IJICTR International Journal of Information & Communication Technology Research ITRC

Volume 9 - Number 2 - Spring 2017(53-62)

A Security Mechanism for Detecting Intrusions in Internet of Things Using Selected Features Based on MI-BGSA

Mansour Sheikhan
Department of Communication Engineering
South Tehran Branch, Islamic Azad University
Tehran, Iran
msheikhn@azad.ac.ir

Hamid Bostani
Research Center of Modeling and Optimization in Science and Engineering
South Tehran Branch, Islamic Azad University
Tehran, Iran
st_h_bostani@azad.ac.ir

Received: January 17, 2017 - Accepted: March 16, 2017

Abstract—**Internet of things (IoT) is a novel emerging approach in computer networks wherein all heterogeneous objects around us, which usually are resource-constrained objects, can connect to each other and also the Internet by using a broad range of technologies. IoT is a hybrid network which includes the Internet and also wireless sensor networks (WSNs) as the main components of IoT; so, implementing security mechanisms in IoT seems necessary. This paper introduces a novel intrusion detection architecture model for IoT that provides the possibility of distributed detection. The proposed hybrid model uses anomaly and misuse intrusion detection agents based on the supervised and unsupervised optimum-path forest models for providing the ability to detect internal and externals attacks, simultaneously. The number of input features to the proposed classifier is reduced by a hybrid feature selection algorithm, as well. The experimental results of simulated scenarios show the superior performance of proposed security mechanism in multi-faceted detection.**

Keywords- Internet of things, intrusion detection, anomaly-based, misuse-based, optimum-path forest

I. INTRODUCTION

During the recent decades, computer networks (and especially the Internet) have been widely used in human's world and various areas such as education, government and business. The Internet of things (IoT) is a novel emerging approach in computer networks in which all heterogeneous objects around us (such as smart phones, laptops or smart sensors) can connect to the Internet by using a wide range of technologies. In other words, large number of smart interconnected devices in IoT results in valuable services to the society and individual citizens [1]. Moreover, IoT can be supported by satellite communication systems for the case of Internet of remote things (IoRT) in which the Internet protocol version 6 (IPv6) should be supported over satellite [2].

Wireless sensor network (WSN) is one of the main components of IoT in which the nodes are able to communicate with each other and also intelligent systems, autonomously [3]. However, one of the major efforts in creating the real IoT is the IPv6 over low-power wireless personal area networks (6LoWPANs) [4] which was introduced and standardized by the Internet Engineering Task Force (IETF) work group. 6LoWPANs is a WSN based on compressed IPv6 which uses the routing protocol for low power and lossy networks (RPL) [5] for routing the packets in a low-power and lossy network. The general architecture of IoT is shown in Fig. 1. As seen in Fig.

1, the resource-constrained devices of 6LoWPANs such as sensor nodes can connect to the Internet through the 6LoWPAN border router (6BR) [5, 6].

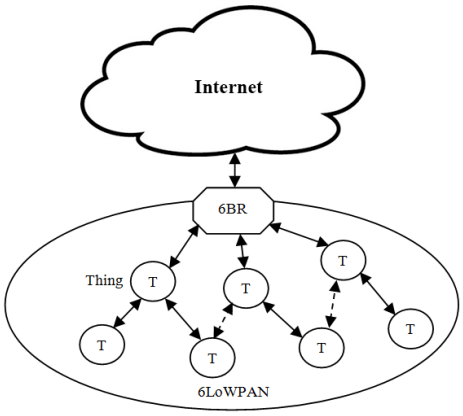

Figure 1. General architecture of IoT.

It is noted that the RPL is a certain routing protocol for 6LoWPAN, which is based on the construction of a destination-oriented directed acyclic graph (DODAG). In other words, RPL is an IP-based distance vector and hop-by-hop routing protocol that enables different operations such as the unidirectional traffic towards a DODAG root, bidirectional traffic between resource-constrained devices (i.e., 6LoWPAN nodes), and bidirectional traffic between resource-constrained devices and the DODAG root [5].

The communication in the IoT can be secured by using standard mechanisms such as cryptography and authentication techniques; however, these preventive mechanisms cannot detect all possible attacks, because of the nature of wireless communication. On the other hand, the resource-constrained devices are connected to the unreliable Internet via IPv6 and 6LoWPAN networks in the IoT; so, they are vulnerable to intrusions (both from the Internet and WSNs) [6]. Therefore, IDS is required for detecting malicious activities in the IoT besides the standard security mechanisms.

Today IDS is one the major and effective solutions for dealing with security threats in computer networks, especially IoT. To put in simply, IDS is looking for some signs that indicate the wrong activities. Therefore, for the reason of specifying these kinds of wrong activities, the information should be filtered. In other words, IDS is an effective tool which gathers system activities or network traffic as input data with the aim of analyzing them for identifying malicious behaviors. From the aspect of the network security researches, IDSs can be categorized into various methods based on (a) analysis methods; (b) data sources; and (c) system architectures. According to the analysis methods, IDSs are classified into the following categories: (a) misuse-based (as the best method for detecting known attacks); (b) anomaly-based (as the best method for detecting unknown attacks); and (c) specification-based detection systems.

In the misuse-based detection systems, predefined attack patterns are modeled and maintained in the database of the attacks' signature. Therefore for detecting intrusion, these signatures are used for matching with collected data. The signature of each attack is required for detecting them; hence, misuse-based intrusion detection systems cannot identify the unknown intrusions. In other words, the detection rate (DR) of this method is lower than other methods. It is noted, DR is an evaluation metric which refers to the ratio of the number of attacks that were detected properly to all occurred attacks. On the other hand, anomaly-based IDSs are focused on normal traffic of the system. In other words, this type of system which usually uses statistical or machine learning methods is based on finding the deviations from normal behavior of network traffic or system activities. Although, anomaly-based IDSs are outperformed in detecting the unknown attacks, but unlike misuse-based IDSs, they suffer from the high rate of false alarm rate (FAR). It is noted that the FAR, as an important evaluation metric of IDSs, refers to the ratio of extracted normal samples (from network traffic) which were classified as attack into the total extracted normal samples. The specification-based systems work by the same way, as well. However, user guidance is required to extract legitimate system activities or network traffic for developing a model of normal behavior.

