Research on Network Attack Effect Evaluation Based on Confrontational Perspective*

Jingxuan Li a, Han Qiu b, Junhu Zhu c
PLA Strategic Support Force Information Engineering University, Zhengzhou, Henan Province, China
a719587438@qq.com, 18695869453.86.
bqiuhan410@aliyun.com
Zhujunhu74@163.com

ABSTRACT. The cyberspace battlefield is full of strategic confrontations between attack and defense sides. However, the existing evaluation methods do not consider the impact of defense strategies on attack effect, and only carry out unilateral attack effect evaluation from the perspective of attacker, which cannot reflect the dynamic changes of attack effect against different defense strategies. To solve this problem, we propose the Confrontation Oriented Attack Effect Evaluation Model (CO-AEEM) to evaluate the attack effect from the perspective of both attack and defense sides. By creating CO-AEEM, we establish a multi-level evaluation index system covering all assets-elements of the target object, and propose an attack effect evaluation algorithm to reduce the impact of subjective factors and adapt to the complex characteristics of network confrontation, then design the comparison and evaluation process of attack test and confrontation test according to the changes of defense measures. Finally, the simulation experiments in various confrontation scenarios are carried out in network range, which show that the model can not only describe the actual attack effect in confrontation with different defense strategies, but also the effectiveness degree of defense strategies against various types of attacks. Thus, the evaluation results of CO-AEEM can provide scientific guidance of action decision making and equipment development in network counterwork.

CCS Concepts
- Security and privacy ➝ Formal methods and theory of security
- Computing methodologies ➝ Modeling and simulation.

1. INTRODUCTION
As human society moves toward a highly informational society, various cyberattack technologies are increasingly diverse, intelligent, and complex under the influence of interests. The evaluation technology of the network attack effect can give qualitative and quantitative evaluation of the effects of cyber attacks in a complex network environment[1], and it is of great significance to both attack and defend sides: on the one hand, it can help the attacker to evaluate the effectiveness and destructiveness of various attack tools, and develop more efficient attack strategies; on the other hand, it can show the damage degree of various attack methods, and help the defender to improve the network security protection capability by discovering the weak links of the system security.
Because the research on network attack effect evaluation is somewhat sensitive, there is little public research results. The existing research can be divided into the following two categories:

1) Research based on certain attack type or application background. The former mainly designs corresponding evaluation models according to the distinct characteristics of different attack types, for example, the effect evaluation of Trojans[2] and DDoS attacks[3][4]. The latter usually incorporates appropriate background knowledge to design a targeted effect evaluation approach, for example, the attack effect evaluation based on wireless sensor network[5], 3G Network[6] and Ad Hoc Network [7].

2) Research based on different mathematical theories. This kind of research mainly uses various mathematical theories to improve the evaluation model, making the evaluation results more accurate and scientific. For example, the attack effect evaluation based on rough set theory[1], grey theory[8], fuzzy theory[2] and network entropy[9].

The real cyber space is full of confrontation between offensive and defensive strategies, and the essence of confrontation is a process of restriction and dependence between strategies. Therefore, the attack effect is not only related to the attacker's ability, but also closely related to the effectiveness of defensive measures [10]. However, the existing evaluation methods only carry out unilateral evaluation of the implementation effect of the attack action from the perspective of the attacker, without considering the impact of different levels of defense measures on the attack effect in the case of confrontation. Such one-sided evaluation results cannot show the dynamic change of the attack effect in the face of different defense mechanisms. After analysis, we found that the existing evaluation methods of attack effect mainly have the following defects:

Firstly, attack actions show different attack effect in the face of different levels of defense measures. The attack effect values obtained through traditional methods by unilaterally characterizing attack capability without considering the impact of defense are not applicable to the real network confrontation scenarios.

Secondly, most existing studies build evaluation index system according to the type of attack or Information Security CIA Attributes, which are lack of generality and extensibility, and easy to cause index redundancy or omission, cannot cover all aspects of effect evaluation.

Thirdly, some indicators are relatively abstract, without good measurability and operability, leading to inaccurate and incomprehensive acquisition of evaluation data and affecting the scientificity of evaluation results.

Aiming at the above problems, the Confrontation Oriented Attack Effectiveness Evaluation Model (CO-AEEM) is constructed. This model establishes an evaluation index system with strong operability and comprehensive coverage towards all-assets-elements, and designs an attack effect quantitative algorithm that can reduce the influence of subjective factors and adapt to the complex characteristics of network confrontation. The model designs the evaluation process of comparing attack test and confrontation test, which can not only reflect the actual effect of attack actions in the face of different defense strategies, but also the effectiveness degree of defense strategies against different attack actions. It overcomes the existing evaluation methods’defect of ignoring the defensive effect and evaluating attack effect only from the perspective of attacker, and can provide practical scientific guidance for decision making and equipment development about network attack-defense combat. At last, the practicability and effectiveness of the model are verified by the simulation experiments of various confrontation scenarios.

