Journal of mathematics and computer Science 8 (2014) 398 - 410

Contents list available at JMCS

Journal of Mathematics and Computer Science

Journal Homepage: www.tjmcs.com

Multi-Level Fusion to Improve Threat Pattern Recognition in Cyber Defense

AliJabar Rashidi
Malek-e-Ashtar University of Technology, Tehran, Iran
Department of computer engineering
Aiorashid@yahoo.com

Kourosh Dadashtabar Ahmadi
Malek-e-Ashtar University of Technology, Tehran, Iran
Department of computer engineering
Dadashtabar@yahoo.com

Ali Jafari
Malek-e-Ashtar University of Technology, Tehran, Iran
Department of computer engineering
IustUser@yahoo.com

Article history:
Received February 2013
Accepted May 2013
Available online August 2013

Abstract

Considering fast growth of internet and related network infrastructures, it is important to detect the intrusion and respond to it in a timely manner. Network intrusion can make vital information systems and communication networks inaccessible and imposes high cost of communication infrastructures. In order to gain high degrees of success in providing services, current and future generation of networking and internet technologies, require a set of tools to analyze the network and to detect the threats and intrusion in network. Due to main weakness in terms of high rate of false alarms and low accuracy of detection, by which cyber space detection and identification systems are opposed, fusion theory in decision level provides a new method for data analysis from multiple nodes in order to increase the possibility of intrusion detection through improving pattern recognition. This paper aims to present a novel method of fusion in decision level based on complex event processing and show how this method would be successful in exposing cyber threats for timely response.

Keywords: Information fusion; complex event processing; cyber defense; pattern recognition

1. Introduction

Currently, due to fast development of cyber threats make against security and defense infrastructures, communication systems and financial markets need to design, produce, domesticate and optimize a new generation of weapons and equipment of cyber defense [10]. Moreover, since current techniques and technologies of attack have become more complex and more consistent, common detection and identification methods are not only unsuitable and not providing necessary security for critical events in dangerous situations but also they can pose a significant risk to important and critical infrastructures such as electricity power transfer network, integrated banking network, work and transportation distribution network [17]. To proactive protection of important, critical and vital infrastructures, application of advanced intrusion detection systems, suitable response systems, creation of event driven infrastructures and application of complex processing solutions with intelligent detection and flexible monitoring of the events are considered as critical problems. Moreover in these systems, application of stored data and fusion of real time acquired data (from cyber space) will end up to improvement of cyber threats recognition [20].

It is too difficult to find the models and techniques in business domain which can recognize the attack with proper and timely information at real place and at short opportunity before the attack. In respect to this, exposing and recognition of events derived from cyber threats based on data fusion are necessary for cyber defense on threshold of attack. Four pivots which provide the foundation for predictive business are: description (what happened?), orientation (focusing on key problems), prediction (future prediction for appropriate action) and consistent action [25];

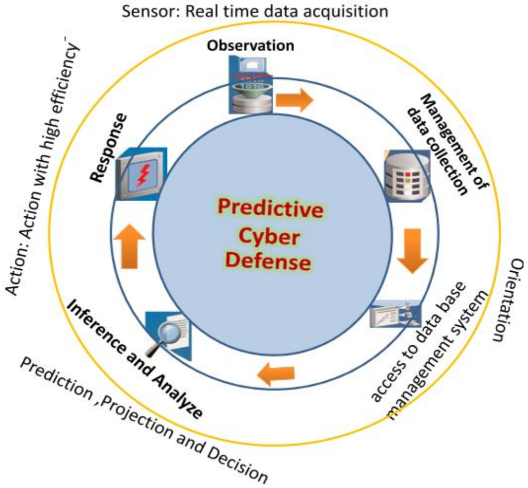
Sensor: Real time data acquisition
Observation
Management of data collection
Orientation
access to data base management system
Prediction ,Projection and Decision
Inference and Analyze
Action: Action with high efficiency
Response
Predictive Cyber Defense

Figure 1.Five Main Phases of Predictive Businesses

As it is shown in Figure 1, four above pivots consist of five main phases in data acquisition (recording, filtering a wide range of data gathered from resources such as sensors or other readers) management (management of large volume of distributed data and events which should be scalable, secure and reliable), access (access to data anytime, from anywhere), analysis (real-time data and event analysis) and response (automatic response to events) [26].

Predictive businesses for detection and recognition of threats pattern, besides application of complex event processing solutions, makes use of management of real-time events derived from distributed sensors and multi sensor data fusion. Therefore, in intelligent defense, it is possible to [25]:

- To detect and identify the threats which are as attack vector or the patterns which have the potential of a percussion attack.
- To forecast the threat patterns and to analyze the network vulnerabilities which are very attractive to attackers [20].

- To apply an in phases–trap system (such as pitfalls /honey pots) to draw attraction and antecedence of attacker in order to increase informational depth for acquiring an understandable behavioral pattern [12].
- To gather the events occurred in large scale in form of a threat event cloud in order to identify the profiles of internal and external multi vector threats which are the most common threats in vital networks [26].

