International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 4 Issue: 2
ISSN: 2454-4248
310 – 314

Probablistic Queuing Model to Detect Covert Channel Communication in Wireless Sensor Network

A. Swaminathan
Research Scholar, Department of Computer Science and Engineering
CEG, Anna University
Chennai, India
linuxswami@yahoo.com

Dr. P. Vivekanandan
Professor of Eminence
A.C Tech, Anna University
Chennai, India
vivek@annauniv.edu

***Abstract*—** The communication through a covert channel is a major milestone in a secret communication. The covert communication has attracted many hackers to find an alternate way to exfiltrate/infiltrate information secretly and for an effective utilization of the network resources. The covert channel in a wireless sensor network (WSN) is an emerging problem in the area of the sensor communication. A novel detection technique is proposed using probabilistic fluid and Random early detection models to detect the covert channel communication in the wireless sensor network. The proposed methodology is capable of dealing with both storage channel, timing channel and also it ables to detect the hybrid channels. The proposed methodology is experimentally tested and the result clearly shows that this method yields better result than the existing methodologies**.**

Keywords- *Covert channel; covert communication; WSN; Probabilistic fluid; Random early detection; data exfiltration; data infiltration.*

*****

I. INTRODUCTION

Recent advancement in Internet of things (IoT) attracts many researchers toward the wireless sensor network. The wireless sensor network plays a major role in the field of agriculture, military, weather forecasting etc. An advancement of present day technology uses this sensor technology as hybrid fashion. These sensors were widely used in the automated cars, drones and special purpose Radio Frequency (RF) card tags in order to automate things and speak via connected devices through a cloud environment. Though the technology grows and usage increases drastically but the flaw and security drawbacks are also increasing side by side.

The covert communication is the one such attack widely used in all infrastructures which are loosely or tightly connected. Nowadays, the covert communication through the connected sensor nodes and network are popular and emerging problem in the area of wireless sensor network. The covert communication possesses a multiple channel deployment in sensor nodes which intentionally compromise a secured and legitimate protocol for communication. It is very hard to detect such kind of the covert communication and the covert channel attacks. A new method for the covert channel attack is proposed. The covert channel uses IP for most of its communication channel. The identification of the covert communication by dissecting and inspecting data packets were hard task. S. Z. Goher et al summarized some literature to detect the covert communication [1].

II. RELATED WORK

A novel scale free network to detect the covert channel communication has been proposed [1]. Here, the author has proposed a high degree searching method and uses Principle Component analysis (PCA) for entropy based subspace method to detect the covert communication [2]. A novel TCP Markov model to detect the covert channels has been proposed. Here, the key idea is to analyze the TCP flag sets for any modification or deviation. Then the TCP properties were modeled using Markov property [3]. The author uses Kull back - Leiblei method to predict the survival of the covert anomalies.

Optimization and simulation techniques for the covert channels are proposed. The author proposes the covert channel detection system along with the firewall and IDS. The key idea is the author integrates all the security solutions into a single package to detect the covert communication happening within and on the connected networks [4]. A novel Initial Sequence Number (ISN) generation model based on the operating system and its associated process has been proposed [5]. Then the author applies some statistical methods to detect whether the packet generated belongs to a particular network or not. If any significant deviation is recorded during the packet arrival then the particular packet generated is flagged as the covert anomaly.

An adaptive detection scheme to detect http based covert communication is proposed [6]. The author introduces a new system called DUMONT which is highly reliable to detect the outbound covert communication. The key idea of DUMONT system is to make initial system profiles using various host and system features.

A new framework called SIDD to detect data leakage is proposed [7]. SIDD framework consists of three main models i) the covert channel detection module ii) an application level detection and iii) signature based detection. The author modeled the traffic flow and processed the model using SIDD framework.

ISSN: 2454-4248
310 – 314

An entropy based approach to detect the covert channel has been proposed [8]. The key idea is to estimate the interval arrival time of packet using conditional correlated entropy. Once the entropy for each packet gets generated, then it uses Support Vector Machine (SVM) classifier to classify the profiles.

