Careful About What App Promotion Ads Recommend! Detecting and Explaining Malware Promotion via App Promotion Graph

Shang Ma, Chaoran Chen, Shao Yang, Shifu Hou,
Toby Jia-Jun Li, Xusheng Xiao, Tao Xie, Yanfang Ye

UNIVERSITY OF NOTRE DAME

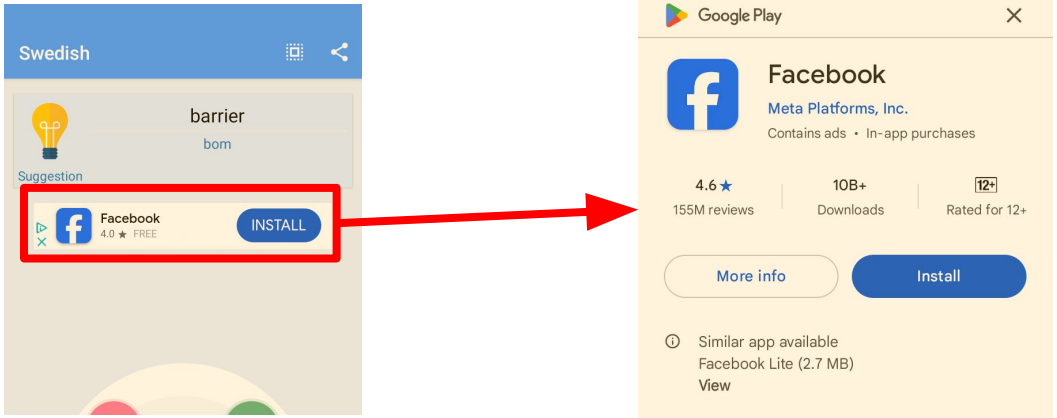
Arizona State University

App Promotion Ads

Over 57% of all apps in Google Play contain advertisements (ads).

App promotion ads are used to promote apps.

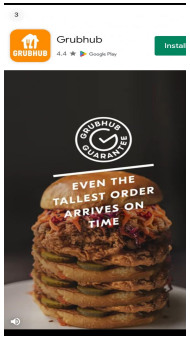
– ⅓ of users discover new apps through app promotion ads