2. Complex event processing in detection and identification systems

Cyber-attacks are currently more complex than anytime and events caused by these types of attacks considering variety and improvement of technologies are changing continuously. Such complexities end up with event cloud. Due to these event clouds, event flows in cyber systems are not transparent and are too difficult to understand, requiring complex event processing [21].

Complex event processing for information extraction, detection and response to abnormalities, threats and distributed businesses opportunities utilizes filtering with very low delay time (particle filtering), correlation(establishment of logical relationship between events), aggregation and computation on event data . Complex event processing can be known as a network based technology which has the operational capability of situational awareness extraction from distributed message based systems, databases and applications in real -time or semi real -time. Complex event processing technology for cyber defense on threshold of attack processes and analyses a large volume of multiple sequence of attacks, high speed attacks and events caused by these attacks and detects the possible opportunities and threats in event occurrence time or near it [21]. The important question about using of complex event processing for intrusion detection systems in cyber space is that what kind of problem would be solved with this kind of processing? The answer to this question can be found in figure 4 which refers to cyber threats detection and identification in shortest time and quick response to the event caused by that [17,23]

2.1. Complex event processing based on data fusion

Defense information operation and intrusion exposing systems had been designed first in 1980 for to support accessibility, confidentiality and integration of vital information infrastructures. These operations were protecting information infrastructures against denial of service, disclosure of confidential information and data modification or extermination attacks. In spite of significant development in the field of identification and detection, yet the majority of security experts believe that current real-time intrusion detection systems are not enough mature from technical aspects for exposure of cyber-attacks. For example, during Longley cyber-attacks, this exposure system didn’t show any action for exposure of massive volume of email bombs which disabled vital email servers [20, 22].

For automatic exposure and instant reporting of events caused by cyber threats, a suitable system is required against computerized and networked attacks. This basic system is equipped with state based exposure, statistical anomaly exposure, traffic analysis, threat behavior pattern and identified pattern template; in order to detect abnormal behaviors in an environment.

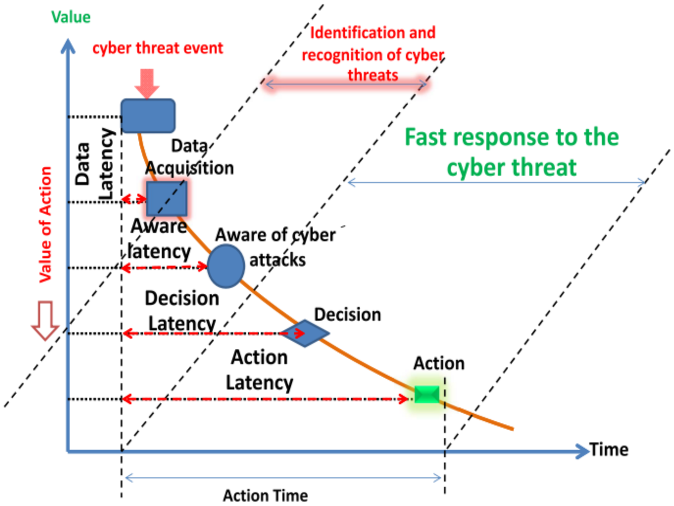

Figure 2.The Value of Timely Recognition and Identification

Such systems permanently are searching and finding abnormal behaviors caused by cyber-attacks and act based on behavior changes during the time and comparison of it with previous behaviors in the system. Currently an important challenge that yet exists in intrusion exposure systems, is data and information aggregation yield from distributed heterogeneous agents in form of inter related processes which is used for cyber space security assessment. Multi sensor/agent data fusion provides an important performance framework to create situational awareness of cyber space and next generation of intrusion detection and identification. Multi sensor data fusion technology is a path which is used for development of intrusion detection with high reliability to identify, detect and evaluate the cyber space situation which is affected by many complex threats. Multi sensor data fusion or distributed observations fuse data from multiple sensors and resources to enable the identification of events, activities and situations of a cyber space with high reliability. This topic can be compared with human cognitive process where human brain mixes (fuses) the sensor information from different body organs to evaluate the situations, make appropriate decisions and run the controlled actions[1,2].

The fusion is done at three levels of signal fusion (low level), attribute/feature fusion (medium level) and decision level (high level). The operations which are done at signal fusion level mostly are related to data collection from sensors, sensor selection, data template transformation, data aggregation and finally data fusion. At attribute fusion level, different features of a goal in here, cyber threats, are extracted and then the main feature of the goal would be determined through their fusion [3, 5].

In this case, the identification process includes knowledge based techniques (such as expert and fuzzy logic systems) or learning based methods (such as Bayesian theory and neural networks). At high level fusion, received results from multiple local decision making positions with some extent of uncertainty are fused together and the final decision is created. at decision level fusion, information volume from signal and attribute levels are less and the processing and final decision making are quicker. So, this level is more suitable for distributed goal detection and processing. There are many advantages for information fusion at different levels. Such as information redundancy, information complements, information and cost reduction. Information redundancy is made through applying information from multiple resources (or from one sensor at different time frames). With complementary information and information fusion, a value added system will be created which had not been possible through individual sensor information. Different models of information fusion have been used in related research literature. Some of these models are functional while others are procedural. The hybrid procedural model based on predictive cyber defense model is shown in Figure 3[1, 4].

