DLI Accelerated Data Science Teaching Kit

Lecture 1.7 - Example Data Science Project 2: NetProbe Auction Fraud Detection

The Accelerated Data Science Teaching Kit is licensed by NVIDIA, Georgia Institute of Technology, and Prairie View A&M University under the
Creative Commons Attribution-NonCommercial 4.0 International License.

NetProbe:
Fraud Detection in Online Auction

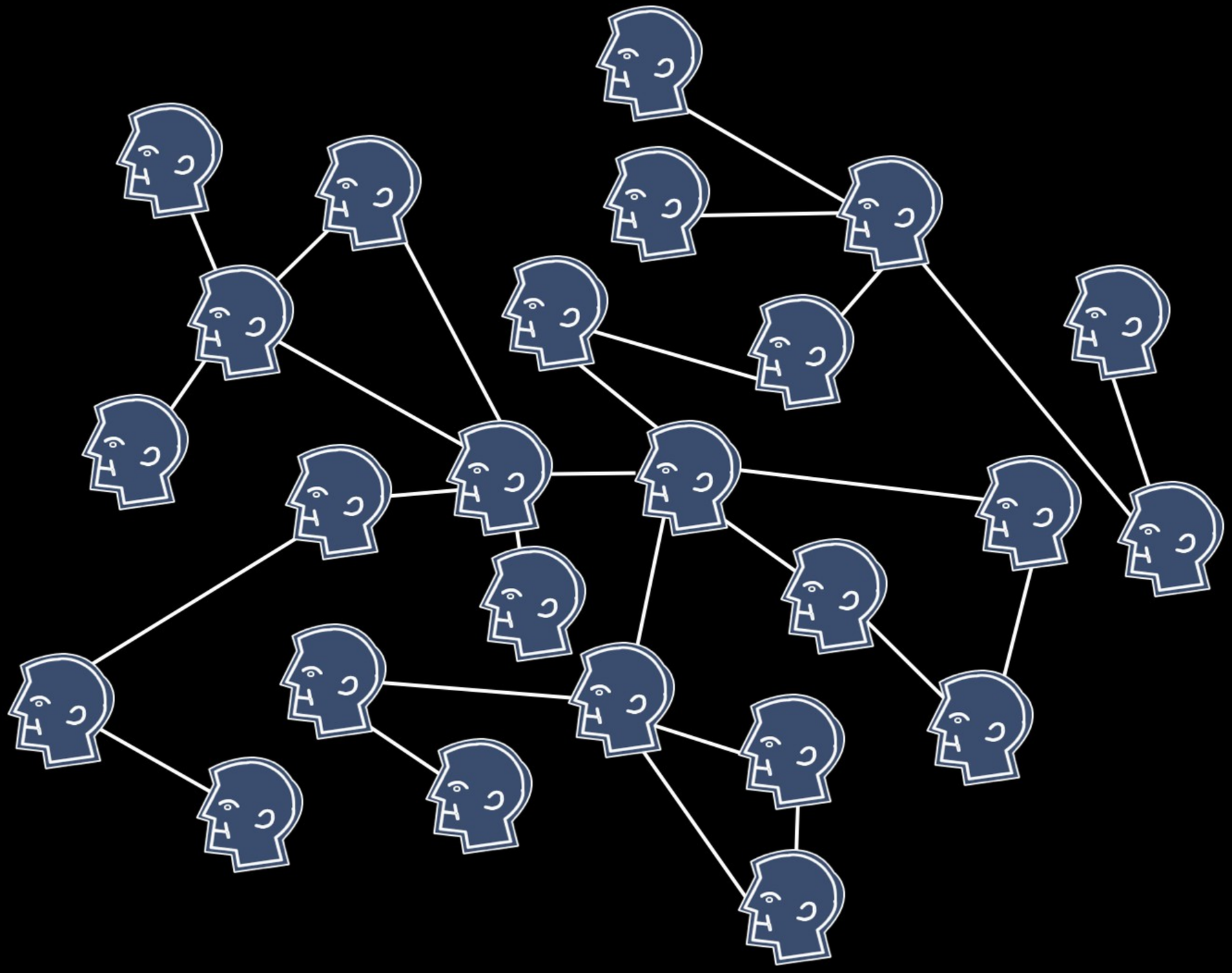
NetProbe: A Fast and Scalable System for Fraud Detection in Online Auction Networks.
Shashank Pandit, Duen Horng (Polo) Chau, Samuel Wang, Christos Faloutsos. WWW 2007

