GRAPH-BASED USER BEHAVIOR MODELING
PREDICTION TO FRAUD DETECTION

Alex Beutel
Leman Akoglu
Christos Faloutsos

bit.ly/kdd2015_userbehavior
KDD 2015 - August 10, 2015 - Sydney, Australia
Thanks to

NSF Grant No. IIS-1408924, IIS-1408287, CAREER 1452425, DGE-1252522, ...
User Behavior Challenges
User Behavior Challenges

Three Main Questions:

1. How can we understand typical/normal user behavior in a graph?
User Behavior Challenges

Three Main Questions:

1. How can we understand typical/normal user behavior in a graph?

2. How can we find suspicious user behavior?
User Behavior Challenges

Three Main Questions:

1. How can we understand typical/normal user behavior in a graph?
2. How can we find suspicious user behavior?
3. How can we distinguish the two?
Graphs of User Behavior

- Undirected graphs

[Diagram showing a network of users with logos for Facebook, LinkedIn, and dblp]
Graphs of User Behavior

- Undirected graphs
- Directed graphs
Graphs of User Behavior

- Undirected graphs
- Directed graphs
- Bipartite graphs
Graphs of User Behavior

- Undirected graphs
- Directed graphs
- Bipartite graphs
- Node attributes
Graphs of User Behavior

- Undirected graphs
- Directed graphs
- Bipartite graphs
- Node attributes
- Edge attributes

A. Beutel, L. Akoglu, C. Faloutsos
KDD 2015

Graphs of User Behavior:

- Undirected graphs
- Directed graphs
- Bipartite graphs
- Node attributes
- Edge attributes
Graphs of User Behavior

- Undirected graphs
- Directed graphs
- Bipartite graphs
- Node attributes
- Edge attributes
Graphs of User Behavior

- Undirected graphs
- Directed graphs
- Bipartite graphs
- Node attributes
- Edge attributes

Examples:
- Netflix
- Amazon
- Yelp
- Google Play
- Newegg
- Reddit
Graphs of User Behavior

- Undirected graphs
- Directed graphs
- Bipartite graphs
- Node attributes
- Edge attributes

Semi-supervised

Logos: Netflix, Amazon, Yelp, Google Play, Newegg, Reddit
Graphs of User Behavior

- Undirected
- Directed
- Bipartite

- Node Attributes
- Edge Attributes
- Unsupervised
  - Semi-Supervised
Graphs of User Behavior

- User icons connected to movie icons, followed by a graph matrix representation.
Graphs of User Behavior
Graphs of User Behavior
Modeling “Normal” Behavior

- Predict edges
Modeling User Behavior

- Predict edges
- Predict node attributes
- Predict edge attributes
Modeling User Behavior

- Predict edges
- Predict node attributes
- Predict edge attributes
  - “Netflix Problem”
Modeling User Behavior

- Predict edges
- Predict node attributes
- Predict edge attributes
  - “Netflix Problem”
- Frequent Itemset Mining & Community Detection
Modeling User Behavior

- Predict edges
- Predict node attributes
- Predict edge attributes
  - “Netflix Problem”
- Frequent Itemset Mining & Community Detection
- Fraud Detection

Deceives users and manipulates recommendations!
Modeling User Behavior

Modeling normal users and detecting anomalies are two sides of the same coin – understanding user behavior.
Modeling User Behavior

Modeling normal users and detecting anomalies are two sides of the same coin – understanding user behavior.

Rough model of normal – detect general outliers
Modeling User Behavior

Modeling normal users and detecting anomalies are two sides of the same coin – understanding user behavior.

More fine grained model of normal can find more subtle outliers
Modeling User Behavior

Modeling normal users and detecting anomalies are two sides of the same coin – understanding user behavior.

More complex model can capture both normal and abnormal patterns – micro-clusters with small variance are particularly suspicious.
Modeling User Behavior

Modeling normal users and detecting anomalies are two sides of the same coin – understanding user behavior.

Sometimes domain experts know a specific pattern is fraudulent, and we can search for exactly that pattern.
THREE MAIN TECHNIQUES

1. Local Subgraph Analysis: Patterns and Features

2. Global: Propagation Methods

FOR ALL THREE PARTS

a) Background

b) Normal

c) Abnormal