Research Article

A Multisource Situation Information Fusion Method Based on Dynamic Evidence Combination

Jing Liu,^{1,2} ChaoWen Chang,^{1,2} Yuchen Zhang,^{1,2} and Yongwei Wang^{1,2}

¹Information Science and Technology Institute, Zhengzhou 450001, China

²Henan Key Laboratory of Information Security, Zhengzhou 450001, China

Correspondence should be addressed to Yongwei Wang; davysecurity@126.com

Received 17 September 2020; Revised 10 April 2021; Accepted 3 May 2021; Published 29 October 2021

Academic Editor: Haipeng Peng

Copyright © 2021 Jing Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To address the problems of fusion efficiency, detection rate (DR), and false detection rate (FDR) that are associated with existing information fusion methods, a multisource information fusion method featuring dynamic evidence combination based on layer clustering and improved evidence theory is proposed in this study. First, the original alerts are hierarchically clustered and conflicting evidence is eliminated. Then, dynamic evidence combination is applied to fuse the condensed alerts, thereby improving the efficiency and accuracy of the fusion. The experimental results show that the proposed method is superior to current fusion methods in terms of fusion efficiency, DR, and FDR.

1. Introduction

In the field of network security, current research on situational awareness technology mainly focuses on individual alert events. The strategy is to calculate the threat level based on statistical analysis, without deeply examining the correlation between security events [1]. In fact, there are often strong correlations, such as causality and sequential relationships, between security events detected across multiple sensors; thus, statistics-based situational analysis cannot fully reflect the true state of the network [2, 3]. Network security situational awareness technologies that are based on multisource fusion can realize complete and in-depth security situational awareness. However, due to various components in the security system, it is often difficult to obtain a comprehensive overview of the entire network, despite the available data being very rich and diverse [4]. Situation awareness is a process of simplifying and processing a large number of situation basic data to obtain more accurate information. Situation information comes from various sources and has the characteristics of multiclassification, multisource, and high conflict. How to deal with multisource and high conflict data quickly puts forward higher requirements for situation information processing

methods [5, 6]. Situation information fusion is the key to solve this problem. From the perspective of the acquisition of situational awareness data, current methods, based on single-source data, fail to consider the multisource nature and heterogeneity of situational information [7]. As the low-level primary alert event data to be fused are of high volume but low precision, this has a significant negative impact on the final outcome of situational awareness systems [8–10].

The main idea of current situation information fusion research is to transplant classical data fusion methods such as association analysis and evidence theory to the field of network security situation information fusion, without considering the multiclassification, multisource, and high conflict characteristics of situation information itself, which leads to low efficiency and low fusion accuracy in the process of fusion. In this study, a multisource situational information fusion method based on dynamic evidence combination (MSF-DEC) is proposed that involves layer clustering and improved evidence theory. The proposed method refines the evidence through layer clustering and evidence elimination. We show that the MSF-DEC algorithm reduces the amount of input data of situation information fusion, improves the quality of input data, and provides more accurate and concise input for situation information fusion. Furthermore,

a rule of combination, based on dynamic construction, can effectively avoid conflicts in situation information synthesis, thus obtaining more accurate situational information fusion results and fusion efficiency.

The remainder of this paper is organized as follows. Section 2 is a summary of current information fusion methods. Section 3 introduces the basic principles of evidence theory. In Section 4, the proposed method is explained in detail. Section 5 presents the experimental simulation and analysis, and conclusions are reported in Section 6.

2. Related Work

Data fusion technology is widely used in the field of network security. Tim Bass introduced the Joint Directors of Laboratories (JDL) data fusion model in the field of network situational awareness [8]. Gucciardi proposed a data fusion method based on prior information credibility to overcome the limitations of Bayesian theory in the fusion of small data samples [9]. Good fusion accuracy can be achieved by measuring the differences between different groups of experimental data based on prior information credibility. But it is less efficient. Alhaj et al. proposed an alert clustering method based on attribute similarity [1]. The Alhaj method first sorted features in descending order based on high information gain entropy and then performed similarity correlation analysis based on the characteristics of high information gain entropy. The Alhaj method can achieve good efficiency, but the accuracy was not ideal [11]. Liu proposed a causality-based alert information fusion method where the causality between each alert event was established to form a knowledge base [4]. Then, associations between alerts were established using the knowledge base to generate what were known as superalerts. The method proposed by Liu has poor fusion accuracy in a complex network environment, with large resource consumption and low fusion efficiency [12]. Zheng et al. used the Pignistic distance to measure the similarity between different pieces of evidence based on sampling past evidence using a sliding window [13]. This established an uncertain state evidence model of the Markov chain and obtained the fusion results using a Murphy combination. However, this method simply averages multiple sets of evidence and does not consider the correlation between the various pieces of evidence, causing some data with large deviations to have a destructive effect on the entire fusion process. Zhao et al. proposed a combined method for conflicting evidence based on inconsistent measurements [14]. This method first measured the degree of conflict between two pieces of evidence and then modified the conflicting evidence by calculating a discount coefficient in order to improve the accuracy of the results of Dempster's combination. This method may overly weaken the weight of some large deviation evaluation information and lose part of the decision information [15]. Those methods can effectively deal with highly contradictory evidence and provide a good idea for evidence fusion, but they will also weaken the weight of some large deviation evaluation information. Pan et al. proposed a data fusion method based on a measure of uncertainty [16], which facilitated dividing the evidence into

