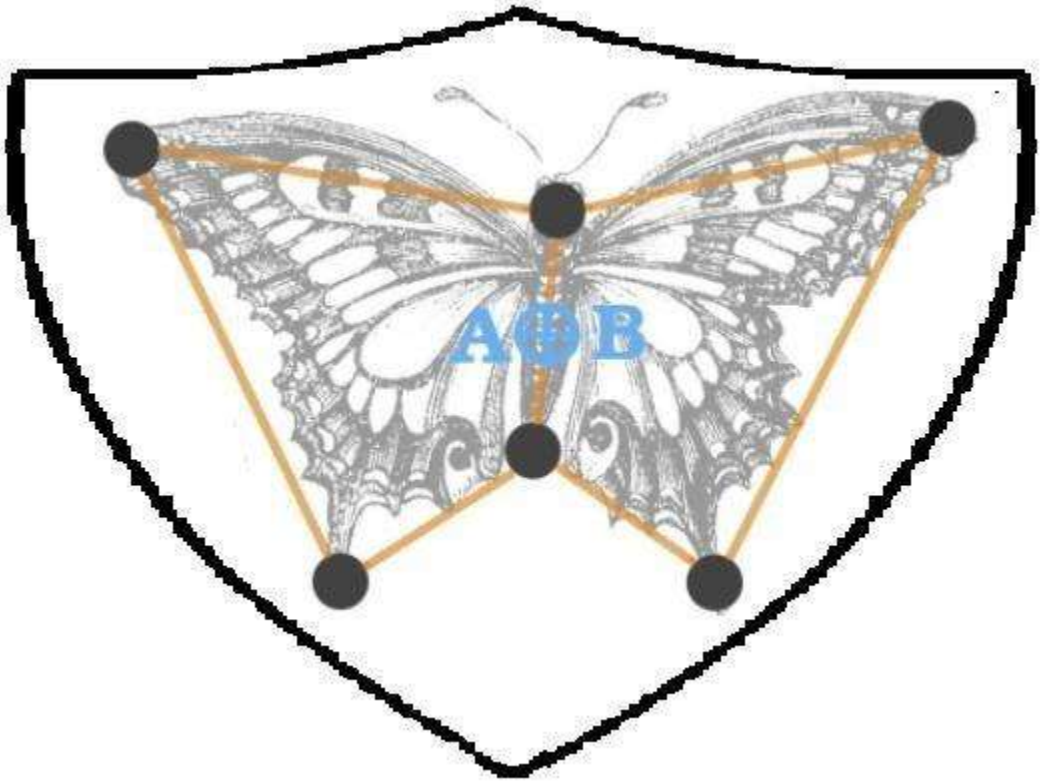
Network Coding: A Practical Application



Final Project Report for Computer Networks (COEN 233)

By,

***Avinash Kumar**
***Anoop Thomas**
***Nithin Narayan**

Acknowledgement

It is our pleasure to be indebted to all the people, who contributed in the development of this work and influenced our thinking, behavior, and acts during the course of study.

We took this opportunity to express our profound gratitude and deep regards to our guide Professor Ming Wang for his guidance, monitoring and constant encouragement and feedback throughout the course of this project.

We got to learn a great deal about the serious drawbacks of Wireless network, how implementing Network Coding helps maximize efficiency, increase bandwidth, and how data is made robust using this technique.

Lastly, we are highly thankful to the almighty God and our friends with whom we shared our day- to-day experience and received lots of suggestions that improved our quality of work.

Table of Content

Output generation
Output analysis
Compare output against hypothesis
Abnormal case explanation (the most important task)
Statistic regression
Discussion

Conclusions and Recommendations
Summary and conclusions
Recommendations for future studies

Bibliography

- # Introduction

· Objective

Network coding has been recently proposed in information theory as a new dimension of the information multicast problem that helps achieve optimal transmission rate or cost Network coding involves coding at intermediate nodes between the source and the receivers in one or multiple communication sessions minimize latency, increased throughput, and bandwidth efficiency.

