

# CHAPTER 6

Chapter-6. Regulating the affects of Topology Mismatch

<table>
<thead>
<tr>
<th>S.NO</th>
<th>Name of the Sub-Title</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Ip Address Distribution</td>
<td>111</td>
</tr>
<tr>
<td>6.2</td>
<td>Related Work</td>
<td>113</td>
</tr>
<tr>
<td>6.3</td>
<td>IP Prefix Match</td>
<td>114</td>
</tr>
<tr>
<td>6.4</td>
<td>Overlay Formation</td>
<td>116</td>
</tr>
<tr>
<td>6.5</td>
<td>Overlay Maintenance</td>
<td>117</td>
</tr>
<tr>
<td>6.6</td>
<td>Proof of Correctness</td>
<td>119</td>
</tr>
<tr>
<td>6.6</td>
<td>Performance Evaluation</td>
<td>120</td>
</tr>
<tr>
<td>6.8</td>
<td>Summary</td>
<td>125</td>
</tr>
</tbody>
</table>
Network studies [1] imply that peer-to-peer applications consistently contribute more than 50% of the total Internet traffic. Even though the increase in the resultant traffic increases the revenue for the internet service providers, it also imposes many traffic control challenges.

Traffic generated by the peer-to-peer systems sometimes stops the functioning of other application due to its excessive loads. This is mainly due to the way peer-to-peer systems function, which depends on the application layer routing based on a virtual overlay over the underlying physical network of the Internet. The logical or virtual links in overlay are formed independently from the physical network. Even though the routing in overlay occurs through these overlay links or paths, the actual transmission of the data occurs in the physical network.

The routing strategy in the overlay involves finding the best routing path from the source to destination along the virtual links connected between the peers, while the service provider tries to route through the best links in the physical network. As a result, the data is routed through sub optimal paths which result in redundant network traffic. This situation between the overlay and underlay network is called the Topology Mismatch Problem [52]. This is depicted in the Figure 6.1.
When the traffic in the Fig 6.1 is considered, to send a data item from A to C involves the path $A \rightarrow B \rightarrow C$ which is the best path in the overlay. This data transmission actually occurs in the physical network which traverses the following path i.e. $A \rightarrow D \rightarrow E \rightarrow B \rightarrow E \rightarrow D \rightarrow A \rightarrow C$. Even though the peer C is present within the same ISP (Internet Service Provider) as that of A, it has to traverse the ISP to send that data. This problem arises due to the fact that the connections in the overlay are formed without the idea of the topology of the physical network.

Since the network traffic cost will increase when there are transmissions across the ISPs to improve performance and reduce cost of communication, peer-to-peer systems have to consider the locality of peers while constructing the logical overlay. The proposed technique in...
this chapter considers the peer’s locality for solving the mismatch problem.

To solve the topology mismatch problem, peer-to-peer network formation involves the following two steps:

a. Constructing an efficient overlay using underlay topology information.

b. Choosing those nodes as neighbors which are close, during communication.

This chapter illustrates a simple method for building a peer-to-peer overlay using the structure of IP addresses and grouping the nodes which are present in the same region. The proposed approach does not introduce costly probe packets to derive or estimate the distances between the nodes [50]. This helps to reduce the communication overhead in the peer-to-peer network. Using the techniques of clustering reduces the query delay time while maintaining the accuracy of the distance estimates.

6.1 IP Addresses Distribution

Before 1999, the responsibility of assigning IP addresses to users was in the hands of the Internet Assigned Numbers Authority (IANA). This organization used to assign IP addresses according to the requirement of the user and they were not structured. From 1999, five Regional Internet
Registries (RIR’s) took over the management of assigning the IP addresses. The RIR’s can be depicted in Figure 6.2.

