CHAPTER 2

LITERATURE REVIEW

2.1. Overview

This chapter combines several aspects of research study. It starts with definition and classification of threats, attacks and intrusions. Then, attacks and intrusions will be classified in determining the classification of both of them in Secured Autonomous Agent-based Intrusion Detection System (SAAIDS). After that, this chapter continues with procedures in defending and recovering steps. Later, this chapter gives definition and classification of. There are ten analysis approaches used in IDS and all of them will be discussed later including signature analysis that is used as data analysis approach in this research. Then the discussion followed on cryptology in IDS including encryption and digital signature. Cryptologies used in this research are Elgamal Encryption and Elgamal Digital Signature which will be given details explanation later in this part. After that, recent trend and previous works on IDS which discuss about recent technologies used in IDS and previous works are discussed. From observation and research done, issues and problems in previous works is discussed. Finally, all the information gathered concluded in the final section.

2.2. Threats, Attacks and Intrusions

Threat is a term used to describe all sort of attempts intending to attack a system through vulnerabilities and weaknesses. Threats begin with an attack either into a network called network-attack or directly into a system called host-attack. There are many ways attacks can be launched through network such as smurf attack and syn flood attack which later lead to Denial of Service (DoS). Besides that, password cracking and access violation are among the most common attacks towards host-attacks. All these kinds of threats intend either to shut down or halt a system or at least steal information and gain access violation.

Due to the increasing usage of Internet, confidentiality is easily to be threatened. With the increasing number of services and business activities which are running through web and network connectivity, the usage of online payment and credit card are unlimited. All threats mentioned above had been big issues in system and network security. These have lead to big loss for organizations or companies involved which uncertainly caused an economic environment chaos.

By referring at MyCert/CyberSecurity Malaysia incident statistics from August 1997 until July 2008, there has been rapid increase of computer crime cases in Malaysia especially in intrusion (MyCert/CyberSecurity Malaysia, 2008). For intrusion cases, from 22 and 37 cases in 1997 and 1998 respectively, it has increased to 897 cases in 2006 and until July 2008 it has touched 143 cases. In 2006, intrusion has been highlighted as the most frequent cases reported which represents 65.38% from all incidents (897 cases from total of 1372 cases) as shown in Figure 2.1. All the incident statistics reported by organizations and public member show that intrusion detection is a major incident in Malaysia.

Statistic in Figure 2.1 shows the impact of intrusion in system and network security specifically in Malaysia. In determining causes of each incident, this chapter discusses history, definition and classification of attacks as stated in next section. The history, definition and classification are gathered from previous research publication

worldwide and incidents statistics in Malaysia towards providing better understanding about attacks in searching solutions for intrusion detection and prevention.

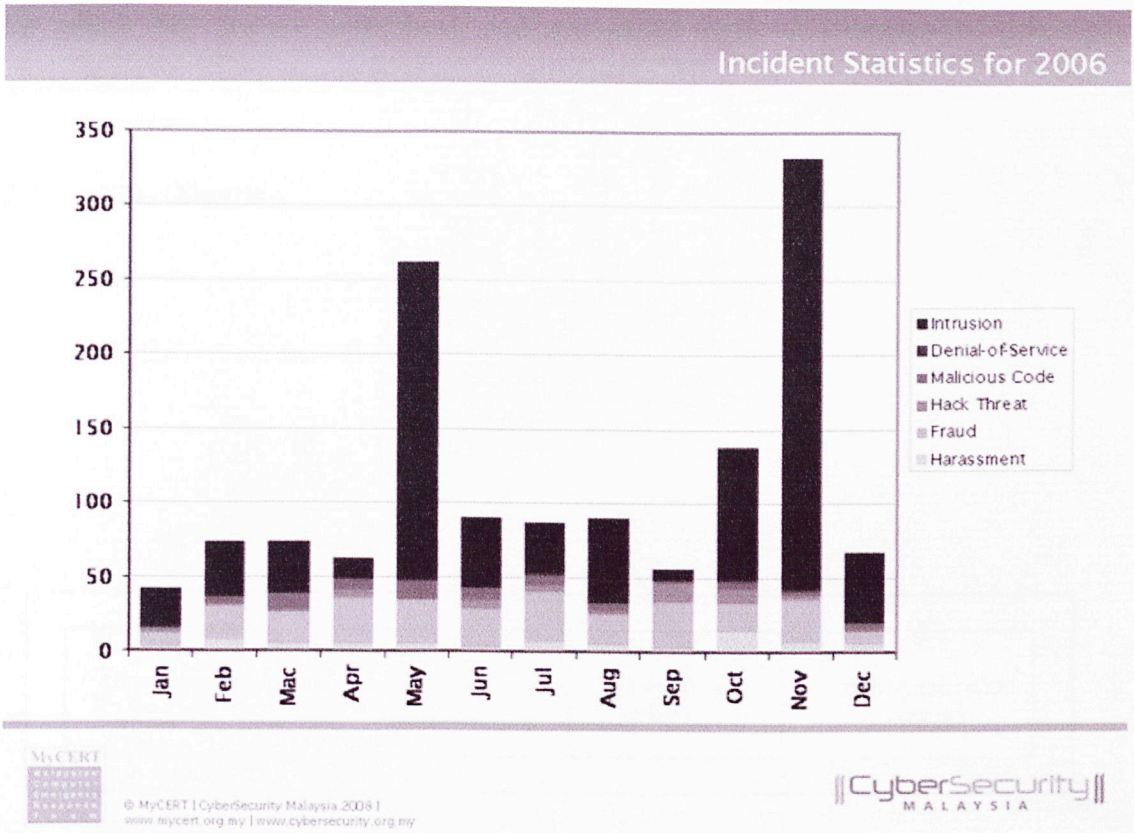

Figure 2.1: Incident Statistics for 2006

2.2.1. Definition of Attack and Intrusion

Attacks in this chapter is a term used to describe all kind of threats in system or network security which more consistently called hacks or cracks. Therefore, for time being, an attacker is widely known as a hacker or cracker instead of attacker. Many early hacks took the form of elaborate practical jokes. In ensuring better understanding about attacks, threats and intrusions, this research discusses taxonomies on several terms used in intrusion detection system cited from several articles and reports from previous researches.

Intrusion is defined as any set of actions that attempt to compromise the integrity, confidentiality, or availability of a computer resource (Heady, 1990). Rebecca Bace and Peter Mell define intrusions as attempts to compromise the confidentiality,

integrity, availability, or to bypass the security mechanisms of a computer or network (Bace and Mell, 2001). Intrusions are caused by attackers accessing the systems from the Internet, authorized users of the systems who attempt to gain additional privileges for which they are not authorized, and authorized users who misuse the privileges given them.

2.2.2. Classification

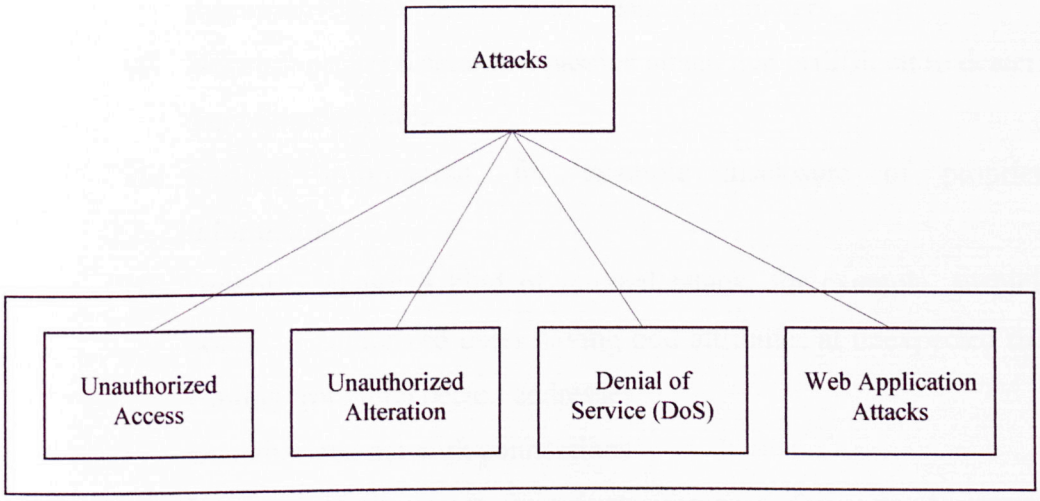

Figure 2.2: Attacks Classification

There are various kinds of attacks identified by source category either internal system; intranet (LAN) or external system; internet. Figure 2.2 shows attacks classification and below are the description of classification of attacks which are usually detectable by IDS (Kazienko et al., 2004).

