On Detecting Growing-Up Behaviors of Malicious Accounts in Privacy-Centric Mobile Social Networks

Zijie Yang, Binghui Wang, Haoran Li, Dong Yuan, Zhuotao Liu, Neil Zhenqiang Gong,
Chang Liu, Qi Li, Xiao Liang, and Shaofeng Hu

Privacy-centric Mobile Social Networks (PC-MSNs)

- A new trend of Online Social Networks (OSNs)

- Strict usage policy
  - Restricted account access
    - search account ID / phone number
    - scan QR code
  - Restricted content access
    - posts and comments can be viewed only by friends

Growing-up Accounts

- Growing-up behaviors
  - disguised as a benign user for a period of time
  - make connections (e.g., making friends and commenting on posts) with other benign users

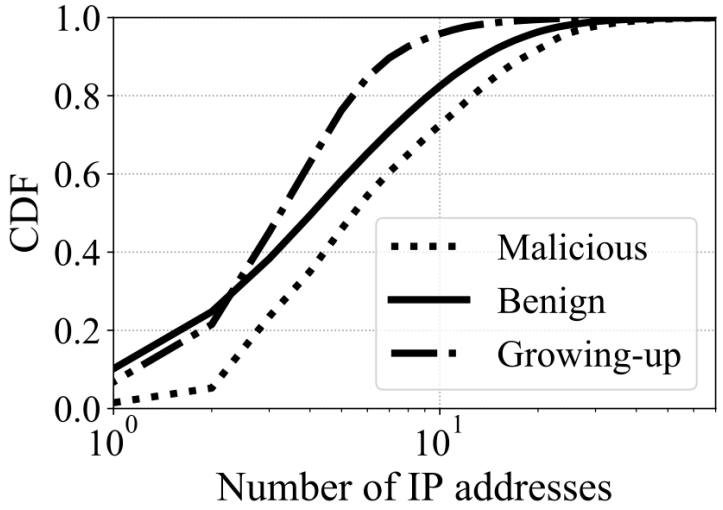
- Growing-up accounts form huge threats to PC-MSNs
  - click farm, spam, phishing
  - over 90% of malicious accounts can be classified as growing-up accounts
  - it is important to detect them before they engage in effective malicious campaigns

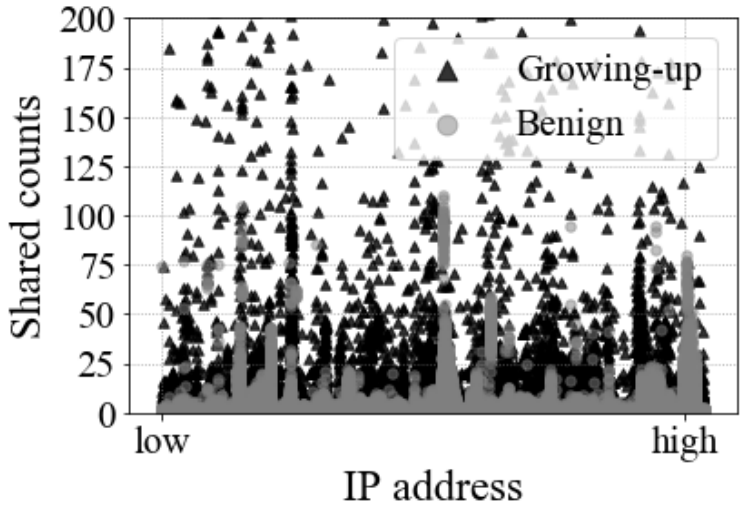
Real WeChat Dataset in Our Study

- WeChat: The largest mobile social network app in China
  - over 1 billion monthly active users

- WeChat Dataset
  - collected real-world data
  - first-week action records after registration
  - around 440k accounts
  - label obtained from WeChat's security team

