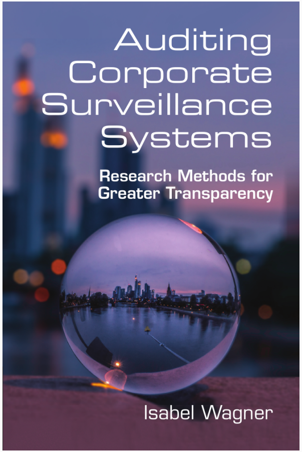
RESULTS FROM TRANSPARENCY RESEARCH

MOBILE SERVICES

Isabel Wagner

De Montfort University



OVERVIEW

- Results for three key parts of the mobile ecosystem:
  - App stores
  - Third-party libraries
  - Individual apps, including ride sharing
- Focusing on characteristics and extent of corporate surveillance

App stores

CHARACTERISTICS OF APP STORES

- Top 1% of free apps cause 81% of downloads from Play store[1]
- 3.5 times more free apps than paid ones
- Ad library use: two-thirds of popular apps on Google Play, 55% on iOS[2]
- Comparing Google Play with 16 Chinese app markets[3]
  - Google Play not accessible in China
  - Apps with update in last 6 months: 5% on Chinese stores, 23% on Google Play
  - Ad libraries: in 53% of apps (Chinese stores), in 70% of apps (Google Play)
  - Permission use: apps on Chinese stores use more permissions
  - Malicious apps: 12% on Chinese stores, 2% on Google Play
  - Removal of malicious apps: 0.01%–34% (Chinese stores), 84% on Google Play

[1] N. Viennot, E. Garcia, and J. Nieh, "A Measurement Study of Google Play," in *The 2014 ACM International Conference on Measurement and Modeling of Computer Systems*, ser. SIGMETRICS '14, Austin, Texas, USA: ACM, 2014, pp. 221–233. DOI: 10.1145/2591971.2592003.

[2] M. Egele, C. Kruegel, E. Kirda, *et al.*, "PiOS: Detecting Privacy Leaks in iOS Applications," in *NDSS Symposium*, San Diego, California, USA: Internet Society, Feb. 2011, p. 15.

[3] H. Wang, Z. Liu, J. Liang, *et al.*, "Beyond Google Play: A Large-Scale Comparative Study of Chinese Android App Markets," in *Proceedings of the Internet Measurement Conference 2018*, Boston, MA, USA: ACM, Oct. 2018, pp. 293–307. DOI: 10.1145/3278532.3278558.

APP UPDATES

- Frequency of updates[4]
  - Less than half of apps updated within the last year, less than 25% updated within last 6 months
  - Updates correlate with decrease price for apps, increased permission use
- Effect of reviews that mention security or privacy[5]
  - 60% of apps with security/privacy relevant review made corresponding updates
  - Privacy reviews are significant predictor of apps becoming more privacy-friendly (e.g., developer was not aware of privacy practices of third-party library)

[4] B. Carbunar and R. Potharaju, "A Longitudinal Study of the Google App Market," in *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*, ser. ASONAM '15, Paris, France: ACM, 2015, pp. 242–249. DOI: `10.1145/2808797.2808823`.

[5] D. C. Nguyen, E. Derr, M. Backes, *et al.*, "Short Text, Large Effect: Measuring the Impact of User Reviews on Android App Security & Privacy," in *IEEE Symposium on Security & Privacy*, San Francisco, CA, USA: IEEE, May 2019, p. 15.

Third-party libraries

THIRD-PARTY LIBARIES

- Large variety of third-party libraries: 60,000 different libraries in 2017[6]
- Ad libraries and analytics libraries most common
- Libraries may leak user data: device identifiers, location, sensor data[7]
- Updates to third-party libraries[8]
  - Developers have to monitor libraries for new versions and update their apps manually
  - Outdates libraries (with security vulnerabilities) are common
  - Of 90 libraries included in 1.2 million apps: 85% of inclusions were outdated, 48% could be upgraded without changing app code

