Multiple Simultaneous Threat Detection in UNIX Environment

Zafar Sultan†

School of Science and Technology, University of New England, Armidale, NSW, Australia

Summary
Although UNIX is considered a very stable and secure platform, the development of Intrusion Detection Systems (IDS) is essential as current and future generations of hackers are continuously attempting to undermine its integrity. The empirical experiment of multiple simultaneous threat detection system proved that use of hybrid data fusion model of Bayesian, Dempster Shafer and extended Dempster Shafer increased an average 20% threat detection rate. The false positive rate also went down by 51%. The use of Extended Dempster Shafer to combine probability mass of 4 intrusion detection (Multisensor) systems increased precision of threat detection by 36% whilst the initial probability mass of the Dempster Shafer of Multisensor was only 0.03.

Set Cover as a middle tier data fusion tool produced incredible results, particularly in data grouping by reducing the population size from 2273 to 429 that amazingly minimise the computational processing cpu and memory overhead cost and time. In order to improve the results of the precision of the multiple simultaneous threat detection system, as a next step of my research is that is an extension to the Bayesian and Dempster Shafer theory. GEP presents a better evidential combination and separate propositions and the decisions.

Key words:
Multiple Simultaneous Threat Detection; Intrusion Detection Systems; Bayesian Theory; Dempster Shafer, Multisensor Data Fusion; Extended Dempster Shafer, Set Cover; Set Packing; GEP; UNIX .

1. Introduction

A large number of Intrusion Detection Systems have been developed for computer security but more development is required as attackers are very shrewd these days and have developed different approaches and programmes to penetrate into computer systems and have succeeded many times in breaking all security walls. Thus hackers, in fact, not only have stolen valued and critical business data but also forced computer industry and businesses to develop advance software to monitor and block their attacks. As a result the companies have to spend billion dollars to develop preventive codes for this purpose [4]. For example Microsoft spent $1.2 billion to stop Sapphire/Slammer worm in 2003 [17] [22].

Integration of UNIX with Firewall protection and CISCO technology were considered very secure systems but hackers have also broken such security measures. The way the security field is progressing, it looks like this is a continuous battle between security professionals and hackers. Hackers are in reality people familiar with all types of computer systems like cyberspace, networks, operating systems and their thousands of applications. Hackers know the loopholes of information technology systems, they exploit system weaknesses and misuse their expertise to perform illegal functions on business critical systems such as stealing important information, business secrets, damaging data or systems etc. etc. The hardest problem in tracking these types of attack, their origin and quantity of damage depends upon attacker's software and techniques. Hackers may attack from multiple sites and hide their identity by continuously changing their IP addresses. Sometimes they do physical damage to the systems or their applications, but if they just steal important information, the security experts may be unaware of it for many months until they apply a new security update or hackers does any physical damage to any process or data of that particular business [5] [21].

False positives and false negatives are additional issues in computer security and also in UNIX systems. False alarm results because alarms are set at low levels of security. The present monitoring and IDS analyse data taken from system processes, memory, CPU, disk utilization and log messages and track or batch or log files. Attacks are checked based on pattern matching with existing situations of the processes and systems attributes [1] [7].

The aim of this paper is the preliminary experimental evaluation of the multiple threat detection system using Multisensor data fusion, its various approaches and techniques. However, the main emphasis of my research is to detect multiple simultaneous attacks in UNIX environments. My research will help in building multiple simultaneous threat detection system for computer security in general and for UNIX environments in particular.

2. Existing Threat Detection Approaches

2.1 Data Fusion Approaches in UNIX

Bayesian, Dempster Shafer, fuzzy rules, parametric / non parametric and Kalman Filter are widely used data fusion techniques [5] [4] [11]. Chapman-Kalmogorov prediction model has also been used as an integral model with

Manuscript received February 5, 2009
Manuscript revised February 20, 2009

Bayesian and Dempster Shafer [9] [14]. Inferences regarding threats, location and other attributes are made from these models. These models fuse data from the Multisensor systems on the same or different networks. Fusion model behave exactly the same way like human brain process data and take actions or decisions. UNIX system's IDS get their data from different sensor created by systems commands and networks packets. Data may be sniffer packets; sys log files, SNMP traces, system messages and other similar activities of the network [6]. Data fusion model after processing this information send its outputs in form of alarms to security people and system engineers and warn of any expected threat on a particular subnet. Though data fusion models work like cognitive approach but in fact they are not really intelligent enough to cope with different type of changes or attacks if their information does not already exist in the IDS database. The Langley attack lost million dollars and they could not find email bombs until their business server crashed.

The current Multiple Intrusion Detection Models are unable to auto track, identify, and block all suspected threats. Advance IDS are required to deliver enhanced reliability and precision in threat detection [1]. Thus additional development is required in the field of multiple sensor data fusion models of IDS in UNIX [16] [19].

2.2 Other Data Fusion Approaches

A large amount of research work and literature is available on Multisensor data fusion of IDS in defense and other related fields. However, there is a little work in the field of UNIX, only few scientists worked on multiple simultaneous threat detection in UNIX. It is, therefore, a relatively new area to work on. Though a few yeas back UNIX was one of the secure environments from outside hackers but intruders now have broken many business applications and databases in UNIX network whilst all critical business like credit cards, client profiles and financial transactions are online and need more security ever than before.

Majority of the workers used Bayesian, Dempster Shafer, parametric / non parametric and few others inference engines for Multisensor data analysis in their IDS.

Dong and Deborah [10] worked on DARPA IDS evaluation data set show in their experiments that improved threat detection rates from 75 to 94 % with their hybrid models. In another study, Dong and Deborah emphasized that hybrid model of Bayesian is the best technique to improve the intrusion detection precision for IDS.