a. Those related to unauthorized access to the resources, often as introductory steps toward more sophisticated actions:
    i. Password cracking and access violation.
    ii. Trojan horses.
    iii. Interceptions; most frequently associated with TCP/IP stealing and interceptions that often employ additional mechanisms to compromise operation of attacked systems. For example, by flooding; man in the middle attacks.

iv. Spoofing deliberately misleading by impersonating or masquerading the host identity by placing forged data in the cache of the named server such as DNS spoofing.

v. Scanning ports and services, including ICMP scanning (Ping), UDP, TCP Stealth Scanning TCP that takes advantage of a partial TCP connection establishment protocol.

vi. Remote OS Fingerprinting, for example by testing typical responses on specific packets, addresses of open ports, standard application responses (banner checks) and IP stack parameters.

vii. Network packet listening, a passive attack that is difficult to detect but sometimes possible.

viii. Stealing information, for example disclosure of proprietary information.

ix. Authority abuse; a kind of internal attack, for example, suspicious access of authorized users having odd attributes at unexpected times, coming from unexpected addresses.

x. Unauthorized network connections

xi. Usage of IT resources for private purposes, for example to access pornography sites.

xii. Taking advantage of system weaknesses to gain access to resources or privileges.

b. Unauthorized alteration of resources (after gaining unauthorized access):

i. Falsification of identity, for example to get system administrator rights

ii. Information altering and deletion

iii. Unauthorized transmission and creation of data set, for example arranging a database of stolen credit card numbers on a government computer. For example, the spectacular theft of several thousand numbers of credit cards in 1999.

iv. Unauthorized configuration changes to systems and network services (servers)

c. Denial of Service (DoS)

i. Flooding – compromising a system by sending huge amounts of useless information to lock out legitimate traffic and deny services:

- Ping flood (Smurf) – a large number of ICMP packets sent to a broadcast address
- Send mail flood - flooding with hundreds of thousands of messages in a short period of time; also POP and SMTP relaying
- SYN flood – initiating huge amounts of TCP requests and not completing handshakes as required by the protocol
- Distributed Denial of Service (DDoS); coming from a multiple source

ii. Compromising the systems by taking advantage of their vulnerabilities:

- Buffer Overflow, for example Ping of Death - sending a very large ICMP (exceeding 64 KB)
- Remote System Shutdown

d. Web Application attacks; attacks that take advantage of application bugs may cause the same problems as described above.

i. Viruses
ii. Worms
iii. Spamming
iv. Phishing

2.3. Attacks Classification in SAAIDS

For this project, based on researcher observation and research, to ensure the system working effectively, two categories of attacks had been classified which are attacks against agent communication and attacks against agent itself. Figure 2.3 the overview of attacks classification in SAAIDS.

2.3.1. Attacks Against Agent Communication

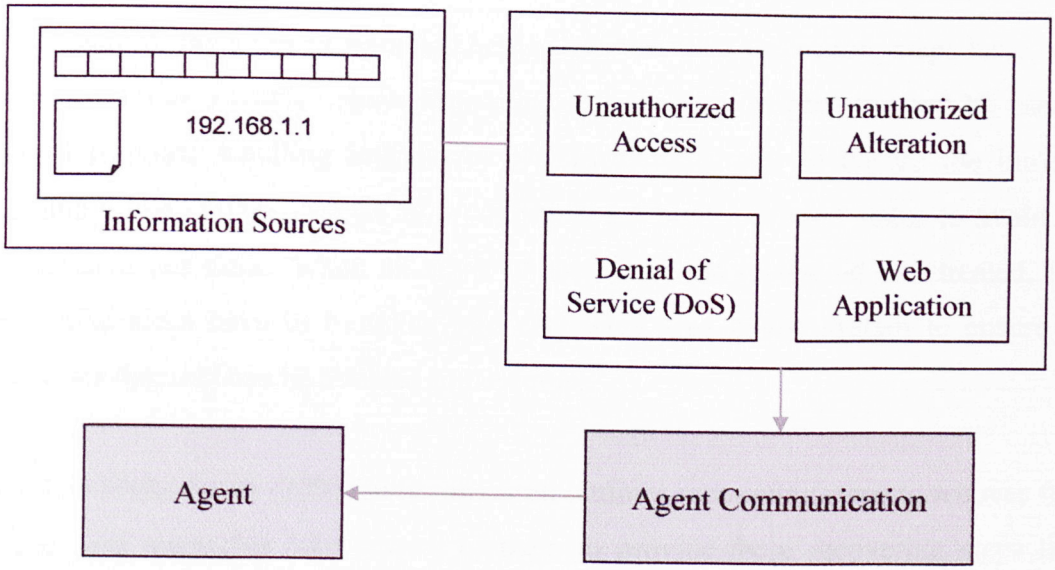

Figure 2.3: Attacks Classification

Agent communication is defined as a process of communication between agents in agent-based IDS. Communication allows agents to send messages to each others. Agent communication plays an important role to ensure a successful intrusion detection process. Attacks against agent communication will put agent-based IDS in risk of failing in doing intrusion detection which may come from all four categories of attacks; unauthorized access, unauthorized alteration, denial of service (DoS) and web application stated above.

2.3.2. Attacks Against Agent

Attacks against agent communication leads to attacks against agent itself. When an attacker gets into network, it opens chances for the attacker to take control of agent and or doing damages to the agent. Besides that, an agent in agent-based IDS is open to risks of an attack from other agents too. The attacker may use the damaged or the unauthorized agent to attack other agents running in the system.

2.4. Recovering From Intrusion

When there is any signs of intrusion being detected, then recovering steps have to be done as soon as possible. These recovering steps is very helpful and can be used as manual response handling actions for system administrator to ensure the intruded machine in the system or network are being under highly caution status to avoid any further damages done. When all signs of intrusions are recovered and treated, then prevention steps have to be taken into correlated area in the system to ensure the intrusions detected can be avoided in the future.

Madihah Mohd Saudi (2005) in her research outlines recovering steps to recover from viruses and researcher take several of them to provide these recovering steps from intrusions combined with other researches views in these actions. These recovering steps may vary to certain signs of intrusion to minimize the infection of the attacks from spread into other side of the system or network.

2.5. Intrusion Detection System

Since 1980, many researches had been done in IDS and various kinds of technology have been produced to detect intrusion. These technologies are growing rapidly and being enhanced to be immune with new ways of attacks. Therefore, here the author is trying to gather as many as information about IDS from technical papers, journals and books and giving an understanding through definition and classification of intrusion detection system.

Figure 2.4: Fundamental of IDS

Figure 2.4 shows the fundamental of IDS which includes:

a. Information Sources

The different sources of event information used to determine whether an intrusion has taken place. These sources can be drawn from different levels of the system with network, host and application monitoring.

b. Analysis

The part of intrusion detection systems that actually organizes and make sense of the events derived from the information sources, deciding when those events indicate that intrusions are occurring or have already taken place. The most common analysis approaches are *misuse detection* and *anomaly detection.*

c. Response

The set of actions which system applied once intrusions detected. These are typically grouped into active and passive measures, with active measures involving some automated intervention on the part of the system, and passive measures involving reporting IDS findings to Network Security Administrator, who are then expected to take action based on those reports.