IoT is a hybrid network which is composed of the Internet and the networks with heterogeneous nodes (e.g., 6LoWPAN). The traffic pattern of these networks is completely different. Moreover, 6LowPAN is an IP-based WSN which consists of 802.15.4 and IPv6 networks; so, the traffic pattern of this network is also different. Therefore, traditional IDSs for WSN and IP-based networks are not appropriate for IoT. In other words, IDS for IoT should be able to control all different traffic patterns of IoT. With the aim of providing multi-faceted detection (considering both traffic patterns from the Internet and WSNs sides), a hybrid distributed IDS based on MapReduce approach is proposed for intrusion detection in this study for IoT. In fact, a novel real-time intrusion detection framework is proposed which is based on anomaly-based detection for detecting insider (internal) attacks that happen in 6LoWAPN. The proposed method is focused only on detecting the malicious behaviors of sinkhole and selective-forwarding attacks in 6LoWPAN; however, it can be extended for detecting other attacks. On the other hand, a misuse-based intrusion detection engine is provided that is responsible for detecting cyber (external) attacks that occur from the Internet (or LANs) side.

The rest of this paper includes the following sections: Section II reviews related work. The foundations of preliminaries are introduced briefly in Section III. In Section IV, the proposed model is introduced in detail and the performance of it based on simulated scenarios is reported in Section V. Finally, the paper is concluded in Section IV.

II. RELATED WORK

One of the main challenges in IoT is providing an efficient security mechanism for IoT. Generally, the following studies are example researches of security issues in IoT: Security may be addressed as an important factor in the integration of low-power WSN with the Internet. Granjal et al. [7] reviewed the

proposals supporting this integration. Cloud computing technology can be used to enhance the function of the IoT. So, a new paradigm is termed as cloud of things (CoT) or CloudIoT [8]. A survey of IoT and cloud computing with an emphasis on the security issues of both technologies was presented by Stergiou et al. [9]. The security and privacy requirements for the IoT applications such as personal and home, government and utilities, and enterprise and industry were analyzed by Ouaddah et al. [10]. Alcaraz et al. [11] analyzed the security requirements of industrial sensor network-based remote substations [12] in the context of IoT. Airehrour et al. [13] analyzed routing protocols for secure routing communications in IoT. Services in IoT can be provisioned by two architectures: (a) centralized and (b) distributed. In the distributed architecture, the entities at the edge of the network exchange information and collaborate with each other in a dynamic way. Roman et al. [14] analyzed the security and privacy issues in the distributed IoT. Data mining and computational intelligence methods can also play an important role in creating smarter and more secure IoT [15].

Designing IDS for IoT is still a new and on-going research subject and to the best knowledge of the authors, a few researchers in the security field work on this context. For example, Raza et al. [6] proposed a real-time IDS for IoT called SVELTE. They showed that SVELTE has a small overhead to deploy on the constrained nodes and can detect most of malicious nodes that launch sinkhole and/or selective-forwarding attacks. Kasinathan et al. [16] introduced a DoS attack detection architecture for 6LoWPAN. Their simulation results showed the capability of the proposed architecture in detecting DoS attacks. One of the main goals followed by employing IDS in the IoT is fast security event-processing that results in detecting network attacks, immediately. For this purpose, Jun and Chi [17] designed a complex event-processing (CEP)-based IDS in the IoT environment to achieve better performance in real-time. Weber and Studer [18] examined the changing legal cyber-security environment in the IoT context. In this way, selected applicable international regulations and alternative approaches to address the security issues in the IoT were discussed.

A hybrid IDS model consists of anomaly-based and specification-based intrusion detection modules using unsupervised optimum-path forest (OPF) is recently proposed by the authors [19]. It is noted that the proposed system in the present study has three main differences as compared to the proposed system in [19]: (a) the present hybrid IDS model consists of anomaly-based and misuse-based intrusion detection modules which can detect simultaneously both insider and cyber attacks of IoT (instead of a specification-based module used in [19]), (b) the proposed misuse-based detection module employs a modified OPF (MOPF) [20] as an efficient graph-based machine learning, and (c) a hybrid feature selection (FS) algorithm based on mutual information and binary gravitational search algorithm (MI-BGSA) [21] is used in the proposed misuse-based intrusion detection module to reduce the number of input attributes of the Internet traffic samples. The simulation results of the proposed system show that the classifier model used in the misuse-based detection module of proposed system outperforms support vector machine (SVM), naïve Bayes (NB), and classification and regression tree (CART) classifiers. Furthermore, the proposed hybrid FS model decreased the number of input features from 41 to 29. So, the computational complexity of the system is reduced.

III. PRELIMINARIES

In this section the foundation of RPL, supervised and unsupervised OPF, MapReduce architecture, and IoT's attacks, as the main concepts which are used in the proposed model, are reviewed briefly. The main symbols used in this paper and their meanings are listed in Table I.

A. RPL

RPL is an IP-based distance vector and hop-by-hop routing protocol, which uses DODAG for routing the packets in 6LoWPAN. RPL supports one-to-one, one-to-many, and many-to-many traffic patterns by enabling different operations such as the unidirectional traffic towards a DODAG root, bidirectional traffic between resource-constrained devices, and bidirectional traffic between resource-constrained devices and the DODAG root [5].

According to the DODAG architecture, for creating an acyclic graph, the nodes are organized into a hierarchical tree structure which has a unique root named 6BR (as destination). Figure 2 shows the general schema of DODAG wherein different nodes connect to each other based on DODAG topology. As seen in Fig. 2, each node has a unique IPv6 address, a specific rank, a set of neighbors and at least one parent in DODAG. Node's rank shows the position of nodes related to other nodes and 6BR [5]. DODAG's nodes use some optimum criteria (e.g., link reliability, latency, hop count, and node energy) for transition the packet toward the root [22]. It is noted that the nodes' rank is specified based on one the mentioned criteria as objective function.

TABLE I. DESCRIPTION OF MAIN SYMBOLS

Symbol	Description
Z	Training set
S^*	Optimum set of prototypes
$f_{\max}$	Path-value function
S	Prototype set
s	A training sample
t	An unlabeled sample
$P^*(t)$	Optimum path from S^* to t
(G_{k-nn})	k-nearest neighbors graph
d_f	Maximum arc weight in G_{k-nn}
$A_k(s)$	Neighbor set of $s \in Z$
R	Set of OPF's roots
L	Set of source nodes in the network
P^i	Set of received packets from the i-th source in a time-slot

DAG Information Object (DIO) is an important control message that is used by RPL for constructing DODAG. This message carries the needed information (e.g., DAG-ID and node's rank) which is required for DODAG construction. In the first stage of DODAG construction, the 6BR assigns "1" to itself as its rank. Then it starts to broadcast the DIO message which includes DAG-ID, objective function (for finding the optimum path), and its rank [4]. When the neighbor nodes receive this message, they select 6BR as their preferred parent. Then, they increment their parents' rank and select it as their rank. Next, they start to broadcast their DIO messages. This process will be continued until the DODAG construction is completed.

