Design of TCP SYN Flood DDoS Attack Detection Using Artificial Immune Systems

Gilang Ramadhan[#*1], Yusuf Kurniawan[#2], Chang-Soo Kim[*3]

[#]*School of Electrical Engineering and Informatics, Institut Teknologi Bandung*
Jalan Ganesha no.10, Bandung, Indonesia
[1]gilangramadhn@gmail.com
[2]ysfk2002@yahoo.com
[*]*Pukyong National University*
45 Yongso-ro, Namgu, Busan, South Korea
[3]cskim@pknu.ac.kr

Abstract— **TCP Flood DDoS attack is one of the most commonly-used attacks. DDoS has a huge impact on the victim because DDoS attack can multiply the power of attack and makes the victim server unavailable for the normal user. Recently, there are many types of research about DDoS attacks and the detection or mitigation techniques. One of the novel methods to detect DDoS is using dendritic cell algorithm (DCA). DCA is a kind of artificial immune system in the evolutionary algorithm that can be used as an anomaly detection. The DCA is also designed to solve the problem in network intrusion detection. In this paper tries to make a design of TCP Flood DDoS attack detection using artificial immune systems, especially dendritic cell algorithm.**

Keywords— **Artificial Immune System, Dendritic Cell Algorithm, DDoS, Transport Layer**

I. Introduction

Information Security has three main important aspects that are confidentiality, integrity, and availability. However, there is much threat that can interfere it. Distributed Denial of Service (DDoS) is one of the causal factors that can break the availability aspect. DDoS makes the server unavailable for the user by making a huge traffic until the server cannot handle another connection request [1].

Nowadays, there are many types of research that have been presented about DDoS attacks as well as the detection or mitigation techniques. Many methodologies have been proposed to prevent and detect DDoS attack. However, the detection and mitigation of DDoS attack is still a complicated task [1].

The artificial immune system is a computational system inspired by human immunology system. In the artificial immune system, there are many algorithms based on human immune functions, principles, and models, which are applied to problem-solving[2]. Dendritic Cell Algorithm is one of an artificial immune system based on danger theory. It proposes the detection of the immune system by the presence of danger signals. The Danger signal is released by tissues as a result of necrotic cell death (cellular damage caused by pathogenic infection or exposure to extreme condition) within host tissue [4].

The ultimate purpose of dendritic cell algorithm in the artificial immune system is to solve the problems in network intrusion detection. The dendritic cell algorithm has a promising result in reducing the high rates of false positive [4]. This paper tries to show a design of TCP SYN Flood DDoS detection using artificial immune systems, especially dendritic cell algorithm.

II. Literature Review

A. TCP SYN Flood DDoS Attack

In normal TCP three-way-handshake, the connection starts with the SYN packet sent by the client to the server. After received the SYN packet, the server reply with SYN/ACK packet to the client. Last, the client should reply with SYN packet to the server. After all of these process, the connection is established.

One of the shortcomings of TCP connection is the limitation in maintaining half-open connections in TCP. The half-open connection is the state when the server is waiting for the acknowledgment from the client in three-way handshake. Because of this shortcoming, the attacker can abuse the limitation to conduct TCP SYN flood attack. Normally, the server should reply the client with SYN/ACK packet while the server receives an SYN request. Subsequently, the server is waiting for the acknowledgment from the client. Nevertheless, the attacker generates TCP SYN request packets with spoofed source IP. In this situation, the ACK packet never arrives at the server. Because of it, the server machine with TCP module will be in LISTENING state. The SYN/ACK packets are destined for the spoofed host, the mechanism of the 3-way handshake is never completed, and connection entry remains in the connection backlog queue until time expires [5]. The illustration of TCP SYN flood attack can be seen in Figure 1

978-1-5090-5089-5/16/$31.00 ©2016 IEEE

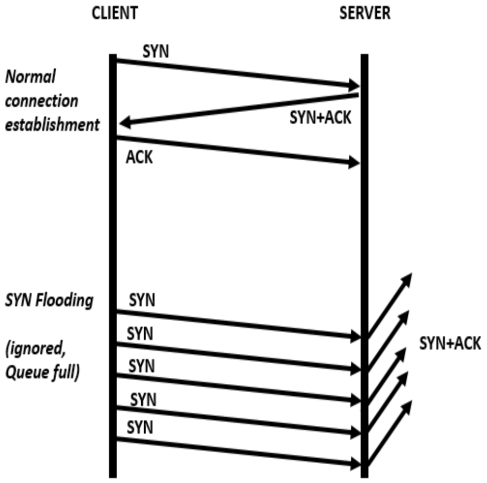

Fig. 1 TCP SYN Flood attack process

B. *Artificial Immune System*

The artificial immune system (AIS) is a computational paradigm that is inspired by immunology concepts, such as immune function, principles, and mechanism. In the initial development of AIS, most of AIS based on self-non-self-discrimination. The self-non-self-discrimination is an immunology concept to distinguish between foreign cells entering the body (non-self-element) and the body's cells [2].

