A NOVEL USER ACTIVITY PREDICTION MODEL FOR CONTEXT AWARE COMPUTING SYSTEMS

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF INFORMATICS
OF THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

SERHAT PEKER

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF INFORMATION SYSTEMS

SEPTEMBER 2011

Approval of the Graduate School of Informatics

____

Prof. Dr. Nazife Baykal
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

____

Prof. Dr. Yasemin Yardımcı Çetin
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

____

Assoc. Prof. Dr. Altan Koçyiğit
Supervisor

Examining Committee Members

Assist. Prof. Dr. Sevgi Özkan	(METU, II)	____________________
Assoc. Prof. Dr. Altan Koçyiğit	(METU, II)	____________________
Assist. Prof. Dr. Aysu Betin Can	(METU, II)	____________________
Dr. Nail Çadallı	(KAREL A.Ş)	____________________
Assist. Prof. Dr. Pekin Erhan Eren	(METU, II)	____________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Serhat Peker

Signature : __________________

ABSTRACT

A NOVEL USER ACTIVITY PREDICTION MODEL FOR CONTEXT AWARE COMPUTING SYSTEMS

Peker, Serhat
M.Sc., Department of Information Systems
Supervisor: Dr. Altan Koçyiğit

September 2011, 77 Pages

In the last decade, with the extensive use of mobile electronic and wireless communication devices, there is a growing need for context aware applications and many pervasive computing applications have become integral parts of our daily lives. Context aware recommender systems are one of the popular ones in this area. Such systems surround the users and integrate with the environment; hence, they are aware of the users' context and use that information to deliver personalized recommendations about everyday tasks. In this manner, predicting user’s next activity preferences with high accuracy improves the personalized service quality of context aware recommender systems and naturally provides user satisfaction. Predicting activities of people is useful and the studies on this issue in ubiquitous environment are considerably insufficient. Thus, this thesis proposes an activity prediction model to forecast a user’s next activity preference using past preferences of the user in certain contexts and current contexts of user in ubiquitous environment. The proposed model presents a new approach for activity prediction by taking advantage of ontology. A prototype application is implemented to demonstrate the applicability of this proposed model and the obtained outputs of a sample case on this application revealed that the proposed model can reasonably predict the next activities of the users.

Keywords: Activity Prediction, Pervasive/Ubiquitous Computing, Context-aware Recommender Systems, Context Prediction, Ontology Based Systems.

ÖZ

BAĞLAM BİLİNÇLİ SİSTEMLER İÇİN YENI BİR KULLANICI AKTİVİTESİ TAHMİN MODELİ

Peker, Serhat
Yüksek Lisans, Bilişim Sistemleri Bölümü
Tez Yöneticisi: Dr. Altan Koçyiğit

Eylül 2011, 77 Sayfa

Son 10 yılda, mobil elektronik ve kablosuz ağ iletişim araçlarının yaygın kullanımıyla birlikte, bağlam bilinçli uygulamalara artan bir gereksinim vardır ve birçok yaygın hesaplama uygulamaları günlük hayatımızın tamamlayıcı bir kısmını oluşturmaktadır. Bağlam bilinçli öneri sistemleri bu alandaki en yaygın olanlardan bir tanesidir. Bu tarz sistemler kullanıcıların etrafını çevreler ve çevreyle bir bütün oluştururlar; bunun sonucunda kullanıcıların durum bilgisinden haberdar olurlar ve bu bilgiyi günlük işlerle ilgili kişiye özel öneri vermek için kullanırlar. Bu bağlamda, kullanıcıların bir sonraki aktivite tercihlerini yüksek kesinlikle tahmin etmek bağlam bilinçli öneri sistemlerinin kişiye özel servis kalitesini geliştirir ve doğal olarak kullanıcı memnuniyeti sağlar. Kullanıcıların aktivitelerini tahmin etmek yararlıdır ve yaygın çevrelerde bu konuda yapılan çalışmalar oldukça yetersizdir. Bu yüzden, bu tez yaygın ortamda kullanıcının geçmişteki belli bağlamlardaki tercihlerini ve mevcut bağlamlarını kullanarak kullanıcının bir sonraki aktivite tercihini tahmin eden bir aktivite tahmin etme modeli önerir. Bu önerilen model ontoloji avantajını kullanarak aktivite tahmini için yeni bir yaklaşım sunar. Önerilen modelin uygulanabilirliğini göstermek için bir prototip uygulama geliştirilmiştir ve bu uygulama üzerindeki örnek bir vakadan elde edilen sonuçlar açığa çıkarmıştır ki önerilen model, kullanıcıların bir sonraki aktivitelerini mantıklı olarak tahmin etmektedir.

