A REVIEW OF HUMAN IMMUNE INSPIRED ALGORITHMS FOR INTRUSION DETECTION SYSTEMS

Chukwuemeka Duru
chdur8@morgan.edu
Department of Electrical and Computer Engineering, Morgan State University, Baltimore, MD 21251 USA

Jumoke Ladeji-Osias
jumoke.ladeji-osias@morgan.edu
Department of Electrical and Computer Engineering, Morgan State University, Baltimore, MD 21251 USA

Ketchiozo Wandji
ketchiozo.wandji@morgan.edu
Department of Electrical and Computer Engineering, Morgan State University, Baltimore, MD 21251 USA

Toutsop Otily
ottou1@morgan.edu
Department of Electrical and Computer Engineering, Morgan State University, Baltimore, MD 21251 USA

Rachida Kone
rakon1@morgan.edu
Department of Electrical and Computer Engineering, Morgan State University, Baltimore, MD 21251 USA

Abstract— **Security and trust of Information Systems are critical in its design as they directly influence users' view and acceptance of such systems. Security can be said to be a contextual and dynamic term as there has not been a holistic, universal, and eternal security measure to date. Recent years have seen a lot of confidential and sensitive information being sent, received, and analyzed on the Internet, and a plethora of investigations on ways of developing comprehensive security solutions like encryptions, pattern recognition, and anomaly detection. This work reviews the human inspired algorithms that are particularly employed in pattern recognition and anomaly detection problems. The work discusses the components of the immune system that inspired the artificial Immune System (AIS) based algorithms for pattern and intrusion detection (IDS) problems. A detailed comparison is made between negative selection, clonal selection, and dendritic cell algorithms (danger theory) which are the three major AIS algorithms. AIS is ubiquitous in computer and information security because it is based on the theories developed through years of study and understanding of the human immune system by immunologist. The strengths and weaknesses of these algorithms are also discussed, and possible improvement suggested.**

Keywords –immune system, negative selection, clonal selection, dendritic cell, intrusion detection, pattern recognition.

I INTRODUCTION

A 'good' information system must address the cardinal points of information system security which includes confidentiality, integrity, and availability. Toutsop et al. [1] investigated ways of applying machine learning algorithms like decision trees, random forest, logistic regression for solving security challenges in medical and home-based Internet of Things (IoT). Recent research around information security has focused majorly on Intrusion Detection Systems (IDS) and pattern recognition problems and have classified their solution set as anomaly and signature-based algorithms [2]. Toutsop et al. [3] showed that Hackers may be able to get unauthorized access to user network and may compromise user data from different IoT devices. They came up with an excellent IDS method that combines machine learning classifiers with deep learning. The anomaly-based intrusion detection system can be likened to an unsupervised learning approach where the normal profile of the system is observed (no attack condition) and patterns outside this observed system functionality are classified as an intrusion. The signature-based IDS system can be likened to a supervised leaning process where a system is pre-trained with a wide range of normal traffic conditions or patterns and tested with abnormal traffic profile (usually obtained from databases of previous attacks provided by experts). These kinds of systems are only able to identify the abnormal traffic profile shown prior to deployment. These abnormal traffic patterns are known as attack signatures. When deployed, any traffic or pattern matching these signatures is classified as an intrusion or anomaly. AIS Algorithms are employed in solving the two problem sets (i.e., anomaly and signature-based IDS, and pattern recognition problems) described in section 1.0.

.

II THE HUMAN IMMUNE SYSTEM

Humans are composed of complex systems that constantly process information, from the central nervous systems to the human Immune system. Unlike the Central Nervous System which has a centralized information processing scheme, the Human Immune System (HIS) employs a decentralized information processing scheme.

HIS is composed of cells forming a network-like structure and interacting with each other to provide protection to the human body against a variety of viruses and bacteria. The skin, mouth and tears are part of the innate or intrinsic immune system and provide the first level of defense for humans while the phagocytes, T-cells, B-cells provide the last level of defense and are categorized under adaptive immune system [4].

Table 1.0 shows the components of the immune system. The white blood cell is the major player in the human immune system and has the lymphocyte as a key component. The B cells and T Cells are the two major types of lymphocytes. B cells are responsible for producing antibodies that are used in wading off intruders (bacteria and virus) from the body while the T cells are responsible for destroying corrupted body cell (self-matching cells). The molecules of the B cells are designed to latch (detection phase in AIS) to the molecules of an invader. T Cells produce cytokines that activates the immune system in the face of an attack. The part of the immune system activated by the cytokines in the face of an attack is the macrophages which are responsible for removing remains of the invader and dead tissues after a successful defense of the body [5].The T Cells

978-1-6654-8453-4/22/$31.00 ©2022 IEEE

and B Cells are the major components of the adaptive immune system [6]. They provide defense against invaders based on their previous encounter with such invaders.