· What is the problem

For a single source multiple terminal network, 2 bits of data from the source node cannot reach the destination node in a single transmission Here we assume the channels have unit capacity Hence at least two separate transmissions are required.

· Why this is a project related to this class

Since we study the various modes of communication namely Unicast, Broadcast and Multicast; and the advantages and drawbacks of each type here we determine that using Linear coding within multicast communication increases efficiency, robustness and throughput within the network

· Why other approach is no good

Within a single source multiple terminal network, 2 bits of data cannot reach the destinations using the Hop by Hop mode of transmission Hence multiple transmissions are required to get this done.

· Why you think your approach is better

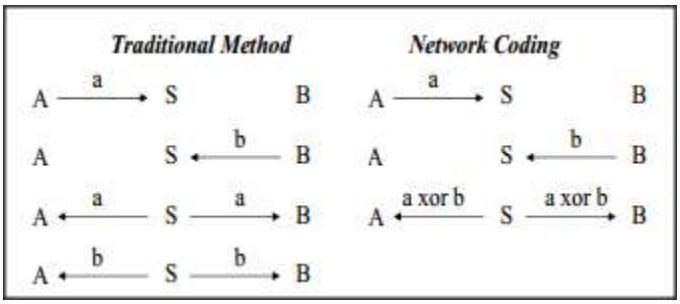
Using a simple combination method, two bits can be transmitted to both the destinations simultaneously by sending the sum through the center in the final destination; one information bit can be calculated by subtracting the other one from the combination

The network coding helps to reduce the transmission cost and improve the security. Network bandwidth is improved since the number of packet retransmission is reduced for a single hop wireless multicast transmission. It is considered as an alternative to forward error correction and ARQ in traditional and wireless networks and is resilient to network attacks like snooping and eavesdropping.

· Statement of the problem

To use network coding techniques to improve the channel efficiency, reduce the transmission cost and increase the security in a network.

· **Area or scope of investigation**- Data transmission and improvement of data stream communication.



Theoretical bases and literature review

· Definition of the problem

Network coding is a networking technique in which transmitted data is encoded and decoded to increase network throughput, reduce delays and make the network more robust In network coding, algebraic algorithms are applied to the data to accumulate the various transmissions The received transmissions are decoded at their destinations This means that fewer transmissions are required to transmit all the data, but this requires more processing at intermediate and terminal nodes

· Theoretical background of the problem:-

A communication network can be considered as a finite directed graph, where multiple edges from one node to another are allowed and the edges represent pathways for information A node without any incoming edges is called a source node Destination nodes are known as terminal / sink nodes All other nodes are intermediate nodes Using the max-flow min-cut theorem, one can calculate the maximum amount of information that can be pushed through this network between two nodes For a single-source multiple-terminal network, the information rate to each terminal is the minimum of the individual max-flow bounds over all source-terminal pairs under consideration and that in general we need to code over the links in the network to achieve the maximum capacity It is proved that this max-flow value is achievable, but the traditional routing mechanism is not capable of achieving it.

Network coding is a field of information theory and coding theory and is a method of attaining maximum flow in a network The core notion of network coding is to allow mixing of data at intermediate network nodes A receiver sees these data packets and deduces from them the messages that were originally intended for that data sink The fundamental concept of network coding was introduced for satellite communication networks The theory of network coding has been developed in various directions, and new applications of network coding continue to emerge Network coding can be implemented in both cyclic and acyclic networks.

· **Related research to solve the problem-** It is proved that using linear network coding, the upper bound in multicast problem can be solved Let a group of nodes **P** are involved in moving the data from **S** source nodes to **K** sink nodes Each node receives new packets which are linear combinations of the earlier received packets, multiplying them with the coefficients chosen from a finite field of size GF(2s).

· **Advantage/Disadvantage of those research-** Linear Network Coding is not sufficient to solve the case of a multi-source, multi-sink with arbitrary demands.