App Promotion Ecosystem

Ad Provider

Ad

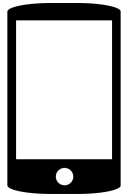


Ad Library

Google AdMob

App Developer

App

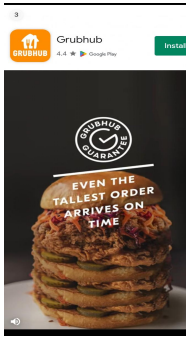


App Promotion Ecosystem

Ad content is vetted by the ad library

Ad Provider

Ad



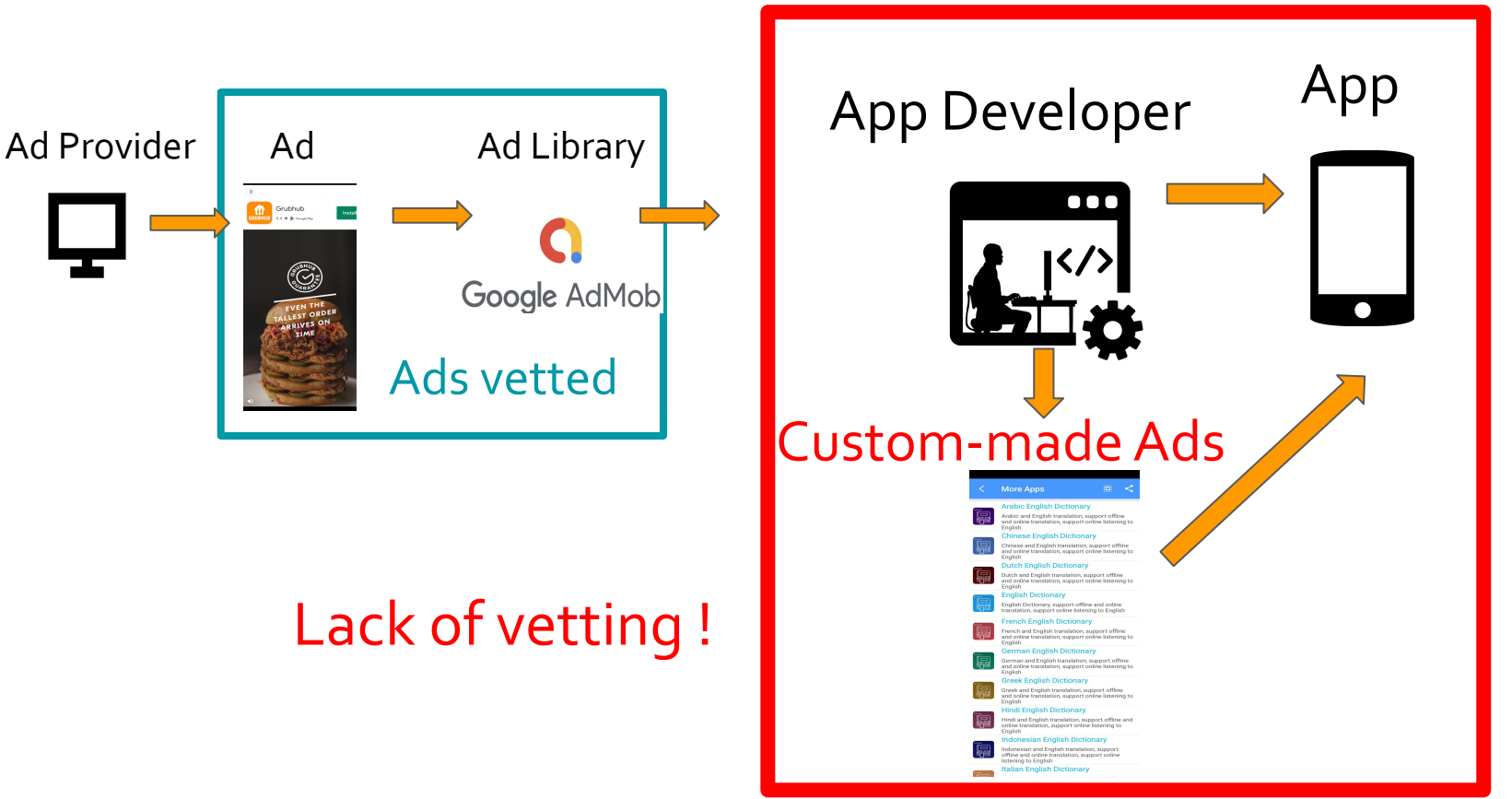
Ad Library

Google AdMob

App Developer

App

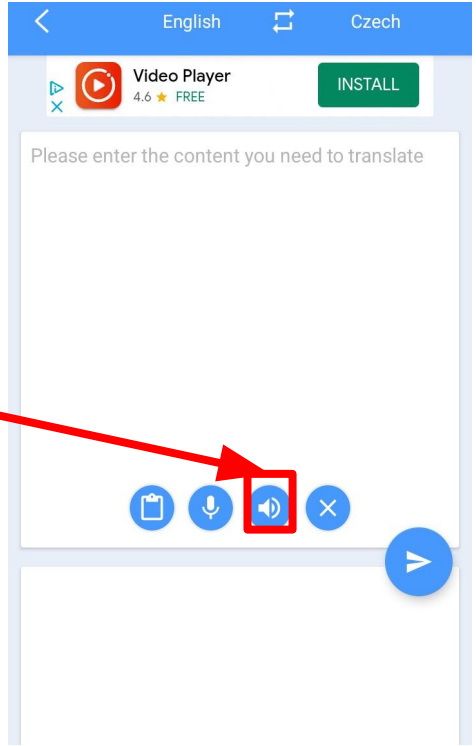
App Promotion Ecosystem



Lack of vetting !