Table 1.0 Components of the immune system [7]

Innate Immune System	Adaptive Immune System
Macrophages	T Cell CD4+ T cell CD8+ T cell
Natural Killer Cells	B cell
Basophil	Natural Killer T cells
Dendrite Cells	
Eosinophil	
Neutrophil	

Some of the most popular HIS inspired algorithms in the field of computer and network security includes Negative Selection Algorithm, the Clonal Selection Algorithm, and the Dendritic Cell or Danger Theory.

This study reviews the different techniques employed by AIS algorithms for application in intrusion detection and pattern recognition problems. It compares their strengths, weaknesses, and best domain of application. A comparison of their popularity in the last few years, and recommendations on ways of improving their performance are equally discussed.

III ARTIFICIAL IMMUNE SYSTEM

An Artificial Immune System (AIS) is used to model the complex and adaptive nature of the human immune system for application in information security domain. An AIS is employed in optimization problems, anomaly and attack detection in networks, noise detection, pattern recognition, and document clustering problems [8]. As an anomaly-based IDS system, AIS algorithm can be developed by adopting the major approaches for development of anomaly-based systems which includes:

i. Pattern identification with respect to the presented data.
ii. Applying some defined rule sets to identify potential attacks presented by suspicious network traffic.
iii. Examining series of event to detect attacks, and
iv. Use of heuristics [9].

Immunity based models are domain specific and care should be taken in understanding the domain of application prior to choosing the model [10].

The requirements of a complete and effective artificial immune system are listed below [11].

a. An AIS should have a distributed architecture for distributed analysis of data.
b. An AIS should have learning capability notwithstanding the level of computational capability of the host node.
c. The system should have real time data analysis capability implying that it should be able to go beyond analysis of offline data.
d. The system should have the ability to analyze a broad range of weakly represented invaders. This feature is comparable to the innate immune system of humans.
e. The AIS should detect the smallest change in the pattern of the analyzed invader thereby behaving like the adaptive immune system.
f. The system should have memory capabilities to enable fast detection of previously encountered anomaly.
g. An AIS should react to an intrusion or anomaly.
h. The system should have the capacity to be improved through external knowledge. This is equivalent to the application of vaccines to boost the human immune system.
i. An AIS should adopt analysis of specific environments or sets of environments to enable them produce efficient detectors in a new environment.

For an effective AIS design, there should be a defined way of representing the components of the system. The component may be represented as real-valued vectors, set of strings {0,1} produced out of finite alphabets set {ABCDE…}, integers, or symbolic representation like "color". AIS should also provide a way to evaluate the interaction of the components of the system (detectors) with the environment and with each other. This interaction is known as the affinity measure and should be carefully chosen as it directly affects the efficiency of the algorithm. Finally, there should be a defined adaptation scheme which underpins the learning process and the operation of the system [12]. From the foregoing, the design of an AIS can be grouped into three main steps, choice of suitable shape space, choice of one or more measure of affinity and selecting a suitable immune algorithm.

These steps are described below [13];

i. Choice of suitable shape-space: An AIS is an abstract representation of the adaptive human immune system, and a shape space provides a way to represent this abstraction. All the indices needed to quantify the interaction of an immune molecule with its environment and other molecules are contained in the shape of the molecule. AIS employs four major types of shape spaces viz: - Euclidean or real-valued, hamming, Integer, and symbolic. While Euclidean shape-space depicts the components of the system as real valued vectors, hamming and symbolic shape space uses a set of strings from a finite alphabet. On the other hand, integer values are used for the integer shape-space.

ii. Choice of one or more affinity measure: The components of an AIS are involved in two major types of interactions. They can interact with their environment hence, providing a good application for pattern recognition problems. For this purpose, a degree of affinity is calculated that relates the similarity of the shape of a particular immune molecule to that of its environment. Being able to identify patterns through its interaction with the environment, an AIS provide a satisfactory solution to intrusion detection problems. For instance, using hamming distance shape space we can find the Hamming distance hence the affinity between an immune molecule with X = [1001101] and an antigen with Y = [10001] to be 2. Aside from the Hamming distance, Euclidean distance given by the following equation can also be used to evaluate the affinity measure.

$$d(X,Y) = \sqrt{\sum_{i=1}^{n}(Xi - Yi)^2} \qquad (1)$$

X, Y = Two Points in Euclidean n-space
Xi,Yi = Initial points in the Euclidean Vector Space
n = dimension or size

iii. Choice of suitable immune algorithm: Immune algorithms are classically subdivided into population-based algorithm (negative selection and clonal algorithm) and network based algorithm (discrete and continuous algorithm). While population based algorithms interact with only the environment, network-based algorithm has both intra-molecular and environmental interaction.

