IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 18, 2018, accepted August 17, 2018, date of publication August 28, 2018, date of current version September 21, 2018.

*Digital Object Identifier 10.1109/ACCESS.2018.2866962*

# A Multi-Level Intrusion Detection System for Wireless Sensor Networks Based on Immune Theory

**VISHWA TEJA ALAPARTHY AND SALVATORE DOMENIC MORGERA, (Life Fellow, IEEE)**
Department of Electrical Engineering, University of South Florida, Tampa, FL 33613, USA
Corresponding author: Vishwa Teja Alaparthy (vishwateja@mail.usf.edu)

**ABSTRACT** The human body has been, and will continue to be, a source of inspiration for researchers across various disciplines owing to its robustness and myriad of functions. While some of these advancements include the attempt to replicate the entire body to create an artificial self, some tend to use a few characteristics and theories and build upon an artificial subsystem. In this paper, an effort is made to secure a wireless sensor network (WSN) using an immune theory technique called Danger Theory. In other words, a multi-level intrusion detection system (IDS) is designed based on the functions of various immune cells. This is realized by monitoring WSN parameters, such as energy, volume of data and frequency of data transfer and developing an output based on their weights and concentrations which is a suitable basis for IDS design in WSNs.

**INDEX TERMS** Wireless sensor network, intrusion detection, security, Danger theory, artificial immune system, Internet of Things.

## I. INTRODUCTION

Wireless sensor networks (WSN) are witnessing a rapid growth both in the volume of their usage and research conducted. This, coupled with their necessity in mission critical applications, makes data security a significant area of concern. Due to their limited resources and the harsh environments in which they are sometimes deployed, it is a challenging task to secure them from attack.

Security to be provided to WSNs falls into different levels. While the first level deals with evading intrusions, the second one deals with detecting an intrusion. The third level provides an Intrusion response which can be of varied approaches. Cryptography and firewalls are the most preferred forms of securing a WSN from an attack. Intrusion detection is obligatory when an adversary manages to penetrate the firewall and causes complications with the privacy, confidentiality, and authenticity of the information in the network.

Several IDS have been designed based on biological models [1], [2] and Artificial Immune Systems (AIS) [3] is one such model. AIS draws inspiration from the immune system and the theories associated with it. Although there are quite a few IDS methods designed for computer networks and MANETs based on the immune theory, WSN's are not prominently featured in that research. This is due to the fact that these Immune inspired techniques are mostly centralized as an IDS and they consume more energy and memory resources. So, we used a distributed and a light weight approach which distributes the tasks between different nodes and also decreases the energy overhead and the packet overhead to a considerable extent. Features such as size, power, density, and the scale of deployment make WSN's appear to be more similar to the cells in the human body which communicate among themselves to form an extremely large network. Hence, taking inspiration from the white blood cells which act as the immune cells and relate them to WSNs to create an artificial Human Immune System (HIS) is a practical idea.

HIS-based IDS designs are mostly centered around two theories, Negative Selection (NSA) [4] and Danger Theory [5]. While Negative Selection deals with identifying self and non-self-entities, Danger Theory revolves around danger signals which are emitted by the Dendritic cells when an intrusion or an anomaly is detected. Clonal Selection [6] is also used to devise an Artificial immune system; however, Clonal Selection is primarily used in conjunction with Negative Selection. Positive Selection [7] is another theory which also identifies the self and non-self entities; however, the censoring is done from a randomly generated data set by eliminating the non-self entities as opposed to negative selection, which bases itself on recognizing the self entities

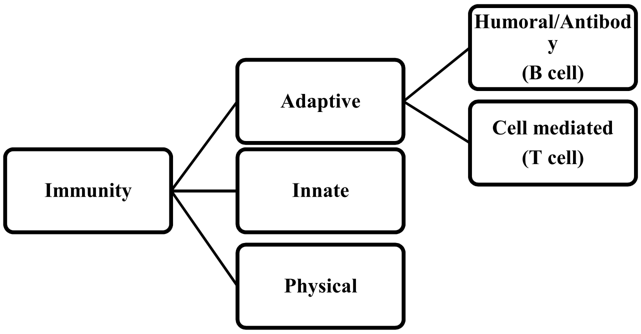


**FIGURE 1.** Classification of the immune system.

and editing them out. The remainder of the data set is stored as a detector set. NSA is bound to produce a large number of detector sets and is generally not suitable for dense environments. Danger Theory does not necessarily generate a large dataset and has fewer false positives than the NSA and Clonal Selection. Hosseinpour *et al.* [8] introduced a detection mechanism based on innate immune properties using an unsupervised machine learning approach. In [9], danger theory is used to build an artificial immune system to help detect intrusions. As mentioned earlier, these approaches require a lot more computations and memory, leading to a much higher energy consumption.

This paper tries to replicate segments of the HIS in a WSN and then subsequently attempts to derive an IDS scheme. More definitively, this IDS design has three layers. To begin with, the whole network is treated as a HIS and parallels are drawn between the nodes and the immune cells and certain tasks are assigned to the nodes. Second, an innate response which is analogous to signature analysis is performed based on the strings obtained. Third, Danger theory is applied as an adaptive immune response which is similar to anomaly-based techniques. These techniques will be studied in the sequel. As mentioned earlier, the concentrations of some features, such as energy dissipated, number of packets sent and received, frequency of data transfer are matched with their respective weights and a variable called aggregator output is generated. In several key aspects, the work presented here differs from that found in the literature.

The remainder of this paper is organized as follows: Section II provides a brief overview of the HIS and the theories put forward to counter the antigens. Section III deals with a basic classification of IDSs in WSNs and draws an analogy with the immune system. Section IV describes the manner in which an analogy is derived and puts forward the functional mechanism of the IDS. Section V presents the simulations and discussions, and Section VI concludes this paper and presents topics for future research.

## II. AN OVERVIEW OF IMMUNE SYSTEM AND DANGER MODEL

Immunity in the human body is supported by white blood cells/Leucocytes. Different types of Leucocytes have different properties and collectively work towards a three-stage immune mechanism which involves a physical barrier, innate immune system, and adaptive immune system. The adaptive immune system is further classified as Humoral or Cell mediated based on the cells that are associated with the task performed [10].

The physical layer involves physical barriers such as skin which acts as the first layer of defense, protecting the body from various antigens. The innate response refers to the pro-active defense mechanisms which causes reactions such as inflammation to protect the body from external pathogens. The adaptive immune system, as the name suggests, learns, responds, and adapts as a result of previous attacks.