· **Your solution to solve the problem** Coefficients chosen from a Galois field can be used for the nodes to transmit random linear combinations of the packets they receive If the field size is sufficiently large, the probability that the receiver(s) will obtain linearly independent combinations (and therefore obtain innovative information) approaches 1.

Why your solution is better Linear Network coding cannot be applied to the case of a multi-source, multi-sink with arbitrary needs but the proposed random network coding scheme allows close to optimal throughput The random network coding has excellent throughput performance but if the sink node obtains insufficient number of packets, then it is extremely difficult to recover the original packets.

- **Hypothesis (or goals)**

If random linear combinations of the packets are performed in the intermediate nodes, then separate transmissions for each and every message (like that in the normal store-and-forward case) is not required These transmissions, when network encoded, will prove to be more robust and therefore allow for less error and a reduced need to retransmit packets, when compared to the traditional, single-connection method implemented in most networking applications today We hope to be able to prove that network coding is a more versatile, reliable, and efficient means of communication than sending data from node to node by using “traditional” (single data stream) methods utilized today.

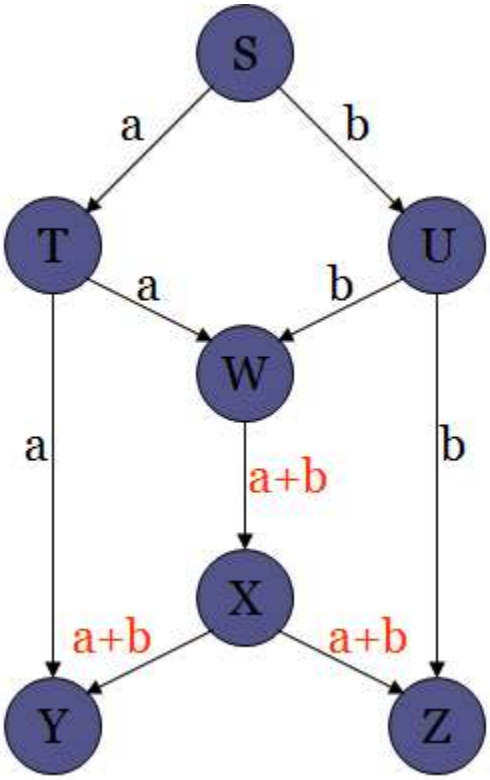
- **Methodology**

· **How to generate/collect input data:-**

Data will be analyzed from a broad variety of sources Using NS2 or a similar simulation program, we will be able to map out various network topologies, looking for layouts that compare with pre-existing, conventional ones that rely on the standard TCP/IP and OSI stack models.

We will also be taking measurements of network lag in both traditional and network coding topologies, analyzing all causality that we can infer.

A large component of our analysis will be to measure packet loss, as well as the ability to recover those packets This correlates directly to network lag We will then analyze the number of packets dropped relative to the amount of data trafficked, and compare them between the two topologies.



· **How to solve the problem**

We plan on setting up connections between multiple source and destination (sink) nodes in order to create a well-established lattice network, as per network coding topology Each set of nodes contains a subset of nodes that we will hereafter refer to as intermediate or bridge nodes: these

intermediate nodes have access to the network beyond the LAN, and source intermediate nodes will be able to communicate with the sink only through it's intermediate nodes Non-intermediate source and sink nodes will only have access to source or sink nodes, respectively Each data packet that is transmitted will be a Galois-Field linear combination of all data being sent from the various source non-intermediate nodes.

· Algorithm design

Data will be received from the source nodes at the source-intermediate nodes These nodes will then compute a new stream of data, to be sent out through the connections that are made to the intermediate sink nodes A list of coefficients will also be transmitted, and will be used to decode the linearly-combined data at the sink intermediate node level This data will then be distributed to all sink nodes.

· Language used

The language we intend to use is Java We feel that it is more robust compared to most other options, and has good support for network coding and general networking libraries.

· Tools used

We intend to use eclipse as a code developing tool, as well as the labs in the computing center to implement our experiment.