Anahtar Kelimeler: Aktivite Tahmin Etme, Her yerde Bilişim, Bağlam Bilinçli Öneri Sistemleri, Bağlam Tahmin Etme, Ontoloji Tabanlı Sistemler

ACKNOWLEDGEMENT

First of all, I would like to thank my supervisor Assoc. Prof. Dr. Altan Koçyiğit for being a great advisor throughout my thesis study. His ideas and directions helped me stay on the right track all the time. I have learned a lot from him during my mater thesis. I am proud of being one of his students and I owe him my eternal gratitude.

I would like to express my deepest gratitude to my mother for her endless love, encouragement, and support in every part of my life. This thesis would never have been completed without her endless love and support.

I am forever grateful to Gaye Alkan for her continuous motivation, endless patience, support and understanding during my thesis.

I am especially thankful to İ.Güven Arkılıç, Davut Çavdar and Özge Gürbüz who share the same destiny with me. They are always with me whenever I need them. Thanks for your endless motivation, support, encouragement and help throughout the thesis study.

I also want to thank my close friends Ayşegül Özkaya, Erhan Uyar and Gökçen Yılmaz for their precious friendships, motivations and help throughout the my thesis.

I would also like to thank Oğuzhan Alaşehir, Mahir Kaya, Süleyman Özarslan and Emre Sezgin for their support and help during the thesis study.

I would like to express my thanks to my thesis committee, Assist. Prof. Dr.Sevgi Özkan, Assist. Prof. Dr. Erhan Eren, Assist. Prof. Dr. Aysu Betin Can and Dr. Nail Cadallı for their inspiring questions and suggestions.

I would like to thank to administrative personal of institute, Sibel Gulnar, Necla Işıklar and Ali Kantar for their kindness and help since the beginning of my M.Sc. study.

Finally, I want to thank to The Scientific and Technical Research Council of Turkey (TÜBİTAK) for supporting me with scholarship during my MSc study.

To my mother, Gülfem Arslanoglu and grandmother, Hatice Arslanoglu
for bringing me to these days
And
being always there for me with their endless love.
I have learnt the most important things in life from you.
I always owe you my endless gratitude…

TABLE OF CONTENTS

LIST OF TABLES

TABLE OF FIGURES

LIST OF ACRONYMS

RS	Recommender Systems
CARS	Context-aware Recommender Systems
CP	Context Prediction
DFD	Data Flow Diagram
EWMA	Exponential Weighted Moving Average
PT	Prediction Threshold
CT	Occurrence Count Threshold
AC	Activity Appearance Count
DC	Disable Activity Appearance Count
IT	Interval Time
IOT	Interoccurrence Time
$\mathrm{I\widetilde{O}T}$	Estimated Interoccurrence Time
$\widetilde{\mathrm{Var}}$	Estimated variance
$\tilde{\sigma}$	Estimated standard deviation
$\tilde{D}$	Estimated D score

CHAPTER 1

INTRODUCTION

Today we have limitless information available everywhere and this huge amount of information provides unlimited choices to us, so the actual problem is making smart decisions with the plenty of alternatives (Anderson, 2006). In this manner, the recommendations play important roles in our daily lives in this era.

With the rise of the Internet, we have too much information on our fingertips. The rapid development in the mobile and wireless communication technology lowers size and costs of the computers and allow them to become available everywhere. People are also enthusiastic to use these technologies such as notebooks, smart phones, PDAs, etc.

Besides these improvements, with the use of sensor technology mobile devices will be used to interact with the appliances and environment and so ubiquitous computing, also generally referred to as pervasive computing technologies have been growing tremendously in the last two decades. Ubiquitous computing means infusing the computers into the physical environment while making them invisible to the humans (Weiser, 1991). It enables development of pervasive technologies that are highly accessible and usable by the humans (Poslad, 2009).

Soon the environments are going to be active by equipping with the sensors (McCarthy, 2001). These environments are called ubiquitous computing environments that support their individuals in a personalized manner and context aware systems are one of the popular applications in ubiquitous computing environments. Such systems sense the individuals' contexts and act proactively to present appropriate service or information for the user (Dey & Abowd, 2000). Context is any information describes the situation of an entity, where an entity can be a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves (Dey et al., 2000).

All of these technological improvements enable limitless information source everywhere and this leads us to information overload problem. We have difficulty in making choices on any issue that satisfy our needs, and we do not have enough time to compare and evaluate the huge amount of information. We need filtered information and so this makes the recommendations more valuable. Recommendations from any reliable source ease our decision making process and direct us to the most appropriate choice. As a result, recommender systems are designed to overcome the information overload problems. Such systems help individuals make decisions quickly and efficiently by providing the most valuable and appropriate recommendations that suit their situation and needs.