B. *Supervised and unsupervised OPF*

Supervised OPF [23] algorithm is a graph-based machine learning method which reduces a pattern recognition problem into an optimal graph partitioning in a given feature space. In the OPF, each training sample is shown as a node in a complete weighted graph. The weighted arcs, which are defined by adjacency relations between samples, link all pairs of nodes in this graph.

Suppose $G=(Z_1,A)$ as a complete weighted graph where Z_1 denotes the training dataset. The samples in the training dataset are represented by the nodes of *G*, and each pair of samples is defined by its arc as $A=Z_1\times Z_1$. In the training phase, some key samples from the training set, called prototypes, should be identified for each class in the classification problem. The closest nodes in the minimum spanning tree of *G* which have different labels in Z_1 are the prototypes of OPF, wherein some prototypes that minimizes the classification error make the optimum set of prototypes ($S^*\subset Z_1$) [23].

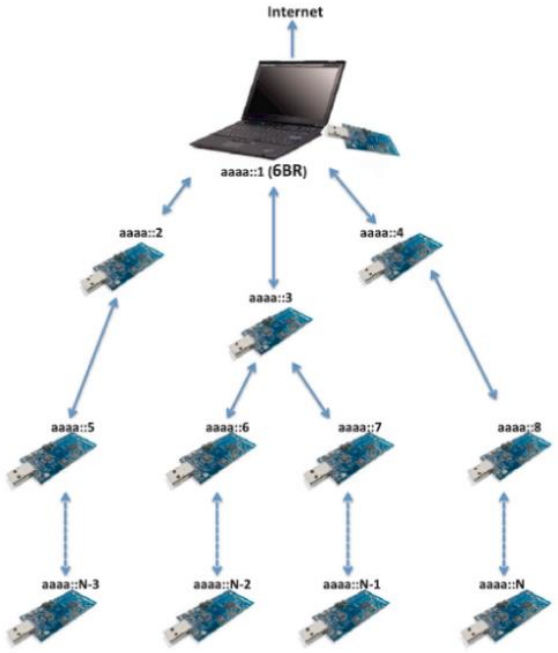

Figure 2. General schema of DODAG wherein each node has a unique IPv6 address and rank [6].

Then, the complete weighted graph will be partitioned into optimum-path trees (OPTs) by a competitive process between prototypes (as the roots of the OPTs) which introduces optimum paths to the remaining nodes of the graph [24]. The nodes of OPT will be strongly connected to their prototypes as compared to other prototypes in the OPF and consequently they have the same label as the OPT's root. Partitioning the OPF for computing OPTs is performed by minimization of $f_{\max}$ which assigns an optimum path $P^*(t)$ from the set of prototypes to every sample $t\in Z_1$ whose minimum cost $C(t)$ is calculated by Eq. (1) [23]:

$$C(t)=\min_{\forall \pi_t\in(Z_1,A)}\{f_{\max}(\pi_t)\} \quad (1)$$

where $f_{\max}$ is a path-value function that assigns a path cost to each path π_t and is calculated using expressions given below [23]:

$$f_{\max}(\langle s\rangle)=\begin{cases}0 & if \quad s\in S\\ +\infty & otherwise\end{cases} \quad (2)$$
$$f_{\max}(\pi_s.\langle s,t\rangle)=\max\{f_{\max}(\pi_s),\mathrm{d(s,t)}\}$$

where *s* is a training sample and *S* represents the prototype set. To classify each unlabeled sample such as *t*, we assumed *t* as a part of the training set. The purpose of the classification phase is to find an optimum path $P^*(t)$ from S^* to *t*, and then labeling *t* with the class of its root. The optimum path can be found incrementally by evaluating the optimum cost as shown in Eq. (3) [23]:

$$C(t)=\min\{\max\{C(s),d(s,t)\}\};\forall s\in Z_1 \quad (3)$$

where $C(s)$ is the minimum cost of *s* and $d(s,t)$ is the Euclidean distance from *t* to *s*. Notably, the learning phase that uses the classification error on the evaluation set can be used for improving the accuracy of OPF. In other words, during the learning phase, the misclassified samples of evaluation samples should be replaced by some non-prototype training samples that their labels are same as misclassified samples. To see more details about the training and classification phases of OPF algorithm, refer to [23].

Unsupervised OPF, which is also called optimum-path forest clustering (OPFC) [25], is almost same as supervised OPF. However, in the OPFC, each sample in the dataset (represented by a feature vector) is shown as a node in the k-nearest neighbors graph (G_{k-nn}) that is connected with its *k* best neighbors in a given feature space [25]. In OPFC, the arcs are weighted by the distance between each pair of nodes and the nodes are weighted by the probability density function (pdf) of each node that is based on the distance between the samples and their k-nearest neighbors [26]. When G_{k-nn} is created, the OPFC algorithm will find one sample (node) at each maximum pdf as a root of a dome or cluster which includes dense samples in the feature space. Then, an OPT will be created from each root to every node in the influence zone (cluster) such that each OPT node

will be strongly connected to its root as compared to other obtained roots in the G_{k-nn} [25]. Notably, as mentioned earlier, each $s \in Z$ (where Z is the training set) is weighted by a pdf that is defined in Eq. (4) [25]:

$$p(s) = \frac{1}{\sqrt{2\pi\sigma^2 |A_k(s)|}} \sum_{\forall t \in A_k(s)} \exp\left(\frac{-d^2(s,t)}{2\sigma^2}\right) \quad (4)$$

where $|A_k(s)| = k$, $\sigma = d_f / 3$, and d_f is the maximum arc weight in G_{k-nn}. It is noted that $A_k(s)$ is the neighbor set of $s \in Z$. Notably, in OPTs construction, OPFC assigns an optimum path to each node $t \in Z$ such that the minimum density value along the path is maximum (Eq. 5) [25]:

$$v(t) = \max_{\forall \pi_t \in (Z, A_k)} \{f_{\min}(\pi_t)\} \quad (5)$$

where $f_{\min}(\pi_t)$ is defined as Eq. (6) [25]:

$$f_{\min}(\langle t \rangle) = \begin{cases} p(t) & s \in S \\ p(t) - \delta & otherwise \end{cases} \quad (6)$$
$$f_{\min}(\pi_s . \langle s,t \rangle) = \min\{f_{\min}(\pi_s), d(s,t)\}$$

where $\delta = \min_{\forall (s,t) \in A_k | p(t) \neq p(s)} |p(t) - p(s)|$ and R is the set of OPF's roots. An OPFC model classifies a new sample to a special cluster (that was created in the OPFC algorithm), by finding a root which provides the optimum path to the new sample. To see more details about the OPFC algorithm refer to [25, 26].