The rest of the paper is organized as follows. In Section 2 we give the definition of our model—CO-AEEM and introduce the logical relationship among the various parts of CO-AEEM. In Section 3 we introduce the attack effect evaluation index system based on all-assets-elements, including how to acquire and process the index values. The attack effect evaluation algorithm from the confrontational perspective is described in Section 4. In Section 5, the simulation experiments carried out in various confrontation scenarios verify the practicability and effectiveness of the proposed model. At last, we summarize our work in Section 6.
2. CONFRONTATION ORIENTED ATTACK EFFECT EVALUATION MODEL

In the process of network counterwork, the specific attack and defense actions adopted by both sides to achieve their respective goals compose the attack-defense scene. Although there are many kinds of network attacks and their principles are different, their effects are all about destroying the safety features or resources of the target object [11], which are ultimately reflected by the changes of various attribute indexes of the target. The implementation of the defensive action can relieve and even eliminate the negative influence of attack action, reduce the degree of changes in related indicators of the victim system. It can be seen that the change of relevant indicators of the target object can reflect the attack effect under confrontation conditions. Therefore, two key problems need to be resolved for the network attack effect evaluation based on confrontational perspective: Firstly, to get the network index data which can reflect the attack effect overall by some technical means. This process is the confrontation simulation experiments carried out in network range. Secondly, to convert the acquired index data into the attack effect value under the confrontation condition through the evaluation and quantification algorithms, which will be described in detail in Section 4.

To measure the impact of different kinds of defense measures on the attack effect in confrontation scenario, it is necessary to take the attack effect when the defense is not implemented as a comparison, that is, to quantify and compare the attack effect from the attacker's perspective and the confrontation perspective respectively. Therefore, the simulation experiment in each counterwork scenario is divided into two parts. One is the pure attack test, under the circumstance of no defense deployment, only attack actions are applied and the network index data reflecting its impact on the target is collected to calculate the attack effect from the perspective of attacker. The second is confrontation test, in which both defense action and attack action to be evaluated are loaded to measure the degree of the target affected in the context of confrontation, and compares it with the attack test where no defense is applied, so as to reflect the attack effect from the perspective of confrontation.

Combined with the above analysis, we propose the Confrontation Oriented Attack Effect Evaluation Model $CO - AEEEM = \{A, D, R, U, \gamma, \beta\}$, where,

1) $A, D$ represent the behavior space of attacker and defender, $A = \{a_1, a_2, \ldots, a_m\}, \quad D = \{d_1, d_2, \ldots, d_n\}, m, n \in \mathbb{N}^+$. While, $a_i (i = 1, 2, \ldots, m), d_j (j = 1, 2, \ldots, n)$ respectively refer to the attack action and defense action to be evaluated.

2) $I$ represents the set of index values acquired in the confrontation simulation experiments carried out in network range. $I(a_i) \in A$ indicates the set of index values acquired in the attack test ($a_i, \emptyset$) that performs the attack action $a_i$ but does not load defense action; $I(a_i, d_j) \in A, d_j \in D$ indicates the set of index values acquired in the confrontation test ($a_i, d_j$) that performs the actions $a_i, d_j$. The attack effect evaluation index system based on all-assets-elements is established as a guide for acquiring index values in confrontation experiments, which will be described in detail in Section 3.

3) $\gamma$ is the mapping relationship between index set and attack effect, $E_a(a_i, \emptyset) = \gamma(I(a_i, \emptyset)), E_a(a_i, d_j) = \gamma(I(a_i, d_j))$. It represents the attack effect evaluation algorithm from the confrontational perspective, which will be described in detail in Section 4.

4) $E$ represents the set of attack effect values evaluated in all confrontation scenarios. In the attack test where no defense actions are deployed, the attack effect value of $a_i$ is expressed as $E_a(a_i, \emptyset)$. In the confrontation test, the attack effect value of $a_i$ when against $d_j$ is expressed as $E_a(a_i, d_j)$. When the behavior space of both sides is $A = \{a_1, a_2, \ldots, a_m\}, D = \{d_1, d_2, \ldots, d_n\}$ respectively, the attack effect evaluation results from confrontational perspective can be represented by a $n \times m$ matrix $\{E_a\}_{n \times m}$.

5) $\beta$ refers to defense effective rate , when the behavior space of both sides is $A = \{a_1, a_2, \ldots, a_m\}, D = \{d_1, d_2, \ldots, d_n\}$ respectively, it can be represented by a $n \times m$ matrix $\{\beta_{ij}\}_{n \times m}$. Where, $\beta_{ij}$ represents the effectiveness of defense action $d_j$ when it confront with attack $a_i$. We use the changes of attack effect to measure the effectiveness of defense, the evaluation results of attack test...
and confrontation test show that, the defense effect of $d_j$ when it confront with attack $a_i$ can be expressed as $E_d(a_i, d_j) = E_a(a_i, \emptyset) - E_a(a_i, d_j)$. Thus,

$$\beta_{ij} = \frac{E_d(a_i, d_j)}{E_d(a_i, \emptyset)}$$

When $E_d(a_i, \emptyset) > E_d(a_i, d_j)$, $\beta_{ij} \in (0,1]$, it indicates that defense $d_j$ can effectively reduce the attack effect of $a_i$. When $E_d(a_i, \emptyset) = E_d(a_i, d_j)$, $\beta_{ij} = 0$, it shows that the attack effect of $a_i$ does not change with or without defensive measure $d_j$, so $d_j$ is noneffective against $a_i$; When $E_d(a_i, \emptyset) < E_d(a_i, d_j)$, $\beta_{ij} \in (-\infty, 0)$, it indicates that $d_j$ can not effectively defend against $a_i$, on the contrary, it has the negative effect of increasing the damage effect.

