substantiate the security threats and thereby reducing the number of false positives.

The early warning system (EWS) [3] used in this research includes a packet capturing module, an event engine, an anomaly detection module and an intrusion prevention module. The packet capturing module monitors and captures the raw wireless traces. The event engine extracts the management frames associated with the IEEE 802.11i security association. The anomaly detection module uses both timing and/or behavior analysis to detect anomalies. The detected anomalies are validated by the intrusion prevention module using outlier based data association techniques. During IEEE 802.11i security association there can be different types of events that can be grouped depending on the authentication mechanism used. Hence, based on our previous work reported in [4], in this paper we propose a novel measure called the group confidence level (GCL) for estimating the cohesiveness of group events. The EWS uses the GCL to substantiate anomalies in group events during EAP-LEAP [5], PEAP [6] and TLS [7] authentication schemes. The experimental results show that EAP group events exhibiting high confidence than the IEEE 802.11 group events in all three authentication mechanisms.

This paper is organized as follows. Section II gives a brief overview of related work on intrusion and anomaly detection. Section III introduces the RSN framework. In Section IV the concept of OLAP and data views are explained. The matrices used to substantiate the security threats are derived in Section V. Details of our proposed anomaly detection system is illustrated in Section VI. Experimental results are presented in Section VII. The results are discussed and analyzed in Section VIII. Finally, Section IX concludes the paper.

II. Related Work

Our EWS uses anomaly detection techniques to discover anomalies in IEEE 802.11i wireless environment and thereafter uses outlier based techniques to substantiate the detected anomalies. In this view we first studied the reported security issues in IEEE 802.11i wireless environment. We then examined the different techniques, including anomaly based techniques used to detect security threats in both wired and wireless networking environment. Finally, we reviewed a number of data mining techniques that are used to detect outliers in large data sets.

Analysis of IEEE 802.11i by Sithirasenan et al. [8] identifies a number of weaknesses in the standard together with some solutions from the software implementation perspectives. A similar analysis by He et al. [9] on IEEE 802.11i wireless networking further highlights the weaknesses of the standard. They have discussed the possibilities of several attacks on poorly configured 802.11i networks. Further, they state that although the new security standard offers sufficient protection to the wireless environment it is up to the implementer to ensure that all issues are addressed and the appropriate security measures are deployed. For instance, a single misconfigured station could lead the way for a attack and expose the organizational network. Lynn et al. [10] discuss that if no authentication mechanisms are implemented an adversary could establish two separate connections to the supplicant and the authenticator to construct a Man-in-the-Middle (MitM) attack. Furthermore, if mutual authentication mechanism is not appropriately implemented an adversary will be able to launch a MitM attack and learn the Pairwise Master Key (PMK) as illustrated by Asokan et al. [11]. Although these vulnerabilities are not directly connected with 802.11i, any implementer of 802.11i needs to consider these problems warily and keep monitoring the wireless hosts to guarantee proper integration of the security mechanisms.

Barbeau et al. have analyzed impersonation attacks in future wireless and mobile networks [12]. They have performed a risk analysis on the possible types of impersonation attacks in wireless networks. Summing up their findings they stress that the risk of impersonation in wireless networks is critical since the threat can be materialized into several forms of attacks. Hence they highlight the need for effective countermeasures to address the threats.

The countermeasure proposed by Barbeau et al. [12] includes the use of Radio Frequency Fingerprinting (RFF) and User Mobility Profiles (UMP) for Anomaly Based Intrusion Detection (ABID). At the same time the use of device and user profiles to detect anomalies has been studied by Hall et al. [13]. They incorporate RFF into an ABID system to associate a MAC-address of a device with the corresponding transceiver profile. Hence, if an observed transceiver print from a claimed MAC-address matches the corresponding transceiver profile, then the MAC-address has not been spoofed. They make use of this idea to verify the credibility of their UMP based ABID system. They discussed enhancing the capability of their system by supplementing existing user and device-based profiles, with those based on mobility. However, this system is more suitable for addressing the problem of stolen cell phones, given that the mobility behavior of the thief and the user are likely to be different. In the case of wireless networks the attacker needs to be in the same domain as the user to carry out an attack. Therefore the use of mobility profiles will not be suitable for intrusion

detection in infrastructure WLANs.

Further enhancing their work, Gill et al. studied the anomaly detection response using correlation techniques in their later study [14]. Here, having detected an anomaly from either mechanism; Round Trip Timing (RTT) or Received Signal Strength (RSS), they wait for a confirmation from the other. For example, if the RSS detection algorithm registers an abrupt spike in RSS value for a particular MAC address, the correlation engine would register this and wait for the next RTS-CTS event for this MAC address and check the results of the RTT detection algorithm. If both algorithms register an alert for that MAC address, then an alarm would be raised. From a number of experiments performed they claim nil false negative and a low number of false positives. Hence they claim anomaly detection systems could not rely only on one observation and that multiple observations is necessary together with effective correlation techniques to verify the legitimacy of each alarms raised. However, this technique is not suitable in every situation since the legitimate user and the intruder can exhibit the same RSS and RTT.

Maxion et al. [15] apply benchmarking to prove that differences in data regularities influence anomaly detector performance, and such differences are found in natural environments. All anomaly-detection algorithms operate on data captured from some kind of computing domain or environment. Embedded in each type of environment is a particular structuring of the data that is a function of the environment itself. The researchers provide quantitative results of running an anomaly detector on various data sets containing different structure. The results on different data regularities proved data consistency influences the anomaly detector performance. Hence they emphasize the need for anomaly detection systems that can handle differences in data regularities effectively.

Most data mining based intrusion detection systems detect unusual events or anomalous behavior in networks effectively [16]–[19]. However, the major problem is validating the legitimacy of abnormal events. Almost all of the work found in the literature present techniques to detect abnormal or rare events in networks. However, having found an anomalous or rare event how to validate it as legitimate or illegitimate is the major challenge yet to be addressed. In this view, discovery of outliers to extract a few data objects with abnormal behavior patterns, which are more imperative than common patterns in some cases, can be of practical significance in intrusion detection systems, credit fraud detection, etc [20]. Hence, in this study we investigate the use of outlier mining techniques to address the problem of validating the legitimacy of abnormal events.

III. Robust Security Network

The IEEE 802.11i standard [1] defines two classes of security framework for IEEE 802.11 WLANs: RSN (Robust Security Network) and pre-RSN. A station is called RSN capable equipment if it is capable of creating

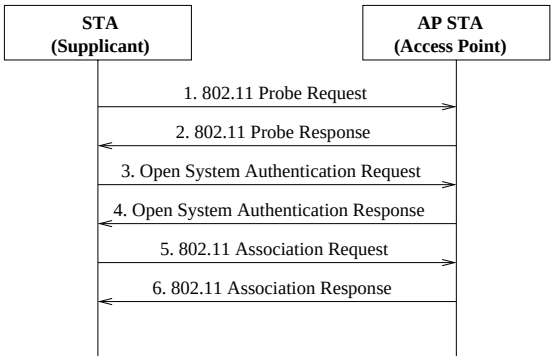

Fig. 1. RSNA - Discovery Phase

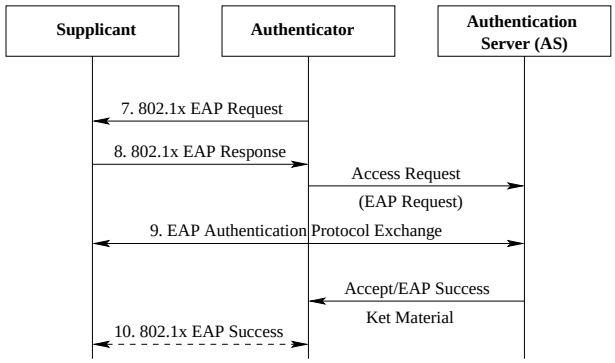

Fig. 2. RSNA - Authentication Phase

RSN Associations (RSNA), otherwise, it is a pre-RSN equipment. In an Extended Service Set (ESS), during the RSNA a number of messages are exchanged between the supplicant (STA) and the authenticator (AP). The network that only allows RSNA with RSN-capable equipments is called a RSN security framework. The major difference between RSNA and pre-RSNA is the 4-way handshake. If the 4-way handshake is not included in the authentication / association procedures, stations are said to use pre-RSNA.

Figure 1 shows the first phase in RSNA - the discovery phase. During this phase both the STA and the AP agree on a common security policy. Flows 1-6 are the IEEE 802.11 [21] association process prior to attaching to the AP. During this process, security information and capabilities are negotiated using the RSN Information Element (IE). The Authentication in flows 3 and 4 refer to the IEEE 802.11 open system authentication. On successful completion of the discovery phase the AP initiates the authentication phase by starting the IEEE 802.1X [22] authentication as shown in Figure 2. If the STA and the authentication server authenticate each other successfully, both of them independently generate a Pairwise Master Key (PMK). Depending on the type of authentication mutually agreed between the STA and the authentication server, there could be several messages exchanged between them (flow 9) before the PMK is generated. The authentication server then transmits the PMK to the AP through a secure channel (for example, IPsec or TLS).