reliable and unreliable types. This method then used information entropy to measure the amount of information contained in the evidence, which was regarded as the weight of each item of evidence. However, this method has a slower convergence rate and requires more evidence accumulation. Mihai established an evidence combination rule according to the degree of conflict [17]. When the conflict was large, Mihai assigned more trust to the combination rule based on a set union. When there was less conflict, more trust was assigned to the combination rule based on set intersection. Mihai's method is rational; however, it fails to address the highly conflicting evidence often involved in the fusion process. Jing et al. proposed a probability transformation method based on the correlation coefficient of belief function, which maximizes the correlation coefficient between the given BPA and the transformed probability distribution and can reflect the original information of the given BPA to the maximum extent [18]. Wu et al. proposed a method of using belief entropy to calculate the uncertainty of each evidence body, which can better improve the classification or fusion effect [19, 20]. Tang et al. analyzed the relationship between conflict caused by incomplete information and information fusion and proposed a generation method generalized basic probability assignment (GBPA) based on the triangular fuzzy number model under the open-world assumption [21]. Therefore, in general, past research has only transplanted traditional data fusion methods, such as association analysis, Bayesian networks, and evidence theory, to the problem of network security situational information fusion, without comprehensively considering the multiclassification, multisource, and highly conflicting nature of situational information. This results in an inefficient fusion process with low fusion accuracy [22, 23].

To address these problems, we propose a multisource situational information fusion method based on dynamic evidence combination (MSF-DEC). The MSF-DEC algorithm first clusters and classifies the alert information hierarchically, based on features of the information, to reduce the range of situational information fusion. Second, an evidence distance metric and bitmap method are used to identify and eliminate highly conflicting evidence to further reduce the amount of alert data and decrease the number of alerts to be fused. Third, the reduced alerts are used as the fusion evidence of MSF-DEC and the combination rule is dynamically adjusted based on the conflict information. The method not only fully considers the hierarchical and highly conflicting nature of alert data but also makes full use of the complementarity between alert data sources to achieve accurate fusion of multisource situational information.

3. Evidence Theory

As an indeterminate reasoning theory [24], evidence theory was first proposed by Dempster and later improved by Shafer [12]. Evidence theory is an extension of probability theory. In classical probability theory, $1 - p$ indicates the negation of proposition p . However, in evidence theory, $1 - p$ denotes the probability of the entire set except for proposition p . The core of evidence theory is Dempster's rule

of combination, through which the effective fusion of multisource information can be achieved [25].

Definition 1. When $\forall A \subseteq \Theta$, m_1 and m_2 are BPA functions in the identification framework Θ . If m_1 and m_2 are two independent basic plausibility assignments in 2Θ , then the combination rule can be expressed as [26]

$$m(A) = \frac{\sum_{A_1 \cap A_2 = A} m_1(A_1) m_2(A_2)}{1 - K}, \quad A \neq \phi. \quad (1)$$

The Dempster rule of combination for n items of evidence is as follows [26]:

$$m(A) = \frac{\sum_{\cap A_i = A} m_1(A_1) m_2(A_2), \dots, m_n(A_n)}{1 - K_n}, \quad A \neq \phi. \quad (2)$$

Whether there are two pieces of evidence or n pieces, it is first necessary to define normalization factors, K or K_n , to reflect the degree of conflict between the pieces of evidence to be combined. The greater the value of K or K_n , the greater the conflict in the evidence. As stipulated by evidence theory, $m(\phi) = 0$; however, when the conflict is too large, the probability of the null set ϕ during the evidence combination process is not always equal to 0, which may lead to unreasonable results. This is the classic Zadeh paradox [24]. Therefore, when evidence theory is used to solve the fusion problem with highly conflicting data sources, the accuracy of the fusion results may be poor.

4. Multisource Information Aggregation Method Based on Dynamic Evidence Combination

4.1. Basic Principles. In view of the problems with the application of evidence theory, many solutions have been put forward to improve the theory's adaptability for multisource fusion. There are three ways to improve the theory [27]. The first approach is to estimate the importance of pieces of evidence and modify their weight in the combination of the results. The second is to improve the combination rule. The third is to use methods, such as neural networks, to optimize the combination rule. Since the security components are affected by factors including the data collection method and the application environment, the security incidents obtained vary greatly in sensitivity [28]. Also, the evidence provided by the less reliable security components should be discarded because they often conflict with the evidence generated from more reliable sources [29].

We propose MSF-DEC to address these problems. The MSF-DEC algorithm first eliminates redundant and misleading data using a conflict evidence elimination method. On this basis, the evidence combination rule is constructed dynamically to improve the multisource information fusion capability of evidence theory. The fusion process is divided into three phases: alert aggregation, evidence elimination, and evidence fusion (Figure 1).

FIGURE 1: Multisource situational information fusion based on dynamic evidence combination.

- (1) Alert aggregation: the original alert data are clustered and classified to reduce the quantity of alerts and improve the efficiency of fusing the alert information
- (2) Evidence elimination: the distance between multiple sources of evidence and degree of conflict between pieces of evidence are measured so that the influence of highly conflicting alerts on the fusion is eliminated, thereby improving the accuracy of situational information fusion
- (3) Evidence fusion: the evidence is fused based on the dynamically constructed combination rule, which effectively avoids the Zadeh paradox and improves the accuracy of situational information fusion

4.2. Alert Aggregation. As mentioned in Section 4.1, there is a problem with fusing the alert information. Specifically, that there are too many alerts to process efficiently. By analyzing the alert data, we found that when a target was being attacked, a large number of identical or similar alerts occur in a short time, but these repeated alerts describe the same event or attack and therefore provide little new information [30–32]. To address this issue, classification based on spatial features (CSF) and alert layer clustering based on multi-feature similarity (ALC-MFS) are proposed in this study, the basic principle of which is to divide the aggregation process into two levels. Firstly, the alerts are classified using the spatial features of the IP address by the CSF algorithm. Then, for each classification, the ALC-MFS algorithm is used to cluster the alerts. The ALC-MFS algorithm can reduce the volume of the original alert information, thereby improving the efficiency of the alert fusion.