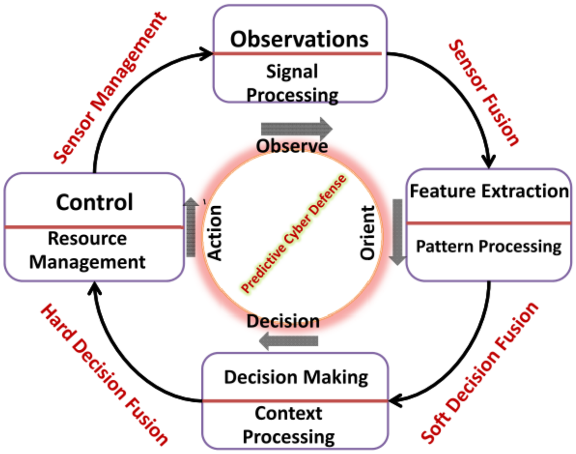

Figure 3.Hybrid Procedural Model of Information Fusion

The inputs of a recognition and identification system equipped with data fusion module are raw data of database, sensor data, comments and events. For example, we can refer to the data resulted from distributed packet nodes id, system files registration, queries and SNMP traps, user profiles database, system messages and operator commands. The outputs of the above system include an estimate of threat (intruder) position and its activities, observed threats, attack rate and cyber-attack intensity evaluation. In military command and control systems, data fusion sensors receive electromagnetic radiations, thermal and acoustic energy, nuclear emission, noise and other signals. Whereas in cyber space threat, identification and identification systems, sensors are of different types because their environmental type and size is different. Since instead of launching rockets to atmosphere, the information, flows in cyber space. While command and control commanders mostly focus on threat origin, speed, kind and the goals of missile cyber defense authorities are interested in identification of attack rate, threat kind and the hostile intruder goals. Therefore, special sensors or agents are required to be used in cyber space [6, 8].

As interpreted here, complex event processing deals with event gathering from multiple resources, filtering, format conversion, correlation, aggregation and complex events detection based on predefined rules and patterns. Complex event processing based on fusion at three levels of signal, attribute and decision besides high advantages of more information extraction and cyber defense recognition and identification system improvement, fuses the information from different resources. This will end up with uncertainty reduction in threat pattern identification and increasing understanding of the system. In figure 8, three phases of recognition, identification, representation and real-time processing of complex event in an advanced cyber threat recognition and identification system are shown [9, 11].

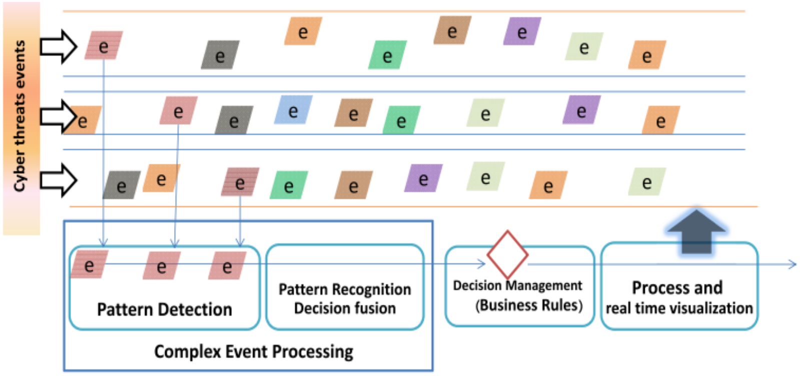

Figure 4.Cyber Threat Pattern Detection and Recognition from Different Resources based on Fusion

2.2. Fusion strategy for use in complex event processing based on data fusion

In recent years, multi-level information fusion drew the attention of many researches around the world. Since through multi-level information fusion, more complete information will be gained which in return will cause with increased accuracy of decision making process. There are different methods for multi-level fusion such as feature level fusion, decision level fusion and hybrid fusion [1].

At feature level, extracted attributes from input data first are fused and then are sent as an input to signal analysis unit in order to be analyzed. Since these attributes refer to some of distinguishable features of a stream or object, attribute fusion unit fuses the multi aspect attributes resulted from an object as an attribute vector which is received from detection unit [7].

Figure (5) shows the fusion unit which takes a set of features or decisions and after analyzing them, makes a decision at semantic level. Figure (5) shows the feature fusion unit which receives a set of attributes from F_1 to F_n and fuses them as a feature vector$\mathrm{F}_{1,\mathrm{n}}$. Figure (5) represents a view of a multi aspect fusion which first fuses the extracted attributes with feature fusion unit and then passes the fused feature vector on to the analysis unit. The advantage of feature level fusion is correlation of multiple features from different aspects which serve to improve the performance although it only needs one phase of learning in fused feature vector. Anyway, the synchronization of multi aspect features seems to be the difficult part, because the features from different aspects and different times must be extracted and fused. Moreover the features should be converted into the same format before the fusion. Finally, increasing the number of aspects for learning of diagonal correlation between heterogeneous attributes would be considered as another problem of this method [1, 7].