Generally, supervised and unsupervised OPF are simple and fast classifiers which are parameter-independent and originally support multi-class problems [23]. Moreover, OPF does not make any assumption about the shape of classes; so, partial overlapping among the classes can be handled by the OPF [23]. In this study, the OPFC algorithm is used as an anomaly detection engine for detecting the insider attacks in IoT which may be happened by malicious things from WSN sides. Moreover, a new version of supervised OPF, called MOPF [20], is used in a misuse-based detection engine for identifying the cyber (external) attacks from the Internet side. More details about the proposed method will be discussed in Section IV.

C. MapReduce approach

MapReduce approach is an efficient solution for the big data problem [27]. This approach employs algorithms that have parallelism capabilities in a parallel space. In this approach, a big dataset is split to smaller datasets and stored on different machines. These machines process smaller datasets in parallel and finally, the results will be integrated. In the Map phase, input data is partitioned to smaller segments named chunk. Then, they are delivered to some machines (called mappers) that are responsible for the mapping operation [27]. Each mapper converts the content of the chunk to a sequence of key-value pairs by using the user-defined "map" function. On the other hand, in the Reduce phase, MapReduce framework performs sorting based on the keys and collects each key-value pair with the same key and sends them to the reducer node. Then, the user-defined "reduce" function accepts the mediate keys with a set of values representing the dimension of keys and merges the values by converting them to a smaller value [27].

D. IoT's attacks

Generally, IoT is a hybrid network which consists of the Internet and IP-based WSN. So, it can be threatened from both sides of the Internet and IP-based WSN. For example, suppose an end user in the Internet who can access to the things' information of 6LoWPAN, illegally. This user can threat an object in 6LoWPAN by denying its service. As mentioned earlier, we just focus on selective-forwarding and sinkhole attacks as insider attacks in this study. It is noted that in selective-forwarding attacks, which primarily disrupt the routing path, malicious nodes forward packets selectively to remove some packets based on the importance of data or randomly [5]. However, a malicious node represents itself to others as an optimal routing path in sinkhole attacks for attracting nearby nodes to route traffic through it [6].

Moreover, we assumed that the Internet traffic patterns were same as NSL-KDD [28] dataset. NSL-KDD includes 41 different features which are categorized into three categories: (a) basic features; (b) traffic features; and (c) content features. However, noisy NSL-KDD dataset can be employed for evaluating the performance of IDSs in real-world environments [29]. In this way, Karkouch et al. [30] reviewed a set of generic and domain-specific data quality (DQ) dimensions in IoT. For this purpose, the DQ enhancement techniques were presented with an emphasis on data cleaning methods. In this study, the proposed misuse-based detection focuses on detecting anomalous traffic as cyber attacks.

IV. Proposed Model

In this section, the proposed model is introduced. As mentioned earlier, with the aim of providing multi-faceted detection which can identify internal and external attacks, we proposed a flexible model which can detect simultaneously both malicious behavior of 6LoWPAN and the Internet (or LANs) sides. Figure 3 shows the general schema of the proposed model. As seen in Fig. 3, the proposed model consists of two modules:

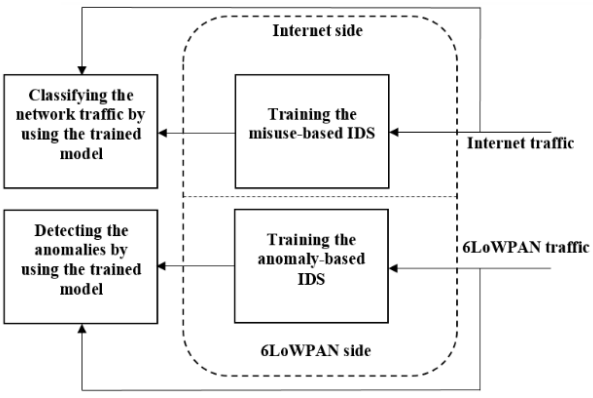

Figure 3. General schema of proposed model.

a) The Internet side module; which is called misuse detection module and provides a misuse-based IDS for classifying the Internet (or LAN) traffic and consequently detecting cyber attacks.

b) 6LoWPAN-side module; which is called anomaly detection module for identification of the insider attacks (i.e., sinkhole and selective-forwarding) by using the anomaly-based IDS that was projected based on 6LoWPAN traffic.

In the 6LoWPAN side, the anomaly detection module creates a sample for each source node by extracting four traffic-related features from the raw received packet of the source node at each time-slot Δw : (a) packet receiving rate; (b) packet dropping rate; (c) average latency; and (d) maximum hop-count. Then, the module projects a clustering model based on an unsupervised OPF algorithm for each source node by using its generated samples. The algorithm selects a cluster (or clusters) including a few samples and then labels the samples as anomalous for each projected model.

Notably, we implemented a WSN based on the RPL routing protocol in this study for simulating the 6LoWPAN functionality (see Section V). Generally, the structure of data packets in the simulations consists of two main parts: (a) data (includes *SrcID*, *SrcTimeStamp*, *SequenceNumber*, *RouterID*, *RouterTimeStamp*, *ReceiveTimeStamp*, and *HopCount*) and (b) data access interfaces which can access and manipulate data (e.g., *incHopCount*( ), *getReceiveTimeStamp*( ), and *getSrcTimeStamp*( )). *SrcID* and *SrcTimeStamp* fields represent the ID of source node and the time of packet sending by the source node, respectively. *RouterID* and *RouterTimeStamp* fields represent the ID of the last router node (before the current node) and its forwarding packet time, respectively. *HopCount* field shows the number of hops taken by the packets and each router node increments it by *incHopCount*( ) function. We assumed that the router node cannot access the data with the aim of manipulating values. In extracting features for producing a new sample for each source node, such as *A*, the packet dropping rate is computed based on the following steps at each time-slot:

Step 1- Sort the received packets from node *A* based on its *SequenceNumber* field.