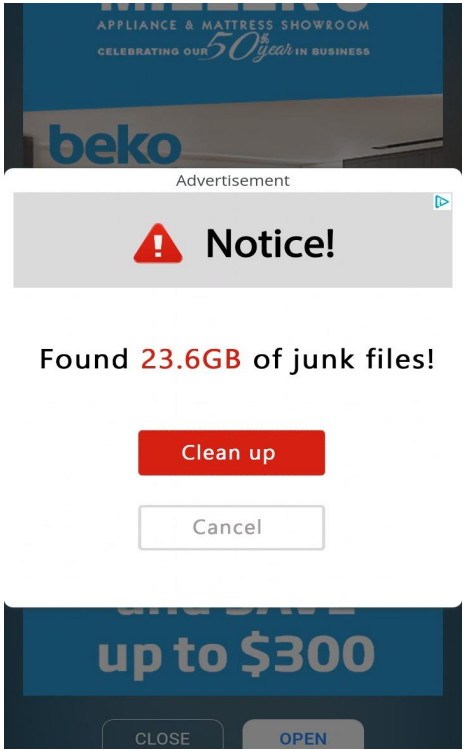
Malware Promotion Example

A malicious ad hides in the “Sound” Icon



Malware Promotion Example

When clicked, a full-screen ad pops up

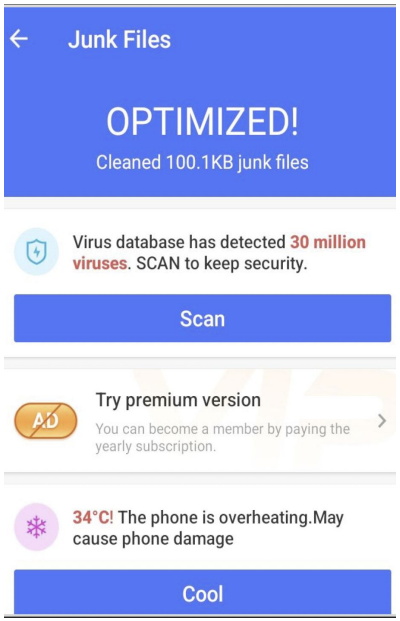


Malware Promotion Example

Redirect to Google Play to install the app

Malware Promotion Example

Scamware



Ursh anabi

Aggressive ads

★☆☆☆☆ October 12, 2023

Too many ads and new version looks and feels terrible. It's much harder to use

Fake functionality

★☆☆☆☆ September 30, 2023

Downloaded an update and it is horrible. Only takes you to game Is not cleaning phone just clogging it up. Very dissatisfied

Fake subscription

★☆☆☆☆ September 24, 2023

Are you serious!!! paid for premium a few months ago and now I lost it since the update!! I tried to restore and it says there is no account found 😡😡

Research Problem

Can we automatically detect the malware promoted by app promotion ads ?

Preliminary Study

Dataset: sampled from AndroZoo[1] (200 apps), Rico[2] (405 apps)

Findings:

- **Custom-made ads** are
  - prevalent: **23%** app promotion ads are custom-made ads
  - risky: **51%** of them promote malware
- Ad content are requested from the **server at runtime**

Challenge: Applying **static analysis** on **ad libraries** is not sufficient to detect malware promotion

[1] Androzoo: Collecting millions of android apps for the research community." Proceedings of the 13th international conference on mining software repositories. 2016.

[2] "Rico: A mobile app dataset for building data-driven design applications." Proceedings of the 30th annual ACM symposium on user interface software and technology. 2017.

Insight: *App Promotion Graph*

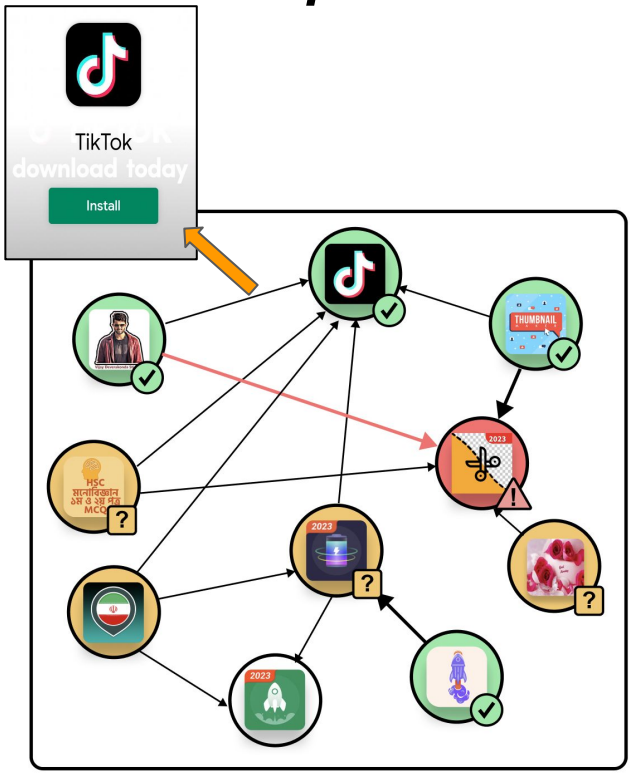
The ecosystem can be modeled as a graph: *app promotion graph*

- Edges: app promotion ads
- Nodes: apps

We use this graph to capture

- **app promotion relations** among apps
- **app attributes** derived from *app markets*, *security vendor*, and *binary code*

Malware detection→ **Node classification**



Our Approach

Part 1: *UI exploration* to **collect app promotion ads** to construct an app promotion graph

Part 2:*Graph learning* to **detect malware promotion** based on the constructed app promotion graph

Motivation for combining *UI exploration* and *graph learning*:

1.UI exploration alone can collect app promotion ads but **cannot determine the maliciousness** of the promoted apps.