The adaptive immune system is further classified based on two vital immune cells. The B cell is involved in humoral immunity which tends to generate antibodies to kill the pathogens/antigens that enter the body. These antibodies are produced to be antigen specific. Previous knowledge of the attacks is used by the B cells to produce antibodies and thereby enable the immune system to learn and adapt to new attacks. Some of the B cells are memory cells which store the information regarding a pathogen. Other kinds of B cells include plasma, B1 and B2 cells. T cells are the cells responsible for cell mediated immunity and for secreting lymphokines. T cells act as natural killer cells and kill the intruders themselves and keep a memory of them. Other kinds of T-cells include helper, cytoxic, and memory cells [11]. As mentioned earlier, each leucocyte has a different role. For instance, basophil and eosinophils are the cells responsible for innate response. B and T cells are the important cogs of the adaptive Immune system as mentioned above. Dendritic Cells act as a link between innate and adaptive immune systems and present the antigens to the antibodies secreted by B cells. A detailed classification of Leucocytes is given in Fig. 2 The work in this paper, is mostly centered around Dendritic, B and T cells.

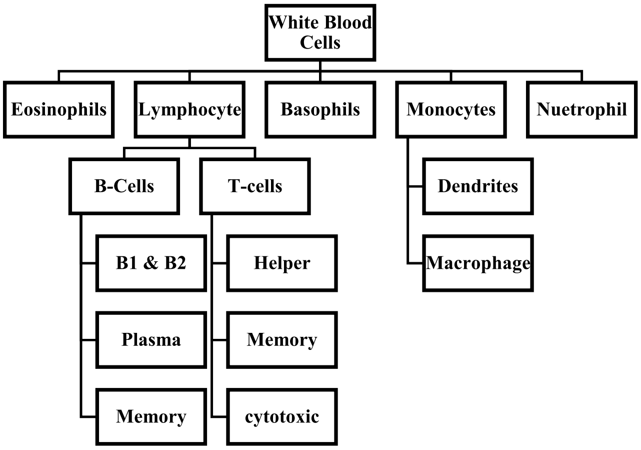


**FIGURE 2.** Classification of the white blood cells.

### A. DANGER THEORY

The Danger model states that the tissues and cells are responsible for immune response when they are subjected to stress or abnormal cell death [12]. These tissues generate different kinds of danger or alarm signals after they identify an antigen presence.

The signals generated generally include Damage-associated molecular patterns (DAMP) and pathogen associated molecular pattern (PAMP).Safe signals are also introduced to make sure there are no major false positives that might disrupt the functioning of the network. Although this model has both its merits and demerits when compared to the classical self and non-self model, Danger theory is generally more preferable to serve as a model for an Artificial Immune System for WSNs owing to its centralized organization and the low number of false positives.

## III. CLASSIFICATION OF INTRUSION DETECTION SYSTEMS

A brief classification of intrusion detection systems based on the methods of detection [13] is given in Fig. 3. The detection method used in this paper is a hybrid of anomaly and signature-based detection methods. Anomaly-based detection involves generating a network profile. These methods include statistical modelling which builds statistical models of the features that the normal network possesses as a reference and comparing them to the actual parameters generated by the network under test. The extent of anomaly is calculated, and an attack is flagged when the anomaly reaches beyond a threshold. Knowledge-based detection draws a profile based on the previous knowledge of the network under different test cases and uses it to detect intrusions. The third type of anomaly detection is based on machine learning. This method generally uses the previous states of the network to define the current state and compare it with the actual state of the network. Fuzzy learning, Neural Networks, Bayesian Networks and Markov models are a few examples of machine learning techniques. Supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning are the types of machine learning techniques employed. Signature based detection has a predetermined set of anomalous profiles which will be used to compare with the network under test. This method is known to produce a very limited number of false positives and can detect any known attack which is previously fed to the detector. Specification based intrusion detection is based on a set of rules or specifications developed by a user. These subsets of IDS design are able to identify known and a variety of unknown attacks.

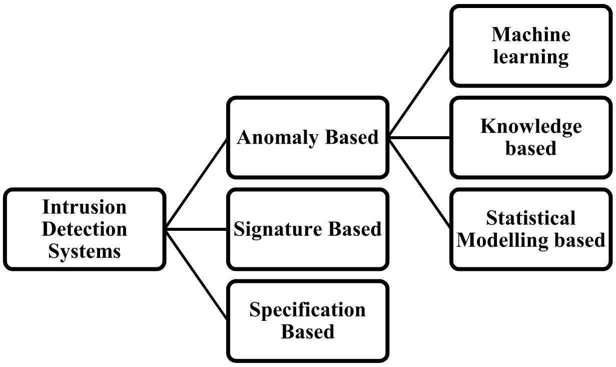


FIGURE 3. Classification of the intrusion detection system.

## IV. ENGINEERING AN ARTIFICIAL IMMUNE SYSTEM

The IDS developed will be studied in three phases. While the first phase includes drawing an analogy using different leucocytes, assigning node functionalities and giving an overview of the system, the second and third phases include establishing the innate and adaptive response systems, respectively. While Fig. 4 demonstrates the second and third phases of the system through a flow chart, Fig. 5 gives an illustration of the first phase, which adapts the immune cell properties into a wireless sensor Network. Each corresponding node adapts the designated immune properties of these cells. For instance, a Dendritic cell is mapped to a detection node which has a higher priority than the other nodes. Similarly, we have a B-cell and a T-cell mapped to certain nodes in a WSN which are the part of the detection process and are designated to be the detection nodes.

### A. DRAWING AN ANALOGY AND MODELLING THE IDS

The IDS proposed is built around an immune model called Danger Theory. Dendritic cells act as a bridge between innate and adaptive immune systems and act as the immune cell responsible for activating immune response by sending alerts. These alerts, in the form of danger signals and PAMP signals are exploited by correlating them with a handful of WSN features which are vital for the functioning of the network. Features such as energy, packets transmitted, and time duration are obtained, and a statistical change is noted for anomaly detection through the alerts. A few nodes mimicking specific immune cells are strategically placed in a cluster and certain tasks are assigned to each of those specialized nodes along with providing some computational abilities. These nodes, which are ideally placed close to root or sink nodes, include Dendritic, B, T and a Basophil cells. These nodes combine to create a private network in order to communicate among each other. Fig. 6 gives an illustration of this network which is the mainstay of the proposed AIS. As a part of the preliminary detection process or the innate immunity process, the network profile is synthesized as artificial peptides and fed to the B-cell, which tries to match them to the PAMP's generated as a part of the Adaptive immunity system discussed in the sequel. This phase is called PAMP Analysis, as seen in Fig. 4 and it helps in reducing the energy consumption of the IDS and also the time required to detect an attack. If an attack is not detected during this stage, it is turned over to the dendritic cell for further analysis. This stage is the Adaptive immunity phase, which generates the Danger