SVM based system to detect the covert communication has been proposed [9]. The author employs the machine learning algorithm to detect whether the traffic pattern belongs to normal or abnormal. The author performed some significant works to train and test the traffic pattern. However, the mechanism proposed by the author is generic and applicable only to the trained network.

It is clearly shown from the literature that signature based and protocol specific detection mechanism is not capable to detect these kinds of the covert attacks. Since the covert channel uses legitimate protocols to establish the communication between the victim and attacker. Hence, it is very tough to detect such a communication. Security tools like IDS, IPS, firewall sitting at choke point of enterprises can inspect the anomalies and filter out the suspicious traffic. When the generated traffic is from the legitimate system which resides within the enterprises, these security systems fail. Hence, an effective security system is in need to detect such a kind of communication. In this work, a novel covert channel communication system is proposed using a mathematical queuing model. The model works fine and highly reliable to detect any form of the covert communication.

III. THE PROPOSED MODEL

The proposed model is pictorially represented in Figure 1. In the cluster phase, the cluster head is selected based on the energy level or certain parameter which influences the node capability in terms of power, resource, node ability to move frequently. However, the heterogeneous fashion of the deployment makes the process of cluster head selection very tedious. Hence, the acceleration and velocity are major terms considered for a node which is in mobility. In this paper, the velocity estimation is carried out in order to estimate the node mobility based on the formulae as stated in the equation (2).

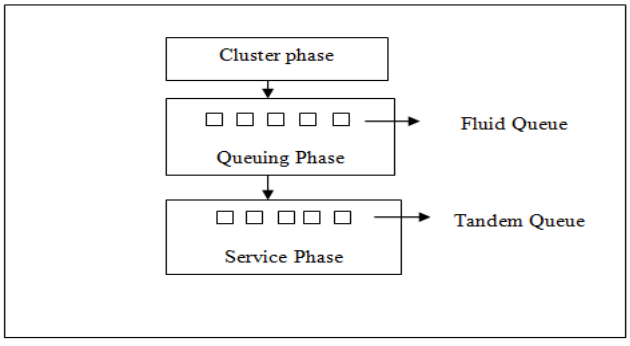

Figure 1. The Proposed Model

Figure 2 shows the flow diagram of the proposed model. The flow starts with the incoming fluid queue where the queue model is deployed and monitored before the base station. The extracted features are correlated in order to obtain the feature set relation and followed by the algorithmic processing is depicted.