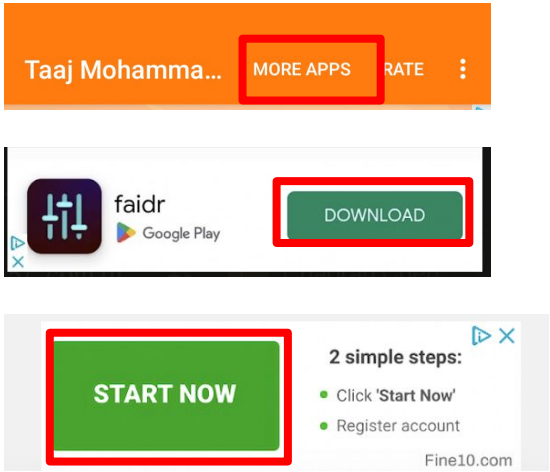
2. Effectiveness of graph learning depends on the **features of the app promotion graph** built by UI exploration

Part 1:
App Promotion Graph Construction

Ad-oriented UI exploration

- **Depth first search** to navigate to the UIs containing ads
- **Text patterns** (empirically crafted ad-related string) to detect ad content
- **Iteratively restarting** the app to capture the periodically changing ad content.

Examples of ad patterns

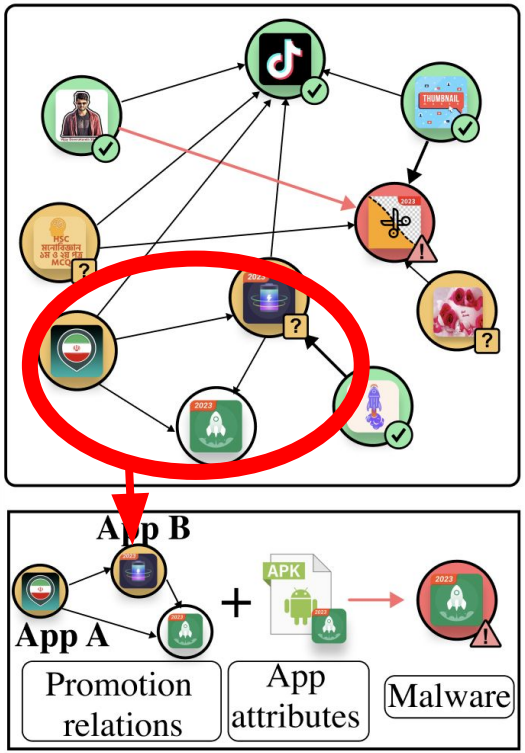


Part 2: Malware Detection

Features for malware detection

- Existing work:
  app attributes
  (single app features)

- Our approach
  app attributes
  + promotion relations
  (graph features)



Effectiveness of App Promotion Graph Construction

Starting from 36,000 seed apps,
we construct an app promotion graph consisting of:

- **18, 627** app promotion ads (edges)
- **6, 008** apps (2420 source nodes, 3859 target nodes)

Approaches	Ad Units	Ad Types		
		Inherent	Pop-up	Custom-Made
Droidbot [31]	76	27	38	9
Monkey [32]	71	26	34	11
DARPA [33]	8	8	0	0
MadDroid [8]	75	32	39	6
ADGPE (bfs)	131	52	58	15
ADGPE	**165**	**76**	**71**	**17**

Effectiveness of Malware Promotion Detection

- Overall performance: **97.74% accuracy, 95.31% F1 score**
- Performance gain brought by ***promotion relations***: **5.17%** (90.14% to 95.31% F1 score)