People in different contexts (such as time, location, weather, etc.) may have different preferences and needs and so contextual information plays a crucial role in the preferences of people (Yujie & Licai, 2010). In this manner, context aware recommender systems (CARS) are developed by incorporating available contextual information into the recommendation process. Taking contextual information into the consideration, CARS help people in decision making process by presenting more personalized recommendations.

Various types of CARS (such as device, location, place, movie, etc.) have been developed for different domains. Most of them require too many inputs to provide recommendations. On the other hand, mobile devices have smaller screens, limited keypads, and supporting a conversation on such a device is extremely difficult. Thus, users usually do not like too much interaction with the systems by entering lots of data and understanding the questions and requests of the application may be challenge for some of the users.

There are some CARS that work implicitly in the literature. However, environments may be complex for some domains and CARS do not satisfy the user needs in such environments. For example, shopping centers are rich and heterogeneous environments and people can perform many different activities (such as shopping, eating, cinema, etc.) in them. Suppose you are in a shopping center and recommender application in your PDA recommends a movie for the cinema while you are thinking to go shopping. The recommended movie may be appropriate for your pleasure and profile and you may want to watch this movie in another time but now you want to do shopping. As a result, this recommendation is unnecessary and inappropriate for you at this time. Hence, it can be concluded that recommendations are valuable and appropriate for the individuals only when they are offered in the right times.

Outside activities take too much time of the people and making preference for these activities has always been an important issue in complicated or heterogeneous environments having various alternatives. CARS usually do not exactly meet the actual user needs in such environments. Because of that, there is a need to predict the next activities of the individuals to get the most valuable and appropriate recommendations from any recommender systems that work in heterogeneous environments. With such a solution, the user’s need is satisfied well as the exact matching is provided between the user’s intention and the recommended activity.

The purpose of this thesis is to develop a prediction model that predicts the next preferences of the individuals by using past preferences of individuals in certain context conditions and current context information of individuals and the environment. Context history of individuals’ previous activity preferences is used as input in our model. The individuals perform the activities within the specific time intervals and the context affects the preferences of them. In this manner, our model proposes a new prediction model based on the interoccurrence times of the past activities of the individuals in similar contexts.

Moreover, humans generally exhibit behavior through their habits (Lee & Tran, 2010) and they can temporarily interrupt or they can permanently change their behaviors (Petzold, Bagci, Trumler, & Ungerer, 2003). In order to overcome this challenge in the prediction, an activity management mechanism is proposed to realize the seasonal or periodic patterns in the activity occurrences, and take into account the exceptions, changed or interrupted activities of the user. The proposed model also uses ontology in the prediction process to take into account the changing but similar situations and for the challenge of inadequate historical context data.

The proposed model only provides prediction for shopping (only clothing), entertainment/enjoyment (cinema, theatre, bowling etc.), eating/drinking (restaurant, fast-food, café, etc.) type of outside activities. Moreover, the model in this thesis only predicts the activities but it does not suggest the details of them. As an example, the developed model predicts that the user will go to the cinema or she/he will eat in a restaurant. The detail of the activity is not in the scope of this thesis. For example, proposed model does not make a prediction that the user will go to ‘film X’ or ‘Restaurant Y’. Therefore, the predictions by our model can be used as a basis by the context aware recommender systems but making specific recommendations for any activity is not in the scope of this study.

This thesis is organized as follows. Context aware recommender systems, context prediction, various context prediction techniques in ubiquitous computing environments are discussed in chapter 2. Chapter 3 describes our activity prediction model. In this chapter, the context dimensions and the main terms used in the study are introduced, then the proposed prediction approach, ontology structures for each context dimensions and the algorithms of the proposed model are described. Usage of the model is also illustrated through the use of some sample scenarios in Chapter 3. A prototype application that makes use of the proposed model is introduced and a sample case which examines the proposed model on this application is proposed in Chapter 4. Finally, concluding remarks and directions for the future work are given in Chapter 5.

CHAPTER 2

LITERATURE REVIEW

In early 1990s, Mark Weiser who is the pioneer of the field of ubiquitous or pervasive computing technologies described the vision of these technologies as “The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it.”(Weiser, 1991). Context aware computing which is one of the major fields of ubiquitous computing has been a popular research field during recent years. The applications of this field provide proactivity by being aware of the user’s contexts and adapting the services according to the user needs.

Context awareness is the most important feature of these applications, since the context information (such as current location, current action and the surrounding environment) is extracted and used to serve user needs by such applications. Context aware recommender systems (CARS) are developed with the aim of using context information to recommend proper alternatives to the user. However, most of these systems do not meet the expectations of the users in the heterogeneous environments. Therefore context prediction is a new popular research filed nowadays.

