INFORMATION RETRIEVAL IN NETWORKS USING BFS

CSCE 5150

Sai Divya Gunuguntla
Computer Science
University of North Texas
Denton, Texas

Raji Kovvuri
Computer Science
University of North Texas
Denton, Texas

Project Objective

The major objective of the project is to find the shortest path on a network graph through searching all nodes and edges using Breadth First Search (BFS). The algorithm, BFS can be used in P2P Networks, Web Crawlers, Network Broadcasting, Navigation Systems and Un-Weighted Graphs. In our scenario, we pegged our research discussion on P2P networks in a learning institution. The learning institution has a departmental computer connected together via a switch, referred herein as Local Area Networks (LAN), connection of departments referred to as Metropolitan Area Network (MAN) and the provision of/to external connection to the rest of the world referred herein as Wide Area Network (WAN).

Applications of BFS

BFS can be used in scenarios such as:

1. GPS Navigation system to find all neighboring locations.
2. Packets broadcasting in networks
3. Location of nearby individuals from the social media user in social networking websites.
4. Finding of paths
5. Shortest path and minimum spanning tree for unweighted graph.
6. Testing of bipartite nature of graphs.
7. In Garbage Collection: Breadth First Search is used in copying garbage collection using Cheney's algorithm.

These and many more others depicts scenarios where BFS can be used.

Experiments

Datasets Identification

Our datasets involve networks that we encounter in our daily activities. We have identified both large and smaller networks within our institutional environment. Each department/school within the institution has its own local LAN that can only be accessed by the authorized users, equally, the LANs of these departments/schools are connected to form the faculty network, MAN, which are further connected to form the institution's WAN network.

Graph	# of nodes	# of edges
Departmental LAN	300	500
Faculty MAN	20	40
Institution's WAN	30	60

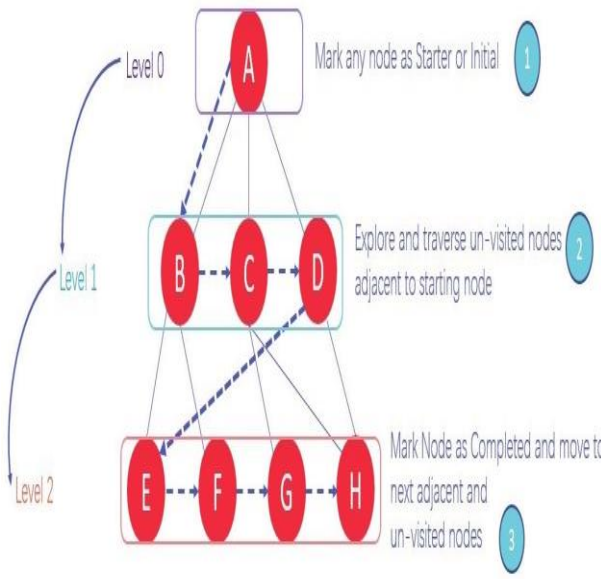
Departmental LAN: This is a computer network within 100m radius and encompasses a single unit that needs to share same immediate resources. Departmental LAN contains client machines for both the students and departmental staff, the wireless router (for internet provision within the department), the web and application server and a gateway to the rest of the networks within the institution.

Faculty MAN: It is viewed as a Metropolitan Area Network (more than 100m distance coverage) that has a composition of several faculty departments' networks. Each faculty has designated network parameters including IP addressing, servers and the necessary applications. Faculty network often has its nodes and edges originate from the departmental LAN networks. The gateways of these departmental networks act as the nodes to the faculty's network.

Institution's WAN: A group of the faculties MAN that are managed at a centralized pool. The institutional WAN provides gateway to the rest of the world on behalf of all the networks configured within it. It is the source of all internet provision within the institution and logically has several nodes embedded unto it. Physically, it only broadcasts the signals to respective faculty and departmental gateways thus making it have even fewer node.

Algorithm:

Breadth first search algorithm scans the entire network route by vising all the nodes and paths to ascertain the shortest path that packets can use to traverse from the source to the destination.



Breadth first search

The algorithm efficiently visits and marks all the key nodes in a graph in an accurate breadthwise fashion. This algorithm selects a single node (initial or source point) in a graph and then visits all the nodes adjacent to the selected node. BFS accesses these nodes one by one.

Once the algorithm visits and marks the starting node, then it moves towards the nearest unvisited nodes and analyses them. Once visited, all nodes are marked. These iterations continue until all the nodes of the graph have been successfully visited and marked.

Pseudo code:

```
BFS (v):
    For each vertex n
        Do g[l] = inf
    G[k] = 0
    Queue q
    q.push (k)
    while q is not empty
        z = q.front ()
        q.pop
        for each t in adj[u]
            if g[t] ==inf
```

Conclusion:

The BFS algorithm is of beneficial to those scenarios that need to find the shortest paths that network packets can traverse from source node to the destination node. In this setup scenario we have provided an implementation of four nodes being traversed by the BFS in an institution's departmental network layout.

The new routers and switches have BFS algorithm installed into them especially at routing level/configuration.

Reference:

1. https://www.hackerearth.com/practice/algorithms/graphs/breadth-first-search/.
2. https://ieeexplore.ieee.org/abstract/document/130694.
3. https://www.researchgate.net/figure/Routing-Query-Messages-in-a-P2P-network-using-a-BFS