![Regional Internet Registries](image)

**Fig: 6.2, Regional Internet Registries**

In 2003, all the five RIR’s were united under Number Resource Organization (NRO). The addresses assigned within the organization are such that the nodes which are close to each other geographically are assigned the address from the same block. The presented system uses the IP allocations to these regions to cluster the nodes according to the region to which the node belongs. The details regarding the blocks belong to the RIR’s are given in [10]. Nodes which don’t fall into any of these RIR’s may affect the accuracy of estimates. But experiments showed that even with these nodes the performance of the peer-to-peer network can greatly enhanced.
6.2 Related Work

To solve the topology mismatch problem overlay links must match with the links of the physical network. But obtaining the information about the physical network during runtime involves complex and difficult issues. So, the peers are modeled to select those peers as neighbors that are in close proximity based on a network metric. This method of selecting the nodes is called proximity neighbor selection.

While solving the topology mismatch problem the following issues are considered:

a. Precision of the distance estimation between the nodes.

b. Communication overhead introduced.

c. Number of neighbors does a node have.

The methods described in [2] [3] [4] performs clustering based on the network infrastructure. The usability of this solution is dependent on the information available about the network infrastructure. For example, the node does not have the idea about its Autonomous System (AS) number. But the information like the nodes, IP address or DNS (Domain Name Service) servers address can be obtained easily to improve the locality awareness of the system. The earlier work considering the IP addresses was used in clustering of web servers [4]. In this the nodes were grouped based on closeness in their topology and nodes which are under the control of a common administration. Since it was used for
clustering general nodes it does not consider the properties of the peer-to-peer networks.

The initial approach for building a peer-to-peer network was made in Topology-centric Look-up service (TOPLUS) [2]. In this nodes are grouped according to their topological closeness i.e. common IP prefix. But the problem is that the solution requires global knowledge of the system which is not desirable in unstructured peer-to-peer networks. The IPv6 address structure was used in implementing the Chord overlay [6]. The resulted overlay was effective but the system does not specify how the IP addresses are assigned to the nodes.

Another solution presented in [5] uses the IP prefixes to derive the distance between the nodes and accordingly clusters them. The drawback of this approach is that the bootstrap peer maintaining the cluster routing table will become a bottleneck and a single point of failure. Other notable solutions specified to solve topology mismatch problem include LTM [7], SBO [8] and AOTO [9] use RTTs as the metric to calculate distance between nodes. But to obtain these values the nodes have to disseminate additional probe packets which add to the network traffic in the peer-to-peer systems.

6.3 IP Prefix Match

The match between two IP addresses indicates the closeness of the nodes in physical network. The match refers to the relative physical distance
between two nodes. The metric considered is the longest prefix or the IP segment which is common among two nodes IP addresses. The process starts by comparing the highest segment of the IP addresses and then goes on to compare the next segment if the first segment matches. The process will stop when the segments does not match. The longest prefix match is used to connect a node to other nodes which are physically near to each other. This process can be presented as in Algorithm 6.1.

**Step 1.** \(n1.n2.n3.n4/m\) \(\leftarrow\) Format of the IPv4 address

**Step 2.** \(N\) \(\leftarrow\) IP address of the node

**Step 3.** \(H\) \(\leftarrow\) File retrieved from the Gnutella Web Cache (GWC).

**Step 4.** \(M\) \(\leftarrow\) File containing prefix matches

**Step 5.** \(\forall I \in H\) Compare \((I, N)\)

**Step 6.** \(M \leftarrow (PM, I)\);

**Step 7.** Return \(M\);

**Step 8.** Compare \((I, N)\)

**Step 9.** For \((t=0; t<4; t++)\)

**Step 10.** Compare segment \((I_t, N_t)\)

**Step 11.** If (match) \(PM++\);

**Step 12.** End Compare segment, End if

End For

End Compare

**Algorithm 6.1: Prefix Match**
Based on the number of matched segments in the IP address the nodes can be placed at closer to each other in the peer-to-peer network.

6.4 Overlay Formation

Formation of overlay in a general Gnutella [11] network involves the following steps:

a. The Gnutella client is initiated and host cache is loaded.

b. Connect request is sent to the Gnutella network using the host cache.

c. If there is no reply and connection to the network is not established after ‘T’ seconds, a query is sent to a random GWC (Gnutella Web Cache).

d. After waiting for ‘R’ seconds another query is sent to a random GWC.

e. Wait for ‘R’ seconds from the initial GWC, if no reply comes then call a new GWC for every ‘X’ seconds.

f. After getting the replies the nodes choose the random nodes from the cache as neighbors.

