Ubiquitous and Mobile Computing
CS 528: *MobileMiner*
Mining Your Frequent Behavior Patterns on Your Phone

Muxi Qi

Electrical and Computer Engineering Dept.

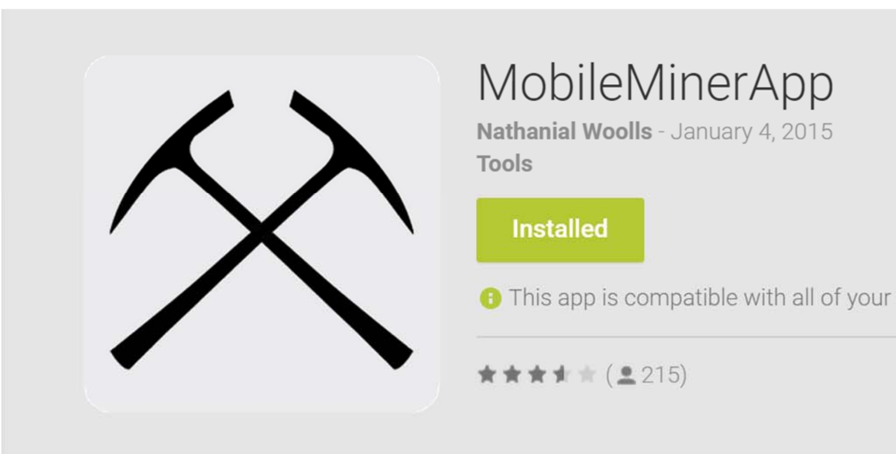
Worcester Polytechnic Institute (WPI)

OUTLINE

- Introduction
- System Design
- Evaluation
  - Performance
  - Pattern Utility
- Example Use Cases: App and Call Prediction
- Related Work
- Conclusion

INTRODUCTION

- The Goal:
  - Long Term: Novel middleware and algorithms to **efficiently** mine user **behavior patterns** entirely on the **phone** by utilizing **idle processor cycles**.
  - In This Paper: MobileMiner on the phone for frequent **co-occurrence patterns.**



INTRODUCTION

- Idea Inspiration:
  - We can log raw contextual data.
  - Previous:
    - Location & physical sensor data

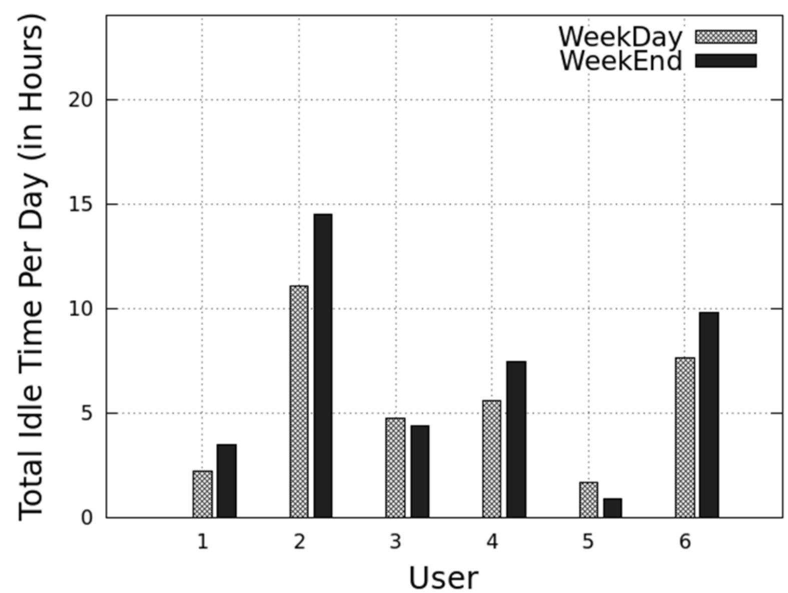
    -> higher level user context
  - Now:
    - Higher level **behavior patterns**

    from a long term
  - Why **Behavior Patterns**?
    - Personalize & improve user experience.

INTRODUCTION

- How to Achieve
  - Co-occurrence Patterns & Their Utility
    - Useful
    - In association rules: easily used & *if-this-then-that*
      - {Morning; Breakfast; At Home} -> {Read News}
  - Smartphone Computing Potential
    - Powerful quad-core processors

  & unused for a majority of time
    - Privacy guarantees (not cloud)
    - Cloud connectivity constrain

