Mimicry Detection Method Based on Immune Danger Theory
Min-sheng Tan, Chang Cai, Huan Zhou and Lin Ding

ABSTRACT

Considering that the traditional detector has lower detection rate, resource utilization and higher false alarm rate, this paper proposes a mimicry algorithm and a mimicry detector based on Immune Danger Theory (IDTMD). The detector adjusts the types and number of detection organizations and detection cells dynamically, according to the types and number of the detected signal and the detected dangerous signal, as to realize the dynamic, diversity and randomness of mimicry detection. The KDD-CUP99 dataset was used to test the comprehensive detection performance of the IDTMD detector, the detector based on Negative Selection Algorithm (NSA), the detector based on Dendritic Cell Algorithm (DAC). The results show that the detector improves the detection rate and the resource utilization, decreases the false alarm rate, and when testing amount reaches a certain scale, its response time is significantly less than other detectors’.

INTRODUCTION

Mimicry is an ecological adaptation phenomenon, which refers to the special behavior of a creature that imitates the feature or form of another organism and is formed into the long-term evolution process of nature to protect itself from invasion [1]. This is similar to intrusion detection system in defending against attacks and protecting network security.

In 1922, the U.S. news scholar Walter Lippmann through long-term to public opinion transmission phenomenon is observed and summarized, put forward the concept "mimetic environment" [2]. In 1968, Tanjug, a transmission of Japanese scholars, proposed a question “mimetic environment of environment” based on Walter Lippmann’s point [3], namely the realistic environment of information trend. Wu Jiangxin, a scientist in China, combined bionics with cognitive science and modern information technology, proposed new ideas [4] about safety and mimetic mimicry computation for the first-time, and successfully developed the first simulation computer with dynamic and variable structure in 2013.

Studies have found that the characteristics of the biological immune system, such as immune recognition, distribution detection of immune cells, immune adaptive regulation and immune memory, which are very similar to those required...
by the intrusion detection system. Therefore, the research on intrusion detection based on reference and simulation of biological immune mechanism has been widely concerned [5]. However, the traditional method of intrusion detection based on artificial immunity is mainly based on autogenous/non-autogenous models [6]. Generally, low detection rate and high false alarm rate make it difficult to meet the complex and changeable network security requirements. Immunological risk theory is based on the immune response based on the danger signal. The key factor to activate the immune response is the danger signal generated by the invader resulting in the death or pathological changes of the body cells. As long as the danger signal is generated, the antigen presenting cells will be activated and the immune response will be activated, without considering the problem of autologous and autologous tolerance [7, 8]. Therefore, compared with the traditional artificial immunity, the detection method based on immune risk theory can not only save the overall detection time, but also improve the utilization rate of resources.

Immune danger theory is applied to the mimicry defense, this paper proposes the mimicry detection method based on immune danger theory, and mimicry detector detection algorithm is given, and to enhance the defense system of unknown network attacks in adaptive ability, improve the detection rate of network attacked and resource utilization, reduce the rate of false positives provides a new method.

1 IMMICY DETECTOR BASED ON IMMUNE DANGER THEORY

1.1 Mimicry Detector Structure

Mimicry detector based on immune danger theory (IDTMD detector) mimics the immune system of organisms, each IDTMD detector is composed of different detection organizations, each organization is made up of several detection cells (as shown in figure 1), different detection organizations contain different numbers and types of detection cells. The amount and type of detection organizations according to the amount of signal detected and the test results for dynamic adjustment.

![Figure 1. Mimicry detector structure.](image)

1.2 Detection Organizations’ Structure

Each detection organization is the basic independent testing unit, consists of several different types of detection cells. The organization adjusts the amount and
type of detection cells according to test results dynamically to realize the dynamic and diversity of the organization. Different colors say different kinds of test cells.

1.3 Detection Cells’ Structure

Different organizations are randomly assembled from different detection cells, each cell can only detect a type of intrusion behavior, is the smallest detection unit, it must be combined with other detection cells to complete the detection task.

(1) The feature set unit: mimicry detector is based on the basis of all the dangerous signals are divided into several classes, if a detection cell is used to detect a certain type of dangerous signals, the feature set unit of the detection cell unit is placed with the types of dangerous signals’ feature set, each characteristic signal is composed of this kind of dangerous properties of corresponding binary code, the feature set appropriately updates according to the new detected danger signals. (2) The detection unit: it mainly realizes the function of detection cells. Call features concentrated code matches with the input signal, if matching success, the counting unit and memory unit will be adjusted. (3) The memory unit: whenever a new dangerous signal is detected, the memory unit adds a new signature similar to the signature corresponding to the dangerous signal in the feature set unit to record the dangerous signal which has been detected. (4) The counting unit: when the detection cell finds a dangerous signal, counting cell counter pluses one accordingly.