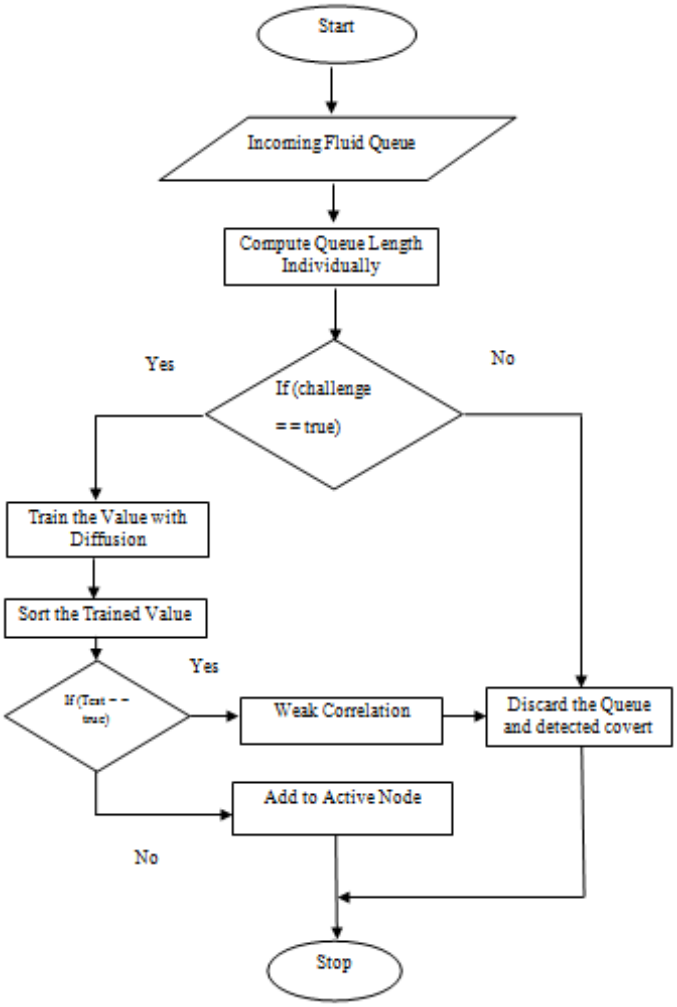

Figure 2. System Flow Diagram of the Proposed Model

A. *The Clustering Phase*

Let us assume that the proposed test bed consists of 100 sensor nodes. The K-means clustering algorithm is applied to separate the nodes with the identical features in terms of values. Hence in this research work such a process is carried out in order to filter out the outliers among the nodes. The sensor nodes are grouped into different clusters by fixing the centroid using the default nominal function. Once the centroid is fixed, the nodes with the identical features are grouped [11]. Then the process of separating outliers is carried out using linear separation methodology among the clusters [12].

 ISSN: 2454-4248
310 – 314

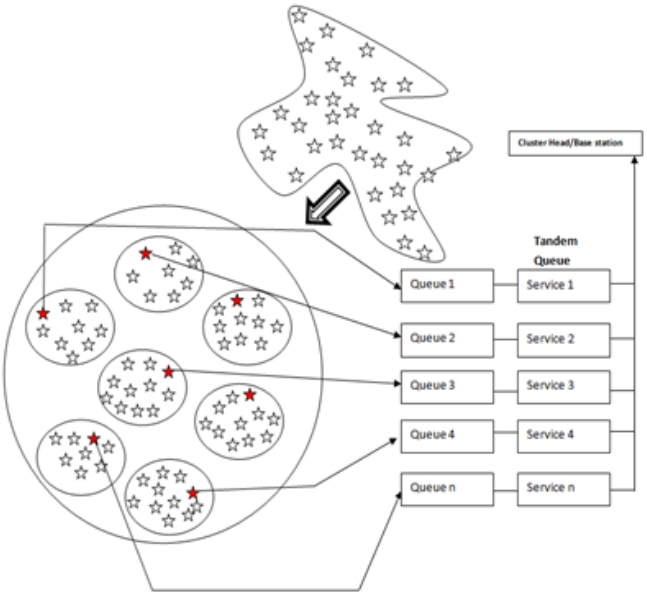

Figure 3. The Clustering Model

Now the cluster assigns the cluster head based on the above mentioned method. Then the meta information among the nodes are fetched or saved in the head node based on the node id. The meta information includes node id, sensor type, resources and sensor information. After the successful clustering, the clusters start to index the coordinating nodes using metadata.

B. *The Fluid Queues*

The cluster head starts to communicate with the base station using the separate channel which is deployed using fluid queues. The fluid queues are responsible to control the entire flow within the defined buffer. A clear illustration on the buffer, service rate, arrival rate in the queue are given in the equation (1). Let us assume that the network is a large container C with an infinite room, where the nodes are represented as n within the container and clustered into cluster head (*Cl*). The buffer is represented as B in the queue as in the equation (1), where the behavior is Jockeying method and a connection tunnel T is built from each cluster head to base station (BS). The busy period distribution $W^*(S)$ is computed as follows:

(1)

where λ and μ are arrival rate and service rate respectively, α and β are fluctuating parameters. Let the traffic generated from each node will be redirected to *cl* and *cl* transmits the data to base station through the tunnel T. Each arrival rate from *cl* to BS is stated as:

$$\frac{dX(t)}{dt} = \begin{cases} cl_n \ if \ X(t) > 0 \\ \max(cl_n, 0) \ if \ X(t) = 0 \end{cases} \quad (2)$$

where X is the process and t is the fluid level time. The traffic is transmitted through T which controls the flow at B and it is a continuous time Markov chain as stated in equation (2). This transmitted traffic will possess a solid data structure with a point scale probability. These probabilistic data structures have the significant value and q is transmitted by a node. A run time challenge between the particular node n, cluster head *cl* and *cl* to BS is generated in order to overcome the node spoofing attacks.