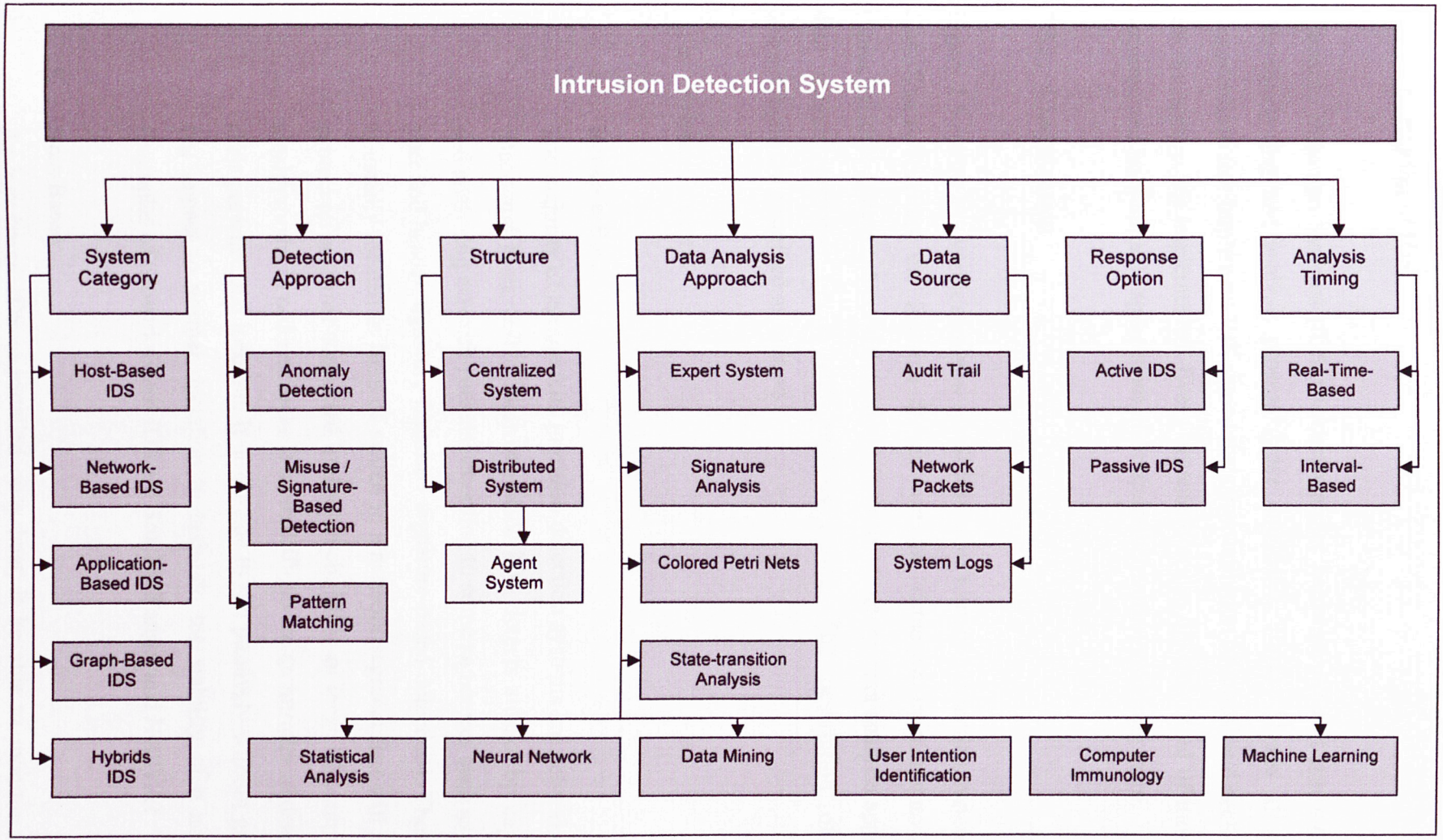

Figure 2.5: An Overview of Classification of Intrusion Detection System

2.5.1. Definition of IDS

Intrusion Detection Systems (IDSs) are software or hardware product that automate this monitoring and analysis process (Bace and Mell, 2001). Almost all of the organizations that implement IDS shared the same definition of intrusion detection but not the approaches implemented. Every organization has different kind of intrusions occurred and need suitable approaches and requirements in overcoming them.

2.5.2. Classification

Figure 2.5 shows an overview of new classification of intrusion detection system. The classification divides the features of an intrusion detection system into seven characteristics; system category, detection approach, structure, data analysis approach, data source, response option and analysis timing. The description of each characteristic are described as follows:

a. System Category

i. Network-Based

The majority of commercial intrusion detection systems are network-based (Bace and Mell, 2001). Network-based IDS detects attacks by capturing and analyzing network packets according to some signatures such as string, port and header signature (misuse / signature-based detection). These IDS is usually located in front of switch in a network segment that will able to protect an area that connected to the switch. Many of these IDS running in stealth mode, in order to make it more difficult for an intruder to determine their presence and location. It also made to run passively without giving a big impact to normal operation to maintain the stability of a network. Examples of network-based IDS are Bro, NetRanger and NetSTAT.

ii. Host-Based

As host-based IDS is installed on a host, it be able to monitor system resources as well as look at operating system audit trails and application

logs (anomaly detection). This process allows host-based IDS to analyze activities with great reliability and precision by determining which processes and users are involved in an attack. Host-based IDS is operated in a host, so it will not burden network traffic in a segment. Examples of host-based IDS are Emerald-Expert BSM, NFR HID, Snort and Dragon Squire.

iii. Application-Based

Application-based IDS is a subset of host-based IDS that analyze an attack with the same functionality but within a software application. So the most common information sources of these IDS are application's transaction log files. The ability to operate through its interface to the application's log files, the main functionality of this IDS is to detect suspicious behavior due to authorized unauthorized users or exceeding their authorization. Examples of application-based IDS are WebLoad.

iv. Graph-Based

Graph-based IDS is a subset of hybrid IDS which uses activity graph to detect distributed attacks against large networks. It constructs *activity graphs* to represent connections between hosts (network traffic) in its domain (Albag, 2005). A single host or a set of hosts (domain) and every edge in the graph represents the network traffic between nodes. GrIDS uses network sniffers and also host based ID systems as its data source. At this point it is said to be a combined approach of host based and network based IDS. Collected data will be the input of the graph engine, which is responsible for creating activity graphs. Then, the activity graph will be a guide to determine certain class of attacks and analyzed for detection. As a result, the incident can be ensured as an attack or just a false alarm. GrIDS is a graph-based IDS.

v. Hybrid

Hybrid IDS is a combination of any system category abovementioned. Certain authors consider a blend of host-based IDS and network-based IDS as a separate class so called Network Node IDS (NNIDS) which has its agents deployed on every host within the network being protected (Kazienko and Dorosz, 2004). Recently, agent is now very popular in intrusion detection process called agent-based IDS. Agent-based IDS is a combination between host-based IDS and network-based IDS. In agent-based IDS, agent will be deployed in every host and will detect certain class of attacks and collect all the information and send it to transceiver and monitors which maintaining network area being protected. AAFID is known as agent-based IDS but also categorized as Hybrid IDS.

This research proposed agent-based IDS as system category in implementing research objectives since agent-based IDS have many advantages in terms of mobility, scalability and reliability. Please refer to section 2.10 for further explanation.

b. Detection Approach

i. Anomaly Detection

In anomaly detection, detectors identify abnormal unusual behavior on a host or network. Here, anomaly detectors construct profiles that represent normal usage and then use current behavior data to detect a possible mismatch between profiles and recognize possible attack attempts. Anomaly detection will be able to detect the possibility of attacks as intrusion without getting inside their causes and characteristics and less dependence of IDS's operating environment.

ii. Misuse / Signature-Based Detection

Misuse detectors analyze system activity, looking for events or set of events that match a predefined pattern of event that describe a known attack. As the patterns corresponding to known attacks are called signature,

therefore misuse detection also called signature-based detection. There are two common approaches associated with signature detection: verification of pathology of lower layer packets and verification of application layer protocol. Any action is the system that is not clearly considered prohibited is allowed, therefore their accuracy is very high and low number of false alarms produced.

iii. Pattern Matching

In pattern matching detection, administrator monitor various systems and network attributes and information gathered obtained to be a specific environment. Any mismatch occurred during operational time will be able detected similar to the two abovementioned detections but the pattern or profile in this detection is part of human knowledge. This detection uses day-to-day operational experience to detect anomalies. This is a very powerful technique, since it able to detect unknown type of attacks.