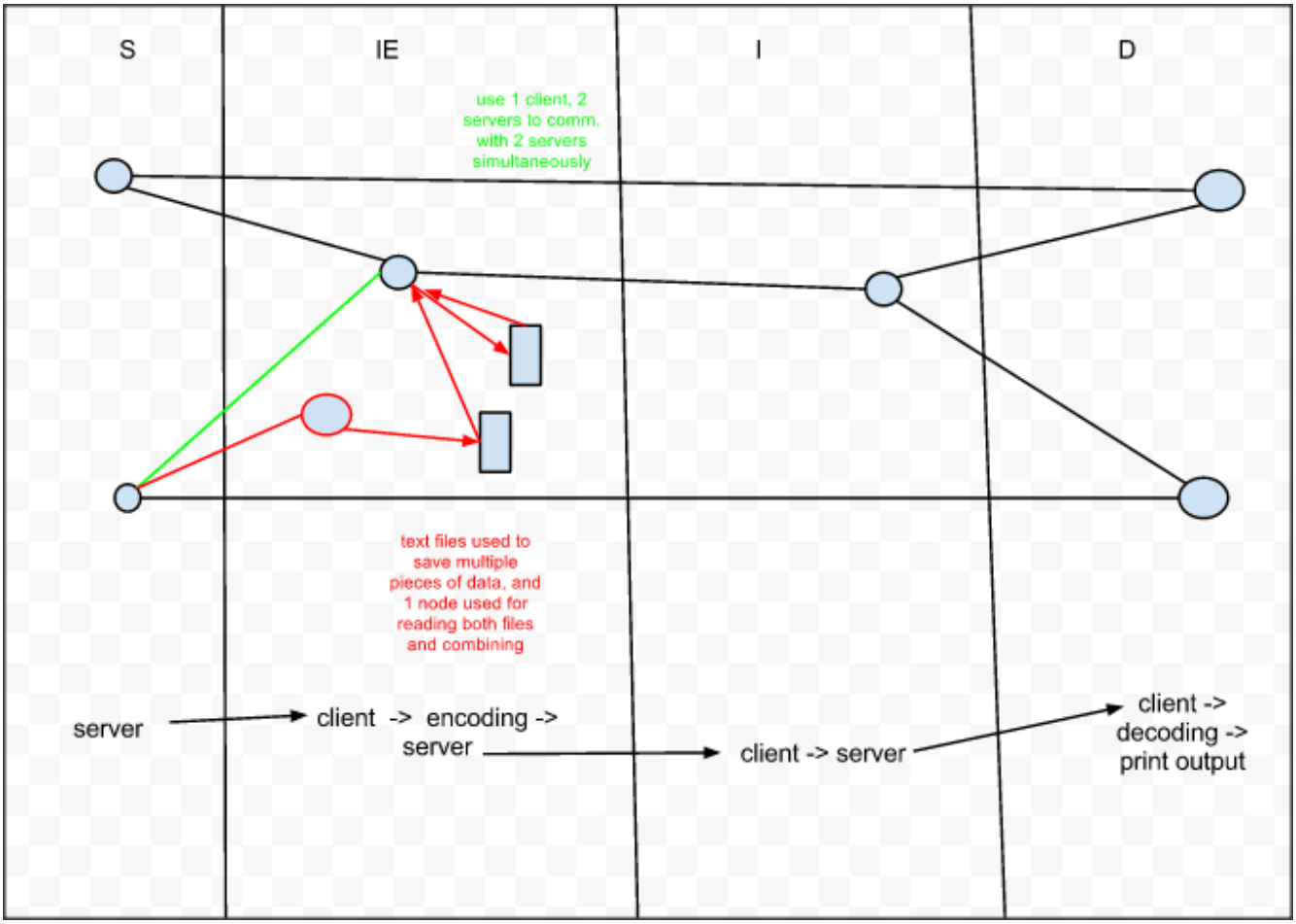
. How to generate output

The output will be generated at the sink intermediate nodes Once coefficients have been solved for, we can utilize matrix row reduction to solve the given system of equations and extract the data encoded within the stream The original data is then transmitted to the non-intermediate sink nodes.