Christos and Basil [21] worked on multiple data fusion model and concluded that the use of Multisensor data analysis increases threat detection accuracy. They used a Bayesian and Dempster Shafer detection engine.

Huadong Wu, Mel Siegel and Rainer [11] identified relationship between Bayesian and Dempster Shafer theory and compared with the probability method and concluded that combined mathematical inference models will be a promising area for Multisensor data analysis in IDS.

A. Habib, M Hefeeda [2] and Christos [21] worked on DoS in an IDS and found a increase in precision by using classical Bayesian methods for data analysis.

Diego Zamoni [8] used a pattern matching detection model to detect new attacks, however, he did not mention any particular fusion model in his experiment.

V. Chatzigiannakis, A. Lenis, C. Siaterlis, M. and Grammatikou D [24] found that their fusion model is more effective than single metric analysis. They used Principal Component Analysis for Multisensor data fusion for intrusion detection.

Vladimir G, Oleg K, and Vladimir S [25] suggested that combining a decision model is better in thereat detection precision than a Meta model in IDS.

Kapil K S [15] worked on IDS architecture and found that rule set knowledge, expert systems state models and string match are useful parameters in the development of an advance threat detection model.

Hugh Durrant-Whyte and Mike Stevens [12] described mathematical model for their fusion model. They analysed data using Kalman Filter and theoretical methods derived from Bayesian theorem.

S Terry Brugger, [20] worked on offline data fusion model, used data mining approach in her IDS. However, she did not produce any particular model during her experiment.

In the view of all above literature reviews, it is obvious that there is enough material on Multisensor data fusion models of IDS. However, very little was reported in the UNIX. And almost negligible work was found if we search material or study on multiple simultaneous threat detection in the field of UNIX.

2.3 Research Directions

In this research, I'll identify a multiple simultaneous threat detection model. This model will be a hybrid of Bayesian and Dempster Shafer theory of inferences with Set Cover theory. The new model will increase the precision in threat detection and reduce the volume of false alarms in UNIX environment. The use of the model will assist in decreasing the data security expenses, particularly web based businesses. Researchers will get also benefit for future IDS developments in UNIX.

The new multiple simultaneous threat detection model will be able to detect more than one threat simultaneously. Another advantage is that the results of this research can be applied in high speed networks like cyberspace. There are also some additional situational parameters that will be generated as a result of this work such as high level architecture of multiple threat detection model, identification of proper Multisensor environment based on hybrid model, and identification of middle tiers of the research.

3 Original Contributions

In order to detect multiple threat detection, researchers are making efforts in order to develop suitable data fusion model based on advanced mathematical and statistical techniques. However, most of the models detect single threats, few models are advanced but the work in multiple threat detection is rare in UNIX environment [13].

3.1 Novelties of multiple simultaneous threat detection

This research, in fact, is a step forward that addresses the additional precision in multiple threat detection process as compared to the existing threat detection approaches in UNIX and it is different in many ways from other's work in Multisensor data fusion in IDS development.

1) I used hybrid model of Multisensor data fusion comprised of basic Bayesian, Dempster Shafer, and Extended Dempster Shafer theory of inference in multiple simultaneous threat detection of UNIX environment.

2) Set Cover as a middle tier data fusion tool in hybrid Bayesian and Dempster model is a novel approach as no one has used it before.

3) Generalized Evidential processing (GEP) presents a better evidential combination and separate propositions and the decisions. GEP will be implemented very first time in a distributed Multisensor network of an UNIX environment.

3.2 Set Cover

Set Cover is a branch of mathematics and in this research I deal with sets, subsets and their interaction sets. Set Cover is the basic system of mathematics. Simple facts of set union and its subsets are used in cover sets of multiple simultaneous threat detection system that is a basic branch of mathematics [3].

In multiple simultaneous threat detection system the total numbers of elements were 2274 denoted by:-

$$\bigcup = \sum_{i=1}^{n} u_i \qquad (1)$$

Where $\bigcup$ is the universal set and $\sum_{i=1}^{n} u_i$ is sum of all the elements in the universal set

In the experiment, the types of threats represented by subsets $S_1, S_2, S_3, ..., S_n \subseteq \bigcup$ And the cost of each set is $C_1, C_2, C_3, ..., C_n$.

In our case threat(s) are present in different data substrings from any of the 4 different intrusion detection systems of a distributed Unix Network. The target is to find the sets P = {1, 2,…, n} that must contain minimum number of strings having threats so that each set have all the relevant strings of data and summation of sets will have all the strings of the inputs. Cover set using greedy algorithm also provides minimum cost represented by Q.

$$Q = \sum_{i=1}^{n} C_i \qquad (2)$$

Where $\sum_{i=1}^{n} C_i$ is the sum of the costs in selecting a new node of the experiment

The cost effectiveness to select computer node is denoted by β

$$\beta = \frac{C(Q)}{Q - P} \qquad (3)$$

Where $C(Q)$ is the initial cost for selecting the nodes for each intrusion detection system and P is the set with minimum elements and Q is the minimum cost in selecting the new node.

4 Approaches and Methodology

In a large number of Multisensor data fusion model, Bayesian and Dempster Shafer have been used for data analysis. Most of the existing work was in single threat detection [11]. Only couple of researchers tried to focus on multiple threat detection without using Set Cover theory. Set Cover has been identified as a new area which can be used to prioritize and schedule rule set on certain criteria in the fusion process. On this topic there are only a few papers available in the UNIX environment, therefore, it is difficult to compare literature on Multisensor threat detection in UNIX [5] [8].

Statistical /mathematical models such as Parametric / Non Parametric, Bayesian, Dempster are the most commonly used theory of estimation in data fusion in UNIX environment. These experimental models were used in detection of DoS, email bombs and buffer overflow attacks with many limitations.