The first five steps involve the general bootstrapping process in the Gnutella network. The existing implementations of Gnutella provide enough GWC servers to support the bootstrapping process. In this manner a general Gnutella node will connect to the network and choose the random nodes as neighbors.

In the proposed implementation, the Gnutella overlay formation process is modified to imbibe the changes proposed in the system. Firstly
the process change is made in step (c) where instead of connecting to the random GWC server, it let the node to find the closest default GWC server in a region to acquire the list of nodes querying the server for the nodes having longest prefix match. Next the step (e) is modified in such a way that if no response comes from a GWC server then next closest GWC server is queried for every ‘R’ seconds.

In this manner, the node will query the closest GWC server to retrieve the nodes which are present within the same region and derive the neighbors which are physically closer to it.

6.5 Overlay Maintenance

During the formation of overlay the nodes, will choose the neighbors according to the prefix match with the other nodes. First a node will select those nodes with the longest prefix match and if no such node exists then it will try to connect to the next lower prefix match nodes. In the similar fashion a node will try to get the maximum neighbors it can have. Once it reaches the max number of connections it will disconnect the node with lowest prefix match value if a better node tries to connect to the node. This process is depicted in Fig 6.3.

In the Figure 6.3, if the node F sends a connect message to the node A which can have a maximum of 4 neighbors then it will check the prefix match with E and remove one of the existing connections i.e. node D which has smaller prefix match than the new node.
Step 1. \( \text{Nbr}_A \leftarrow \text{Neighbors of } A \).

Step 2. \( \text{Max}_\text{Nbr}_A \leftarrow \text{Maximum Neighbors } A \text{ can have} \).

Step 3. \( \text{PM}_{AB} \leftarrow \text{Prefix Match of } A \text{ and } B \).

Step 4. If \( Nbr_A + 1 \leq \text{Max}_\text{Nbr}_A \) then

Step 5. Accept \( F \); Return

Else

Step 6. \( \text{Candidates} \leftarrow C \forall C \in \text{Nbr}_A \text{ with } PM_{FA} > PM_{CA} \)

Step 7. If No Candidates Exists

Step 8. Reject \( F \); Return

Else

Step 9. Accept \( F \);

Step 10. Drop \( C \) with least \( PM \) (Prefix Match) value;

Algorithm 6.2: Topology maintenance using prefix match
Along these lines the node will keep changing its connections while further closer nodes try to connect. Even in the case where some of the neighbors move away from the network, the node will still find the other closer nodes using the GWC servers. Algorithm 6.2 shows the topology maintenance using prefix match.

6.6 Proof of Correctness

**Theorem 6.1:** In the inter Autonomous System (AS) the traffic will be reduced when the overlay formation methodology takes into account the structuring of IP addresses.

**Proof:** Let us assume the inter AS traffic will be quite low, during the overlay formation without using IP address structuring. Without considering the IP address Structuring, the nodes may choose the neighbors randomly and these neighboring nodes may or may not fall within the same AS. Hence we consider a overlay network \( O(N,E) \) in which the cost of communication for a intra AS link be \( I \), and for a inter AS be \( I_1 \), where \( I >> I_1 \). The cost of communication for any five hop length path \( p_1 \) in the overlay network is

\[
\text{cost}_{\text{random}}(p_1) = \begin{cases} 
5 \times I \text{ when intra AS} \\
5 \times I_1 \text{ when all inter AS links}
\end{cases}
\]
When the random choice of neighbors occurs, there will be chances of large number of inter AS links, so by considering the structuring of IP addresses the cost of communication will be \( \text{Cost}_{IP}(P_i) = 5 * I \). Hence by using IP address structuring, we reduce the inter AS traffic and its Communication cost.

### 6.7 Performance Evaluation

The implementation results presented for the overlay are launched into the Gnutella nodes by changing the specifications based on operations performed by the node such as contacting the GWC server and choosing its neighbors based on the prefix match of the nodes. The implementation is carried out on a general purpose java simulator. A peer-to-peer system is simulated using a 500 node network and the IP addresses are randomly assigned to nodes by the simulation module. To compare the accuracy and effectiveness of this proposal first a general Gnutella network is implemented which chooses the neighbors randomly and then simulate the improved Gnutella network formation using the prefix match to find the neighbors.