This chapter presents a review of the related literature. The first section introduces traditional recommender systems and the context aware recommender systems and sample applications developed are presented. Finally, the second section introduces context prediction and different approaches and techniques proposed for context prediction are summarized.

2.1. Recommender Systems and Context Aware Recommender Systems

With the rapid development of information technology, there are many information services provided to the users and mobile users are able to access the information via such services wherever they are and whenever they want. These technological improvements have led us to the information overload problem. In order to overcome this problem, recommendations play an important role to make decisions quickly and efficiently. As a result, recommender systems emerged as an independent research area in the mid-1990s.

(Resnick & Varian, 1997) first introduced the recommender systems that help individuals to identify content of interest from various alternatives by using opinions of a community of users. Different methods have been proposed for the recommender systems. These recommendation methods are usually categorized into three main types which are collaborative filtering, content-based filtering and hybrid filtering (Balabanović & Shoham, 1997). These filtering strategies are defined by (Adomavicius & Tuzhilin, 2005) as:

- *Content-based recommendations* recommend the products similar to the previous preferences of the user.
- *Collaborative recommendations* recommend the products that people having similar tastes and preferences with user liked in the past.
- *Hybrid approaches* combine collaborative and content-based filtering strategies.

In other words, collaborative filtering assumes that the users have similar preferences with the people who has similar profile with the user, whereas content based filtering assumes that the users have similar preferences with their preferences in the past. Most of the recommender systems focus on recommending items to users based on customer's historical preferences and others are based on the available data on users (such as demographics) and items (such as content descriptions) (Lombardi, Anand, & Gorgoglione, 2009).

Personalized services help the organizations to build customer loyalty and increase the competitiveness in the marketplace by providing services to unique needs and preferences of individuals (Lombardi et al., 2009). In this manner, recommender systems filter the related information and support the individuals in decision making process by presenting the recommendations personalized to their needs.

Recommender systems have been implemented in several domains on the web especially for E commerce web sites such movie, music, books, news, etc. On the other hand, the mobile technologies have been growing tremendously in recent years and the demand for them continues to grow. Therefore, they are becoming the primary resources of information and recommender systems have been widely applied on the mobile technologies.

Furthermore, context-awareness is a growing subject in the ubiquitous computing environments and numerous context aware applications in different domains have been developed in recent years. However the traditional recommender systems only deal with the user personal needs and interests. They do not take into account any contextual information (such as time, location, weather, etc.) in the recommendation process.

Prahalad in (Prahalad, 2004) emphasized the importance of context in the marketplaces. Companies must deliver unique, real-time customer experiences shaped by customer context in order to satisfy the expectancies of the customers at any time. Delivering recommendations perfectly matching with users' preferences at the right moment, in the right place and on the right media is essential for a better recommender system and this can only be achieved by taking the all contextual information of the users (Abbar, Bouzeghoub, & Lopez, 2009). Moreover, individuals' needs and preferences may change according to contexts they are in. As a result, contextual information plays an important role in delivering the better recommendation for the needs of individuals (Yujie et al., 2010).

All of these studies mention about the importance of the context in recommender systems and incorporating contextual information into the recommendation process leads a new type of recommender system which is called context aware recommender system. (Van Setten, Pokraev, & Koolwaaij, 2004) propose a mobile tourist application that provides context-aware recommendations by integrating a recommender system with a context aware application.

(Adomavicius, Sankaranarayanan, Sen, & Tuzhilin, 2005) proposed an approach by incorporating available contextual information into the recommendation process and (Adomavicius et al., 2005) discusses the next generation of recommender systems. (Oku, Nakajima, Miyazaki, & Uemura, 2006) also incorporated the contextual information into the recommendation process in a restaurant recommender system. (Tuzhilin & Adomavicius, 2008) emphasized the importance of the contextual information and take this information into account when providing recommendations.

Although CARS help people in decision making process by presenting more personalized recommendations; the service quality of existing CARS are seriously restricted. They only consider the current context and they do not predict the future contexts. However, predicting the future contexts is required to get the most valuable and appropriate recommendations from CARS. Therefore, context prediction which is to predict the future contexts and proactively recommend high-quality services for users in ubiquitous computing environments has become popular in recent years (S. Lee, Lee, & Cho, 2010).

2.2. Context Prediction and Context Prediction Techniques

As emphasized in the previous section, many context aware recommender systems and applications for ubiquitous environments are proposed in the literature and these systems are aware of the context of the individuals. However, only context awareness does guarantee proactiveness (Tennenhouse, 2000). Information about users’ future needs is required for proactiveness (K. C. Lee, Cho, & Lee, 2010). Because of that, most of the existing CARS do not exactly meet the user actual needs especially in heterogeneous environments which offer plenty of alternatives.