· How to test against hypothesis

To test against our hypothesis, we need to analyze packet loss and lag time for both a network-coded topology and a similar, single-connection network that is identical in function We hope to show that network coding needs to account for less packet loss when compared to a "traditional" network, and can thus be exposed as the superior method of network data transmission.

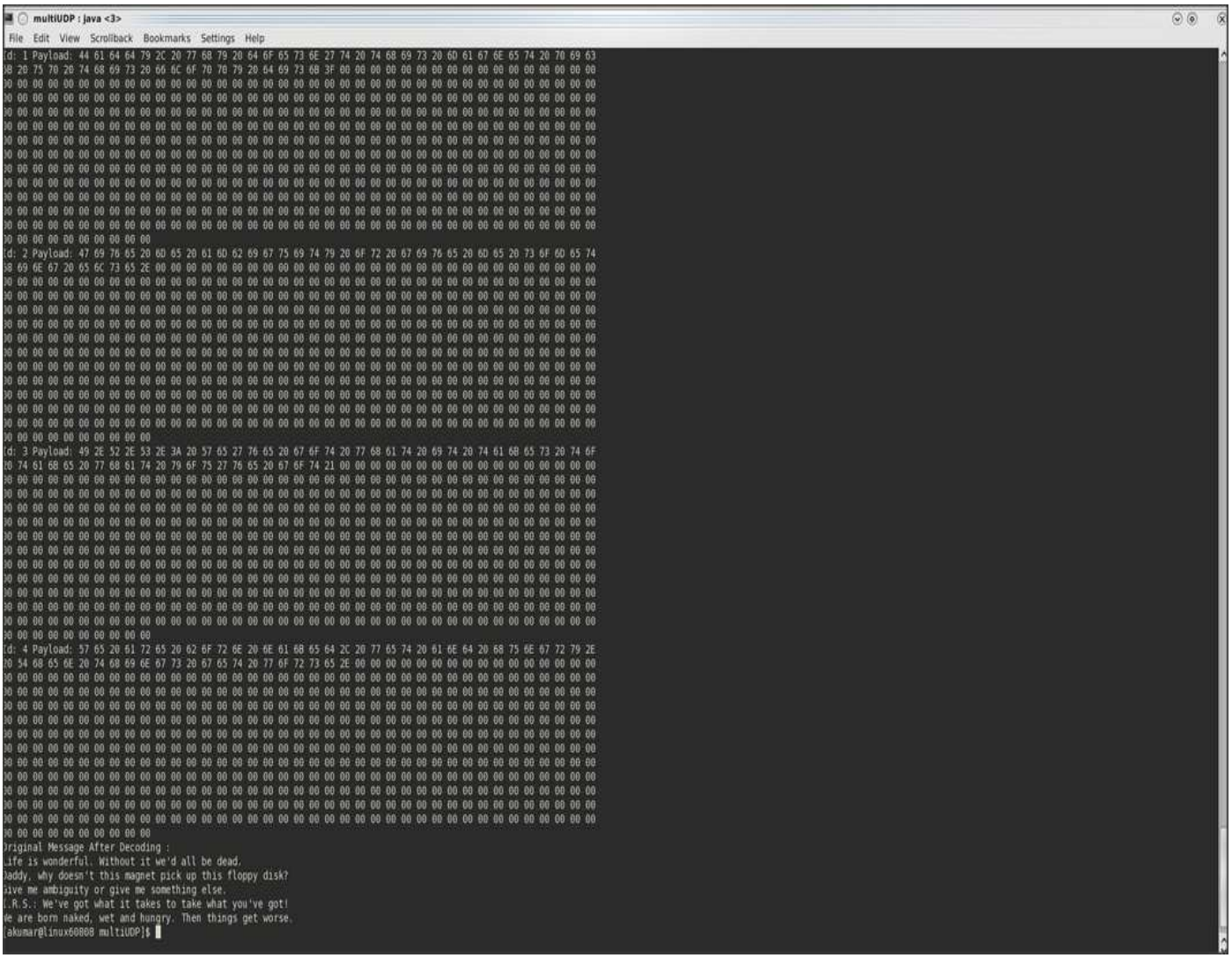
Implementation