At decision level, fusion unit first creates the local decisions $\mathrm{D}_1$to$\mathrm{D}_\mathrm{n}$base on features F_1toF_n. These local decisions are then fused as decision vector by decision analysis unit, composed to make the final decision. Finally the output of a semantic level analysis unit will be a decision. A view of decision fusion unit is shown in figure (5). Also, figure (5) represents decision level multi aspect analysis from decisions gained from different fusion unit. Decision level strategy has more advantages compared with feature level. For example, representation format at feature level based on extracted aspects are different while representation format at decision level is often the same. Therefore fusion at this level is easier. Moreover, at decision level, a composition of different methods of fusion can be used which is not possible at feature level. One of important problems of decision level multi aspect fusion is about correlation between different aspects features at feature level because in case of a fault in this phase, there is no feedback for compensation/recovery. In addition, different classifiers which are used in learning process for local decision acquisition are too time- consuming and tedious [1, 8].

In order to make use of feature and decision fusion strategies, multi aspect hybrid fusion has been used by different researchers in different domains. A view of hybrid strategy is shown in Figure (5). In this strategy, first, attributes are fused together by different feature fusion units and then the feature vector will be analyzed by analysis unit. In a similar way, other attributes are analyzed by different analysis units and then their decisions are fused by decision fusion unit. Finally, all decisions resulted by decision fusion unit from previous phases, will be fused and final decision will be extracted.

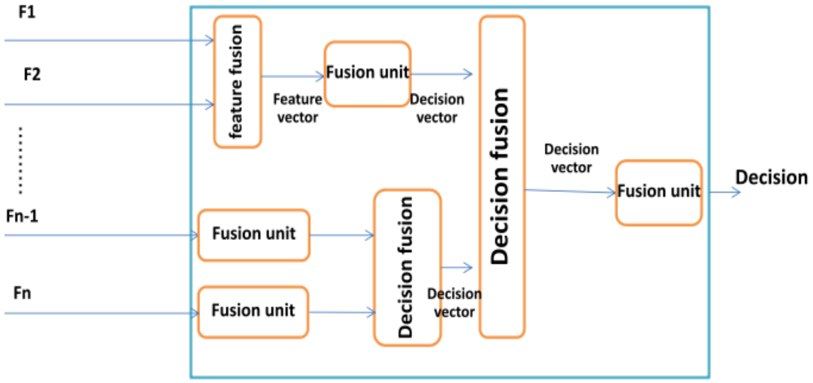

Figure 5.Hybrid Multi-level Fusion

2.3. Evidential fusion for improvement of cyber threat pattern recognition

At lowest level of induction which is data aggregation process, the methods are used that enable unrelated events elimination. This would decrease the number of network event pairs for evaluation. At this level, some metrics are needed to show the similarity and adjacency of the events.

The other activity of this level of fusion is to select the events which clarify the relevance through intrusion theories and theories processing.

At higher level, parametric data are used for network events parameters estimation and estimation theory is used for determination of attacks rate, attack goals, attack origin and situational parameters of cyber space. Detection and estimation processes need a strong processor and advanced mathematics and normally apply optimization, least square approximation and sequential approximation. Intrusion detection systems require complex fault analysis algorithms and statistical models for false and noisy alarms estimation.

Pattern recognition and identification in fusion model is too difficult and complex because its induction level is too high. This phase often is done with extraction of abstract attributes of raw data. For advanced intrusion detector systems, networks which are faced to attacks from different areas require clustering analysis techniques, adaptive neural network and rule-based knowledge systems. So, how pattern recognition systems learn is based on four learning pivots: pattern adaption, hybrid adaption, neural networks and statistical identification which are not necessarily independent. In some cases, a composition of these methods is used [14].

Presumptive theories used in pattern recognition will cause some kind of uncertainty in results of decisions. Currently multi-level fusion as a tool for improvement of pattern recognition and analysis quality can overcome this uncertainty. In this kind of fusion which is the high level of fusion, decision composition is done by multiple fault independent classifiers (with the same identification problem). Also, decision fusion using multiple individual attributes which reflect different features of a threat and are gained by different sensors from different cyber space or by different agents in cyber defense systems, improve the identification accuracy of a single classifier [11, 13].

The most important part of a pattern identification system is assigning an instance to a class and tagging the instances based on extracted model from a group of instances. Considering the importance of this issue, regardless of how the tagging is done, the other problem is the necessity of assigning an instance just to one special class which will cause ambiguity in tagging some of the instances. This problem is more likely when some information related classes overlap in used attribute space. For example, in Bayesian method, an instance is purely assigned to the class resulted from class probabilities. i.e. The instance would be assigned to the class which has the most class possibilities. In cases where some class possibilities are near each other or their probabilities are not recognizable, Bayesian theory cannot be responsive, alone. Moreover, in some cases, a special classification due to learning across a special set of data does not have the ability to make distinction between all information. Therefore application of hybrid methods such as creating a classifier resulted from several classifiers and composition of the results of different attribute classifiers with different methods can solve some of these problems. To compose the results of different attributes classifiers, we can use statistical methods, fuzzy theory and evidence fusion theory which are known as Dempster-Shafer reasoning [1].