This research proposed misuse/signature-based detection as detection approach which use network packet, IP address and system log as information sources.

c. Structure / Control Strategy

i. Centralized

Under centralized control strategy, all monitoring, detection and reporting is controlled directly from a central location (Figure 2.6).

ii. Partially Distributed

Monitoring and detection is controlled from a local control node, with hierarchical reporting to one or more central location (Figure 2.7).

iii. Fully Distributed

Monitoring and detection is done using an agent-based approach, where response decisions are made at the point of analysis (Figure 2.8).

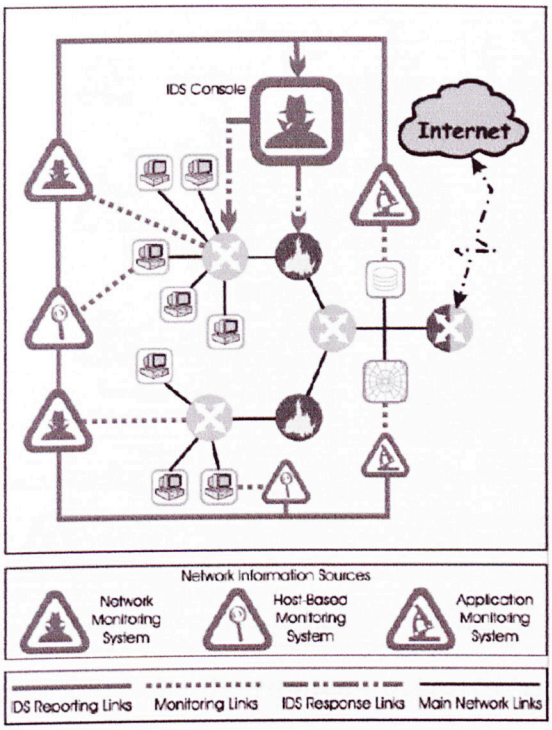

Figure 2.6: Centralized Control Strategy

(Source from: (Bace and Mell, 2001), Page 12, Figure 1)

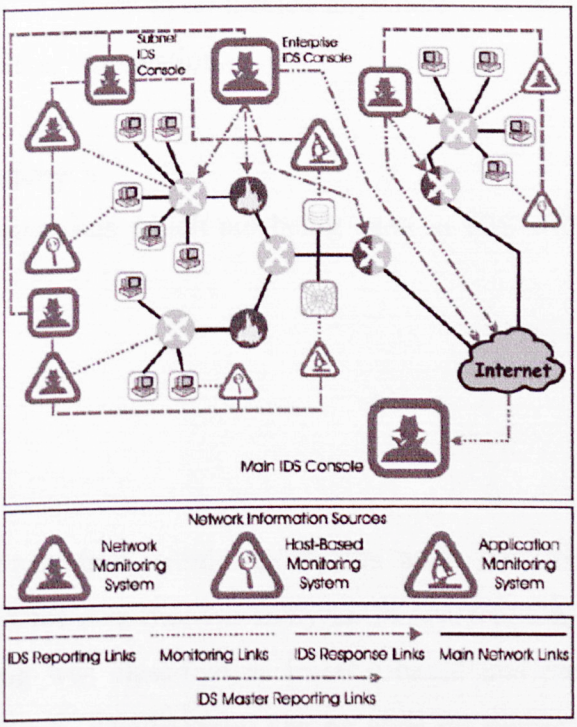

Figure 2.7: Partially Distributed Control Strategy

(Source from: (Bace and Mell, 2001) , Page 13, Figure 2)

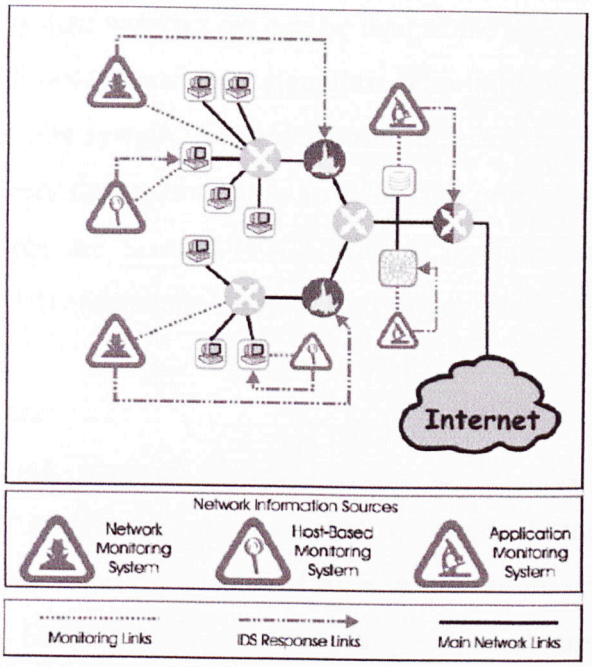

Figure 2.8: Fully Distributed Control Strategy

(Source from: (Bace and Mell, 2001), Page 14, Figure 3)

Agent-based IDS is a fully distributed structure or control strategy because of its scalability which enable agents to be located anywhere in a network. Please refer to section 2.10 for further explanation.

d. Data Analysis Approach

Data analysis approaches which are being used in IDS will be described in next section 2.6.

e. Data Source

i. Audit Trail

Audit trail or system event log reports audit for detection of attack manifestations for post-mortem analysis. In analysis timing classification, this processing was classified as interval-based and passively response. From the analysis, identification can be done on successful intrusion and use it for definition of network traffic rules later. From the log files, it has huge abilities in recurring intrusion activity such as unauthorized privileges or abuse and attack attempts for further action. Identification of successful

intruders and system weaknesses can be lead to the prevention of the same attacks through access and user signature. Through its details in logging every events in the system, it can detect all intrusion occurred in a system and can be a very dependable tools in IDS. IDS products using audit trail as a data source are SecureView, CMDS (Computer Misuse Detection System) and ACID (Analysis Console for Intrusion Databases).

ii. Network Packets

Through network packets filtering, system give active and real-time response and information of intrusion detected. Filter and response processes are based on the identified rules or procedure and prolonging to an algorithm. Basically, every event has a special algorithm and the algorithm will be updated every time an intrusion detected. There are several methods used in network packets processing such as by looking for character strings in transport layer in header by IP address that initiates connections or checking for misappropriate TCP/IP flag combination. This processing is the most suitable for a widely exposed to the external sources or internet. Example of IDS used network packets processing are NNIDS, NetSTAT, Snort and Bro.

SAAIDS used network packets as data sources in doing intrusion detection. Please refer to section 2.10 for further explanation.

f. Response Option

i. Active

Active response detects and responses to an attack instantly. It has a chance to protect a system from being hacked and avoiding from halt by blocking all intruders' activities or logging out potential intruders from further action.

ii. Passive

Passive response will log intrusion event and used it for incoming potential attacks.

SAAIDS used active response as response option since SAAIDS run without human intervention or automated system. Please refer to section 2.10 for further explanation.

g. Analysis Timing

i. Real-time-Based

Analysis being done instantly after an intrusion occurred and usually the results of the analysis are known in time of intrusion. This kind of analysis is under active response classification.

ii. Interval-Based

Analysis being done in certain identified time as schedule in the procedure of a system.

This system used real-time based in response to intrusion to avoid any further damage done in the system. Please refer to section 2.10 for further explanation.

2.5.3. Snort as Intrusion Detection Tool

The Snort IDS is one of the most popular intrusion detection platforms available and is an open-source IDS solution, meaning that the source code is free for anyone to use or modify. Snort is capable to detect and response to worms, system crackers and other kinds of attacks (Scott et al., 2004). It use WinPcap to grabs packets from the network and pitches them to Snort. Besides that, Snort is configurable which means that all processes, configuration and rules are depended on user's requirement and can fit all network architecture. New rules can be created for new attacks based on user's new definition of attacks. The process of intrusion detection is concluded in an output

which gathered all needed information. Whenever a network packet hits an Ethernet wire the Snort is sniffing, it takes the following path:

a. Packet capture library

   WinPcap is the packet capture library that gather network packet information that extracts unprocessed Data-Link Layer packets such as Ethernet frames.

b. Packet decoder

   The packet decoder takes the Layer 2 data sent over from packet capture library and takes it apart. First it decodes the Data Link frame, the IP protocol and TCP or UDP packet. When finished decoding, Snort has all the protocols information in all the right places for further processing.

c. Preprocessor

   Snort's preprocessor has several plug-ins that can be turn on or off. Preprocessing operates on the decoded packets, performing a variety of transformations making the data easier for Snort to digest. Preprocessors can alert on, classify or drop a packet before sending it on to the more CPU-intensive detection engine.

d. Detection engine

   The detection engine is the heart of Snort. It takes information from the packet decoder and preprocessors ad operates on it at the transport and application layers, comparing what's in the packet to information in its rules-based detection plug-in which is contains signature for attacks.

e. Output

   When a preprocessor or rule is triggered, an alert is generated and logged. Snort supports a variety of output plug-ins, including its own text and binary-based logging formats, a number of databases and syslog.

