ADAPTIVE PROFILE DRIVEN DATA CACHING AND PREFETCHING IN MOBILE ENVIRONMENT

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the School of Information Technologies at The University of Sydney

Omer Mahmood
December 2005
ABSTRACT

This thesis describes a new method of calculating data priority by using adaptive mobile user and device profiles which change with user location, time of the day, available networks and data access history. The profiles are used for data prefetching, selection of most suitable wireless network and cache management on the mobile device in order to optimally utilize the device’s storage capacity and available bandwidth.

Some of the inherent characteristics of mobile devices due to user movements are – non-persistent connection, limited bandwidth and storage capacity, changes in mobile device’s geographical location and connection (e.g. connection can be from GPRS to WLAN to Bluetooth). New research is being carried out in making mobile devices work more efficiently by reducing and/or eliminating their limitations. The focus of this research is to propose, evaluate and test a new user profiling technique which specifically caters to the needs of the mobile device users who are required to access large amounts of data, possibly more than the device storage capability during the course of the day or week. This work involves the development of an intelligent user profiling system along with mobile device caching system which will first allocate weight (priority) to the different sets and subsets of the total given data based on user’s location, user’s appointment information, user’s preferences, device capabilities and available networks. Then the profile will automatically change the data weights with user movements, history of cached data access and characteristics of available networks.

The Adaptive User and Device Profiles were designed to handle broad range of the issues associated with:

- Changing network types and conditions
- Limited storage capacity and document type support of mobile devices
- Changes in user data needs due to their movements at different times of the day
Many research areas have been addressed through this research but the primary focus has remained on the following four core areas. The four core areas are: selecting the most suitable wireless network; allocating weights to different datasets & subsets by integrating user’s movements; previously accessed data; time of the day with user appointment information and device capabilities.
ACKNOWLEDGEMENTS

It is not possible to thank everyone who contributed assistance, encouragement and advice, but the success of this project would not have been possible without anyone of them.

Particular thanks go to my supervisor Professor Albert Zomaya, for his support and encouragement during the course of this work and for helping to bring it all to an end. I am very grateful to him for all his ideas, his continuous interest in my work and most of all for his never ending optimism.

Further, special thanks to MySql, Netbeans and Java community for giving me access to excellent free softwares.
DEDICATION

I dedicate this thesis to my father, Dr. Safdar Mahmood, a distinguished scholar and writer, who has always been a source of advice, inspiration, encouragement and motivation. I thank him for helping me to develop independent, creative and critical thinking.
## Contents

**LIST OF FIGURES** ix

**EQUATIONS** x

**LIST OF ABBREVIATIONS** xi

1. **INTRODUCTION** 1
   1.1. Motivation 2
   1.2. Summary and Thesis Overview 4

2. **BACKGROUND** 6
   2.1. Prefetching Background 6
   2.2. Profile Driven Prefetching 9
   2.3. Cache Management and User Profile 9

3. **APPLICATION DOMAIN** 15

4. **IMPLEMENTATION** 18
   4.1. Brief Overview 18
   4.2. Broad Objective of the System 19
   4.3. Core Objectives of the System 20
   4.4. Device Profiling System 21
      4.4.1. Document (MIME) Support System 21
      4.4.2. Cache Management System 22
      4.4.3. Automatic Network Selection System 23
   4.5. **USER PROFILING SYSTEM** 23
   4.6. **CONTEXT-AWARE PREFETCHING SYSTEM** 24
   4.7. **SIMULATOR ARCHITECTURE AND PERFORMANCE** 27
      4.7.1. Static Data Layer 28
      4.7.2. Random Information Generator 36
      4.7.3. Runtime Environment Layer 45
         4.7.3.1. Conventional System Simulation 51
         4.7.3.2. Proposed System Simulation 54
      4.7.4. Simulator Performance 57

5. **EVALUATION AND CONCLUSION** 61
   5.1. Evaluation 61
   5.2. Conclusion 64

6. **Future Directions** 66
# APPENDICES

A. **APPENDIX A: Simulation Diagrams**  
   67

B. **APPENDIX B: Source Code of the System**  
   B I) Random Information Generator  
   89
   B II) Runtime Environment Simulator  
   160