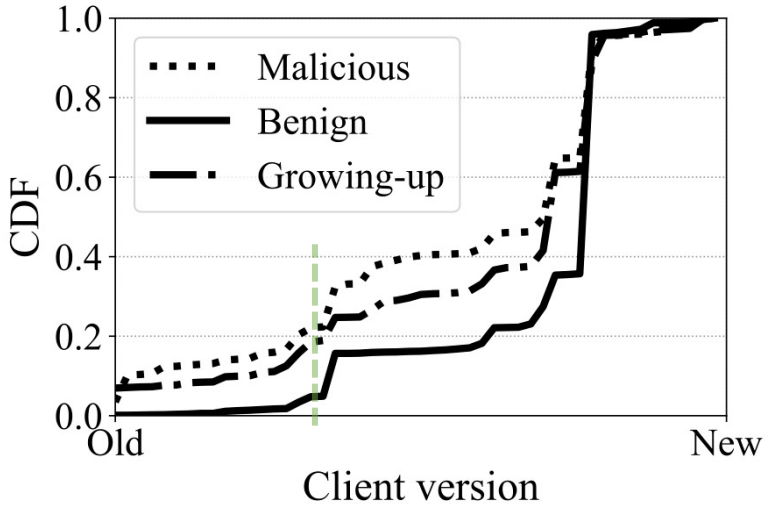
Measurement: IP Address

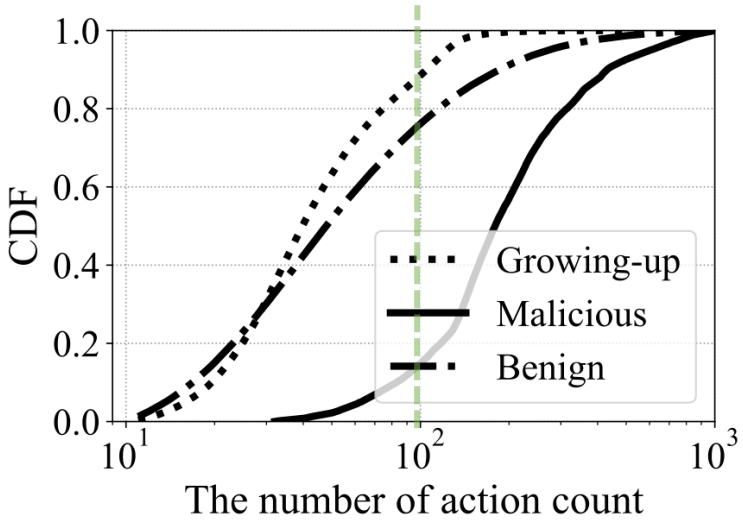




- CDF of the number of used IP addresses
- The number of accounts sharing each IP address

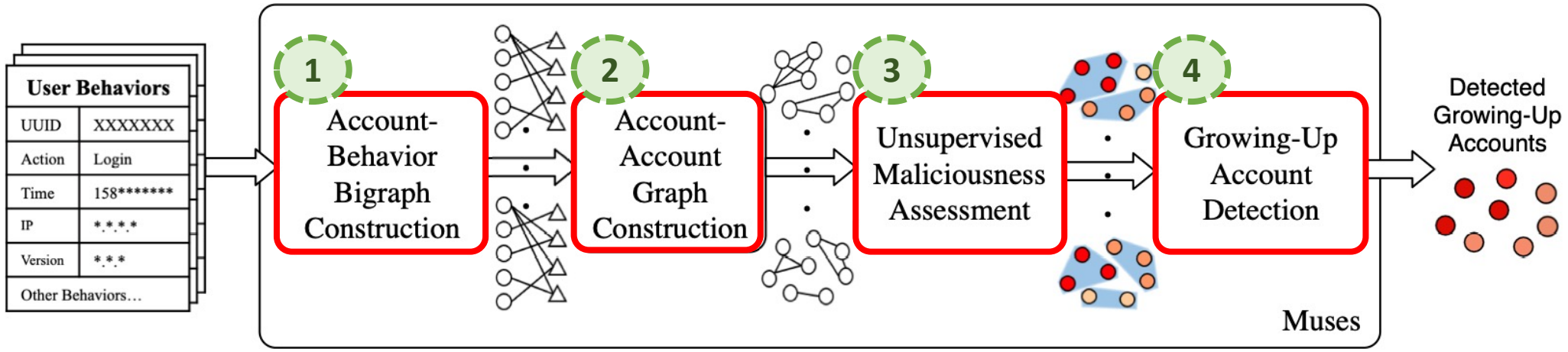
Measurement: Client Version & Action Count