In this paper, in order to resolve the restriction of instance assignment to one class in classifier systems, the information is classified in a way that all possible compositions from classes and a suitable criterion are considered. This process is done for every different attribute and individually. Considering hybrid classes, system response to different attributes comes with uncertainty. So, by using belief accumulation theory with the aim of decreasing uncertainty level, different responses are fused with each other. In other word, notwithstanding real classes overlapping, belief fusion theory, removes the created uncertainty.

2.4. Belief accumulation theory in data fusion

Bayesian reasoning is often efficient in classic sensor fusion methods and is used more than other fusion methods. This method more than being efficient at decision level fusion is effective at attribute level fusion. Observations obtained from different aspects or decisions gained from different classifiers are fused and the common probability of an observation or decision is extracted. In Bayesian method, the probability of a belief is measured. Normally, because it is difficult or impossible to compute the possibility of belief A, it is computed considering another event B. Computation of this conditional probability is easier and is called as main Bayesian formula [1, 3].

$$P(A|B) = \frac{P(B|A\ ) \times P(A)}{P(B)} \tag{1}$$

If feature or decision vector is from multiple aspects as:

$$(I_1, I_2, \dots, I_n) \tag{2}$$

And given that multi-aspect observations are statistically independent from each other, common probability of a theory H based on fused feature vectors or decisions can be computed as:

$$P\left({}^{H}/_{I_1, I_2, \dots, I_n}\right) = \frac{1}{N}\prod_{k=1}^{n} P\left({}^{I_k}/_{H}\right)^{w_k} \tag{3}$$

In above formula, N is used for normalization of post probability estimation $\mathrm{P}\left({}^{\mathrm{H}}/_{\mathrm{I_1, I_2, \dots, I_n}}\right)$. $\mathrm{w_j}$ is considered as the weight of k th aspect and $\sum_{\mathrm{k=1}}^{\mathrm{n}} \mathrm{w_j} = 1$. This post probability is computed for all possible theories. Bayesian theory is also used for join possibility reasoning of input classifiers. Assuming that:

$$\Omega = \{\mathrm{w_1, w_2, \dots., w_M}\} \tag{4}$$

Is the set of classes of data. In this case, data X belongs to class $\mathrm{w_i}$ if:

$$X \in w_i \;\; if \;\; P\left({}^{w_i}/_{X}\right) > P\left({}^{w_j}/_{X}\right) \quad \forall_{j\neq i} \tag{5}$$

One of the problems of this method is that post probabilities of $\mathrm{P}\left({}^{\mathrm{w_i}}/_{\mathrm{X}}\right)$ is not easily comparable.

Dempster-shafer theory is a tool for displaying and fusion of heuristic measurements. This theory is extended theory of Bayesian reasoning which was published in 1967 by Dempster and his student Shafer. When the knowledge regarding the goal is not complete and we face to uncertainty and ignorance in information, the mentioned theory is more flexible that Bayesian theory and is more efficient in checking and mixing of evidences.

In belief structure of Dempster-Shafer, the assumption is that there is a question and θ is a set of answers and suggestions related to that question. θ elements include all possible answers and are mutually exclusive(every element include unique information). For example, "pattern x belongs to class $\mathrm{w_i}$ ". In this structure, set of distinct (without overlapping) and complete (including all possible events) suggestions is called observation framework and each of the elements of this set is called a simple suggestion.

The belief structure of Dempster-Shafer is composed of a number of non-empty subsets of θ called $\mathrm{A_i}$ which are called focal elements and a set of weight functions $\mathrm{m(A_i)}$ (instead of assigning possibilities to suggestions) are used as evidence values which must have these conditions:

$$\theta = \{A_1, A_2 \quad \ldots. \quad A_k\} \tag{6}$$

θ is a set of all possible assumptions for a K classes classifier in an observation framework in which

$$\mathrm{A_i = x \in w_i} \tag{7}$$

$$m: P(\theta) \rightarrow [0{,}1] \qquad m(A_i) \in [0, \quad 1] \tag{8}$$

$$m(\emptyset) = 0 \;\; m(A_i) \neq 0 \tag{9}$$

$$\sum m(A_i) = 1 \tag{10}$$

Function $\mathrm{m(A_i)}$ is called as Mass Probability Assignment Function. This function is a mapping from exponential set θ in 0 to 1 interval (i.e. a value between 0 and 1 is assigned to every member of θ). The value of this function m for empty set is zero and for the set including all assumptions is one. Belief function is defined in relation with basic probability assignment. And assigns a value between zero and one to every non empty subset $\mathrm{B}$of reference set θ. This function represents the degree of belief in B and is defined by following relation:

$$Bel(B) = \sum m(A_i) \tag{11}$$

Actually, $\mathrm{Bel(B)}$ consists of our belief in that the answer of the question is a somewhere in$\mathrm{B}$.

While probability theory assign a probability value to an atomic assumption $\mathrm{\theta_i}$(for example, this pattern belongs to class n), Dempster –Shafer theory assigns Measures of Support or evidence that we tend to assign to a composite assumption which is a composition of several atomic one. (For example this pattern belongs to class n or class K). This evidence measure shows our uncertainty resulted from inability in more division of evidence among atomic assumptions constitute a composite assumption. If for an observation framework $\mathrm{\theta}$, for all atomic assumptions

$\theta_n$we consider $m(\theta_n) \neq 0$ and for all composite assumptions A, we consider $m(A) = 0$, then we meet a situation that probability theory is $\sum_n m(\theta_n) = 1$ and $m(\theta_n)$ which can be considered as probability θ_i.

