AI Powered Security and Securing AI

Dr. Jimmy Su
Head of JD Security Research Center, Senior Director
Outline

- Intro of JD
- Black market in e-Commerce
- AI vs AI
- AI platform security
JD.COM Introduction

700 Million Items Sold

June Sales Event

Massive Scale

301.8M
Active customer accounts

160K
Active third-party vendors on JD platform

190K
Full-time employees

1.59B
Orders fulfilled in 2017
Redefining Retail Through Technology

Our First 12 Years

- Lowering costs
- Enhancing efficiency
- Improving user experience

The Next 12 Year: Leveraging AI

Creating a truly intelligent business

- Natural Language Processing
- Smart logistics
- Smart supply chain
- Financial technologies
- Cloud computing
How large is the black market in e-Commerce?

- 51.8% of the network traffic are bot traffic
- 28.9% of them are malicious bot traffic
- Billions of dollars lost for companies
- Many cases: DIDI, Uber, even Apple iOS has been targets in this black market
- E-Commerce companies have been largely targeted
  - Large promotions
  - Many coupons
  - Flash sales
Overview of This Black Market

- Traditionally, the black market full of manual labor and low technology.
- Now it is a complete industrial chain with AI driven technology and automated tools.
- Greatly undermines the reputation of e-Commerce companies
- Impact normal customers' shopping behaviors
- Ecosystem consists of upstream, midstream and downstream.
Upstream of the Black Market

- Verification code/image platform
  - Automate registration and login processes
- Account take-over data
  - Individually targeted high usage passwords
- Automated software
  - Timers for flash sales
- Proxy tools
  - Defeat IP based risk control strategies
Midstream of the Black Market

- Provides various accounts related services such as Instant Messaging (IM) groups or online forums.
- Fake account registration
- Account take-over
- Account washing
- Information exchange platform
- Trading platform
Downstream of the Black Market

- Gain profits and incur losses to normal users and e-Commerce companies
- Theft
- Fraud
- Blackmailing
- Click farming
- Scalper
Scalpers

- Scalping is a common threat to E-Commerce Platforms in China
  - Huge promotions
  - Alibaba, VIP, Suning, etc
  - User experience
- Monitoring Scalper activities and notify their activities in advance
- New AI based monitoring system is a strong plus to the legacy rule-based system
  - Accuracy of the information
  - New threat info acquisition never discovered before
What do Scalpers do?

Bot Detection  NLP  Reverse Engineering  Adversarial Machine Learning  Address Clustering
Fraudulent Cell Phone Orders
Bulk Registration

- Using tools or simulators
- Underground economy chain
  - Access code service
  - SMS verification service
  - Fake ID
- Features to detect
  - Behaviors of bots
  - Rush registration
  - Fake information
Account Trading

- Price by account types
- Trading platform
  - E-commercial websites
  - IM Groups (QQ, Wechat)
  - Personal Websites
Account Security: AI Empowered

- APP security SDK
  - User biometrics
  - Bot detection during registration and login
  - Device fingerprints
- Anomaly detection
  - Threat intelligence
  - User behaviors
- Business validation and feedback
  - Order risk control
Bot Detection Using Biometrics

• Many scenarios at JD would benefit from distinguishing between bot and human
  . Bot account registration
  . Bot placing an order
  . Bot crawling our site to extract pricing info

• Exploring biometrics features including mouse movement and keyboard

POP QUIZ: Can you identify the bot mouse movement graphs from the human ones?
Bot Detection
Attackers using AI to Lower Costs

- Intermediate level CAPTCHA is a solved problem in the Black Market.
CAPTCHA Solving AI Platforms

Tested on eight most popular CAPTCHA solving platforms in the Black Market

- 65% accuracy
- 71% accuracy
- 42% accuracy
- 49% accuracy
- 76% accuracy
- 68% accuracy
Based on CAPTCHA solving platform accuracies, we use GAN to generate adverserial samples.

GAN model combines the best features of various CAPTCHAS
GAN Adversarial Samples

- Lower CAPTCHA platform accuracy to 12%!
- Must balance user experience with CAPTCHA difficulty
AI平台安全隐患

恶意数据类型
(Malformed Image
Malicious PDF)

AI平台在提取特征时
需要Parse这些文件
从而引发漏洞

框架平台漏洞
(开发代码依赖库)

Tensorflow – 887K行代码 – 97个依赖库
Caffe – 127K行代码 – 137个依赖库
Torch – 590K行代码 – 48个依赖库

包含Heap Overflow, DOS, integer overflow等漏洞

模型攻击
(第三方模型
模型重用攻击)
AI平台数据隐私保护

- GDPR (Global Data Protection Regulation) 从2018/05起在EU生效
  - 对包含用户数据的训练集的使用有更加严格的限制
  - 对AI做出的决定的可解释行有更加严格的限制
  - 对数据中出现的歧视现象的限制
  - 需要相应的机制去监控这些条例没有被违反

- 原始训练数据的泄漏
  - AI平台的Attack Surface之广 包含漏洞之多 隐患很大
  - JD的隐私数据种类多 数量大 用于AI的更不少

- 挑战
  - Privacy Preserving Data Release (PPDR) vs AI/ML的矛盾
  - Differential privacy: 通过增加Noise的办法
  - RAPPOR (Google开发的)
AI平台中的BlockChain

- JD内部使用AI的部门很多，数据集重用情况复杂，追责困难
  - BlockChain记录了谁用了什么数据，做了什么model
- Trained AI Model on Blockchain
  - Model信息放到chain上，这样保证不被attacker恶意修改
AI Security Future

- Black market in e-Commerce in China is one scenario of AI vs. AI

- More AI security problems to solve
  - Openness and collaboration
Thank you!