Step 2- Calculate the sum of the distances between each two consecutive packets (based on *SequenceNumber* field) and return the result as the packet dropping rate. For simplicity, we assumed that each packet is sent only once.

Moreover, other features such as the maximum hop-count and the average latency are computed as follows:

$$\begin{aligned} &\forall i \in L : MaxHopCount = Max\left(packet_j^i.getHopCount(\,)\right) \mid j \in P^i \\ &\forall i \in L : AverageLatency = \\ &\frac{\sum_{j \in P^i} packet_j^i.getReceivingTimeStamp(\,) - packet_j^i.getSrcTimeStamp(\,)}{\left\|P^i\right\|} \end{aligned} \quad (7)$$

where L is the set of source nodes in the network. P^i is the set of received packets from the i-th source in time-slot Δw and $\left\|P^i\right\|$ is the number of its members.

By increasing the number of source nodes, the sequential projection of clustering models will be time-consuming that is not acceptable for a real-time model. The proposed anomaly detection method has the capability of parallelism, because projecting and using clustering models are independent processes. In this study, we inspired from MapReduce approach to improve the speed of projecting models and anomaly detection. In fact, we proposed an anomaly detection method based on the MapReduce architecture. In other words, if an appropriate platform (hardware/software) is prepared, then the model can run in parallel on a distributed space based on the MapReduce architecture. In this approach, the root node sends the values of extracted traffic features of the source nodes to corresponding reducer nodes for anomaly detection. According to this approach, we can also add a new reducer node to the proposed architecture that is a host of the proposed misuse-based detection module. Therefore, by sending the values of the Internet traffic features from root node to the new reducer node, our framework can be employed to detect the cyber attacks from the Internet side. Figure 4 shows the general architecture of the proposed IDS which is based on the MapReduce approach.

As seen in Fig. 4, the root node in anomaly-based detection module (i.e., the 6BR) extracts mentioned traffic-related features from the receiving raw packets (i.e., 6LoWPAN traffic) in each time-slot and creates a new sample for source nodes. Then, it sends the sample's information with key-value pair format to a node (i.e., the reducer) that is responsible to work with a special key. This format includes source ID as the key and feature vector as the value. Then, the reducer node projects a clustering model by using its samples which are received from the mapper node. As mentioned earlier, we assumed that a cluster with fewer samples is anomaly; hence, if the new sample belongs to this cluster, it is classified as anomalous and otherwise, it is classified as a normal sample. So, the reducer node returns a new key-value pair with ⟨SID,Label⟩ format (in response to the incoming key-value pair) to the root node. It is noted that the key and the value are source ID's sample and its label (i.e., anomalous/normal), respectively. Notably, this scenario is also repeated for the misuse-based detection module. As mentioned earlier, the proposed misuse-based intrusion detection is based on MOPF [20] which is supervised graph-based machine learning. In this study, this machine works on NSL-KDD dataset with the aim of simulating the Internet traffic. It is noted that each feature in this dataset may not be relevant to the anomaly detection task. On the other hand, some features which are noise and irrelevant features, can lead to undesirable effect on the performance of MOPF. Hence, for improving the performance of MOPF, selecting an optimum subset of features by using a suitable feature selection can improve the performance of MOPF. In addition to improve the performance of classification models,

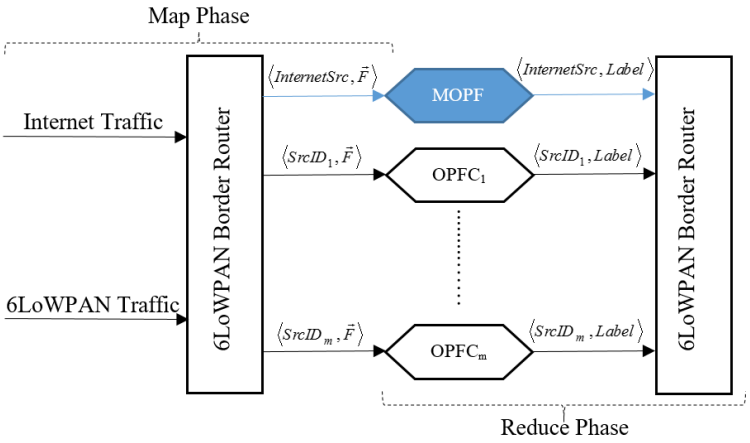

Figure 4. General architecture of IDS model based on the MapReduce approach.

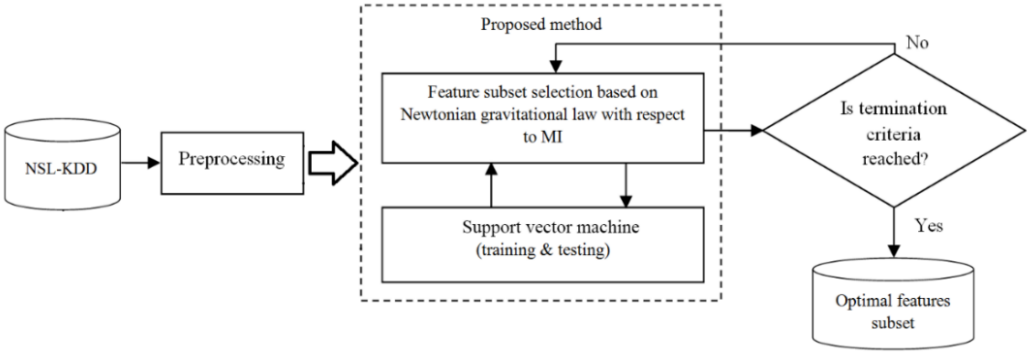

Figure 5. Block diagram of feature selection method based on MI-BGSA.

feature selection is useful for reducing the training and classification time, improving stability against noise, and reducing the measurement and storage requirements [21].

It is noted that the MI-BGSA feature selection method [21] is used for selecting the best subset of NSL-KDD features. MI-BGSA is a population-based heuristic search based on BGSA and MI [21]. This method employs BGSA as a global search method to find the optimum subset of features as wrapper-based feature selection. Moreover, with the aim of improving the quality of selected features, MI is used as a local search for finding the best features among all selected features. The block diagram of MI-BGSA is shown in Fig. 5.

