Mini Project – 1

Shreya Bhatia
110619150

Introduction

We have been given the network flow traffic data (Dataset 1). It basically contains details about the packets being transferred between source and destination at a particular time stamp. We need to analyze the dataset for anomalies and other network attacks like intrusion, failure. In a network, failure of a node occurs when it receives large number of requests which it cannot handle. We are going to do the temporal analysis of the data on day basis to check for any abnormal pattern for the requests in the network. Also we will be doing node analysis to verify what is its behavior and is it deviating from the regular pattern.
We will also be studying behavior between two nodes in a network and analyze it over time.

Problem Statement

The task is to analyze the network flow traffic data. For network Traffic Data 1, to look for the possible anomalies in the network. To represent the data such that we can look for abnormal patterns. Also doing the temporal analysis and looking for possible intrusion or failure in the network.

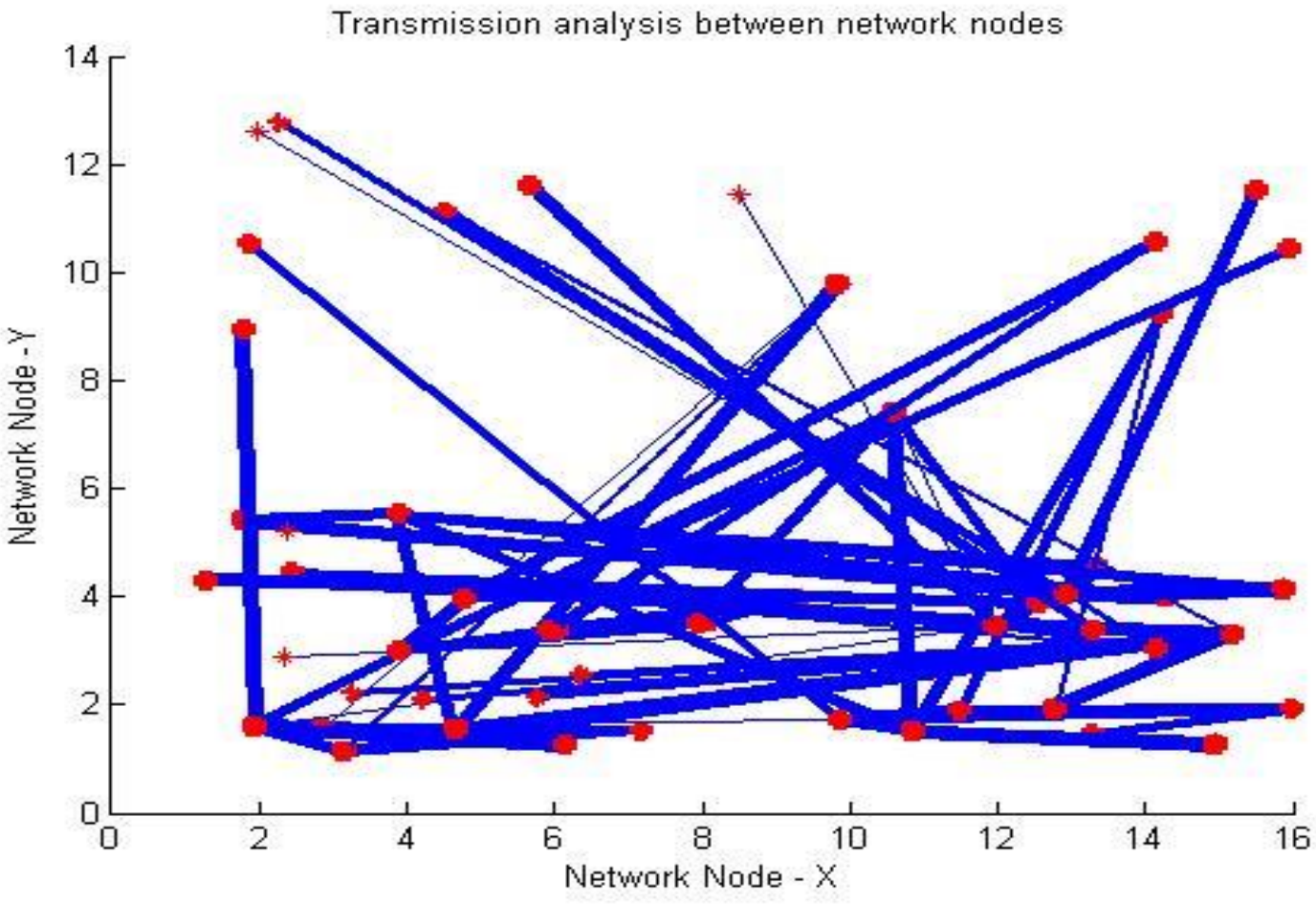
Result and Discussion

In Network Traffic Data1, information is given about the transmission between the nodes in a network.

