Collaborative adversary nodes learning on the logs of IoT devices in an IoT network

Sandhya Aneja, Melanie Ang Xuan En, Nagender Aneja
Faculty of Integrated Technologies, Universiti Brunei Darussalam
School of Digital Science, Universiti Brunei Darussalam

COMSNETS 2022
14th International Conference on COMmunication Systems & NETworkS
January 3 - 8 | Hybrid Conference | Bengaluru, India

Presentation Preview

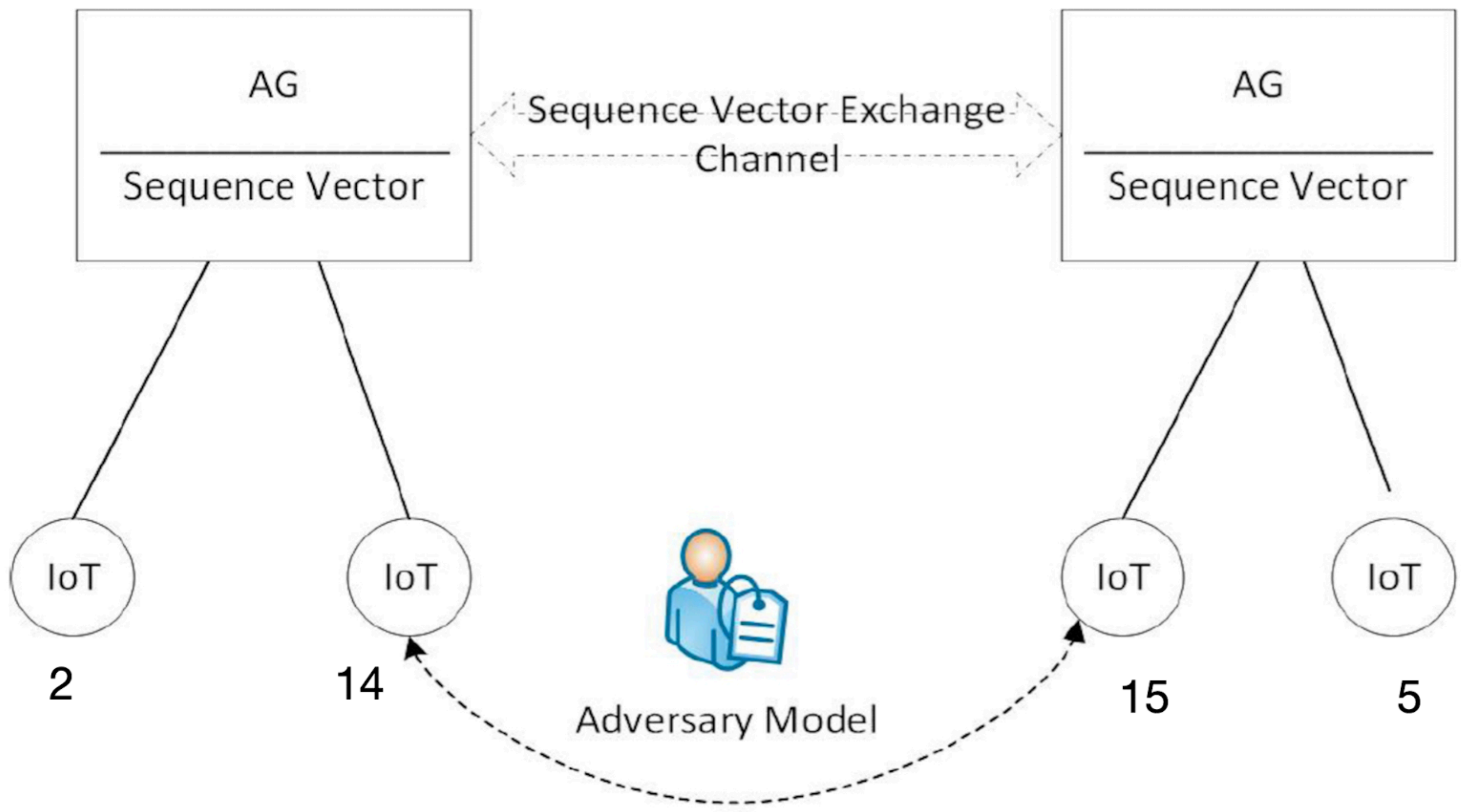
- Details on the collaborative attack- Deep Adversary Architecture
- System and Attack Model
- Predictive model to detect adversarial attack
- Experiment Setup, Results, and Analysis
- Simulated traffic- for example, what % of traffic constitutes the attack
- RNN to analyze the network traffic and detect collaborative adversarial attacks

