Market Fragmentation and the Latency Arms Race
Chris Rockwell, Erik Brinkman, Elaine Wah and Michael P. Wellman
University of Michigan, Computer Science and Engineering

INTRODUCTION

High frequency trading (HFT) is the rapid trading of securities executed by computers. It is important today as it controls over 50% of total trading volume in stock markets. Our research examines a specific type of HFT prevalent in markets today, latency arbitrage.

Latency arbitrage takes advantage of the latency (delay) across markets in attempt to generate income. In other words, Latency arbitrageurs get access to information from multiple markets faster than other traders, and can exploit this to find small price disparities between markets to make nearly risk-free profit. Latency arbitrage has been estimated to account for $21 billion in profit per year.

Currently there are over 40 trading venues in the U.S. used to trade stocks, each keeping track of its highest bid and lowest sell prices. Because of this market fragmentation, there is the possibility of price disparities across markets. Regulation NMS was created to mitigate this, as it routes orders for best execution. This means upon submitting orders, traders have access to all other markets through the National Best Bid and Offer (NBBO).

We study the phenomenon of a latency arms race, where HFTs try to keep reducing their latencies, approaching none at all. In reality, HFT firms spend millions of dollars investing in technology to reduce latency, and we model this by looking at revenues in our model to predict how much arbitrageurs value decreasing latency.

Currently there are over 40 trading venues in the U.S. used to trade stocks, each keeping track of its highest bid and lowest sell prices. Because of this market fragmentation, there is the possibility of price disparities across markets. Regulation NMS was created to mitigate this, as it routes orders for best execution. This means upon submitting orders, traders have access to all other markets through the National Best Bid and Offer (NBBO).

Background

The SIP has the aggregated lowest sell, or ask price among all markets, as well as the highest bid price, called the National Best Bid and Offer (NBBO).

Methodology

We simulate a two-market model with two latency arbitrage agents, measuring revenues and varying latencies between simulations. The two general scenarios are:

[a]. Both agents at an equal latency
[b]. One agent at a latency less than the other agent

We will compare LA revenue between different latencies and will use Empirical Game Theoretic Analysis (EGTA) to make sure fundamental agents respond appropriately to varying HFT strategies.

Our focus is on what reaction each HFT agent will take to different situations and how these decisions will change the shared revenue of the two agents.