1. **Representing The Data**

For representing the data, Graph has been used. We created an adjacency matrix from the CSV file.

For each network Node, its IP address was used as a node unique identification number. Example for node: 10.51.16.43, 43 was used as its ID number. And Number of Packets sent between two nodes was stored in the matrix.



In this Graph,
X -axis - > Network Node- X coordinate
Y -axis -> Network Node- Y coordinate
Line Width -> proportional to count of packet between nodes

2. Node Analysis (Central Nodes)

For Each node, Graphs were plotted for the number of packets being sent and number of packets being received. This was done only for one day first..

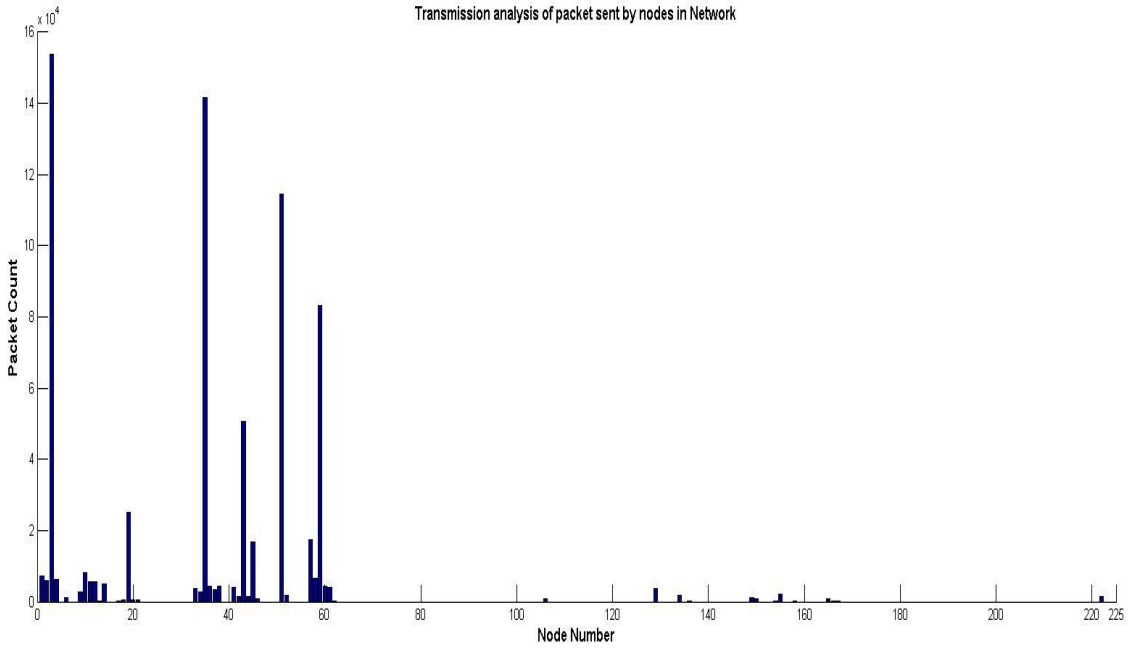


Fig 1: Packet sent frequency per node

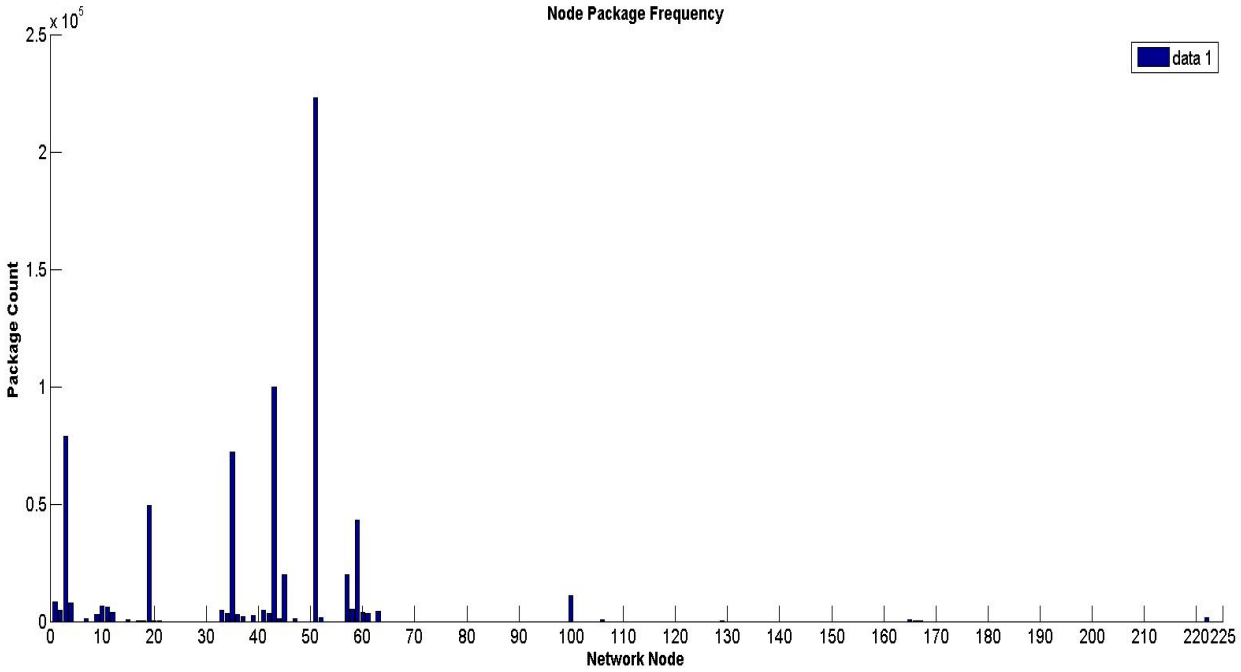


Fig. 2: Packet received frequency per node.

Observation 1: Some of the nodes were very active and were transferring large number of packets, compared to other nodes in the network.

3. Day based Network Analysis

To check whether, Obs1. Was an anomaly or not, Network analysis was done on per day basis for AllNetwork.csv File for all 9 days.

Among these, an abnormal pattern was noticed in the network for second, third and eighth day. There were huge number of requests and only some nodes were very active.