Deep Adversary Architecture

- The problem of detecting a set of malicious nodes in an IoT network by analysing the networks logs at the gateways that are between the IoT devices and the servers

- One of the scenarios is wherein IoT devices in two different LANs or locations can collaborate using a high transmission antenna to exchange data say temperature, pressure, and humidity.

- The collaborating IoT devices can then upload the distant location data to the server.



IP Add	MAC Add	S/D
15	aa-bb-cc-dd-10	S
2	aa-bb-cc-dd-11	D

ARP Table of IoT Device -14

IP_A	MAC_A	S/D
14	aa-bb-cc-dd-22	S
5	aa-bb-cc-dd-21	D

ARP Table of IoT Device -15

- The collaborating IoT devices have their high range channel collaborators into ARP table. The trace driven event log comprises the communication across the IoT network.

Details on the collaborative attack

- While capturing data through IoT, metadata can also be captured to apply AI techniques for IoT network security.
- Traditional AI techniques were about centralized data.
- Federated learning (FL) model trained from distributed systems over the cloud.
- Here interesting observation for FL is that the learned model over distributed systems can be secured like other encrypted numbers communicated over the Internet [6].
- A simple/low complexity resource allocation algorithm is proposed for a wireless network to support multiple FL groups [7].
- IoT devices may be compromised. We propose in this paper to analyze network traffic logs of IoT devices distributed in a network behind the application gateways.
- This network traffic logged at application gateways can be used to identify compromised devices as well as collaborative adversaries.

Sensors generate tremendous amount of data and analyzing this data is nearly impossible manually. Automating the analysis by applying AI/ML is latest trend and this paper goes a step further and not only suggests analysis of sensor logs for potential threats but also have themselves created training model based on negative loss likelihood model

System Model

This paper presents an approach, called Adversary Learning (AdLIoTLog) to use deep learning on IoT data to detect behavior of malicious, collaborating IoT devices.

AdIoTLog uses packet event sequence of protocols such as TCP, UDP, HTTP to identify collaborating nodes in an IoT network that can connect through hidden channels for adversary behavior to other nodes.

Let S be a set of p malicious nodes represented by $m_1; m_2; \ldots; m_p$. Let S_1 is the set of events of m_1 malicious nodes and S_p is the set of events of m_p malicious nodes. The node m_1 perform l sequence of events $e_{11} \rightarrow e_{12}, e_{12} \rightarrow e_{13}, \ldots, e_{1l-1} \rightarrow e_{1l}$. The node m_p perform t sequence of events $e_{p1} \rightarrow e_{p2}, e_{p2} \rightarrow e_{p3}, \ldots, e_{pt-1} \rightarrow e_{pt}$.

Therefore it is required to learn a function that can be used
for any given source malicious events of ms to predict the targeted
coordinated malicious events of mt. AdIoTLog collects
IoT log over the LAN therefore AdIoTLog comprises of let
m1;m2; ::mp nodes over one application gateway say AG1
while n1; n2; ::nq nodes over another application gateway say AG2. AdIoTLog computes the probability of possible events in
the sequence
P($(m_p : e_{p1} \rightarrow e_{p2}, e_{p2} \rightarrow e_{p3} \ldots e_{pl-1} \rightarrow e_{pl}) \rightarrow (n_q : e_{q1} \rightarrow e_{q2}, e_{q2} \rightarrow e_{q3} \ldots e_{ql-1} \rightarrow e_{ql})$.