The next phase in the RSNA is the key distribution phase as illustrated in Figure 3. The PMK generated during the authentication phase is used to derive and

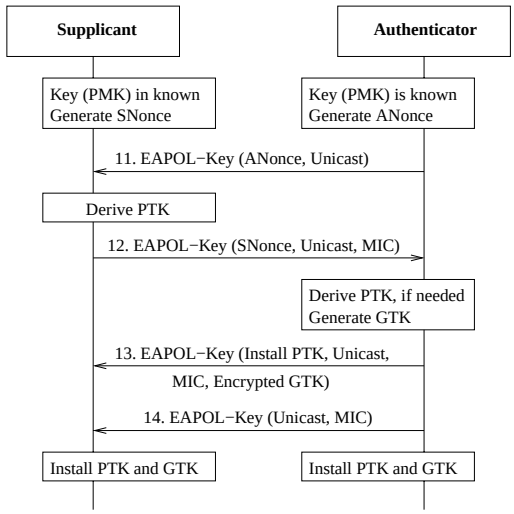

Fig. 3. RSNA - Key Distribution Phase

verify a Pairwise Transient Key (PTK), guaranteeing fresh session key between the STA and the AP. This is called the 4-way handshake phase as shown by flows 11 - 14. Next, the group key handshake is initiated. The group key handshake is used to generate and refresh the Groupwise Transient Key (GTK), which is shared between a group of STAs and APs. Using this key, broadcast and multi-cast messages are securely exchanged in the air.

The anomaly detection modules in our Early Waring System (EWS) tracks all messages discussed above to make an assessment on the level and nature of an anomaly. The software model of the IEEE 802.11i security architecture, which was developed and analyzed by [8] is used as the prototype for detecting behavioral anomalies. This model has been formally verified for consistency and completeness [23]. In the next section we introduce partial data cubes and the concept of surrogate views.

IV. OLAP AND DATA VIEWS

A popular model for Online Analytical Processing (OLAP) applications is the multidimensional database also known as the data cube [24]. A data cube consists of two kinds of attributes: dimensions and measures. The set of dimensions may consist of elements like IP addresses, port addresses, event identities etc. that together form a key. Measures are typically numeric elements like packet size, duration etc. Data cube queries represent an important class of OLAP queries in decision support systems. The pre-computation of the different views of a data cube (i.e., the forming of aggregates for every combination of GROUP BY attributes) is critical to improving the response time of the queries [25]. However, in many cases not all views are needed for decision making, therefore it is advantageous to use only selected views. Such cubes with only selected views are referred to as partial data cubes [26].

Figure 4 shows a data cube with thirty two views made up of five attributes. In an actual application although the number of attributes can be many, the required number of

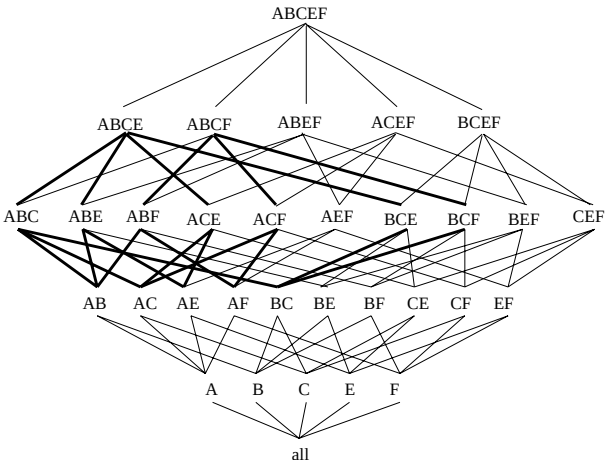

Fig. 4. Data Cube Views

TABLE I
DATA CUBE ATTRIBUTES

Reference	Attribute	Type
A	Source ID	Dimension
B	Destination ID	Dimension
C	Event ID	Dimension
D	Time of Day	Dimension
E	Day of Week	Dimension
F	Protocol ID	Dimension
G	Cipher ID	Dimension
-	Event count	Measure

views may be vary few depending on the requirements. For example, if we assume that attribute “A” and “B” represent the identities of APs and STAs respectively, then for data association analysis between the APs and STAs we will consider only those views that consists of attributes “A” and “B”, i.e. views “ABCD”, “ABCE”, “ABDE”, “ABC”, “ABD”, “ABE” and “AB”. Furthermore, in some cases we may also need to establish associations with child views to further strengthen our validation (see views connected by dashed lines in Figure 4). In this manner, for intrusion detection purposes, we selected the appropriate attributes bearing in mind the various security threats on the wireless networking environment.

Table I presents the list of attributes that are stored in the data cube for our analysis. Since we are concerned with wireless traces during RSNA, we store only those attributes that are necessary for this purpose. In this respect the identities (Source and Destination IDs) of the two communicating hosts, the current message (event) passed, the time during the message is passed, the day on which the message is passed, the protocol used and the cipher used are stored in the data cube.

The attribute “Event ID” can take values 0 to 56, representing the 57 different messages exchanged during an RSNA [5]–[7], [27]. The attribute Time of Day represents one of twenty four time periods during the day. It starts from value 0 (for time period midnight - 1.00 am) and continues up to value 23 (for time period 11.00 pm - midnight). Day of Week is from Monday to Sunday. Protocol ID ranges from 0 to 5, representing the six protocols; IEEE802.11, EAP, TLS, PEAP, LEAP and EAPOL considered in our study. The Cipher ID is 0 for

TABLE II
SELECTED VIEWS

ABCDFG	BCDFG	ABCFG	BDF	BDFG
ABDFG	AF	BF	BEF	ABDE
ADE	BCFG	AC	ABEF	ACE
AE	ACDEFG	ACD	ABEG	BDG
BCEFG	AB	BDEF	ACEFG	ABF
ABDEF	ABCE	ADG	ACDE	BDEG
ABC	BE	ABCDEFG	AEF	ABDEG
ABCEFG	BEG	BD	ADF	AD
AG	ABD	BCE	AEG	BC
ADEF	ABDG	ABDF	ABG	ACFG
BDE	BG	BCDEFG	ABE	ABCDE
ADEG	A	BCDE	BCD	B
ACDFG	ABCD			

TABLE III
VIEWS ASSOCIATED WITH SECURITY THREATS

No.	Threat	Attributes
1	Replay attack	ABCDE, ABCE, ABCD, ABC, ACD, ACE, BCD, BCE, AC, BC
2	Message deletion	same as above
3	Masquerading and Malicious AP	BCDEFG, BCFG, BCDF, BCDG, BCEF, BCEG, BCD, BCE, BCF, BCG, BC, BD, BE, BF, BG
4	Session Hijack	ABCFG, ABCF, ABCG, ABC ABF, ABG, AB
5	Man-In-The-Middle (MitM)	same as above
6	Denial of Service (DoS)	ABC, AB

AES/CCMP, 1 for TKIP and 2 for WEP. In addition to these attributes we can also store information related to the network identities and traffic related information if desired. The examples considered in our study are mainly focusing on wireless networks that adopt IEEE 802.11i security mechanism.

Having decided on the attributes, next, we established the views that are necessary to substantiate the security threats. Table II shows all of the views considered in our analysis. As mentioned earlier, almost all of the views selected include either one or both attributes A and B. Using these views we can query the data cube on any of the associations related to access points and/or hosts. However, since we are interested in substantiating security threats it is important to establish a relationship between these views and security threats.

Table III shows the relationship between data cube views and some common security threats. This table was established considering each threat and identifying the attributes that are associated with the threat. Having identified the attributes we then categorized the views that are vital for substantiating the security threat. For example in the case of Threat 1 - Replay Attack, we need to track the source and destination of every message in addition to the event identity. Hence, attributes A, B and C are important. However, attributes D (Time of Day) and E (Protocol ID) can be used to further refine the analysis. Therefore, to substantiate this threat we consider all views that are associated with these attributes. These same attributes will be sufficient to substantiate Threat 2 - Message Deletion as well.

In the case of Threat 3 - Masquerading and Malicious AP, we considered attributes B (Destination ID) and C (Event ID), since we need to track messages directed towards the Malicious AP. For these type of threats attributes F (Protocol ID) and G (Cipher ID) are also significant. Hence, views BCFG, BCF and BCG, BF and BG will be important.

Threat 4 - Session Hijack is a more advanced attack, where the association between a legitimate station and the access point is hijacked by an illegitimate user. In this case the illegitimate user will force a channel change with the access point/station and masquerade as a legitimate access point/station. Hence, in this kind of a threat we need to track the source, destination, event, protocol and the cipher of the messages exchanged. By tracking the protocol and cipher we can establish whether the illegitimate session establishes a different kind of association. Therefore, for this type of a threat, views associated with attributes A, B, C, F and G are are considered.