2 MIMICRY RANDOMIZATION AND DYNAMISM, DIVERSIFICATION OF THE DETECTORS

2.1 Mimicry Detector, Dynamism and Variety

2.1.1 DETECTION ORGANIZATION, DYNAMISM AND VARIETY

Mimicry detector can according to the detected dangerous signal to the testing organization for dynamic adjustment. If detected a dangerous signal from something to nothing or never existed, IDTMD detector will remove part of the original contains the danger signal detection cell in a fixed detection organization, keep a small percentage of the class organization, and portfolio risk frequency of 0 detection cells to produce new organization; If a certain period of time did not go to detect the danger exists, all organizations which contain this detection cell will be deleted; If found a new dangerous signal in the process of testing, testing organizations which not contain the dangerous signal detection cell will be replaced by the testing organizations which include the dangerous signal detection cell.

Each IDTMD detector consists of many types of testing organization, namely, \( O = \{ O_1, O_2, ..., O_m \} \), among them, \( m \) says types of testing organizations in the detector. The number of testing organizations of each mimicry detector satisfies
The number of detection cells in organizations is determined according to formula 4 and formula 5. If some time a certain value of formula 4 is 0, namely, the dangerous signal not exists, the detector will generate a new detection organization, which is composed of only the detection cells corresponding to the non-zero frequency of the signal. The danger signal is divided into n class, D={D_1, D_2, ..., D_n}(D_1 says a type of dangerous signals, the latter is represented by the dangerous signal 1, The representation of D_2, ..., D_n in turn is analogous), its corresponding number is R={R_1, R_2, ..., R_n}, if the frequency of a dangerous signal i is changed from non-zero to zero, and the number of other dangerous signals is non-zero, the detector generates a detection organization that not contains the detection cell i according to formula 4 and formula 5. At the same time, the detector becomes another detector because of the change of the type of the detection organizations.

2.1.2 DETECTION CELL, DYNAMISM AND VARIETY

The detection organizations dynamically generate detection cells that can detect the relative dangerous signals based on the types and frequency of the dangerous signals detected before, and produce the same proportion of detection cells according to the proportion of the dangerous signals. For instant, the same type of danger signals can be divided into one group, there is N groups, D= {D_1, D_2, ..., D_n}, among them, D is used to distinguish between dangerous signals, D_1 shows the dangerous signal 1, D_2 shows the dangerous signal 2, and so forth. Different types of attacks correspond to different risk repository, and also correspond to different detection cells, this fully reflects the diversification of detection cell.

According to the number R (R= {R_1, R_2, ..., R_n}, n is the kinds of dangerous signals, R_1 shows the rate of dangerous signal 1, R_2 shows the rate of dangerous signal 2, and so forth) of dangerous signals detected by the detection organization for every interval T, the type and quantity of the detection cells in the detection organization is determined. To a certain extent, this reflects the dynamic of testing organizations, but also reflects the diversity of testing organizations.

The number R_i (1<=i<=n) of each dangerous signal is the sum of the dangerous signals of class i detected by all the detection cells within a certain period of time. Among them, C_d says the number of a detection cell’s detected dangerous signals, m says the number of detection organizations, n says the number of class i detection cell in testing organization.

\[
\sum_{i=m}^{m} O_i \leq C_o
\]  \hspace{1cm} (1)

\[
C_o = kD_v
\]  \hspace{1cm} (2)
Adjust the number of various kinds of cells in each detection organization when meet formula 4 and formula 5. Among them, \( C_i \) says the number of detection cells \( i \) in each detection organization, \( C_c \) is detection cells’ number threshold.

\[
C_1 : C_2 : \ldots : C_n = R_1 : R_2 : \ldots : R_n
\]  

(4)

\[
C_1 + C_2 + \ldots + C_n \leq C_c
\]  

(5)

Finally, according to the \( R_1, R_2, \ldots, R_n \) to adjust the number of each testing organization’s testing cells, the number meet formula 6.

\[
C_i = \frac{e}{e^{R_1} + e^{R_2} + \ldots + e^{R_n}} \quad (1 \leq i \leq n)
\]  

(6)

### 2.2 Mimicry the Randomization of the Detector

Before IDTMT detector detects the detected signals, according to the numbers of the detected signals, the different detection cells are dynamically combined, several different kinds of detection organizations are randomly generated. Each detector has a certain amount testing organization which contain all the types detection cells, namely fixed testing organization, and the number of the testing organizations meets the formula 1. Then the detector will randomly divide signals for each testing organization for testing, which reflects the randomization and the diversity of mimicry detector.