These fluid queues services are evaluated based on the Random early detection model using the phase type distribution method. Here, each fluid queue is individually processed by Random early detection model at a constant service rate μ and arrival rate λ [10]. The processing type of the model is represented in the Figure 2. The average rate and level of the fluid queues are represented as δ [10]. Random early detection model is widely used in order to optimize the fluid queue flow rate and also to avoid the congestion. It keeps track of Buffer B size and verifies the run time challenge components. The average threshold estimation can be calculated using the Carl Pearson correlation. From the assumption, the model doesn't suffer any busy period or waiting period in order to avoid any sort of delay and congestion within the network.

C. *Tandem Queue*

Tandem Queue represented in fig 2 is the sequence of queuing models which possess the stochastic behavior of the networks of single server queues when the successive service times of a given customer are highly correlated. The study is conducted in two particular cases: 1) networks in heavy traffic and 2) networks in which all the successive service times have the same value (for a given customer) in order to avoid the possibility of breaking up the busy periods. It is shown that how the local queuing delay (for an arbitrary customer) can be derived through an equivalent tandem queue on the condition that one other local queuing delay is added to the jitter delay due to the independence of the partial traffic streams.

D. *The Training Phase*

In the training phase, the nodes were trained based on the various network features. Some of the prominent features from the system and network were taken into considerations. Network features include IP packets, packet and byte transmitted, received, data packets, control packets, management packets, broadcast packets, Source IP, Source MAC, Destination IP, Destination MAC, TCP flag, flag sequence, TCP streams, TCP stream flow, Gateway address. Node level feature includes RAM usage, power in battery, RAM utilization, energy spent at packet level, energy consumed by the cluster head, average transmission ratio, transmission rate, transmission probability, next hop address, node id, node address (MAC), next hop id. All the systems and network level features were taken and some of the prominent, significant variant among the features were considered into the selection. These features were trained to the system with the

 ISSN: 2454-4248
310 – 314

calculated threshold using the probabilistic schemes and manual weights. The weight factors were manually assigned based on some of the influencing parameters such as bandwidth, node count etc. These parameters are on demand, hence automating weight or threshold or entropy based methods may lead to a failure whenever there is a change in these parameters. If the threshold value leads the variation in the network features, the network is reformed. This leads to false positive in detection. To train the system, the presented features were taken into consideration and the network setup, on-going traffic information, node information etc were monitored and recorded for more than 10 days. This leads to active stabilization of the setup.

Algorithm for the Training phase

For all the nodes in the network
Extract the possible system and network features
Apply fluid queue
Transform each feature components into a probabilistic data structure
Apply Random early detection
Verify the challenge
if challenge accepted
Enqueue the value
Store the trained values
End

E. The Detection Phase

The process carried out for the feature extraction and selection are explained in the section "training phase" which is adapted and used in the testing phase for generating the probabilities. These probabilities for both the system and network features were correlated with the trained and stored values. If the correlation between the live feed and trained values are fluctuated then the correlation results are weak, there may be a chance for the covert communication.

Algorithm for the Detection phase

For all the nodes in the network
Extract the feature values
Apply the correlation (live feed, trained feed)
If the correlation < threshold
Label: The Covert communication
Else if the correlation>threshold
Label: Normal transmission

F. The Validation Phase

In order to explain in deep, a system feature "RAM utilization" and a network feature "packet transmitted" were taken into consideration for the evaluation. In the training phase, the probabilistic value is generated for both the features and it is stored.

In the detection phase, the probabilistic value is generated and correlated with the trained values. If the correlation of the packet transmitted exceeds the actual training values, then the system raises the flag as it is the covert communication.