Recently, there is a new concept in AIS called the danger theory that concern with the damage instead of only with the foreignness. The Danger uses alarm signals from injured tissues rather than by recognition of non-self or infectious non-self elements to determine anomaly situation [3]. In the computational point of view, the immune system is the source of inspiration as it displays learning, adaptability is self-organizing, highly distributed and displays a memory. There are many reasons why the immune system is of interest to computing [6]: recognition, feature extraction, diversity, learning, memory, distributed detection, self-regulation, metadynamics, and immune network.

C. *Dendritic Cell Algorithm*

Dendritic cell algorithm (DCA) is a kind of artificial immune systems that use danger theory concept. The DCA is a population-based algorithm that consists many individual dendritic cells as an agent. Each cell can collect and represent data items [4]

Figure 2 represents data structure in DC and its environment. Each DC in DC population store four kinds of data, which are DC input signal matrix that stores 4 type of signal, antigen, 3 type of output signals, and migration threshold. An individual DC in DC population gets the antigen and signal from antigen vector and signal matrix that are located in tissue (environment) [7].

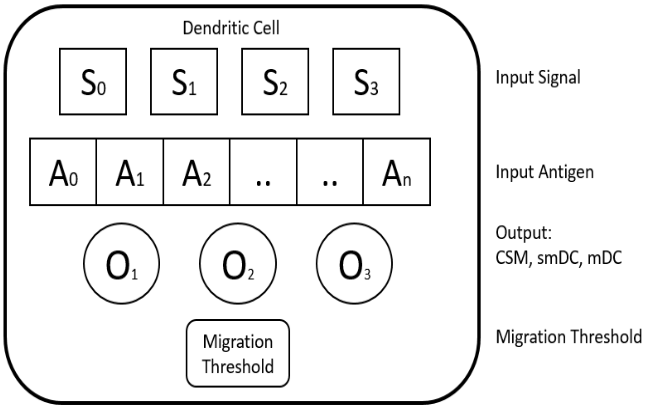

Fig. 2 Dendritic cell algorithm data structure

The DCA has two types of input data: signal and antigen [3]. Signal acts as a 'proof' that there is an anomaly on the body. There are four types of signals, i.e. PAMP, danger, safe, and inflammation. Meanwhile, the antigen acts as a 'suspect' indicating an activity which is not normal. Both of the inputs will mutually strengthen each other in determining the existence of an anomaly activity [7].

Here is the explanation of the type of signal:

a. PAMP: A signal which is a confident indicator of an abnormality
b. Danger signal: Indicator of abnormality but have a lower value of confidence than associated with the PAMP.
c. Safe signal: data which indicates normal system/data behavior.

The general dendritic cell algorithm as follow.

```
Input : signals from all categories and antigens;
(SS, PAMP, DS and antigens).
Output : antigen context values; (0/1).
for each DC do /* Preprocessing & Initialization
phase */
  initialize-DC();
end for
while CSM < mt do /* Detection phase */
  get-antigens();
  get-signals();
  calculate-inter();
  update-cumul();
end while
if smDC > mDC then /*Context Assessment phase */
  cell-context = 0;
else
  cell-context = 1;
end if
for each antigen do /* Classification phase */
  if cell-context == 1 then
    Nb-mature ++;
  end if
end
for each antigen do
  MCAV = Nb-mature / Nb-antigen ;
end
```

At last, Dendritic Cell Algorithm generates mature context antigen value (MCAV). The MCAV defines as the severity of the anomaly of antigen. The value of MCAV is between 0 to 1.

If the value is closer to 1, it means that the antigen is anomalous [3]. The MCAV is generated by each antigen. In the implementation of the DCA, there is four main phase as follow:

1) *Pre-processing and initialization phase*

The implementation of dendritic cell algorithm sometimes needs a data pre-processing phase. In dendritic cell algorithm, the mapping process is the preliminary step in the implementation. Mostly, the input data based on the mapping process has a various format. Because of it, before further processing, the input data should be refined into the appropriate format. The pre-processing stage has two main steps, the first is feature reduction and then signal categorization. Specifically, dendritic cell algorithm selects the most important and proper attributes from data. Then, the selected attributes are assigned to a specific signal category, either as a safe signal, danger signal or as PAMP signal [4].