This research is about identification of multiple simultaneous threat detection models. Data fusion will be done using hybrid model of Bayesian and Dempster. Set Cover will be used to identify data groups and scheduling [3]. The hybrid model will provide an increase in precision of threats detected and additional theoretical and technical

knowledge about multiple threat detection for computer security, especially for UNIX [9].

As the multiple simultaneous threat detection system is a future prospect for IDS development in the UNIX environment. In order to make it a useful work, I'll identify the following parameters in my research;-

.Identify and implement multiple simultaneous threat detection model targeting future Intrusion Detection systems

.identify high-level model/architecture that can address Multiple Simultaneous attacks in UNIX

.Identify proper Multisensor data environment to use in the fusion model

.Identify and Implement and run testing environment for the data fusion algorithms in multiple simultaneous threat detection systems

.Identify if my new research on multiple simultaneous threat detection model works well or not? And provide all possible reasons in any case

.I'll provide an excellent comparison of models based on different mathematical inferences

4. 1.Multiple simultaneous threat detection system

The main target of this research is to identify the exact threat(s) with a high degree of precision by using hybrid data fusion model comprised of Set cover, Bayesian theory of estimation, Dempster and Extended Dempster Shafer theory. The origin and directions of the threats are exclusive of this research as that includes complicated, extensive and separate research.

In this distributed test environment which is conceptually the same as server client environment, a multiple simultaneous threat detection system has been set up on different nodes across the distributed subnets. Computer nodes are comprised of multiple operating systems and located at different networks, predominantly UNIX though include Wintel machines as well. Each computer node has different intrusion detection system that filters all the network data and collects threat related information and transfers them to the computer node hosting multiple simultaneous threat detection system for further accuracy and precision of the threat detection results. The computer nodes across different subnets receive different threats. As this is a controlled experiment, 4 types of threats mainly denial of service, man-in-the-middle, buffer overflow and Trojan will be initiated from one of the experimental computer node.

4. 2 Architecture of Multiple simultaneous threat detection system

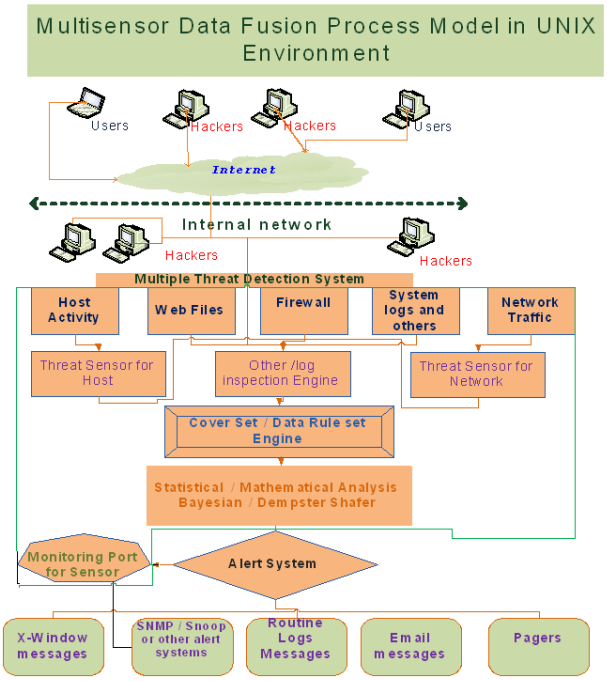

Fig 1 showing the architecture of the multiple simultaneous threat detection system

In this experimental test environment, 4 independent intrusion detection systems work as a separate Multisensor observers on different subnets. In order to monitor all data packets in the test environment, I used a switch on test network and configured a monitoring port to replicate all packets of the data traffic passing through the switch. Network data of layer 2 and 3 was also gathered. Data collecting software then decodes and analyse the data. I used following software for data collection:- MARS, Sniffers, Snoop and Wireshark. Four types of threats:- DoS, Denial of Service, Mom- man-in-the-middle attack or bucket-brigade attack or Janus attack, Buffer overflow or buffer overrun and Trojan Horses

Each intrusion detection system collects network data and filters it using Cover set theory. Data may contain a single, two, three, four or any combinations of the above 4 x threats or false alarms and then move the data to next level of the data fusion within the multiple simultaneous threat detection system. The Multiple simultaneous threat detection system processes the data through different statistical and mathematical techniques and makes decision about the threats.

Then multiple simultaneous threat detection system's client nodes that exist on each computer uses the Set Cover Model as a middle tier data fusion tool which refines the data into small group of sets and schedules these groups of data for onward statistical and

mathematical data fusion. Another benefit of the Set cover model is to choose computer nodes that cover all the anticipated threats at a minimum cost.

5. Results of the Test Experiment

Set Cover Fusion Model

The initial cost for selecting the nodes for each intrusion detection system is:-

$C_{(A)}$ = 8 = cost of the 1st Node

$C_{(B)}$ = 5 = cost of the 2nd Node

$C_{(C)}$ = 12 = cost of the 3rd Node

$C_{(D)}$ = 8 = cost of the 4th Node

The sets with minimum number of elements denoted by P and set with minimum cost Q for each node were determined during the experiment. The total number of sets whose cost was lowest and set with minimum number of elements covered by node A, B, C and D of the experiment are given as:-

$P_{(A)} = 0,\ P_{(B)}=3,\ P_{(C)}=3,\ P_{(D)}=4$

$Q_{(A)}=8,\ Q_{(B)}=4,\ Q_{(C)}=7,\ Q_{(D)}=8$

Here I like to make it clear that the above values are the number of the sets not the elements of the sets, therefore, it should not cause any problem or mix up whilst reading Table 1. The cost effectiveness to select computer node A, B, C and D are calculated using equation (3)

Selecting A: $$\beta_A = \frac{C(A)}{Q(A)-P(A)} = \frac{8}{8-0} = 1 \quad (4)$$

Selecting B: $$\beta_B = \frac{C(B)}{Q(B)-P(B)} = \frac{5}{4-3} = 5 \quad (5)$$

Selecting C: $$\beta_C = \frac{C(C)}{Q(C)-P(C)} = \frac{12}{7-3} = 3 \quad (6)$$

Selecting D: $$\beta_D = \frac{C(D)}{Q(D)-P(D)} = \frac{8}{8-4} = 2 \quad (7)$$

As per above equations, the cost effectiveness of the node are A, D, C and B respectively. Initial total cost of selecting these nodes was 8+5+12+8=33 that is not optimal. The optimal cost as per cost effectiveness of the nodes would be A+D+C=8+8+12=28.