Threat 5 - Man-In-The-Middle attack is not very much significant in the context of effective confidentiality measures. However, if a MitM attack turns into a session hijack attack then our detection mechanism can be of use. Hence, we can make use of the same attributes as for session hijack attack. Furthermore, if a MitM is actively participating in the communication between the two legitimate hosts, then the timing anomaly detector will be able to detect some form of timing anomaly. But, it will again be the intrusion prevention module that will substantiate the anomaly.

For Threat 6 - Denial-of-Service attack we consider the events associated with the source and the destination. Hence, attributes A, B and C will be considered to substantiate this attack.

The rationale for using data cubes is (i) we can readily utilize the advantages of OLAP, such as systematic storage of historical data and fast querying (ii) by linking the threat to the data cube views we can provide a suggestion about the type of threat and the possible solutions and (iii) the scaling of the entire system to meet ever growing needs.

V. QUANTIFYING SECURITY THREATS

In order to substantiate the security threats (using the data cube views) we established the notion of confidence level using principles of information theory. The following derivation describes how the frequency measure stored in data cube views is used to establish the confidence level of an anomaly.

Let $A_1, A_2, \ldots, A_m$ be the m attributes considered in our study, and $D_1, D_2, \ldots, D_m$ be their domains respectively. Let $x_k = \{B_1, \ldots, B_k\}$ be a subset of $\{A_1, A_2, \ldots, A_m\}$ with dimension k, where $k \leq m$. If $R_i(k)$ is the *relative frequency* for itemset x_k with respect to attribute B_i, then

$$R_i(k) = \frac{freq\ (B_1, B_2, \ldots, B_k)}{freq\ (B_1, \ldots, B_{i-1}, B_{i+1}, \ldots, B_k)} \qquad (1)$$

In the above equation $freq\ (B_1, B_2, \ldots, B_k)$ represents the frequency count of itemset x_k and

TABLE IV
RISK VS CONFIDENCE

Risk Level	Confidence Level
Trivial	$\geq$ 95.00%
Low	85.00% - 94.99%
Moderate	75.00% - 84.99%
High	65.00% - 74.99%
Major	$\leq$ 65.00%

$freq\ (B_1, \ldots, B_{i-1}, B_{i+1}, \ldots, B_k)$ represents the frequency count of its parent. The relative frequency in this context is said to be the support of the itemset with respect to its parent view. We next define the *normalized support* for itemset x_k as follows

$$N_i(k) = -R_i(k) \log[R_i(k)] \tag{2}$$

Now, if itemset x_k is assumed to cover n possible disjoint events related to a specific security association, such that

$$\sum_{i=1}^{n} prob(R_i(k)) = 1 \tag{3}$$

In equation 3, $prob(R_i(k))$ represents the probability of the relative frequency for itemset x_k with respect to B_i. The task of measuring the confidence level involves translating the $apriori$ probabilities to a single real number representing the units of uncertainty, which we define as the *uncertainty level* $H(k)$.

$$H(k) = -\sum_{i=1}^{n} prob(R_i(k)) log[prob(R_i(k))] \tag{4}$$

From this equation we can deduce the group confidence level for a normally behaving host with n disjoint events as $log(n)$. Once the uncertainty level is established, we could compute the group confidence level (GCL) of the security association represented by itemset x_k. This is defined as the ratio between the abnormal behavior and the normal behavior during a similar security association process. Hence, the *group confidence level* $C(k)$ is given by the following equation as a percentage.

$$C(k) = \left[\frac{H(k)}{log(n)}\right] * 100\% \tag{5}$$

The group confidence level $C(k))$, thus obtained could be used as a scale to indicate the threat level of the respective association.

Table IV classifies the risk levels, as an example, in relation to the group confidence level values. This classification was made on the assumption that the number of abnormal events are significantly high. However, the threshold levels at which we set the risk can vary from one environment to another. Hence proper tuning procedures needs to be adopted to fix the right confidence levels for each environment. This method of establishing the threat levels will be effective mainly when the number of abnormal events are relatively high. Such situations can normally arise during most common security breaches.

To be more specific, consider a wireless host that always connects to a particular access point with a specific protocol, during week days. Assuming that the credentials of this wireless host have been compromised, an intruder can use them to gain access to the network. If the intruder connects to the network during weekends and/or any other unusual times, analyzing the network traces in the time domain will highlight the unusual behavior of the host. Hence, the information that this wireless host is connecting during unusual times will be of significance. However, the legitimate user may also have seldom used the network during weekends. Therefore, in such situations we need even more fine mechanism to differentiate the legitimate user from the intruder.

In such a scenario, examining the association details of this particular host in a different view may provide useful indicators to differentiate the legitimate user from the illegitimate user. For example, as an organizational policy the legitimate users may always use a particular protocol to communicate with the access point. Thus, analyzing the network traces in the protocol domain may present with abnormal conditions, if the intruder uses a different protocol to communicate with the same access point. This type of analysis is useful in session hijack attacks where rouge access points forcing misconfigured legitimate hosts to associate with them. In order to substantiate this type of security breaches we propose the notion of mutual outliers between different views of data.

VI. THE EARLY WARNING SYSTEM (EWS)

The EWS [3] includes a packet capturing module, an event engine, a timing anomaly detection module, a behavioral anomaly detection module, an intrusion prevention module and a data mining engine. The intrusion prevention module is the main component of our system with the ability of processing outlier based data association requests in real time. Our system has multiple levels of defense offering improved reliability for anomaly detection. The first level of defense is the detection of timing anomalies followed by the discovery of behavioral anomalies. If an event is detected with either one or both of the anomalies a third level of defense is triggered to validate the legitimacy of the event based on different views of data. Since our system needs to search enormous amounts of historical data (in real time) we use parallel processing techniques to query the database. In the following section we describe the process of anomaly detection in detail.

As shown in Figure 5 anomaly filters detect anomalies that are outside the theoretical or practical behavior region of a protocol. Anomalies in these regions result in a large number of false positives in wireless networks due to the inherent qualities of the wireless environment. Anomaly detection systems that operate within this region do not effectively detect anomalies that follow theoretical or practical behavior of the protocol. Anomalies outside the theoretical or practical behavior regions occur often and hence are usually easy to detect and substantiate. On the

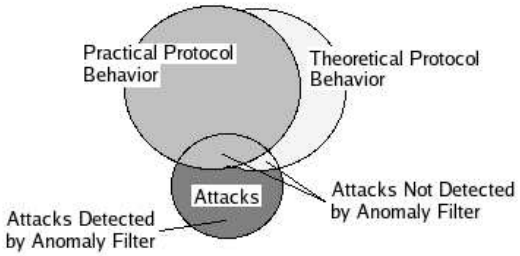

Fig. 5. Anomaly Filter

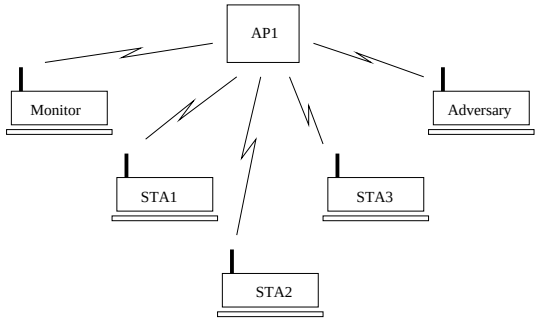

Fig. 6. Test Wireless Environment

other hand, anomalies within the theoretical or practical behavior regions are rare and lead to false negative reporting. Detecting these anomalies is challenging and requires analyzing the protocol behavior from different perspectives. Our proposed system detects such anomalies using timing inconsistencies and then substantiates the anomaly by analyzing its behavior from different view points. A wireless host which behaves normally from one view point can behave differently from another view point.

The first stage in our anomaly detection process is the examination of timing anomalies. This is achieved by maintaining a timing profile for every participating host in the wireless network. For example, we maintain the mean and the standard deviation for each round trip message transfer. A timing anomaly is triggered when a host exhibits an abnormal timing value during a message exchange. Further, a normally behaving host has to traverse all of the legitimate states of the RSNA process. If anomalies occur there can be situations where hosts fall into illegitimate states and fail to match the projected behavior. The EWS does not instantly classify anomalies as illegitimate, instead, anomalous events are forwarded to the intrusion prevention module for further processing.

The important part of the EWS is the intrusion prevention module. This module has to arrive at important decisions to verify the legitimacy of the anomalies discovered by the previous modules. The intrusion prevention module executes a number of data association analysis to decide whether anomalies detected by the anomaly modules are legitimate or not. This is achieved by analyzing the detected anomaly from different view points. Wireless attacks such as ”Session Hijack“ or ”Man-In-The-Middle” do exactly follow the protocol but can exhibit differences in timings. Hence, such attacks cannot be substantiated merely by looking at just one characteristic of the wireless environment. Instead, analyzing anomalies caused by these attacks from different view points can reveal better results.