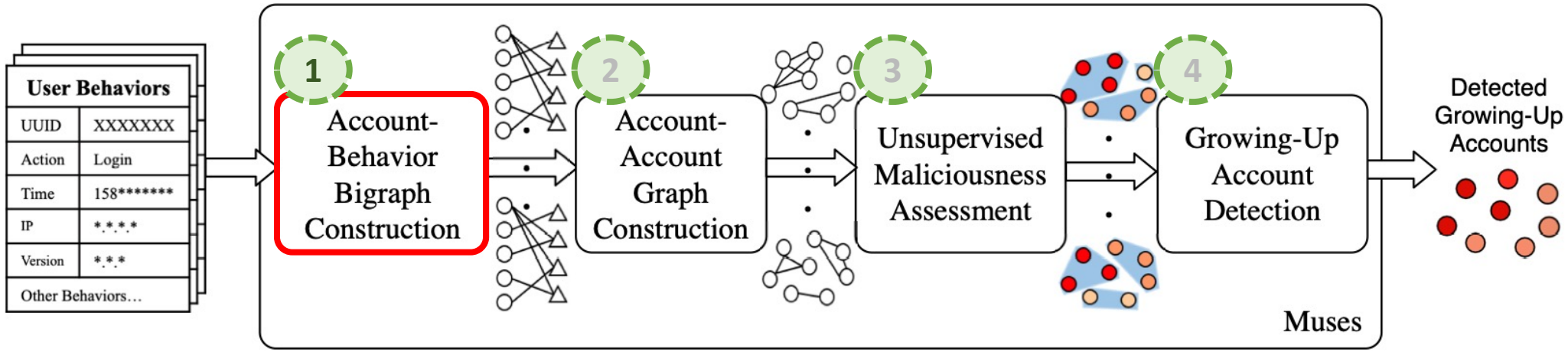
- The number of accounts that use each client version
- The CDF curves of the number of actions

Overview of Muses

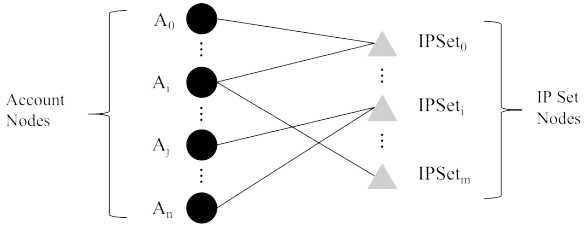


- Muses: Detect growing-up accounts in an unsupervised fashion using behavior data