5.1 Set Cover's set generation

In order to collect 4 types of threats, 4 intrusion detection systems collected 2274 malicious substrings of 15 and above bytes from the experimental network. Set covering is a complex problem in information technology because of the complexity of NP-complete problems. However, it was not very hard in my research as I already had used well known intrusion detection systems to collect threat data and each of them gathered data containing all of the threats. The threat data was a mixture of all 4 types of generated threats. The second issue was cost effectiveness in choosing the computer node. Here are the big benefits that I achieved were minimizing the size of the sets and cost effectiveness by using Set Packing and Greedy algorithm respectively. Set Packing provided me the ability to select the K = 4 number of subsets out of the union set N of 2274 such that each subset is a pair wise disjoint to other subsets. Thus each subset now has similar strings of the threat data whose union is N.

In order to find out pair wise disjoint subsets, I analysed N=2274 threat data using a small pearl script. The script separated pair wise disjoint strings of the threat data into 4 subsets of total 429 substrings. The number of elements or substring in each subset is given as:-

Threat Data of Intrusion Detection Systems		
IDS	Before Set Cover	After Set Cover
Wireshark	128	24
Sniffers	439	83
Snoop	646	122
MARS	1061	200
Total	2274	429

Table 1 showing the set cover subsets reduced the sizes of the sets

The client node sends all the above filtered data to next level of the data fusion system of the multiple simultaneous threat detection system. The multiple simultaneous threat detection system combines all the multisensor threat data that has already been filtered into different sets of minimum size using Set Cover model. In order to detect the real threats, improve the accuracy and precision in threat detection; the multiple simultaneous threat detection system fuses the multisensor data with Bayesian theory of estimation, Dempster and Extended Dempster Shafer theory.

5.2 Dempster Shafer Theory to fuse Data

In this experiment, frame of discernment θ will be a set of elemental propositions or combinations of the hypothesis statements. Threats denoted by T may be overlapping or different to each other. In the set of n mutually exclusive and exhaustive set of hypotheses about threat(s) $T \ldots T_n$.

$$\Theta = \{T_1, T_2 \ldots T_n\} \quad (8)$$

If θ have set of n hypotheses, Boolean combination of the set will be θ^n hypotheses.

Dempster Shafer theory does not calculate the probability of a hypothesis but helps in finding out the probability of the evidential support for a hypothesis.

Unlike Bayesian and classical theory of inference, Dempster Shafer theory of inference helps in developing probability mass m (θ) by assigning evidence to each propositions or general propositions. Each intrusion detection system can assign evidence via probability mass to each of the 4 threats, e.g. M1 (T1), M2 (T2), M3 (T3) and M4 (T4). The total probability masses of all the propositions including general propositions will be equal to 1. The probability mass is represented as:-

$$m(\theta) \leq 1 \quad (9)$$

$$\sum_{i=1}^{n} m(\theta) = 1 \quad (10)$$

m (θ) is the probability mass of any possible hypotheses. In this experiment that may be a single threat or combinations of the 4 threats.

5.3 Propositions / Hypothesis

A hypothesis may be a proposition whilst a proposition can be a hypothesis or combinations of hypotheses. In this experiment, I've 4 sensors (intrusion detection systems) and 4 different types of threats. Sensors can receive a single threat or any possible combinations of the 4 generated threats. The total possible base propositions using mathematical theory of combinatorics with and without repetitions are 340 and 15 respectively. As repeated threats are of no significance during hypothesis testing and will also unnecessary increase processing cost and time. Therefore, I'll only concentrate on the propositions without repetitions.

Only 6435 (15 x 429) non repetitive propositions will be processed and tested by MTDS engine as compared to 773160 (340 x 2274) with repetitive propositions.

The general Combinations and Permutations formula is:

$$P(n,r) = \frac{n!}{r!(n-r)!} \quad (11)$$

Where n is the number of sensors (Intrusion Detection System), r is the number of threats to be selected ($0 \leq r \leq n$), where n =4 in this experiment, if r = n, P(n, r) = n!

Case 1: when single threat detected by each sensor, total # of hypothesis / propositions with and without repetitions would be 4 and 4

Case 2: when two threats detected by each sensor, total # of hypothesis / propositions with and without repetitions would be 6 and 16

Case 3: when three threats detected by each sensor, total # of hypothesis / propositions with and without repetitions would be 4 and 64

Case 4: when four threats detected by each sensor, total # of hypothesis / propositions with and without repetitions would be 1 and 256

Limitation

Due to high complexity of the probability mass and weights calculations, it is not possible for me to cover all the 15 non repetitive hypotheses during my research. Therefore, I'll test only four elementary hypotheses as mentioned in section 5.3.

5.4 Fusion without using the Weights of the intrusion detection systems

This experiment has 4 types of intrusion detection systems and each has its own way of threat detection. This means each intrusion detection system has different perception and reliability that it provides to multiple simultaneous threat detection system.