Context prediction uses context information acquired from various types of sensors and observed context history in the prediction process of the next situation of the individuals in ubiquitous computing environment. It aims to predict the next context that users will likely enter (for the location and situation contexts) or perform (for the action contexts) on the basis of a context history (S, Lee et al., 2010). (Byun & Cheverst, 2004) emphasized the importance of context history usage in ubiquitous computing environments. Since the context history usually embodies patterns, the more personalized services can be provided to the user according to such patterns.

In the literature, numerous context-prediction approaches and techniques for various domains such as, location, movement, action, etc have been proposed. One of the early studies is The Adaptive House Project (Mozer, 1998). It is a smart house and the lifestyle and desires of the inhabitants are observed and learned to forecast their future needs. Neural networks are used to predict user needs and desires. (Davison & Hirsh, 1998) is also one of the early studies. This study proposed a method to predict the next Unix commands of the user based on the frequency of last two succeeding commands.

(Kaowthumrong, Lebsack, & Han, 2002) proposed an automatic active device selection model and Markov chains are used to predict the user’s next device choice. (Patterson, Liao, Fox, & Kautz, 2003) presented a model to predict the travel destinations of user on a city map. Dynamic Bayesian network is used in this study. Another study for the location prediction was proposed by (Laasonen, Raento, & Toivonen, 2004) that use a Markov predictor and a weighted graph to forecast the users’ future locations.

(Petzold, Bagci, Trumler, Ungerer, & Vintan, 2004) proposed a context prediction approach based on previous behavior patterns of the users. A new approach with global first-level histories and two-state predictors is proposed to predict the next movement of a user in an office environment. (Byun, et al. 2004) proposed a model that predicts the preferences of the user in an intelligent office environment by using decision trees.

A detailed overview on context prediction is given by Mayrhofer. (Mayrhofer, 2005) proposed an architecture for context prediction. Aspects and important issues of context prediction, some of the benefits and challenges are also given in this study. Moreover, this study presents comparison of performances of different methods such as neural networks, Markov models, autoregressive moving average model (ARMA) forecasting, and support vector regression.

(Sigg, Haseloff, & David, 2006) introduced a context prediction approach based on time series based local alignment methods. Moreover, (Sigg, Haseloff, & David, 2010) proposed the alignment prediction approach based on a time-series-estimation technique. In this study, the proposed approach is compared with the prediction algorithms of Markov, ARMA, PCA, and ICA (independent component analysis).

(K. C. Lee et al., 2010) proposed a new type of ubiquitous decision support system that uses a General Bayesian Network (GBN) for context prediction. It improves decision support by taking advantage of the what-if analysis. Furthermore (S. Lee et al., 2010) proposed a Dynamic Bayesian Network (DBN) approach to location prediction for ubiquitous computing environments.

(S. Y. Lee & Tran, 2010) proposed a user's habits-based context prediction algorithm to predict users' next location in order to provide users with suitable services. In this approach users' behaviors are monitored and their habits are learned from acquired context information to predict their next locations in a building. Hybrid P2P technique is also used to share and learn context information from other members.

Although there are many context prediction models and approaches for different domains, activity or service prediction is one of the least studied problems in this research field. In recent years, some studies based on activity or service prediction in context aware computing environment have been proposed.

(Hong, Suh, Kim, & Kim, 2009) proposed an agent based framework for predicting the preferences of users and providing the personalized services by using context history. In this study, a context history consists of users’ profile, current context of users and the services selected by the users is used. Apriori algorithm is used to extract the relationship among the services for predicting and recommending the next service after offering the previous service.

(Chen, Shao, Xie, & Huang, 2010) proposed an attribute-based scheme to predict the service preferences of the users and provide recommendations. In this approach the services are classified into several service clusters, and the apriori and colony algorithms are used in the service recommendation process.

As seen so far, context prediction is useful for predicting the next contexts of the individuals. Context prediction can be employed in CARS to more effectively satisfy the users' needs by providing more customized recommendations in ubiquitous environments. Many different techniques and methods have been applied in several different context prediction models and approaches on various domains.

However, the research concerning activity prediction techniques is a relatively insufficient. Moreover, there is no activity prediction model considering the time intervals between the occurrences of the activities. Time intervals of occurrences of activities which we call interoccurrence times in this study affect the individuals' preferences. Therefore, in this study, we propose a new prediction approach based on the interoccurrence times of the user's previous activities. Section 3 describes our new prediction model to forecast the next activity preferences of the individuals in context aware computing environments.