The largest randomly generated testing organizations are: (\( C_{o\text{-max}} \) says the maximum number of organizations’ types, \( C_c \) says the number of cells’ types.)

\[
C_{o\text{-max}} = A_{Cc}^C
\]  

(7)

### 3 A MIMICRY DETECTION ALGORITHM BASED ON IMMUNE DANGER THEORY

In the mimicry detection algorithm based on immune danger theory (IDTMDA, Mimicry Algorithm Based on Immune Danger Theory), the amount and type of testing organizations is dynamically adjusting according to the detected signals and detected dangerous signals, the detected signals are randomly assigned to each detection organization. Each organization dynamically adjusts testing cells’ types and numbers according to the detected signals’ types and numbers, determines each cell’s completing test order (number of cells first for testing) according to different detection cells’ numbers, and each cell is assigned a detected signal. For example:
1000 input signals will be detected, they are divided into 10 testing organizations evenly. If a detection organization has four types of cells, detection cell 1, detection cell 2, detection cell 3, detection cell 4, in turn the number of each cell reduces, 10, 8, 6, 4 respectively, the detection organization will first arrange all the detection cells 1 to test the input signals and assign 10 input signals to each test cell 1 for detection. Each detection cell calls CCDA algorithm (Classification of Cells Detection Algorithm) for testing. By the first quarter, the feature set of each detection cell only has one type of dangerous signals, in the CCDA algorithm, every dangerous signal has appropriate two thresholds a and b (a is dangerous threshold, b is security threshold), the signals match different thresholds according to different dangerous types. By matching the input signals with the detection cells’ feature code, the affinity of input signals and dangerous signals are calculated, the signals are hierarchically stratified according to the affinity, different levels’ signals have different processing. If affinity p reaches a (p>=a), it shows p is dangerous signal, the memory unit is updated in the detection cell corresponding to affinity p.

In IDTMDA algorithm, when the detection cell 1 not detects dangerous signals, the input signals will be transmitted to the detection cell 2, which is next only to the detection cell 1. (each detection cell’s counting unit contains the assigned number of detections, detection organizations according to this quantity for each cell 2 evenly distribute detection signals). When the detection cell 2 not detects the dangerous, according to the above rules the subsequent detection cell goes on to test. If the input signal has been matched all types detection cells in the detection organization (each test cell is taken one), and each matching affinity is smaller than the security threshold b, that the input signal is the normal signal, its corresponding binary signature is added to the normal feature library of normal detection cell; Otherwise, the input signal may be normal signals can also be dangerous, the input signal shall be further matched with the feature library of the normal cell library, it is required to get the biggest affinity P’, according to the formula 8 the relative affinity V_1, V_2, …, V_n between the input signal and the feature library of the detected cell in the detection organization is given, this signal is added to the corresponding feature library with the highest relative affinity.

\[ \lambda = \frac{P'}{a} \]  

(8)

4 EXPERIMENTAL PROCESS AND RESULTS

4.1 Experimental Data Set

According to the assessment of the KDD-CUP99 dataset of Srinivas Mukkamala and Andrew H. Sung, 41 attributes record has different effects on different kinds of attacks, the greater impact properties respectively are: The attributes corresponding to normal data are: 1, 3, 5, 6, 10, 17, 23, 27, 28, 29, 33, 36, 39; The attributes
corresponding to Dos attacks data are: 1, 5, 6, 23, 24, 26, 32, 36, 38, 39; The attributes corresponding to Probing attacks data are: 3, 5, 6, 23, 24, 32, 33; The attributes corresponding to R2L attacks data are: 3, 5, 6, 32, 33; The attributes corresponding to U2R attacks data are: 5, 6, 32.

4.2 Arrangement of Experiment

Firstly the KDD-CUP99 dataset collected by MIT Lincoln Laboratory was copied to excel table, the excel table was modeled as a database for operation, the dataset’s data was read into the detector as input, secondly its 41 bit string data was encoded into binary code of 293 bits, thirdly the binary code of dangerous signals and input signals were sent to the detector, according to the Hamming Distance Matching Rule matching, different dangerous signals match different attributes corresponding to the binary number, such as DoS attacks match the 1, 5, 6, 23, 24, 26, 32, 36, 38, 39 corresponding to the binary code, Probing, R2L, U2R analogy.