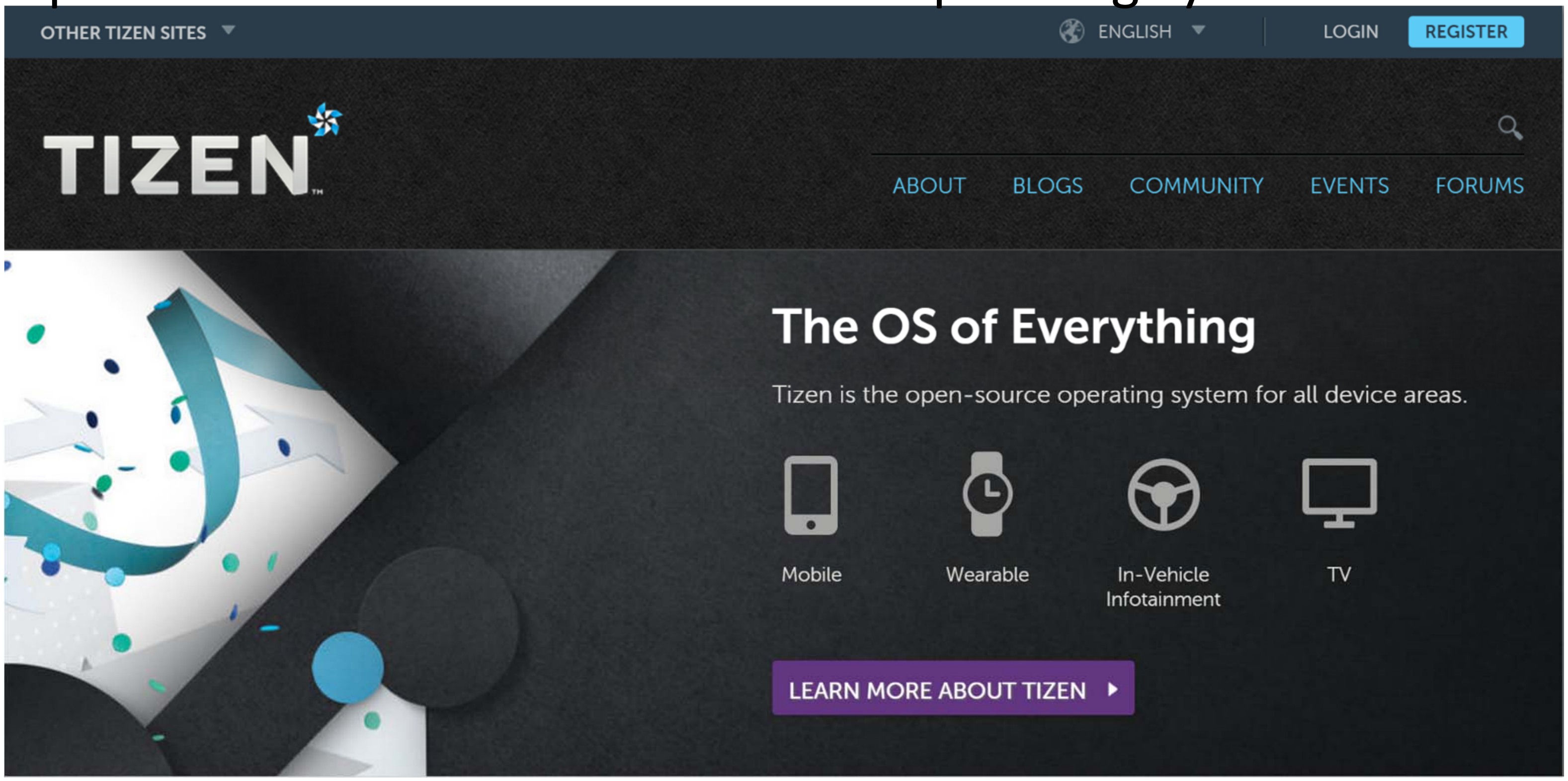


INTRODUCTION

- Main Contributions:
  - System Design
  - System Performance
  - Patterns' Utility Analysis
  - UI Improvement Implementation

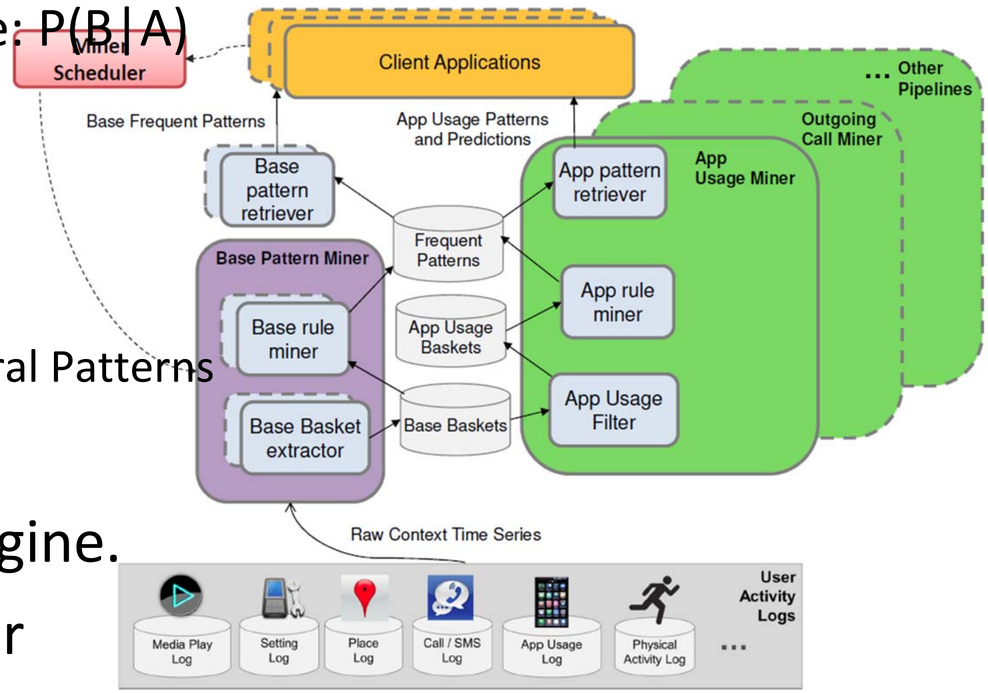
SYSTEM DESIGN

- Platform: Tizen Mobile
  - Tizen:
    - Open and flexible Linux Foundation operating system.