These steps are shown in Fig 1.

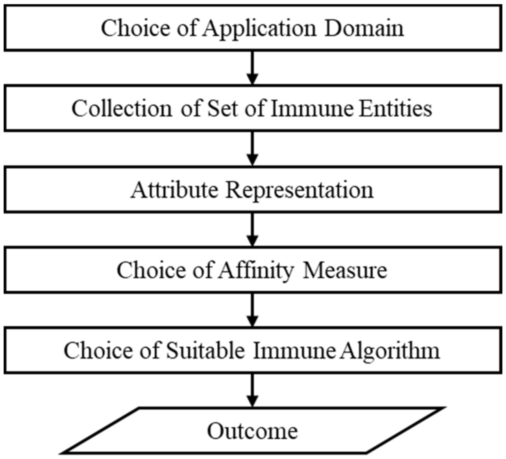
Choice of Application Domain
Collection of Set of Immune Entities
Attribute Representation
Choice of Affinity Measure
Choice of Suitable Immune Algorithm
Outcome

Fig 1: Steps involved in Artificial Immune System based computation [14]

A. Negative Selection Algorithms

T cells which are critical in preventing auto-immune systems in humans are the foundations of the Negative Selection NS algorithms.

The main function of the immune cells is to respond to invaders but sometimes, lymphocytes can erroneously detect the body immune cells (self), if this continues, this can be disastrous to the host organism, resulting in a condition known as autoimmunity. This situation is possible because antibodies produced by B Cells are composed of dissimilar gene segments each having a random composition and going through somatic hypermutation. To avoid this, the body is protected against immune cells that have high affinity to self through the process of negative selection. Negative selection of immature T Cells called thymocytes are performed in the thymus in a protected and controlled environment. The Thymus is protected from foreign interference by the blood-thymic barrier that is impermeable by macromolecules. These thymocytes are presented with a self-peptide by Antigen Presenting Cells. The thymocyte that binds strongly to any of the self-peptide are eliminated through a process known as apoptosis (controlled cell destruction) leaving only those that are non-reactive to self-cells but with high affinity to nonself cells [12]. This is the foundation of the research carried out by Forrest et al on "Self-Nonself Discrimination in a Computer" [15].

Figure 2 describes the change detection algorithm proposed by forest et al. In their work, indication of information change like user log details, systems call, or network traffic are converted into a set of strings (S) referred to as Self. A randomly generated set of detectors are applied to S. Detectors that detect self are rejected while those that failed to match with self are forwarded and added to the set of detectors to serve as a matching set for incoming unknown traffic (Nonself). A match of an incoming traffic with the detector set R, is classified as an anomaly depending on a chosen threshold. Once there is a match, a human operator may be notified to check the validity of the anomaly alert. The set of detectors R that matched the nonself string are then promoted to memory detector with reduced detection threshold and infinite lifespan if the anomaly detection was a true positive.

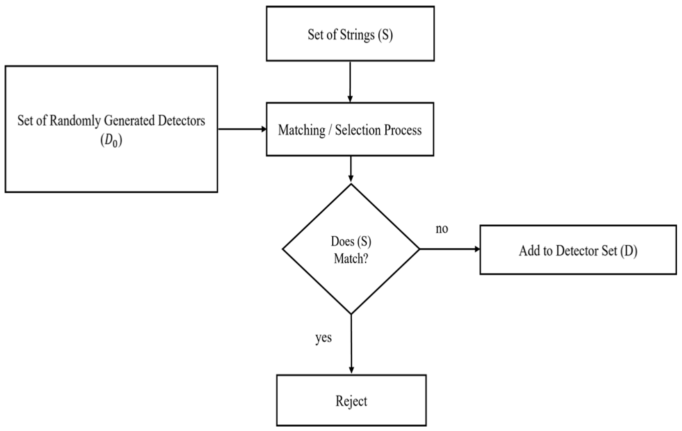
Set of Strings (S)
Set of Randomly Generated Detectors (D_0)
Matching / Selection Process
Does (S) Match?
no
Add to Detector Set (D)
yes
Reject

Fig 2: A description of the Forest et al change detection model [15]

Negative Selection Algorithm (A1): Detector Generation – Exhaustive Approach

Input: *ls*, *rt*, *Tr* ϵ *N* where 1 ≤ *rt* ≤ *ls* and S ⊂ U; *ls* = string length, *rt* =
matching threshold, *Tr* = repertoire size, *Do* = Set of Detectors , S = set of self strings
// Data here is represented as binary thereby making Hamming distance a good affinity measure candidate binary
Output: Set *Do* ⊂ U detectors generated using *r*-contiguous bits (*rcb*) matching rule

```
begin
    D: = φ
    while |Do| < Tr do
        Generate random bit string D ϵ U
        if D does not match any string in S then
            Do : = D ∪ {Do}
end
```

Algorithm A1 above is the bedrock of the work of Forest et al.

