Industry Practice Expo Invited Talk

Adaptive Adversaries: Building Systems to Fight Fraud and Cyber Intruders

Ari Gesher
Palantir
regs@palantir.com

ABSTRACT

Statistical machine learning / knowledge discovery techniques tend to fail when faced with an adaptive adversary attempting to evade detection in the data. Humans do an excellent job of correctly spotting adaptive adversaries given a good way to digest the data. On the other hand, humans are glacially slow and error-prone when it comes to moving through very large volumes of data, a task best left to the machines.

Fighting complex fraud and cyber-security threats requires a symbiosis between the computers and teams of human analysts. The computers use algorithmic analysis, heuristics, and/or statistical characterization to find interesting 'simple' patterns in the data. These candidate events are then queued for in-depth human analysis in rich, expressive, interactive analysis environments.

In this talk, we'll take a look at case studies of three different systems, using a partnership of automation and human analysis on large scale data to find the clandestine human behavior that these datasets hold, including a discussion of the backend systems architecture and a demo of the interactive analysis environment.

The backend systems architecture is a mix of open source technologies, like Cassandra, Lucene, and Hadoop, and some new components that bind them all together. The interactive analysis environment allows seamless pivoting between semantic, geospatial, and temporal analysis with a powerful GUI interface that's usable by non-data scientists. The systems are real systems currently in use by commercial banks, pharmaceutical companies, and governments.

Categories and Subject Descriptors

H.2.8 [Database Management]: Data Mining.

General Terms

Algorithms, Performance

Keywords

Intelligence Augmentation, Machine Learning, Knowledge Discovery, Cyber Security.

Bio

Ari Gesher is a senior engineer and Engineering Ambassador at Palantir Technologies. An alumnus of the University of Illinois computer science department, Ari has worked in the software industry for the past fifteen years, including a stint as the lead engineer for the SourceForge.net open source software archive. At Palantir Technologies, Ari has split his time between working as a backend engineer on Palantir's analysis platform, thinking and writing about Palantir's vision for human-driven information data systems, and moonlighting on Palantir's Philanthropic engineering team. Ari is in demand as a speaker on the topic of big data and the limits of automated decision- making. In the past year, he's spoken at Harvard Business School, the Institure for the Future's Tech Horizons conference, multiple O'Reilly Strata Big Data Conferences, the Economist Future Technologies Summit, and PayPal's TechXploration series.

KDD'13, August 11–14, 2013, Chicago, Illinois, USA.
ACM 978-1-4503-2174-7/13/08.