CHAPTER 3

ACTIVITY PREDICTION MODEL

Most of us prefer to do any behavior or activity according to the context that we are in. In other words, the contexts of individuals affect their preferences and individuals do similar activities in the similar contexts. For example, a user can do similar activities every Saturday or when the weather is rainy. Therefore, context information plays an important role on the decisions of the people and the user's future preferences can be forecasted by the means of the user context history.

Time interval is another important issue that affects the preferences of the individuals. Individuals do similar activities within the specific time intervals. As an example, a user can go to cinema once a week or go to shopping once a month. As a result, the activity preferences of any user can be predicted by the means of this information.

As the above examples suggest, the future activity preferences of the users can be predicted based on the preferences in the specific context information in the past and also the information of time interval they do these activities. In this thesis, we use this idea and proposed a new prediction approach that uses the context history and the current context of the user to predict the user's next activity. The prediction algorithms devised basically matches the current context with the entries in the context history and makes use of activity interoccurrence times computed from the returned entries to rank possible activities.

However, there are two fundamental problems with the above idea. First, humans tend to change their preferences. They sometimes interrupt an activity that they usually do. For example, a person who often goes to cinema does not want to prefer this activity in summer seasons and s/he can interrupt this activity in that season. Moreover, people may cease an activity and replace it with another activity preference. As an example, while a person usually go to fast food restaurant, s/he starts a diet or wants to have healthy foods and then s/he gives up to go fast food restaurants. As a result of these, seasonal or periodic patterns, changes and exceptions are important to be recognized by the prediction approach to fulfill the users' expectations and the future needs.

Encountering new context information that is not in the context history of the user is always possible for any user because of the variety of user contexts. For example, location as a context information can always change and a user might not have appeared in a location before. Therefore, it is difficult to predict the preferences of that user at that location as there is no similar record in the context history.

In order to solve the first problem, this study provides a model that manages status of each activity for each user. The proposed model recognizes the interrupted and changed activities of the users by keeping interoccurrence time, estimated interoccurrence time and estimated variance of interoccurrence time values for each activity for each user. Moreover, exceptions are considered in the prediction process by the use of the prediction threshold for each activity. The second problem is solved by the use of the ontology and when new context data occurs, the model uses the ontology and finds the similar contexts from the activity history.

This chapter consists of five main subsections. First section defines the context dimensions considered in this study. In the second section, main terms used in the model are introduced. Prediction approach is outlined in the third section. Next, ontology models for each context dimension are given and the algorithms used in the model are explained. Finally last section gives the typical scenarios to illustrate how to apply the proposed algorithms in the model.

3.1. Context Dimensions

Context has numerous dimensions and different dimensions of context have been defined in various studies. Since context is the core part of the prediction model proposed in this thesis, it is important to define the context dimensions that are considered. In this study, five context dimensions are considered: location, time frame, day, nearby person and weather. We prefer to incorporate these context dimensions in our model as they are the mostly used ones and they play important roles on the preferences of the people. The information for each of these context dimensions can be retrieved easily for any mobile device user from the service provider or the sensors on the mobile device. The context dimensions considered in this study are (these are summarized in Table 1):

- *Location:* It is the context that gives the location of the user. In this study we consider the outside activities of the people. Therefore the location is an important factor that affects the preferences of the users.

- *Time Frame of Day:* It is the context that provides time frame of the moment. In general, people's activity preferences change according to time frames. Therefore instead of the time instant, time frame of the moment is considered in this study. The time is partitioned into two-hour periods and these two hour time periods are used in this study as we assumed that the outside activities generally last two hours and the people usually finish an activity and start new one at most within two hours. Besides, the time between 08.00 and 02.00 are considered for the prediction in this study, because the outside activities are generally performed in this time period. For example, the extracted data that has time attribute 12.45 is interpreted as the time frame of 12.00-14.00 in this study.

- *Day of Week:* It is the context that provides day information. People's activity preferences change in terms of the days, if it is a weekday or weekend.

- *Nearby Person:* It is the context that provides the information about the person who accompanies to the user at that moment. The nearby person has an effect on the user's decision, because of that, this context information is considered in this study.

- *Weather:* It is the context that provides weather information of the moment. Weather also affects the user's personal mood and so the outside activity preferences change according to weather conditions. The sense of temperature values could change from user to user. For example, 25˚C might be hot for some users but mild for other users. Because of these reasons, specific temperature values are not considered. Only general weather conditions are used in this study and so the weather conditions given in Table 1 are considered in this study.