The logical relationship among the elements of CO-AEEM is shown in Figure 1.

\[\text{Figure 1. Logical relationship of elements in CO-AEEM.}\]

In the real network confrontation environment, effective defense measures can reduce the damage effect brought by the attack and thus reduce the risk faced by the target system, while improper defense measures will increase the risk of the system and even endanger the normal operation of the network [12]. Therefore, CO-AEEM model can describe the actual attack effect and confrontation result of both attack and defense strategies in the case of counterwork through evaluation and comparison of attack test and confrontation test. This model has stronger practical significance compared to the traditional methods which evaluates attack effect only from the perspective of attackers: in terms of decision support, it enables both sides to choose more pertinence solutions in the real network confrontation environment and make more effective and scientific action decisions. In terms of equipment development, after both sides understand the confrontational effectiveness of each other’s strategies, they can focus on the research and development of their own shortcomings, and develop more pertinence and groundbreaking attack tools and defense mechanisms.

3. NETWORK ATTACK EFFECT EVALUATION INDEX SYSTEM BASED ON ALL-ASSETS-ELEMENTS

In this section, we introduce the network attack effect evaluation index system based on all-assets-elements, and the method of acquiring and processing index values in the confrontation simulation experiment.

3.1 Establishing Index System

The rationality of the index system is an important guarantee for the scientific and accurate evaluation results. Therefore, the construction of a scientific index system needs to follow the principles of pertinence, completeness, operability and independence [7].

No matter how the principle and form of network attacks evolve, the final effect will be reflected as the reduction of security attribute of the target equipment or the destruction and illegal control of
various resources. Moreover, compared with the complex and changeable attackers, the changes of target object are easier to acquire. Therefore, we establish a multi-level evaluation index system based on all-assets-elements. Firstly, the highest layer is called object layer, which expresses comprehensive evaluation result of attack effect. Then, we give a new definition of all the assets elements of the target object, including intangible assets and tangible assets: target resources are divided into host resources and network resources. According to the different attributes of the attack target, it is refined into five categories of assets: host information, host permissions, host performance, host data, and network performance. They form the criteria layer of the indicator system as a macro factor that affects the effectiveness of the attack. Finally, indexes that can describe the impact of various attack on each type of asset are selected to form the index layer. Furthermore, the value of each index in the index layer can be measured directly in the confrontation simulation experiments carried out in network range. The network attack effect evaluation index system based on all-assets-elements is shown in Figure 2.

![Figure 2. Network attack effect evaluation index system based on all-assets-elements.](image)

On the one hand, compared with the traditional method of constructing index systems based on the attack types, the index system in this paper does not have redundant indexes, and the indexes can be directly acquired in the network range simulation experiment, with good accuracy and operability. On the other hand, compared with the traditional method of establishing index system based on Information Security CIA Attributes classification, this method gives consideration to both tangible assets and intangible assets, and describes target assets in a more comprehensive way without omission of any indicator which can reflect attack or defense effect. In addition, the proposed index system also has strong expansibility. Since the changes caused by all threats that the target’s resources may face are included in the index system, even if a new attack mode emerges in the future, the proposed index system can comprehensively depict attack effect from the perspective of the system being attacked, and will hardly bring changes in indicators.

### 3.2 Acquiring Index Values

Since the effect of network attack is measured by the impact on targets indicators, the reliability of the original data acquisition is the key factor affecting the accuracy of the evaluation results. On the one hand, the more times simulation experiments are performed, the higher the reliability of the acquired index data will be. On the other hand, non-inductive data acquisition without disturbing the experimental scene can improve the accuracy of evaluation results. However, most of the existing researches focus on the construction of index system and the improvement of evaluation method, and few of them give a detailed introduction to the acquisition method of indicators, which has poor practical guidance. According to the characteristics of the underlying indicators and network range [13], the methods we used to acquire the original indicator data are summarized as follows:

1) View directly in the network range background management system: network range can provide approximate actual combat simulation environment and the conunterwork platform, but unlike the
opacity of information in real network environment, evaluators can get kinds of indicator data directly from the network range background management system, such as software and hardware configuration information of nodes in the topology, detailed logs of network traffic and equipment, etc. For example, as for detection accurate rate index \( (I_{12}, I_{14}, I_{16}, I_{18}, I_{1}, I_{10}) \), we can check the corresponding nodes’ configuration information in the background management system, so as to get the denominator data needed to calculate these ratio-type indexes.

2) Manual input: since individual qualitative indicators cannot be quantified by direct measurement or query, it is necessary for experts to conduct comprehensive evaluation according to the actual situation and historical experience of network confrontation, and then evaluators import indicator data manually, such as the relevant indicators \( (I_{23}, I_{43}) \) describing the importance of various assets.