4.2.1. Classification Based on Spatial Features. Since IP addresses have regional features, the closer the destination IP addresses to the attack event, the stronger is the possibility of association between events. In other words, there is a high correlation between alert events covered by the same IP address. Therefore, based on the type of IP address, the alert information is divided into five categories (A, B, C, D, and others) according to the destination IP address in the alert information. The numbers of the five categories of addresses are denoted by A_0, A_1, A_2, A_3 , and A_4 , respectively, where the subscript numbers represent the address ranges.

The pseudocode description of classification based on spatial features is shown in Algorithm 1. Assuming that the number of IP addresses is n , the computational complexity of classification based on Algorithm 1 is $O(n)$.

```

Input: original alert information
Output: classification result
(1)  Begin
(2)     $n = \text{AlertsNum}(\text{alerts});$  //alert numbers
(3)     $q = 0;$  //set the counter
(4)    While ( $q < n$ )
(5)      IP = GetIP(alerts); //get IP in alerts
(6)      If IP in [10.0.0.0 10.255.255.255]
(7)        return class  $A_1$ ;
(8)      Else if IP in [128.0.0.1 191.255.255.254]
(9)        return class  $A_2$ ;
(10)     Else if IP in [192.0.0.1—223.255.255.254]
(11)       return class  $A_3$ ;
(12)     Else if IP in [224.0.0.1—239.255.255.254]
(13)       return class  $A_4$ ;
(14)     Else
(15)       return class  $A_0$ ;
(16)     End if
(17)      $q = q + 1;$ 
(18)   End while
(19) End

```

ALGORITHM 1: Classification based on spatial features (CSF).

4.2.2. Alert Layer Cluster Based on Feature Similarity. Based on the classification of spatial features, the alert layer clustering method based on feature similarity is applied in the second layer. To further improve the alert analysis efficiency of the system, similarity fusion is conducted on the security events to remove redundancy and reduce the amount of alert data. Security events in the network are not commonly isolated but related to each other. In this case, redundancy can be reduced by combining information belonging to the same type of alert into one record. Using the above method, similar alert information is correlated and merged according to specific rules and the original alerts are filtered, which not only greatly reduces the number of alerts and relieves the processing burden, but also improves the accuracy of the fusion.

(1) *Feature Selection.* The similarity between alerts is estimated according to their attributes. The attributes of an alert record in this study include source IP, destination IP, source port, destination port, protocol type, and time. The similarity is expressed as a value in the range [0, 1]. The larger the value of the similarity, the closer the attributes. A value of 1 means that the attributes are exactly the same while 0 indicates no similarity.

The various attributes of the alerts are first compared, and then the similarity between the alerts is calculated based on the alert similarity function. Finally, the similarity between the alert and the overall cluster is calculated to determine whether to generate a new cluster or add the alert to the old cluster.

The similarity function of each attribute is defined as follows:

- (1) Similarity of IP addresses: IPV 4 and IPV 6 can both be represented by a binary number. For IPV 4, 1 byte

contains 8 binary numbers, while 2 bytes can represent 16 binary numbers in the case of IPV 6. The similarity of the corresponding part of each piece of alert information is calculated separately before calculating the overall similarity of the entire IP address.

Definition 2. The similarity of IP addresses refers to the proportion of similarities between two IP addresses. This is calculated using the following equations:

$$S(a_i) = \frac{\varepsilon}{\omega}, \quad (3)$$

$$S(\text{IP}) = \sum_{i=1}^n (\delta_i * S(a_i)). \quad (4)$$

In the above equations, $S(a_i)$ represents the similarity of each part of the IP address. ε represents the number of bits with a similarity of 1 in IP address, ω represents the total number of bits in IP address, and $S(\text{IP})$ denotes the overall similarity of the IP address. The value of $S(a_i)$ ranges from 1 to n . n can be 4 or 8, which represents each component of the IP address, and δ_i denotes the weight of the component.

- (2) Similarity of ports: an attack on a port often follows a regular pattern, and the attacked port generally has continuous characteristics. Hence, the possibility of being attacked can be assessed by the similarity of ports which is defined as follows.

Definition 3. Similarity of ports refers to the distance between them which can be calculated based on the following equation:

$$S(\text{Port}) = \begin{cases} \frac{\beta - |X_p - Y_p|}{\beta}, & (|X_p - Y_p| \leq \beta), \\ 0, & (|X_p - Y_p| > \beta), \end{cases} \quad (5)$$

where $S(\text{Port})$ denotes similarity of ports and β is a preset threshold. When the distance between ports is greater than β , the probability that the alerts are associated is low, and the similarity of ports is set to 0. X_p and Y_p represent the port values of the alerts X and Y .

- (3) Similarity in time: for some attacks like DOS or DDOS, when a network attack is launched, many data packets will be generated within a short time. The alert records captured by the security sensor are also densely distributed over time. Meanwhile, signals not associated with the attack are more widely distributed. Therefore, when calculating the similarity of alert information, time factors can provide useful information.