2) *Detection Phase*

In the detection phase, dendritic cell algorithm will make a storage in the form of database that contains the combination of signals and antigens. The dendritic cell algorithm is a kind of population-based-algorithm. Therefore, based on the population concept, the same antigen will be sampled by different DCs multiple times [4]

Every time the dendritic cell sample the antigen and signal, the interim value of outputs are calculated. Eventually, the individual dendritic cell produces three cumulative output signal values known as the costimulatory molecule signal (CSM), the semi-mature signal value (smDC), and the mature signal value (mDC). To calculate output signal, the DCA applies the following weighted sum equation where C = C(CSM, smDC, mDC) [7]:

$$C = \frac{((W_{PAMP} * \sum_i PAMP_i) + (W_{SS} * \sum_i SS_i) + (W_{DS} * \sum_i DS_i))}{W_{PAMP} + W_{SS} + W_{DS}} * \frac{1+i}{2}$$

Based on the equation above, $PAMP_i$, DS_i, and SS_i are the input signal values of category PAMP, danger signal, and safe signal for all signals (i) of that category. W_{PAMP}, W_{SS}, W_{DS} represent the weights used for PAMP, SS, and DS. To calculate the output value, the output calculation is using weighting value based on Table 1. The weighting values are obtained from the concept of the human immune system [9].

TABLE I
WEIGHTING VALUES FOR SIGNAL PROCESSING

Signal	CSM	smDC	mDC
PAMPs	2	0	2
Danger Signals (DS)	1	0	1
Safe Signals (SS)	2	3	-3

3) *Context Assessment Phase*

When the dendritic cell has exceeded the migration threshold value, the dendritic cell will transform into another maturation state, from the immature state into either mature state of the semi-mature state. The transformation is based on the output value of the dendritic cell. If the output value of semi-mature dendritic cell (smDC) is higher than the output value of the mature dendritic cell, the immature dendritic cell will be transformed into the semi-mature state, and vice versa. Semi-mature state dendritic cell has context value 0 and mature state dendritic cell has context value 1. The context value is important to classify the antigen that is stored in every individual dendritic cell [4].

4) *Classification phase*

In the previous phase, the cell context of each dendritic cell has been known. Every dendritic cell stores many antigens. The number of antigen in every dendritic cell is different because each antigen has different migration threshold. An antigen can be sampled by multiple dendritic cells. In this phase, the presentation of each antigen by the individual dendritic cell is calculated. Moreover, the number of antigens that has context value one is also calculated.

The calculation of the MCAV of each antigen as follow:

$$at = \frac{an}{tn}$$

Based on the equation, *an* is the number of anomalous data items that has context cell value 1; *tn* is the total number of data items or some antigen presentation by the individual dendritic cell. *At* refers to the MCAV value of each antigen

The dendritic cell algorithm was used in many application. Paper [4] describe the main works on the DCA application domains. Related to this research, the DCA can produce a high rate of true positive and low rates of false positive in comparison with statistical techniques.

III. PROPOSED DESIGN OF DDOS DETECTION USING DENDRITIC CELL ALGORITHM

A. *System Mapping*

Dendritic cell algorithm is based on human immunology system. Because of it, to make a DDoS detection based on dendritic cell algorithm, the mapping between human immunology system and computer security must be defined as shown in Table 3.

In dendritic cell algorithm, the PAMP signal refers to the existence of the attack and the Safe signal refers to normal condition. On the mapping process, the PAMP signal and safe signal have the same parameter, that is the ratio of incoming SYN packets and outgoing SYN_ACK packet. However, it is different based on the value. If the ratio is higher than median value, it means that there is a lot of incoming SYN packet and a little SYN_ACK packet. This situation indicates that there are an SYN Flood attack. In safe signal, the ratio must be lower than the median because the amount of SYN packet and SYN_ACK almost balanced.

TABLE II
MAPPING BETWEEN HUMAN IMMUNOLOGY SYSTEMS AND COMPUTER SECURITY

No	Immune System	Computer Systems
1	PAMP signal	The ratio of incoming SYN packets and the outgoing SYN_ACK packets (when ratio is higher than median value of data input)
2	Danger signal	Number of incoming SYN packet
3	Safe signal	The ratio of incoming SYN packets and the outgoing SYN_ACK packets. (when ratio is lower than median value of data input)
4	Antigen	IP Address of attacker or user

B. *System Architecture*

The main function of the system is to detect the TCP Flood DDoS Attack and distinguish between the attacker and legitimate user. Based on Figure 4, there is two main component of the systems:

Figure 4, there is two main component of the systems:

1) *Collecting data*

In this component, data is collected. An individual dendritic cell is collecting data from the environment. The data includes signal and antigen based on the system mapping. Each antigen has a lifetime or threshold for collecting the data. If the costimulation values (CSM) exceed the threshold value, the individual dendritic cell will move to analysis data phase.

2) *Analysis data*

After all data have been collected, the antigen and signal must be analyzed to determine the anomaly of the antigen. For each antigen, the MCAV is calculated. The detection process of TCP DDoS attack using dendritic cell algorithm can be seen in Figure 5.