The Dempster Shafer model to combine the probability masses of the threats from more than two independent intrusion detection systems:-

$$\sum_{i=0}^{n} Mi(Ti) = \sum_{i=0}^{n} \frac{p(\{Ti\})}{p(\{Ti\}) + p(\{\neg Ti\})} \quad (12)$$

Where $Mi(Ti)$ is probability mass function, T is the threat(s) and $P(\{Ti\})$ is the probability of an ith threat of the jth Intrusion Detection System for a particular type of the threat?

The calculation of the combined probability mass functions will be followed as:-

$$P(\{T_1\}) = \frac{Detected \quad Alerts}{Observed \quad Alerts} = \frac{1}{20} = 0.05 \quad (13)$$

$P(\{T_1\})$ is the probability assigned to the 1st threat by 1st Intrusion detection system.

$$P(\{T_2\}) = \frac{Detected \quad Alerts}{Observed \quad Alerts} = \frac{30}{43} = 0.697674419 \quad (14)$$

$P(\{T_2\})$ is the probability assigned to the 2nd threat by 2nd Intrusion detection system.

$$P(\{T_3\}) = \frac{Detected \quad Alerts}{Observed \quad Alerts} = \frac{9}{95} = 0.094736842 \quad (15)$$

$P(\{T_3\})$ is the probability assigned to the 3rd threat by 3rd Intrusion detection system.

$$P(\{T_4\}) = \frac{Detected \quad Alerts}{Observed \quad Alerts} = \frac{73}{103} = 0.708737864 \quad (16)$$

$P(\{T_4\})$ is the probability assigned to the 4th threat by 4th Intrusion detection system.

Putting the above values in the combined probability mass formulas!

$$M_{1,2}(T_{1,2}) = \frac{P(\{T_1\})P(\{T_2\})}{P(\{T_1\})P(\{T_2\})P(\{T_3\}) + P(\{\neg T_1\})P(\{\neg T_2\})} \quad (17)$$

$M_{1,2}(T_{1,2})$ is the combined probability mass of the intrusion detection system 1 and 2 assigned to threat 1 and 2.

$$M_{1,2}(T_{1,2}) = 0.108303249 \quad (18)$$

$$M_{1,2,3}(T_{1,2,3}) = \frac{P(\{T_1\})P(\{T_2\})P(\{T_3\})}{P(\{T_1\})P(\{T_2\})P(\{T_3\}) + P(\{\neg T_1\})P(\{\neg T_2\})P(\{\neg T_3\})} \quad (19)$$

$M_{1,2,3}(T_{1,2,3})$ is the combined probability mass of the intrusion detection system 1, 2 and 3 assigned to threat 1, 2 and 3.

$$M_{1,2,3}(T_{1,2,3}) = 0.012551134 \quad (20)$$

Similarly the probability mass of the 4 intrusion detection system would be:-

$$M_{1,2,3,4}(T_{1,2,3,4}) = \frac{P(\{T_1\})P(\{T_2\})P(\{T_3\})P(\{T_4\})}{P(\{T_1\})P(\{T_2\})P(\{T_3\})P(\{T_4\}) + P(\{\neg T_1\})P(\{\neg T_2\})P(\{\neg T_3\})P(\{\neg T_4\})} \quad (21)$$

$M_{1,2,3,4}(T_{1,2,3,4})$ is the combined probability mass of the intrusion detection system 1, 2, 3 and 4 assigned to threat 1, 2, 3 and 4.

$$M_{1,2,3,4}(T_{1,2,3,4}) = 0.03000137 \quad (22)$$

In this experiment only 4 x threats and 4 x intrusion detection systems are participating in data gathering, therefore this combined probability mass formula for two, three and four threats will be calculated.

5.5 Data Fusion using Weights of the intrusion detection systems

Bayesian decision theory cannot differentiate between uncertainty and ignorance, plus it needs to assign evidence to a hypothesis. The Dempster Shafer theory of inference that is an extension of the Bayesian decision theory overcomes this issue and presents mathematical approach that can assign evidence to a single or group of propositions in an experiment and can combine probability masses of the propositions emerging from more than two sources but its self-evident definition of evidence (probability mass) is not very accurate. The Dempster Shafer theory of inference also has some issues in renormalization of the probability mass during probability masses combinations.

Thus it has become one of the most challenging tasks to find out the ways to perfect the evidential or probability mass combination techniques to increase the accuracy of the statistical decisions.

In my research, I used two different ways to improve decision making.

1. Weights of the observations
2. Generalized Evidence Processing (not done yet)

These methods minimised the effect of probability assignments to the propositions and renormalization of the rule of combinations of the probability masses of the preposition(s).

The assumption I made, in the above data fusion model is that all Intrusion Detection Systems have same weights or degree of accuracy in detecting a type of threat. This is not valid in this particular case because those 4 intrusion detection systems are different products and obviously have different level of accuracy in threat detections. Thus each of these Intrusion Detection Systems in detecting the same type of a threat may provide different level of precision. If an Intrusion Detection System is better than others in determining a particular type of threat(s) so it will be unfair to give the same weights to all the Intrusion Detection Systems in this particular case.

Therefore, I need to measure the weight of each Intrusion Detection Systems that determines its level of precision and reliability for particular threat detection. There are many methods to find out the weights. I used Maximum Entropy method to calculate the weight of the Intrusion Detection Systems in threat detection. (Graham Wallies derivation)

Weights of the four Intrusion Detection Systems are calculated by using Max Entropy (MaxEnt)

As each computer node has different intrusion detection system, it is quite obvious that the reliability of each intrusion detection system is different. The Dempster-Shafer theory is considered to be an excellent mathematical model to measure uncertainty in threat detection. Dempster-Shafer theory and Extended Dempster-Shafer models provide numerical methods for multiple threat analyses of the data collected from different intrusion detection systems whilst each intrusion detection system have different reliability.