- Third Day Network Analysis

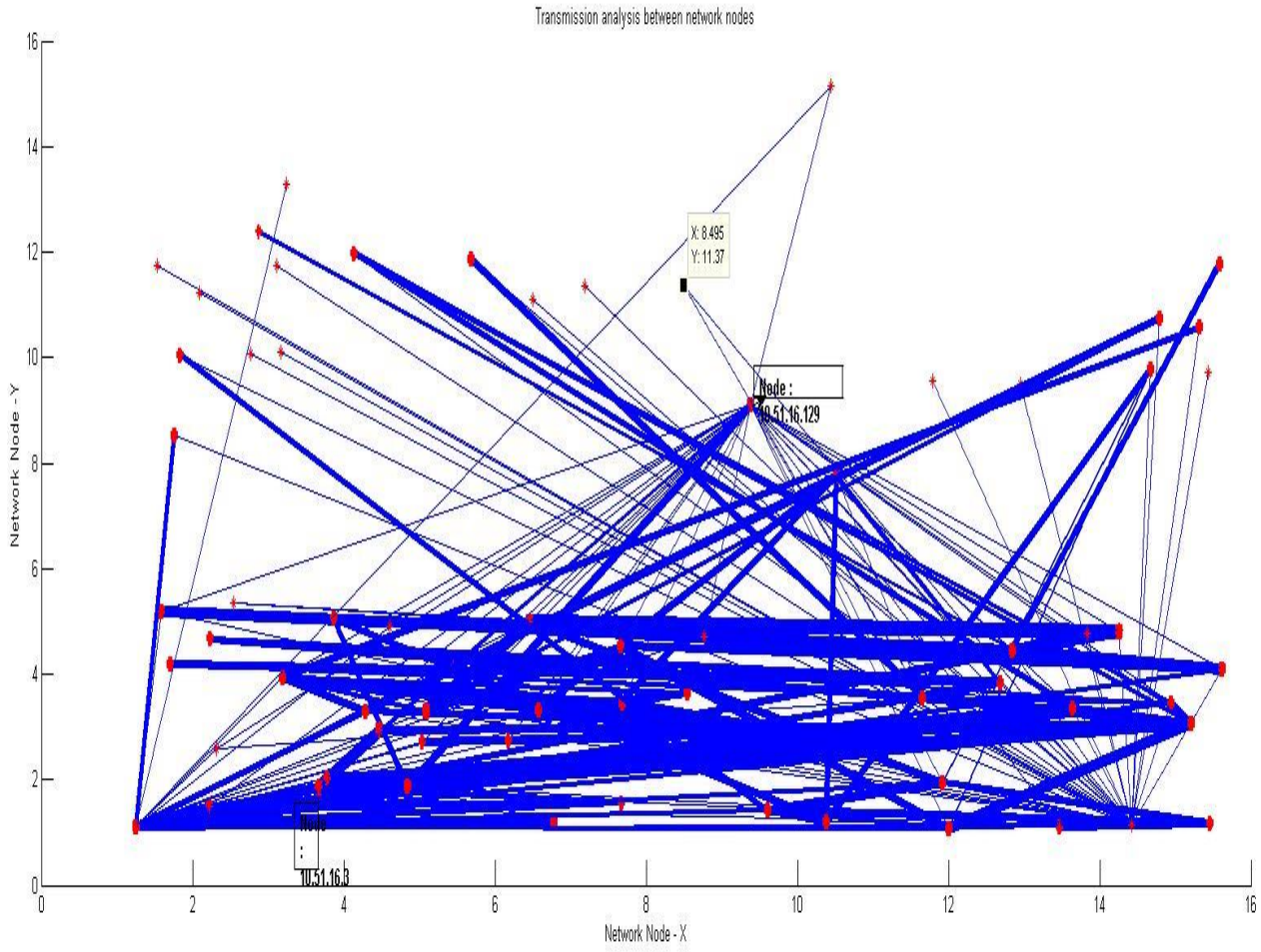