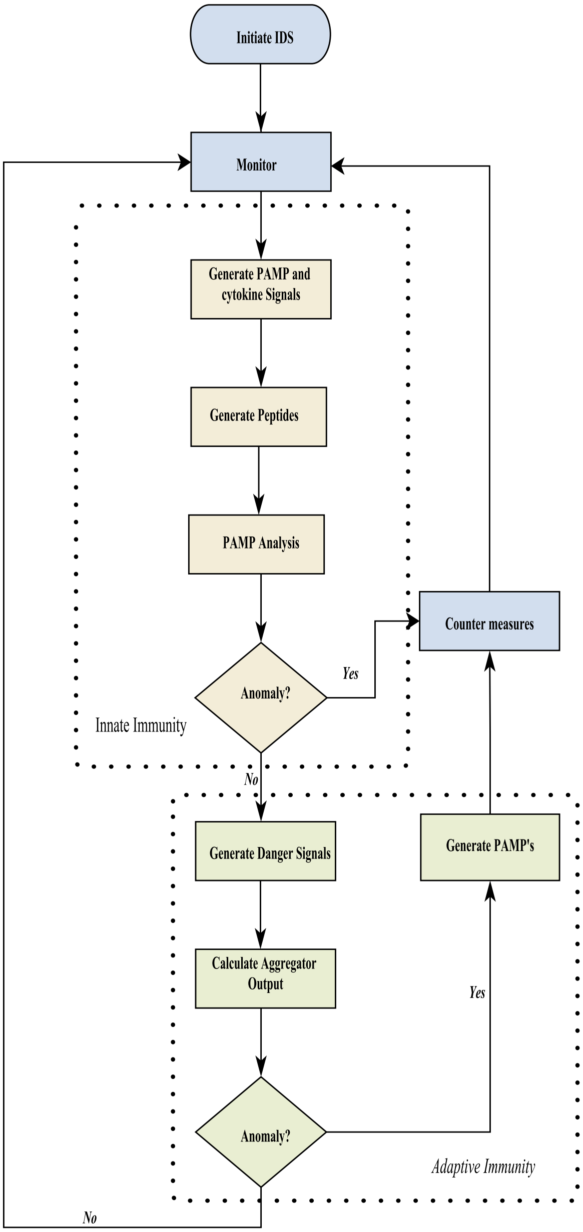


**FIGURE 4.** **Flowchart of the IDS.**

signals and calculates the aggregator output to determine an Anomaly and is shown in Fig. 4. The dendritic cell possesses the computational abilities to determine an anomaly and alert the network by sending danger and PAMP signals along with cytokine signals.

The quantified signals are matched with their respective weights and an aggregator output is obtained. It should be noted that this aggregator output does not necessarily specify the extent of anomaly.

The aggregator output determines whether the node under test is sent to the basophil or the T node/cell. If the aggregator output is greater than the threshold ($\delta_1$), an anomaly is noted. If the aggregator output is below the predefined threshold ($\delta_2$), the node under test will be sent to the basophil which initiates and performs intrusion response. These response techniques can be generic or focused towards a particular type of attack. They can include limiting the data sent from the anomalous node, discarding redundant packets from the affected node, restricting the rate of transmission, use of various error correction techniques and barring connection requests from the malicious node. Basophil implementation is not studied in this paper. If the aggregator output is greater than the threshold ($\delta_2$), the node under test will be sent to the T node/cell, which shuts down or isolates the anomalous node and caches the information regarding the node in its own reserved memory which can be used for future analysis. B cell has its own reserved memory which contains the PAMPs that is, the anomalous signatures which are used for PAMP analysis using bit matching.

A total of nine alert signals are generated out of which two of them are danger signals, three are PAMP signals, one is a safe signal, and the other three are cytokine signals. DS1 is the first danger signal and is based on probabilistic energy comparisons. Another danger signal (DS2) is generated when there is statistical anomaly in the data sent and received. PS1 is attained by monitoring the frequency of the data transfer from a particular node to the sink. PS2 is produced by looking at the duration of the connection each node has with its sink or the cluster head. PS3 is based on the time of transfer, rather the time interval between the transfers. Cytokine signals, IC1, IC2 and IC3 are derived based on the node status, number of hops and the type of connection respectively.

Type of connection (local/remote) is determined by monitoring whether the node is transmitting in the same cluster. Node status and type of connection are binary value representations during the detection process. A Safe Signal (SS) is also incorporated to make sure the network wouldn't start responding for intrusion even when there is no intrusion. A safe signal is sent when the sink receives all the data it is supposed to, in a timely fashion. If a safe signal is received after the basophil initiates intrusion response, it triggers immunosuppression, which means that the response mechanism is driven down. It should be noted that there can be more than one signal at any given instant of time. A cumulative output is determined by assigning weights to these signals resulting in the aggregator output. The weights of the signals are determined based on their impact on the network behavior and their likeliness to predict an intrusion. A graphical representation of these weights in a descending order is given in Fig. 7. Aggregator output, calculated based on the signals generated is given by equation 1.

$$Aggregator\ Output = \left|\sum_{ps1}^{ps3}(W_p{}^{*}C_p)\right| + \left|\sum_{ds1}^{ds3}(W_d{}^{*}C_d)\right| + \sum_{ic1}^{ic3} C_{ic} \tag{1}$$