Definition 4. Similarity of time refers to the distance in time between alerts in terms, which can be calculated based on the following equation:

$$S(\text{Time}) = \begin{cases} \frac{\lambda - |X_t - Y_t|}{\lambda}, & (|X_t - Y_t| \leq \lambda), \\ 0, & (|X_t - Y_t| > \lambda), \end{cases} \quad (6)$$

where $S(\text{Time})$ is similarity in time and λ is a threshold. Only when the time interval between alerts is less than λ can a correlation be established. In this study, it is assumed that alerts with a time interval exceeding this range are unlikely to be correlated. Thus, the similarity of events exceeding the threshold is set to 0. X_t and Y_t denote the occurrence time of alerts X and Y , respectively.

According to the experiment, in order to achieve a better clustering effect, the values of β and λ in equations (5) and (6) are in the range of [0.7, 0.9]. Specifically, according to the actual needs, when the amount of alarm is large and the amount of alarms needs to be compressed to a large extent, β and λ use smaller values to divide the alarm information into larger clusters. When the number of alarms is small and the number of alarms only needs to be compressed to a small extent, β and λ adopt larger values to divide the alarm information into multiple clusters.

Definition 5. A feature vector refers to the vector formed by the similarity of each feature attribute in the alert.

- (4) Similarity of alerts is defined as follows.

Definition 6. The similarity of alerts is the weighted sum of the similarity of all feature attributes in the alerts.

The overall similarity function between alert events is defined as shown in the following equation:

$$S(X, Y) = \sum_{i=1}^n w^{(i)} S_{xy}^{(i)}, \quad \left(\sum_{i=1}^n w^{(i)} = 1 \right), \quad (7)$$

where $w^{(i)}$ denotes the weight corresponding to the i th attribute, n is the number of attributes considered, and $S_{xy}^{(i)}$ is the similarity of the i th attribute between alerts X and Y .

- (5) The similarity between a new alert and an existing cluster of alerts is defined as the weighted sum of the similarity of the existing cluster and the similarity between the new alert and the most recent alert, which can be calculated using the following equation:

$$S(n+1) = a * S(n) + (1-a) * S(X, Y), \quad (8)$$

where $S(n+1)$ is the similarity of the new cluster, $S(n)$ is the similarity of currently existing cluster, $S(X, Y)$ denotes the similarity between the most recent alert and the new alert in the cluster, and a is the weight. In general, because the number of alarms to be classified is relatively large, according to the experiments, the value of a is generally taken as the larger value in [0.7, 0.9], which can keep the balance between retaining the historical similarity calculation results and adding the new similarity calculation result.

(2) *Algorithm Description.* The pseudocode description of the aggregation procedure is shown in Algorithm 2.

The correlation algorithm determines whether a new alert and the original alert are correlated and whether a cluster relationship should be established between the two by comparing the similarity value with the preset threshold. If the similarity between the alerts is greater than the similarity threshold, there is a strong correlation between the alerts, and the newly received alert will be merged into the established alert class. Otherwise, if the similarity is less than the set threshold, the new alert is not strongly correlated and a new alert class should be created and added to the database. Hence, threshold setting is fundamental to determining the clustering effect in this algorithm. The threshold in this study was established by repeated experiments in a specific experimental environment.

4.3. *Elimination of Conflicting Evidence.* Based on the method used by Pan and Deng to measure the relationship between pieces of evidence [19], we assume that there are n evidence items in the identification framework. If m_1 and m_2 are two BPAs in the framework, then the distance between m_1 and m_2 can be expressed as

```

Input: alert information
Output: association result
(1) Begin
(2)   Set ( $T_i, T_p, T_T, T$ ); //initialization, set the similarity threshold
(3)   Set ( $S$ ) = 0; //initialization, set the initial similarity as 0
(4)   {Alert1, Alert2, Alert3, ...}; //initialize the alert records in time t
(5)    $q = 1$ ; //set the counter
(6)   While ( $q \leq n$ )
(7)     GetAlert ( $A_k$ ); //select the  $q_{th}$  alert
(8)      $d = \text{Calculate } S$  (); //calculate the similarity using equation (7)–(11)
(9)     If ( $d > S$ )
(10)       $S = d$ ;
(11)     Else
(12)       $q = q + 1$ ;
(13)     End if
(14)   End while
(15)   If ( $S < T$ ) then
(16)     Add (new_class) //add a new class of alert
(17)   Else
(18)     Add (new_alert) //add a new class of alert
(19)   End if
(20) End

```

ALGORITHM 2: Alert layer cluster based on multifeature similarity (ALC-MFS).

$$d_{\text{BPA}}(m_1, m_2) = \sqrt{\frac{1}{2}(\vec{m}_1^2 + \vec{m}_2^2 - 2\vec{m}_1 \cdot \vec{m}_2)}, \quad (9)$$

where $\vec{m}^2 = \vec{m} \cdot \vec{m}$ is the inner product of the two vectors. This method for calculating the distance between pieces of evidence is applicable in the case of multisource evidence fusion. A compatibility matrix is obtained using this method. Assuming there are n data sources and each source collects information independently to obtain n items of evidence, each piece of evidence constitutes an m -dimensional vector and the distances between evidence items m_i and m_j forms the matrix:

$$\text{DM} = \begin{bmatrix} 0 & d_{12} & \dots & d_{1n} \\ d_{21} & 0 & \dots & d_{2n} \\ \dots & \dots & \dots & \dots \\ d_{n1} & d_{n2} & \dots & 0 \end{bmatrix}. \quad (10)$$