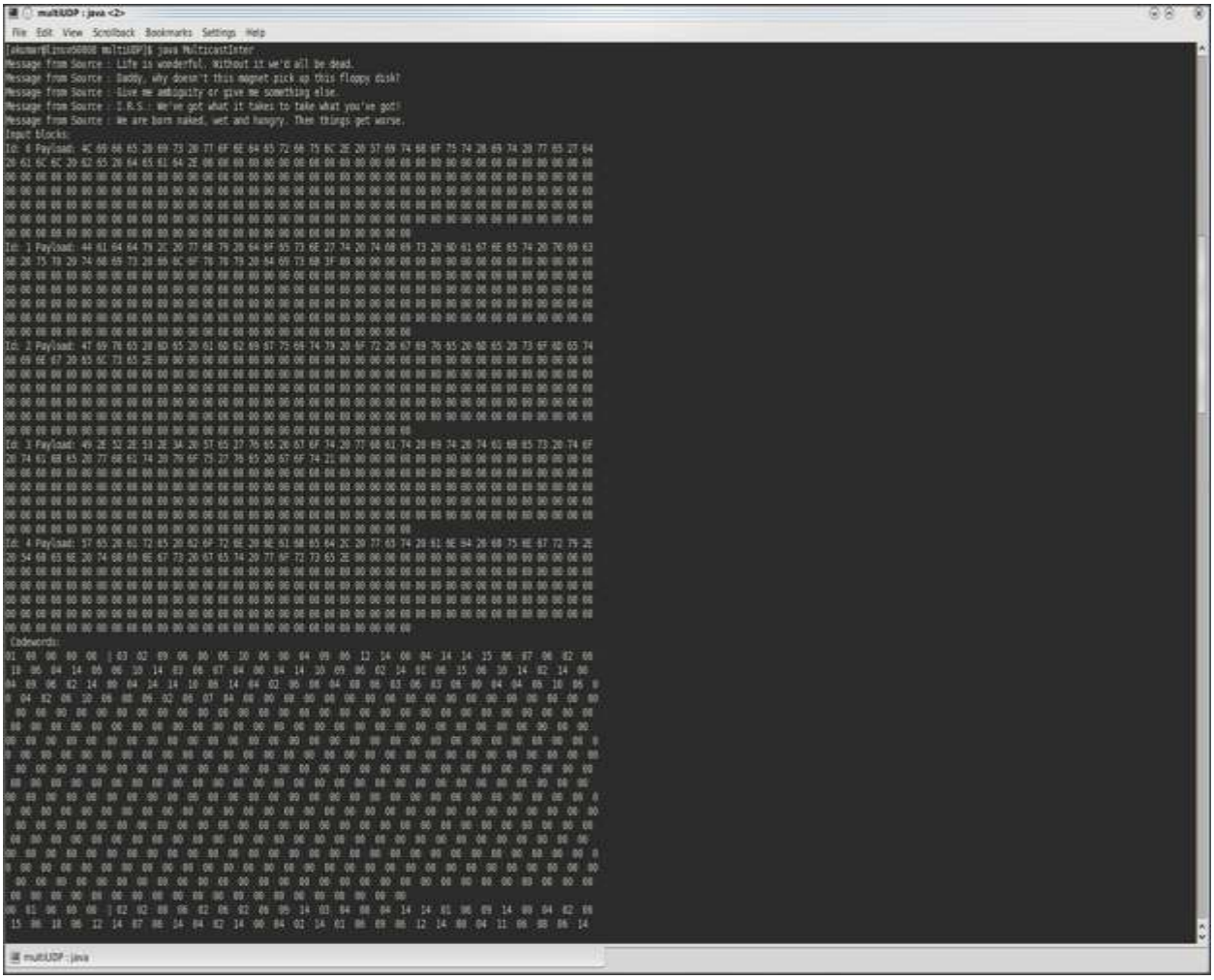
Data Analysis and Discussion

. Output generation and Analysis

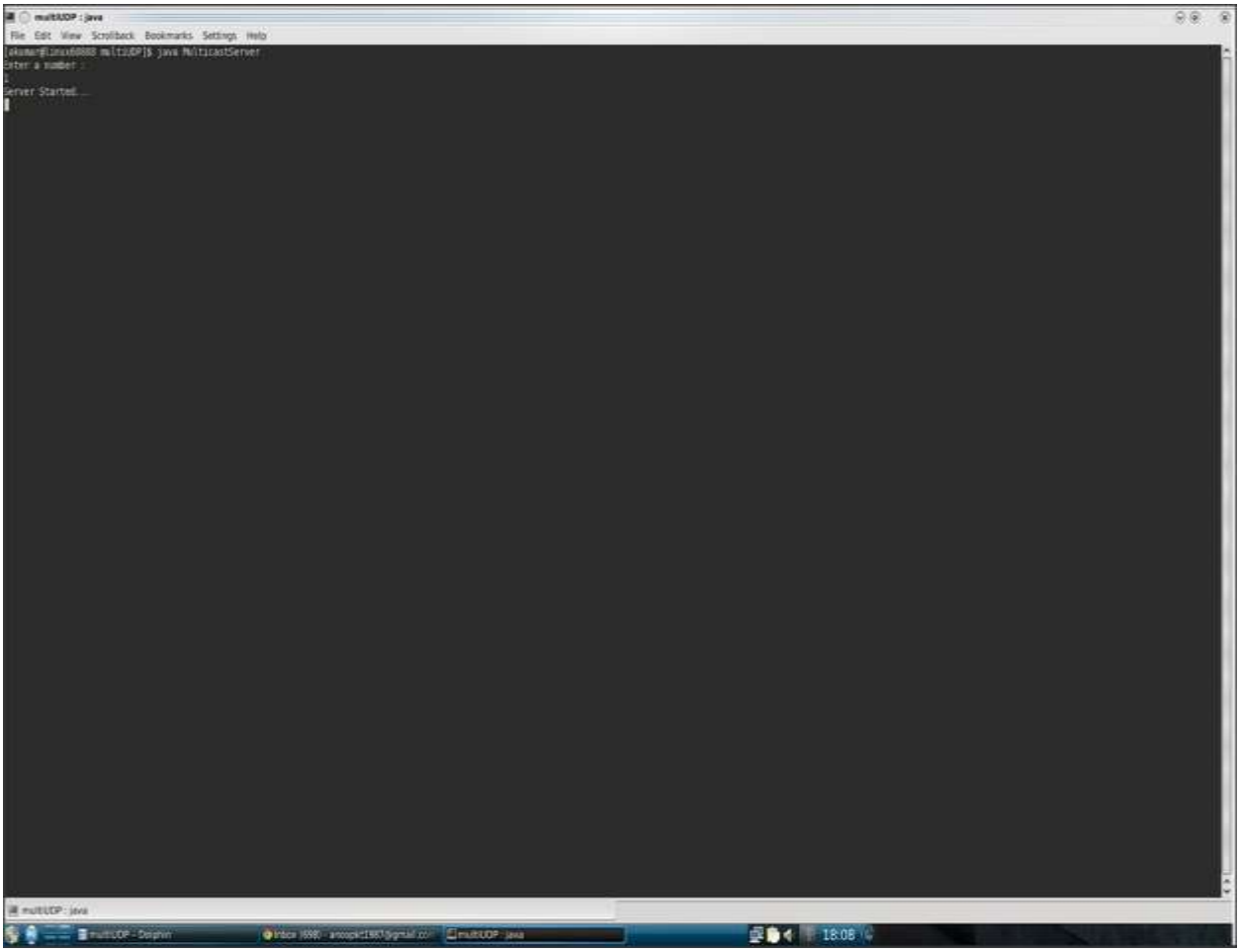
Snapshot for Client



Snapshot of Intermediate



Snapshot of Server



Compare output against hypothesis

We started off with the following Hypothesis:- If random linear combinations of the packets are performed in the intermediate nodes, then separate transmissions for each and every message (like that in the normal store-and-forward case) is not required These transmissions, when network encoded, will prove to be more robust and therefore allow for less error and a reduced need to retransmit packets, when compared to the traditional, single-connection method implemented in most networking applications today We hope to be able to prove that network coding is a more versatile, reliable, and efficient means of communication than sending data from node to node by using “traditional” (single data stream) methods utilized today.