IV. REULTS AND DISCUSSION

The proposed approach is systematically implemented in a simulated test bed using python 2.7. The experiments were carried out in a controlled environment with the maximum number of 100 nodes. The general behavioral properties of sensors were clearly defined in a config file. Figure 4 to Figure 6 show the experimental setup and the simulation initialization is carried out for the experimentation. The general traffic is simulated using inetsim and fakenet packages. Ostinato simulator is also used to generate general TCP and UDP traffic. The communication path between the nodes was analyzed by tapping the general packet info over the simulated interfaces.

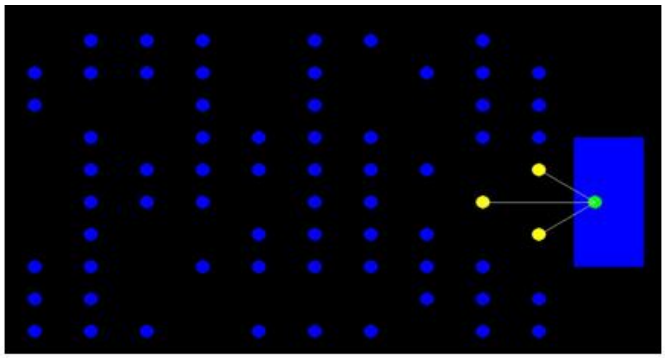

Figure 4. Simulated Test bed

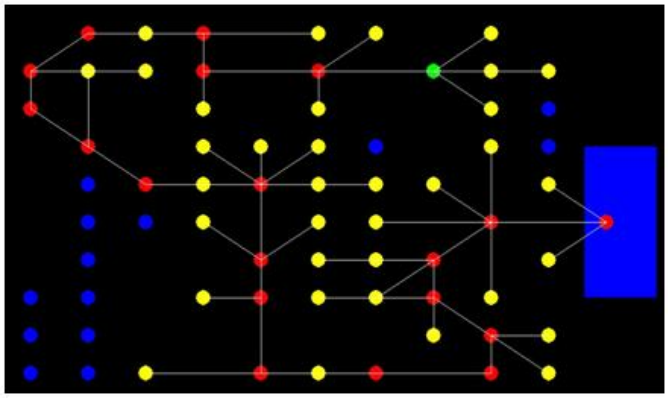

Figure 5. Simulation Results for the Cluster head and its metrics

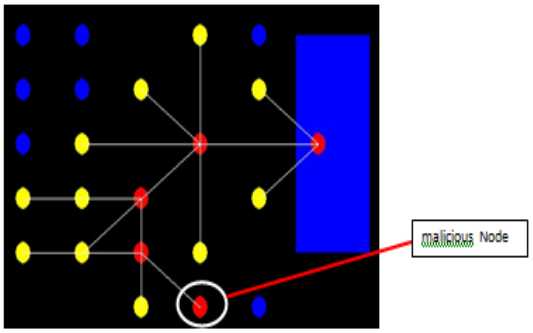

Figure 6.Detection of Malicious Node trying to establish the covert channel

 ISSN: 2454-4248
310 – 314

V. THE PERFORMANCE ANALYSIS

The performance of the proposed model was tested using two optimal parameters. From the Table 1 and Figure 7, one can conclude that the evaluation of a detection rate is better. The observation was made from the simulated test bed. It is clearly shown that the false positive rate of detecting the covert communication was significantly very low and detection results were also promising.

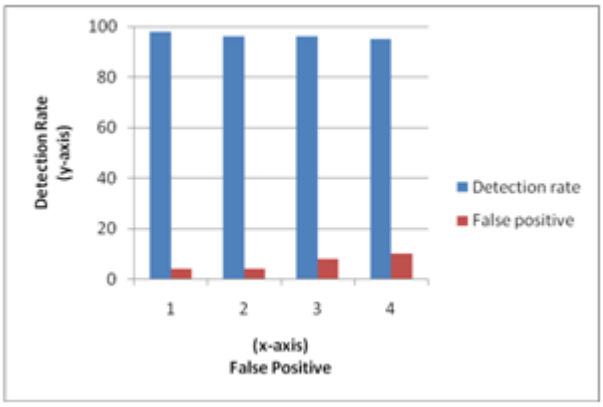

Figure 7.Detection Rate Vs False Positive Rate

TABLE I. DETECTION RATE VS FALSE POSITIVE RATE (%)