4.3 Important Parameter Determination Method

Before the experiment, it is necessary to determine the main parameters of the algorithm, which are the threshold value $C_0$ of the number of organization and the threshold value $C_c$ of the number of cells in each detection organization. And the higher the value $C_0$ and $C_c$, the greater the number of cells in the detection organization and organization of the detector.

$C_0$ and $C_c$ were determined by a number of experiments. According to the formula 2, $C_0$ just is determined by K. $D_v$ is the amount of test data (in this paper, $D_v = 1000$), the effects of K and $C_c$ changes on resource consumption and detection time are obtained through experiments. And the time of detection is setting $C_c$ to a fixed value when to determine the effects on resource consumption and detection time by the values K, similarly in determining the influence of parameters $C_c$ on resource consumption and detection time set K to a fixed value. Through many experiments and the analysis shows that, the value of K in about 0.003 and value of $C_c$ in about 15 system, the resource consumptions and the detection time reach the best state, the

| K         | 0.001 | 0.002 | 0.003 | 0.004 | 0.005 |
|-----------|-------|-------|-------|-------|-------|
| count_c   | 5     | 5     | 5     | 5     | 5     |
| Resource consumption | 15 | 30 | 45 | 60 | 75 |
| Detection time      | 7581 | 4821 | 1889 | 1668 | 1469 |
TABLE II. RESOURCE CONSUMPTION AND DETECTION TIME OF DIFFERENT COUNT_C VALUES.

| K   | 0.003 | 0.003 | 0.003 | 0.003 | 0.003 |
|-----|-------|-------|-------|-------|-------|
| count_c | 5     | 10    | 15    | 20    | 25    |
| Resource consumption | 15    | 30    | 45    | 60    | 75    |
| Detection time       | 8231  | 5021  | 1889  | 1552  | 1249  |

comprehensive detection performance of the system is best. The experimental results are shown in TABLE I and TABLE II.

4.4 Experimental Results and Analysis

In order to determine the total affinity threshold $\theta$, four kinds of attacks such as smurf, ipsweep, buffer_overflow and ftp_write were selected from the data sets to conduct experiments, 1000 signals were selected from the data sets randomly, among which danger signals for 40, smurf, ipsweep, buffer_overflow, ftp_write and were 10. When the signal matches the dangerous and normal database, the dangerous threshold and security threshold are expressed in $\theta (a, b), \theta_1, \theta_2, \theta_3, \theta_4, \theta_5$.

TABLE III. COMPARISON OF FALSE ALARM RATE OF DETECTION RATE.

| Type | Num | Detection rate /% | False alarm rate /% |
|------|-----|-------------------|---------------------|
| NSA  | 1   | 65.80             | 0.75                |
|      | 2   | 64.75             | 0.81                |
|      | 3   | 65.51             | 0.85                |
|      | 4   | 66.19             | 0.67                |
|      | 5   | 64.37             | 0.90                |
| DCA  | 1   | 77.73             | 0.60                |
|      | 2   | 75.49             | 0.58                |
|      | 3   | 74.62             | 0.62                |
|      | 4   | 77.25             | 0.57                |
|      | 5   | 75.50             | 0.55                |
| IDTMD| 1   | 87.26             | 0.42                |
|      | 2   | 91.26             | 0.39                |
|      | 3   | 86.68             | 0.46                |
|      | 4   | 88.36             | 0.48                |
|      | 5   | 84.36             | 0.43                |

respectively is Dos attacks dangerous library, Probing attacks dangerous library, R2L attacks dangerous library, U2R attacks dangerous library, and normal library, suppose $\theta_1=(0.85,0.68), \theta_2=(0.83,0.66), \theta_3=(0.80,0.64), \theta_4=(0.78,0.68), \theta_5=(0.86,0.66)$. The IDTMD detector, the NSA detector, and the DCA detector were compared, the parameters of the three detectors are selected as parameters for achieving the best results. The samples of generating random are detected, and three
detector of the detection rate and the false positive rate of the comparison results are shown in TABLE III shows.

As can be seen from the results, the detection rate of the IDTMD detector is 22.26 percentage points higher than that of the NSA detector based on the traditional artificial immune, 11.466 percentage points higher than that of the DCA detector. The false alarm rate of the IDTMD detector is 0.36 percentage points lower than the NSA detector, 0.148 percentage points lower than the DCA detector.