NetProbe: The Problem

Find **bad sellers** (**fraudsters**) on eBay who don't deliver their items

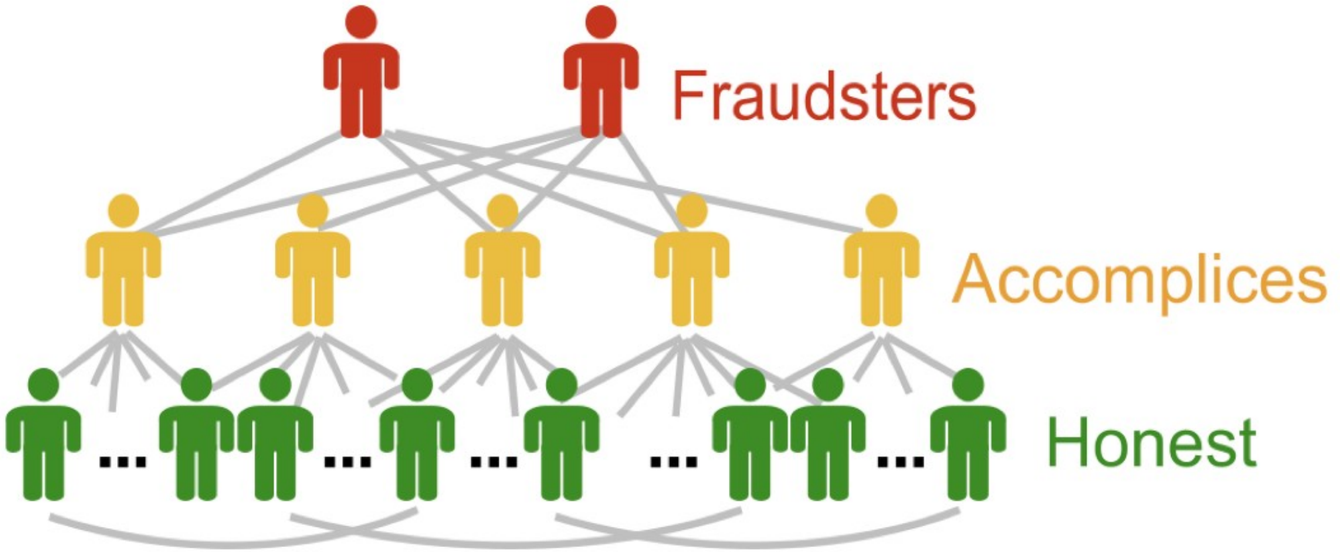
Non-delivery fraud is a common auction fraud

source: https://www.fbi.gov/contact-us/field-offices/portland/news/press-releases/fbi-tech-tuesday---building-a-digital-defense-against-auction-fraud