Snort is useful in doing information filtering, data analysis and response for SAAIDS. With all services provided as mentioned above, it produces an output which contains

all needed information from data sources by doing signature analysis and enable to alert with real-time response based on rules created on user's requirements. In general mode, when an alert generated, following form of output are produced (this example alerted a Worm propagation attempt):

02/26- 17:59:01635549 [**] [1:2003:2] MS-SQL Worm propagation attempt [**] [Classification: Misc Attack} [Priority: 2] [UDP] Y.Y.250.124:1162 -> X.X.2.27:1434

The alert notified as following:

a. 02/26- 17:59:01635549 - The date and time of the attack.
b. [1:2003:2] - SID
c. MS-SQL Worm propagation attempt – Description of the alert.
d. Classification: Misc Attack – Type of attack
e. Priority: 2 – The alert's priority
f. UDP – Protocol of the attack
g. Y.Y.250.124:1162 -> X.X.2.27:1434 – Source IP address and port number and destination IP address and port number.

The general output of Snort shown above indicates all important information needed in intrusion detection which involves each process in IDS's fundamental as mentioned earlier. Snort is configured to fit SAAIDS's intrusion detection requirements in accomplishing intrusion detection to ensure that SAAIDS research project's objective achieved. Intrusion detection requirements are further discussed in next Chapter 4.

2.6. Data Analysis Approaches

An IDS is depending on data analysis approaches in processing data to provide information neither an attack occurred or not (Kazienko and Dorosz, 2004). Various data processing mechanisms used in IDS recently as described in follows:

2.6.1. Expert System

These work on a previously defined set of rules describing an attack. All security related events incorporated in an audit trail are translated in terms of if-then-else rules. Examples are Wisdom & Sense and ComputerWatch which developed at AT&T.

2.6.2. Signature Analysis

Similarly to Expert System approach, this method is based on the attack knowledge. They transform the semantic description of an attack into the appropriate audit trail format. Thus, attack signatures can be found in logs or input data streams in a straightforward way. An attack scenario can be described, for example, as a sequence of audit events that a given attack generates or patterns of searchable data that are captured in the audit trail. This method uses abstract equivalents of audit trail data. Detection is accomplished by using common text string matching mechanisms. Typically, it is a very powerful technique and as such very often employed in commercial systems. Examples are Stalker, Real Secure, NetRanger, Emerald eXpert-BSM.

2.6.3. Colored Petri Nets

The Colored Petri Nets approach is often used to generalize attacks from expert knowledge bases and to represent attacks graphically. Purdue University's IDIOT system uses Colored Petri Nets. With this technique, it is easy for system administrators to add new signatures to the system. However, matching a complex signature to the audit trail data may be time-consuming. The technique is not used in commercial systems.

2.6.4. State-Transition Analysis

Here, an attack is described with a set of goals and transitions that must be achieved by an intruder to compromise a system. Transitions are represented on state-transition diagrams.

2.6.5. Statistical Analysis

This is a frequently used method such as SECURENET. The user or system behavior or set of attributes is measured by a number of variables over time. Examples of such variables are: user login, logout, number of files accessed in a period of time, usage of disk space, memory and CPU. The frequency of updating can vary from a few minutes to, for example, one month. The system stores mean values for each variable used for detecting exceeds that of a predefined threshold. Yet, this simple approach was unable to match a typical user behavior model. Approaches that relied on matching individual user profiles with aggregated group variables also failed to be efficient. Therefore, a more sophisticated model of user behavior has been developed using short- and long-term user profiles. These profiles are regularly updated to keep up with the changes in user behaviors. Statistical methods are often used in implementations of normal user behavior profile-based Intrusion Detection Systems.

2.6.6. Neural Network

Neural networks use their learning algorithms to learn about the relationship between input and output vectors and to generalize them to extract new input/output relationships. With the neural network approach to intrusion detection, the main purpose is to learn the behavior of actors in the system such as users and daemons. It is known that statistical methods partially equate neural networks. The advantage of using neural networks over statistics resides in having a simple way to express nonlinear relationships between variables, and in learning about relationships automatically. Experiments were carried out with neural network prediction of user behaviors. From the results it has been found that the behavior of UNIX super-users (*roots*) is predictable because of very regular functioning of automatic system processes. With few exceptions, behavior of most other users is also predictable. Neural networks are still a computationally intensive technique, and are not widely used in the intrusion detection community.

2.6.7. Data Mining

Generally refers to a set of techniques that use the process of extracting previously unknown but potentially useful data from large stores of data. Data mining method excels at processing large system logs or audit data. However they are less useful for stream analysis of network traffic. One of the fundamental data mining techniques used in intrusion detection is associated with decision trees. Decision tree models allow one to detect anomalies in large databases. Another technique refers to segmentation, allowing extraction of patterns of unknown attacks. This is done by matching patterns extracted from a simple audit set with those referred to warehoused unknown attacks. A typical data mining technique is associated with finding association rules. It allows one to extract previously unknown knowledge on new attacks or built on normal behavior patterns. Anomaly detection often generates false alarms. With data mining it is easy to correlate data related to alarms with mined audit data, thereby considerably reducing the rate of false alarms.

2.6.8. User Intention Identification

This technique models normal behavior of users by the set of high-level tasks in relation to the users' functions and they have to perform on the system. These tasks are taken as series of actions, which in turn are matched to the appropriate audit data. The analyzer keeps a set of tasks that are acceptable for each user. Whenever a mismatch is encountered, an alarm is produced.

2.6.9. Computer Immunology

Analogies with immunology have lead to the development of a technique that constructs a model of normal behavior of UNIX network services, rather than that of individual users. This model consists of short sequences of system calls made by the processes. Attacks that exploit flaws in the application code are very likely to take unusual execution paths. First, a set of reference audit data is collected which represents the appropriate behavior of services, and then the knowledge base is added with all the known good sequences of system calls. These patterns are then used for

continuous monitoring of system calls to check whether the sequence generated is listed in the knowledge base; if not an alarm is generated. This technique has a potentially very low false alarm rate provided that the knowledge base is fairly complete. Its drawback is the inability to detect errors in the configuration of network services. Whenever an attacker uses legitimate actions on the system to gain unauthorized access, no alarm is generated.

2.6.10. Machine Learning

This is an artificial intelligence technique that stores the user-input stream of commands in a vectorial form and is used as a reference of normal user behavior profile. Profiles are then grouped in a library of user commands having certain common characteristics.

SAAIDS proposed signature analysis as data analysis approach in doing intrusion detection and used pattern matching analysis in doing agent verification process. These approaches will be further discussed in next chapter.

2.7. Cryptology In Intrusion Detection System

Cryptology plays an important role in system or network security since long time ago, but still being the most practical way in defending a system. There are various cryptosystems has been developed and used in IDS, combined with other security mechanisms such as System Vulnerabilities Checkers and Packet Filter. Although existing cryptology has been proved as a most practical and successful mechanism, researchers is still working on enhancement in ensuring stability and reliability of their algorithms. SAAIDS integrates Elgamal encryption and digital signature algorithm which are one of well known Public Key Infrastructure (PKI) cryptosystem. The integration process will be further discussed in next chapter.