C. **APPENDIX C: Database Scripts**  
   C I) Runtime Layer Database  
   242
   C II) Static Layer Database  
   251

D. **APPENDIX D: Installation and Execution**  
   264

E. **BIBLIOGRAPHY**  
   300
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Figure 1: Architecture Data Recharging Service by Mitch</td>
</tr>
<tr>
<td>11</td>
<td>Figure 2: Profile use in a Proposed System</td>
</tr>
<tr>
<td>13</td>
<td>Figure 3: The MobiScape Model</td>
</tr>
<tr>
<td>27</td>
<td>Figure 4: Layered Structure of Simulator</td>
</tr>
<tr>
<td>29</td>
<td>Figure 5: Tables and their Relations in Static Layer</td>
</tr>
<tr>
<td>31</td>
<td>Figure 6: Suburb Structure and Shape</td>
</tr>
<tr>
<td>31</td>
<td>Figure 7: Inter Suburb Links and User Movements</td>
</tr>
<tr>
<td>37</td>
<td>Figure 8: Table and their Relations in Runtime Layer</td>
</tr>
<tr>
<td>49</td>
<td>Figure 9: UML - Object Diagram of Runtime Environment</td>
</tr>
<tr>
<td>58</td>
<td>Figure 10: Profiling Results for Run Time Environment</td>
</tr>
<tr>
<td>59</td>
<td>Figure 11: Profiling Results of Product Data Method</td>
</tr>
<tr>
<td>60</td>
<td>Figure 12: Profiling Results of Without Profile Environment Simulation Method</td>
</tr>
<tr>
<td>60</td>
<td>Figure 13: Profiling Results of Profile Environment Simulation</td>
</tr>
</tbody>
</table>
# Equations

<table>
<thead>
<tr>
<th>Equation</th>
<th>Page</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equation 1</td>
<td>25</td>
<td>Cached data weight equation</td>
</tr>
<tr>
<td>Equation 2</td>
<td>26</td>
<td>Prefetch data packet weight equation</td>
</tr>
<tr>
<td>Equation 3</td>
<td>52</td>
<td>Utility per data object</td>
</tr>
<tr>
<td>Equation 4</td>
<td>62</td>
<td>Utility from Access Time (Per Second) without using Profiles</td>
</tr>
<tr>
<td>Equation 5</td>
<td>63</td>
<td>Utility from Downloaded Data per byte without using Profiles</td>
</tr>
<tr>
<td>Equation 6</td>
<td>63</td>
<td>Utility from Access Time (Per Second) by using Profiles</td>
</tr>
<tr>
<td>Equation 7</td>
<td>64</td>
<td>Utility from Downloaded Data per byte by using Profiles</td>
</tr>
<tr>
<td>Equation 8</td>
<td>64</td>
<td>Percentage Change in Utility Derived from Access Time per Second</td>
</tr>
<tr>
<td>Equation 9</td>
<td>65</td>
<td>Percentage Change in Utility Derived from Downloaded Data per Byte</td>
</tr>
</tbody>
</table>
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVP</td>
<td>Adaptive Value-Based Prefetch Scheme</td>
</tr>
<tr>
<td>MARS</td>
<td>Mobility-Aware Cache Replacement</td>
</tr>
<tr>
<td>GPRS</td>
<td>General Packet Radio Service</td>
</tr>
<tr>
<td>WWW</td>
<td>World Wide Web</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless Local-Area Network</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>SS</td>
<td>Support Station</td>
</tr>
<tr>
<td>MH</td>
<td>Mobile Host</td>
</tr>
<tr>
<td>HTML</td>
<td>Hyper Text Markup Language</td>
</tr>
<tr>
<td>HTTP</td>
<td>Hyper Text Transfer Protocol</td>
</tr>
<tr>
<td>PDA</td>
<td>Personal Digital Assistance</td>
</tr>
<tr>
<td>AVP</td>
<td>Adaptive Value-Based Prefetch Scheme</td>
</tr>
<tr>
<td>GSM</td>
<td>Global System for Mobile Communication</td>
</tr>
<tr>
<td>UMTS</td>
<td>Universal Mobile Telecommunications System</td>
</tr>
<tr>
<td>MULTICS</td>
<td>Multiplexed Information and Computing Service</td>
</tr>
<tr>
<td>OBL</td>
<td>One Block Look-Ahead</td>
</tr>
<tr>
<td>MSHR</td>
<td>Miss Information Status Handling Register</td>
</tr>
<tr>
<td>MIME</td>
<td>Multipurpose Internet Mail Extensions</td>
</tr>
<tr>
<td>UML</td>
<td>Unified Modeling Language</td>
</tr>
<tr>
<td>URL</td>
<td>Uniform Resource Locator</td>
</tr>
<tr>
<td>ERD</td>
<td>Entity Relationship Diagram</td>
</tr>
<tr>
<td>OO</td>
<td>Object Oriented</td>
</tr>
<tr>
<td>LRU</td>
<td>Least Recently Used</td>
</tr>
<tr>
<td>DOM</td>
<td>Document Object Model</td>
</tr>
<tr>
<td>UP</td>
<td>Uniprocessor</td>
</tr>
<tr>
<td>TIP</td>
<td>Transparent Informed Prefetching</td>
</tr>
<tr>
<td>PDF</td>
<td>Portable Document Format</td>
</tr>
<tr>
<td>PPT</td>
<td>Power Point Document</td>
</tr>
<tr>
<td>UML</td>
<td>Unified Modeling Language</td>
</tr>
</tbody>
</table>