NetProbe: Key Ideas

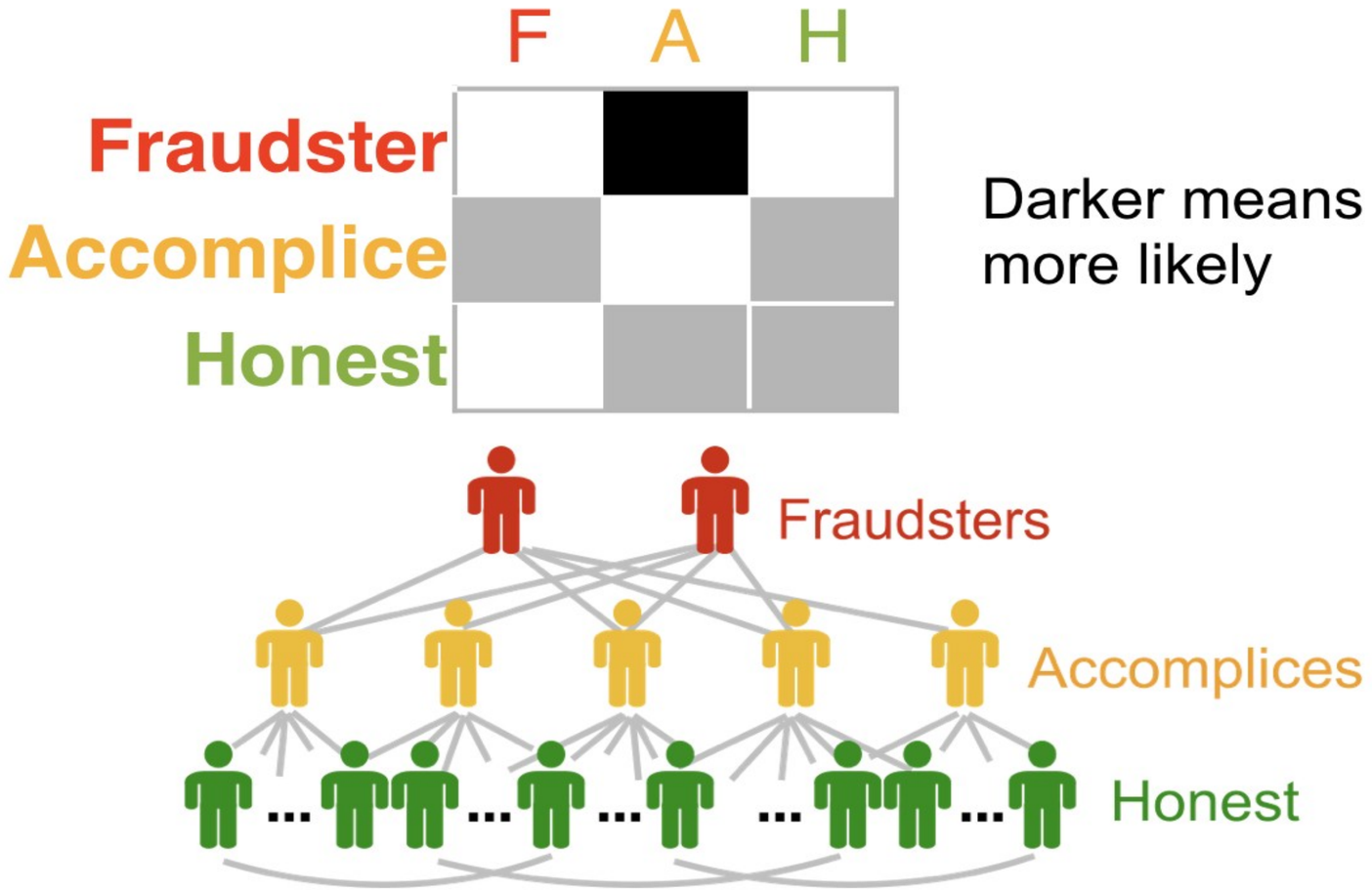
- Fraudsters **fabricate their reputation** by "trading" with their accomplices
- Fake transactions form **near bipartite cores**
- How to detect them?



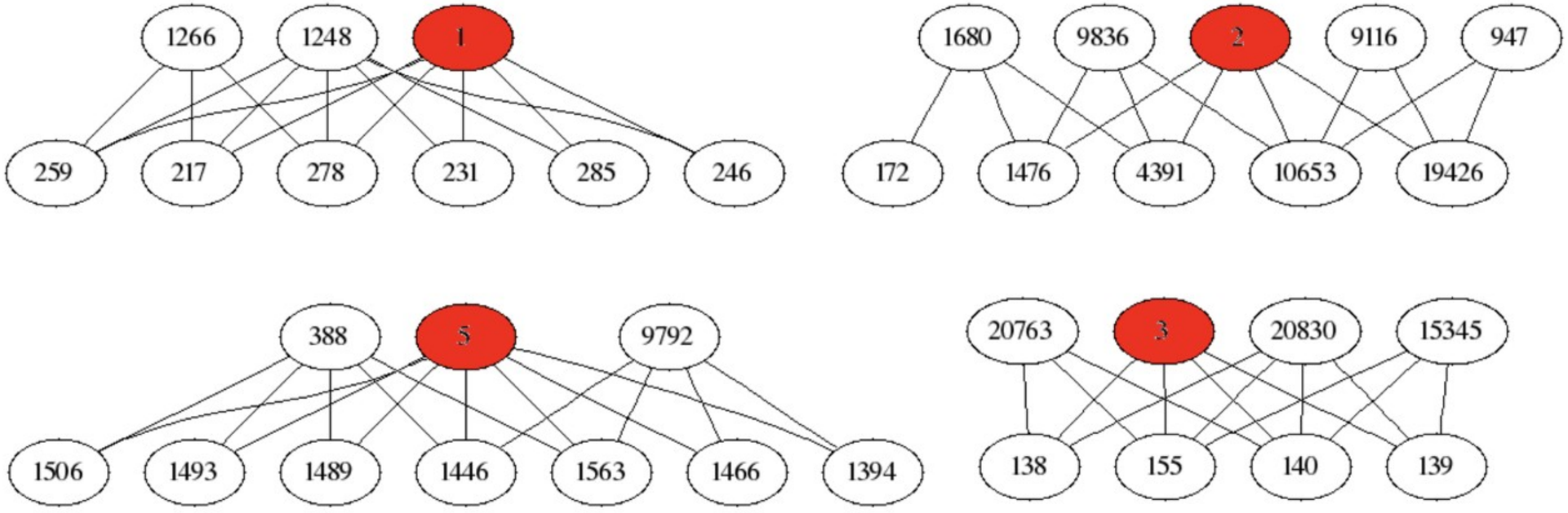
NVIDIA DEEP LEARNING INSTITUTE | Georgia Tech | PRAIRIE VIEW A&M UNIVERSITY