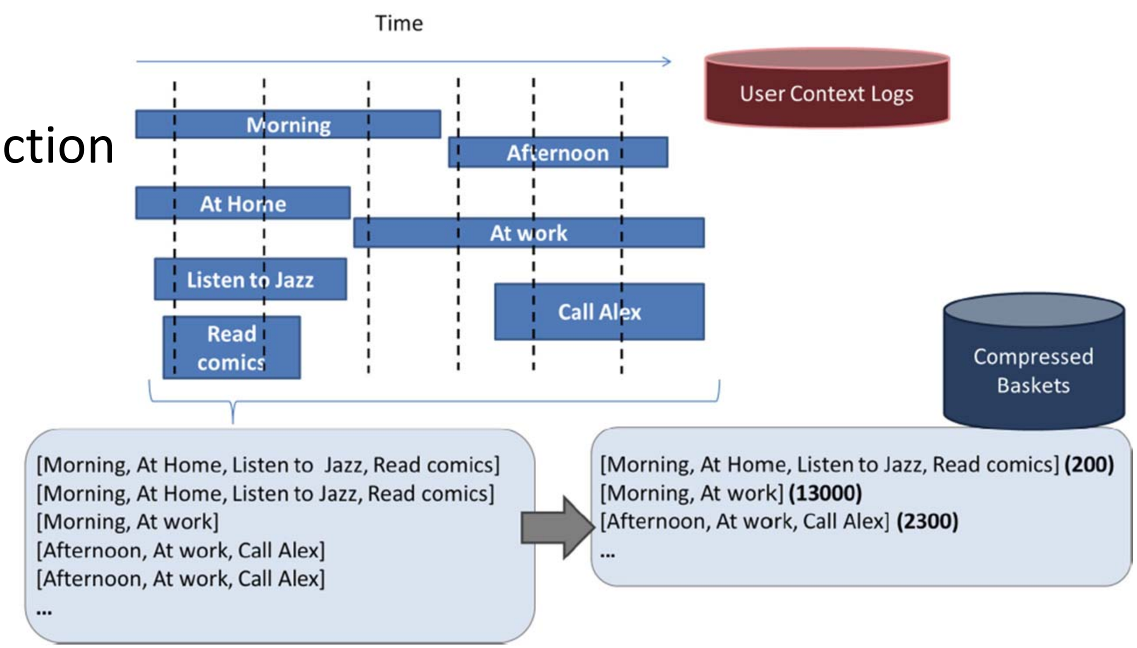
SYSTEM DESIGN

- System Architecture
  - Frequent Pattern Formulation:
    - Association Rule. {A: Antecedents} -> {B: Consequence}
  - Threshold:
    - Support: P(AB); Confidence: P(B|A)
  - Baskets: Time Stamped
  - Mining Algorithm:
    - WeMiT, not Apriori
      - Weighted Mining of Temporal Patterns
  - Filters
  - Predictions: Prediction Engine.
  - Schedule: Miner Scheduler



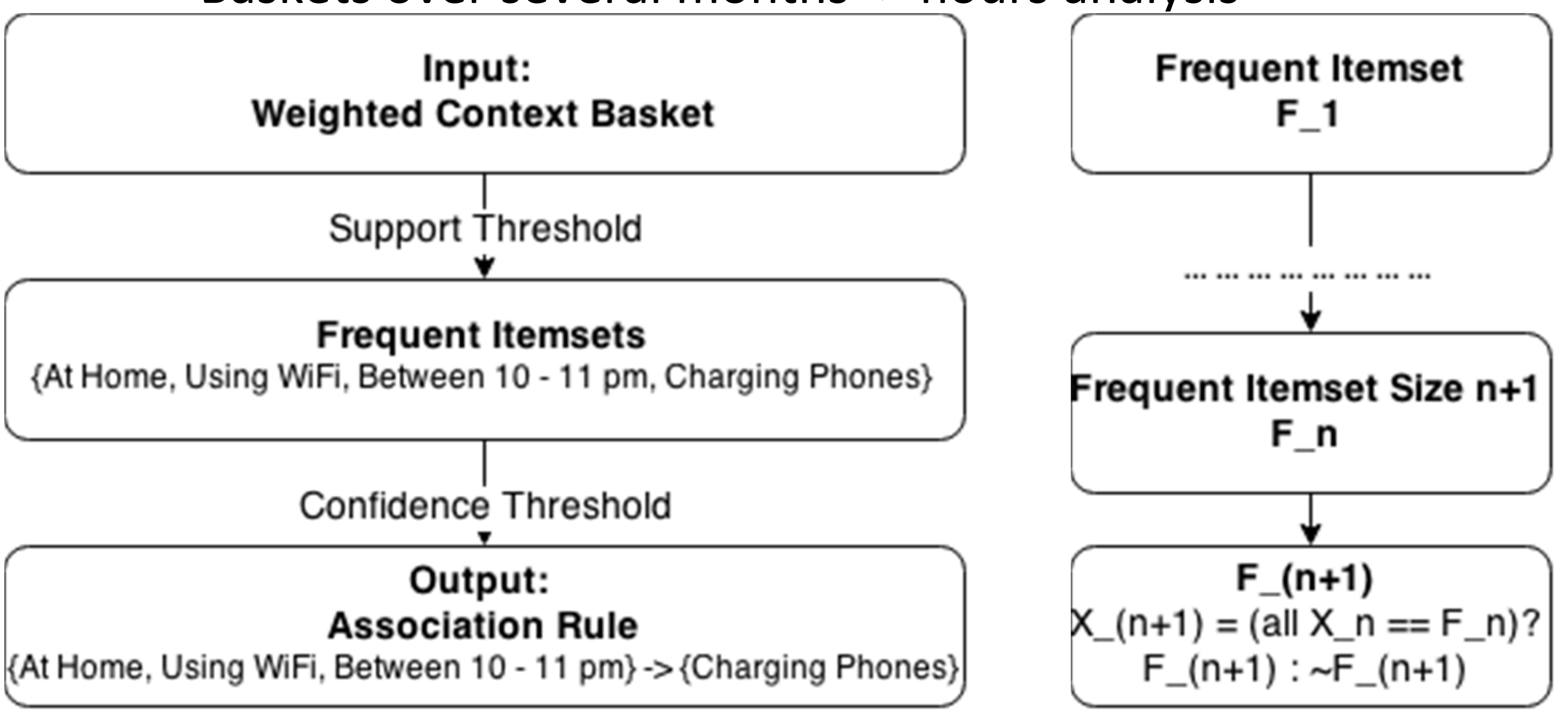
SYSTEM DESIGN

- Basket Extraction:
  - Discretization (Categorical Data) => Baskets Extraction
- Basket Filtering
  - Using Boolean expression, utility functions
  - Benefits:
    - More accurate prediction
    - Faster
    - free of noise



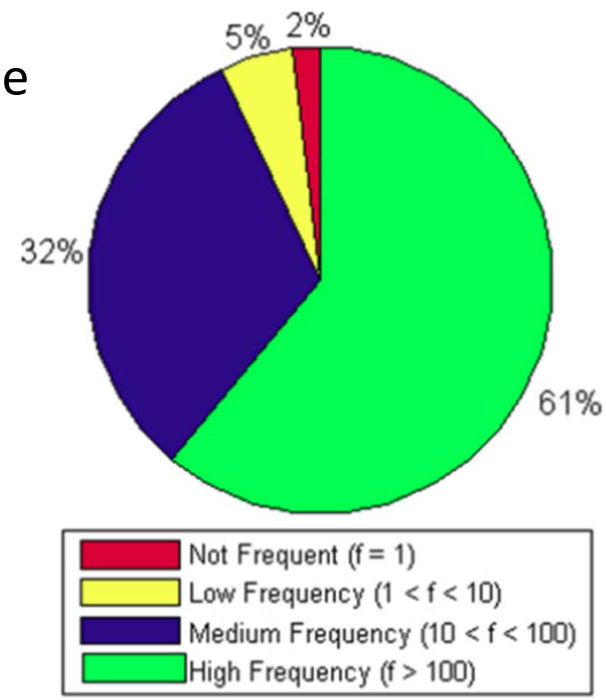
SYSTEM DESIGN

- Rule Mining:
  - Apriori Algorithm: “Bottom Up”
    - All subsets of a frequent itemset are also frequent itemsets.
    - Baskets over several months -> hours analysis