The Probability formula for calculating the probability mass and weights of an Intrusion Detection System for a particular threat is:-

$$\sum_{i=0}^{n} Mi(Ti)^{W_i^n} = \sum_{i=0}^{n} \frac{P(\{Ti\})^{W_i^n}}{P(\{Ti\})^{W_i^n} + P(\{\neg Ti\})^{W_i^n}} \quad (23)$$

Where T is the threat and W is the weight of the intrusion detection system and P is the probability of the ith threat of jth Intrusion Detection System.

And $$P(\{Ti\})^{W_i^n} = 1 - P(\{\neg Ti\})^{W_i^n} \quad (24)$$

$P(\{Ti\})^{W_i^n}$ is the probability assigned to the threat by Intrusion detection system with weight.

The Probability formula for calculating the weights of an Intrusion Detection System for a particular threat:-

$$W_i^n = -\sum_{i=1}^{n} P_i \log P \quad (25)$$

Where W is the weight of the Intrusion Detection Systems (sensors) and P is the probability of an ith threat of jth Intrusion Detection Systems.

Calculations of the Extended Dempster Shafer will be as followed:-

$$P(\{T_1\})^{W_1^n} = \frac{Detected\ \ Alerts}{Observed\ \ Alerts} = \frac{9}{23} = 0.391304348 \quad (26)$$

$P(\{T_1\})^{W_1^n}$ is the weighted probability assigned to the 1st threat by 1st Intrusion detection system.

$$P(\{T_2\})^{W_2^n} = \frac{Detected\ \ Alerts}{Observed\ \ Alerts} = \frac{38}{46} = 0.826086957 \quad (27)$$

$P(\{T_2\})^{W_2^n}$ is the weighted probability assigned to the 2nd threat by 2nd Intrusion detection system.

$$P(\{T_3\})^{W_3^n} = \frac{Detected\ \ Alerts}{Observed\ \ Alerts} = \frac{16}{98} = 0.163265306 \quad (28)$$

$P(\{T_3\})^{W_3^n}$ is the weighted probability assigned to the 3rd threat by 3rd Intrusion detection system.

$$P(\{T_4\})^{W_4^n} = \frac{Detected\ \ Alerts}{Observed\ \ Alerts} = \frac{61}{69} = 0.884057971 \quad (29)$$

$P(\{T_4\})^{W_4^n}$ is the weighted probability assigned to the 4th threat by 4th Intrusion detection system.

The weights of the intrusion detection systems:-

$$W_1^n = -\sum_{i=1}^{n} P1 \log P1 = 0.15945078 \quad (30)$$

W_1^n is the weight of the 1st intrusion detection system

$$W_2^n = -\sum_{i=1}^{n} P2 \log P2 = 0.068543933 \quad (31)$$

W_2^n is the weight of the 2nd intrusion detection system

$$W_3^n = -\sum_{i=1}^{n} P3 \log P3 = 0.128507117 \quad (32)$$

W_3^n is the weight of the 3rd intrusion detection system

$$W_4^n = -\sum_{i=1}^{n} P4 \log P4 = 0.047314125 \quad (33)$$

W_4^n is the weight of the 4th intrusion detection system

$$M_{1,2}(T_{1,2})^{W_i^n} = \frac{P(\{T_1\})^{W_1^n} P(\{T_2\})^{W_2^n}}{P(\{T_1\})^{W_1^n} P(\{T_2\})^{W_2^n} + P(\{\neg T_1\})^{W_1^n} P(\{\neg T_2\})^{W_2^n}} \quad (34)$$

$M_{1,2}(T_{1,2})^{W_i^n}$ is the weighted combined probability mass of the probability assigned to 1st and 2nd threat by 1st and 2nd intrusion detection system.

$$= 0.753303965 \quad (35)$$

Similarly the other weighted combined probability masses of the other intrusion detection systems would be:-

$$M_{1,2,3}(T_{1,2,3})^{W_i^n} = 0.373362445 \quad (36)$$

$$M_{1,2,3,4}(T_{1,2,3,4})^{W_i^n} = 0.819596134 \quad (37)$$

5.6 Threat Results based on Dempster Shafer Theory of Inference

After Set Cover data fusion we had a total of 429 threats (Table1). This threat data is now further processed by the next part of the multiple simultaneous threat detection system that is Dempster Shafer as shown in Fig 1. In order to increase the precision of each threat was passed through multiple hypotheses testing as proposed in sec 5.3. The intrusion detection system classified the Dempster Shafer inferences into 4 types. Observed threats, Observed Alerts, Detected Alerts and Real Alerts that helped in determining

the real threat detection and false positive rates. The final results of this part of the fusion have been given in Table 2.

False Positive rates are determined using the formula:-

$$\text{False Positive Rates} = 1 - \frac{Real\ \ Alerts}{Observerd\ \ Alerts} * 100 \quad (38)$$

And Threat Detection rate is calculated using the equation:-

$$\text{Threat Detection Rate} = \frac{Detected\ \ Alerts}{Observerd\ \ Threats} * 100 \quad (39)$$

Threat observations by the multiple simultaneous threat detection system using Dempster Shafer					
IDS		OT	OA	DA	RA
Wireshark	22	20	3	1	1
Sniffers	83	43	9	30	1
Snoop	122	95	9	9	9
MARS	200	103	21	73	3
Total		260	42	113	14
where OT stands for Observed threats, OA for Observed Alerts, DA for Detected Attacks and RA for Real Alerts					

Table 2 Threat Results based on Dempster Shafer Theory of Inference

5.7 Threat Results based on Extended Dempster Shafer Theory of Inference

Just like the Dempster Shafer inference, 429 threats data analysed by the Extended Dempster Shafer inference and Intrusion Detection System then grouped as given in Table 3.