$$W_{ds1} + W_{ds2} + W_{ps1} + W_{ps2} + W_{ps3} = 1 \tag{2}$$

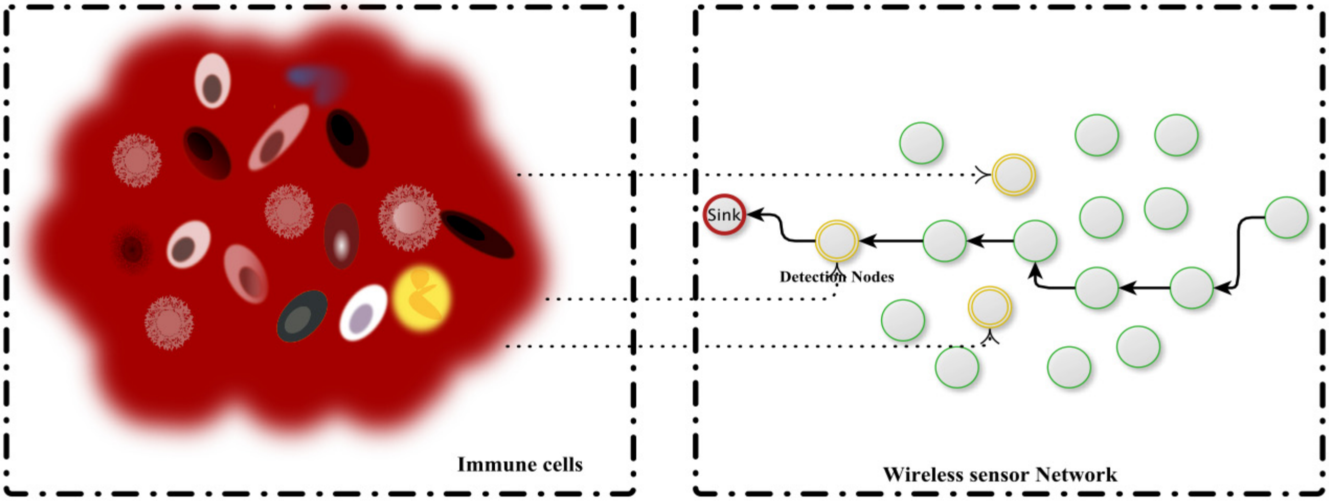


**FIGURE 5.** Deriving an analogy from HIS to WSN.

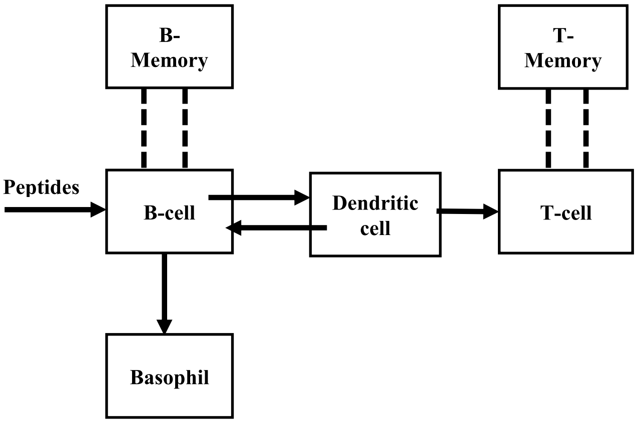


**FIGURE 6.** Block diagram of IDS.

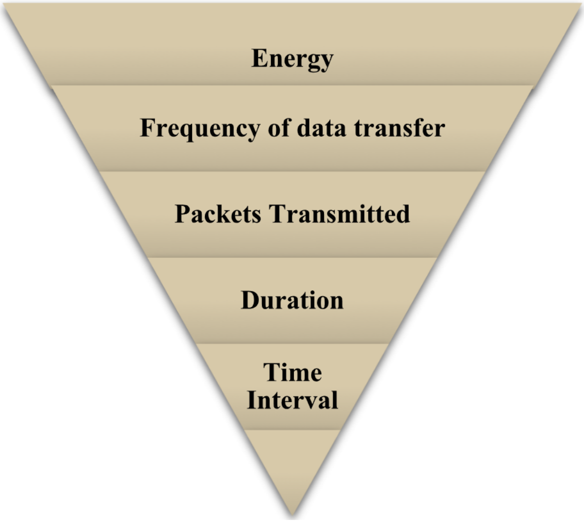


**FIGURE 7.** Weights of the Parameters from High to Low.

Here, Wp, Wd and Wic represent the weights of the parameters used. $C_p$, $C_d$, $C_{ic}$ are the concentrations of PAMP, danger and the cytokine signals respectively. The weights the system assumes are arbitrary. They can be varied and reconfigured during the execution. The rationale behind using the proposed weighing scheme is the impact of those features over the network performance and the number of false positives these parameters generate when they are used as a part of an IDS. So, effectively, the parameters with higher influence on the network's performance such as energy and data transmitted have higher weights. As mentioned earlier, these weights can be varied depending on the user requirements and the type of attack the system is typically prone to. $C_p$, $C_d$, $C_{ic}$ are calculated based on the statistical analysis of the various danger and PAMP signals. Cytokine concentrations assume binary values as mentioned above and only appear in the integer part. A cytokine signal confirms an anomaly. When there are no cytokine signals, the fraction part will be the measure of anomalous activity. Concentrations, $C_p$ and $C_d$ are assigned values in the range of 0.1 to 1 with a step increase of 0.1. The signal outputs obtained will be sampled into multiple range of values and each range is represented with multiples of 0.1.

### B. GENERATION OF PEPTIDES AND PAMP ANALYSIS

A set of peptides are derived from each node on different time intervals. These peptides are then fed to the B-cells to perform PAMP analysis. PAMP's are the attack signatures that are derived previously and are stored in the B-memory. B-cell compares these artificial peptides to the PAMP's using a bit matching algorithm and certifies an intrusion when a match is found. The peptides derived based on network profile are in the form of binary strings and have a length of 121 bits. The order of the string would be source address, destination address, duration, time interval, bytes received, bytes sent, Hop Count, protocol type, Connection type and node status.

A typical peptide would look like (e,1, 6, c,7, a, c, e, 1, 6, c, 7, a, 1, 2, 0, 0, 0, 7, 6, 5, 7, 0, 0, 0, 2, 4, 5, 6, 0, 0, 8, 9, 0, 1, 2, 1, 5, 7, 1, 1). The string contains the information that the source address 225.108.122.12 is sending and receiving 2456 and 7657 bytes respectively to and from 225.108.122.12 for a duration of 86 seconds with a hop count of 5. The last couple of binary representations mean that the node is transmitting locally and to an active node. If a match is found for this peptide, in the B-memory, then it signals an anomaly. The Algorithm 1 below gives an overview of a version of PAMP analysis and the subsequent detection process.

**Algorithm 1** PAMP Analysis & Consequent Steps

```
1: Input: Artificial Peptides,
2: Output: Aggregator Output
3: Result: Intrusion detection.
4: while $\text{Source}_{pe} = \text{Source}_{pa}$ && $\text{Destination}_{pe} =$
     $\text{Destination}_{pa}$ do
5: if $\text{PS}_{pe} = \text{PS}_{pa} || \text{IC}_{pe} = \text{IC}_{pa}$, then
5: confirm Intrusion
6: else compute Aggregator Outputs
7: while $(\text{AO} > \delta_1)$ do
8: Confirm Anomaly
9: if $(\text{AO} < \delta_2)$ then
10: Send to Basophil for response
11: else Send to T-cell
12: end if
13: end while
14: end if
15: end while
```

### C. GENERATING DANGER AND PAMP SIGNALS

Anomalies in energy consumed are obtained based on the difference between the predicted and the residual energies. Probabilities of state transitions between sensing, calculating, sending, receiving and sleep modes are determined by using Chapman-Kolmogorov equations [14] based on Markov chain modelling [15]. The system assumes a Markovian process, when the probability of transitioning from one state 'i' to the next state 'j' is not dependent on any other states, which are in operation prior to the preceding state 'i'. When a sequence of random variables, $\{x_1, x_2, x_3, x_4, x_5, \ldots\}$, which denote the state of the network satisfy the Markovian process, along with assuming the properties $P_{\text{ij}} = 0$ and $\sum_{\text{j}=0}^{\infty} P_{\text{ij}} = 1$ they are said to adopt a Markovian chain. Chapman-Kolmogorov equations can be realized when the network or a system which satisfies the Markovian property and follow a finite discrete time process. These equations can help determine the probability of the system to move from one state to another in 'n' time steps. Using these equations, the probability of state transitions from i to j in 't' time slots is given in equation (3)

$$P_{ij}^{t} = \sum_{k=0}^{t} P_{ki}^{r} P_{kj}^{(t-r)} \quad (0 < r < t) \tag{3}$$