S.No	Number of Nodes	Detection scheme	Detection rate	False positive
1	25	Queuing model	98	4
2	50	Queuing model	96	4
3	75	Queuing model	96	8
4	100	Queuing model	95	10

VI. CONCLUSION

A novel queuing model based technique to detect the covert channel communication is proposed. From the experiment, the results obtained are promising and method is more reliable and adoptable. It can be easily implemented in both homogenous and heterogeneous sensor network. The presented model is robust in detecting hybrid covert channel which is a challenging problem in sensor network to detect inside the node communication through the covert channels. But the advancement in technologies and Cloudification of IoT devices is an emerging feature in the field of the sensor network. Adoption of protocols for such an advanced technology leads to the security weakness. Attacker exploits such vulnerabilities possessed by these protocols. Hence in future, a new technique or enhancement of the presented model can be developed to detect the covert communication in future protocols.

REFERENCES

[1] S. Z. Goher, B. Javed and N. A. Saqib, "Covert channel detection: A survey based analysis," High Capacity Optical Networks and Emerging/Enabling Technologies, Istanbul, 2012, pp. 057-065.

[2] L. Lan, X. Linglin and W. Wenhong, "Covert Channel Detection Based on Scale-Free Networks Theory," Computational Intelligence and Design, 2009. ISCID '09. Second International Symposium on, Changsha, 2009, pp. 378-380.doi: 10.1109/ISCID.2009.103

[3] J. Zhai, G. Liu and Y. Dai, "A Covert Channel Detection Algorithm Based on TCP Markov Model," 2010 International Conference on Multimedia Information Networking and Security, Nanjing, Jiangsu, 2010, pp. 893-897.

[4] L. Frikha, Z. Trabelsi and S. Tabbane, "Simulation, optimisation and integration of Covert Channels, Intrusion Detection and packet filtering systems," 2009 Global Information Infrastructure Symposium, Hammemet, 2009, pp. 1-4. doi: 10.1109/GIIS.2009.5307102.

[5] E. Tumoian and M. Anikeev, "Network Based Detection of Passive Covert Channels in TCP/IP," The IEEE Conference on Local Computer Networks 30th Anniversary (LCN'05)*l*, Sydney, NSW, 2005, pp. 802-809.

[6] G. Schwenk and K. Rieck, "Adaptive Detection of Covert Communication in HTTP Requests," Computer Network Defense (EC2ND), 2011 Seventh European Conference on, Gothenburg, 2011, pp. 25-32.

[7] Yali Liu, C. Corbett, Ken Chiang, R. Archibald, B. Mukherjee and D. Ghosal, "SIDD: A Framework for Detecting Sensitive Data Exfiltration by an Insider Attack," System Sciences, 2009. HICSS '09. 42nd Hawaii International Conference on, Big Island, HI, 2009, pp. 1-10.

[8] L. Shrestha, M. Hempel, F. Rezaei and H. Sharif, "Leveraging Statistical Feature Points for Generalized Detection of Covert Timing Channels," 2014 IEEE Military Communications Conference, Baltimore, MD, 2014, pp. 7-11.

[9] P. L. Shrestha, M. Hempel, F. Rezaei and H. Sharif, "A Support Vector Machine-Based Framework for Detection of Covert Timing Channels," in IEEE Transactions on Dependable and Secure Computing, vol. 13, no. 2, pp. 274-283, March-April 1 2016.

[10] https://en.wikipedia.org/wiki/Fluid_queue.

[11] L. Tan and S. Tang, "Energy Harvesting Wireless Sensor Node With Temporal Death: Novel Models and Analyses," in IEEE/ACM Transactions on Networking, vol. 25, no. 2, pp. 896-909, April 2017.

[12] M. Gautam and V. Sejwar, "A brief review of WSN energy and routing protocols," 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, 2017, pp. 849-855.