3) Third-party tools: attack effect evaluation requires real-time monitoring and perceiving of dynamic changes in scene information, so it can use mature third party tools for automatic data acquisition and analysis, to meet the real-time requirements of some special indicators in the attack effect evaluation process. For example, Tsar, an open source acquisition tool, is often used to monitor and acquire information about system and network performance, while Nagios is often used to detect performance of hosts and services in real-time.

4) Flag: in the network range of confrontation simulation experiment, we can learn from The CTF (Capture The Flag) [14] information security competition mode, and set assets such as account keys and protected data as strings or other contents with a certain format in advance, which are called "Flag". Attackers get "Flags" by implementing attack actions, and take them as evidences of successfully stealing account or file information, gaining control of the target, thus providing evaluators with indicator information of the relevant attacked assets, such as indicators related to damage extent of host permissions and host data assets.

3.3 Standardizing Index Values

Due to the different physical significance and dimensions of each indicator, the original indicator data acquired in the simulation confrontation experiments cannot be used directly for evaluation and calculation, which needs to be standardized to make all kinds of indicators comparable at the same level [15]. According to the form of assignment, indicators are divided into qualitative indicators and quantitative indicators [1]. As for qualitative indicators, since the influence of all quantitative indicators on evaluation results is basically linear in our index system, the threshold method is selected for index standardization. Firstly, we divide the quantitative indicators into positive and negative ones according to the actual impact of indicators. The larger the positive indicator value is, the higher the efficiency will be (such as indicators of detected information amount and attacked data amount); the smaller the negative indicator value is, the higher the efficiency will be (such as the network throughput indicator). Supposing a group of original indicator data \( x_1, x_2, \cdots, x_l \) are acquired in a confrontation simulation experiment, two types of quantitative indexes are processed with Equations 2 and 3 respectively to obtain a group of standardized index data \( y_1, y_2, \cdots, y_l \), \( y_l \in [0,1] \):

For positive index:

\[
y_l = 0.9 \cdot \frac{x_l - \min(x_l)}{\max(x_l) - \min(x_l)} + 0.1
\]

For negative index:

\[
y_l = 0.9 \cdot \frac{\max(x_l) - x_l}{\max(x_l) - \min(x_l)} + 0.1
\]

The threshold \( \min(x_l) (\max(x_l)) \) is often the minimum (maximum) value of class \( i (i = 1, 2, \cdots, l) \) index in the data collected from the multiple simulation experiments, and may also be determined according to the evaluator's satisfactory value and impermissible value of the index.
4. NETWORK ATTACK EFFECT EVALUATION METHOD FROM THE CONFRONTATIONAL PERSPECTIVE

In this section, we introduce the attack effect evaluation method that converts the index value \( I \) into effect value \( E \). This section corresponds to the mapping \( y \) in CO-AEEM, which is the key link connecting the confrontation simulation experiment part with the evaluation and quantification part.

4.1 Combination Weights of Indexes

Since different indicators have different contributions to the effect evaluation results, it is necessary to design a reasonable method to determine the index weight. The commonly used methods of indicator weighting mainly include subjective and objective weighting method, but both have their own advantages and inherent defects: Subjective weighting method can reflect the historical experience and the intention of evaluators to some extent, but introduces more subjective factors. The objective weighting method takes the effective data collected in the experiment as the evaluation basis, but ignores the accumulation of previous experience knowledge and the will of the evaluator. Therefore, it is difficult to get accurate indicator weight values by relying only on a single method among them.

Aiming at the above problems, we design the combination weighting approach based on Fuzzy AHP in the subjective weighting method and Entropy Weight Method (EWM) in the objective weighting method, organically combining subjective experience and objective data to obtain more scientific combination weights of indicators.

Supposing \( w'_i \) and \( w''_i \) be the weight of index \( i \) determined by Fuzzy AHP and EWM respectively, and combine the weights by Equation 4 to obtain the final combination weight results \( w_l (w_l \in [0,1]) \):

\[
    w_l = \frac{w'_i \cdot w''_i}{\sum_{i=1}^{l} w'_i \cdot w''_i}, l = 1,2,\ldots, l
\]

4.2 The Non-linear Index Aggregation Method

The index system based on all-assets-elements is a multi-level structure, to get the final performance evaluation results, it is necessary to aggregate the underlying indicator data layer by layer to the top level. Most of the current evaluation methods adopt linear weighting method for indicator aggregation, which oversimplifies the influence relationship among index levels. Due to the dynamic nature of confrontation in real network environment, the indicators of effectiveness evaluation aggregate into diversity characteristics: the influence of some key indicators on the upper indexes is often non-linear [15], and the change in their values will cause sharp changes of the upper indicators. However, some indicators can compensate each other with those at the same level, and the changes of individual indicators will not directly lead to great changes of the upper indicators. Therefore, this paper takes into account the characteristics of these two types of indicators and proposes the Non-linear Index Aggregation Method (NIA) to achieve multi-level index synthesis. It applies the Weighted Sum Method (WS) to the indicators with linear additivity, the Weighted Product Method (WP) to the uncompensable indicators, and making combination of two aggregation methods organically according to the practical significance of the indicators in each level. After analysis, the specific index aggregation relationship between the layers of our index system is as follows:

1) Aggregation relationship between object layer and criteria layer

To fully reflect the attack effect of an action on the target object, it is necessary to synthesize the damage degree of various assets of the criterion layer. Therefore, the aggregation relationship between the criterion layer and the object layer is expressed as:

\[
    R_a = \sum_{i=1}^{5} I_i \cdot w_i
\]

Where, \( R_a \) is the evaluation result of attack effect, \( I_i \) is the value of index \( i \) in the criterion layer, and \( w_i \) is the weight value corresponding to this indicator.

2) Aggregation relationship between criteria layer and index layer

Taking host information asset index \( I_1 \) as an example, its corresponding underlying indexes can be divided into two categories: detected information amount \((I_{11}, I_{13}, I_{15}, I_{17}, I_{19})\) and detection accurate rate \((I_{12}, I_{14}, I_{16}, I_{18}, I_{1,10})\). Although the same kind of indexes can compensate each other, there is no
linear compensation between these two kinds of indexes, the aggregation relationship between $I_1$ and the underlying indexes is expressed as:

$$I_1 = (C_{11} + C_{13} + C_{15} + C_{17} + C_{19}) \cdot (C_{12} + C_{14} + C_{16} + C_{18} + C_{1,10})$$

Where, $C_{ij} = I_{ij} \cdot w_{ij}$, $i = 1, j = 1, 2, \cdots, 10$. $I_{ij}$ is the standardized index value, $w_{ij}$ is the index weight obtained by using the method in Section 3.3.2. Similarly, the aggregation relationship between other indicators in the criterion layer and their underlying indexes can be expressed as:

$$I_2 = C_{21} \cdot C_{22} \cdot C_{23}$$

$$I_3 = C_{31} + C_{32} + C_{33} + C_{34} + C_{35}$$

$$I_4 = C_{41} \cdot C_{42} \cdot C_{43}$$

$$I_5 = C_{51} + C_{52} + C_{53} + C_{54} + C_{55}$$

4.3 The Network Attack Effect Evaluation Algorithm from the Confrontational Perspective

On the basis of the previous content, the evaluation process of network attack effect under the confrontational perspective is given: Firstly, multiple attack tests and confrontation tests are performed with the actions to be evaluated in the network range, and the corresponding indicators are collected as the input of the evaluation algorithm according to the index system shown in Figure 2. Afterwards, the multi-level index synthesis is carried out by using the combination weighting approach and the NIA method we proposed. Finally, the evaluation result of the attack effect is obtained. The specific process is shown in Algorithm 1:

Algorithm 1. Attack effect evaluation algorithm from the confrontational perspective

Input: a group of original indicator data $x_1, x_2, \cdots, x_l$

Output: the attack effect evaluation result $R_a$

1. Standardizing the original indicator data $x_1, x_2, \cdots, x_l$ by Equations 2 and 3 to get $y_1, y_2, \cdots, y_m, y_l \in [0,1]$;
2. Using the method in Section 4.1 to determine the weight $w_3^{(2)} = \{w_{ij}|i = 1,2,\cdots,5; j = 1,2,\cdots,10\}$ of the underlying indicators relative to the criterion layer;
3. Combining the attack mechanism and the importance of various assets to the target, using the method in Section 4.1 to determine the weight $w_2^{(1)} = \{w_i|i = 1,2,\cdots,5\}$ of the criterion level indicators relative to the object layer;
4. Substituting the index value obtained in step 1 and the corresponding weight obtained in step 2 into Equations 7–10 respectively, and obtaining the evaluation values $I_1, I_2, \cdots, I_5$ of the criterion layer indexes by NIA method in Section 4.2;
5. Substituting the evaluation values obtained in step 4 and the corresponding weight obtained in step 3 into Equation 6, and obtaining the attack effect evaluation result $R_a$.

5. EXPERIMENTAL ANALYSIS AND VERIFICATION

This section verifies the practicability and effectiveness of CO-AEEM by carrying out multiple confrontation simulation experiments in various attack-defense scenarios.

5.1 Simulation Experiments of Network Confrontation

In this section, we use the typical enterprise network structure shown in Figure 3 to simulate the confrontation deduction process of multiple attack-defense behaviors in network range.
Since the experiment focuses on simulating the confrontation deduction process related to target breakthrough, it omits the attack preparation stage of collecting the target’s information through phishing, social engineering, scanning and other means in the early stage, saving unnecessary simulation time and steps. It is assumed that the attacker has Root privilege on the invading host and cannot execute multiple intrusion actions simultaneously. The vulnerability information of each device in the topology is obtained through attack preparation stage, as shown in Table 1.