2.7.1. Public Key Infrastructure (PKI)

Besides the symmetric cryptosystem, there is asymmetric cryptosystem which overcomes the weakness of symmetric cryptosystem. Asymmetric cryptosystem is also known as Public Key (PK) cryptosystem or Public Key Infrastructure (PKI). PK cryptosystem uses two different keys; a public key to encrypt the message and private key to decrypt it. The public key can only be used in encryption message and private key can only be used to decrypt the sent message. PK cryptosystems are based on concept that public key is known by everyone but private key is only known by each sender and recipient (Kamaruzzaman et al., 2007).This method allows a user to freely distribute the public key to peoples who want to communicate without worry of compromise because only who has the private key can decrypt the message sent. The general concept of PK cryptosystem is shown in Figure 2.9.

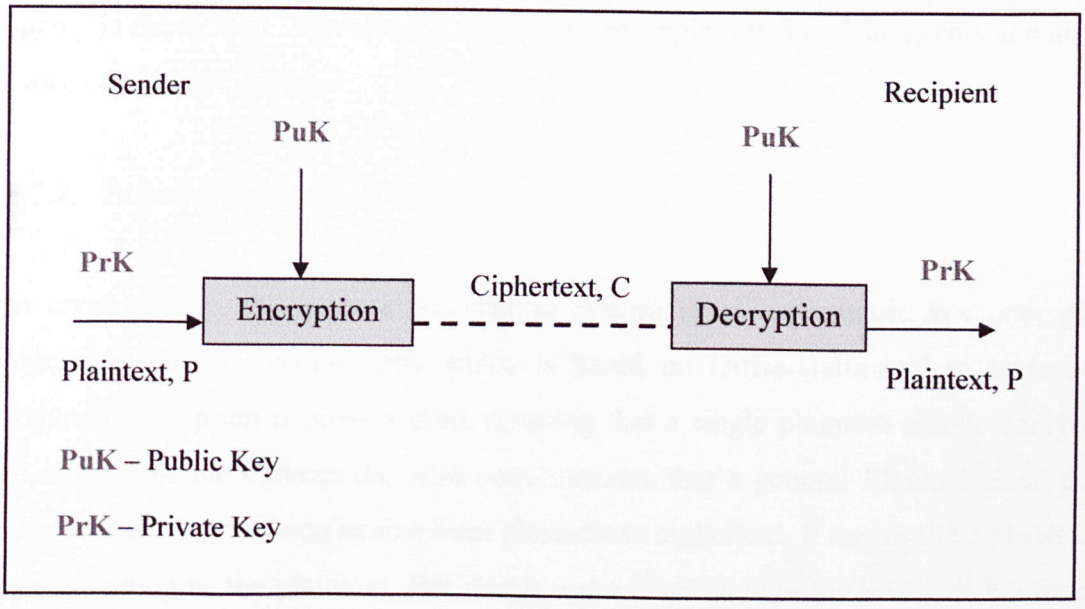

Figure 2.9: Public Key System Concept

Based on the advantage of PK cryptosystem mentioned above, in this project, PK cryptosystem has been chosen to be implemented. There are many PK cryptosystem that are widely used to provide a secured communication such as RSA, Diffie-Hellman and Elgamal. Based on researcher observation and research, to ensure the system is structured and easy to be implemented, Elgamal Encryption has been chosen as method for implementation of PK cryptosystem for this project. Elgamal

Encryption used in agent communication to ensure secured message sending process between agents in the system. The messages are containing with several important contents in running system and have to be sent without risks of potential attacks against it. When agent communication is secured, agents can get optimum functionality in doing their tasks.

In agent communication, Elgamal Encryption used to secure messages sending process in ensuring that the messages sent safely to the destination. The secured channel can provide a safe communication between agents but the other issue is how to ensure that the agent itself is a safe or authorized agent running in the system. With various kind of attacks described before, there is a high risk for agent to be opened for attacks. Either the attacks are coming from other agents or from outsider, it put the agent in danger of being damaged or controlled by attacker. Besides that, after an agent being attacked and controlled by attacker, the damaged agent may attacks other agents in the system. This situation will ruin the original tasks of the agents and all the processes in the system.

2.7.2. Elgamal Encryption

In cryptography, the Elgamal encryption system is an asymmetric key encryption algorithm for PK cryptography which is based on Diffie-Hellman key agreement. Elgamal encryption is probabilistic, meaning that a single plaintext can be encrypted to many possible ciphertexts, with consequences that a general Elgamal Encryption produces a 2:1 expansion in size from plaintext to ciphertext. It means the ciphertext is twice as long as the plaintext. But this is not a disadvantage for SAAIDS because of messages sent between agents just using a small size of contents. It has been proved that Elgamal Encryption is a reliable encryption system which has been based on several other cryptosystem Cramer-Shoup system and DHAES. Besides that, other the advantage of Elgamal Encryption is its algorithm produced its digital signature algorithm which is known as Elgamal Signature Scheme or Elgamal Digital Signature.

2.7.3. SHA1 with Elgamal Digital Signature

In this research project, Elgamal Digital Signature completed with SHA1 hash function which called SHA1 with Elgamal Digital Signature along with Elgamal Encryption will provide a verification system for each agents in the system called agent verification process. Agent verification is designed to verify that each agent in the system is an authorized agent. Besides that, agent verification detects if there is any unauthorized agent or damaged agent running in the system and response to it. The description of Elgamal Encryption and SHA1 with Elgamal Digital Signature using will be further discussed in next chapter. Figure 2.10 shows the concept of SHA1 with Elgamal Digital Signature.

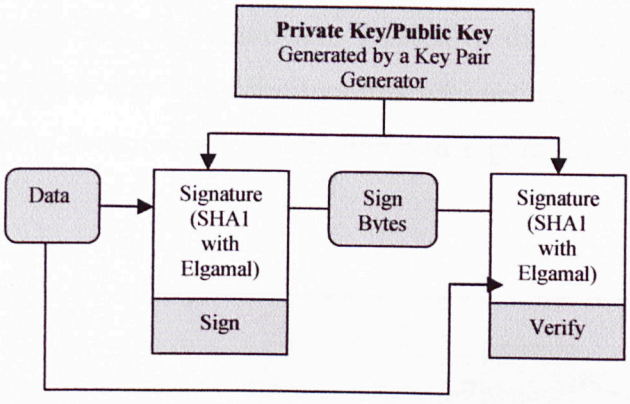

Figure 2.10: SHA1 with Elgamal Digital Signature

2.8. Related Works

Agent-based Intrusion Detection System (IDS) is one of famous IDS technologies for time being. The progress of agent-based IDS is very impressive since then many researches related done because of its advantages such as overcoming network latency, system scalability reduce network load and platform independence (Albag, 2005). Many of the researches conducted with various mechanisms in building agent-based IDS. In this research, there are 9 researches has been studied; AAFID (Autonomous Agent for Intrusion Detection), PAID (Probabilistic Agent-Based Intrusion Detection System), IDA (Intrusion Detection Agent System), INDRA,

Hummingbird, JAM (Java Agents for Meta Learning), Intelligent Agents for Intrusion Detection, Advanced Telecommunications/Information Distribution Program and MA (Mobile Agent). This section discusses three of them in term of recent trends in several previous works in agent-based IDS as follows:

2.8.1. AAFID (Autonomous Agent for Intrusion Detection)

AAFID (Autonomous Agent for Intrusion Detection) uses autonomous agents at lower levels to collect, analyze and filter the data. It consists three essential components in the architecture; agents, transceivers and monitors. These agents are run and controlled at every host by local transceivers. All the agents in a host report their findings to a single transceiver. Transceivers in turn report to monitors. A single monitor may control several transceivers located at different hosts and may itself report to other monitors above it. Monitors this way get a global view of the network that facilitates in the decision process as shown in Figure 2.11 (Balasubramaniyan et al., 1998).