The BGSA in MI-BGSA works as a wrapper-based feature selection method; so, a fitness function is required for evaluating the performance of it. In MI-BGSA, the fitness function is a two-objective function based on DR and FAR. As seen in Fig. 5, SVM as a fast binary classifier is used for evaluating each subset of features in terms of DR and FAR. According to the MI-BGSA feature selection method, the optimum subset of features of NSL-KDD dataset are given in Table II.

TABLE II. SPECIFICATION OF SELECTED FEATURES

# Feature subset	Selected Features
29	{2, 3, 4, 5, 7, 8, 10, 11, 12, 14, 15, 16, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 34, 37, 38, 40}

V. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, the simulation results of the proposed model are presented to show the robustness of the proposed method in IoT's intrusion detection field. Notably, for evaluating the anomaly detection module which is used in 6LoWAN side, we developed a special WSN that is based on the RPL protocol using .Net Framework technology and C#.Net programming. So, a flexible evaluation platform was provided for developing the proposed intrusion detection method and simulating the selective-forwarding and sinkhole attacks, as well. The assumptions in the developed simulator are given in Table III.

Moreover, as mentioned earlier, the NSL-KDD dataset [27] was used instead of the Internet (or LANs) traffic for evaluating the misuse-based detection module in the proposed architecture (Fig. 3). In this study, 7,000 and 3,000 instances were randomly selected from the NSL-KDD as the training and test datasets, respectively. Table IV lists the number of training and test instances. Notably, the NSL-KDD features have significantly different ranges and various resolutions; therefore, most of the classifiers are not able to process data in this format. Therefore, it is essential to normalize the value of each feature to avoid data imbalance. In this study, as seen in Fig. 5, in the preprocessing stage of the feature selection method, we normalized each feature of the NSL-KDD using Eq. (8) [21]:

TABLE III. ASSUMPTIONS IN THE DEVELOPED SIMULATOR IN THIS STUDY

Parameter	Value/Type
Network scale	100 m × 100 m
Routing protocol	RPL
Transmission range	10 m
Packet size	127 bytes
DIO size	24 bytes
Δw	30 sec

TABLE IV. SIZE OF TRAINING AND TEST DATASETS

Type of dataset	Total number of instances	Number of normal instances	Number of anomaly instances
Training	7,000	3,490	3,510
Test	3,000	1,526	1,474

$$x_i = \frac{v_i - \mu}{\sigma}; \quad 1 \le i \le N \tag{8}$$

where N and v_i are the number of samples and the value of i-th sample in the dataset for the given feature, respectively. It is noted, μ and σ are the mean and standard deviation of the given feature, respectively.

In the 6BR implementation, which was based on the MapReduce architecture, the MATLAB server was used as the reducer node for projecting the clustering and classification models with the aim of anomaly-based and misuse-based detection, respectively. We implemented anomaly-based and misuse-based detection methods that were based on the OPFC and MOPF algorithms, respectively, using MATLAB R2014a on a PC with an Intel(R) Core (TM) i5-4460, CPU 3.20 GHz, and 8 GB RAM.

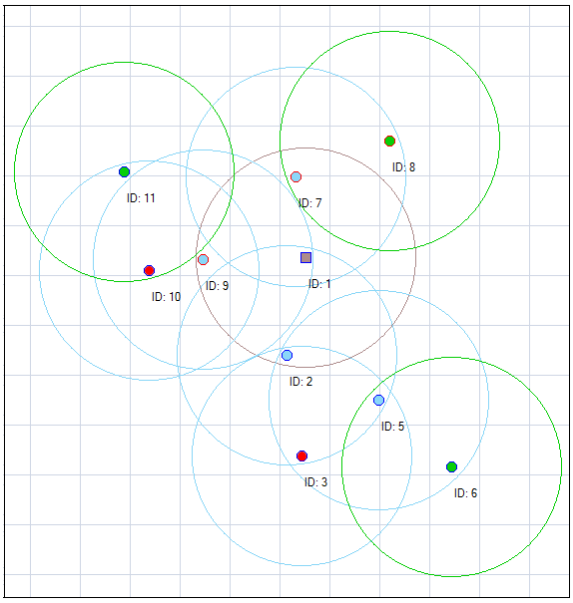

Figure 6. Screenshot of 6LoWPAN simulation.

In this study, the performance of the proposed method was evaluated in terms of DR and FAR. As mentioned earlier, for evaluating the performance of the proposed model in terms of detecting the insider attacks, a simulated WSN's traffic was used in 6BR. The screenshot of simulated WSN is shown in Figure 6. In Fig. 6, sinkhole and selective-forwarding attacks were launched simultaneously by node 3 and node 10, respectively. Through this simulation, the performance of the proposed model was evaluated to deal with the joint occurrence of sinkhole and selective-forwarding attacks in the 6LoWPAN. The assumptions in these simulations are given in Table V. Moreover, the performance of the proposed model, in terms of detecting cyber attacks (as anomaly detection) and specifying the type of detected attacks, was evaluated simultaneously by using NSL-KDD dataset. It is noted that experimental results are reported when feature selection method was used or not.

The performance of the proposed model in the mentioned experiments is reported in Table VI. As seen in Table VI, the performance of the anomaly-based detection module (as compared to other proposed methods such as SVELTE [6]) and misuse-based detection module is appropriate in terms of DR and FAR. The experimental results show that the acceptable performance of the proposed model in detecting both insider and cyber attacks, simultaneously. As seen in Table VI, the performance of the misuse-based detection module when using feature selection unit is worse than the performance of this module when feature selection unit was not used. However, as mentioned earlier, using an FS algorithm in the proposed misuse-based detection module has improved the execution time of MOPF (as a key engine in misuse-based detection module) from 1121 seconds to 837 seconds.

TABLE V. ASSUMPTIONS IN THE 6LOWPAN SIMULATION

Feature	Value
Number of source nodes	3 (IDs: {6, 8, 11})
Number of router nodes	6 (IDs: {2, 3, 5, 7, 9, 10})
Root's ID	{1}
Malicious nodes' ID	{3} (as the sinkhole agent) and {10} (as the selective-forwarding agent)
Simulation time (min)	30
Number of source nodes	3 (IDs: {6, 8, 11})

TABLE VI. PERFORMANCE OF THE PROPOSED MODEL IN SIMULATED IOT

Method	DR (%)	FAR (%)
Anomaly-based detection module	80.95	5.92
Misuse-based detection module without using feature selection	97.88	1.96
Misuse-based detection module when using MI-BGSA feature selection method	95.48	4.39

TABLE VII. PERFORMANCE COMPARISON OF DIFFERENT CLASSIFIERS IN MISUSE DETECTION MODULE

Classifier (used in misuse-based IDS)	DR (%)	FAR (%)
SVM	95.05	2.10
NB	81.00	9.37
CART	97.15	2.75
MOPF	97.88	1.96

Notably, for evaluating the performance of MOPF in detecting cyber attacks in the proposed model, the proposed misuse-based detection module (employing an MOPF) was compared with other misuse-based detection systems that used the following classifiers (instead of MOPF): (a) SVM; (b) CART; and (c) NB (Table VII). These classifiers have been already implemented in MATLAB. As seen in Table VII, the MOPF classifier achieves better DR and FAR as compared to other classifiers (i.e., SVM, NB, and CART).