Attack Model

- The attack model - When two hosts communicate with one another, it does not always indicate malicious activity; however, if those nodes are not in range, then it indicates malicious behavior, which is modelled as nodes in two distinct AGs.

- The trace is generated using ns-2. The two scenarios differ by having one additional communication. There is comparison with other baseline i.e. networks with hidden channel and network without hidden channel were input to the model

What constitutes communication over a "hidden" channel?

Communications over a "hidden" channel include data packets as well as control packets, such as ARP packets, TCP/UDP packets, and other sorts of packets, among other things.

If two hosts are not in range of one another, it is not expected that those nodes will interact; nevertheless, this identification is not straightforward.

Two malicious IoT devices communicating with high transmission power. if direct communication in the IoT network is allowed, then it would be through AGs

- The logged network events were paired following the order of timestamps of network events one after the other including the collaborative attack itself. The input file included 12,236 network sequence pairs with 4170 unique elements that comprised different types of packets, protocols, sequence numbers, and flags.

Predictive model to detect adversarial attack

The
TCP packet, UDP packet, and
HTTP packet were considered
in different contexts. This
potentially can reduce the gap
between training and inference
by training the model to handle
the situation, which will
appear during test time.

Recurrent Neural
Network models - LSTM, GRU etc.
are learned with less
execution time and better
predicting for network problems
in addition to language
translation, emotion detection,
and fake news detection
problems.

Literature review

// Robust Model
Almiani et al. [16] for intrusion detection system in an IoT network was trained on the NSL-KDD dataset for different types of attacks.

// Distributed Model
Shen et al. [4] for collaborative nodes trying to use vulnerabilities of intrusion protection system also used RNN. The negative shift in model prediction was used to detect the attack, however, the attack considered was not over distributed machines

// Authentication
A similar method is presented at the application gateway to authenticate the IoT device by analyzing the 212 features like TCP src port and TCP dst port from the packet headers of IoT devices in the logged network traffic.