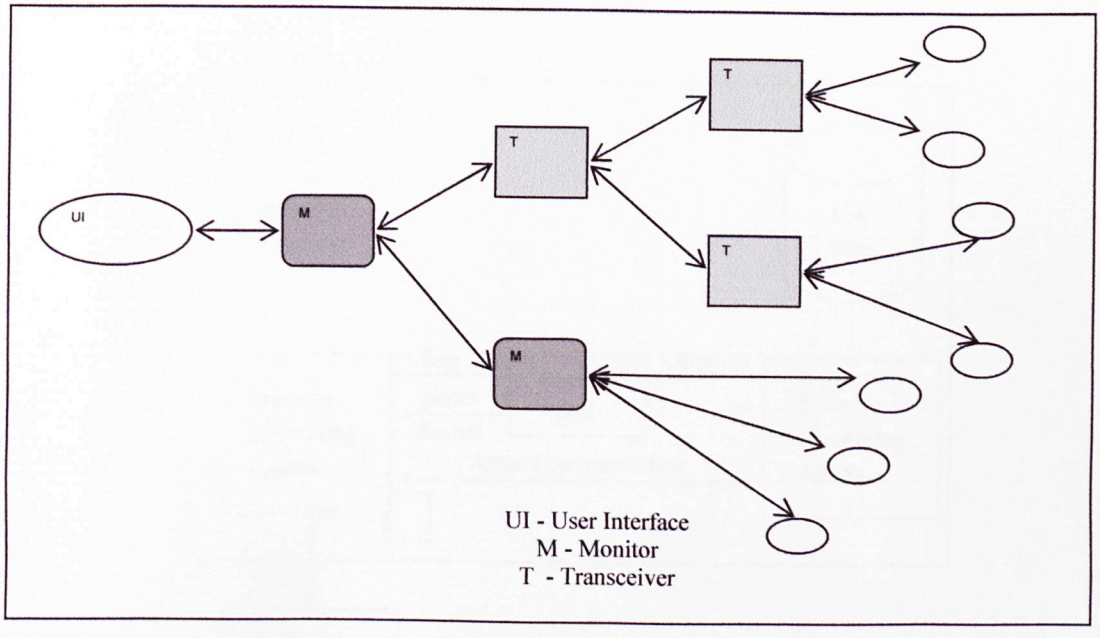

Figure 2.11: AAFID Overview

(Source from: Balasubramaniyan et al., 1998)

2.8.2. PAID (Probabilistic Agent-Based Intrusion Detection System)

PAID (Probabilistic Agent-Based Intrusion Detection System) allows agents to share their beliefs, such as the probability distribution of an event occurrence. Agents are capable to perform soft-evidential update, thus providing a continuous scale for intrusion detection along with methods for modeling errors and resolving conflicts among beliefs (Gowadia et al., 2005).

PAID components are consisting of intrusion monitoring agents, Registry Agents and System Monitoring Agents. Intrusion Monitoring Agents collect analyze and filter data in networks and communicates with System Monitoring Agents which has connection with log files. Between both agents is Registry Agent as registrar of all agents. Intrusion Monitoring Agents always connected with system monitoring agents through Agent Communication to send messages between them including filtered data. System Monitoring System stores all needed information into Log Files and Intrusion Monitoring System sends all filtered data to Intrusion Probability in detecting intrusion process. Figure 2.12 shows PAID Overview.

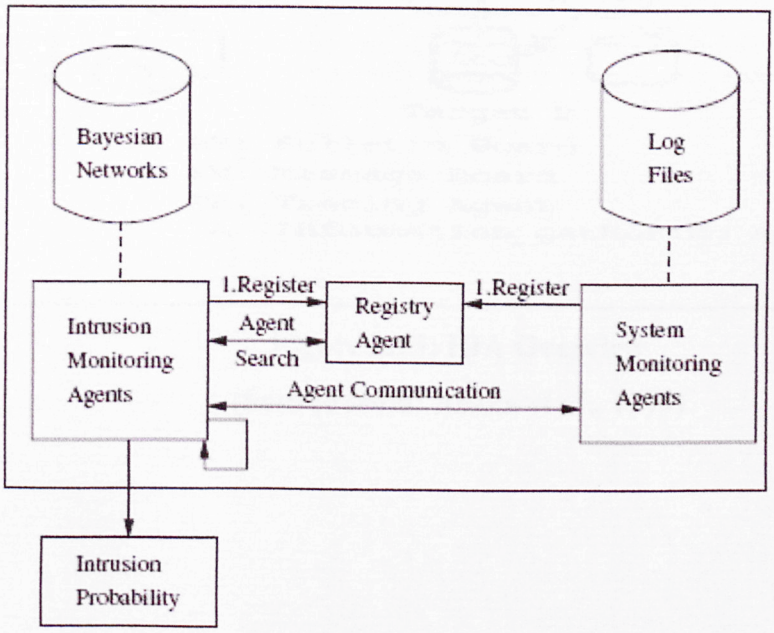

Figure 2.12: PAID Overview

(Source from: Gowadia et al., 2005)

2.8.3. IDA (Intrusion Detection Agent System)

IDA (Intrusion Detection Agent System) was developed by Information-technology Promotion Agency (IPA). The IDA is a multi host-based IDS. Instead of analyzing all of the users' activities, IDA works by watching specific events that may relate to intrusions, referred to as Marks Left by Suspected Intruder (MLSI). If an MLSI is found, IDA gathers information related to the MLSI, analyzes the information, and decides whether or not an intrusion has occurred (Asaka et al., 1999). The architecture of IDA consists Bulletin Board, Message Board, Tracing Agent and Information_gathering Agent. Figure 2.13 shows IDA overview.

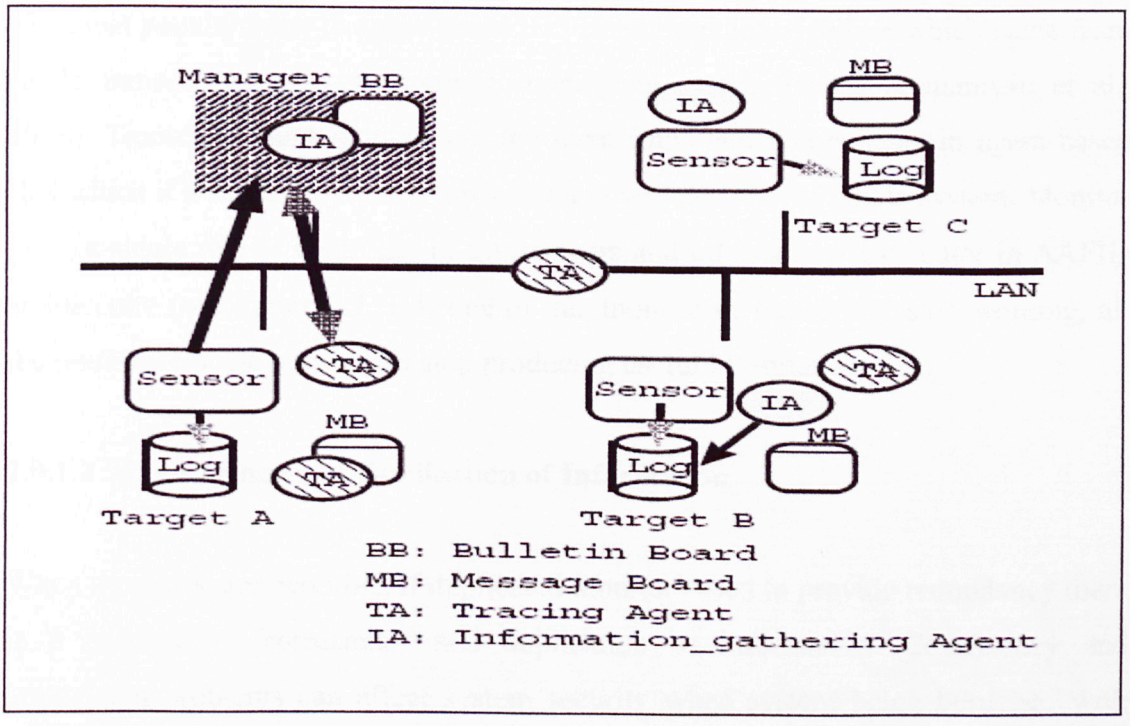

Figure 2.13: IDA Overview

(Source from: Asaka et al., 1999)

2.9. Issues and Problems in Previous Works

2.9.1. Architecture Issues

Existing architecture of agent-based IDS has several weaknesses as determined by several researchers. Follows are the most mentioned weaknesses include single point of failure, consistency and duplication of information and delay in information sending.