NetProbe: Key Ideas

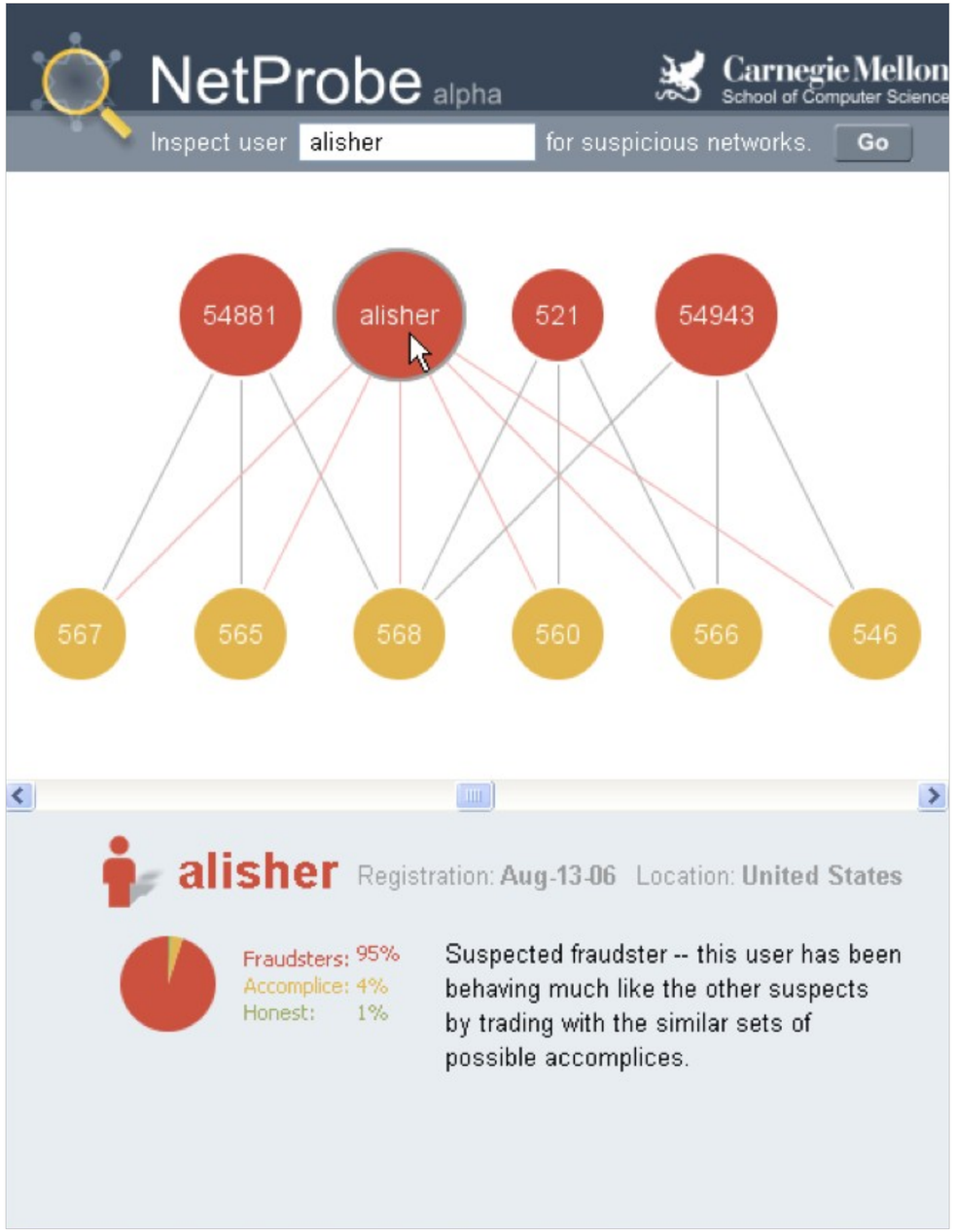
Use **Belief Propagation**



NetProbe: Main Results



NVIDIA DEEP LEARNING INSTITUTE | Georgia Tech | GTx



NVIDIA DEEP LEARNING INSTITUTE
Georgia Tech
GTx

What did **NetProbe** go through?

Collection	Scraping (built a “scraper”/“crawler”)
Cleaning	
Integration	
Analysis	Design detection algorithm
Visualization	
Presentation	Paper, talks, lectures
Dissemination	Not released

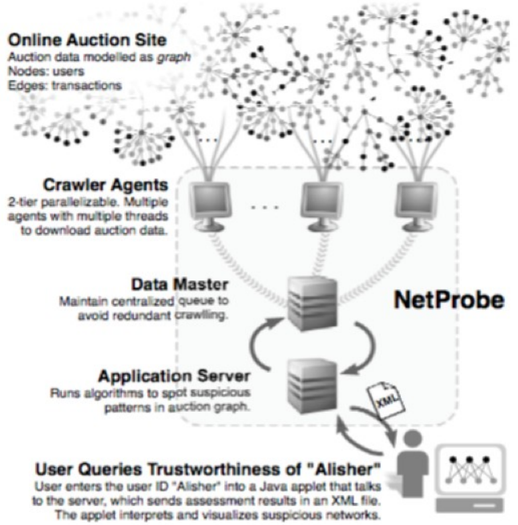
NVIDIA DEEP LEARNING INSTITUTE | Georgia Tech | GTx

NetProbe: A Fast and Scalable System for Fraud Detection in Online Auction Networks

Shashank Pandit, Duen Horng Chau, Samuel Wang, Christos Faloutsos
Carnegie Mellon University
Pittsburgh, PA 15213, USA
{shashank, dchau, samuelwang, christos}@cs.cmu.edu

ABSTRACT

Given a large online network of online auction users and their histories of transactions, how can we spot anomalies and auction fraud? This paper describes the design and implementation of NetProbe, a system that we propose for solving this problem. NetProbe models auction users and transactions as a *Markov Random Field* tuned to detect the suspicious patterns that fraudsters create, and employs a *Belief Propagation* mechanism to detect likely fraudsters. Our experiments show that NetProbe is both efficient and effective for fraud detection. We report experiments on synthetic graphs with as many as 7,000 nodes and 30,000 edges, where NetProbe was able to spot fraudulent nodes with over 90% precision and recall, within a matter of seconds. We also report experiments on a real dataset crawled from eBay, with nearly 700,000 transactions between more than 66,000 users, where NetProbe was highly effective at unearthing hidden networks of fraudsters, within a realistic response time of about 6 minutes. For scenarios where the underlying data is dynamic in nature, we propose *Incremental NetProbe*, which is an approximate, but fast, variant of NetProbe. Our experiments prove that Incremental NetProbe



NetProbe: A Fast and Scalable System for Fraud Detection in Online Auction Networks. Shashank Pandit, Duen Horng (Polo) Chau, Samuel Wang, Christos Faloutsos. *International Conference on World Wide Web (WWW) 2007*. May 8-12, 2007. Banff, Alberta, Canada. Pages 201-210.

Categories and Subject Descriptors

1. INTRODUCTION

NVIDIA. DEEP LEARNING INSTITUTE | Georgia Tech | GTx

DLI Accelerated Data Science Teaching Kit

Thank You