Simulated data and the Simulation process

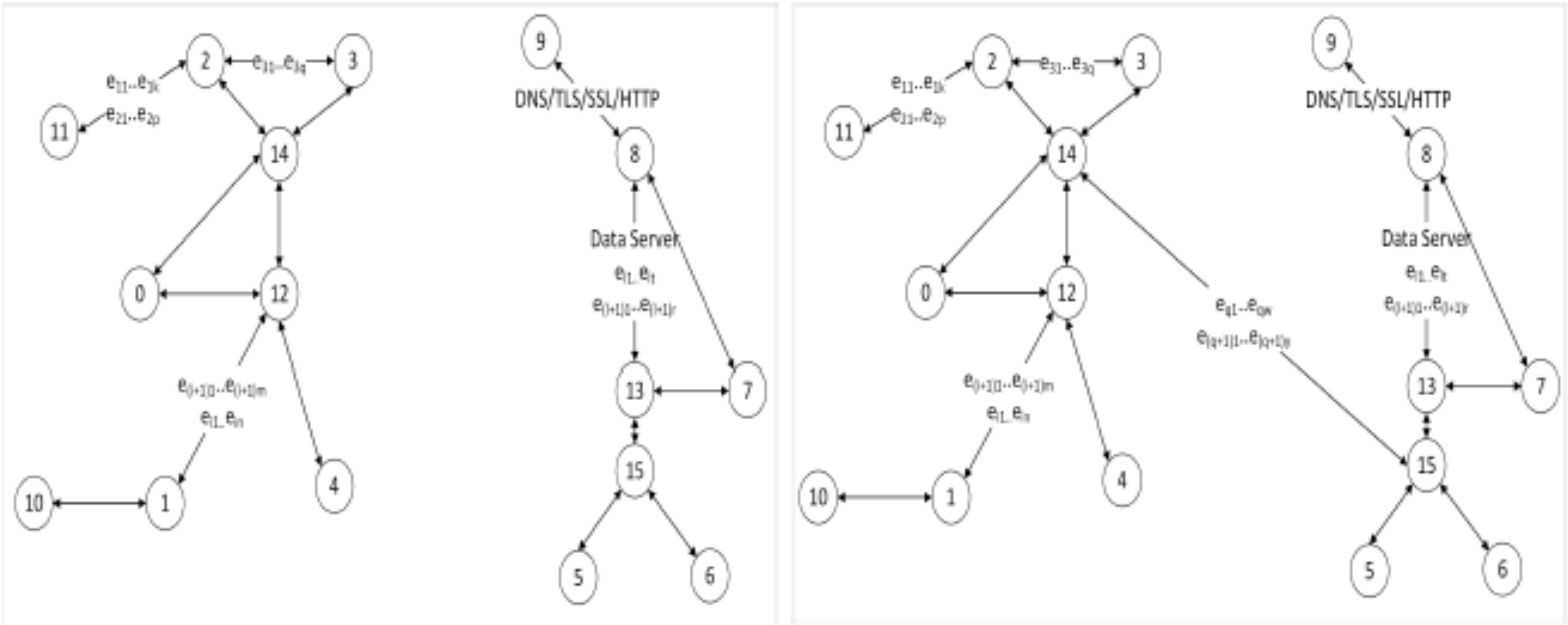


Fig. 2: (a) IoT network without collaborating nodes in ns-2 (b) IoT network with collaborating nodes in ns-2

Eight UDP/TCP communication node pairs were used with a 1500 byte packet at the rate of 1 mbps to generate the data.

Adversary Node pair(14,15) used High Range Antenna identified in trace through ARP protocol at link layer

The training dataset in ns-2 was created with 16 nodes.
For example, node pair (14, 2) was simulated for node 14 to upload data to node 2.

In the first case, when there is no collaboration, node 14 will upload the data to node 2.

In the second case, when nodes can collaborate using hidden channel, node 14 will upload the data to node 15.

- *One pair (14,15) was collaborating adversary nodes which means12.5\% of simulated traffic constitutes the attack.*

Experiment Setup, Results, And Analysis

```
// The BLEU score
The BLEU score was computed by
comparing the predicted network event
sequence with the
ground truth network sequence using 1-gram
(single words).
```

- Definition: Let $tp_1, tp_2, \ldots, tp_n$ testing pairs and are testing pairs and their respective bleu scores are $b_1, b_2, \ldots, b_n$ then accuracy of model output will be $\frac{\sum_{i=1}^{n} b_i}{n}$.

- The generated sequence numbers were much easier to keep track of relatively small, predictable numbe rather than the actual numbers.