Table 1: Description of Context Dimensions

Dimension	Value	Example
Location	all types of indoor and outdoor locations.	Cepa, Bilkent, etc.
Day of Week	all days of the week.	Monday, Saturday, etc.
Time frame of Day	10.00-12.00, 12.00-14.00, 14.00-16.00, 16.00-18.00, 18.00-20.00, 20.00-22.00, 22.00-00.00, 00.00-02.00	10.00-12.00, etc.
Weather	weather conditions.	sunny, clear, cloudy, rainy, snowy
Nearby Person	any person accompanies to the user.	Tom, Marry, etc.

3.2. Main Terms Used in the Model

The terms of estimated interoccurrence time and estimated variance used in this study are inspired from the TCP work in (Jacobson, 1988). All the concepts related to the proposed model are the following:

- ***Interval Time (IT):*** Interval time is one of main variables used in this study. We can define the interval time in this study as: *the amount of time in day between the last occurrence date of an activity and today' date*. As an example, assume that last occurrence date of an activity is on 18 March and today's date is 25 March. If the last occurrence date of the activity is subtracted from today's date, the interval time of the activity is obtained according to today's date.

  (25 March)-(18 March) = 7, so the interval time of the activity is 7 days.

  Interoccurrence Time (IOT): Interoccurrence time is an essential variable used in both the management of activities' statuses process and the process of activity prediction algorithms and calculations related to these processes are based on it. We can define the interoccurrence time in this study as: *the amount of time in day* *between the occurrence date of an activity and the next occurrence date of the same activity*.

This variable is calculated and kept separately for each activity to manage the activities' statuses. As an example, assume that the last occurrence date of Activity A is 17 March and this activity occurs on 23 March for the last time, then if the date of activity's last occurrence is subtracted from that of the activity's current occurrence, the current IOT for activity A is obtained as:

(23 March)-(17 March) = 6, so the current IOT for activity A is 6 days.

On the other hand, IOT is also calculated during the data filtering stage. Since the prediction algorithms are based on the estimated value of this variable, it is calculated for each consecutive occurrence of each selected activity for a specific context. In this stage, matching entries for the each activity is sorted in ascending order according to the date of occurrences and IOT is calculated for each consecutive record pair of the each activity. As an example, assume that a selected activity in the prediction process occurs in the following days:

{14 March, 19 March, 22 March, 27 March,..}

According to this example data, if the date of activity's first occurrence is subtracted from that of the activity's second occurrence, the IOT between the activity's first and second occurrence is obtained as:

(19 March)-(14 March) = 5, so the IOT between the activity's first and second occurrence is 5 days.

This calculation process for each data of the selected activity is repeated for all the interoccurrence times.

- ***Estimated Interoccurrence Time ($I\widetilde{OT}$):*** In this study, an estimated interoccurrence time of interoccurrence time values is used in the management of activities' statuses process and the process of activity prediction. IOT values vary according to user's past preferences. In order to get a typical IOT value, our model takes the weighted average of IOT values. Our model maintains and keeps the estimated interoccurrence time ($I\widetilde{OT}$) value for each activity after the activity occurs for the first time. When an activity occurs, current IOT is calculated for this activity and $I\widetilde{OT}$ is updated in the process of management of activities' statuses. Moreover, this value is also calculated for the occurrences of each selected activity in the prediction process. In this process after the first entry of each activity, the new $\widetilde{\mathrm{IOT}}$ is updated and this calculation is repeated for all selected data of each activity. The formula for $\widetilde{\mathrm{IOT}}$ is inspired from TCP's timeout calculations (Jacobson, 1988) and it is given as:

$$\widetilde{\mathrm{IOT}}_{\mathrm{n}} = (\alpha * \mathrm{IOT}_{\mathrm{n}}) + ((1-\alpha) * \widetilde{\mathrm{IOT}}_{\mathrm{n-1}})$$

The weighting factor (α) is constant smoothing factor and it must be between 0 and 1. Its value is important factor that determines the decay velocity of the effect of old samples in $\widetilde{\mathrm{IOT}}$ value. A large α value indicates higher decay which means that the weights fall off more quickly. As a result a rapid decay is seen in the effect of old samples in the $\widetilde{\mathrm{IOT}}$ value. On the other hand, a small α value indicates lower decay which means that the weights fall off more slowly and the effect of old samples in $\widetilde{\mathrm{IOT}}$ value also decays slowly.

As seen in the formula, $\widetilde{\mathrm{IOT}}$ is a weighted average of the IOT values and as a result it depends on all previous IOT values. This weighted average puts more weight on recent data than old ones and such an averaging operation is called an exponential weighted moving average (EWMA) in statistics. The use of EWMA enables us to incorporate all related historical records in predictions.