VII. Anomaly Detection

Our main data processing unit is made up of a small Beowulf cluster. It is used to capture and analyze the wireless traces. The cluster has four nodes including the head node. The head node is a Pentium 4 machine with 1GB internal memory and 120GB secondary storage. The back nodes are all Pentium III machines with 512MB internal memory and 40GB secondary storage. The head node also includes a Dlink AG530 wireless adapter configured to capture wireless network traffic in promiscuous mode. The captured packets are passed to the event engine for further processing. One of the back nodes runs the RADIUS server software and acts as the authentication server. Our present setup has only one monitoring device. However, in a practical situation there may be a number of monitors talking to the central intrusion prevention module. Thus a distributed monitoring setup would further enhance the capabilities of our EWS.

The advantage of using a Beowulf cluster in our system is its scalability and performance. Depending on the size of the networks and the amount of wireless traces to be processed the Beowulf cluster could be expanded from one node to many nodes without any additional software changes. Further, querying of different views is made faster, since individual views are held on a number of different nodes enabling parallel processing. Thus, delivering better response times.

Using this data processing unit we captured and analyzed large number of wireless traces from our test wireless environment shown in Figure 6. Here, STA1 and STA2 are both Linux machines and STA3 is a Windows XP machine. STA1 is configured for EAP-LEAP authentication, STA2 is configured for EAP-PEAP authentication and STA3 for EAP-TLS authentication. The Monitor is part of the data processing unit which captures wireless network traffic in promiscuous mode. The Adversary is capable of introducing various anomalies into the wireless network. During the experiments were carried the test setup was exposed to the normal wireless environment of the university attracting traffic from various unspecified sources.

Using this setup a number of experiments were carried out to validate the proposed EWS. Two different types of security threats; Denial-of-Service (DoS) attack and Replay attack were executed in order to verify the detection capabilities of Timing and Behavior Anomaly modules. The proposed substantiating mechanism was used to quantify the security threats.

TABLE V
EAP-LEAP TIMING PROFILE

Event	Min (ms)	Max (ms)	Mean (ms)	StDev
OPEN AUTHENTICATION	0.11	13.27	1.22	2.79
OPEN ASSOCIATION	0.09	14.37	2.93	4.48
EAP 1	0.06	17.45	4.15	5.87
EAP 2 (PEAP)	0.06	11.84	2.37	2.99
EAP 3 (LEAP)	0.08	26.42	6.00	8.32
EAP 4 (LEAP)	2.57	41.85	17.90	9.75
EAP Key Exchange	5.19	38.08	13.58	7.78
Overall	90.41	205.07	122.98	27.26

TABLE VI
EAP-PEAP TIMING PROFILE

Event	Min (ms)	Max (ms)	Mean (ms)	StDev
OPEN AUTHENTICATION	0.38	1.39	0.62	0.29
OPEN ASSOCIATION	0.09	32.42	4.22	8.02
EAP 1	0.06	17.93	2.54	4.47
EAP 2 (PEAP)	0.08	13.04	3.26	3.78
EAP 3 (PEAP)	0.05	5.31	1.05	1.55
EAP 4 (PEAP)	0.06	32.56	6.55	7.78
EAP 5 (PEAP)	0.06	16.24	2.61	4.33
EAP 6 (PEAP)	0.05	17.79	3.19	5.43
EAP 7 (PEAP)	0.05	19.14	4.99	5.04
EAP 8 (PEAP)	0.07	19.74	2.44	4.61
EAP 9 (PEAP)	0.07	25.05	3.91	8.16
EAP Key Exchange	0.16	19.46	7.68	7.39
Overall	113.01	334.23	235.88	50.09

A. Timing Anomaly

In order to study the behavior of wireless hosts during different EAP type specific authentication process, we collected 802.11 management traces from different wireless hosts configured to authenticate using LEAP, PEAP and TLS authentication mechanisms. In this case the same access point was used with the three different stations ST1, STA2 and STA3.

Tables V and VI show the timing profiles for stations STA1 and STA2 respectively. These results were obtained from over forty five trials conducted on our experimental setup. The timing profiles show the allowable timings to complete a particular round trip event during different type specific authentication processes: EAP-LEAP for STA1 and EAP-PEAP for STA2. The round trip event is considered to be the completion of two messages; a request and the corresponding response. The $Mean$ timing values together with the standard deviation ($StDev$) is used to determine the range of timings allowed for the round trip event during normal operation. We have approximated the round trip timing values to a normal distribution. Although, the wireless hosts demonstrate a very small Min time and a very high Max time, the timing anomaly module considers about 68% (within one $StDev$ away from the $Mean$) of values drawn from the standard normal distribution as normal.

Table VII shows the timing profile for station STA3. In this case the timing profile gives the allowable timings during EAP-TLS authentication process. Experiments show that out of the timing values for LEAP, PEAP and TLS authentication, TLS method demonstrates a more close to normal distribution with consistent timing values. Hence, maintaining such profiles for every station and access point duo can reveal any timing anomalies that may arise due to abnormal conditions between two wireless host pairs.

TABLE VII
EAP-TLS TIMING PROFILE

Event	Min (ms)	Max (ms)	Mean (ms)	StDev
OPEN AUTHENTICATION	0.33	0.47	0.41	0.03
OPEN ASSOCIATION	0.49	1.26	0.66	0.21
EAP 1	3.81	9.46	5.19	1.21
EAP 2	4.66	7.95	6.04	0.82
EAP 3 (PEAP)	0.10	8.39	2.30	2.11
EAP 4 (TLS)	6.51	16.37	9.76	2.79
EAP 5 (TLS)	0.06	6.47	2.32	2.66
EAP 6 (TLS)	7.64	17.04	14.82	2.33
EAP 7 (TLS)	0.08	12.30	5.67	4.52
EAP Key Exchange	13.92	21.13	16.80	1.74
Overall	147.94	220.25	176.75	19.23

TABLE VIII
EAP-LEAP TIMINGS DURING DOS ATTACK

Event	Time (ms)
OPEN AUTHENTICATION	0.41
OPEN ASSOCIATION	0.16
EAP 1	5.65
EAP 2 (PEAP)	3.04
EAP 3 (LEAP)	2.29
EAP 4 (LEAP)	36.55
EAP Key Exchange	9.11
Overall	157.00

Next, we injected some abnormal management frames forcing the wireless hosts to deauthenticate and reassociate frequently. This was done by injecting "Deauthentication" and "Association Request" frames using the "airreplay" [28] tool.

The timing values shown in Table VIII is the minimum of several RSNA associations during a DoS attack. Here neither the access point nor the station was targeted. However we injected several "Deauthentication" frames with a fake MAC address jamming the medium. During this period we initiated a EAP-LEAP authentication process and the timings were noted. To make sure the timing anomaly module detects every possible anomaly, we considered the minimum timings. From the timing measurements in Table VIII it can be seen that although the timings values for "EAP 1, 2 and 3" messages are comparable to the normal timings, "EAP 4" and the "Key Exchange" messages show significantly high timing values resulting in high overall timing. Although, these timing values fall within the Max timings, it will be considered abnormal since they do not fall within the 68% of the standard normal distribution.

The timings shown in Table IX were obtained during a DoS attack. As in the above case while we injected several "Deauthentication" frames with fake MAC addresses we initiated EAP-PEAP authentication and recorded the round trip times. Here, "EAP 7" and the "Key Exchange" messages show very high timing values resulting in very high overall timing.

The timings shown in Table X above were obtained during a DoS attack. Here, like in the other cases we injected some "Deauthentication" frames with a fake

TABLE IX
EAP-PEAP TIMINGS DURING DOS ATTACK

Event	Time (ms)
OPEN AUTHENTICATION	0.63
OPEN ASSOCIATION	5.36
EAP 1	0.27
EAP 2 (PEAP)	7.44
EAP 3 (PEAP)	2.52
EAP 4 (PEAP)	5.30
EAP 5 (PEAP)	0.45
EAP 6 (PEAP)	0.20
EAP 7 (PEAP)	37.63
EAP 8 (PEAP)	0.47
EAP 9 (PEAP)	4.56
EAP Key Exchange	79.07
Overall	426.70

TABLE X
EAP-TLS TIMINGS DURING DOS ATTACK

Event	Time (ms)
OPEN AUTHENTICATION	0.43
OPEN ASSOCIATION	1.28
EAP 1	2.54
EAP 2	0.49
EAP 3 (PEAP)	4.73
EAP 4 (TLS)	58.01
EAP 5 (TLS)	0.36
EAP 6 (TLS)	14.51
EAP 7 (TLS)	4.87
EAP Key Exchange	85.87
Overall	276.97

MAC address jamming the medium. During this period we initiated an EAP-TLS authentication process and the timings were noted. From the timing measurements in Table X it can be seen that the values shown for some "EAP TLS Request/Response" messages are significantly higher resulting in a high overall timing. Thus our Timing Anomaly module detects that the hosts are experiencing some abnormal condition.

