# Research Philosophy

Andrew Rosen

## 1 Executive Summary

### 1.1 Motivation

I research and develop fault-tolerant systems because they are absolutely critical to the Internet and networking, which have become an essential part of our lives. Fault-tolerant and robust systems are what hold a complex environment like the Internet together. The Internet is made up of millions of computers/ billions of devices, all going wrong all the time. The packets are dropped, the routers are poisoned, the hard drives keep crashing, some suit/poor sod tripped over a power cable, the software is full of bugs, we meant to add security but it was adopted before we got the chance, and nobody can connect to the printer. It is, in a word, chaos. Complete chaos that works because smart people decided that failure was an assumption baked into every layer of every protocol.

Fault-tolerance is in many ways the most important part of what makes the Internet and networking work in general. As companies grow and create larger and larger intranets and datacenters, as more people gain access to the Intenet, adding billions of handheld devices that rival the power of last decade's consumer desktop, this problem is only going to get magnified.

### 1.2 Broad Overview

My research primarily focuses on Distributed Hash Tables (DHTs), with applications towards peer-to-peer (P2P) networks, Big Data, and Distributed Computing. DHTs are primarily used in P2P applications due to their decentralization and robustness . These qualities make them well suited for applications such as file-sharing, content distribution, multi-player video games, botnets, and video chat.

My past and current research focuses on using DHTs for distributed computing [2] and defining a DHT at its most abstract level. There are many different types of DHTs, but all DHTs share very important qualities. They are *scalable*, which means that each additional node in the network minimally impacts the cost of keeping the network organized. DHTs are also highly *fault-tolerant*. Unlike many other systems, DHTs assume that nodes will be continuously entering and leaving the network. Because of the way DHTs are organized, they can handle large scale failures, such as a power outage affecting an entire city. The last quality of DHTs is that they are *load-balancing*. This means the data stored in a DHT is evenly distributed among nodes in the network.

I found that DHTs can be mapped to Voronoi Tessellation and Delaunay Triangulation, and my coauthors and I created an efficient greedy heuristic for calculating the Voronoi Tessellation.

## 2 Past and Future Work

#### Ramifications

Our research would have many applications within the field of Computer Science. Traditional MapReduce frameworks and other distributed computing solutions are limited to being deployed solely in a datacenter [1]. Distributed computing frameworks built with UrDHT would be able to be deployed any context, be it a data center, a P2P network, or a volunteer computing pool. Furthermore, the autonomous load-balancing feature would allow frameworks to automatically give nodes with more power more work.

Our framework would not only be of interest to computer scientists. The "plug-and-play" nature we envision would make it easy for experts in other fields to use our framework to solve computationally intensive problems. Some example problems are Monte-Carlo Methods (of interest to mathematicians and economists) and machine learning (of interest to biologists).

This project would allow both organizations with large amounts of computing power and average developers with fewer resources to spend less time setting up and configuring their hardware to work together. The goal is for our software to make distributed computing more of matter of "plug-and-play," allowing researchers to spend less time setting up and maintaining computing platforms.

# 3 Misc

I find great value in abstracting techniques and platforms even further, as that makes it possible to see novel applications for various technologies. This can be seen in my work on Distributed Hash Tables, where I showed that DHTs can be mapped to Voronoi Tessellations and Delaunay Triangulation.

We used this insight to create UrDHT [3], a generalized framework for building distributed hash tables. UrDHT essentially acts as a "fill-in-the-blanks" for creating distributed hash tables. The novelty of UrDHT is that it makes it easy for scientists in any field to construct a DHT based application. For computer scientists, UrDHT provides a means of rigorously comparing different architectures.

I believe that collaboration is vital to academic growth. I have worked with researchers in both Computer Science and other scientific fields, including Astronomy, Biology, and Psychology.

A large part of why I do research is that I want to help people. When I started doing undergraduate research, I worked in the Sonification Lab at Georgia Tech. There, I helped PhD students implement their research in creating non-traditional user-interfaces geared towards the visually impaired community.

# 4 Future research

# References


[1] Virtual hadoop. http://wiki.apache.org/hadoop/Virtual

[2] Andrew Rosen, Brendan Benshoof, Robert W Harrison, and Anu G. Bourgeois. Mapreduce on a chord distributed hash table. In *2nd International IBM Cloud Academy Conference*.

[3] Andrew Rosen, Brendan Benshoof, Robert W Harrison, and Anu G. Bourgeois. Urdht. `https://github.com/UrDHT/`.