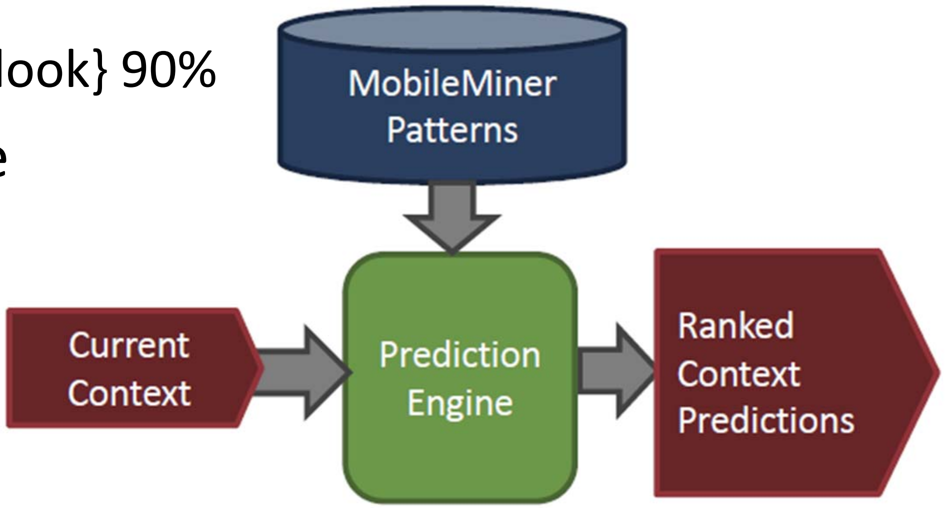
SYSTEM DESIGN

- Rule Mining:
  - WeMiT: “Repeated Nature”
    - $B = \{b_1^{w_1}, b_2^{w_2}, , b_n^{w_n}\}, \sum_{i=1}^{n} contain(b_i^{w_i}, X).w_i$,
    - 92.5% reduction by compression
    - 15 times reduction in average running time



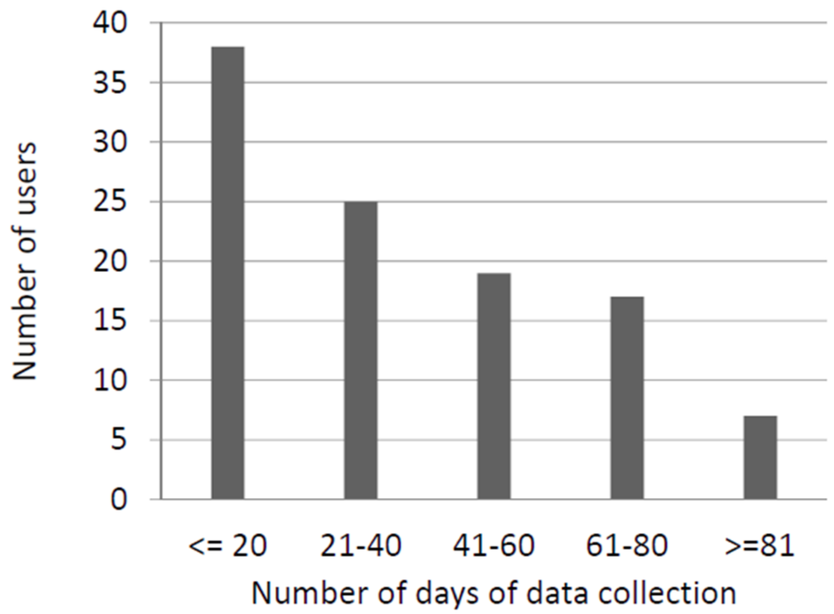
SYSTEM DESIGN

- Context Prediction
  - Novelty: 1 second return prediction
  - Input: {Morning; At Work} & {Using Gmail; Using Outlook}
  - Rule:
    - {Morning} -> {Gmail} 90%
    - {At Work} -> {Gmail} 80%
    - {Morning; At Work} -> {Outlook} 90%
  - Ranking Order: Confidence
  - Same target?
  - Same confidence?



EVALUATION - Context Data

- Participants:
  - 106 (healthy mix of gender and occupation), 1 - 3 months
- Collector: EasyTrack using Funf sensing library
- Results:
  - 440 Unique Context Events
  - Active participants?



EVALUATION - Context Data

- Focused Context Events
  - <call type=“” duration=“” number=“”>
  - <SMS type=“” number=“”>
  - <placeIdentifier place=“home”>
  - <location clusterLabel=“”>
  - <charging status=“”>
  - <battery level=“”>
  - <foreground app=“”>
  - <connectivity type=”WiFi”>
  - <cellLocation id=“”>
  - <movement status=“1”>