β_{ij} ($1 \leq i, j \leq n$), the boundary of d_{ij} , was set based on a combination of expert knowledge and the results of our experiments. The relationship between d_{ij} and β_{ij} can be utilized as follows:

$$\text{tc}_{ij} = \begin{cases} 1, & d_{ij} \leq \beta_{ij}, \\ 0, & d_{ij} > \beta_{ij}, \end{cases} \quad (11)$$

Equation (11) is converted into the evidence similarity matrix TS in the following equation:

$$\text{TS} = \begin{bmatrix} 1 & \text{tc}_{12} & \dots & \text{tc}_{1n} \\ \text{tc}_{21} & 1 & \dots & \text{tc}_{2n} \\ \dots & \dots & \dots & \dots \\ \text{tc}_{n1} & \text{tc}_{n2} & \dots & 1 \end{bmatrix}. \quad (12)$$

In equation (12), if $\text{tc}_{ij} = 0, 1 \leq i, j \leq n$, the similarity between the evidence displayed by the i th sensor and the j th sensor is low, that is, there is a conflict. If $\text{tc}_{ij} = 1, 1 \leq i, j \leq n$, then the evidence submitted by the i th sensor and the j th sensor is in agreement. In other words, there is a slight or even no conflict within the evidence, which will not reduce the fusion accuracy.

In general, the fusion results obtained based on the evidence theory are reliable. The paradox mainly occurs when there is conflicting evidence. Thus, by properly eliminating highly conflicting evidence, the fusion results can be optimized.

If pieces of evidence in equation (12) conflict with each other, the following actions are taken. First, the norm is defined as follows:

$$Y_i = \|m_i\| = \left(\sum_{j=1}^m m_i(A_j) \right)^{1/2}, \quad 1 \leq i \leq n, \quad (13)$$

where m is the number of individual pieces of evidence. The norm Y_i is ranked from small to large as follows: $Y_1 \leq Y_2 \leq \dots \leq Y_n$. Then, Y_n is the upper limit value, Y_1 is the lower limit value, and the median is obtained:

$$M = \begin{cases} Y_{(n+1)/2}, & n \text{ is an odd integer,} \\ \frac{(Y_{(n+1)/2} + Y_{n/2})}{2}, & n \text{ is an even integer.} \end{cases} \quad (14)$$

Similarly, the upper quartile is the median of the interval $[M, Y_n]$ and the lower quartile is the median of the interval $[Y_l, M]$, which are denoted by F_u and F_l , respectively. The dispersion of the quartile can be expressed as

$$dF = F_u - F_l, \quad (15)$$

If the distance between the norm Y_l and the median is larger than βdF , then the gross error is produced. When $|Y_l - M| > \beta dF$, Y_l is considered invalid and the elimination point is

$$\rho_1 = F_l - \frac{\beta}{2} dF, \quad (16)$$

$$\rho_2 = F_u + \frac{\beta}{2} dF,$$

where $\beta = 1, 2, \dots$. The elimination point is effective in the interval $[\rho_1, \rho_2]$, and by adding evidence to the fusion engine, the result will become more accurate and robust.

4.4. Evidence Combination. Current combination rules can be divided into two types:

$$m(A) = \begin{cases} p(k) \sum_{B \cap C = A} m_1(B)m_2(C) + q(k) \sum_{B \cup C = A} m_1(B)m_2(C), \\ 0, & A = \phi, \end{cases} \quad (17)$$

where $p(k)$ is a monotonically decreasing function of k , $q(k)$ is a monotonically increasing function of k , and the range of k is $[0, 1]$. When $k = 1$, the conflicts are at their maximum and all the conflicts are assigned to the union proposition. When $k = 0$, the conflicts are at their minimum and they are all assigned to the intersection proposition. The sum of the weights of the disjunction conjunction rule is 1. That is, the following conditions are satisfied:

$$\begin{aligned} k &= 0, \\ p(k) &= 1, \\ q(k) &= 0, \end{aligned} \quad (18)$$

$$\begin{aligned} k &= 1, \\ p(k) &= 0, \\ q(k) &= 1, \end{aligned} \quad (19)$$

- (1) Assign the conflict trust to the union of conflict-related elements using the conjunction method
- (2) Assign the conflict trust to the intersection of conflict-related elements with the disjunction method

The first method retains the conflict trust in the union and waits for more evidence to be gathered so as to draw a more reliable conclusion. This is suitable for highly conflicting situations. The second method is suitable for cases with good consistency. In this study, an adaptive conflict evidence combination method is put forward based on the above two approaches. The main idea is that if the conflict within the evidence is small, the evidence will be combined using the conjunction rule by assigning the conflicts into the intersection propositional space. If the conflict is large, the evidence will be combined according to the disjunction rule by assigning the conflicts into the union propositional space. In other words, the weights of the disjunction rule and the conjunction rule vary as the amount of conflict changes.

4.4.1. Combination of Two Pieces of Evidence. Assume that the identification framework $\Theta = \{B, C\}$ consists of two items of evidence, and the corresponding BPA functions are $m_1(B)$ and $m_2(C)$. Then, the belief assignment function after evidence combination is as follows:

$$p(k) + q(k) = 1. \quad (20)$$

Values of $p(k)$ and $q(k)$ that satisfy equations (18)–(20) can both be used as the coefficient of the conjunction rule and the disjunction rule:

$$\begin{aligned} p(k) &= 1 - k, \\ q(k) &= k. \end{aligned} \quad (21)$$

$p(k)$ and $q(k)$ satisfy conditions (18)–(20).