Table 6.1 shows comparison of Gnutella and proposed technique Gnutella-PM (prefix match) with respect to the communication cost and time.
Table 6.1: Communication Cost

<table>
<thead>
<tr>
<th>Time(Sec)</th>
<th>Communication Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gnutella</td>
</tr>
<tr>
<td>50</td>
<td>17000</td>
</tr>
<tr>
<td>100</td>
<td>16000</td>
</tr>
<tr>
<td>200</td>
<td>19000</td>
</tr>
<tr>
<td>300</td>
<td>19000</td>
</tr>
<tr>
<td>400</td>
<td>19000</td>
</tr>
<tr>
<td>500</td>
<td>20000</td>
</tr>
</tbody>
</table>

To access the cost effectiveness of the solution, communication cost of the two networks is compared as shown in Figure 6.4. It implies that overlay formed using the prefix matching method reduces the cost of communication drastically as the nodes chooses the neighbors within the same region.

Fig: 6.4, Communication Cost
Next, the simulation considers the delay involved to reply the queries sent by the nodes. The delay is considered in time (ms) taken by the network to reply queries sent by a set of nodes.

Table 6.2 shows comparison of existing and proposed techniques with respect to the delay and no of nodes.

Table 6.2: Query Delay

<table>
<thead>
<tr>
<th>No. of Nodes</th>
<th>Query Delay (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gnutella</td>
</tr>
<tr>
<td>50</td>
<td>17000</td>
</tr>
<tr>
<td>100</td>
<td>16000</td>
</tr>
<tr>
<td>200</td>
<td>19000</td>
</tr>
<tr>
<td>300</td>
<td>19000</td>
</tr>
<tr>
<td>400</td>
<td>19000</td>
</tr>
<tr>
<td>500</td>
<td>20000</td>
</tr>
</tbody>
</table>

The comparison of existing and proposed systems is depicted in Figure 6.5. It shows that delay in the presented system is far less compared to Gnutella network as the network connections across an Autonomous System are very less.
Fig: 6.5, Query Delay

Table 6.3 shows comparison of existing and proposed techniques with respect to the average no of messages and no of nodes.

Table 6.3 Average no of Messages

<table>
<thead>
<tr>
<th>No. of Nodes</th>
<th>Average No. of Messages</th>
<th>Gnutella</th>
<th>Gnutella-PM</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>100</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>100</td>
<td>150</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>200</td>
<td>300</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>300</td>
<td>350</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>400</td>
<td>400</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>500</td>
<td>450</td>
<td>200</td>
<td>200</td>
</tr>
</tbody>
</table>

The average number of messages generated by a set of nodes in both the networks is given Figure 6.6.
Table 6.4 shows comparison of existing and proposed techniques with respect to the average hop length and no of nodes.

Table 6.4: Average hop length

<table>
<thead>
<tr>
<th>No. of Nodes</th>
<th>Average Hop Length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gnutella</td>
</tr>
<tr>
<td>50</td>
<td>4.5</td>
</tr>
<tr>
<td>100</td>
<td>5</td>
</tr>
<tr>
<td>200</td>
<td>5</td>
</tr>
<tr>
<td>300</td>
<td>5.5</td>
</tr>
<tr>
<td>400</td>
<td>5.3</td>
</tr>
<tr>
<td>500</td>
<td>6</td>
</tr>
</tbody>
</table>

The average hop length of the query is taken to derive the effectiveness of the prefix match method that is helpful in solving the topology mismatch problem. This comparison is shown in Fig 6.7.
Fig: 6.7, Average Hop Length

The simulation shows that the proposed overlay formation technique using the prefix matching method greatly reduces the redundant traffic caused by topology mismatch problem without introducing additional probe packets and providing faster query responses.

6.8 Summary

This chapter illustrates a simple but effective approach to regulate the effects of the topology mismatch problem by constructing a peer-to-peer overlay using the IP Prefix matching between the nodes. The system does not generate any additional messages to determine the distance between nodes and provides a better performing peer-to-peer system than the existing peer-to-peer systems.