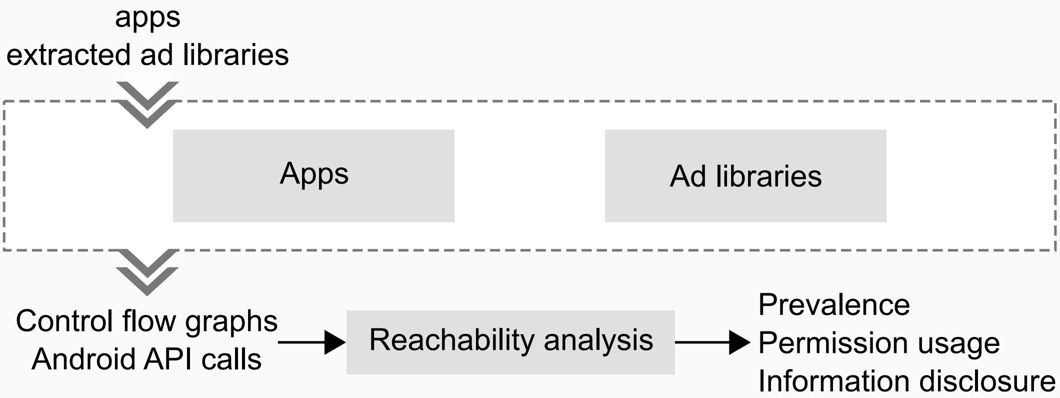
[6] M. Li, W. Wang, P. Wang, *et al.*, "LibD: Scalable and Precise Third-Party Library Detection in Android Markets," in *2017 IEEE/ACM 39th International Conference on Software Engineering (ICSE)*, May 2017, pp. 335–346. DOI: `10.1109/ICSE.2017.38`.

[7] Y. He, X. Yang, B. Hu, *et al.*, "Dynamic privacy leakage analysis of Android third-party libraries," *Journal of Information Security and Applications*, vol. 46, pp. 259–270, Jun. 2019. DOI: `10.1016/j.jisa.2019.03.014`.

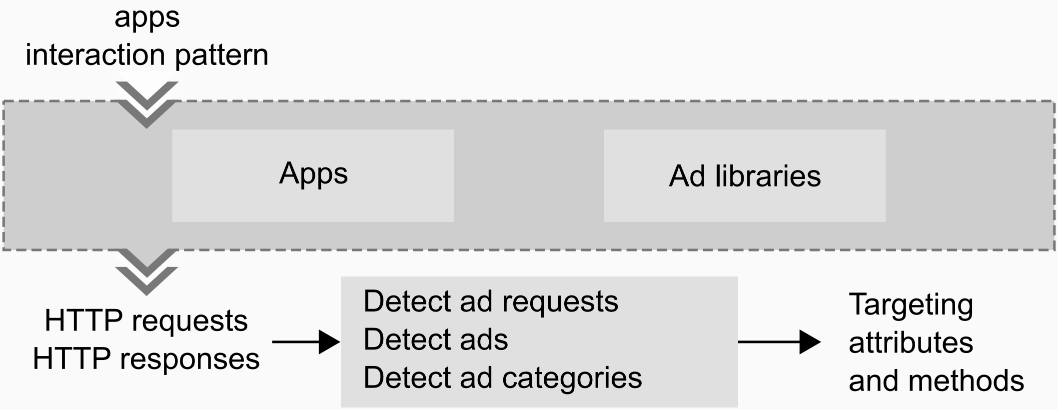
[8] E. Derr, S. Bugiel, S. Fahl, *et al.*, "Keep Me Updated: An Empirical Study of Third-Party Library Updatability on Android," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '17, Dallas, Texas, USA: ACM, 2017, pp. 2187–2200. DOI: `10.1145/3133956.3134059`.

Apps

DESIGN FOR STUDYING MOBILE SERVICES WITH STATIC ANALYSIS



DESIGN FOR STUDYING MOBILE SERVICES WITH DYNAMIC ANALYSIS



PRIVACY LEAKS FROM MOBILE APPS

Platform	Number of apps	Device identifiers	User identifiers	Locations	Contacts	Media	Year
iOS	1400	50%	—	3%	0.4%	—	2011
iOS	225,000	48%	–	13%	6%	—	2013
iOS	100	47%	14%	26%	2%	—	2016
Windows	100	55%	3%	8%	—	—	2016
Android	100	52%	14%	14%	1%	—	2016
Android	100	42%	—	1%	—	—	2017
Android	9,100	—	—	—	—	0.2%	2018
Android	524	62%	8.5%	53%	—	—	2018
Android	46	33%	4%	2%	—	—	2020

PRIVACY LEAKS FROM MOBILE APPS

- Leaks from paid apps[9]
  - If free version uses third-party library or dangerous permission: 45% of paid apps retain libraries/permissions
  - Only 3.7% of privacy policies differentiate free/paid versions
- Leaks from desktop versions of apps[10]
  - 2016: PII leaked from 90% of apps and 76% of web versions
- Cloud use of apps[11]
  - 90% of apps use cloud services
  - Each app connects to 3.2 cloud services
  - 50% of apps send 95% of traffic to cloud services
  - Most popular cloud services: Google (54% of apps), Amazon (65% of apps)

[9] C. Han, I. Reyes, Á. Feal, *et al.*, "The Price is (Not) Right: Comparing Privacy in Free and Paid Apps," *Proceedings on Privacy Enhancing Technologies*, vol. 2020, no. 3, pp. 222–242, 2020. DOI: 10.2478/popets-2020-0050.