	Approaches	Accuracy	Precision	Recall	F1 score
Baselines	Symantec	96.99	81.66	69.01	74.80
	Lionic	96.72	74.64	74.64	74.64
	McAfee	95.99	69.56	67.60	68.57
	Avira	94.26	53.57	84.50	65.57
	K7GW	93.63	50.41	85.91	63.54
	DroidEvolver [29]	$75.48_{\pm 7.12}$	$72.92_{\pm 7.96}$	$70.93_{\pm 11.39}$	$71.21_{\pm 6.96}$
	MaMaDroid [28]	$79.38_{\pm 7.33}$	$75.48_{\pm 6.32}$	$78.41_{\pm 9.54}$	$76.58_{\pm 6.14}$
	ANDRUSPEX [30]	$95.15_{\pm 1.24}$	$95.32_{\pm 1.14}$	$88.79_{\pm 3.19}$	$92.48_{\pm 3.19}$
Ablation Study	− promotion	$96.29_{\pm 1.07}$	$95.27_{\pm 3.68}$	$86.01_{\pm 7.23}$	$90.14_{\pm 7.23}$
	→DGI [74]	$97.47_{\pm 0.61}$	$99.10_{\pm 2.19}$	$91.43_{\pm 6.82}$	$94.96_{\pm 6.82}$
	→GRACE [75]	$97.45_{\pm 0.66}$	$\mathbf{99.82}_{\pm 0.55}$	$91.43_{\pm 6.64}$	$95.30_{\pm 6.64}$
	→MVGRL [76]	$97.38_{\pm 0.65}$	$98.57_{\pm 2.48}$	$90.90_{\pm 7.27}$	$94.40_{\pm 7.27}$
	ADGPE	$\mathbf{97.74}_{\pm 0.62}$	$99.44_{\pm 1.67}$	$\mathbf{91.78}_{\pm 7.02}$	$\mathbf{95.31}_{\pm 7.02}$

+5.17%

Interesting Findings

1. **Prevalence**:
   Popular *ad networks* are exploited to spread a variety of malware (520 malware promotion ads): adware, trojan, and fleeceware…

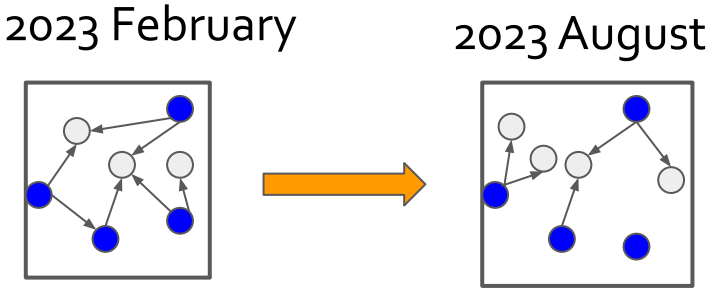
2. **Risk:**
   2.64% of apps promoted by app promotion ads are malware.
   Extremely risky given the large user base
   "⅓ of users discover new apps through app promotion ads"

3. **Promotion Tactics**
   - Promotion Chain: ~~Benign apps→Malware~~
     Benign apps→PUAs (Potentially Unwanted Apps) →Malware
   - Flagship: A popular app to attract downloads and promote malware
   - App Waves: No-code app makers to create massive adware

Temporal Analysis

Method

- Re-construct an app promotion graph from the same dataset 6 months later



Findings

1. **Zero-day apps**
   - Definition: new nodes/apps
   - 190 found, **18** are **malware** with million downloads

2. **Late-detection malware**
   - Definition: benign in February, malware in August.
   - **28** found. **All detected by our approach** early in February.

	# VirusTotal flags in February	# VirusTotal flags in August
Zero-day apps	**N/A**	**n**
Late-detection malware	**0**	**≥1**

