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
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

![Graph showing idle time per day (in hours) with WeekDay and WeekEnd bars for different users.](image-url)
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. \( \{A: \text{Antecedents}\} \rightarrow \{B: \text{Consequence}\} \)
  - **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”
    - \[ B = \{ b_1^{w_1}, b_2^{w_2}, \ldots, b_n^{w_n} \}; \sum_{i=1}^{n} \text{contain}(b_i^{w_i}, X), w_i, \]
  - 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?

![Graph showing number of users by number of days of data collection]

- Number of users
- Number of days of data collection
- <= 20, 21-40, 41-60, 61-80, >=81
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

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

![Graphs showing recall vs precision for different data sets and algorithms.](attachment:graphs.png)

(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
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

<table>
<thead>
<tr>
<th>Precision</th>
<th>No. of recommendations</th>
<th>Recall</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>3</td>
<td>35%</td>
<td>30.95%</td>
</tr>
<tr>
<td>80%</td>
<td>3</td>
<td>51%</td>
<td>16.67%</td>
</tr>
<tr>
<td>80%</td>
<td>5</td>
<td>68%</td>
<td>23.81%</td>
</tr>
<tr>
<td>80%</td>
<td>7</td>
<td>80%</td>
<td>11.90%</td>
</tr>
<tr>
<td>75%</td>
<td>3</td>
<td>66%</td>
<td>4.76%</td>
</tr>
<tr>
<td>75%</td>
<td>5</td>
<td>87%</td>
<td>11.9%</td>
</tr>
<tr>
<td>75%</td>
<td>7</td>
<td>100%</td>
<td>19.05%</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Precision</th>
<th>No. of recommendations</th>
<th>Recall</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>3</td>
<td>35%</td>
<td>30.95%</td>
</tr>
<tr>
<td>80%</td>
<td>3</td>
<td>51%</td>
<td>16.67%</td>
</tr>
<tr>
<td>80%</td>
<td>5</td>
<td>68%</td>
<td>23.81%</td>
</tr>
<tr>
<td>80%</td>
<td>7</td>
<td>80%</td>
<td>11.90%</td>
</tr>
<tr>
<td>75%</td>
<td>3</td>
<td>66%</td>
<td>4.76%</td>
</tr>
<tr>
<td>75%</td>
<td>5</td>
<td>87%</td>
<td>11.9%</td>
</tr>
<tr>
<td>75%</td>
<td>7</td>
<td>100%</td>
<td>19.05%</td>
</tr>
</tbody>
</table>
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


References


References

References


References


References


QUESTIONS AND DISCUSSION

Thank you!