[10] C. Leung, J. Ren, D. Choffnes, *et al.*, "Should You Use the App for That?: Comparing the Privacy Implications of App- and Web-based Online Services," in *Proceedings of the 2016 Internet Measurement Conference*, ser. IMC '16, Santa Monica, California, USA: ACM, 2016, pp. 365–372. DOI: 10.1145/2987443.2987456.

[11] M. Henze, J. Pennekamp, D. Hellmanns, *et al.*, "CloudAnalyzer: Uncovering the Cloud Usage of Mobile Apps," in *Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, ser. MobiQuitous 2017, Melbourne, VIC, Australia:

RIDE-SHARING APPS (1)

- Popular ride-sharing apps (Uber, Lyft) not willing to share data with local transport planning authorities[12]
- Rider demographics: average Uber rider is white (81%), mid 20-s, above-average income[13]
- Driver demographics: white (60%), male (76%), only 17% are also riders
- Temporal availability: supply peaks during rush hour[14]
    - Compare to taxis: constant supply throughout day
    - Different employment model: gig economy (pay-per-ride) vs. regular employment (pay-per-hour)

[12] D. Cooper, J. Castiglione, A. Mislove, *et al.*, "Profiling Transport Network Company Activity using Big Data," *Transportation Research Record*, vol. 2672, no. 42, pp. 192–202, Dec. 2018. DOI: *10.1177/0361198118798459*.

[13] F. Kooti, M. Grbovic, L. M. Aiello, *et al.*, "Analyzing Uber's Ride-sharing Economy," in *Proceedings of the 26th International Conference on World Wide Web Companion*, ser. WWW '17 Companion, Perth, Australia: International World Wide Web Conferences Steering Committee, 2017, pp. 574–582. DOI: *10.1145/3041021.3054194*.

[14] S. Jiang, L. Chen, A. Mislove, *et al.*, "On Ridesharing Competition and Accessibility: Evidence from Uber, Lyft, and Taxi," in *Proceedings of the 2018 World Wide Web Conference*, ser. WWW '18, Lyon, France: International World Wide Web Conferences Steering Committee, 2018, pp. 863–872. DOI: *10.1145/3178876.3186134*.

RIDE-SHARING APPS (2)

- Spatial availability[15]
  - Decreasing availability with distance to city center
  - Reduced wait times in areas with: high population density, high median income
- Negative externalities[16]
  - Increase in congestion: ride-sharing apps draw commuters away from public transport
- Driver privacy[17]
  - Adversarial tracking of drivers' routines, working patterns, home location, employment status
  - By running copies of ride-sharing app in grid of geolocations

[15] S. Jiang, L. Chen, A. Mislove, *et al.*, "On Ridesharing Competition and Accessibility: Evidence from Uber, Lyft, and Taxi," in *Proceedings of the 2018 World Wide Web Conference*, ser. WWW '18, Lyon, France: International World Wide Web Conferences Steering Committee, 2018, pp. 863–872. DOI: `10.1145/3178876.3186134`, J. Thebault-Spieker, L. Terveen, and B. Hecht, "Toward a Geographic Understanding of the Sharing Economy: Systemic Biases in UberX and TaskRabbit," *ACM Trans. Comput.-Hum. Interact.*, vol. 24, no. 3, 21:1–21:40, Apr. 2017. DOI: `10.1145/3058499`.

[16] S. Agarwal, D. Mani, and R. Telang, "The Impact of Ride-hailing Services on Congestion: Evidence from Indian Cities," Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 3410623, Jun. 2019.

[17] Q. Zhao, C. Zuo, G. Pellegrino, *et al.*, "Geo-locating Drivers: A Study of Sensitive Data Leakage in Ride-Hailing Services," in *Proceedings 2019 Network and Distributed System Security Symposium*, San Diego, CA: Internet Society, 2019. DOI: `10.14722/ndss.2019.23052`.

SUMMARY

RESULTS FOR MOBILE SERVICES

- Study designs for static and dynamic analysis
- Overview of results for app stores, third-party libraries, individual apps

ABOUT THIS SLIDE DECK

- These slides are designed to accompany a lecture based on the textbook “Auditing Corporate Surveillance Systems: Research Methods for Greater Transparency” by Isabel Wagner, published in 2022 by Cambridge University Press.
-
-
- Please leave this slide intact, but indicate modifications below.
    - Version 2022-04
    - Improved version for release on book website (Isabel Wagner)
- Updated versions of the original slide deck are available online: corporatesurveillance.org