Although the DoS attack did not affect the authentication process itself the timing anomaly detection module detects a time delay in the authentication process. Such delays could be due to other qualities of the wireless environment as well, and hence, needs to be further investigated. Therefore, without merely reporting this as an anomaly we need to investigate the legitimacy of the anomaly. Hence, our EWS passes this information to the third phase - the intrusion prevention module, where the anomaly is substantiated using our proposed outlier based data association techniques.

B. Behavior Anomaly

The next stage in our anomaly detection process is the behavioral analysis of the participating hosts. The events shown in Table XI represent the normal behavior of a station during an EAP-TLS association process. Here the first four events represent 802.11 open system authentication. The next event initiate the 802.1x mutual authentication with the "EAP REQUEST IDENTITY 1" message. The 802.1x authentication begins with the access point requesting the wireless station to identify itself. The response from the wireless station is redirected to the authentication server which in turn initiates the

TABLE XI
EAP-TLS EVENTS DURING NORMAL BEHAVIOR

ID	Event
0	OPEN AUTHENTICATION
0	OPEN AUTHENTICATION
3	OPEN ASSOCIATION REQUEST
4	OPEN ASSOCIATION RESPONSE
6	EAP REQUEST 1 (IDENTITY)
7	EAP RESPONSE 1 (IDENTITY)
20	EAP REQUEST 2 (IDENTITY)
21	EAP RESPONSE 2 (IDENTITY)
40	EAP REQUEST 3 (PEAP)
41	EAP RESPONSE 3 (NAK)
42	EAP REQUEST 4 (TLS)
43	EAP RESPONSE 4 (TLS)
44	EAP REQUEST 5 (TLS)
45	EAP RESPONSE 5 (TLS)
47	EAP REQUEST 6 (TLS)
48	EAP RESPONSE 6 (TLS)
49	EAP REQUEST 7 (TLS)
50	EAP RESPONSE 7 (TLS)
51	EAP SUCCESS
52	EAPOL KEY
52	EAPOL KEY
52	EAPOL KEY
52	EAPOL KEY

TABLE XII
EAP-PEAP EVENTS DURING NORMAL BEHAVIOR

ID	Event
0	OPEN AUTHENTICATION
0	OPEN AUTHENTICATION
3	OPEN ASSOCIATION REQUEST
4	OPEN ASSOCIATION RESPONSE
6	EAP REQUEST 1 (IDENTITY)
7	EAP RESPONSE 1 (IDENTITY)
20	EAP REQUEST 2 (PEAP)
22	EAP RESPONSE 2 (PEAP)
23	EAP REQUEST 3 (PEAP)
25	EAP RESPONSE 3 (PEAP)
26	EAP REQUEST 4 (PEAP)
27	EAP RESPONSE 4 (PEAP)
29	EAP REQUEST 5 (PEAP)
30	EAP RESPONSE 5 (PEAP)
31	EAP REQUEST 6 (PEAP)
32	EAP RESPONSE 6 (PEAP)
33	EAP REQUEST 7 (PEAP)
34	EAP RESPONSE 7 (PEAP)
35	EAP REQUEST 8 (PEAP)
36	EAP RESPONSE 8 (PEAP)
37	EAP REQUEST 9 (PEAP)
38	EAP RESPONSE 9 (PEAP)
39	EAP SUCCESS
52	EAPOL KEY
52	EAPOL KEY
52	EAPOL KEY
52	EAPOL KEY

EAP type specific authentication process. Depending on the security configuration appropriate EAP type specific authentication will commence. Table XI lists the ten EAP-TLS authentication events (40-45, 47-50). Event 40 - "EAP REQUEST 3 (PEAP)" is initiated by the authentication server since we have set the default EAP type to PEAP in our RADIUS authentication server.

Table XII lists the EAP-PEAP authentication events. Here also the association begins with open system authentication followed by the 802.1x authentication. Unlike EAP-TLS authentication EAP-PEAP authentication consists of sixteen EAP type specific events (20, 22, 23, 25-27, 29-38). It can be noticed that this list does not have event "EAP RESPONSE (NAK)", since the wireless station is configured for EAP-PEAP authentication itself.

Table XIII lists the EAP-LEAP authentication events. Here again we have events 20 and 21 since the wireless

TABLE XIII
EAP-LEAP EVENTS DURING NORMAL BEHAVIOR

ID	Event
0	OPEN AUTHENTICATION
0	OPEN AUTHENTICATION
3	OPEN ASSOCIATION REQUEST
4	OPEN ASSOCIATION RESPONSE
6	EAP REQUEST 1 (IDENTITY)
7	EAP RESPONSE 1 (IDENTITY)
20	EAP REQUEST 2 (PEAP)
21	EAP RESPONSE 2 (NAK)
11	EAP REQUEST 3 (LEAP)
12	EAP RESPONSE 3 (LEAP)
13	EAP SUCCESS
14	EAP REQUEST 4 (LEAP)
18	EAP RESPONSE 4 (LEAP)
52	EAPOL KEY
52	EAPOL KEY
52	EAPOL KEY
52	EAPOL KEY

TABLE XIV
EAP-TLS EVENTS DURING REPLAY ATTACK

Raw Event	ID
STA3 - AP1 444.803634 360 AUTHENTICATION	0
STA3 - AP1 444.805244 362 ASSOCIATION REQUEST	3
AP1 - STA3 444.873872 813 DEAUTHENTICATION	1
STA3 - AP1 444.888390 374 ASSOCIATION REQUEST	3
AP1 - STA3 444.888758 814 DEAUTHENTICATION	1
STA3 - AP1 444.902434 376 ASSOCIATION REQUEST	3
AP1 - STA3 444.902874 815 DEAUTHENTICATION	1
STA3 - AP1 447.425796 053 AUTHENTICATION	0
AP1 - STA3 447.426138 883 AUTHENTICATION	0
STA3 - AP1 447.427177 054 ASSOCIATION REQUEST	3
AP1 - STA3 447.427769 884 ASSOCIATION RESPONSE	4
AP1 - STA3 447.428256 490 EAP REQUEST IDENTITY 1	6
STA3 - AP1 447.435340 056 EAP RESPONSE IDENTITY 1	7
.... adopts normal behavior from here	-

station is configured for EAP-LEAP authentication. On all of the three types of authentication processes the last four messages (52) represent the key distribution phase. EAP-LEAP authentication process consists of five events (11-14, 18) only.

Tables XI, XII and XIII all list the EAP type specific events during normal behavior. However, when anomalies arise this behavior can change. There can be situations where the number of events are extraordinarily high or low. There can also be situations where events can totally disappear from the receptive range of the monitoring devices. Therefore tracking these events and analyzing them appropriately can reveal vital information regarding any abnormality in the wireless environment. As discussed before here again we hold behavior profiles for every participating station - access point pair. Therefore, when stations roam and re-associate via another access point we will then be analyzing the behavior with a profile specific to that access point.

In order to study the behavior of the wireless stations during abnormal conditions, next, we injected some abnormal management frames forcing the wireless hosts to deauthenticate and reassociate frequently. This was done by injecting “Deauthentication” and “Association Request” frames simulating a Replay attack.

Table XIV lists the traces obtained during EAP-TLS authentication between station STA3 and the access point. Each line of the table corresponds to a single message passed between two wireless hosts, indicating the direction of message transfer, the time of transfer, the sequence number and the event.

TABLE XV
EAP-PEAP EVENTS DURING REPLAY ATTACK

Raw Event	ID
AP1 - STA2 85.025390 0509 AUTHENTICATION	0
STA2 - AP1 85.025936 0117 ASSOCIATION REQUEST	3
AP1 - STA2 85.026522 0510 ASSOCIATION RESPONSE	4
STA2 - AP1 85.027379 3694 ASSOCIATION REQUEST	3
AP1 - STA2 85.027884 1371 EAP REQUEST IDENTITY 1	6
STA2 - AP1 85.042119 3694 ASSOCIATION REQUEST	3
STA2 - AP1 85.042220 3694 ASSOCIATION REQUEST	3
AP1 - STA2 86.016005 0846 DEAUTHENTICATION	1
STA2 - AP1 90.743014 0216 AUTHENTICATION	0
AP1 - STA2 90.743440 0912 AUTHENTICATION	0
STA2 - AP1 90.744476 0217 ASSOCIATION REQUEST	3
AP1 - STA2 90.745070 0913 ASSOCIATION RESPONSE	4
AP1 - STA2 90.745531 1386 EAP REQUEST IDENTITY 1	6
STA2 - AP1 90.763463 0218 EAP RESPONSE IDENTITY 1	7
..... adopts normal behavior from here	-

TABLE XVI
EAP-LEAP EVENTS DURING REPLAY ATTACK

Raw Event	ID
AP1 - STA1 205.665782 1789 AUTHENTICATION	0
STA1 - AP1 205.666533 2725 ASSOCIATION REQUEST	3
AP1 - STA1 205.667928 1790 ASSOCIATION RESPONSE	4
AP1 - STA1 205.668035 1409 EAP REQUEST IDENTITY 1	6
AP1 - STA1 235.657322 2145 DEAUTHENTICATION	1
STA1 - AP1 241.369603 2824 AUTHENTICATION	0
AP1 - STA1 241.370002 2213 AUTHENTICATION	0
STA1 - AP1 241.370995 2825 ASSOCIATION REQUEST	3
AP1 - STA1 241.371589 2214 ASSOCIATION RESPONSE	4
AP1 - STA1 241.372046 1424 EAP REQUEST IDENTITY 1	6
STA1 - AP1 241.373288 2826 EAP RESPONSE IDENTITY 1	7
..... adopts normal behavior from here	-

Tables XIV, XV and XV list the wireless traces obtained from a number of RSNA during EAP-TLS, EAP-PEAP and EAP-LEAP authentication process between the access point and stations STA3, STA2 and STA1 respectively. In all three scenarios we see a behavioral anomaly where the access point deauthenticate the station without responding to the replayed messages sent by the adversary. This is due to the adversary sending an Association Request message without the correct credentials to meet the requirements of the access point.

