Data Mining in Mobile Telecom

Dr. Chris Groër
**LINK ANALYTICS**

**WE ARE BIG DATA**
AN ANALYTICAL SOLUTIONS COMPANY, WE KNOW THE KEY TO YOUR SUCCESS – DATA.

**WE KNOW ANALYTICS**
A TEAM OF MATHEMATICIANS, DATA SCIENTISTS & MARKETERS, WE PRODUCE USEFUL SOLUTIONS FOR MOBILE, GOV’T, ENERGY & AUTOMOTIVE.

**WHAT WE PRODUCE**
WE PROVIDE ANALYTICAL CONSULTING SERVICES, MODELING & SEGMENTATION. WE ALSO BUILD CUSTOMIZED SOLUTIONS FOR OUR CLIENTS TO ENABLE REPEATABILITY & SCALE.

**OUR FOUNDATION**
A JOINT VENTURE WITH MARKET STRATEGIES, LINK WAS BUILT FROM THE GROUND UP TO INTEGRATE DATA- ATTITUDINAL, BEHAVIORAL, & TRANSACTIONAL- ALL CRITICAL FOR SMARTER ORGANIZATIONAL DECISION MAKING.
Problems in Mobile Telecom

• Domestic mobile phone market is completely saturated and extremely competitive

• Carriers can gain new customers only by taking subscribers from other networks and they must protect their current customer base from competitive threats

• How am I doing against my competitors in various markets?
• Where are my new customers coming from?
• Where did my customers go?
• The domestic prepaid market is growing – can it be understood?
• Who are my important customers (influencers)?
Problems in Mobile Telecom

- One huge advantage for us is that the data is exceptionally “clean”
- Call Detail Record (CDR) processing has been refined over the years as it serves as the underlying data source for billing
- We often find anomalies in the data that are reported to fraud departments
The Mobile Social Graph

- A primary object of study is the *Mobile Social Graph* $G(t)$ – nodes are subscribers, edges are relationships or links (voice, SMS)
- $G(t)$ has 300 million nodes and tens of billions of edges
Data Mining Efforts

- Our Data Engineers post-process these CDRS to create a proprietary labeled version of $G(t)$ that summarizes many features and greatly reduces the size.
- We attempt to analyze this object to produce answers to very important “CMO” level questions:
  - Why did I lose 200,000 subscribers last month in Chicago?
  - What happened when Sprint got the iPhone in October?
  - Did my last 5 million dollar ad campaign in Southern California have an impact?
  - What happened when my competitor offered this new trial family plan in New England?
Data Mining Efforts

• Using data that has been “available” for decades, we perform a few fundamental calculations each month:
  • Estimation of competitor’s subscriber counts
  • Prediction of subscriber movement amongst carriers
  • Social Fingerprinting: Each mobile subscriber has a unique pattern of calling behavior
    – Identifying an individual after he changes his number (and/or carrier) has a variety of applications: Fraud detection, Prepaid movement, Churn movement, Cannibalization, what else?
Some Final Thoughts

- FedEx likes us: “Never underestimate the bandwidth of a station wagon full of tapes hurtling down the highway” (Andrew Tanenbaum)
- Privacy and data security are paramount concerns – all data is encrypted in transit and at rest
- These concerns limit what we can do with the data and where we can do it
- Mobile industry is full of massive, extremely rich data sets, and extremely difficult, interesting, and important problems