B. Clonal Selection Algorithm

The Clonal Selection Algorithm (CSA) was inspired by the Darwin's theory of evolution which includes diversity, genetic variation, and natural selection [16]. They are based on the B cells of the HIS. The CSA employs specificity, proliferation, and variation strategies for information systems tasks like optimization and pattern recognition problems [17]. In [18] and [19], the authors described the process of CSA to include initialization of a set of antibody population, selection, cloning, hypermutation, and receptor editing. In clonal selection when an immune cell detects an antigen and depending on its frequency of antigen detection (affinity measure), that cell can replicate itself (asexually). A common practice is to generate these antibodies in a random manner. It is important to note that if the population of antibodies it too large, then there would be lot of computational requirements. On the other hand, a small sample of antibodies makes the algorithm to have a premature convergence. The CSA algorithm is described below [20].

a. Initialization: Randomly initialize a population of individuals {N}.
b. Evaluation: Given a set of patterns to be recognized {P}, the affinity or match for each pattern with respect to each element of {N} is determined.
c. Selection and Cloning: A set {n} which is a subset of {N} is generated based on the level of affinity with the presented patterns as described in step b. The elements of {N} with the highest affinity is selected and becomes members of {n}. The individual element of {n} is then regenerated according to their affinity level. The rate of regeneration is directly proportional to their individual affinity to the presented pattern {P}.
d. Hypermutation: This is the mutation stage. Each element of (n) is mutated at rate that is directly proportional to its affinity with the input pattern.
e. Receptor editing: The mutated set is then re-added to the population and a specific amount is reselected based on predetermined number (d) and used as memory set.
f. Repeat steps b–e until a termination criterion is met.

C. Danger Theory Algorithm

Danger Theory (DT) is based on the Dendritic Cell Algorithm (DCA) that mimics the behavior of the human dendritic cell. The dendritic cells are professional antigen presenting cells that reside in peripheral tissues like the skin and serve as monitors for the immune system. Though they are part of the innate immune system, they serve as a link between the innate immune systems and the adaptive immune system. Through them, the immune system is notified of the presence of an antigen thereby enabling timely activation of the antigen's specific T cells (T cells are part of the adaptive immune system). DCA also increases self-tolerance (which is quite low if left to Negative Selection process alone). By notifying and activating the T cells only when an antigen is present, false positives and autoimmunity are significantly reduced.

The proportional relationship between Pathogen-Associated Molecular Patterns (PAMP), Danger Signal (DS) and Safe Signals (SS) determines to a large extent the state of the Dendritic Cell. The Dendritic cell can be in the following states [21]:

a. Immature DC: Immature Dendritic Cells (iDC) is the natural state of DC and are found in tissues. In this state, DCs collects PAMP, DS and SS signals. Depending on the collected signal with the highest proportion, the DC can either become mature or semi-mature.
b. Mature DC: When a larger number of PAMP and DS are presented to iDC, the iDC becomes mature (mDC). Maturity of the DC causes its migration from the tissue to the lymph node in other to present the antigen and activate the immune system.
c. Semi-mature DC (smDC): When the number of SS signals presented to the immune system is larger than

PAMP and DS signal, the DC remains in the tissues and immune stimulation is not activated at this point.

Adapting the behavior of the dendritic cells for information system security involves dividing the algorithm into four stages that includes [22];

a. A preprocessing stage that classifies the antigen (input data) based on a comparative evaluation of the output signals from PAMP, DS and SS. A detection stage that evaluates the concentration of the co-stimulation (C_{csm}), Semi-mature (C_{smDC}) and mature (C_{mDC}) cytokines with respect to a predetermined output cytokine equation for all the input data with respect to a particular DC.

b. Detection phase involves calculating the concentration of co-stimulation (C_{csm}), Semi-mature (C_{smDC}) and mature (C_{mDC}) cytokines based on an output cytokine equation for every antigen in each DC thus.