However, although all three authentication mechanisms do not demonstrate similar behavior, EAP-PEAP and EAP-LEAP authentication mechanisms show similar behavior (Tables XV and XV) compared to the EAP-TLS authentication process. This is because STA3 is a Windows XP client, whereas the other two are Linux clients. Since the adversary is also a Linux client the replayed Association Request message is acknowledged and the access point initiates EAP authentication with the EAP REQUEST IDENTITY message, which the adversary does not recognize and sends another Association Request message. At this point the access point realizes foul play and disconnects the station which results in the legitimate station also getting disconnected.

The behavior anomaly detector will detect all three scenarios as anomaly, but it will be up to the intrusion prevention module to substantiate the legitimacy of the anomalies. Hence, let us now see how the intrusion prevention module will substantiate these anomalies.

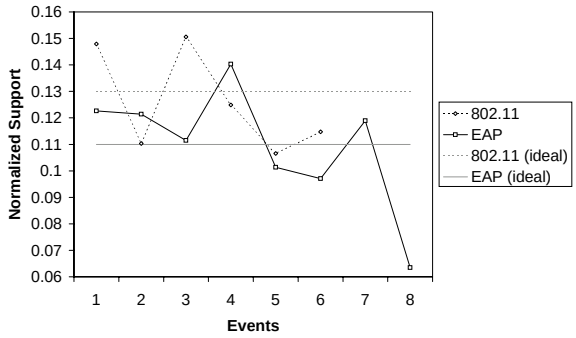

Fig. 7. Normal EAP-LEAP Authentication

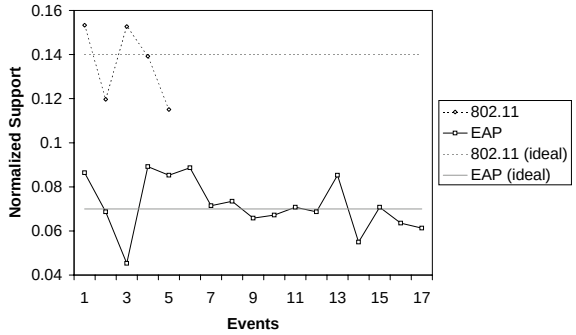

Fig. 8. Normal EAP-PEAP Authentication

VIII. SUBSTANTIATION

Figures 7, 8 and 9 show the normalized support (Equation 2) for EAP type specific events on wireless stations configured for EAP-LEAP, EAP-PEAP and EAP-TLS authentication respectively. These readings were obtained from our experimental setup during normal operations. Here, we use the term normal operation meaning that there was no abnormal events introduced explicitly. However, the wireless environment was exposed to the regular campus wide wireless traffic. The diagrams also indicate the ideal normalized support (Table XVII) for both IEEE 802.11 and EAP events. The event numbers used in these diagrams are serial values that correspond to the event IDs in Tables XI, XII and XIII in the same order.

On the above diagrams that show the behavior of wireless hosts under normal operations, the normalized support values of most events have support values close to the ideal normalized support. Although some EAP-LEAP and EAP-PEAP events show larger deviation from the ideal support values, the overall confidence is guaranteed as shown in Table XVIII. The table lists the uncertainty level (Equation 4) and the confidence levels (Equation 5) of IEEE 802.11 and EAP group events under EAP-LEAP, EAP-PEAP and EAP-TLS authentication.

To calculate these values, the outlier detection process made at most two drill-down queries on the data cube and the average time taken to execute a single result was approximately 0.0324 ms. We made use of the frequency measure stored on each cell to calculate the individual frequencies of the events and their aggregate frequencies. In many data warehousing applications the measure values stored in cells are fixed. In our case since

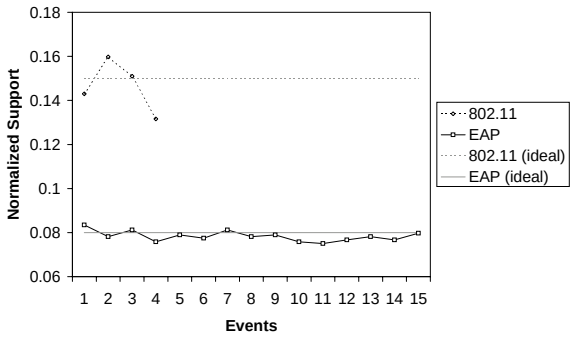

Fig. 9. Normal EAP-TLS Authentication

TABLE XVII
IDEAL NORMALIZED SUPPORT

EAP-LEAP		EAP-PEAP		EAP-TLS	
802.11	EAP	802.11	EAP	802.11	EAP
0.43	0.37	0.46	0.24	0.50	0.27

we need real time response the measure values need to be incremented on-the-fly when new legitimate states are identified.

From these results it is evident that all three authentication schemes demonstrate high confidence during normal operations irrespective of those events deviating from ideal support values. This is due to the fact that all IEEE 802.11 and EAP events, as a group, behave in a predictable manner under normal conditions. However, under abnormal conditions their group behavior cannot be guaranteed. To further emphasize our claim we consider two abnormal situations in the wireless environment and discuss their results.

Figures 10, 11 and 12 show the normalized support for EAP type specific events on wireless stations configured for EAP-LEAP, EAP-PEAP and EAP-TLS authentication respectively, during a DoS attack. In all three cases the support values of IEEE 802.11 events deviate considerably from the ideal support values. This is because of the "Deauthentication" frames injected. Consequently, this demonstrates the unreliable nature of 802.11 open system association.

During the DoS attack, events "ASSOCIATION REQUEST" (5) and "ASSOCIATION RESPONSE" (6) have almost equal support in both EAP-LEAP and EAP-PEAP authentication. This is due to the legitimate reassociation after a deauthentication. Similarly, the "DEAUTHENTICATION" (3,4) and "AUTHENTICATION" (1,2) events also have almost equal frequencies. However, "DEAUTHENTICATION" events have high frequency because of the injected "Deauthentication" frames. Furthermore, the EAP events on all three stations show less support compared to the 802.11 events, particularly the "DEAUTHENTICATION" events. This is because of the large number of "DEAUTHENTICATION" events injected and the corresponding response. However, reassociation does not take place at the same frequency, since the channel is busy with "DEAUTHENTICATION" events. In some cases when 802.11 open system authentication is in progress there has been interruptions due to the injected

TABLE XVIII
QUANTIFYING NORMAL OPERATION

Parameter	Authentication Type					
	EAP-LEAP		EAP-PEAP		EAP-TLS	
	802.11	EAP	802.11	EAP	802.11	EAP
$H(k)$	2.51	2.91	1.95	3.85	1.93	3.74
$C(k)$	97.02	97.07	97.39	98.45	96.57	98.11

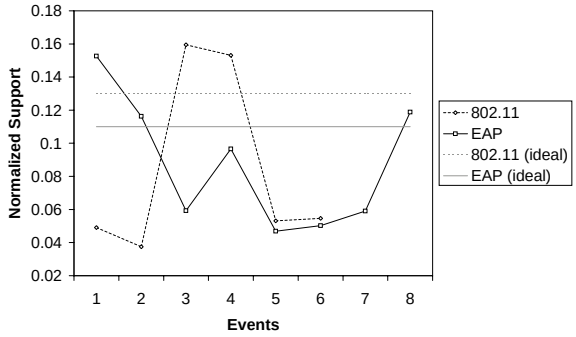

Fig. 10. DoS Attack during EAP-LEAP Authentication

"DEAUTHENTICATION" events resulting in fresh reassociation.

On the other hand, if we look at the EAP events, except for the first "EAP REQUEST IDENTITY 1" event all other EAP events have support values close to ideal. This is because "EAP REQUEST IDENTITY 1" event is triggered by the access point and at this point the actual EAP authentication has not commenced. Therefore, a bogus "DEAUTHENTICATION" event can still interrupt the association process resulting in fresh reassociation.

