FraudVis: Understanding Unsupervised Fraud Detection Algorithms

Jiao Sun¹, Qixin Zhu¹, Zhifei Liu¹, Xin Liu¹, Jihae Lee¹, Lei Shi², Zhigang Su³, Ling Huang¹ and Wei Xu¹

¹ Institute of Interdisciplinary Information Sciences, Tsinghua University
² SKLCS, Institute of Software, Chinese Academy of Sciences
³ JD.com
Motivation

- Great loss caused by fraud users
  - Various kinds of fraud behavior
- Difficulty in distinguishing fraud users from normal users
  - Camouflage in individuals
  - Collaborative behavior
- Algorithm itself is hard to interpret
  - Feature selection is a black-box to user
  - The origin of “group” and why a group is abnormal
What are fraudsters?

<table>
<thead>
<tr>
<th>id</th>
<th>ostype</th>
<th>ip_prefix_22</th>
<th>phone_city</th>
<th>phone_prefix_3</th>
<th>duration_quantized_20</th>
<th>label</th>
<th>id</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>good</td>
<td>101.90.252.0</td>
<td>上海</td>
<td>180</td>
<td>0</td>
<td>good</td>
<td>30</td>
</tr>
<tr>
<td>40</td>
<td>good</td>
<td>114.95.156.0</td>
<td>上海</td>
<td>131</td>
<td>0</td>
<td>good</td>
<td>40</td>
</tr>
<tr>
<td>55</td>
<td>good</td>
<td>101.229.16.0</td>
<td>上海</td>
<td>137</td>
<td>0</td>
<td>good</td>
<td>55</td>
</tr>
<tr>
<td>63</td>
<td>good</td>
<td>1.197.176.0</td>
<td>上海</td>
<td>138</td>
<td>0</td>
<td>good</td>
<td>63</td>
</tr>
<tr>
<td>200</td>
<td>good</td>
<td>117.136.8.0</td>
<td>上海</td>
<td>134</td>
<td>0</td>
<td>good</td>
<td>200</td>
</tr>
<tr>
<td>299</td>
<td>good</td>
<td>180.172.168.0</td>
<td>上海</td>
<td>156</td>
<td>0</td>
<td>good</td>
<td>299</td>
</tr>
<tr>
<td>458</td>
<td>bad</td>
<td>113.122.4.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>458</td>
</tr>
<tr>
<td>474</td>
<td>bad</td>
<td>119.122.96.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>474</td>
</tr>
<tr>
<td>491</td>
<td>bad</td>
<td>58.54.220.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>491</td>
</tr>
<tr>
<td>498</td>
<td>bad</td>
<td>117.66.192.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>498</td>
</tr>
<tr>
<td>519</td>
<td>bad</td>
<td>114.106.224.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>519</td>
</tr>
<tr>
<td>529</td>
<td>bad</td>
<td>112.113.60.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>529</td>
</tr>
<tr>
<td>531</td>
<td>bad</td>
<td>59.53.140.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>531</td>
</tr>
<tr>
<td>540</td>
<td>bad</td>
<td>27.27.172.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>540</td>
</tr>
<tr>
<td>544</td>
<td>bad</td>
<td>123.161.152.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>544</td>
</tr>
<tr>
<td>572</td>
<td>bad</td>
<td>114.239.0.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>572</td>
</tr>
<tr>
<td>614</td>
<td>bad</td>
<td>117.81.116.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>614</td>
</tr>
<tr>
<td>626</td>
<td>bad</td>
<td>120.32.128.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>626</td>
</tr>
<tr>
<td>639</td>
<td>bad</td>
<td>180.124.44.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>639</td>
</tr>
<tr>
<td>650</td>
<td>bad</td>
<td>117.93.140.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>650</td>
</tr>
<tr>
<td>652</td>
<td>bad</td>
<td>220.179.212.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>652</td>
</tr>
<tr>
<td>677</td>
<td>bad</td>
<td>180.122.20.0</td>
<td>上海</td>
<td>170</td>
<td>0</td>
<td>bad</td>
<td>677</td>
</tr>
</tbody>
</table>
Challenges

- High-dimensional datasets
  - Various kinds of fraud behavior
  - High-dimensional data for each log
- Selection of features and algorithms
  - Hard to evaluate which ones are useful
  - Heavily depends on the scenario
- No labels
  - No label for evaluation
  - High cost for false positive
Works for both algorithm experts and domain experts

- Why do users belong to the same group?
- What are the important features?
- What did they do as a fraud group?
- Do they have some correlations?
- Is the user good or not?

- What causes the form of a fraud group?
- What are the distributions of the important features?
- Do users in the same group share the same pattern?
- Will members in one group build a strange network?
- Did I make a mistake for this user?
Works for both algorithm experts and domain experts

- Why do users belong to the same group?
- What are the important features?
- What did they do as a fraud group?
- Do they have some correlations?
- Is the user good or not?

- What causes the form of a fraud group?
- What are the distributions of the important features?
- Do users in the same group share the same pattern?
- Will members in one group build a strange network?
- Did I make a mistake for this user?
Works for both algorithm experts and domain experts

- Why do users belong to the same group?
  - **What are the important features?**
  - What did they do as a fraud group?
  - **Do they have some correlations?**
  - Is the user good or not?

- What causes the form of a fraud group?
  - **What are the distributions of the important features?**
  - Do users in the same group share the same pattern?
  - Will members in one group build a strange network?
  - Did I make a mistake for this user?
Works for both algorithm experts and domain experts

- Why do users belong to the same group?
- What are the important features?
- What did they do as a fraud group?
- Do they have some correlations?
- Is the user good or not?

- What causes the form of a fraud group?
- What are the distributions of the important features?
- Do users in the same group share the same pattern?
- Will members in one group build a strange network?
- Did I make a mistake for this user?
We need VISUALIZATION!
Main contributions

- Comprehensive analysis
  - Inter-group, intra-group, individual
  - Correlation, temporal, spatial
- Visualization interpretation of algorithm result through customized interactions
  - Instructions
  - Different dashboards
- Evaluation through real-world data sets and algorithms
Workflow

Drill Down FraudVis Step By Step

Domain Experts & Customers
Overview Dashboard
Case study 2: E-commerce Website

- The distribution of the most important feature hardly differ from the overall distribution
- Members from different groups are mixed together and hard to separate
- Large fraud group containing many users that are not similar
Conclusion

- We solve two main problems
  - How to explain the fraud behavior to domain users with little technology background?
  - How to test the result of various fraud detection algorithms and discover the fundamental features?
- A fresh view and a working system to display high-dimensional fraud behaviors
- Visually interpret and compare the result of unsupervised fraud detection algorithms
The future of Fraud Detection

Good detection algorithms

+ VISULIZATION
Thanks & QA

Jiao Sun - https://sunjiao123sun.github.io
Email: j-sun16@mails.tsinghua.edu.cn
All kinds of collaboration are welcomed! 😊