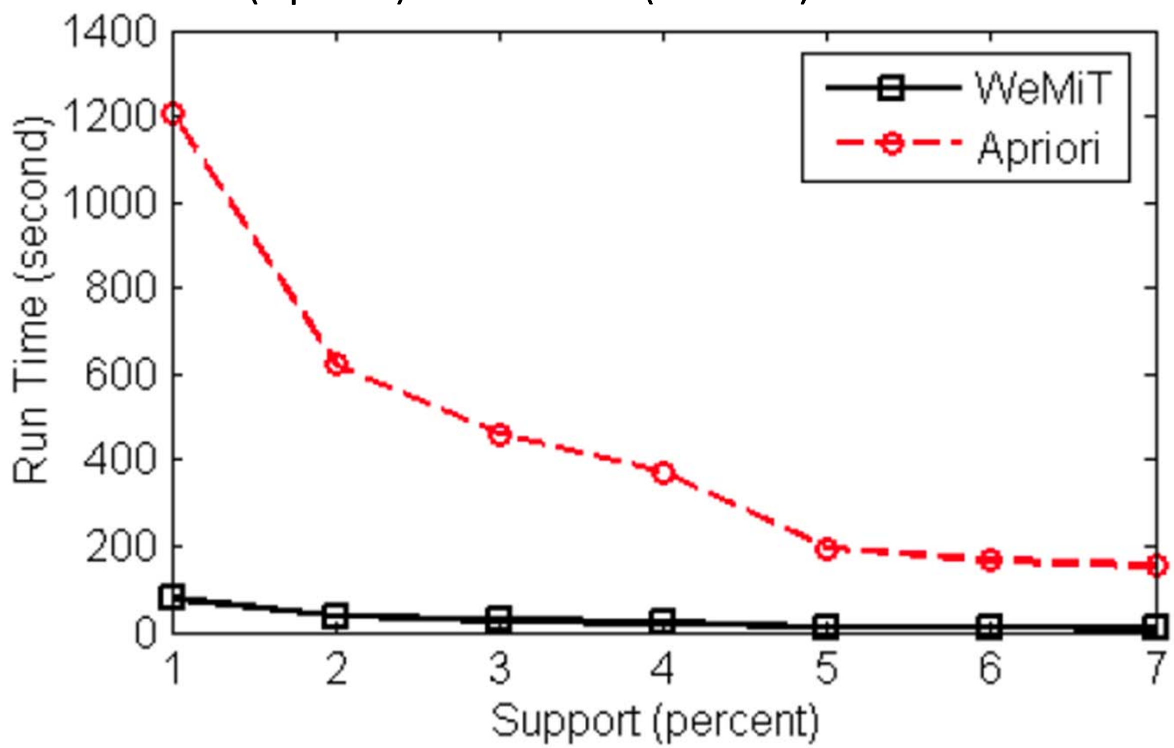
EVALUATION - Performance

- MobileMiner, Tizen phone (==Samsung Galaxy S3)
  - Feasibility
    - Data: 28 representative users, 2 - 3 months.
    - Threshold: Base 1% Support, App 20 Support
    - Compression Reduction: 92.5% and 55%
    - Energy(7.98Wh): 0.45% and 0.01% weekly, 3.09% and 0.05% daily

Performance Metric	Base Basket Extraction	Base Rule Mining	App Usage Filtering	App Usage Rule Mining
Execution time	1.7 seconds	16.5 minutes	1.4 seconds	21.2 seconds
Memory	9.9 MB	44.2 MB	11.6 MB	1.0 MB
CPU Utilization	22.9 %	24.3 %	20.8 %	21.9 %
Number of baskets or rules	114275 baskets 8559 compressed	46675 rules	752 baskets 327 compressed	1062 rules
Energy per day as % of full battery	<0.01 %	0.45 %	<0.01 %	0.01 %

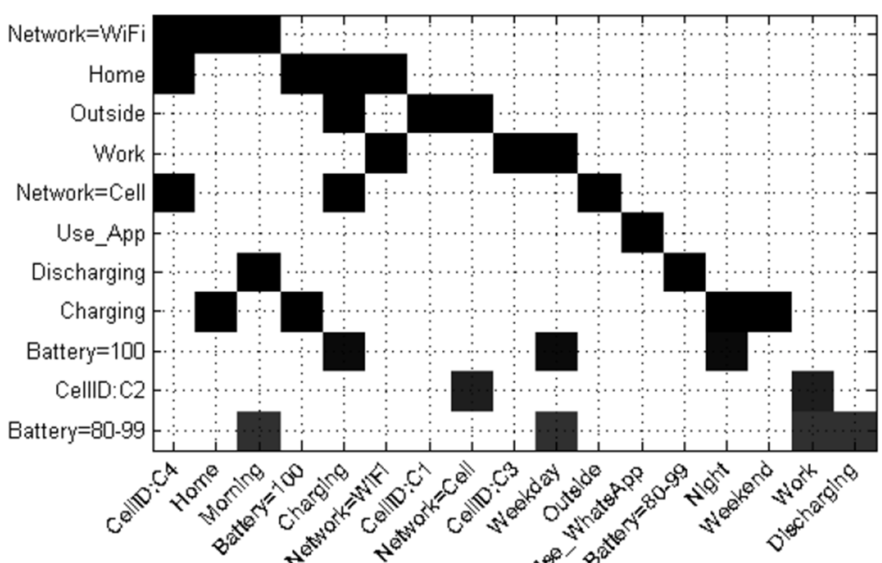
EVALUATION - Performance

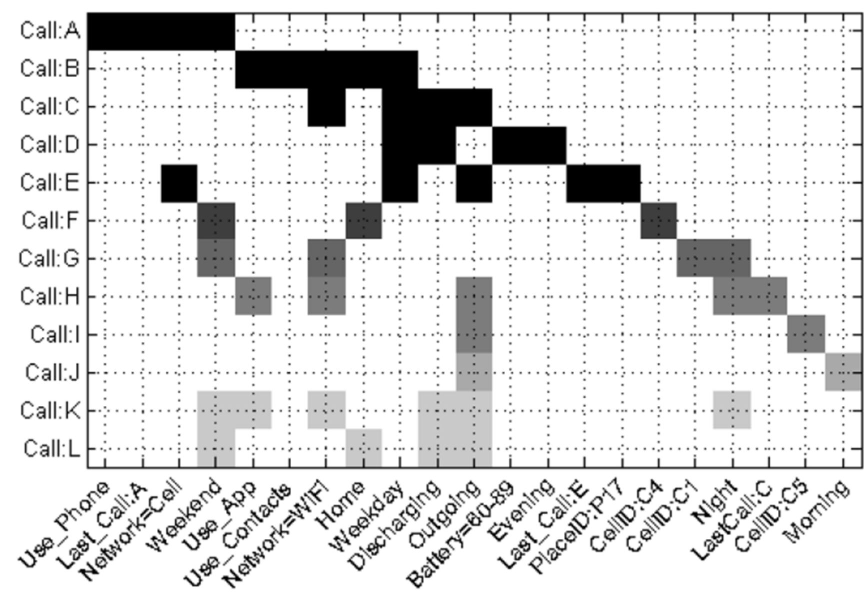
- MobileMiner, Tizen phone (==Samsung Galaxy S3)
  - Comparison:
    - Data: 13 users
    - Short Duration Activities: 20 min (Apriori) vs 78.5 sec (WeMiT)