4.4.2. Combination of Several Pieces of Evidence. Assume that identification framework $\Theta = \{E_1, E_2, \dots, E_n\}$ consists of n pieces of evidence whose BPA functions are $m_1, \dots, m_n$. Then, the belief assignment function after the combination of n pieces of evidence can be expressed as follows:

$$m(A) = \begin{cases} p(k) \sum_{\cap_{i=1}^n A_i = A} \prod_{i=1}^n m_i(A_i) + q(k) \sum_{\cup_{i=1}^n A_i = A} \prod_{i=1}^n m_i(A_i), & A = \phi. \end{cases} \quad (22)$$

By combining two pieces of evidence, the conflict is equal to the sum of the belief products of the elements whose intersections are empty. Correspondingly, when it comes to the combination of n pieces of evidence, the conflict calculation equation is as follows:

$$k = \sum_{\cap_{i=1}^n A_i = \phi} \prod_{i=1}^n m_i(A_i). \quad (23)$$

5. Simulation and Result Analysis

5.1. Simulation Methods. To evaluate the performance of the MSF-DEC algorithm, the DARPA 2000 dataset was used in our experiments. As a set of data collected by the Lincoln Lab to test IDS performance, the DARPA 2000 dataset is currently widely used in academia. The DARPA 2000 dataset includes DDOS attacks, including two attack scenarios: LLDOS1.0 and LLDOS2.0. This makes it suitable for verifying the fusion performance of the alert system and assessing the performance of the situation information fusion method proposed in this study.

Network topology was first set up as shown in Figure 2. The computer mainframe has a frequency of 3.2 GHz and a memory of 8 GB. The data sources are NetFlow Log Analyzer and RealSecure, while the operating systems were Windows 7 and Solaris 9. The LLDOS1.0 data from DARPA 2000 was reset using NetPoke. No other network-related operations were performed in the experiment to simulate real network scenarios.

5.2. Alert Preprocessing Test. In the experiment, 1249 pieces of alert data were collected via NetPoke. The data contain large quantity, false alerts, and repetition. According to the alert information of the NetFlow sensor, the log information flow in the network is approximately 770–860 k/s when there is no attack in progress. This is too high for the fusion algorithm to run efficiently. Therefore, the alert data must be preprocessed before information fusion and the amount of alerts involved must be reduced to improve the efficiency of information fusion.

5.2.1. Alert Reduction Rate. This is completed by clustering the alerts and eliminating those in conflict. Firstly, the multifeature similarity layer clustering algorithm is used to conduct cluster analysis on the original data. In the subsequent evidence fusion process, BPA is evaluated for each evidence source. Since the port changes and the flow-in ratio of the traffic accurately reflect the attack received by the network, the ratio of these two indicators serves as the primary basis to conduct BPA for RealSecure and NetFlow. Table 1 shows the BPAs of the two indicators.

As an indicator to measure the efficiency of reducing the number of alerts, the reduction rate (RR) can be calculated as follows:

FIGURE 2: Experiment topology.

TABLE 1: Basic probability assignment of alert.

Class	Source	
	RealSecure	NetFlow
<i>Sadmin_Ping</i>	0.118	0.322
<i>Stream_Dos</i>	0.295	0.212
<i>Mstream_Zobie</i>	0.063	0.251
<i>Sadmin_Amslverify_Overflow</i>	0.129	0.051
<i>RSH</i>	0.163	0.052
<i>New</i>	0.232	0.112

$$RR = \frac{A_{\text{sum}} - K}{A_{\text{sum}}} * 100\%, \quad (24)$$

where RR denotes the reduction rate, K is the number of alerts after processing, and A_{sum} stands for the number of alerts originally generated.

Results of clustering tests for the effect of different similarity thresholds are shown in Figure 3.

Here, the RR is shown to be strongly affected by the similarity threshold T . When the similarity threshold M is large, there is a high requirement for clustering. Consequently, the lower the RR of the clustering results, the larger the number of alerts finally obtained. Conversely, when the similarity threshold is small, the requirement for clustering is relatively low. Thus, as the RR in the experimental results increases, the number of alerts output is correspondingly reduced. It can also be seen from the figure that a similarity threshold of 0.6 is an inflection point for the change in RR. When the similarity threshold is less than 0.6, the RR changes relatively slowly, but when the value is greater than 0.6, the RR decreases rapidly. After several tests, a reasonable range of the similarity threshold T was established as [0.5, 0.7]. As shown in Figure 3, when the similarity threshold $T=0.5$, the RR is 93.8%, which can greatly reduce the number of alerts. This is beneficial for the later evidence fusion, laying the basis for accurate alert information fusion.

5.2.2. Clustering Stability. Stability is a key index to evaluate the effectiveness of the clustering method. To test the stability of the proposed method, six experiments were carried out based on the collected alert data. In the experiments, the

FIGURE 3: Influence of similarity threshold on the reduction rate.

similarity thresholds were varied from 0.1 to 0.6 in increments of 0.1. An experimental comparison was conducted between the ALC-MFS algorithm proposed in this study and single-layer clustering based on attribute similarity, as proposed by the Alhaj method. The results are shown in Figures 4 and 5.

Figure 4 shows the experimental results obtained by the Alhaj method. It can be seen that the number of alerts in each cluster shows no clear trend as the threshold varies. This indicates that particular alerts were not clustered consistently according to their attributes during the allocation process. Figure 5 shows the experimental results obtained by the ALC-MFS algorithm. Here, the number of alerts in each cluster changes steadily. This shows that the alerts within each cluster are highly correlated, according to their attributes, while the correlation between clusters is comparatively weak. Therefore, the clustering results are quite effective.