As discussed above, Figures 10, 11 and 12 all demonstrate the behavior of RSNA events during the DoS attack. It shows that 802.11 events are vulnerable to the attack. To further investigate this behavior we calculated the uncertainty level and the confidence level for each case. As shown in Table XIX, the confidence levels for IEEE 802.11 events are comparatively low for all three stations. Whereas, the confidence level for EAP events are high. However, station STA3 configured for EAP-TLS authentication demonstrates very high confidence (90.53) for EAP events and very low confidence (40.36) for 802.11 events.

[t]

Figures 13, 14 and 15 show the normalized support for EAP type specific events on wireless stations configured for EAP-LEAP, EAP-PEAP and EAP-TLS authentication respectively, during a Replay attack. Here again, the support for 802.11 events are considerably high. This is because of the "Authentication" and "Association Request" frames injected. Consequently, as in the case of DoS attack this demonstrates the unreliable nature of 802.11 open system association. In this case the events "AUTHENTICATION" (1) and "ASSOCIATION REQUEST" (4), both replayed messages, have very high frequency indicating the Replay attack. It can also be noted that in these figures we have six 802.11 events compared to the five events found during the DoS attack. The additional event found here is the "DISASSOCIATION" (3) message

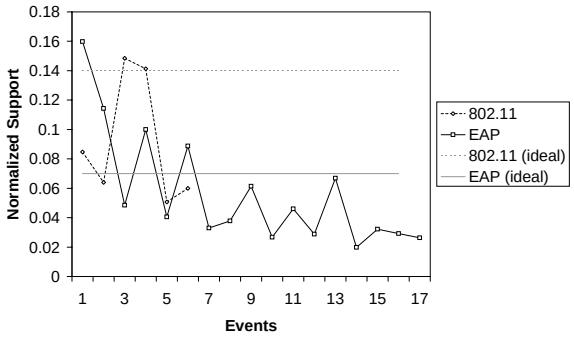

Fig. 11. DoS Attack during EAP-PEAP Authentication

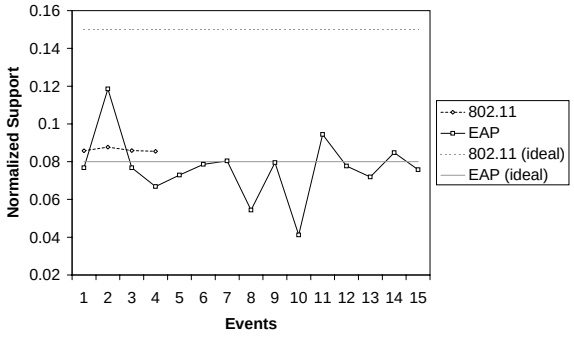

Fig. 12. DoS Attack during EAP-TLS Authentication

fired by the access point. The access point rejects the adversary when it does not respond favorably to the "EAP REQUEST IDENTITY 1" message. Hence, it is evident that the behavior of the EAP configured stations differ depending on the type of attack they experience.

Further, it can be noticed that in both attack scenarios the support for "EAP REQUEST IDENTITY 1" event is high compared to the other EAP events. This is because at this stage the authentication server has still not come into action. It is only after the authentication server issues the "EAP REQUEST PEAP 2" (since default EAP type is set to PEAP on the authentication server) message the EAP events become stable. In this study the "Replay" or the "DoS" attack executed do not interfere with the EAP encapsulated messages, however, a rogue access point can replay the first "EAP REQUEST IDENTITY 1" message.

Next, let us consider the EAP type specific events. The support values of the type specific events are almost constant in both attack scenarios confirming the effectiveness of the EAP authentication process. As soon as the EAP type specific authentication begins the association process becomes very reliable. Although, the DoS attack can jam the authentication process delaying certain message exchange (as seen in Table VIII), the process itself is robust due to the effectiveness of the EAP.

As discussed above, Figures 10, 11 and 12 all demonstrate the behavior of RSNA events during the Replay attack. It confirms that 802.11 events are vulnerable to the attack. To further investigate, as in the case of DoS attack, we calculated the uncertainty level and the confidence level for each case. Table XX shows these values. Here again the confidence levels for 802.11 events are low for

TABLE XIX
QUANTIFYING A DOS ATTACK

Parameter	Authentication Type					
	EAP-LEAP		EAP-PEAP		EAP-TLS	
	802.11	EAP	802.11	EAP	802.11	EAP
$H(k)$	1.68	2.51	1.82	3.38	1.04	3.91
$C(k)$	65.17	79.33	70.57	79.56	40.36	90.53

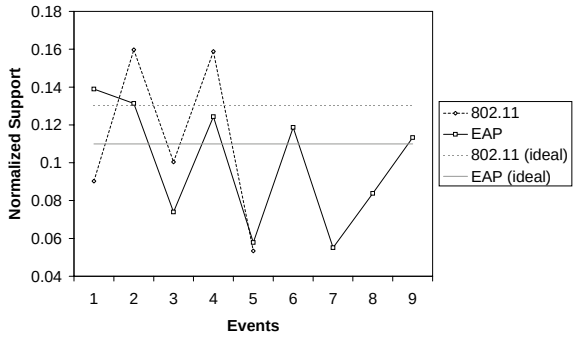

Fig. 13. Replay Attack during EAP-LEAP Authentication

all three stations (71.52 - 80.52). The confidence level for EAP events are comparatively high (94.03 - 94.77).

The above examples show promising results for our proposed concept of using timing and/or behavioral analysis with outlier based confidence measure for detecting and substantiating abnormal conditions. Although in our experiments we have not tested the vulnerabilities of the EAP authentication process, the same concept can be used to detect future exploits on the EAP process. Further, in most attack scenarios the attacker needs to inject many messages before he could actually compromise the credentials of a legitimate station. Therefore, as soon as we detect an abnormal condition we can continue to monitor the suspicious stations and once sufficient traces are collected we can substantiate the abnormality using outlier based data association techniques. In this manner we can either raise an alarm that a security breach is on the verge or give an indication of the level of threat for the exposed station.

In the foregoing paragraphs we have discussed the vulnerabilities of the various protocols and how they affect the overall confidence of the RSNA. However, in the case of RSN, it is important to note, that although IEEE 802.11 and EAP protocols may demonstrate low confidence levels, it is up to the security policy maker to substantiate it appropriately depending on user privileges. Because in the two attack scenarios discussed even though the behavior indicates that the RSN is unstable during the 802.11 and EAP protocol phases, it only affects the "availability" of the stations and does not compromise the "authenticity" of the participating stations. Hence, when formulating the security threat levels one has to consider a range of factors, mainly the user priorities, level of sensitivity of the network etc. and make appropriate decisions.

The data cube results discussed above are obtained in an offline mode. We believe that this system is scalable and can be extended for any type and size of networks

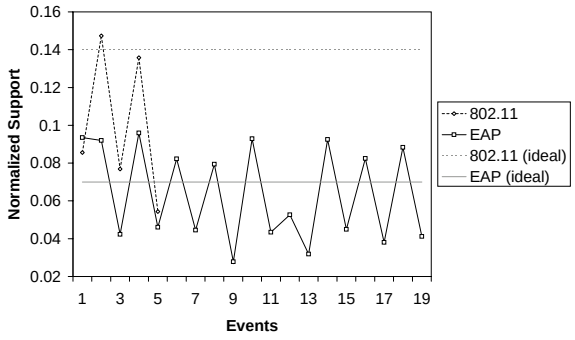

Fig. 14. Replay Attack during EAP-PEAP Authentication

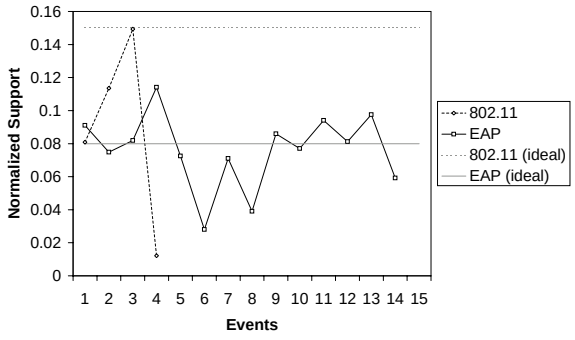

Fig. 15. Replay Attack during EAP-TLS Authentication

in Online mode since the capabilities of the data cube algorithms were rigorously tested [29]. The timing results obtained from our small test cluster is also encouraging.

IX. CONCLUSIONS

In this paper, we have discussed the use of timing and/or behavioral analysis for detecting anomalies in the wireless environment. Further, we have reported the results obtained from our novel substantiation mechanism (GCL) in validating the abnormal conditions. The experimental results obtained with the 802.11i based network are promising and confirming the concept of the proposed detection mechanism. The GCLs obtained for EAP-LEAP, PEAP and TLS show that the EAP events as a group demonstrate high confidence than the IEEE 802.11 events. This confirms the effectiveness of IEEE 802.1x based authorization, authentication and key distribution adapted by the IEEE 802.11i security mechanism.