VI. CONCLUSION

The IoT, as an emerging concept in computer networks is a worldwide network in which all heterogeneous objects around us can connect to the unreliable Internet by using a wide range of technologies (e.g., radio frequency identification, embedded sensors, and miniature actuators). IoT is composed of the Internet and the networks with heterogeneous nodes; so, it provides accessibility to the Internet for all physical objects. Because of the insecure nature of the Internet and WSNs, implementing security mechanisms in IoT is necessary [31]. This paper proposed a novel hybrid architecture as security mechanism for detecting both of insider and cyber attacks in IoT, simultaneously. The proposed model used a real-time anomaly-based intrusion detection module based on unsupervised OPF for detecting insider (internal) attacks which may be happened in 6LoWAPN. On the other hand, a misuse-based intrusion detection engine, which was based on supervised OPF, was responsible for detecting cyber (external) attacks that may be occurred from the Internet (or LANs) side. In addition, a hybrid feature selection algorithm based on mutual information and binary gravitational search algorithm was employed to reduce the number of attributes to the proposed hybrid IDS. The experimental results showed that the classifier model used in the misuse-based detection module of proposed system outperforms SVM, NB, and CART classifiers. Furthermore, the proposed hybrid FS model decreased the number of input features from 41 to 29.

REFERENCES

[1] E. Borgia, "The Internet of things: key features, applications and open issues," Computer Communications, vol. 54, pp. 1-31, Dec. 2014.

[2] M. De Sanctis, E. Cianca, G. Araniti, I. Bisio, and R. Prasad, "Satellite communications supporting Internet of remote things," IEEE Internet of Things Journal, vol. 3, pp. 113-123, Jan. 2016.

[3] S. Raza, Lightweight security solutions for the Internet of things. Ph.D. Thesis, School of Innovation, Design and Engineering, Mälardalen University, Västerås, Sweden, 2013.

[4] T. Kushalnagar, G. Montenegro, and C. Schumacher, "IPv6 over low-power wireless personal area networks (6LoWPANs): overview, assumptions, problem statement, and goals," RFC 4919, 2007.

[5] L. Wallgren, S. Raza, and T. Voigt, "Routing attacks and countermeasures in the RPL-based Internet of things," International Journal of Distributed Sensor Networks, Article ID 794326, pp. 1-11, Jun. 2013.

[6] S. Raza, L. Wallgren, and T. Voigt, "SVELTE: real-time intrusion detection in the Internet of things," Ad Hoc Networks, vol. 11, pp. 2661-2674, Nov. 2013.

[7] J. Granjal, E. Monteiro, and J.S. Silva, "Security in the integration of low-power wireless sensor networks with the Internet: a survey," Ad Hoc Networks, vol. 24, pp. 264-287, Jan. 2015.

[8] A. Botta, W. de Donato, V. Persico, and A. Pescapé, "Integration of cloud computing and Internet of things: a survey," Future Generation Computer Systems, vol. 56, pp. 684-700, Mar. 2016.

[9] C. Stergiou, K.E. Psannis, B.G. Kim, and B. Gupta, "Secure integration of IoT and cloud computing," Future Generation Computer Systems, pp. 1-12, Published online 1 Dec. 2016. DOI: 10.1016/j.future.2016.11.031

[10] A. Ouaddah, H. Mousannif, A.A. Elkalam, and A.A. Ouahman, "Access control in the Internet of things: big challenges and new opportunities," Computer Networks, vol. 112, pp. 237-262, Jan. 2017.

[11] C. Alcaraz, R. Roman, P. Najera, and J. Lopez, "Security of industrial sensor network-based remote substations in the context of the Internet of things," Ad Hoc Networks, vol. 11, pp. 1091-1104, May 2013.

[12] H.R. Ghaeini and N.O. Tippenhauer, "HAMIDS: hierarchical monitoring intrusion detection system for industrial control systems," In: Proceeding of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy, Vienna, Austria, 2016.

[13] D. Airehrour, J. Gutierrez, and S.K. Ray, "Secure routing for internet of things: a survey," Journal of Network and Computer Applications, vol. 66, pp. 198-213, May 2016.

[14] R. Roman, J. Zhou, and J. Lopez, "On the features and challenges of security and privacy in distributed Internet of things," Computer Networks, vol. 57, pp. 2266-2279, Jul. 2013.

[15] F. Alam, R. Mehmood, I. Katib, and A. Albeshri, "Analysis of eight data mining algorithms for smarter Internet of things," Procedia Computer Science, vol. 98, pp. 437-442, 2016.

[16] P. Kasinathan, C. Pastrone, M.A. Spirito, and M. Vinkovits, "Denial-of-service detection in 6LoWPAN based Internet of things," In: Proceedings of 9th International Conference on Wireless and Mobile Computing, Networking and Communications, Lyon, France, 2013.

[17] C. Jun and C. Chi, "Design of complex event-processing IDS in Internet of things," In: Proceedings of the 6th International Conference on Measuring Technology and Mechatronics Automation, Zhangjiajie, China, 2011.

[18] R.H. Weber and E. Studer, "Cybersecurity in the Internet of things: legal aspects," Computer Law & Security Review, vol. 32, pp. 715-728, Oct. 2016.

[19] H. Bostani and M. Sheikhan, "Hybrid of anomaly-based and specification-based IDS for Internet of things using unsupervised OPF based on MapReduce approach," Computer Communications, vol. 98, pp. 52-71, Jan. 2017.

[20] H. Bostani and M. Sheikhan, "Modification of supervised OPF-based intrusion detection systems using unsupervised learning and social network concept," Pattern Recognition, vol. 62, pp. 56-72, Feb. 2017.