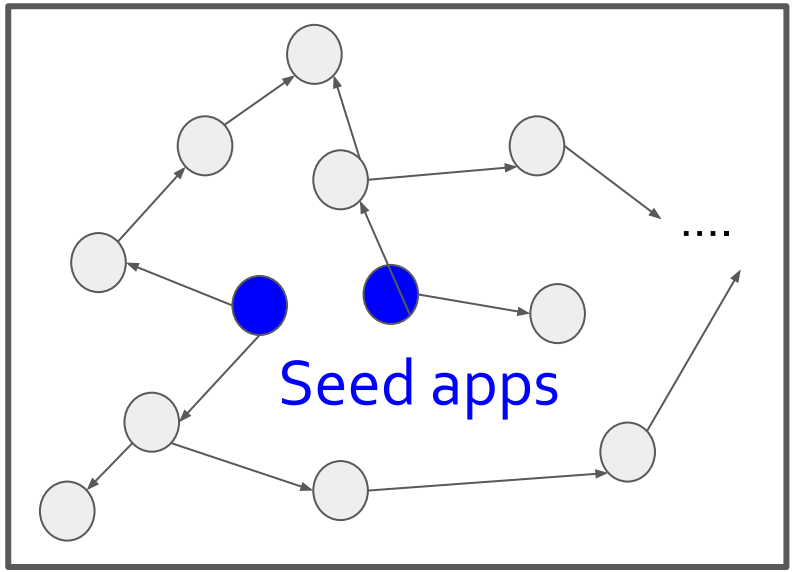
In-the-wild Case Study

Method

- Construct an app promotion graph from **2 seed apps (pirated video apps)**

Findings

- 37-nodes app promotion graph
- **21 malware**: 5 gambling, 11 pornographic, 1 trojan, 4 adware ***All promoted by custom-made ads***
- Potential to study underground economy



Thank you! Questions?

Key Takeaways

1. 2.64% app promotion ads promote malware
2. Graph learning on ad promotion relations helps detect malware promotion

Full paper

UNIVERSITY OF NOTRE DAME

Yes-Lab

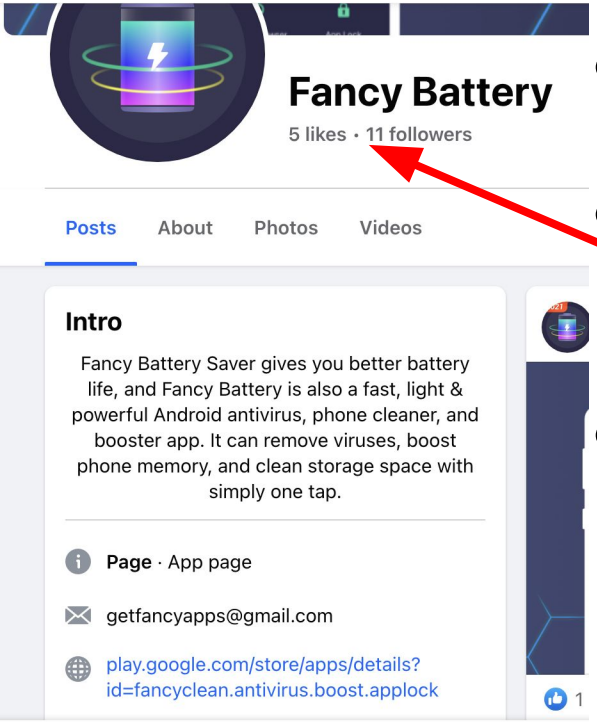
Arizona State University

RISE Lab@ASU

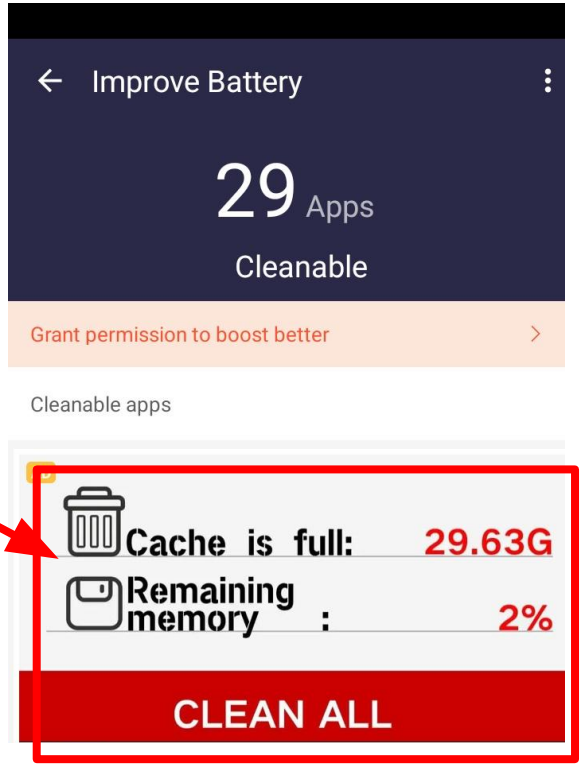
Code and Dataset:
https://github.com/AppPromotionAdsResearch

Shang Ma
sma5@nd.edu

Promotion Tactic: Flagship



- Build a high-quality **flagship app**
- Leverage social media (e.g., Facebook) to boost the flagship app
- Use custom-made ads to promote other adware, scamware



Promotion Tactic: App Wave

- No-code app maker to create massive free apps
- Similar appearance and content: wallpaper of popular anime
- Though most get removed, some remains with high downloads