It is obvious that real alerts have gone up from 14 to 33 that is a significant indication that False Positive Rates have reduced as compared to Dempster Shafer. Likewise there is an obvious improvement in threat detection rate as well.

Threat observations by the multiple simultaneous threat detection system using Extended Dempster Shafer				
IDS	OT	OA	DA	RA
Wireshark	23	12	9	12
Sniffers	46	3	38	3
Snoop	98	12	16	12
MARS	69	9	61	6
Total	236	36	124	33
where OT stands for Observed threats,OA for Observed Alerts DA for Detected Attacks and RA for Real Alerts				

Table 3 Threat Results based on Extended Dempster Shafer Theory of Inference

5.8 Performance of the multiple simultaneous threat detection system

The multiple simultaneous threat detection system is a multisensor data fusion system. Its major components statistical and mathematical set covers, Dempster Shafer and extension Dempster Shafer are the main data processing cores and heart of the data processing unit for the system. The larger the number of sensors the greater should be the accuracy and precision in the results. Although Bayesian and Dempster Shafer provide best processing model in multisensor data fusion but involve too much complex iteration of the data fusion process in terms of its probability mass and weight calculations. Therefore, in real life, it would be a very hard task to use Bayesian and DS model for combining probability masses of an experiment having a more than four sensors, particularly in case of overlapping and conflicting propositions. The greater the number of sensors, larger would be precision in threat detection, that's why I'm looking into possibility of using more than 4 sensors in my next step and will use Generalized Evidence Processing (GEP) theory .

In my experiment, I performed experiment in three steps using evidences of 2nd, 3rd and then 4th sensors (intrusion detection systems) to the 4 type of threats. The sensors were 4 intrusion detection systems. I compared their results and have proved the obvious fact that the combined results of the 4 sensors have improved threat detection significantly.

Bayesian and Dempster Shafer theory of inferences provided me tools to combine evidences of these sensors and measure the uncertainty of a hypothesis or to gain better confidence in the combined probability measurements to the evidences or propositions.

The following are the graphs drawn in Microsoft Excel to display the results of multiple simultaneous threat detection system.

Comparing efficiency of the Demspter Shafer and Extended Demspter Shafer data fusion techniques, figure 1 and 2 are showing a significant increase in the combined probability masses in case of Extended Demspter Shafer Theory. That is a good indication of enhanced precision, accuracy and better performance of Extended Demspter Shafer data fusion in threat detection over the Dempster Shafer Data fusion techniques.

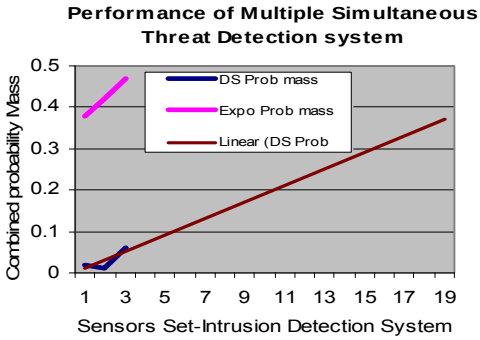

Fig 2 Performance of the multiple simultaneous threat detection system

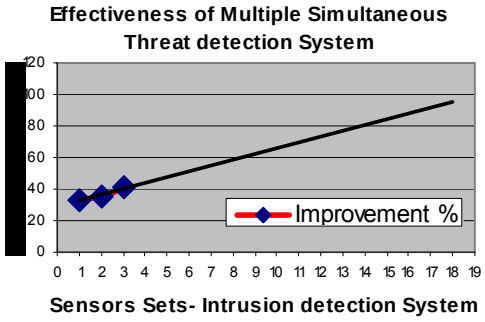

Fig 3 Effectiveness of the multiple simultaneous threat detection System

6. Conclusions

The empirical experiment of multiple simultaneous threat detection system proved that the hybrid model had significant increase in precision in threat detection. Dempster Shafer inference produced 39% detection rate whilst extended Dempster Shafer had 59% detection rate. So on an average, four Intrusion Detection Systems increased 20% detection rate. The false positive rate also went down from 62 % to 11 %. (Detection rate is calculated by dividing detected alerts by observed alerts and false positive rate is derived by dividing real alerts by observed alerts.) Thus there was a net improvement of 51 % in getting rid of false positive alarms and that is highly significant achievement.

Another achievement of the multiple simultaneous threat detection system was its better performance to join probability masses from 4 different Intrusion Detection. The combined probability mass of the Dempster Shafer was 0.06 whilst Extended Dempster Shafer had combined probability mass 0.43, so there was 36 % increase in determining the combined probability masses.

The Multiple simultaneous threat detection system also proved that increase in number of sensors continuously improving the threat detection. However there are calculation complexities involved in determining joined probability masses using Bayesian and Dempster Shafer that made impossible to have more than 4 sensors (intrusion Detection systems).

Set Cover as a middle tier data fusion tool produced incredible results, particularly in data grouping that amazingly minimise the computational processing cpu and memory overhead cost and time. Set Cover reduce data population (from 2274 to 429) to the level that it became possible to detect more than 2 simultaneous threats with less computational efforts whist that was almost impossible with the existing threat detection approaches and others that used Bayesian and Dempster Shafer. Set Cover also determined the cost effectiveness of choosing a computer node for the multiple simultaneous threat detection system. Thus the Set cover played a vital role to assist Multiple simultaneous threat detection system to improve its ability to increase precision of threat detection results.

Looking into the results, it is obvious that results of experiment has proven that my proposed threat detection system “multiple simultaneous threat detection system” remained successful to achieve my research goals.