**Table 1. Vulnerability information of simulation experimental network**

| Host         | Service | Vulnerability  |
|--------------|---------|----------------|
| Web Server   | SQL     | CGI            |
| Web Server   | IIS     | CVE-2017-7269  |
| Bastion Host | SMTP    | CVE-2014-8517  |
| Host1        | None    | Weak password  |
| Database Server | Oracle | CVE-2012-1675 |

Combined with the vulnerability information given in Table 1 and the attack-defense behavior database of MIT [16], the behavior space of both sides are given in Table 2.

**Table 2. Behavior space of attacker and defender**

| Seq | Attack Action    | Seq  | Defense Action           |
|-----|------------------|------|--------------------------|
| $a_1$ | SQL injection    | $d_1$ | Exception field recognition |
| $a_2$ | Code injection   | $d_2$ | Limit SYN/ICMP packets      |
| $a_3$ | Remote buffer overflow | $d_3$ | Reinforce security of IIS |
| $a_4$ | Brute force attack | $d_4$ | Renew root data         |
| $a_5$ | Oracle TNS listener | $d_5$ | Delete suspicious account |
|      |                   | $d_6$ | Patch Oracle TNS Listener |
|      |                   | $d_7$ | Access control TNS Listener |
|      |                   | $d_8$ | No action                |
Table 3. Detailed information of confrontation simulation experiments

| Attack Phase | Target Type   | Attack Test \( (a_i, \emptyset) \) | Confrontation Test \( (a_i, d_j) \) |
|--------------|---------------|-----------------------------------|-----------------------------------|
|              |               | Behavior Space of Attacker \( A(a_i, \emptyset) \) | Behavior Space of Defender \( D(a_i, \emptyset) \) |
|              |               | Behavior Space of Attacker \( A(a_i, d_j) \) | Behavior Space of Defender \( D(a_i, d_j) \) |
| 1            | Web Server    | \( \{a_1, a_2\} \)                 | \( \{a_1, a_2\} \)                 |
|              | Bastion Host  | \( \{a_5\} \)                     | \( \{a_3\} \)                     |
|              | Host1         | \( \{a_4\} \)                     | \( \{a_4\} \)                     |
| 2            | Database Server | \( \{a_5\} \)                 | \( \{a_3\} \)                     |
|              |                | \( \{d_8\} \)                     | \( \{d_7\} \)                     |

An attacker from the Internet attempts to obtain the Root privilege of the database server in intranet through a multi-step attack, thereby stealing, tampering and other malicious operations are performed on the sensitive information stored thereon. For the attacker, there are four attack behavior to choose in the first attack phase: Use the security vulnerability 1 or 2 to launch \( a_1 \) or \( a_2 \) respectively, and obtain the control of the Web server; Make use of vulnerability 3 to launch \( a_3 \) and compromise the bastion host; Exploit vulnerability 4 to launch \( a_4 \) and get control of Host 1. In the second attack phase, the device successfully controlled in the first step is regarded as a springboard, and the network connection relationship is used to launch the attack \( a_5 \) against the vulnerability 5 to obtain the Root privilege of the database server. From the perspective of the defender, there are corresponding defense mechanisms involved in the confrontation process in each attack phase, and any of \( d_1 \sim d_7 \) can effectively defend at least one attack action. Based on the above analysis, Table 3 summarizes the whole process of confrontation simulation experiment according to to and target's type, and shows the action space of both sides in each phase of attack tests and confrontation tests.

5.2 Experimental Results and Analysis
In the simulation experiment shown in Table 3, both sides select an action from their own behavior space each time to form different confrontation scenarios (among which, there are 5 \times 1 kinds of attack tests and 5 \times 7 kinds of confrontation tests), then carry out repeated experiments for each confrontation scenario. Finally, the effect evaluation results are obtained successively through the CO-AEEM. We give the effect evaluation process by taking the attack test \( (a_1, \emptyset) \) and the confrontation test \( (a_1, d_4) \) in the first attack phase as an example: After performing 6 times repeated experiments of attack test \( (a_1, \emptyset) \) and confrontation test \( (a_1, \emptyset) \) in network range, the raw index data of these two scenarios obtained by the data acquisition method of Section 3.2 are shown in Table 4 and Table 5 respectively (the original index data with the value of 0 has been omitted). Then, we input the original index data in Table 4 and Table 5 into Algorithm 1 and get the evaluation results. It should be noted that, since there are criteria layer indexes (e.g., \( I_1, I_4 \)) with all underlying index data of 0, considering the attack mechanism of the target breakthrough attack and the assignment principle of EWM in the comprehensive weighting method, all the underlying index weights corresponding to such criteria layer indexes are set to be 0. The arithmetic average of the evaluation results of multiple experiments shows that the attack effect \( E_a(a_1, \emptyset) \) of \( a_1 \) is 0.8245 when the defense measures are not deployed, however, it reduces to 0.2801 when confronts with defensive action \( d_1 \), and the attack effect of \( a_1 \) under confrontational conditions expressed as \( E_a(a_1, d_4) \). According to Equation 1, the defense effective rate is 66.03%, and \( d_4 \) has great defense efficiency against \( a_1 \).