$$C_{[csm,smDC,mDC]} = \frac{((Wp * Cp) + (Ws * Cs) + (WD * CD))}{Wp + Ws + WD} \qquad (2)$$

Where W_P, W_S W_D are the weights of PAMP signals (C_p), safe signals (C_s) and danger signals (C_D) respectively taken from Table 1 in [23] while the output denotes the cytokine concentration.

c. The concentration of these three variables determines whether to classify an input data as either normal or an anomaly. Based on this, if C_{csm} exceeds a migration threshold (M) (the point where the DC leaves the tissue) then it moves to a decision stage that is based on the total concentrations of C_{smDC} and C_{mDC}. Antigens carried by a particular DC is assigned a value of 0 if the semi-mature concentration of cytokines for the entire antigens exceed mature concentration else, they get assigned a value of 1. This is the context assessment phase that assesses the context of the migrating DC.

d. The final stage is the evaluation of all the antigens processed by the migrated DCs and analyzing the total number of times each antigen was assigned a value of 1 using Mature Context Antigen Value (MCAV). This stage involves the sampling of data combined with context information that was received during the antigen collection process. Semi-mature and mature antigen contexts are the two major antigen contexts. While the former represents antigen data collected under normal (no attack) conditions, mature antigen context denotes anomaly. The ratio of this contexts is represented by the MCAV. An accurately functioning DCA employs the MCAV value to determine the probability (threshold) of the input data being anomalous (closer to 1) [24].

Dendritic Cell Algorithm (A2) [20]**:**

Input: *the dataset D, the Dendritic Cell (DC) pool size n, sampling ratio s,*
migration threshold 0, anomaly-threshold th
Output*: normal or anomalous for data items*
Initialise immature DC pool P; with n DC cells; Initialize migration DC pool Pm with unlimited size.
// The string representation for this algorithm is binary
Output: Set $D \subset U$ detectors generated using *r*-contiguous bits (*rcb*) matching rule

```
begin
  $D := \varphi$
  while $|D| < T$ do
    Generate randomly bit string $d \in U$
    if $d$ does not match any string in S then
      $D := D \cup \{d\}$
end
```

Feng et al. [25] employed statistical learning of PAMP, DS and SS to classify available data as either normal or anomalous. In information systems, PAMP signals can be likened to the number of connection error message generated per unit time in a particular system due to presence of a jammer as an indication of intrusion that eventually elevates the DC to maturity. Danger Signal denotes an abnormal behavior in the network but with lower confidence than PAMP. Average latency of the network can also be likened to DS signals in the DCA algorithm. Finally, SS signals indicates normal behavior of the network. The adherence to network allocation vector in Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) by participating nodes in a wireless network is a good analogy of SS signal [26].

In Danger Theory Algorithm, the main stimulator of the immune system is the presence of danger signal and not the presence of nonself (foreigners or antigens). The most daunting task is quantifying what constitutes danger with respect to normal operations of the information system. This critical component of the algorithm can be chosen based on the problem context, human experience, environmental conditions, or through weighting danger signal levels of different danger indicators in the system [27].

Aickelin et al. [28], presented one of the pioneer research projects on danger theory. Their research was on the two major kinds of cell deaths, necrotic or bad and apoptotic or good cell death with respect to the dendritic cells thus applying this same concept to intrusion detection systems. In terms of the human immune system, apoptosis exhibits a suppressive effect while necrosis constitutes immune simulation. A correlation of these two effects can be used to develop an attack scenario hence, serving as a measure of danger signal [28]. Different low-level alerts from sensors can be classified into either necrotic or apoptotic alerts which help in the correlation problem. Apoptotic alert in computer systems can be legitimate ping requests since ping can also be used by an attacker in the initial attack phase to identify the availability of the target system. A typical instance of a necrotic alert can be an actual denial of service attack like jamming attack [26].

IV STRENGTHS AND WEAKNESSES

The popularity of the three algorithms discussed in this paper are shown in figure 3 representing the effectiveness and applicability of the respective algorithms in the domain of interest.