Based on further analysis of the clustering results, we found out that cluster 1 mainly includes two types of alert events: Echo Reply and ping scanning; the number of alerts was 421 and 419, respectively. Echo Reply is a message generated in reply to ping scanning; therefore, the Echo Reply alert is generated in the process of replying to the ping scanning. Hence, the two types of alerts are strongly correlated with each other. Moreover, the results for cluster 1 are reasonable, with an accurate response to the attack behavior. When $T=0.6$, almost all alert events are assigned to the same cluster, which indicates that the cluster size monotonously and stably changes with the change of cluster granularity. Hence, the clustering process of the method proposed in this study is stable.

5.3. Fusion Accuracy Test. To verify the validity and fusion efficiency of the MSF-DEC algorithm, it was compared with

FIGURE 4: Single-layer clustering results based on attribute similarity.

typical alert information processing methods proposed by other researchers [1, 14]. In order to facilitate the comparison of the above methods, the following detection indicators were used, such as detection rate (DR) and false detection rate (FDR).

Definition 7. DR can be expressed as the ratio of the number of actual attacks to the number of observed attacks, which can be expressed as

$$DR = \frac{\# \text{actual attacks}}{\# \text{observed attacks}}, \quad (25)$$

where # represents the number of attacks.

Definition 8. FDR refers to the ratio of the number of false alerts to the total number of alerts, which can be expressed as

$$FDR = 1 - \frac{\# \text{alerts reflecting actual attack}}{\# \text{whole alerts}}. \quad (26)$$

As shown in Figures 6–8, the DS method, Xiao method, Mihai method, and the method proposed in this study have obvious advantages in terms of DR and FDR. The method proposed by Alhaj et al. has low accuracy but high processing efficiency, mainly because the evidence fusion method was not applied, and therefore, evidence theory consumed more time. The MSF-DEC algorithm has the optimal overall performance because it reduces the volume of evidence by removing conflicting pieces of evidence and therefore greatly reduces the amount of evidence involved in the fusion, thereby improving the fusion efficiency and having similar performance with the Alhaj method in fusion time. Moreover, the dynamic evidence combination rule adopted by the MSF-DEC algorithm effectively reduces the influence of highly conflicting evidence on the fusion results, improving the overall fusion accuracy.

FIGURE 5: Alert layer clustering results based on feature similarity.

FIGURE 6: Comparison of different methods in terms of detection rate (DR).

6. Conclusion

The MSF-DEC algorithm was proposed in this study to overcome the problems associated with current situational information fusion algorithms in terms of fusion efficiency, DR, and FDR. The method divides the situational information fusion process into three stages. First, the ALC-MFS algorithm is used to cluster the original alert information to reduce the volume of alert data. Next, evidence distance and bitmap methods were adopted to eliminate highly conflicting evidence to further reduce the alert data. Finally, with the reduced alerts as the fusion evidence, the multisource situational information was accurately and efficiently fused based on dynamically constructed evidence combination rules.

One of the limitations of this method is that, during information fusion, the remaining evidence, after removing the conflicting alerts, is given with the same weight. In other

FIGURE 7: Comparison of different methods in terms of false detection rate (FDR).

FIGURE 8: Comparison of different methods in terms of fusion efficiency.

words, each data source is assigned with the same belief. This method cannot reflect the importance of each data source. In future research, evidence will be evaluated to improve the impact of more reliable evidence on situation information fusion. Evidence reliability measurement based on entropy and the conflict caused by incomplete information will be significant research directions. Additionally, future studies will weaken the influence of evidence that has low credibility, further improving the accuracy of multisource situational information fusion.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 61902427).