Table 4. Raw index data acquired in the attack test \( (a_1, \emptyset) \)

| Index Seq | \( I_{21} \) | \( I_{22} \) | \( I_{23} \) | \( I_{31} \) | \( I_{32} \) | \( I_{51} \) | \( I_{52} \) |
|-----------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| 1         | 1           | 0.8         | 0.8         | 3.7%        | 2.2%        | 5.2%        | 4.1%        |
According to the evaluation process introduced in the previous content, we conduct a number of simulation experiments on each of the experimental scenarios listed in Table 3. Then we obtain the attack effect evaluation results \( \{E_a\}_{n \times m} \) of both sides' actions under confrontation conditions and pure attack conditions, and the defense effective rate \( \{\beta_{ij}\}_{n \times m} \) of all defense strategies against each type of attack. In order to distinguish the target of various attack actions, we use the tabular form to represent matrix \( \{E_a\}_{n \times m} \) and \( \{\beta_{ij}\}_{n \times m} \), as shown in Table 6 and Table 7.

### Table 6. Attack effect evaluation results \( \{E_a\}_{n \times m} \)

| Index Seq |  |  |  |  |  |  |
|-----------|---|---|---|---|---|---|
| 1         | 0.5 | 0.5 | 0.8 | 3.4% | 3.0% | 4.5% | 3.8% |
| 2         | 0   | 0   | 0   | 4.4% | 5.5% | 4.9% | 3.7% |
| 3         | 0.5 | 0   | 0.5 | 4.1% | 3.5% | 4.8% | 3.1% |
| 4         | 0.5 | 0.5 | 0.8 | 5.2% | 4.8% | 4.7% | 3.9% |
| 5         | 0   | 0   | 0   | 6.9% | 5.5% | 5.7% | 4.3% |
| 6         | 0.5 | 0   | 0.5 | 1    | 3.8% | 3.3% | 4.6% | 3.9% |

### Table 7. Defense effective rate \( \{\beta_{ij}\}_{n \times m} \) of all defense strategies

| Target | Web Server | Bastion Host | Host1 | Database Server |
|--------|------------|--------------|-------|-----------------|
|        | \( a_1 \)  | \( a_2 \)    | \( a_3 \) | \( a_4 \)        |
| \( d_1 \) | 66.03%     | 0.77%        | 3.92% | 0.08%           | 42.67% |
| \( d_2 \) | -0.02%     | 61.42%       | 0.63% | -0.03%          | 0.04%  |
| \( d_3 \) | 11.58%     | 53.27%       | 0.72% | 0.13%           | 0.01%  |
the pertinence of defensive measures, which can not only reflect the actual attack effect in the face of the attack test. Since its evaluation idea is consistent with the traditional evaluation method, so we use the effect evaluation result of the attack test $E_a(a_i, \emptyset)$ to represent the traditional method to compare with the method in this paper. By comparing the evaluation results of attack tests and confrontation tests in Table 6 and 7, we can find that:

1) Attack effect changes with or without the participation of defensive action, which indicates that defensive strategy does have impact on attack effectiveness. Taking attack $d_4$, $d_5$ as an example, Table 6 shows that the attack effect under attacker’s perspective is 0.7946 and 1.3125 when the traditional evaluation method is used, however, when using our method to evaluate the attack effect under the confrontation condition, it is found that when $a_4$ and $a_5$ fight against defensive actions $d_5$ and $d_7$, respectively, their attack effect reduce to 0.1162 and 0.2004, and the defense effective rate of $d_5$ and $d_7$ against $a_4$ and $a_5$ is as high as 85.37% and 84.73%.

2) The effectiveness of attack and defense actions is greatly affected by the pertinence of the actions of both sides. The same type of attack has different attack effect in the face of different levels of defense mechanisms, as shown in Table 6, the attack effect of $a_5$ against defensive measures $d_5$, $d_6$ and $d_7$ is $E_a(a_5, d_5) > E_a(a_5, d_6) > E_a(a_5, d_7)$; Furthermore, the same defense measure have different effects on different types of attacks, too. As shown in Table 7, $d_1$ can filter the input and data packets containing exception fields, thus has better defense effectiveness against attacks caused by such causes, such as $a_4$ and $a_5$. However, it is less effective against other types of attacks in the table.

The simulation results of various counterwork scenarios show that, the attack actions considered by traditional evaluation methods with high attack effectiveness and strong destructiveness may lose its effectiveness seriously when fighting against the effective defense action, thus failing to achieve the expected attack effect in the actual combat environment. Moreover, the effectiveness of attack and defense actions is greatly affected by the pertinence of both sides’ actions. Therefore, the traditional evaluation method cannot accurately determine the actual attack effect in confrontation situation. In contrast, our method remedies the shortcoming of the traditional method of ignoring the impact of defensive actions on attack effectiveness, so that both sides can be informed of the actual performance of each other’s strategies in confrontation situation and the weakness of their respective actions.