The analysis of the test results demonstrate the effectiveness of our proposed system in detecting various anomalies including security threats. Substantiation of the detected anomalies using normalized support values and GCL has been tested on real data. Nevertheless, we have not tested our system with abnormalities that can directly interfere with the EAP authentication process. As of to date EAP vulnerabilities on the wireless environment has not been reported. However, the results presented here shows that our concept is capable of detecting even future exploits on EAP vulnerabilities. Furthermore, although this technique is expected to provide very promising results in cooperate networks, their applicability on smaller networks may be limited due to resource requirements.

TABLE XX
QUANTIFYING REPLAY ATTACK

Parameter	Authentication Type					
	EAP-LEAP		EAP-PEAP		EAP-TLS	
	802.11	EAP	802.11	EAP	802.11	EAP
$H(k)$	1.87	2.98	1.66	4.03	1.71	3.85
$C(k)$	80.52	94.03	71.52	94.77	73.60	94.29

ACKNOWLEDGMENT

We thank Prof. Frank Dehne of Carlton University, Ottawa, Canada for his advice on the use of data cubes on this work. We also thank Prof. Todd Eavis at Concordia University, Montreal, Canada for providing us with the initial software for generating and querying the data cube.

REFERENCES

[1] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Amendment 6: Medium Access Control (MAC) Security Enhancements*, IEEE Standard 802.11i Part 11, July 2004.

[2] A. Bittau, M. Handley, and J. Lackey, "The Final Nail in WEP's Coffin," in *Proceedings of the 2006 IEEE Symposium on Security and Privacy (S&P'06)*, vol. 00, 2006, pp. 386–400.

[3] E. Sithirasenan and V. Muthukkumarasamy, "An Early Warning System for IEEE 802.11i Wireless Networks," in *AusWireless'06: Proceedings the of 1st Australian Conference on Broadband Wireless and Ultra Wideband*, March 2006, pp. 25–30.

[4] ——, "Detecting Security Threats in Wireless LANs Using Timing and Behavioral Anomalies," in *ICON'07: Proceedings of the 15th IEEE International Conference on Networks*, November 2007, pp. 66–71.

[5] C. S. Inc., "Cisco LEAP (LEAP)," http://ietfreport.isoc.org/all-ids/draft-josefsson-pppext-eap-tls-eap-07.txt, October 2003.

[6] A. Palekar, D. Simon, G. Zorn, J. Salowey, H. Zhou, and S. Josefsson, "Protected EAP Protocol (PEAP)," http://ietfreport.isoc.org/all-ids/draft-josefsson-pppext-eap-tls-eap-07.txt, October 2003.

[7] B. Aboba and D. Simon, "PPP EAP TLS Authentication Protocol," http://tools.ietf.org /wg/pppext/draft-ietf-pppext-eaptls/draft-ietf-pppext-eaptls-06.txt, August 1999.

[8] E. Sithirasenan, V. Muthukkumarasamy, and D. Powell, "IEEE 802.11i WLAN Security Protocol - A Software Engineer's Model," in *AusCERT '05: Proceedings of the 4th Asia Pacific Information Technology Security Conference*, May 2005, pp. 39–50.

[9] C. He and J. C. Mitchell, "Security Analysis and Improvements for IEEE 802.11i," in *Proceedings of the 12th Annual Network and Distributed System Security Symposium*, February 2005. [Online]. Available: http://www.isoc.org/isoc/conferences/ndss/05/proceedings/index.htm

[10] M. Lynn and R. Baird, "Advanced 802.11 attack." Las Vegas, USA, July 2002.

[11] N. Asokan, V. Niemi, and K. Nyberg, "Man-in-the-Middle in tunneled authentication protocols," IACR ePrint archive, United Kingdom, Tech. Rep. 2002/163, Cotober 2002. [Online]. Available: http://eprint.iacr.org/2002/163/

[12] M. Barbeau, J. Hall, and E. Kranakis, "Detecting impersonation attacks in future wireless and mobile networks," in *Proceedings of the Workshop on Secure Mobile Ad-hoc Networks and Sensors*, September 2005.

[13] J. Hall, M. Barbeau, and E. Kranakis, "Anomaly-based intrusion detection using mobility profiles of public transportation users," in *Proceedings of the IEEE International Conference on Wireless And Mobile Computing, Networking And Communications*, vol. 2, August 2005, pp. 17–24.

[14] R. Gill, J. Smith, and A. Clark, "Experiences in passively detecting session hijacking attacks in ieee 802.11 networks," in *AISW-NetSec '06: Proceedings of the 4th Australasian Information Workshop*, vol. 54, January 2006, pp. 26–38.

[15] R. A. Maxion and K. M. C. Tan, "Benchmarking anomaly-based detection systems," in *Proceedings of the International Conference on Dependable Systems and Networks*, June 2000, pp. 623–630.

[16] S. Manganaris, M. Christensen, D. Zerkle, and K. Hermiz, "A data mining analysis of rtid alarms," *Comput. Networks*, vol. 34, no. 4, pp. 571–577, 2000.

[17] C. Clifton and G. Gengo, "Developing custom intrusion detection filters using data mining," in *MILCOM 2000: Proceedings of the 21st Century Military Communications Conference*, vol. 1, October 2000, pp. 440–443.

[18] D. Barbar, N. Wu, and S. Jajodia, "Detecting novel network intrusions using bayes estimators," in *Proceedings of the First SIAM Conference on Data Mining*, April 2001, pp. 182–191.

[19] G. Florez, S. Bridges, and R. Vaughn, "An improved algorithm for fuzzy data mining for intrusion detection," in *NAFIPS 2002: Proceedings of the Annual Meeting of the North American Fuzzy Information Processing Society*, June 2002, pp. 457–462.

[20] E. Bloedorn, A. D. Christiansen, W. Hill, C. Skorupka, L. M. Talbot, and J. Tivel, "Data mining for network intrusion detection: How to get started," http:// www.mitre.org/work/tech_ papers/tech_ papers_01/bloedorn_ datamining/bloedorn_ datamining.pdf, August 2001.

[21] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Standard 802.11, June 1999.

[22] *Local and Metropolitan Area Networks, Port-Based Network Access Control*, IEEE Standard 802.1X, June 2001.

[23] E. Sithirasenan, S. Zafar, and V. Muthukkumarasamy, "Formal Verification of the IEEE 802.11i WLAN Security Protocol," in *ASWEC'06: Proceedings of the 21st Australian Software Engineering Conference*, April 2006, pp. 181–190.

[24] J. Gray, S. Chaudhuri, A. Bosworth, A. Layman, D. Reichart, M. Venkatrao, F. Pellow, and H. Pirahesh, "Data cube: A relational aggregation operator generalizing group-by, cross-tab, and sub-totals," *J. Data Mining and Knowledge Discovery*, vol. 1, no. 1, pp. 29–53, 1997. [Online]. Available: citeseer.ist.psu.edu/gray97data.html

[25] F. Dehne, T. Eavis, S. Humbrusch, and A. Chaplin, "Parallelizing the Data Cube," *The Journal of Distributed and Parallel Databases*, vol. 11, no. 2, pp. 181–201, 2002.

[26] F. Dehne, T. Eavis, and A. Chaplin, "Top Down Computation of Partial ROLAP Data Cubes," in *HICSS-37: Proceedings of the 37th Annual Hawaii International Conference On System Sciences*, January 2004, pp. 181–193.

[27] L. Blunch and J. Vollbrecht, "Extensible Authentication Protocol (EAP)," http:// www.ietf.org/rfc/rfc2284.txt, March 1998.

[28] "Aircrack: WEP and WPA-PSK keys cracking program," http://www.aircrack-ng.org/doku.php, July 2006.

[29] E. Sithirasenan, Y. Chen, F.Dehne, T. Eavis, A. Chaplin, and D. Green, "cgmOLAP: Efficient Parallel Generation and Querying of Terabyte Size ROLAP Data Cubes," in *ICDE '06: Proceedings of the 22nd International Conference in Data Engineering*, April 2006, pp. 164–167.

Elankayer Sithirasenan is currently a Ph.D. candidate at Griffith University, Australia. He received his MSoftEng from Griffith University, Australia in 2004 and BScEng from University of Peradeniya, Sri Lanka, in 1991. His current research interests include wireless/wired network security, intrusion detection and prevention techniques, outlier detection on multi-level, multivariate data sets and software requirements analysis. He was a lecturer at the University of Peradeniya, Sri Lanka, 2001-2003.

Vallipuram Muthukkumarasamy received the BScEng(Hons) from the University of Peradeniya, Sri Lanka in 1984 and the PhD from the Cambridge University, England in 1990. He is currently a Senior Lecturer of Information and Communication Technology at Griffith University, Gold Coast, Australia. His current research interests include intrusion detection and prevention techniques, secure key establishment in wireless sensor networks, preventing denial of service attacks in 802.11 networks and security and privacy issues in e-government systems.