where i is the initial state of the sensor node and j is the state of the node after 't' time slots. By determining the initial state of the node under test, we can effectively deduce the node's current state in its succeeding iterations. Before predicting the energy dissipated, the time slots the node remains in state j will be determined using the expression,

$$\text{T}_{\text{j}} = \sum_{t=1}^{T} P_{ij}^{t} \tag{4}$$

If we consider that the energy consumed at each time step as $\text{E}_{\text{t}}$, the energy predicted, $\text{E}_{\text{p}}$ can be calculated by using

$$E_p = \sum_{j=1}^{5} (\sum_{t=1}^{T} P_{ij}^{t})^* E_t \tag{5}$$

The concentration of DS1 is obtained by calculating the difference between the energy predicted and the actual energy consumed, i.e., $E_a = (\text{E}_{\text{i}} - \text{E}_{\text{r}})$

$$\text{C}_{\text{ds1}} = \frac{\left|(\text{E}_{\text{i}} - \text{E}_{\text{p}} - \text{E}_{\text{r}})\right|}{\text{E}_{\text{p}}} {}^* \text{n} \tag{6}$$

Here, $\text{E}_{\text{i}}$ is the initial energy $\text{E}_{\text{a}}$ is the actual energy and $\text{E}_{\text{r}}$ is the residual energy of the node under test.

For DS2, data sent and received are predicted based on (3) and by determining the probabilities for state transitions to both sending and receiving states. The number of time slots the node will remain in receive and transmitting state is calculated using the following expressions:

$$\text{T}_{\text{pr}} = \sum_{t=1}^{T} P_{ir}^{t} \tag{7}$$

$$\text{T}_{\text{ps}} = \sum_{t=1}^{T} P_{is}^{t} \tag{8}$$

Here, *ir* and *is* are the changes of states to receive and send phases respectively. The sent and received packets are predicted using the formulas below. $\text{SP}_{\text{p}}$ and $\text{RP}_{\text{p}}$ are the predicted sent and received packets. $\text{SP}_{\text{t}}$ and $\text{RP}_{\text{t}}$ are the packets transmitted for each time slot.

$$SP_p = \sum_{j=1}^{3} (\sum_{t=1}^{T} P_{is}^{t})^* SP_t \tag{9}$$

$$RP_p = \sum_{j=1}^{3} (\sum_{t=1}^{T} P_{is}^{t})^* TP_t \tag{10}$$

Concentrations of DS2 are determined by using the next expression, which is a result of comparison with the actual packets transmitted and the packets predicted to be transmitted.

$$\text{C}_{\text{ds2}} = \frac{\left|(\text{SP}_{\text{p}} - \text{SP}_{\text{a}}\right|}{\left|(\text{RP}_{\text{p}} - \text{RP}_{\text{a}}\right|} {}^* \text{n} \tag{11}$$

**TABLE 1.** Danger and PAMP statistics.

| Attack | DS1 | | | | DS2 | | | | PS1 | | | | PS2 | | | | PS3 | | | | Aggregator Output |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $T_1$ | $T_2$ | $T_3$ | $T_4$ | $T_1$ | $T_2$ | $T_3$ | $T_4$ | $T_1$ | $T_2$ | $T_3$ | $T_4$ | $T_1$ | $T_2$ | $T_3$ | $T_4$ | $T_1$ | $T_2$ | $T_3$ | $T_4$ | |
| Blackhole (1) | 10% | 10% | 10% | 20% | 20% | 40% | 60% | 80% | 20% | 20% | 40% | 50% | 0% | 0% | 0% | 0% | 60% | 60% | 60% | 60% | 0.395 |
| Blackhole (2) | 0% | 10% | 10% | 10% | 20% | 40% | 60% | 60% | 30% | 30% | 30% | 30% | 0% | 0% | 0% | 0% | 0% | 40% | 70% | 70% | 0.31 |
| Blackhole (3) | 10% | 10% | 10% | 10% | 20% | 50% | 50% | 70% | 10% | 20% | 20% | 30% | 0% | 0% | 0% | 0% | 20% | 30% | 50% | 60% | 0.30 |
| Wormhole1 | 10% | 20% | 40% | 60% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 2.21 |
| Wormhole2 | 20% | 30% | 30% | 50% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 40% | 40% | 40% | 40% | 1.175 |
| DDOS1(control packet flood) | 30% | 50% | 60% | 80% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 40% | 40% | 40% | 40% | 0.36 |
| DDOS2(data flood) | 10% | 20% | 40% | 70% | 20% | 40% | 60% | 60% | 0% | 0% | 0% | 0% | 20% | 20% | 20% | 20% | 40% | 40% | 40% | 40% | 0.405 |
| Selective Forwarding 1 | 0% | 0% | 10% | 10% | 30% | 30% | 30% | 40% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0.075 |
| Selective Forwarding 2 | 10% | 10% | 20% | 20% | 30% | 50% | 50% | 50% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0.12 |

PS1 is determined by calculating the difference in the frequency of data transfers in ‘t’ time slots.

$$C_{ps1} = \frac{|ft_t - ft_{th}|}{ft_T} * n \qquad (12)$$

$C_{ps1}$ can be calculated either by comparing the frequency of data transfer against a threshold or by comparing it with the frequency of transmission at earlier timeslots usually through statistical average. $C_{ps2}$ and $C_{ps3}$ can also be calculated in a similar way. It should be noted that the danger signals are not calculated at each collection point to avoid excess energy consumption.

## V. EXPERIMENTAL SETUP AND SIMULATIONS

Simulations are done using a simulator called Cooja in a Contiki environment using a protocol called RPL [16]. 6LowPAN and IPv6 are the protocols used at the network layer. RPL is a protocol to assist routing in low power and noisy networks.

RPL works by making use of a set of Destination Oriented Directed Acyclic Graphs (DODAG’s). All the nodes are assigned ranks based on their proximity to the sink node or the root. RPL has four control messages namely, DODAG Information Solicitation (DIS), DODAG Information Object (DIO), DODAG advertisement Object (DAO) and an Acknowledgement (ACK) to the DAO. DIO is multi casted to lower ranks by a specific node allowing those nodes to sniff the information regarding the multicasting node. DIO helps the lower ranked nodes in determining if they want to join the DODAG. DIS is sent when a node does not notice any DIO. DIS is broadcasted to see if any DODAG is available for the node to join. After the node finds a suitable DODAG either through acknowledging a DIO or sending a DIS, it makes a request to join the DODAG by sending a DAO. ACK is sent as an acknowledgement to DAO.

During this simulation, the number of nodes deployed is varied from 20 to 30 to 40. The nodes are placed in a 100 m*100 m area with a range of 50 m. Simulations and calculations are performed to calculate the Aggregator output, detection rate, packet overhead and energy overhead. During the initial training phase, cytokine values under normal conditions are determined. The attacks employed during this simulation are Blackhole, Selective forwarding, DDoS, and wormhole. These attacks are further categorized and implemented in various ways.