Function $\mathrm{Bel}: 2^\theta \rightarrow [0,1]$ is derived from Mass Probability Assignment Function while a basic probability assignment functions similar to probability density function. Function$\mathrm{Bel}(A) = \sum_{S\subseteq A} m(S)$ is a measure similar to probability density function in probability theory. If A is an atomic assumption that $\mathrm{Bel}(A) = m(A)$.

Since our knowledge about assumption A, is often incomplete and there is some uncertainty, possibility of occurrence of assumption A is more than $\mathrm{Bel}(A)$. To detect this uncertainty, Plausibility Function is used. The value of this function, for non-empty set A is defined considering the degree of belief in non-occurrence of it, means:

$$pl(A) = 1 - Bel(\neg A) = \sum m(B) \qquad (12)$$

Where $\neg A$ is the event of non-occurrence of A. Belief function identifies the lower border of occurrence possibility of event A and supporting function of its upper border i.e. $\mathrm{Bel}(A) \leq \mathrm{pr}(A) \leq \mathrm{pl}(A)$

$\mathrm{pl}(A) - \mathrm{Bel}(A)$ Represents unknown knowledge about assumption A. Dempster-Shafer theory provides a method of knowledge fusion which is gained from different resources. Depending on the fusion rule definition of D-S for probability density assignment, if m_1 and m_2 are Mass Probability Assignment which are defined over θ , for a non-empty subset A , their fusion or Orthogonal Sum $m = m_1 \oplus m_2$ which is defined as:

$$m(A) = \frac{\sum_{B\cap D=A} m_1(B).m_2(D)}{1 - \sum_{B\cap D=\emptyset} m_1(B).m_2(D)}\ ,\ m(\emptyset) = 0, A \neq \emptyset \qquad (13)$$

Fusion rule can also be generalized to collaborative multiple evidences because there is a one to one relation between m andBel. Therefore orthogonal sum of functions $\mathrm{Bel} = \mathrm{Bel}_1 \oplus \mathrm{Bel}_2$ is defined clearly. Two especial types of Bel can be so effective in D-S applications. These two kinds are called simple supporting and separable functions. Bel is a simple supporting function if $F \subseteq \theta$ so that $\mathrm{Bel}(\theta) = 1$

$$\mathrm{Bel}(A) = \begin{cases} S & F \subseteq A \text{ and } A \neq \theta \\ 0 & \text{otherwise} \end{cases} \qquad (14)$$

Where s is called the measure of Support of Bel and F is called a focal element. A support separable function is a simple orthogonal sum of support functions. If Bel is a simple support function with focal element $F \neq \theta$, then we will have:

$$m(A) = \begin{cases} s & A = F \\ 1 - s & A = \theta \\ 0 & Otherwise \end{cases} \qquad (15)$$

Let F be the focal point of two simple support functions with measures of support of s_1 ands_2, then if $\mathrm{Bel} = \mathrm{Bel}_1 \oplus \mathrm{Bel}_2$, we will have:

$$m(A) = \begin{cases} 1 - (1 - s_1)(1 - s_2) & A = F \\ (1 - s_1)(1 - s_2) & A = \theta \\ 0 & otherwise \end{cases} \qquad (16)$$

Since the above fusion rule has association and substitution properties, repetition of it can be used for multi-function fusion, i.e.:

$$m = (\dots((m_1 \oplus m_2) \oplus m_3)\dots) \quad (17)$$

In sum, it can be said, using D_S theory in fusion of evidences, the probability of an event caused by cyber threats is determined and this theory can be used as a method of reasoning under the conditions of epistemic uncertainty of cyber environment. The important part of this theory is the D-S fusion rule which fuses the observations from two or more sources of threat emission [1, 3, 15].

3. Evaluation of proposed architecture

Considering all previous information, in order to define the performance of the proposed design in figure (6) in cyber threats detection and recognition, the measures from paper [15] were used to compare the performance assurance according to attack scenario of figure (7) in four aspects of Recall، Precision،Fragmentation and Mis-Association[2,15].

Ground truth for the test scenarios consisted of Snort alerts, Dragon alerts, and IIS/Apache web log alerts grouped based on the attacker and labeled with attack types such as background scanner, client track, and attacker. The x-axis indicates the ‘threshold’ value used for determining the corresponding metric scores. The use of a threshold comes from the need of matching detected attack tracks to the actual occurred attacks in ground truth.

According to [2, 15] the four metrics are defined as follows:

- Recall measures the percentage of tracks detected in Relation to the ‘‘total known’’ tracks (in the ground truth).
- Precision is the percentage of correct tracks
- Detected in relation to the number of detections or proposed results.
- Fragmentation is defined as the percentage of tracks reported as multiple tracks that should have been reported as a single track
- Mis-Associations are those known tracks that were incorrectly identified by the proposed tracks.

