Lecture 17 — May 2002

Lecturer: None *Boston University* *Scribe: Vishal Sood*

Five presentations were made in this, the third session of class presentations:

1. Vishal Sood : "Connectivity and Routing in Wireless Networks: A Physics Perspective"
2. Trevor Macdowell: "Cabinet: A Topologically Sensitive Peer-to-Peer Construction"
3. SuYan Zeng and Wenge Yang: "Web Information Searching with Category Trees"
4. Anukool Lakhina: "Discovering Graphs from End to End Measurements"
5. Mina Guirguis: "Trade-offs in Routing Inside a d-dimensional Torus"

17.1 Connectivity and Routing in Wireless Networks: A Physics Perspective..

The author talked about various problems in physics that are related to the connectivity problem of wireless networks. He pointed to the presence of disorder in wireless networks and suggested how the methods used in statistical physics can be used to solve this problem. He also suggested how the problem of routing can be analyzed using random walks and path integrals methods used in physics.

17.2 Cabinet : A Topologically Sensitive Peer-to-Peer Construction.

The author's goal was to build a peer-to-peer lookup system that took advantage of topology while maintaining a uniform overlay. He builds a hierarchial structure on the network. Search is done at two levels. Each node belongs to a single CAN. The diameter of each CAN is less than a certain number C, which becomes a parameter for the construction. Each node maintains a link to a node in each neighbouring CAN. A search algorithm and algorithms for entry of a node and leaving of a node were presented. The author simulated the Cabinet and compared it to a CAN. The simulations were incomplete at the time of the presentation and the results inconclusive.

17.3 Web Information Searching with Category Trees.

The authors attempt to combine the search engine and the classified directory methods for search on the net to design a new search method which they call "Search with Category

Tree (SCT) ". Their idea is to create a tree, called the Category Tree, using for example the Look Smart's Web Directory , with each node representing some related web pages. This construction of the tree is done offline. The user's queries are then mapped to these nodes on relevance.
The authors then discussed in detail how to build such a category tree. They suggested using mutual information between the documents and each category. Then they explained the idea Support Vector Machine (SVM), a method for text classification, and how they use it to build the category tree.
Then they showed how to map a query to a category. A page can be associated with a large number of categories. The authors suggest that with high probability tha this number is less than four and that search be stopped when this upper bound is reached.
They pointed to some well known methods for page ranking and then concluded by remarking that their method of combining the search engine and classified directory methods is quite efficient, and pointing to future directions.

17.4 Discovering Graphs from End to End Measurement.

The author claimed that the present methods of measuring the topology of the internet are not efficient and not enough attention has been paid to this problem. The author performed an experiment on a random graph and measured its degree distribution using the methods used to measure the degree distribution on the internet. He showed the graphs of his results which suggested that the measured distributions are far from the actual.
The author introduced the algorithm used for these measurements. The algorithm used should cover all the links in the graph in the ideal case. This number of links $L(k, m)$ covered depends on the number of sources, k and destinations , m used in the measurement. He showed Chaung-Sirbu's fits for $L(1, m)$. He then showed his own numerical results for $k = 1,\ 2,\ 5,\ 10$. He then discussed the problems of analytically understanding $L(k, m)$.
He concluded by discussing the future work that he intends to carry on.

17.5 Trade-offs in Routing Inside a d-Dimensional Torus.

This is an extension of the CAN. The author's idea is to store the address of some long range nodes (LRNs) along with the 2d neighbours and to allow long jumps to these LRNs. The author discussed the advantages of using these LRNs and how to chose them and their optimal number. The choices depend on the kind of requests.
For uniform requests the list of LRNs is kept static. The author discussed various ways of choosing these LRNs. The trade-offs are the decrease in average lenght and the increase in size of a node's state. The author's results for this case suggest that the average path lenght for a query falls to an asymptotic value of 7.
The author then discussed a "dynamic mode", in which the LRN are updated as nodes route requests. This design is for requests genetrated according to Zipf's distribution. There are

some popular zones and it is beneficial if every node has an LRN inside these zones.
The author created a Hotspot phenomena to test this design and discussed an update algorithm. He found that updating the LRN will eventually converge to a node inside the hotspot zone if done in a smart way.
The author concluded that having LRN is good.