Table 1 gives an overview of the attacks implemented and the resultant anomalies quantified and expressed through danger and PAMP signals along with the Aggregator output. Two types of Selective Forwarding attacks implemented in this work are packet delay based and packet loss-based attacks. While one attack buffers the packet at the anomalous node for a given period of time and transmits them after some time, resulting in a compromise in the freshness of the data, the other type of Selective forwarding attack implemented, discards packets at regular intervals. DDoS attacks are implemented as control packet flooding and data packet flooding, both of them tend to overwhelm the network with excess and unwarranted information. A couple of different versions of Wormhole attacks are implemented based on the tunnel formed by the adversary. One of them creates a tunnel to a distant node in the same cluster, whereas the second variant creates a tunnel to a different cluster. Three different Blackhole attacks are implemented by altering the placement of the malicious node in different ranks. Fig. 8 presents the Aggregator outputs of the attacks mentioned above, each of them individually simulated for a simulation time of 600 seconds. A preprocessed version of Fig. 8 with the alarm signals is given in Table 1. Both Table 1 and Fig. 8 are the resultants of an attacker in a 30-node network, the attack

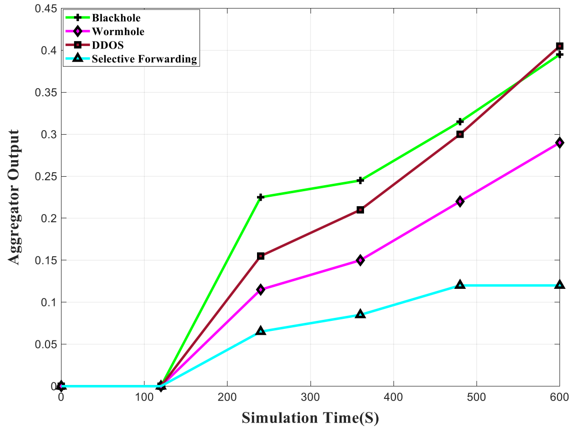


**FIGURE 8.** Aggregator outputs.

starting at 120 seconds. All the danger signals and PAMP signals calculated are presented in the form of percentiles during four time periods, $T_1 = 240$s, $T_2 = 360$s, $T_3 = 480$s and $T_4 = 600$s for all the attacks mentioned earlier. Blackhole attacks generate four different alarm signals, which are DS1, DS2, PS1 and PS3. Anomalies are found in energy, number of packets received or sent, time interval between transfers and the frequency of transfer. Since the three different Blackhole attacks are only a change in the location and RPL rank of the attacker node in the network and all the nodes are transmitting data at a uniform rate, the aggregator outputs do not show any significant variation from one another. There is a proportional increase in PS1 indicating that there exists an anomaly in the number of packets transmitted due to the aggregation of discarded packets as the time progresses. From Fig. 8, we see that the Blackhole curve has the highest output, when seen as a statistical mean of the aggregator outputs at different time slots, thereby it is easier to detect.

Wormhole1 generates DS1, IC2 and IC3. DS1 signal is due to the excessive energy used to forward a packet to a distant node which is exponentially larger than the energy used to send the packet to a neighboring node. IC2 is generated because of the change in the number of hops to the sink. IC3 is due to the connection being remote, that is the node is transmitting to a different cluster. Wormhole2 has similar signals compared to Wormhole1 except the cytokine signal IC3 since it is transmitting locally however with a change in the number of hops. The aggregator output for both Wormhole1 and Wormhole2 has an integer part unlike the rest of the attacks indicating a cytokine presence. As mentioned earlier, a cytokine presence concludes a definite intrusion. DDoS produces the highest final aggregator output.

Both the DDoS attacks witness a substantial anomaly in the energy consumed by the anomalous node. This is due to the flooding of the network and thereby depleting the already constrained resources. DDoS1 floods the network by repeatedly sending DIO's resulting in redundancy. Hence DDoS1 generates only DS1 and PS3. DDoS2 yields DS2 as opposed to DDos1 since it floods the network with UDP packets. Here, the anomalies are found in energy, packets transmitted, duration of the established connection and the time interval (DS1, DS2, PS2 and PS3). The aggregator output of DDoS from Fig. 8 moves past Blackhole at time $T_4$, showing that it is easier to detect DDoS attacks than Blackhole as the simulation time passes a certain threshold. A more detailed comparisons of detection probabilities is given in Fig. 9.

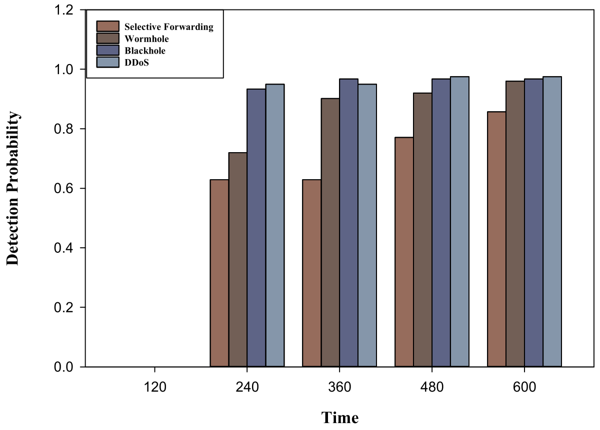


**FIGURE 9.** Probability of detection of each attack.

Selective forwarding produces the lowest aggregator output value. This renders it harder to detect. This is due to the not very substantial change in the energy consumed. On the other hand, this may also be due to the low weights assigned to DS2. More reasons for this behavior are mentioned in the later parts of this paper. Calibrating the weights can result in a better detection probability for Selective Forwarding attacks. However, this leads to an increase in the number of false positives to a considerable extent along with a change in the detection probabilities of other attacks. This is not a desirable trade-off. Both the variants of Selective forwarding attacks have the same signals associated with them (DS1 and DS2) and are comparatively difficult to detect using this detection method. It should be noted that even when some packet delay is associated with one of the attacks, there are no PAMP signals generated.

If we consider $\delta_1$ as 0.1 and $\delta_2$ as 0.3, the intrusion is detected for all the attacks except Selective Forwarding at the first detection point ($T_1$) and is sent to the basophil for countermeasures. However, this is not the case with Selective forwarding. The attack is not detected until the third detection point($T_3$) before it is sent to the basophil. Only the nodes with DDoS and Blackhole after $T_3$ is sent to the T-cell. By varying the threshold, we can determine and set an intrusion tolerance level to all the nodes in the network. Fig. 10 and Fig. 11 provides the packet and energy overhead caused by the IDS respectively for 20,30 and 40 nodes, both of which are vital to the network functioning due to the constraints a WSN possesses. Probability of detection for each attack is derived against the simulated time. Detection rate is measured by varying the seed, placement strategies, number of attackers and as mentioned earlier, number of nodes. From Fig. 9, it can be inferred that Selective forwarding has the least detection probability from this approach. Along with the factors mentioned above, this is also due to the fact that quantum of the packets dropped is not as significant as Blackhole, which discards all the packets traversing through the malicious node and also relatively less energy deviations from the predicted results. Although the detection probability climbs up at a later time due to the aggregation in packets dropped, nevertheless it is lower than the other attacks.