The MIT Lincoln Laboratory under DARPA and AFRL sponsorship has collected and distributed the first standard corpora for evaluation of computer network intrusion detection systems (table 1). This DARPA evaluation data set is used for the purpose of training as well as testing intrusion detectors. These evaluations contributed significantly to the intrusion detection research by providing direction for research efforts and an objective calibration of the technical state-of-the-art (table 1). They are of interest to all researchers working on the general problem of workstation and network intrusion detection [2, 9, and 18].

According to figure 6, each module can be described as below:

- Filtering: Select events that match user-specified criteria using comparison expressions
- Correlation: Join events from different feeds based on common attributes and/or expressions [16].
- Aggregation: Compute various statistics from event data over time- and count-based sliding and jumping windows, including Count, Average, Sum, Min, Max [25].
- Event Pattern Matching: Detect sequential flow of state-changes of one or more event streams over time (i.e., generate an alert if Event A occurred, and then Event B occurred within 60 seconds) [22].
- Enrichment: Merge reference data from external DBMS systems into analytics models to provide full business context for underlying events [24].
- Multi-dimensional Analysis: Compute various statistics from the event data broken down by one or more attributes (dimensions) [20].
- Situational Analysis: Overlay contextual information on event data

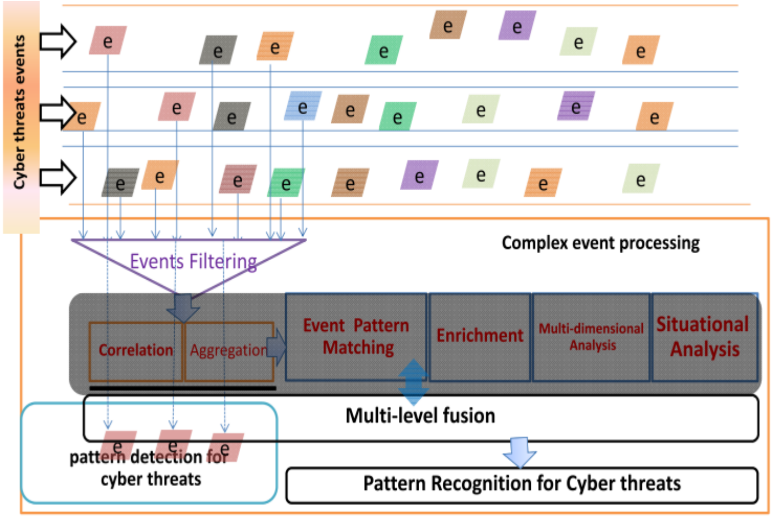

Figure 6.Propose Architecture in Cyber Threat Pattern Detection and Recognition

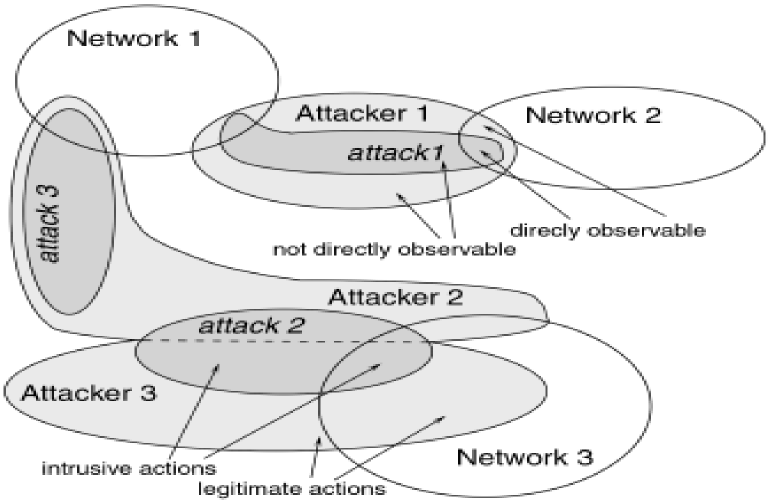

Figure 7.The Determined Attack Scenario for the Design Evaluation

Table 1.Results of Proposed of Detection System

DARPA Data Set	**Intrusion Detection System**		
	IDS Detection Rate(1)%	***IDS Detection Rate(2)%***	***Multilevel Fusion Detection Rate%***
Dataset1(1998)	27.1	56.7	83.2
Dataset2(2000)	61.2	70.8	96.1

4. Conclusion

In this paper, a novel design was presented which can manage the alarms caused by cyber-attacks with improving pattern recognition, while reducing the large number of false alarms which are one of the main problems of intrusion detection system. Moreover by using complex event processing technology, it can correlate the alarms of detection systems to decrease false alarms with higher performance compared to traditional methods while optimizing the detection time of a cyber-attack and reducing its response time compared to other existing architecture.

5. References

[1] M.E. Liggins, D.L. Hall and J. Llinas Handbook of Multisensor Data Fusion: Theory and Practice, CRC Press, 2008.

[2] E. BIasch, I. Kadar, K. Hintz, J. Biermann, C. Chong, and S. Das, Resource Management Coordination with Level 2,3 Fusion Issues and Challenges,IEEE AES Magazine, Vol. 23, No. 3, pp. 32-46, March 2008.