2.9.1.1 Single Point of Failure

The most popular issue in agent-based IDS is single-point of failure which came from single transceiver/analyzer or single monitoring system (Balasubramaniyan et al., 1998). Transceiver and monitor are the most important components in agent-based IDS which if a successful attack towards them will destroy the whole system. Monitor plays a single role as controller of transceivers and information distributor in AAFID architecture (see Figure 2.13). If one of this monitor or transceiver stop working, all the transceivers and agents will stop producing useful information.

2.9.1.2 Consistency and Duplication of Information

When monitors stop working, if duplicated monitors used to provide redundancy there is a problem of consistency and duplication of information. Consistency and duplication problems can affect system security when system being burdened with doubled useless information and causes conflicts and errors in determining suitable information to be used in detection process. System security will be threatened when conflicts and errors occurs which may cause monitor halted and stop system functioning continuously.

2.9.1.3 Delay in Information Sending

The main factor of delay in detection process has been determined as delay information sending due to hierarchical structure of existing IDS. Detection of

intrusions at the monitor level is delayed until all the necessary information gets there from the agents and transceiver (Balasubramaniyan et al., 1998; Douglas et al., 2001). Monitor and Transceiver are situated at higher level in hierarchical structure which as parents they have their own child in lower level. Information of detection will only being sent and broadcast to other monitors when all necessary information needed gathered from agents and transceiver. Delay on detection and information broadcast will open chances for intruder to use the delay as their key or weapon for other attacks.

2.9.2. Hierarchical Structure Caused Multilevel Authorization Problem

Another problem of existing agent-based IDS is on hierarchical structure which causes multilevel authorization problem in monitor, transceiver and agent level (Balasubramaniyan et al., 1998). This architecture controlled by its authorized persons with different hierarchies in each level. When the system deployed into a large area of network, the authorization system will be more complicated in determining:

a. Who are the authorized persons able to be given the authority?
b. Where are the authorized areas able for the authorized persons?
c. When are the authorized persons able to be given authority for certain area authorized?

Questions stated above are among more question in designing authorization for monitor, transceiver and agent. In large area network, for example has 10,000 nodes, 1000 transceiver and 10 monitors, theoretically, system has to manage authorization for 100,000,000 different places and yet to mention amount of the authorized persons for each places.

The process of a big amount authorization may burdensome the system itself. Big size of data store has to be provided to store big size of authorization information. Furthermore, delay on authorization process may occurred by searching process of right information or doing too much authorization concurrently. This multilevel authorization problem which came from hierarchical structure of existing agent-based

IDS may give a big impact to the system performance as well as security might be threatened when delay in information sending problem occurred as stated above.

2.9.3. Attacks Against Agent Communication

Agent-based IDS facing some problems in security, code size and performance issues, see (Balasubramaniyan et al., 1998; Albag, 2005). Actually, these problems came from many kinds of threats and attacks against agent communication such as execution of exploits of different agents, malicious activities against running platform and crash attacks against the system; agents itself (Brown et al., 2001). Execution of exploits might be consisting of viruses, worms or Trojans which are able to do damage and steal classified information in the system.

A worm is a program that self propagates across a network exploiting security or policy flaws in widely used services (Weaver et al., *2003*), while a Trojan is a program that appears to legitimate but it is designed to perform some destructive activities. Trojan is a program that, enter computers appearing to be harmless programs, install themselves and carry out actions that affect user confidentiality (Pandasoftware, 2007). All these threats are called malicious agents (Hao et al., 2006) and attempt to damage and disable the system.

2.9.4. Existence of Unauthorized Agents

Unauthorized agent is a new issue in agent-based IDS. This problem is coming from threats and attacks against agent communication as stated in section 2.10.3. All the threats and attacks attempts could open chances for intruder to get confidential information and lead to the unauthorized agent problem. Malicious agents or stealing information of an agent may lead to duplication of function of an agent and may form a fake or an unauthorized agent running in the system. For example, if information about duplication of agent being sent during agent communication process and attacks occurred tends to steal the classified information, an unauthorized agent might be built and running in system. Addition to this problem, Trojan is a program that, enter

computers appearing to be harmless programs, install themselves and carry out actions that affect agent confidentiality.

Successful attacks against agent which came from either attacks against agent communication or the agent itself, the intruded agent can open chances for attacks against other agents in the system. The intruded agent can be faked to be an authorized agent which either with duplicate the original agent into fake agent or change the roles of the existing agent. Probabilistic Agent-Based Intrusion Detection System (PAID) using a single registry agent in every host to control new agent registration in the system similar as transceiver functionality in AAFID (Gowadia et al., 2005). This issue also can lead to authorized agents running in system if the transceiver or registry agent down although for seconds and may knock down agent-based IDS tremendously.

Mobile Agent (MA) The MA computing paradigm presents a number of security threats that are not addressed by conventional security techniques. Standard security techniques must be modified or new techniques invented to address these threats (Jansen W. et al., 1999). The researchers stated that the security threats can be classified into four broad categories:

a. Agent-to-agent

   Represents the set of threats in which agents exploit security weaknesses of other agents or launch attacks against other agents.

b. Agent-to-platform

   Represents the set of threats in which agents exploit security weaknesses of or launch attacks against an agent platform.

c. Platform-to-agent

   Represents the set of threats in which platforms compromise the security of agents.

d. Other-to-agent platform

   Represents the set of threats in which external entities, including agents and agent platforms, threaten the security of an agent platform.

This research tends to overcome all categories of security threats abovementioned. Several techniques has been suggested by the researchers include mechanisms to control access to computational resources, cryptographic methods to encipher information exchanges, cryptographic methods to identify and authenticate users, agents, and platforms, and mechanisms to audit security relevant events occurring at the agent platform.

All these threats and attacks attempt to do damage and disable the system. In avoiding such damages from happens, two main elements of agent-based IDS have to be considered; agent communication security and agent security (itself) (Brown et al., 2001). To ensure the security of both elements guaranteed, the proposed techniques are including agent communication protocol and agent verification protocol. Proposed agent verification protocol is designed to ensure that each agents running in the system is an authorized agent. Cryptosystem mechanisms used in designing both protocol; Elgamal Encryption for Agent Communication Protocol and Elgamal Digital Signature for Agent Verification Protocol. Both of these protocols will be further discussed in next chapter.

2.10. Proposed Solution

The researcher proposes here an architecture of SAAIDS which contains all important components in an agent. The agents communicate with each other through direct connection which is also known as peer-to-peer (P2P) connection to avoid the existing architecture problems. Then, agent communication protocol and algorithm is proposed to provide secured connection between agents in solving the attacks against agent communication. Finally, the researcher proposed a secure protocol and algorithm that can be used to detect the presence of fake or unauthorized agents in our network infrastructures. Both mechanisms comprised of Elgamal encryption algorithm, digital

signature, and SHA-1 message digest function. The Elgamal encryption algorithm is used to provide message confidentiality by encrypting the messages to be sent, digital signature is used as signing function and the SHA-1 function is to generate a code that is unique for each agent to be used in the verification process.

2.11. Conclusion

From literature review it can be concluded that a depth understanding of attacks, threat and intrusion, and attacks classification in SAAIDS are needed, which later are used for system development for this project. Attacks against agent communication and the agent itself have been suggested as classification of attacks in this system. After that, intrusion detection system has been further discussed in defining intrusion detection and briefly discussed about its classification which is can provide better understanding for anyone interested in studying IDS. Then, complementary tools which are used to do intrusion detection are briefly discussed and assumed to be already equipped in the system. A new classification of intrusion detection system has been produced for this research then followed by details discussion about data analysis approach which is as an important part in IDS. Signature analysis and pattern matching analysis have been chosen as data analysis approaches in providing secured agents and agent communication running in the system. In providing secured agent and agent communication, PK cryptosystem has been determined to be used in the system. Elgamal Encryption and Elgamal Digital Signature are the PK cryptosystems used to ensure the safety of agent and messages sending processes. Three previous works has been discussed through it trends of architecture and methods used, which are AAFID, PAID and IDA. Based on research upon these three researches in IDS, several issues and problems have been determined and tends to be solved with the proposed system; SAAIDS. Finally, to ensure that the research development reaches its objective, proposed solution has been outlined in the last section.