DDoS has the highest detection probability and is one of the easiest attacks to detect using this model, since a noticeable change can be seen in the energy consumed and also since DS1 has the maximum weight among all the signals. Blackhole is another easily detectable intrusion and has one of the highest detection rates due to the number of packets it discards. However, the detection is not probability is not near perfect for a Blackhole attack because of the nodes which are located at a remote location and do not act as intermediate nodes or do not join the DODAG. Detecting an intrusion in these nodes is problematic. A Cytokine presence is typically traced for a Wormhole attack. However, even when there isn't a cytokine occurrence, Wormhole has a decent detection probability as evident from Fig. 8 and Fig. 9. The same scenario with the same attack is implemented twice to check the effectiveness of the B-cell. Iteration1 gives the results of the first experiment and Iteration 2 gives the result of the second experiment, where the B-cell matches the anomalous peptides obtained from Iteration 1 and need not forward it to the Dendritic cell for further analysis, thereby conserving energy and memory. This is evident from Fig. 10 and Fig. 11. We can see a difference of Packet overhead when Iteration1 and Iteration 2 are compared. This shows that the attack is detected early and the dendritic cell is not triggered. Fig. 10 shows that there is at least a 7% difference in the memory consumed by the IDS during Iteration1 and Iteration2. The total memory used by the IDS during iteration1 ranges from 9.2k,13.7k and 18.3k for nodes 20,30 and 40 respectively, which is significantly lower than the 48k ROM allocated in a sensor node. Since the B-memory stores strings of length 121 bits, a total of 3250 detector strings can be stored before the memory depletes. From Fig. 11, we can notice that there is a 14% increase in the energy dissipated when an IDS is in function for a 20-node network. This gets decreased to 8.3% during iteration2, indicating a difference of 5.7%.

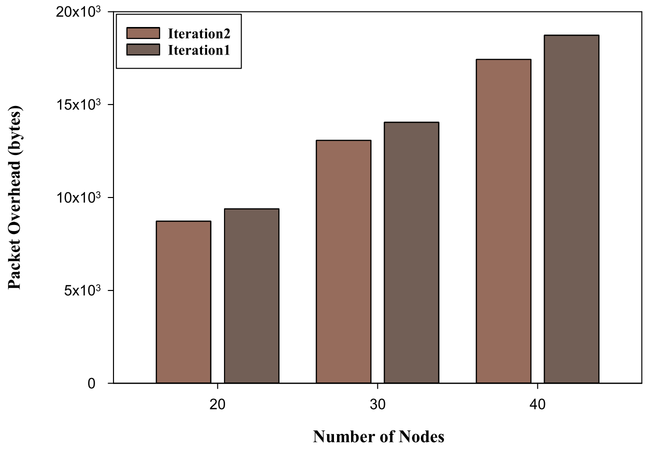


**FIGURE 10. Packet overhead caused by the IDS.**

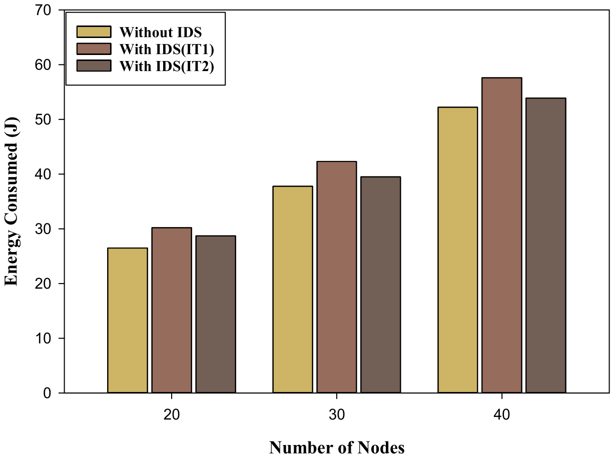


**FIGURE 11. Energy overhead caused by the IDS.**

Similarly, the difference in energies from iteration1 and iteration2, for when the number of nodes is 30 and 40 are 7.5% and 7.7% respectively. This evidently proves that the difference in the energy dissipated is significant and the innate detection phase is efficient thereby saving a considerable amount of energy.

## VI. CONCLUSIONS AND FUTURE SCOPE

An IDS is designed by taking inspiration from the human immune system, while considering the resources that have the potential to affect the performance of a WSN. Although this approach is designed for WSNs, it can be easily modified for use on other resource constrained networks. Modifying the Danger signals generated through the dendritic node and the B-cell can help to adapt and extend this IDS to other networks including ADHOC networks. Although this model can predict different types of attacks, energy depleting attacks of any nature can be predicted more accurately and in a timely manner. It should be noted that any light weight signal generator which can find anomalies in the resource usage can be embedded into the skeleton of this model. Taking this into consideration, further studies will be done to make this model more robust and lightweight. Since this model is not attack specific, efforts are underway to classify and segregate the attacks based on the signals obtained. Further efforts are being made to implement the basophil and the T-cell nodes in order to build a more complete and robust system.

## REFERENCES

[1] K.-M. Kim, H. Kim, and K. Kim, "Design of an intrusion detection system for unknown-attacks based on bio-inspired algorithms," in *Proc. Comput. Secur. Symp.*, 2015, pp. 1–7.

[2] T. Li and N.-F. Xiao, "Novel heuristic dual-ant clustering algorithm for network intrusion outliers detection," *Optic—Int. J. Light Electron Opt.*, vol. 126, no. 4, pp. 494–497, Feb. 2015.

[3] S. Hofmeyr and S. Forrest, "Architecture for an artificial immune system," *Evol. Comput.*, vol. 8, no. 4, pp. 443–473, Dec. 2000.

[4] S. Forrest, A. S. Perelson, L. Allen, and R. Cherukuri, "Self-nonself discrimination in a computer," in *Proc. IEEE Symp. Secur. Privacy*, May 1994, pp. 202–212.

[5] U. Aickelin and S. Cayzer, "The danger theory and its application to artificial immune systems," in *Proc. 1st Int. Conf. Artif. Immune Syst.*, Canterbury, U.K., 2002, pp. 1–8.

[6] J. Kim and P. J. Bentley, "Towards an artificial immune system for network intrusion detection: An investigation of clonal selection with a negative selection operator," in *Proc. Congr. Evol. Comput.*, Seoul, South Korea, May 2001, pp. 1244–1252.