References

- [1] T. A. Alhaj, M. M. Siraj, A. Zainal et al., "Feature selection using information gain for improved structural-based alert correlation," *PloS One*, vol. 11, no. 11, Article ID e0166017, 2016.
- [2] Q. S. Qassim, A. M. Zin, and M. J. A. Aziz, "Anomaly-based network IDS false alarm filter using cluster-based alarm classification approach," *International Journal of Security and Networks*, vol. 12, no. 1, pp. 13–26, 2017.
- [3] A. Javaid, N. Javaid, Z. Wadud et al., "Machine learning algorithms and fault detection for improved belief function based decision fusion in wireless sensor networks," *Sensors*, vol. 19, no. 6, 2019.
- [4] J. Liu, S. Li, and R. Zhang, "Algorithm of reducing the false positives in IDS based on correlation analysis," *IOP Conference Series: Materials Science and Engineering*, vol. 322, no. 6, Article ID 062016, 2018.
- [5] J. Qi, P. Yang, L. Newcombe, X. Peng, and Z. Zhao, "An overview of data fusion techniques for Internet of Things enabled physical activity recognition and measure," *Information Fusion*, vol. 55, pp. 269–280, 2020.
- [6] P. Zuo, Y. Hua, Y. Sun et al., "Bandwidth and energy efficient image sharing for situation awareness in disasters," *IEEE Transactions on Parallel and Distributed Systems*, vol. 30, no. 1, pp. 15–28, 2018.
- [7] L. Snidaro, I. Visentini, and K. Bryan, "Fusing uncertain knowledge and evidence for maritime situational awareness via Markov logic networks," *Information Fusion*, vol. 21, pp. 159–172, 2015.
- [8] T. Bass, "Intrusion detection systems and multisensor data fusion," *Communications of the ACM*, vol. 43, no. 4, pp. 99–105, 2000.
- [9] D. F. Gucciardi and B. Jackson, "Understanding sport continuation: an integration of the theories of planned behaviour and basic psychological needs," *Journal of Science and Medicine in Sport*, vol. 18, no. 1, pp. 31–36, 2015.
- [10] J. An, M. Hu, L. Fu, and J. Zhan, "A novel fuzzy approach for combining uncertain conflict evidences in the dempster-shafer theory," *IEEE Access*, vol. 7, pp. 7481–7501, 2019.
- [11] X. Fan, Y. Guo, Y. Ju, J. Bao, and W. Lyu, "Multisensor fusion method based on the belief entropy and DS evidence theory," *Journal of Sensors*, vol. 2020, no. 10, pp. 1–16, 2020.
- [12] F. Ye, P. Bai, and Y. Tian, "An algorithm based on evidence theory and fuzzy entropy to defend against SSDF," *Journal of Systems Engineering and Electronics*, vol. 31, no. 2, pp. 243–251, 2020.
- [13] X. Zheng, B. Huang, D. Ni, and Q. Xu, "A novel intelligent vehicle risk assessment method combined with multi-sensor fusion in dense traffic environment," *Journal of Intelligent and Connected Vehicles*, vol. 1, no. 2, pp. 41–54, 2018.
- [14] Y. Zhao, R. Jia, and P. Shi, "A novel combination method for conflicting evidence based on inconsistent measurements," *Information Sciences*, vol. 367–368, pp. 125–142, 2016.
- [15] V. Shah, A. Aggarwal, A. Aggarwal, and N. Chaubey, "Alert fusion of intrusion detection systems using Fuzzy Dempster Shafer theory," *Journal of Engineering Science and Technology Review*, vol. 10, no. 3, pp. 123–127, 2017.
- [16] L. Pan and Y. Deng, "A new belief entropy to measure uncertainty of basic probability assignments based on belief function and plausibility function," *Entropy*, vol. 20, no. 11, p. 842, 2018.
- [17] M. C. Florea and A. L. Jousselme, E. Bosse and D. Grenier, "Robust combination rules for evidence theory," *Information Fusion*, vol. 10, no. 2, pp. 183–197, 2009.
- [18] M. Jing and Y. Tang, "A new base basic probability assignment approach for conflict data fusion in the evidence theory," *Applied Intelligence*, vol. 51, no. 2, pp. 1056–1068, 2021.
- [19] D. Wu, Z. Liu, and Y. Tang, "A new classification method based on the negation of a basic probability assignment in the evidence theory," *Engineering Applications of Artificial Intelligence*, vol. 96, Article ID 103985, 2020.
- [20] D. Wu and Y. Tang, "An improved failure mode and effects analysis method based on uncertainty measure in the evidence theory," *Quality and Reliability Engineering International*, vol. 36, no. 5, pp. 1786–1807, 2020.
- [21] Y. Tang, D. Wu, and Z. Liu, "A new approach for generation of generalized basic probability assignment in the evidence theory," *Pattern Analysis and Applications*, vol. 24, pp. 1–17, 2021.
- [22] Y. Song, X. Wang, J. Zhu, and L. Lei, "Sensor dynamic reliability evaluation based on evidence theory and intuitionistic fuzzy sets," *Applied Intelligence*, vol. 48, no. 11, pp. 3950–3962, 2018.
- [23] A. Cheng, X. Jiang, Y. Li, C. Zhang, and H. Zhu, "Multiple sources and multiple measures based traffic flow prediction using the chaos theory and support vector regression method," *Physica A: Statistical Mechanics and its Applications*, vol. 466, pp. 422–434, 2017.
- [24] N. Derrick, L. Renfa, and W. Yongheng, "Particle swarm optimization and dempster shafer approach to achieve Internet of Things context fusion using quality of context," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 11, no. 2, pp. 247–264, 2016.
- [25] W. Li and B. Wang, "A synthetic method for situation assessment based on fuzzy logic and DS evidential theory," *Systems Engineering and Electronics*, vol. 10, 2003.
- [26] F. Yang and X. Wang, *Combination of Conflict for D-S Evidence Theory*, National Defense Industry Press, Beijing, China, 2010.
- [27] X. Yu, F. Zhang, L. Zhou et al., "Novel data fusion algorithm based on event-driven and Dempster-Shafer evidence theory," *Wireless Personal Communications*, vol. 100, pp. 1377–1391, 2018.
- [28] A. Maseleno and M. Mahmud Hasan, "The Dempster-Shafer theory algorithm and its application to insect diseases detection," *International Journal of Advanced Science and Technology*, vol. 50, pp. 111–120, 2013.
- [29] D. Yu and D. Frincke, "Alert confidence fusion in intrusion detection systems with extended Dempster-Shafer theory," in *Proceedings of the 43rd Annual Southeast Regional Conference*, pp. 142–147, New York, NY, USA, March 2005.
- [30] N. Dhanachandra and Y. J. Chanu, "An image segmentation approach based on fuzzy c-means and dynamic particle swarm optimization algorithm," *Multimedia Tools and Applications*, vol. 79, no. 25–26, pp. 18839–18858, 2020.
- [31] A. Mazher, P. Li, T. A. Moughal, and H. Xu, "A decision fusion method using an algorithm for fusion of correlated probabilities," *International Journal of Remote Sensing*, vol. 37, no. 1, pp. 14–25, 2015.
- [32] J. Zhang, B. Yu, and J. Li, "Research on IDS alert aggregation based on improved quantum-behaved particle swarm optimization," in *Proceedings of the International Conference Computer Science and Technology (CST2016)*, pp. 293–299, Shenzhen, China, 2017.