In order to improve precision of threat detection, as a next step of my research, the main task I’m planning is to implement Generalised Evidential Processing (GEP). GEP is an extension of the Bayesian and Dempster Shafer theory that presents a better evidential combination and separate propositions and the decisions. Therefore each proposition or set of propositions can be tested and analysed separately at different levels of the data. In addition to that I’ll focus to improve the quality of the test experiment and write the final thesis.

8. Acknowledgments

I would like to thank Dr. Paul Kwan (my principal supervisor) for his suggestion on the initial research direction and his valuable comments on how to structure this paper.

I also thank to Thomas Kerin, IBM, Capacity Planner, Melbourne, Victoria, Australia for his suggestions.

9. References

[1] A. Bendjebbour, Y. Delignon, et al., “Multisensor Image Segmentation Using Dempster-Shafer Fusion in Markov Fields Context”, IEEE Transaction on GeoScience and Remote Sensing, Volume 39 Issue 8, August 2001; pp. 1-10

[2] A. Habib, M. Hefeeda, and B. Bhargava. Detecting service violations and DoS attacks. In NDSS Conference Proceedings. Internet Society, 2003; pp. 439-446

[3] Aickelin U (2002): 'An Indirect Genetic Algorithm for Set Covering Problems', Journal of the Operational Research Society, 53(10), pp. 1118-1126.
[4] Ben Grocholsky, Alexei Makarenko, Hugh F. Durrant-Whyte: Information-theoretic coordinated control of multiple sensor platforms. ICRA 2003: pp. 1521-1526
[5] Braun, J. (2000) Dempster-Shafer theory and Bayesian reasoning in multisensor data fusion, Sensor Fusion: Architectures, Algorithms and Applications IV; Proceedings of SPIE 4051, pp. 255–266
[6] COMPUTER SECURITY INSTITUTE. Cyber crime bleeds U.S. corporations, survey shows, Apr. 2002. accessed on.. http://www.gocsi.com/press/20020407.html Accessed 16 January 2003.
[7] D. Hall. Mathematical Techniques in Multisensor Data Fusion. Artech House, Norwood, Massachusetts, 1992;pp. 99-105
[8] Diego Zamboni. "Doing intrusion detection using embedded sensors" CERIAS Technical report 2000-21, CERIAS, Purdue University, West Lafayette, IN, Oct, 2000;pp. 1-9
[9] Don Koks and Subhash Challa, 2005, An Introduction to Bayesian and Dempster-Shefer Data Fusion; pp. 1-52
[10] Dong and Deborah (ACM 2005) Alert Confidence Fusion in Intrusion Detection Systems with Extended Dempster-Shafer Theory; pp. 142-147
[11] H. Wu, M. Siegel, R. Stiefelhagen, and J. Yang. Sensor fusion using Dempster-Shafer theory. In Proceedings of IEEE Instrumentation and Measurement Technology Conference, Anchorage, AK, USA, 2002;1-6
[12] Hugh F. Durrant-Whyte: Data fusion in sensor networks. IPSN 2005, pp. 545-565
[13] J. Burroughs, L. F. Wilson and George V. Analysis of Distributed Intrusion Detection Systems Using Bayesian Methods, presented at IPCCC 2002, April 2002; pp. 142-147
[14] J. R. Boston, "A Signal Detection System Based on Dempster-Shafer Theory and Comparison to Fuzzy Detection", IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, Volume 30, Issue 1, February 2000;pp. 45-51
[15] Kapil Kumar S, 2000 Intrusion Detection and Analysis, University of British Columbia
[16] Lawrence A. Klein, "Sensor and Data Fusion Concepts and Applications" (second edition), SPIE Optical Engineering Press, 1999, ISBN 0-8194-3231-8; pp. 1-252
[17] Ma, Bing 2001 "Parametric and Non Parametric Approaches for Multisensor Data Fusion" PhD thesis, University Of Michigan;1-212
[18] Ning P, Xu D, Healey C and Amant R (2004), 'Building Attack Scenarios through Integration of Complementary Alert Correlation Methods', 11th Annual Network and Distributed System Security Symposium, pp. 97-111
[19] Rehman R, (2003) 'Intrusion Detection System with SNORT', accessed on.. http://www.snort.org/ ; pp. 1-288
[20] S Terry Brugger, 2004 Data Mining for Network Intrusion Detection – PP.8/55 accessed on.. www.bruggerink.com/~zow/papers/dmnid_qualpres.pdf
[21] Siaterlis C and Maglaris B (2004), 'Towards Multisensor Data Fusion for DoS detection', Proceedings of the 2004 ACM symposium on Applied Computing; pp.1-8
[22] SPAFFORD, E. H. The Internet worm incident. Tech. Rep. Purdue Technical Report CSD-TR-933, Department of Computer Science, Purdue University, West Lafayette, IN 47907-2004, 1991. LEM OS, R. Counting the cost of slammer, Jan. 2003; pp. 1-19
[23] Tim Bass and Dave Gruber. a glimpse into the future of id. Usenix. 18 Aug 2005. accessed on.. http://www.usenix.org/publications/login/1999-9/features/future.html.
[24] V. Chatzigiannakis, A. Lenis, C. Siaterlis, M. Grammatikou, D. Kalogeras, S. Papavassiliou & V. Maglaris 2002, Distributed Network Monitoring and anomaly Detection as a Grid Application ;pp. 1-13
[25] V.Gorodetski, O.Karsaev, I.Kotenko, and A.Khabalov. "Software Development Kit for Multi-agent Systems Design and Implementation". In B.Dunin-Keplicz and E.Nawareski (Eds.) "From Theory to Practice in Multi- agent Systems". Lecture Notes In Artificial Intelligence, vol. 2296, 121-130, Springer Verlag, 2002;pp. 121-130