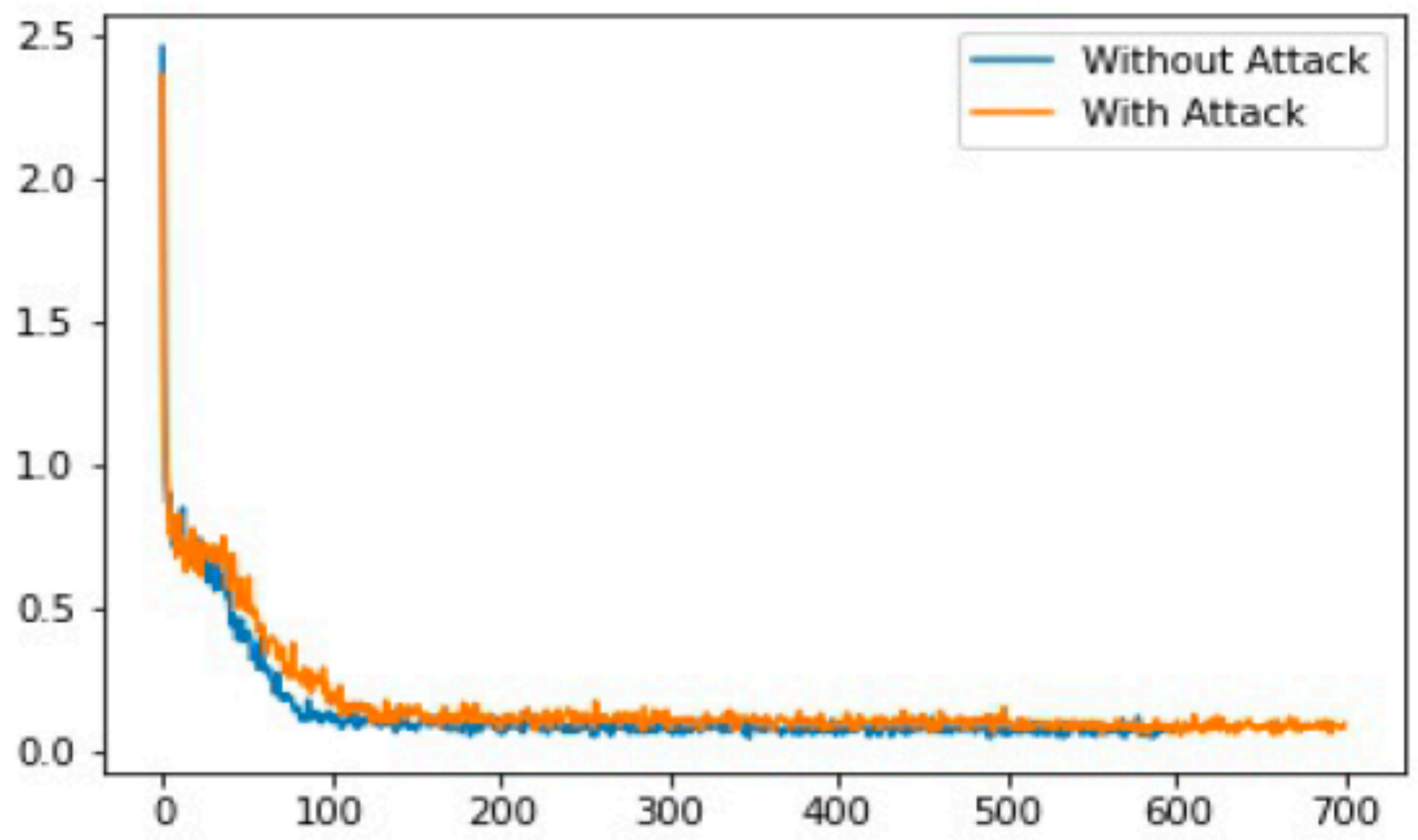


Fig. 3: Variation of model training (NLL Loss) x 100 iteration

TABLE I: Hyper parameters of model

Network	No of hidden layer	No of iterations	Learning rate	Hidden Layer size	optimizer
IoT ns-2	1	70,000	0.01-0.0001	256	SGD

Accuracy with collaborative attack	Accuracy without collaborative attack
89-95%	91-98%

```
The input file included
12,236 network sequence pairs
with 4170 unique elements
that comprise different types
of packets, protocols,
sequence
numbers, and flags.
```

- If we present Node tuple in A as (node, actual data server, pred data server), for tuple (14, 2, 15), the actual node event was by 2 while model predicted node 15 rather than node 2 for source 14.

Performance comparison of their proposed method

The logs are generated using ns-2 simulations, and an existing ML model is used to analyze the logs with and without the malicious nodes. The authors claim that the decrease in accuracy indicates the presence of malicious nodes.

How robust is the proposed algorithm for different IoT traffic patterns? A comprehensive trace-driven simulation is required to evaluate the efficacy of the proposed solution.

- *A network protocol fixes the packet format in the network traffic of the devices.*
- *Model was found more robust for UDP packets in comparison to TCP and HTTP packets.*

Accuracy with collaborative attack	Accuracy without collaborative attack
89-95%	91-98%

THANKS AND QUESTIONS