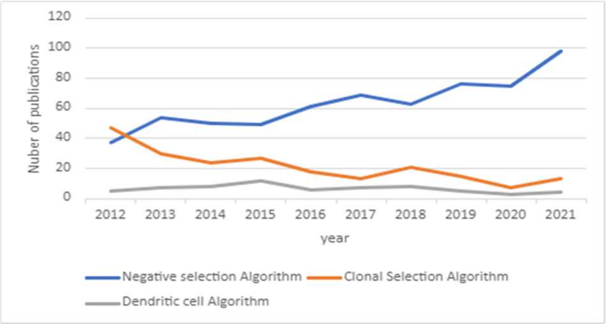

Fig 3: Popularity of NS, CS, and DCA algorithms

A. Self and Nonself Algorithm (Negative Selection Algorithm)

This model has computational constraints and may be inadequate for real life network traffic problems [28]. This is because, as the set of self and nonself data increases, it becomes computationally expensive to match the whole feature space. Moreover, the self and nonself universe is always changing in real life application and in most cases only a subset of nonself is harmful to the network while some self might even cause damage (internal attack). Labelling self and nonself data constitutes a challenge for this algorithm resulting to incorrect label application [29]. The bedrock of the negative selection algorithm is the choice of detector. A problem that should be kept in check when designing this algorithm is "Immunological Hole" [14]. These holes are set of nonself representations that has no matching antibody hence, the immune system is unable to detect it. In information security, this can be patterns or intrusions with no matching detection set. The higher the number of these holes the lower the performance of the model. For instance, taking r-continuous symbols into consideration and assuming a self-string contains CDA and C'DA', where C, C', D, A, and A' are substrings and D contains (r -1) symbols. This implies that there will be no matching detectors for CDA' and C'DA. It is important to perform critical analysis on the presence of holes in the detection algorithm chosen and finding ways to close this gap.

Also, since the length of the strings and the recognition threshold are directly proportional to the reliability of the model, choosing a low threshold to reduce the computational strain is not a good solution.

Recent approaches have been developed to mitigate this challenge and is presented in the table 3:

Table 3 presents the time complexity of different detector generating algorithm. Where:

Z_s: number of self-data
Z_R: number of mature detectors
t: matching threshold
m: alphabet size (for binary representation, b = 2)
l: string length.

From table 3, exponential time complexity is exhibited by the exhaustive search and NSmutation algorithms while the other algorithms exhibit linear time complexity. It is important to keep $r < l$ because as $r \rightarrow l$, the algorithms with linear time complexity would tend towards exponential time complexity. This is due to the exponential nature of m^r in their computation as shown in table 3. In terms of Space complexity, NSMutation shows the maximum space complexity while binary template has the most minimal space complexity.

NSelection algorithm does not perform well as the sets of detectors, self and nonself data increases when evaluated with respect to time and space complexity. This shortcoming makes it inefficient in sensor powered networks that are known to exhibit low computational resources [30].

Table 3: Time and Space Complexity of Different Detector Selection Algorithms [31].

NS Algorithm	Time Complexity	Space Complexity
Exhaustive Approach (Forrest et al.) [15]	$O\ (b^l.Z_s)$	$O(l.Z_s)$
Dynamic Programming (D'haeseleer et al., 1996) [32]	$O((l - t + 1\ .\ Z_sB^r) + O((l - t + 1)\ .\ b^t) + O(l\ .\ Z_R)$	$O((l - t + 1)^2\ .\ b^r))$
Greedy Algorithm (D'haeseleer et al., 1996) [32]	$O\ ((l - t + 1)\ .\ Z_sB^r) + O((l - t +1)\ .\ b^r\ .\ Z_R)$	$O\ ((l - t + 1)^2\ .\ b^t))$
Binary Template (Wierzchon, 2000) [33]	$O\ (b^r\ .\ Z_s) + O((l - t + 1)\ .\ b^r \cdot Z_R)$	$O\ ((l - t + 1) \cdot b^r) + O(Z_R)$
NSMutation (Ayara et al., 2002) [31]	$O\ (b^l \cdot N_s) + O\ (Z_R \cdot m^r) + O\ (Z_R)$	$O\ (l\ (Z_s + Z_R))$

B. Dendritic Cell Algorithm or Danger Theory

The danger theory based dendritic cell algorithm is a supervised learning algorithm with slow training process hence, the issue of overfitting might constitute a major setback in its implementation. Also, the threshold for the proportionality comparison of the three DCA algorithm signals (PAMP, DS and SS) to determine when to mature a dendritic cell vis-à-vis activating immune response is not always an easy decision to make. There is also no clear separation between normal and anomaly output values with respect to context assessment as this is left to the discretion of the designer [34]. A small distinction between what defines a mature and semi-mature context (with respect to the signal received) implies that a little disturbance in the system would alter the output of the algorithm constituting an adverse effect on the accuracy of the classification problem. A highly dynamic data (real time systems) poses a problem for this model. Some researchers have combined this algorithm with clustering algorithms to improve its accuracy [21]. In Danger Theory Algorithm, the main stimulator of the immune system is the presence of danger signal and not the presence of nonself (foreigners or antigens). As noted, the most daunting task here is quantifying what constitutes danger with respect to normal operations of the information system. This critical component of the algorithm can be chosen based on the problem context, human experience, environmental conditions, or through weighting the danger signal levels of different danger indicators in the system [27]. For this purpose, the concept of Alert Correlation (AC) that takes a set of low-level alerts generated by a distributed network intrusion system as input and produces an output of attempted or successful intrusion. AC can be employed to weigh the signals received to determine if it constitutes danger or not. AC should be a critical part of the danger theory algorithm for efficient and effective operation.