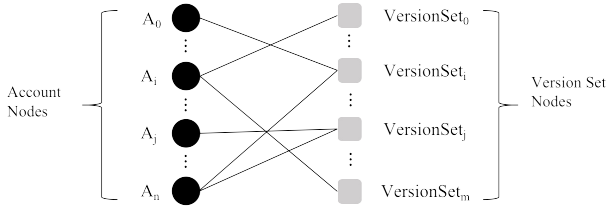
Overview of Muses



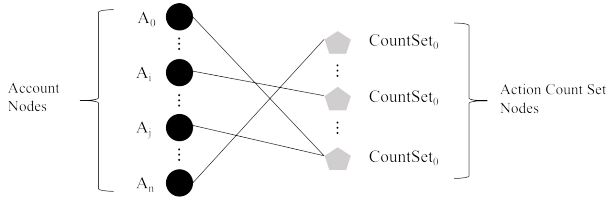
Account-IP Bigraph



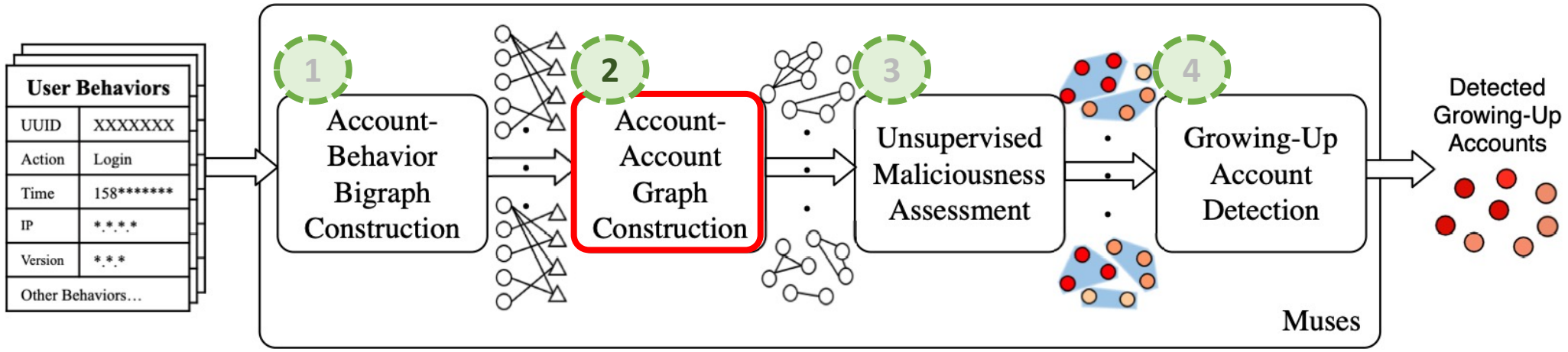
Account-Client Bigraph

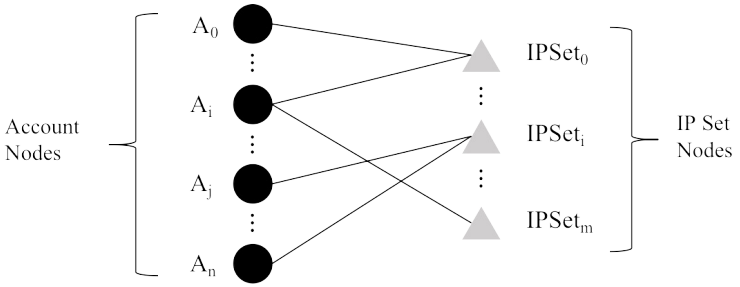


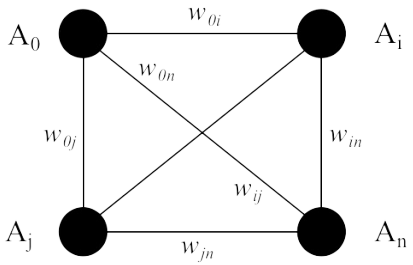
Account-Action Bigraph



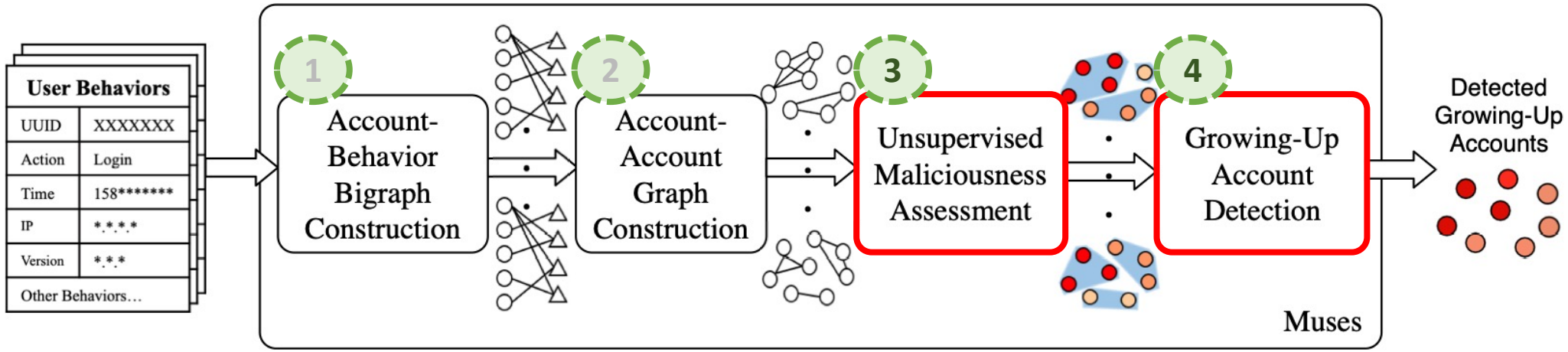
Overview of Muses







Overview of Muses



Community (Account) Maliciousness

$$s(c_i) = \frac{c_{max} - \sigma(c_i)}{c_{max} - c_{min}}$$

$$c_{max} = \max_{c_j \in C} \sigma(c_j)$$

$$c_{min} = \min_{c_j \in C} \sigma(c_j)$$

Final malicious score

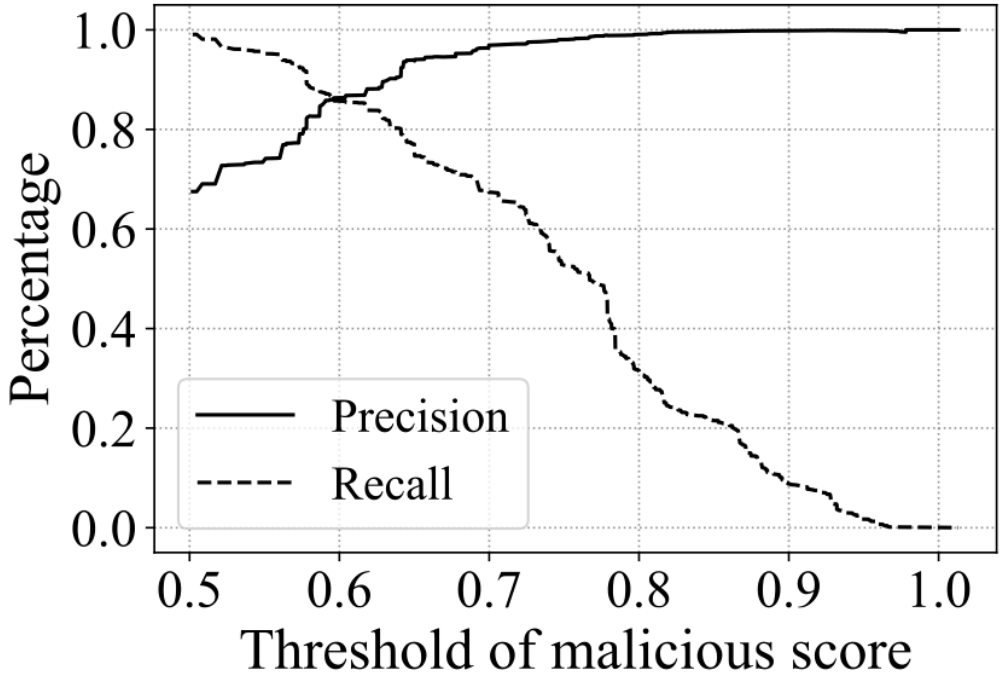
$$s(u) = \sqrt{\frac{s_{IP}(u)^2 + s_{VER}(u)^2 + s_{ACT}(u)^2}{3}}$$

Experimental Setup

- Evaluation with real datasets
  - first-week action records after registration from WeChat
- Evaluation metrics
  - Precision, Recall, aucPR
- Compared methods