[21] H. Bostani and M. Sheikhan, "Hybrid of binary gravitational search algorithm and mutual information for feature selection in intrusion detection systems," Soft Computing, vol. 21, pp. 2307-2324, May 2017.

[22] A. Le, J. Loo, A. Lasebae, M. Aiash, and Y. Luo, "6LoWPAN: a study on QoS security threats and countermeasures using intrusion detection system approach,"

International Journal of Communication Systems, vol. 25, pp. 1189-1212, Aug. 2012.

[23] J.P. Papa, A.X. Falcão, and C.T.N. Suzuki, "Supervised pattern classification based on optimum-path forest," International Journal of Imaging Systems and Technology, vol. 19, pp. 120-131, Jun. 2009.

[24] C.R. Pereira, R.Y.M. Nakamura, K.A.P. Costa, and J.P. Papa, "An optimum-path forest framework for intrusion detection in computer networks," Engineering Applications of Artificial Intelligence, vol. 25, pp. 1226-1234, Sep. 2012.

[25] K.A.P. Costa, L.A.M. Pereira, R.Y.M. Nakamura, C.R. Pereira, J.P. Papa, and A.X. Falcão, "A nature-inspired approach to speed up optimum-path forest clustering and its application to intrusion detection in computer networks," Information Sciences, vol. 294, pp. 95-108, Feb. 2015.

[26] L.M. Rocha, F.A.M. Cappabianco, and A.X. Falcão, "Data clustering as an optimum-path forest problem with applications in image analysis," International Journal of Imaging Systems and Technology, vol. 19, pp. 50-68, Jun. 2009.

[27] J. Dean and S. Ghemawat, "MapReduce: simplified data processing on large clusters," In: Proceeding of 6th Symposium on Operating Systems Design and Implementation, San Francisco, USA, 2004.

[28] M. Tavallaee, E. Bagheri, L. Wei, and A. Ghorbani, "NSL-KDD Data Set" (Available on http://www.unb.ca/research/iscx/dataset/iscx-NSL-KDD-dataset.html), [Accessed on 28 Feb. 2016].

[29] J. Hussain and S. Lalmuanawma, "Feature analysis, evaluation and comparisons of classification algorithms based on noisy intrusion dataset," Procedia Computer Science, vol. 92, pp. 188-198, 2016.

[30] A. Karkouch, H. Mouannif, H. Moatassime, and T. Noel, "Data quality in Internet of things: a state-of-the-art survey," Journal of Network and Computer Applications, vol. 73, pp. 57-81, Sep. 2016.

[31] M. Sheikhan and H. Bostani, "A novel security mechanism for detecting intrusions in Internet of things," In: Proceeding of 8th International Symposium on Telecommunications, Tehran, Iran, 2016.

Mansour Sheikhan is currently an Associate Professor in Electrical Engineering Department of Islamic Azad University-South Tehran Branch. His research interests include speech processing, neural networks, network security, and intelligent systems. He has published about 100 journal papers and 70 conference papers. He is the author of four books in Farsi and seven book chapters for IET, Springer, and Taylor & Francis.

Hamid Bostani is a research assistant in Research Center of Modeling and Optimization in Science and Engineering of Islamic Azad University-South Tehran Branch. He received his B.Sc. and M.Sc. degrees in Computer Engineering from Shiraz and South Tehran Branches in 2008 and 2015, respectively. His research interests are machine learning, artificial intelligence, Internet of things, intrusion detection systems.

IJICTR International Journal of Information & Communication Technology Research

Volume 9 - Number 2 - Spring 2017

Preparation of Papers for IJICTR

Paper Title (use style: *paper title*)

Authors Name/s per 1st Affiliation (Author)
line 1 (of Affiliation): dept. name of organization
line 2: name of organization, acronyms acceptable
line 3: City, Country
line 4: e-mail address if desired

Authors Name/s per 2nd Affiliation (Author)
line 1 (of Affiliation): dept. name of organization
line 2: name of organization, acronyms acceptable
line 3: City, Country
line 4: e-mail address if desired

Abstract—**This electronic document is a "live" template. The various components of your paper [title, text, heads, etc.] are already defined on the style sheet, as illustrated by the portions given in this document. (Abstract should not be longer than 150 words).**

Keywords-component; formatting; style; styling; insert (Include 5 to 10 words .)

I. INTRODUCTION (HEADING 1)

This template, modified in MS Word 2003 and saved as "Word 97-2003 & 6.0/95 – RTF" for the PC, provides authors with most of the formatting specifications needed for preparing electronic versions of their papers. All standard paper components have been specified for three reasons: (1) ease of use when formatting individual papers, (2) automatic compliance to electronic requirements that facilitate the concurrent or later production of electronic products, and (3) conformity of style throughout a conference proceedings. Margins, column widths, line spacing, and type styles are built-in; examples of the type styles are provided throughout this document and are identified in italic type, within parentheses, following the example. Some components, such as multi-leveled equations, graphics, and tables are not prescribed, although the various table text styles are provided. The formatter will need to create these components, incorporating the applicable criteria that follow margins.

(Top:20mm.Bottom:20mm,Left:25mm,Right:25mm)

II. EASE OF USE

Selecting a Template (Heading 2)

First, confirm that you have the correct template for your paper size. This template has been tailored for output on the A4 paper size. If you are using US letter-sized paper, please close this file and download the file for "MSW_USltr_format".

III. PREPARE YOUR PAPER BEFORE STYLING

Space between top of the page and title of the paper has to be 85mm wide with the title centered. Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been

IJICTR International Journal of Information & Communication Technology Research

Volume 9 - Number 2 - Spring 2017

defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

B. Units

- Do not mix complete spellings and abbreviations of units: “Wb/m2” or “webers per square meter”, not “webers/m2”. Spell out units when they appear in text: “. . . a few henries”, not “. . . a few H”.
- Use a zero before decimal points: “0.25”, not “.25”. Use “cm3”, not “cc”. (bullet list)

C. Equations

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation.

IV. USING THE TEMPLATE

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

A. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use “figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract”, will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

B. Figures and Tables

Place figures and tables at the top and bottom of columns. Avoid. placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

ACKNOWLEDGMENT (HEADING 5)

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression, “One of us (R. B. G.) thanks . . .” Instead, try “R. B. G. thanks”. Put sponsor acknowledgments in the unnum-bered footnote on the first page.

REFERENCES

[1] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. (references)

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[3] I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

Note: Please submit your proposals via IJICT website at: http://jounal.itrc.ac.ir