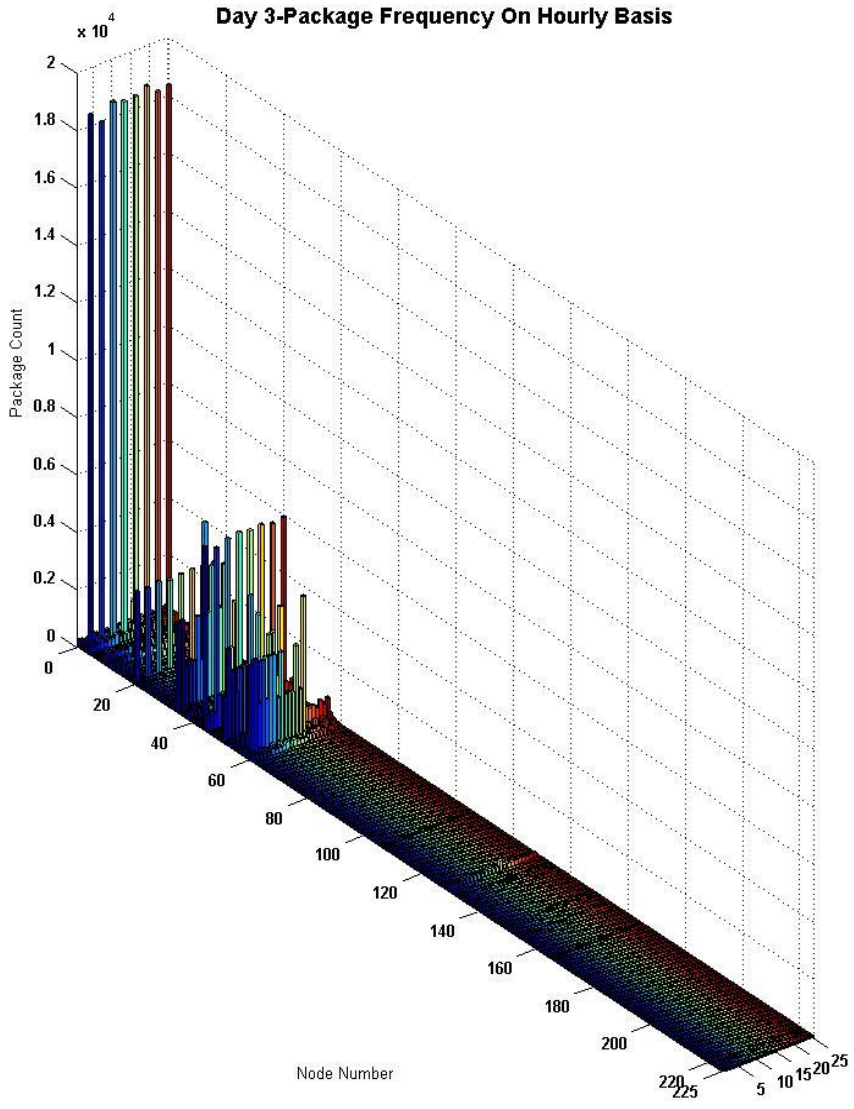