EVALUATION - Pattern Utility

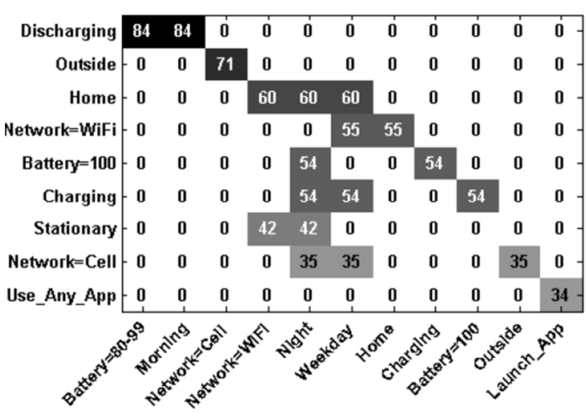
- Sample Patterns
  - Data: sample user #38
  - Threshold: 1% Support
  - Greyscale: Confidence
  - Utility: Provide shortcut for next contact

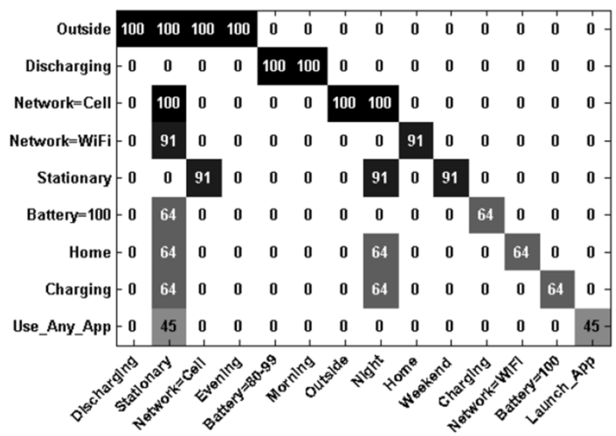


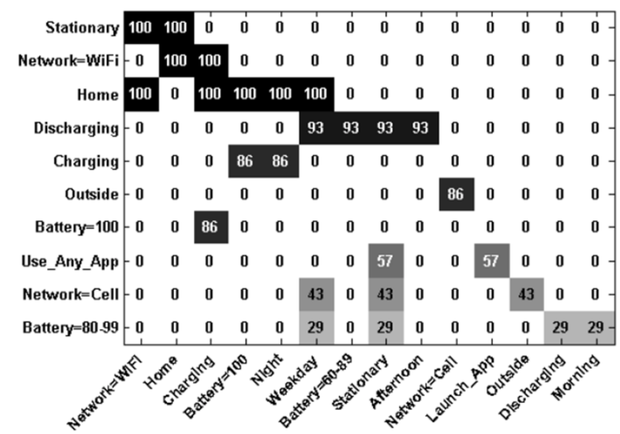


EVALUATION - Pattern Utility

- Common patterns
  - Threshold: 80% confidence 1% support
  - Greyscale: Percentage of users the pattern occurs in
  - Utility:
    - Initial set of patterns while MobileMiner is learning slowly
  - Future:
    - schedule group activity; individual recommendation service