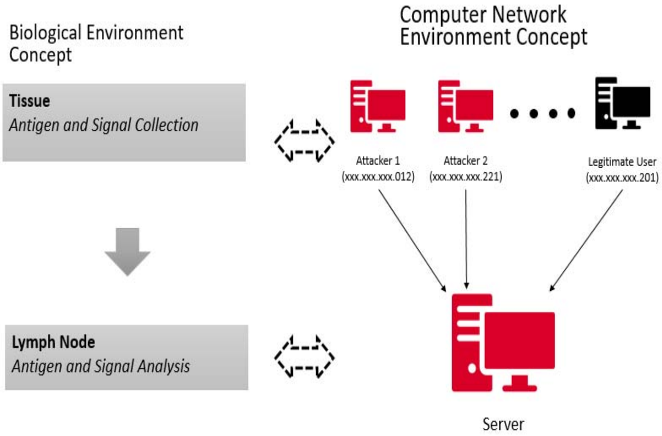
Biological Environment Concept
Computer Network Environment Concept
Tissue
Antigen and Signal Collection
Attacker 1 (xxx.xxx.xxx.012)
Attacker 2 (xxx.xxx.xxx.221)
Legitimate User (xxx.xxx.xxx.201)
Lymph Node
Antigen and Signal Analysis
Server

Fig. 3 Dendritic cell algorithm data structure

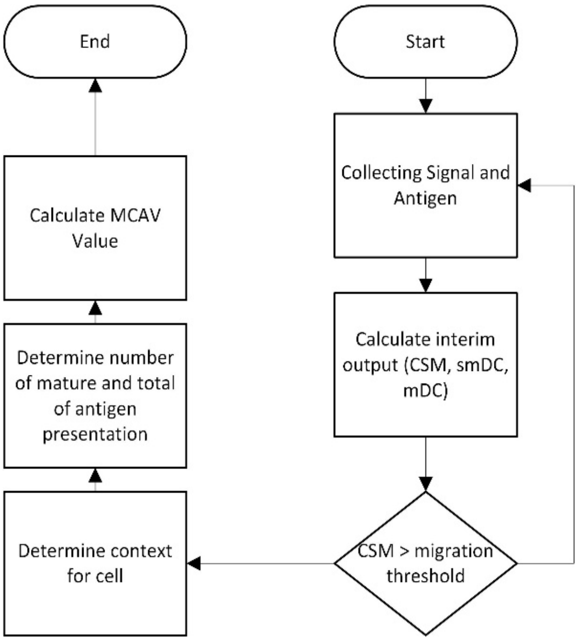
End
Start
Collecting Signal and Antigen
Calculate MCAV Value
Calculate interim output (CSM, smDC, mDC)
Determine number of mature and total of antigen presentation
Determine context for cell
CSM > migration threshold

Fig. 4 Flow Chart of System Implementation

IV. CONCLUSIONS AND FUTURE WORKS

Many types of research have been shown that dendritic cell algorithm can be used in various intrusion detection. This research tries to make a design of dendritic cell algorithm for detecting TCP flood DDoS attack.

The design of TCP flood DDoS attack using Dendritic Cell Algorithm will be implemented into a simple intrusion detection software using Python programming language and its package. Furthermore, there will be a research about the abnormal threshold of this intrusion detection based on MCAV value.

REFERENCES

[1] M.Aamir & M.Arif, “Study and Performance Evaluation on Recent DDoS Trends of Attack & Defense”. In I.J. Information Technology and Computer Science, 2013, 08, pp.54-65

[2] U. Aickelin & D. Dasgupta, "Artificial Immune Systems" in Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques, London, United Kingdom: Springer, 2014

[3] V.R. Kebande & H.S.Venter, “A Cognitive Approach for Botnet Detection Using Artificial Immune System in Cloud”In Proceeding of Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), 2014, pp.52-57

[4] Z. Chelly & Z. Eloudi, “A Survey of The Dendritic Cell Algorithm”. Springer-Verlag London 2015

[5] N.B.I. Al-Dabagh & I. Ali, “Design and Implementation of Artificial Immune System for Detecting Flooding Attacks”. In Proceeding of 2011 International Conference on High Performance Computing and Simulation (HPCS), pp.381-390

[6] Artificial Immune Systems: A New Computational Intelligence Approach by L de Castro, J Timmis, Springer Verlag, 2002

[7] Julie Greensmith, The Dendritic Cell Algorithm, 2007, University of Nottingham

[8] Silvia Anandita, “Implementation of Dendritic Cell Algorithm as an Anomaly Detection Method for Port Scanning Attack”. In Proceeding of 2015 International Conference on Information Technology Systems and Innovation (ICITSI), pp.1-6

[9] Lei Ding, Fei Yu, & Zhenghua Yang, “Survey of DCA for Abnormal Detection”, Journal of Software, Vol.8, No.8, August 2013