  - Clickstream
  - SynchroTrap
  - EvilCohort

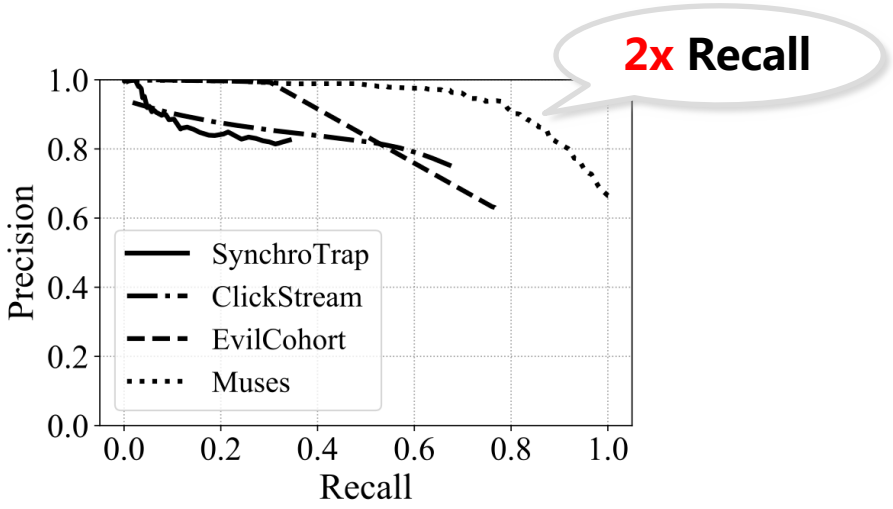
Overall Detection Performance

- Impact of threshold
  - Precision 90%
  - Recall 82%
  - AUC 0.95



Muses vs. Baseline Methods

- Muses outperforms the three baseline methods



Methods	AUC	Recall under different precision		
		0.80	0.90	0.99
SynchroTrap	0.29	0.342	0.067	0.026
ClickStream	0.56	0.571	0.086	-
EvilCohort	0.88	0.398	0.398	0.398
Muses	0.95	0.919	0.818	0.320

Conclusion

- Present the first systematic study of the growing-up behaviors of malicious accounts based on a real-world dataset.
- Propose a novel unsupervised method to effectively detect growing-up accounts.
- Experimental results show that Muses detects more than 82% of growing-up accounts with a precision higher than 90%, achieving 2x recall rate and even better precision compared with existing methods.

Thanks!

Q&A