C. *Clonal Selection Algorithm*

In the CLONALG algorithm for instance, the selection process is done by randomly selecting a predetermined (n) number of elements with best affinity level from a population of {N} elements. Though this guarantees faster convergence of the algorithm, but an eliminated antibody that is assumed to have low affinity level takes some attributes with it thereby reducing diversity of the algorithm [35].

V CONCLUSION

The AIS algorithms described in this work presents good candidates for solving several information security system problems ranging from intrusion detection systems, pattern recognition, and optimization problems. Since no single instance of the AIS algorithm is sufficient to provide an efficient or effective solution to a particular or all problem domains and because most of the algorithms described are domain specific, care should be taken to identify a ‘suitable’ algorithm to employ for a particular problem set. A multi-agent system is also advised when employing AIS algorithms, where each agent can decide to adopt the negative selection, clonal selection, or dendritic cell algorithm (danger theory). For instance, the high false positive rate of the negative selection algorithm can be reduced by combining this algorithm with the dendritic cell algorithm described in this work.

REFRENCES

[1] O. Toutsop, S. Das and K. Kornegay, "Exploring The Security Issues in Home-Based IoT Devices Through Denial of Service Attacks," in *IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Internet of People and Smart City Innovation*, Atlanta, USA, 2021.

[2] A. Khraisat, I. Gondal, P. Vamplew and J. Kamruzzaman, "Survey of intrusion detection systems: techniques, datasets and challenges," *Cybersecurity,* vol. 2, no. 20, 2019.

[3] O. Toutsop, P. Harvey and K. Kornegay, "Monitoring and Detection Time Optimization of Man in the Middle Attacks using Machine Learning," in *IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, Washington DC, 2020.

[4] H. Rathore, M. Guizani and A. Mohamed, "Mathematical Evaluation of Human Immune Systems For Securing Software Defined Networks," in *2018 6th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, Marrakesh, Morocco, 2018.

[5] L. C. Brody, "Genome.gov," National Human Genome Institute, [Online]. Available: https://www.genome.gov/genetics-glossary/Lymphocyte. [Accessed 21 2 2021].

[6] Cancer Treatment Center of America, "B-Cells vs T-Cells: Learn the Difference &amp; Types of T-Cells | CTCA," Cancer Treatment Center of America, 10 January 2022. [Online]. Available: https://www.cancercenter.com/what-are-b-cells-vs-t-cells. [Accessed 25 March 2022].

[7] A. Trivedi, A. Shrivastava, A. Saxena and M. Manor, "Survey Analysis on Immunological Approach to Intrusion Detection," in *2018 International Conference on Advanced Computation and Telecommunication (ICACAT*, Bhopal, India, 2018.

[8] M. Günay, Z. Orman, T. Ensari, S. Oukid and N. Ben, "Diagnosis of Lung Cancer Using Artificial Immune System," Istanbul, Turkey, 2019.

[9] A. Gorbenko and V. Popov, "Abnormal Behavioral Pattern Detection in Closed-Loop Robotic Systems for Zero-Day Deceptive Threats," in *2020 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, Sochi, Rusia, 2020.

[10] J. Timmis, T. Knight, L. d. Castro and E. Hart, "An Overview of Artificial Immune Systems," in *Computation in Cells and Tissues*, Berlin, Heidelberg, Springer, 2004, pp. 51-91.

[11] P. Parrend, F. Guigou, J. Navarro, A. Deruyver and P. Collet, "For a refoundation of Artificial Immune System research: AIS is a Design Pattern," in *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, Bangalore, India, 2018.

[12] J. Timmis, A. Hone, T. Stibor and E. Clark, "Theoretical advances in artificial immune systems," *Theoretical Computer Science,* vol. 403, no. 1, p. 11–32, 2008.

[13] L. N. de Castro, "Immune, swarm, and evolutionary algorithms. Part I: basic models," in *Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02.*, Singapore, 2002.

[14] D. Dasgupta and F. Nino, Immunological Computation Theory and Application, New York: Auerbach Publications, 2008.