In order to further verify the advantages of the IDTMD detector, the resource consumption, the first response time and the second response time were compared under different experimental scale. From TABLE IV it can be seen that because the NSA detector requires to do self tolerance, although its second response time is short, but the first response time is longer, and its response time scale is affected by the autologous set, grows with the increase of the autologous sets. The response time of the DCA detector is closely related to the scale of the experiment, the detection time increases with the increases of the scale of the experiment. The IDTMD detector’s first-time response is longer, with the detection scale increases the detection time will increase slowly, in small scale experiments its advantage compared to above two kinds of detector is not obvious, but when the scale increases to a certain extent, the response time of the IDTMD detector is lower than the two others’.

### TABLE IV. CONSUMPTION OF RESOURCE AND RESPONSE TIME.

| Type  | Size | Resource consumption | The first response time /ms | The second response time /ms |
|-------|------|----------------------|----------------------------|----------------------------|
| NSA   | 100  | 50                   | 3072                       | 589                        |
|       | 300  | 50                   | 5345                       | 898                        |
|       | 500  | 50                   | 7071                       | 1001                       |
|       | 800  | 50                   | 9712                       | 1827                       |
|       | 1000 | 50                   | 10521                      | 2121                       |
| DCA   | 100  | 30                   | 1572                       | 889                        |
|       | 300  | 30                   | 3645                       | 1698                       |
|       | 500  | 30                   | 8471                       | 3301                       |
|       | 800  | 30                   | 15122                      | 5227                       |
|       | 1000 | 30                   | 20520                      | 7021                       |
| IDTMD | 100  | 10                   | 6247                       | 1000                       |
|       | 300  | 15                   | 6335                       | 1101                       |
|       | 500  | 20                   | 6816                       | 1021                       |
|       | 800  | 30                   | 7013                       | 1253                       |
|       | 1000 | 45                   | 9283                       | 1709                       |

### 5 CONCLUSIONS

In view of the traditional detection methods cannot effectively detect the unknown network attacks, Immune Danger Theory and mimicry security concept
were introduced into the network security. IDTMD mimicry detector through the
dangerous signal detection and classification, without the need to establish a set of
huge antibodies, without the need for sample training, it decreases the calculation
scale, enhances the adaptive ability of defense systems to deal with the unknown
network attacks, improves the network attacks detection rate and resource utilization
rate and reduces the rate of false positives. When testing amount reaches a certain
scale, the first response time and the second time response time are significantly
reduced.

The IDTMD mimic detector enables the defender to discover as many unknown
attacks as possible, makes the defense system more effective and credible, makes it
difficult for attackers to find the defender's defense principles and methods, and
increases the difficulty and cost of the attacker's attack.

ACKNOWLEDGMENT

This work was supported by Project 61403183 of the National Science
Foundation of China, Project 2017JJ4048 of the Hunan Provincial Natural Science
Foundation of China, Project 20183350502 of the Hunan Provincial Finance Office
Science Research of China.

REFERENCES

1. Franks, Dw, Noble, J. 2002. “The origins of mimicry rings. In Artificial Life VIII: Proceedings of
the Eighth International Conference on Artificial Life,” MIT Press., pp. 186-191.
2. Wall, Pu-man Li. June 2002. “Public opinion [M],” Shanghai people's publishing house.
3. Qing-guang Guo. November 1999. “Course in communication [M],” Renmin University of
China Press.
4. Jiang-xing Wu. 2014. “Original intention and vision of pseudo computation and pseudo security
defense [J],” Telecommunications Science., 33(7):1-6.
5. Farmer J D, Packard N H, Perelson A S. 1986. “The immune system, adaptation, and machine
learning [J],” Physical D., 2(1-3): 187-204.
6. Hofmeyr S A, Forrest S. 1999. “Immunity by design: an artificial immune system [A],” In: Proc.
Of the Genetic and Evolutionary Computation Conference (GECCO'99) [C]., CA, pp. 1289-
1296.
7. Aickelin U, Gayzer S. 2002. “The danger theory and its application to artificial immune systems
[A],” In: Proc. of the 1st International Conference on Artificial Immune Systems (ICARIS-2002)
[C]., pp. 141-148.
8. Aicklein U, Bentley P, Cayzer S. et al. 2003. “Danger theory: the link between AIS and IDS
[A],” In: Proc. of the 2nd International Conference on Artificial Immune Systems [C]., pp. 147-
155.