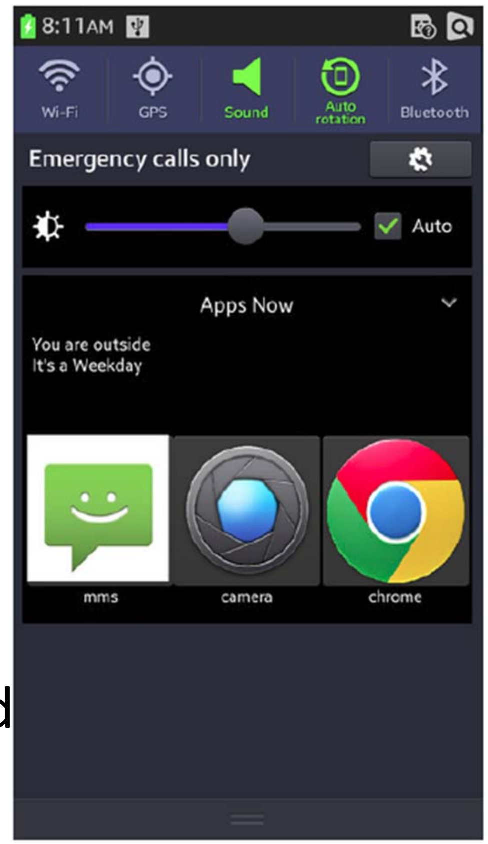


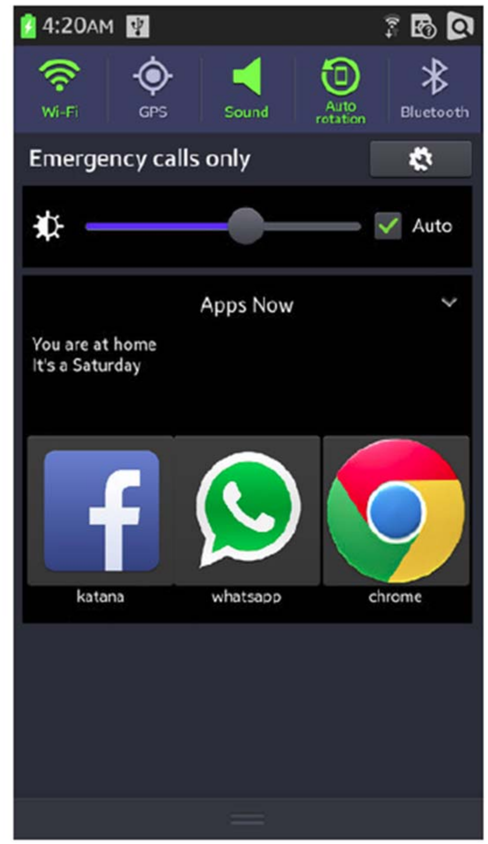




EXAMPLE USE CASE

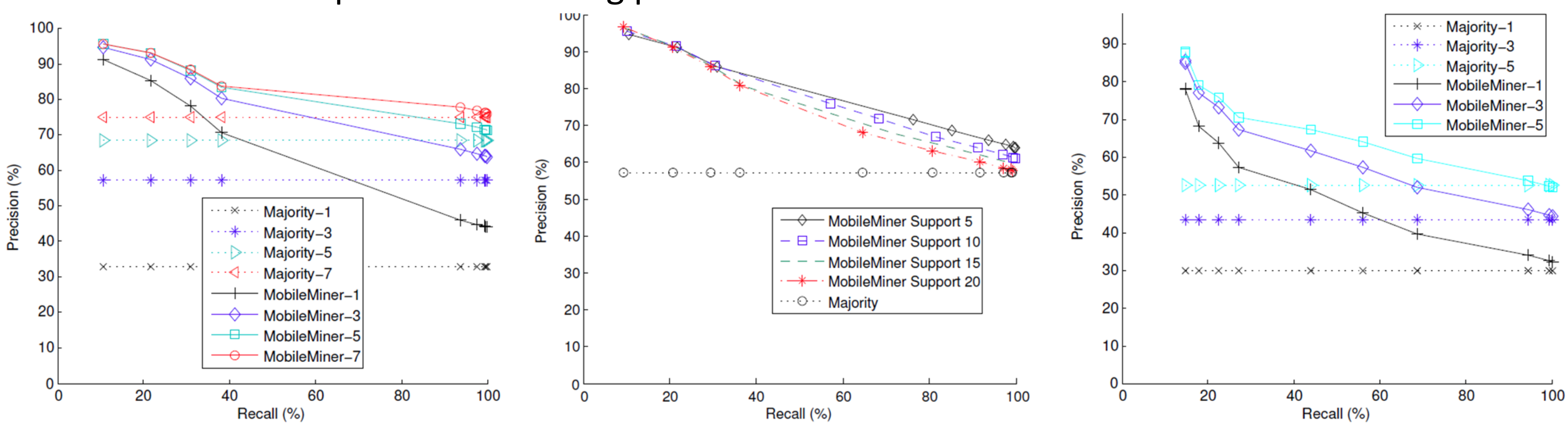
- App and Call Prediction
  - Benefit: Lessen the Burden
  - Feature:
    - Show pattern
  - Evaluation Metrics
    - Recall: of total usage
    - Precision: of popups
  - Setting Parameter:
    - Shortcut #
    - Confidence Threshold





EXAMPLE USE CASE

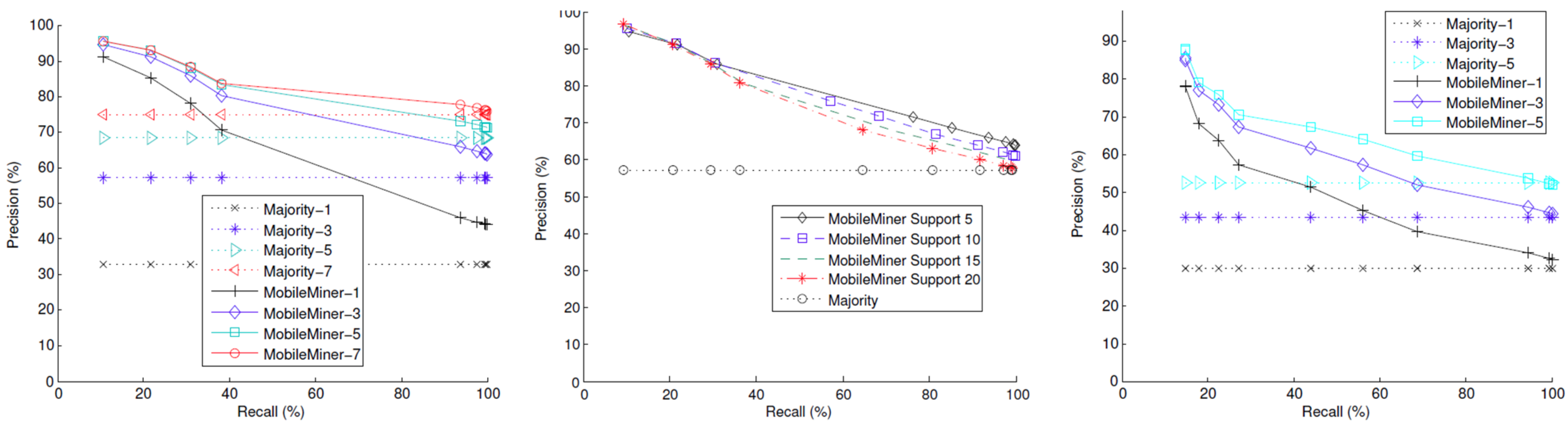
- Recall-Precision Tradeoff
  - Data: 106 for App, 25 for Call
  - MM vs Majority: 89%-184% improvement
  - App vs Call: why?
    - limited data
    - less predictable calling pattern



(a) App prediction. (b) Effect of support on app prediction with 3 recommendations. (c) Call prediction.

EXAMPLE USE CASE

- Recall-Precision Tradeoff
  - Support Threshold
    - Precision: 4-5% improvement
      - Rules of only 5 times may potentially be useful in improving precision
    - Time: 12.4, 37.1, 174.8, 2218.2 sec



(a) App prediction.

(b) Effect of support on app prediction with 3 recommendations.

(c) Call prediction.

EXAMPLE USE CASE

- User Survey
  - Participants: 42 from 106, online
  - Limitation:
    - using not app but explanation with screenshots
  - Conclusion:
    - Positive response
    - Recall - Precision Tradeoff differs

  -> a configurable app

Precision	No. of recommendations	Recall	Responses
90%	3	35%	30.95%
80%	3	51%	16.67%
80%	5	68%	23.81%
80%	7	80%	11.90%
75%	3	66%	4.76%
75%	5	87%	11.9%
75%	7	100%	19.05%

EXAMPLE USE CASE

- User Survey (Detailed Results)
  - Usage Frequency
    - Regularly 57%; Sometimes 42%
  - Shortcut
    - Lock screen 40%; Quick panel 26%; Main tool bar 33%
  - 100% Recall or less for Precision?
    - Recall 9%; Precision 54%; Either 35%
  - Icon Number
    - 4-6 71%; 1-3 26%
  - Tradeoff

Precision	No. of recommendations	Recall	Responses
90%	3	35%	30.95%
80%	3	51%	16.67%
80%	5	68%	23.81%
80%	7	80%	11.90%
75%	3	66%	4.76%
75%	5	87%	11.9%
75%	7	100%	19.05%

RELATED WORK

- Association Rule and Frequent Itemset Mining
  - In the cloud or desktop
  - Our: On-device mining
- Context-ware Computation on Mobile Devices
  - Inferring activity, location, proximity
  - ACE (Acquisitional Context Engine) System:
    - Server-based, without optimized algorithm
    - Privacy, data cost, and latency
  - Our: concerning long term context, on-device