### 6. CONCLUSION

Aiming at the problem that the existing methods ignore the influence of defense strategy and only carry out one-sided attack effect evaluation from the perspective of the attacker, this article takes into account the pertinence influence of different defense measures on the attack effect and constructs the CO-AEEM model to carry out attack effect evaluation under confrontation condition. The main contributions of this paper are summarized as follows: 1) construct an evaluation index system oriented to all-assets-elements and put forward the index data acquisition methods, and the index system has the advantages of comprehensiveness, strong operability and favorable scalability; 2) propose an attack effect evaluation algorithm from the confrontational perspective, which can reduce the impact of subjective factors and adapt to the complex characteristics of network confrontation; 3) design the comparison and evaluation process of attack test and confrontation test according to the changes of the defense measures, which can not only reflect the actual attack effect in the face of different defense strategies, but also give the effectiveness degree of the defense measures against various kinds of attacks. Finally, the practicability and effectiveness of the proposed method and model are verified by simulation experiments in a variety of attack-defense scenarios in network range.
Compared with the traditional attack effect evaluation method, this method can help both sides make more effective and scientific decisions and security deployment under the real network confrontation environment, develop attack tools and defense mechanism with more pertinence and breakthrough, and build a network system that has both attack and defense capabilities very well.

The follow-up research intends to take into account the cost of attack and defensive strategies to further improve the effectiveness quantification of both sides’ strategies, and the quantified results are to be put into the game model to study how the attackers and defenders can dynamically select the strategies and obtain the maximum benefits in the complex and changeable multi-stage network confrontation with limited offensive and defensive resources.

Acknowledgement
Project supported by the National Natural Science Foundation of China (No. 61502528).

REFERENCES
[1] Peng, Z., Zhao, W., & Long, J. (2011). Grey synthetic clustering method for DoS attack effectiveness evaluation. In 2011 8th International Conference on Modeling Decisions for Artificial Intelligence (pp. 139-149). Springer.
[2] Hazra, S., Sattenapalli, J. S., Roy, A., & Dalui, M. (2018, December). Evaluation and detection of hardware trojan for real-time many-core systems. In 2018 8th International Symposium on Embedded Computing and System Design (ISED) (pp. 31-36). IEEE.
[3] Wang, B., & Chen, J. (2019). Research on DOS Attack Effect Evaluation Technology. In Recent Developments in Intelligent Computing, Communication and Devices (pp. 943-949). Springer, Singapore.
[4] Maciel, R., Araujo, J., Dantas, J., Melo, C., Guedes, E., & Maciel, P. (2018, April). Impact of a DDoS attack on computer systems: An approach based on an attack tree model. In 2018 Annual IEEE International Systems Conference (SysCon) (pp. 1-8). IEEE.
[5] Adarkar, O., Mane, R., & Shah, D. (2018, June). Impact of wormhole attack in wireless sensor network. International Research Journal of Engineering and Technology, 5(6), 2237-2241.
[6] Zhao, S., Zheng, K., & Zhao, J. (2012). Evaluation model of 3G network attack effectiveness based on AHP and GRA. In 9th Annual Conference of the Chinese Communication Societ (pp. 77-81).
[7] Guo, L., Wu, B. (2019). Evaluation of attack effect in Ad Hoc networks based on variable weight TOPSIS Method. In Proceedings of the International Conference on Advances in Computer Technology, Information Science and Communications (pp. 213-221). SciTePress.
[8] Yue, C., Tianliang, L., Manchun, C., & Jingying, L. (2018). Evaluation of the attack effect based on improved grey clustering model. International Journal of Digital Crime and Forensics (IJDCF), 10(1), 92-100.
[9] Hamid, T., Al-Jumeily, D., & Mustafina, J. (2018). Evaluation of the dynamic cybersecurity risk using the entropy weight method. In Technology for Smart Futures (pp. 271-287). Springer, Cham.
[10] Musman, S., Tanner, M., Temin, A., Elsaesser, E., & Loren, L. (2011). Computing the impact of cyber attacks on complex missions. In 2011 21st Annual International Symposium of the International Council on Systems Engineering (pp. 973-978). IEEE.
[11] Kott, A., Ludwig, J., & Lange, M. (2017). Assessing mission impact of cyberattacks: toward a model-driven paradigm. IEEE Security and Privacy Magazine, 15(5), 65-74.
[12] Genge, Béla, Kiss, István, & Haller, P. (2015). A system dynamics approach for assessing the impact of cyber attacks on critical infrastructures. International Journal of Critical Infrastructural Protection, 10(C), 3-17.
[13] Ferguson, B., Tall, A., & Olsen, D. (2014). National Cyber Range Overview. In 2014 IEEE Military Communications Conference. IEEE.
[14] Mansurov, A. (2016). A CTF-based approach in information security education: an extracurricular activity in teaching students at Altai State University, Russia. *Modern Applied Science, 10*(11), 159.

[15] Pipyros, K., Thraskias, C., Mitrou, L., Gritzalis, D., & Apostolopoulos, T. (2018). A new strategy for improving cyber-attacks evaluation in the context of Tallinn Manual. *Computers & Security, 74*, 371-383.

[16] Gordon, L.A., Loeb, M.P., Lucyshyn, W., Richardson, R. (2015). 2015 CSI/FBI computer crime and security survey. *In Proceedings of the 2015 IEEE Computer Security Institute* (pp. 48-64). IEEE.