[15] S. Forrest, A. Perelson, L. Allen and R. Cherukuri, "Self-nonself discrimination in a computer," in *Proceedings of 1994 IEEE Computer Society Symposium on Research in Security and Privacy*, Oakland, CA, USA, 1994.

[16] G. Cziko, Without Miracles: Universal Selection Theory and the Second Darwinian Revolution, Cambridge, MA, USA: MIT Press, 1995.

[17] W. Luo and X. Lin, "Recent advances in clonal selection algorithms and applications," in *IEEE Symposium Series on Computational Intelligence (SSCI)*, Honolulu, HI, USA, 2018.

[18] B. H. Ulutas and S. Kulturel-Konak, "A review of clonal selection algorithm and its applications," *Springer Nature,* vol. 36, no. 2, pp. 117-138, 2011.

[19] J. Brownlee, "Clonal selection algorithms," CIS Technical Report, Victoria, Australia, 2007.

[20] L. N. de Castro and F. J. Von Zuben, "Learning and optimization using the clonal selection principle," *IEEE Transactions on Evolutionary Computation,* vol. 6, no. 3, pp. 239-251, 2002.

[21] N. Elisa, L. Yang, Y. Qu and F. Chao, "A Revised Dendritic Cell Algorithm Using K-Means Clustering," in *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, Exeter, UK, 2018.

[22] B. Rizvi, A. Belatreche and A. Bouridane, "A Dendritic Cell Immune System Inspired Approach for Stock Market Manipulation Detection," in *IEEE Congress on Evolutionary Computation (CEC)*, Wellington, New Zealand, 2019.

[23] J. Greensmith, "Dendritic cell algorithm", Ph.D. dissertation, Nottingham, England: University of Nottingham, 2007.

[24] J. Greensmith , U. Aickelin and J. Twycross, "Articulation and Clarification of the Dendritic Cell Algorithm," in *5th International Conference, ICARIS*, Berlin, Heidelberg, 2006.

[25] F. Gu, J. Greensmith and U. Aickelin, "Further Exploration of the Dendritic Cell Algorithm: Antigen Multiplier and Time Windows," in *International Conference on Artificial Immune Systems*, Berlin, Heidelberg, 2008.

[26] C. Duru, A. Aniedu and T. Onyeyili, "Modeling of Wireless Sensor Networks Jamming Attack Strategies," *American Scientific Research Journal for Engineering, Technology, and Sciences,* vol. 67, no. 1, 2020.

[27] W. Said and A. Mostafa, "Towards a Hybrid Immune Algorithm Based on Danger Theory for Database Security," *IEEE Access,* vol. 8, pp. 145332-145362, 2020.

[28] U. Uwe Aickelin, P. Bentley, S. Cayzer, J. Kim and J. McLeod, "Danger Theory: The Link between AIS and IDS?," in *International Conference on Artificial Immune Systems*, Edinburgh, U.K, 2003.

[29] J. Kim, Integrating Artificial Immune Algorithms for Intrusion Detection, PhD Thesis, London: University College London, 2002.

[30] C. Duru, A. Azubogu and A. Aniedu, "Review of embedded systems security," *Journal of Engineering and Applied Sciences,* vol. 17, no. 2, pp. 196-206, December 2020.

[31] M. Ayara, J. Timmis, L. de Castro and R. Duncan, "Negative Selection: How to Generate Detectors," in *1st*, Canterbury, Kent, UK, 2002.

[32] P. D'haeseleer, S. Forrest and P. Helman, "An immunological approach to change detection: algorithms, analysis and implications," in *Proceedings 1996 IEEE Symposium on Security and Privacy*, Oakland, CA, USA, 1996.

[33] S. T. Wierzchoń, "Generating Optimal Repertoire of Antibody Strings in an Artificial Immune System," in *Intelligent Information Systems. Advances in Soft Computing*, Heidelberg, 2000.

[34] Z. Chelly, A. Smiti and Z. Elouedi, "COID-FDCM: The Fuzzy Maintained Dendritic Cell Classification Method," in *International Conference on Artificial Intelligence and Soft Computing*, Zakopane, Poland, 2012.

[35] E. D. ULKER, "An improved clonal selection algorithm using a tournament selection operator and its application to microstrip coupler design," *Turkish Journal of Electrical Engineering & Computer Sciences,* vol. 25, p. 1751 – 1761, 2017.

[36] S. Alhasan, G. Abdul-Salaam, L. Bayor and K. Olive, "Intrusion Detection System Based on Artificial Immune System: A Review," in *2021 International Conference on Cyber Security and Internet of Things (ICSIoT)*, France, 2021.