[3] S. Das, J. LIinas, G. Pavlin, D. Snyder, A. Steinberg and K. Sycar a, Agent Based Inform ation Fusion: Panel Discussion, Int. Conf On Info. Fusion, 2007.

[4] J.J. Salern o, M. Sudit, S.I. Yang, G.P. Tad da, I. Kadar and J. Holsopple, Issues and challenges in higher level fusion: Thre at/ impact assessment and intent modeling (a panel summary), Int. Conf. on Info. Fusion, 2010.

[5] E. Blasch, J. Llinas, D. Lambert, P. Valin, S. Das, C-Y. Chong, M.M. Kokar and E. Shahbazian, High Level Information Fusion Developments, Issues, and Grand Challenges - Fusionl 0 Panel Discussion, Int. Conf. On Info. Fusion, 2010.

[6] E. Blasch, J.J. Salern o and G. Tadda, Measuring the Worthiness of Situation Assessment, IEEE Nat. Aero space Electronics Conf, 2011.

[7] D.A. Lambert, Unification of Sensor and Higher-Level Fusion, Int. Conf on Info. Fusion, 2006.

[8] P. Hilletoft h, S. Ujvari and R. Johansson, Agent-Based Simulation Fusion for Improved Decision Making for Service Operations, Int. Conf. On Info. Fusion, 2009.

[9] A. Patcha and J. Park, An overview of anomaly detection techniques: Existing solutions and latest technological trends, Computer Networks Vol.51 pp. 3448–3470, 2007.

[10] Y. Zhang, S. Huang, S. Guo,J. Zhu,Multi-sensor Data Fusion for Cyber Security Situation Awareness, Procedia Environmental Sciences Vol.10 pp.1029 – 1034,2011.

[11] S. Mathew, C. Shah, S. Upadhyaya,An Alert Fusion Framework for Situation Awareness of Coordinated Multistage Attacks, Proceedings of the Third IEEE International Workshop on Information Assurance,2005.

[12] S. Mukkamala, K. Yendrapalli, R. B. Basnet, A. H. Sung,Detecting Coordinated Distributed Multiple Attacks, 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07),2007.

[13] D. Fava ,J. Holsopple , S. J. Yang , B. Argauer,Terrain and Behavior Modeling for Projecting Multistage Cyber Attacks, Int. Conf. On Info. Fusion. 2008.

[14] F. Alserhani, M. Akhlaq,I. U Awan and A. J. Cullen,Detection of Coordinated Attacks Using Alert Correlation Model, IEEE pp.542-546, 2010.

[15] J. Preden, L. Motus, M. Meriste and A. Riid,Situatinal Awareness for Networked Systems, 2011 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), Miami Beach, FL,pp.123-130,2011.

[16] H. Chai and Y. Du,A Framework of Situation Awareness Based on Event Extraction and Correlation for Military Decision Support, Proceedings of 2012 IEEE International Conference on Mechatronics and Automation, August 5 - 8, Chengdu, China,pp.192-196,2012.

[17] M. E. Kuhl and M. Sudit ,J. Kistner and Kevin Costantin,Cyber Attack Modelling And Simulation for Network Security Analysis, Proceedings of the 2007 Winter Simulation Conference,pp.1180-1188,2007.

[18] A. Arasu, M. Cherniack, E. Galvez, D. Maier, A. S. Maskey, E. Ryvkina, M. Stonebraker, and R. Tibbetts. Linear road: a stream data management benchmark. In *VLDB '04: Proceedings of the Thirtieth international conference on Very large data bases*, pages480–491. VLDB Endowment, 2004.

[19] A. P. Barros, G. Decker, and A. Grosskopf. Complex events in business processes. In W. Abramowicz and W. Abramowicz, editors, *BIS*, volume 4439 of *Lecture Notes in Computer Science*, pages 29–40. Springer, 2007.

[20] G. Jiang, H. Chen, and K. Yoshihira. Modeling and tracking of transaction flow dynamics for fault detection in complex systems. *IEEE Transactions on Dependable and Secure Computing*, 3(4):312–326, 2006.

[21] A. Paschke. A homogenous reaction rule language for complex event processing. In In Proc. 2nd International Workshop on Event Drive Architecture and Event Processing Systems EDA-PS, 2007.

[22] K. U. Schmidt, D. Anicic, and R. St¨uhmer. Event-driven Reactivity: A Survey and Requirements Analysis. In *3rd International Workshop on Semantic Business Process Management*, pages 72–86, 2008.

[23] C. Zang and Y. Fan. Complex event processing in enterprise information systems based on rfid. *Enterp. Inf. Syst.*, 1(1):3–23, 2007.

[24] I. H. Witten and E. Frank. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, San Francisco, 2005.

[25] TIBCO *BusinessEvents User's Guide,* TIBCO® BusinessEvents, Software Release1.2, September 2005.

[26] G.Hermosillo, L.Seinturier, Laurence Duchien"Using Complex Event Processing for Dynamic Business Process Adaptation" "Proceedings of the 7th IEEE International Conference on Services Computing, Miami, Florida : United States (2010).