RELATED WORK

- Prediction Approaches
  - Compare to Others, Ours has:
    - more generalizable approach
    - more configurability
    - more tolerance to missing context events
    - more readable patterns
- A preliminary Version (Poster)

References

1. Aggarwal, C. C., and Yu, P. S. A new approach to online generation of association rules. *IEEE Transactions on Knowledge and Data Engineering 13*, 4 (2001), 527–540.
2. Agrawal, R., and Srikant, R. Fast algorithms for mining association rules in large databases. In *Proceedings of the 20th International Conference on Very Large Data Bases (VLDB '94)*, Morgan Kaufmann (1994).
3. Aharony, N., Pan, W., Ip, C., Khayal, I., and Pentland, A. Social fmri: Investigating and shaping social mechanisms in the real world. *Pervasive and Mobile Computing 7*, 6 (2011).
4. Allen, J. F. Maintaining knowledge about temporal intervals. *Communications of the ACM 26*, 11 (1983), 832–843.
5. Android operating system. http://www.android.com/.
6. Azizyan, M., Constandache, I., and Roy Choudhury, R. Surroundsense: Mobile phone localization via ambience fingerprinting. In *Proceedings of the 15th Annual International Conference on Mobile Computing and Networking (MobiCom '09)* (2009).

References

7. Banerjee, N., Agarwal, S., Bahl, P., Chandra, R., Wolman, A., and Corner, M. Virtual compass: Relative positioning to sense mobile social interactions. In *Proceedings of the 8th International Conference on Pervasive Computing (Pervasive '10)*, Springer-Verlag (2010).
8. Borgelt, C. Efficient implementations of apriori, eclat and fp-growth. http://www.borgelt.net, August 2013.
9. Cheung, D. W., Han, J., Ng, V. T., and Wong, C. Maintenance of discovered association rules in large databases: An incremental updating technique. In *Data Engineering, 1996. Proceedings of the Twelfth International Conference on*, IEEE (1996), 106–114.
10. Samsung galaxy s4. http://www.samsung.com/latin_en/consumer/mobile-phones/mobile-phones/smartphone/GT-I9500ZKLTPA-spec.
11. Samsung gear. http://www.samsung.com/us/mobile/wearable-tech.
12. Han, J., Kamber, M., and Pei, J. *Data Mining: Concepts and Techniques.* Morgan Kaufmann Publishers Inc., 2011.

References

13. Hao, T., Xing, G., and Zhou, G. isleep: Unobtrusive sleep quality monitoring using smartphones. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems (SenSys '13)*, ACM (2013).
14. Ifttt mobile recipes. https://ifttt.com/recipes.
15. ios 7. https://www.apple.com/ios/what-is/.
16. Kwapisz, J. R., Weiss, G. M., and Moore, S. A. Activity recognition using cell phone accelerometers. *SIGKDD Explorations Newsletter 12*, 2 (2011), 74–82.
17. Li, W., Han, J., and Pei, J. Cmar: accurate and efficient classification based on multiple class-association rules. In *Proceedings of IEEE International Conference on Data Mining (ICDM '01)*, IEEE (2001).
18. Lin, K., Kansal, A., Lymberopoulos, D., and Zhao, F. Energy-accuracy trade-off for continuous mobile device location. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services (MobiSys '10)*, ACM (2010).

References

19. Linux foundation. http://collabprojects.linuxfoundation.org/.
20. Liu, B., Jiang, Y., Sha, F., and Govindan, R. Cloud-enabled privacy-preserving collaborative learning for mobile sensing. In *Proceedings of the 10th ACM Conference on Embedded Networked Sensor Systems (SenSys '12)*, ACM (2012).
21. Liu, J., Priyantha, B., Hart, T., Ramos, H. S., Loureiro, A. A. F., and Wang, Q. Energy efficient gps sensing with cloud offloading. In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems (SenSys '12)*, ACM (2012).
22. Lu, H., Pan, W., Lane, N. D., Choudhury, T., and Campbell, A. T. Soundsense: Scalable sound sensing for people-centric applications on mobile phones. In *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services (MobiSys '09)*, ACM (2009).

References

23. Miluzzo, E., Lane, N. D., Fodor, K., Peterson, R., Lu, H., Musolesi, M., Eisenman, S. B., Zheng, X., and Campbell, A. T. Sensing meets mobile social networks: The design, implementation and evaluation of the cenceme application. In *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems(SenSys '08)*, ACM (2008).
24. Monsoon power monitor. https://www.msoon.com/LabEquipment/PowerMonitor/.
25. Nath, S. Ace: Exploiting correlation for energy-efficient and continuous context sensing. In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services (MobiSys '12)*, ACM (2012).
26. Parate, A., B¨ohmer, M., Chu, D., Ganesan, D., and Marlin, B. M. Practical prediction and prefetch for faster access to applications on mobile phones. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, ACM (2013), 275–284.

References

27. Shin, C., Hong, J.-H., and Dey, A. K. Understanding and prediction of mobile application usage for smart phones. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*, ACM (2012).
28. Srinivasan, V., Moghaddam, S., Mukherji, A., Rachuri, K., Xu, C., and Tapia, E. M. On-device mining of mobile users' co-occurrence patterns. In *Proceedings of the 15th International Workshop on Mobile Computing Systems and Applications (POSTER)* (2014).
29. Survey monkey. https://www.surveymonkey.com/.
30. Tizen platform. https://www.tizen.org.
31. Welbourne, E., Wu, P., Bao, X., and Munguia-Tapia, E. Crowdsourced mobile data collection: lessons learned from a new study methodology. In *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications*, ACM (2014), 2.

References

32. Yan, T., Chu, D., Ganesan, D., Kansal, A., and Liu, J. Fast app launching for mobile devices using predictive user context. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*, ACM (2012), 113–126.
33. Yin, X., and Han, J. Cpar: Classification based on predictive association rules. In *Proceedings of the 2003 SIAM International Conference on Data Mining (SDM '03)*, SIAM (2003).
34. Zaki, M. J. Spade: An efficient algorithm for mining frequent sequences. 31–60.
35. Zou, X., Zhang, W., Li, S., and Pan, G. Prophet: What app you wish to use next. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication (UbiComp '13 Adjunct)*, ACM (2013).

QUESTIONS AND DISCUSSION

Thank you!