Fig 3: Transmission analysis between Network nodes.

Observation 2: In the above Figure, we can see there are some node like (10.51.16.3) which are sending large number of requests to other nodes in the network.

Observation 3: We can see one node (10.51.16.129) is sending requests to many nodes in the network compared to other days, when it is almost passive.

For Observation 2, we did the hourly analysis of request per node for Day 3.
In the below graph, we can see, Node 3 (10.51.16.3) is sending around 20,000 requests per hour.
Compared to other days it is very high.



X -axis - > Time (Hourly)
Y -axis -> Node Number
Z-axis -> Packet Count, **Fig 4:** Packet Frequency For each node (Hourly) (Day 3)

- Eighth Day

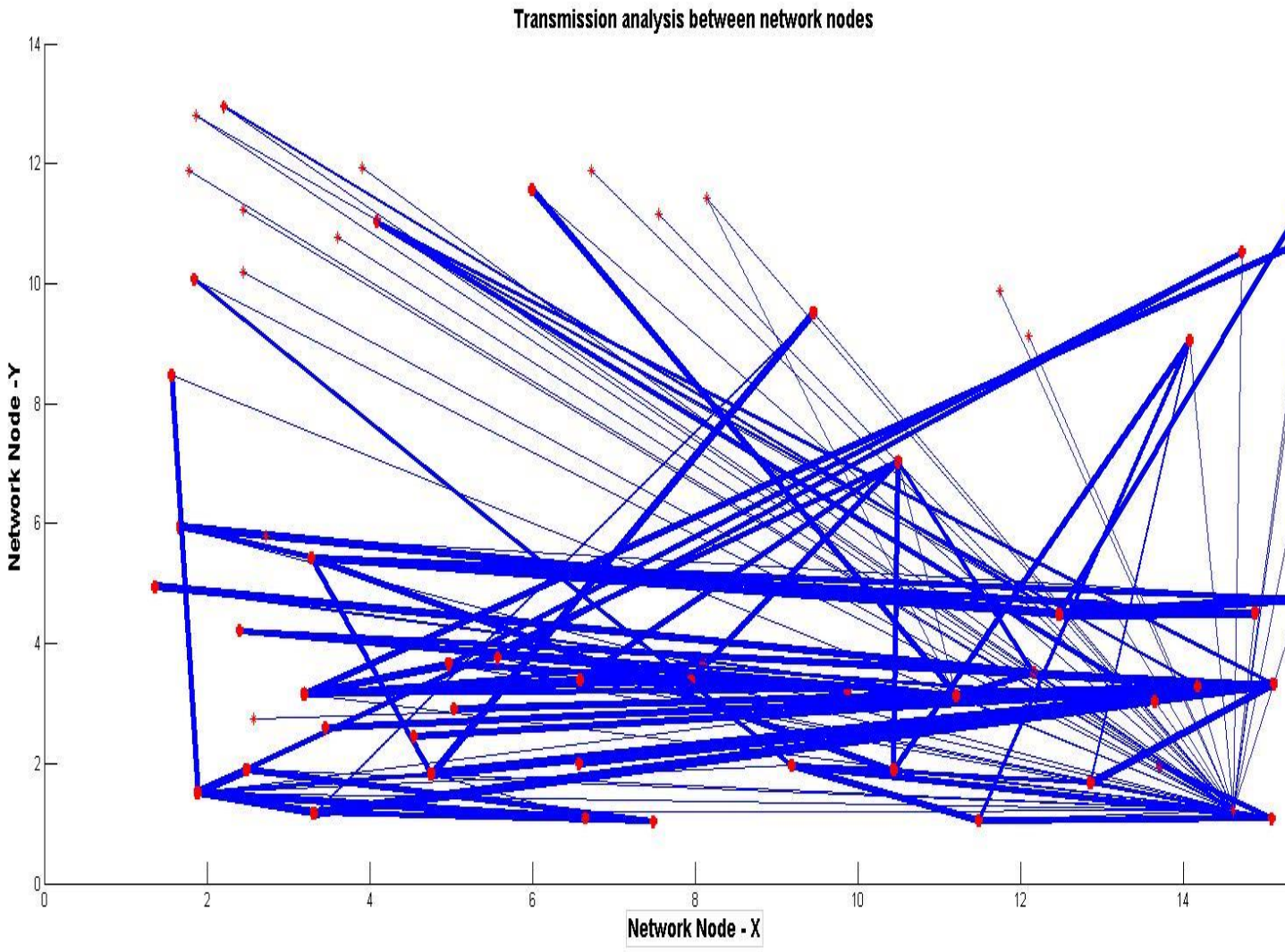