Hence when we run our project we know that our Network is more versatile, more reliable, and efficient because we can now can have data from multiple inputs, reach the destination at the same point in time, without any lag.

. Abnormal case explanation

In case of point to point communication, it does not make sense we implementing the Network coding. Because there would exist no intermediate nodes, for the Linear coding to take place, hence we cannot implement Linear Network coding in Point to Point connections.

- **Conclusions and Recommendations**

Network coding can be used to increase throughput and robustness as well as reduce storage requirements, delay, and energy consumption
The difficulty of P2P content distribution is finding an optimal block scheduling algorithm, which should minimize the file down loading time in a distributed manner. This becomes even more challenging when the overlay network is changing dynamically From the project we see that by introducing Network coding we try to overcome the traditional problems associated with Wireless networks which suffer from a variety of unique problems such as low throughput, dead spots, and inadequate support for mobility.

Without network coding, each peer has to decide which blocks to download from which neighbors, based on local information only. This May be suboptimal due to lack of a global view, since local rarest blocks may not be global rarest blocks In contrast, with network coding, all coded blocks are almost equally useful to any peer, and there is no need to locate and request global rarest blocks in the system. As such, an information Bottleneck is avoided, which in turn reduces the file downloading time.

Another important benefit of network coding is the robustness to peer departures Without network coding, it is possible that a few data blocks are lost, due to departures of the server as well as the peers who have these data blocks In this unfortunate event, the original file cannot be reconstructed by the remaining peers On the other hand, with network coding, the risk of losing certain data blocks is no longer a concern Intuitively, since data blocks are mixed together, each of them is spreading to a large number of code db locks in the system.

In summary, the use of network coding solves the problems with wireless in a surprisingly simple and effective manner, leading to a shorter file downloading time. Moreover; it provides desirable robustness to peer departures. There is no need to worry about losing certain data blocks anymore. All of these contribute to the usefulness of network coding in P2P content distribution.

Linear Network coding is an Alternative to forward error correction and ARQ.

It is Resilient. Even if one of the nodes fails, data still reached destination. Since the throughput is increased the total number of hops reduce.

Recommendations for future studies

Ideas regarding implementing the Linear Network Coding should be implemented in mesh networks.

Importance was given to the resilience in our network. Further studies can be done to incorporate both resilience and throughput within the same network.

- # Bibliography

1 Network Coding Applications by Christina Fragouli and Emina Soljanin
http://arniepflch/contents/courses/monoIIpdf

2 Theory of Secure Network Coding by Ning Cai and Terence Chan
http://ieeexploreieeeorg/stamp/stampjsp?tp=&arnumber=5699897

3 Linear Network Coding : Theory and Algorithms by Shuo-Yen Robert Li, Fellow IEEE, Qifu Tyler Sun, Member IEEE, and
Ziyu Shao, Member IEEE
http://ieeexploreieeeorg/stamp/stampjsp?tp=&arnumber=5682367

4 A random Linear Network Coding Approach to Multicast by Tracey Ho, Member, IEEE, Muriel Médard, Senior Member, IEEE, Ralf Koetter, Senior Member, IEEE,
David R Karger, Associate Member, IEEE, Michelle Effros, Senior Member, IEEE, Jun Shi, and Ben Leong
http://ieeexploreieeeorg/stamp/stampjsp?tp=&arnumber=1705002

5 Wikipedia - http://enwikipediaorg/wiki/Linear_network_coding