[7] S. Forrest, S. A. Hofmeyr, A. Somayaji, and T. A. Longstaff, "A sense of self for unix processes," in *Proc. IEEE Symp. Secur. Privacy*, May 1996, pp. 120–128.

[8] F. Hosseinpour, P. V. Amoli, F. Farahnakian, J. Plosila, and T. Hämäläinen, "Artificial immune system based intrusion detection: Innate immunity using an unsupervised learning approach," *Int. J. Digit. Content Technol. Appl.*, vol. 8, no. 5, pp. 1–12, Oct. 2014.

[9] J. Greensmith, U. Aickelin, and S. Cayzer, "Introducing dendritic cells as a novel immune-inspired algorithm for anomaly detection," in *Proc. 4th Int. Conf. Artif. Immune Syst.*, 2005, pp. 153–167.

[10] K. Hoebe, E. Janssen, and B. Beutler, "The interface between innate and adaptive immunity," *Nature Immunol.*, vol. 5, no. 10, pp. 971–974, Oct. 2004.

[11] L. A. DiPietro, "Wound healing: The role of the macrophage and other immune cells," *Shock*, vol. 4, no. 4, pp. 233–240, Oct. 1995.

[12] P. Matzinger, "The danger model: A renewed sense of self," *Science*, vol. 296, no. 5566, pp. 301–305, 2002.

[13] I. Butun, S. D. Morgera, and R. Sankar, "A survey of intrusion detection systems in wireless sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 266–282, 1st Quart., 2013.

[14] W. Shen, G. Han, L. Shu, J. J. P. C. Rodrigues, and N. Chilamkurthi, "A new energy prediction approach for intrusion detection in cluster-based wireless sensor networks," in *Proc. Int. Conf. Green Commun. Netw.*, Berlin, Germany, 2011, pp. 1–12.

[15] C. Vasar, O. Prostean, I. Filip, R. Robu, and D. Popescu, "Markov models for wireless sensor network reliability," in *Proc. IEEE 5th Int. Conf. Intell. Comput. Commun. Process.*, Aug. 2009, pp. 323–328.

[16] T. Winter *et al.*, *RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks*, document RFC 6550, 2012.

[17] S. Raza, L. Walgreen, and T. Voigt, "SVELTE: Real-time intrusion detection in the Internet of Things," *Ad Hoc Netw.*, vol. 11, no. 8, pp. 2661–2674, Nov. 2013.

[18] D. Dasgupta, "Immunity-based intrusion detection system: A general framework," in *Proc. 22nd NISSC*, 1999, pp. 147–160.

[19] W. Li, "Using genetic algorithms for network intrusion detection," in *Proc. U.S. Dept. Energy Cyber Secur. Group 2004 Train. Conf.*, 2004, pp. 1–8.

[20] V. T. Alaparthy and S. Morgera, "Modelling an intrusion detection system based on adaptive immunology," *Int. J. Interdiscipl. Telecommun. Netw.*, to be published.

[21] P. Pongle and G. Chavan, "Real time intrusion and wormhole attack detection in Internet of Things," *Int. J. Comput. Appl.*, vol. 121, no. 9, pp. 1–9, 2015.

[22] J. Kim, P. Bentley, C. Wallenta, M. Ahmed, and S. Hailes, "Danger is ubiquitous: Detecting malicious activities in sensor networks using the dendritic cell algorithm," in *Proc. Int. Conf. Artif. Immune Syst.*, 2006, pp. 390–403.

[23] M. Zeeshan, H. Javed, and S. Ullah, "Discrete R-contiguous bit matching mechanism appropriateness for anomaly detection in wireless sensor networks," *Int. J. Commun. Netw. Inf. Secur.*, vol. 9, no. 2, pp. 157–163, Aug. 2017.

[24] Y. Maleh and E. Abdellah, "A review of security attacks and intrusion detection schemes in wireless sensor networks," *Int. J. Wireless Mobile Netw.*, vol. 5, no. 6, pp. 1–12, Jan. 2014.

[25] J. M. Vidal, A. L. S. Orozco, and L. J. G. Villalba, "Adaptive artificial immune networks for mitigating DoS flooding attacks," *Swarm Evol. Comput.*, vol. 38, pp. 94–108, Feb. 2018.

[26] P. Mostardinha, B. F. Faria, A. Zúquete, and F. V Abreu, "A negative selection approach to intrusion detection," in *Proc. 11th Int. Conf. Artif. Immune Syst.*, 2012, pp. 178–190.

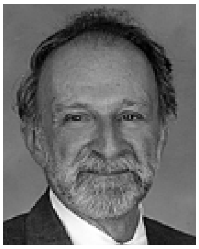
[27] J. Kim, P. Bentley, U. Aickelin, J. Greensmith, G. Tedesco, and J. Twycross, "Immune system approaches to intrusion detection—A review," *Natural Comput.*, vol. 6, no. 4, pp. 413–466, 2007.

[28] H. M. Salmon *et al.*, "Intrusion detection system for wireless sensor networks using danger theory immune-inspired techniques," *Int. J. Wireless Inf. Netw.*, vol. 20, no. 1, pp. 39–66, 2013.

[29] M. Riecker, S. Biedermann, R. El Bansarkhani, and M. Hollick, "Lightweight energy consumption-based intrusion detection system for wireless sensor networks," *Int. J. Inf. Secur.*, vol. 14, no. 2, pp. 155–167, Apr. 2015.

[30] R. Rizwan, F. A. Khan, H. Abbas, and S. H. Chaudhary, "Anomaly detection in wireless sensor networks using immune-based bioinspired mechanism," *Int. J. Distrib. Sensor Netw.*, vol. 11, no. 10, p. 684952, Jan. 2015.

**VISHWA TEJA ALAPARTHY** received the bachelor's and master's degrees in electrical engineering from Jawaharlal Nehru Technological University, Hyderabad, India, in 2011 and 2013, respectively. He is currently pursuing the Ph.D. degree with the University of South Florida. His research interests are network security, IOTs, machine learning, and wireless sensor networks.



**SALVATORE DOMENIC MORGERA** was a Professor and the Director of the Information Networks and Systems Laboratory with the Department of Electrical and Computer Engineering, McGill University, Montreal, Canada. He was also a Professor, the Chair of Electrical Engineering, and the Director of the Bioengineering Program with Florida Atlantic University. He is currently a Professor and the Previous Chair of Electrical Engineering and Biomedical Engineering with the University of South Florida, and the Emeritus Professor at Florida Atlantic University. He is also the Director of the Global Center for Neurological Networks. He has over 40 years of leadership in industry, government, and academia. He was the Major Project Leader at the Canadian Institute for Telecommunications Research, the Government of Canada Network of Centres of Excellence, the President of the Quebec Research Council–Le Fonds Nature et Technologies, and the Special Assistant to the President, Communications Research Center, Industry Canada, Government of Canada.