Fig 5: Packet Frequency For each node (Hourly) (Day 8)

Observation: Node (10.51.16.14) Coordinates: (14.62, 1.22) sending many requests to different node, has been inactive on other days.

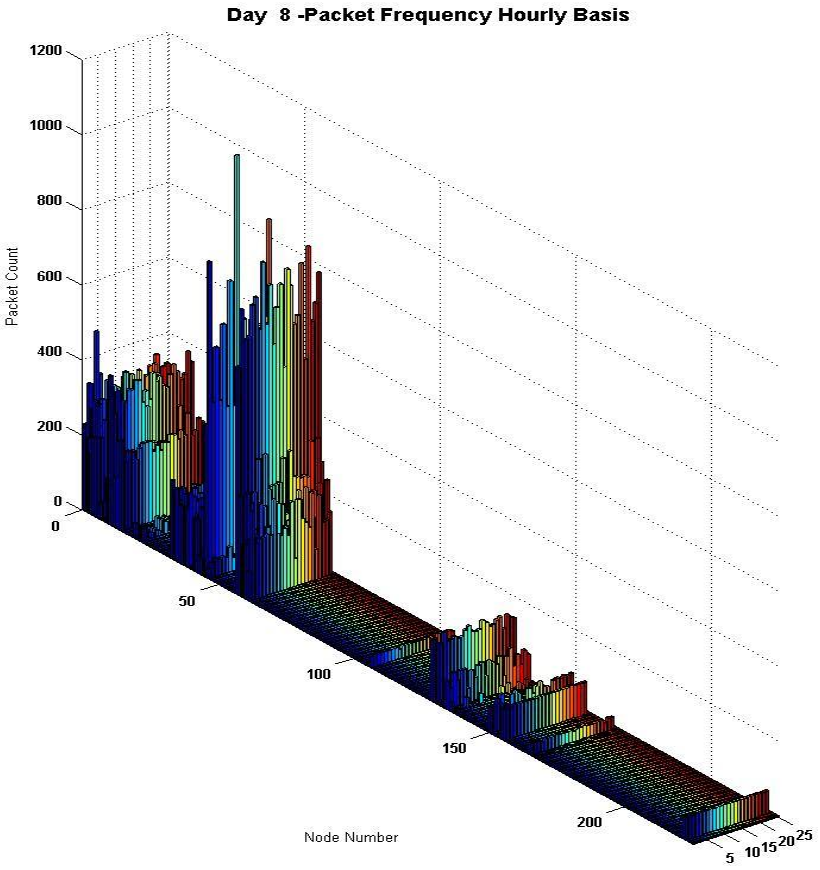


Fig 6: Packet Frequency For each node (Hourly) (Day 8)

Conclusion

There is some unusual activity in network on Day - 2, 3 & 8. Large number of requests. We can conclude this by analysis of packet frequency graph on hourly basis.
We can also see (10.51.16.3) Node sending large number of requests on Day -2&3. Compared to other days and the number is very high. Maybe trying to some nodes and choking them. (10.51.16.129) very active on day 3 and sending requests to almost all nodes in the network. Similarly for (10.51.16.14) on Day 8.