

**A PERVASIVE FRAMEWORK FOR IDENTIFYING
ACTIVITY PATTERNS OF MOBILE USERS AND
PREDICTING ACTIVITIES**

Diyunugalge Chamika Sandun Weerasinghe

168275N

M.Sc. in Computer Science

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

March, 2020

**A PERVASIVE FRAMEWORK FOR IDENTIFYING
ACTIVITY PATTERNS OF MOBILE USERS AND
PREDICTING ACTIVITIES**

Diyunugalge Chamika Sandun Weerasinghe

168275N

This report is submitted in partial fulfillment of the requirements for the Degree of
Master of Science in Computer Science specializing in Mobile Computing

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

March, 2020

DECLARATION

I declare that this is my own work and this report does not incorporate without acknowledgement any material previously submitted for the degree or diploma in any other university or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles of books).

Signature: 

Date: 2020/03/15

Name: D. C. S. Weerasinghe

I certify that the declaration above by the candidate is true to the best of my knowledge and that this report is acceptable for evaluation for the CS-5999 PG Diploma Project.

Name of the supervisor: Dr. Indika Perera

Signature of the supervisor:

Date:

ABSTRACT

Smartphones has become one of the most used devices in day to day life. Even though they already have so many features, they still lack the ability to identify user's context and the intentions. This is important for improving user experience and make existing mobile application more user friendly. The issue is that there is no underlying support either from operating system or software level to predict the user's intensions based on user context.

The main objective of this research is to come up with a framework to predict user intentions based on user context by identifying activity patterns. The framework must be run in-device so that it will function irrespective of the network connectivity.

We selected "clustering" as the approach because it does not involve high computation power or complexities to run in-device. We identify activity patterns by clustering the user's actions and then predict based on the closest cluster for the given time. We have evaluated K-means and Expectation-maximization (EM) clustering algorithms for compatibility for the framework. Unlike computers, mobile devices do not have powerful CPUs or memory. Therefore, we measured CPU time and memory usage of these algorithms to select the best. To maintain low-end device compatibility, we tuned in the algorithm parameters to achieve high accuracy keeping the CPU and memory consumption in low levels.

In conclusion, we have successfully identified that EM clustering is suitable for high-end devices and it gives high accuracy while K-means is suitable for low-end devices with acceptable accuracy. We have implemented the framework as an Android library and developed a proof of concept application by embedding the implemented library to show that this research will actually enables application developers to give better user experience to their applications.

ACKNOWLEDGMENT

I owe my deepest gratitude to my supervisor, Dr. Indika Perera, for his invaluable support in providing relevant knowledge, advice and supervision throughout the project. This would not have been possible without his expertise and continuous guidance.

Finally, I would like to thank my colleagues at CodeGen Int. Pvt Ltd and Bhasha Lanka Pvt. Ltd for covering my work and helping me to balance the workload. Without them, this project would not have been possible.

Last but not least, I am grateful for all the people who supported me throughout this research in various means.

Table of Contents

List of Tables.....	vi
List of Figures	vii
List of Abbreviations.....	viii
1. INTRODUCTION	2
1.1 Background	2
1.2 Problem Statement	3
1.3 Objectives	4
1.3.1 Other Objectives.....	4
1.4 Overview of the Document	4
2. LITERATURE REVIEW.....	7
2.1 Context Identification.....	7
2.1.1 Data collection using mobile device sensors	8
2.1.2 Data collection using mobile device applications.....	10
2.1.3 Preprocessing	10
2.1.4 Feature Extraction	11
2.2 Activity Recognition	12
2.2.1 Activity recognition through learning.....	13
2.2.2 Different classification techniques	15
2.3 Activity Prediction	22
2.3.1 Event-Condition-Action (ECA) Model.....	22
2.3.2 Clustering	24
2.3.3 Analysis.....	25
3. METHODOLOGY.....	27
3.1 Proposed Solution.....	27

3.2	Workflow.....	27
3.2.1	Data Collection.....	28
3.2.2	Clustering Data and Activity Predicting	30
3.2.3	Evaluation	33
3.2.4	Implementing the Framework.....	34
4.	SYSTEM ARCHITECTURE AND IMPLEMENTATION	38
4.1	Overview	38
4.2	Data Collecting Application.....	39
4.3	Prediction Framework	42
4.3.1	Prediction Engine.....	43
4.4	Proof of Concept Application.....	45
5.	RESULTS AND EVALUATION.....	51
5.1	Overview	51
5.2	Testing Process.....	51
5.2.1	Parameter Tuning.....	51
5.2.2	Optimization Testing.....	53
5.3	Evaluation.....	54
6.	CONCLUSION.....	58
6.1	Research Contribution.....	58
6.2	Limitations.....	59
6.2.1	Scalability.....	59
6.2.2	Accuracy	59
6.2.3	Privacy.....	59
6.3	Future Research Directions	59
7.	REFERENCES.....	61

List of Tables

Table 2.1 Classification Techniques	16
Table 3.1 Different test scenarios and the measured metrics.....	34
Table 5.1 Accuracy percentages for different cluster counts.....	52
Table 5.2 Total clustering time in seconds for different cluster counts.....	52
Table 5.3 Results comparison between different clustering algorithms	53

List of Figures

Figure 1.1 Devices used to access the internet, by age group in UK 2018.....	2
Figure 2.1 Generic Process for identifying context continuously.....	7
Figure 2.2 (a) left to right HMM model (b) ergodic HMM model	16
Figure 2.3 Sample decision tree having branching factor of 2	17
Figure 2.4 Structure of a general Neural Network.....	18
Figure 2.5 Placing the new data point in K-Nearest Neighbor Clustering [3].....	21
Figure 2.6 ECA architecture for activity prediction.....	24
Figure 3.1 Research Workflow	28
Figure 4.1 Data Collecting Application System Architecture	39
Figure 4.2 Sample data of events database table.....	41
Figure 4.3 Data Collecting Application Main Screen.....	42
Figure 4.4 Prediction Framework System Architecture.....	43
Figure 4.5 Clustering Data Mapper Interface	44
Figure 4.6 Prediction Engine Interface	45
Figure 4.7 Proof of concept application settings.....	46
Figure 4.8 Application and SMS predictions screens	47
Figure 4.9 Cluster test screen	48
Figure 4.10 Cluster accuracy screen	49
Figure 5.1 Memory usage when regular K-means clustering and testing.....	54
Figure 5.2 Memory usage when K-means optimized clustering and testing	54
Figure 5.3 Different average accuracy percentages	55

List of Abbreviations

Abbreviation	Description
IoT	Internet of Things
HMM	Hidden Markov Model
MSE	Mean Square of classification Error
NN	Neural Networks
KNN	K-Nearest Neighbors
ECA	Event-Condition-Action
GPS	Global Positioning System
API	Application Programming Interface
SDK	Software Development Kit
IDE	Integrated Development Environment
App	Mobile Application
EM	Expectation-Maximization