- ***Estimated Variance ($\widetilde{Var}$):*** Estimated variance that is calculated by using current interoccurrence time, current $\widetilde{\mathrm{IOT}}$ and previous estimated variance values ($\widetilde{\mathrm{Var}}$) is used in the processes of activity management and prediction.

  When identifying dispersion in a set of numbers, sample variance is used and it uses equal weight assignment. With the use of sample variance calculation, very recent IOT, such as the last IOT has no influence on the new variance. Since IOT values vary according to user preferences and sample size is considerably large, the sample variance is not suitable for our study.

  To overcome these challenges, an estimated variance of interoccurrence time values ($\widetilde{\mathrm{Var}}$) that is calculated in the processes of activity management and prediction. As inspired from the TCP's timeout determination mechanism EWMA method is used again in calculation of $\widetilde{\mathrm{Var}}$ like in the calculation of $\widetilde{\mathrm{IOT}}$ value. When a new activity occurs, new $\widetilde{\mathrm{IOT}}$ is calculated for this activity and new $\widetilde{\mathrm{Var}}$ is updated in the process of management of activities' statuses. Furthermore, this $\widetilde{\mathrm{Var}}$ is also calculated for the occurrences of each selected activity in the prediction process. In this process after the first entry of each activity, the new $\widetilde{\mathrm{Var}}$ is updated and this calculation is repeated for all selected data of each activity. The formula for $\widetilde{\mathrm{Var}}$ is:

$$\widetilde{\mathrm{Var}} = \left(\left(|\mathrm{IOT_n} - \widetilde{\mathrm{IOT}}_\mathrm{n}|^2 \right) * \lambda \right) + \left((1 - \lambda) * \widetilde{\mathrm{Var}}_{\mathrm{n}-1} \right)$$

As seen in the formula, $\widetilde{\mathrm{Var}}$ is a weighted average of the square of the difference between the sample IOT and the $\widetilde{\mathrm{IOT}}$ values and previous variance values. Since an exponential weighted moving average (EWMA) is used in the calculations, more recent data have greater weight on the new $\widetilde{\mathrm{Var}}$. The weighting factor (λ) is a constant smoothing factor like the α value used in $\widetilde{\mathrm{IOT}}$ function and λ must also be between 0 and 1. Its value determines the decay velocity of the effect of old samples on $\widetilde{\mathrm{Var}}$ value. Like the α value, a large λ value indicates higher decay which means that the weights fall off more quickly and the effect of old samples on $\widetilde{\mathrm{Var}}$ value decays rapidly. On the other hand, a small λ value displays an opposite attitude of a larger one.

Estimated Standard Deviation ($\tilde{\sigma}$): An estimated standard deviation that is calculated by taking the square root of the estimated variance is used in the process of activity prediction. The equation for the estimated standard deviation is the following:

$$\tilde{\sigma} = \sqrt{\widetilde{\mathrm{Var}}}$$

- ***Activity Appearance Count (AC):*** It is the number that gives how many times the activity has appeared for a specific user. Appearance count of each activity is kept for the user separately and this number is updated, when the activity is performed again.

- ***Disabled Activity Appearance Count (DC):*** It is the number that gives how many times a 'Disabled' activity is performed by a specific user. Disabled appearance count of each activity is kept separately for each user and it is updated when the activity occurs while its status is 'Disabled'. When the status of activity turns to

‘Enabled’, this count is reset and it starts from the one after its status becomes ‘Disabled’ again.

- ***Occurrence Count Threshold (CT):*** It is the threshold and an activity’s occurrence count must exceed this threshold for that activity to be included in the prediction process. As an example, assume that appearance count threshold is n and activity A appears m times (m<n) for a specific user. As a result, this activity is not allowed to take part in the prediction process for that user, since it does not exceed the count threshold. When an activity A is seen again and its appearance count reaches to n, then this activity is included in the next prediction process for the user.

- ***Prediction Threshold (PT):*** Some user activities occur rarely and these are actually exceptions. In this manner, recognizing these exceptions is important to fulfill the users’ expectancies for the future preferences. To get a solution for this problem, a general prediction threshold is defined. PT defines which activities are included in the prediction process.

  The activity’s $\widetilde{\mathrm{IOT}}$ and plus its $\tilde{\sigma}$ as a margin is identified a criterion to compare with the prediction threshold. This sum must be less than the prediction threshold for an activity to include the activity in the prediction process. That is, the condition for any activity to be included in the prediction process is:

$$\widetilde{\mathrm{IOT}} + \tilde{\sigma} < \text{Prediction Threshold}$$

3.3. Prediction Approach

A new prediction approach based on the interocurrence times of the user’s performed activities is proposed in this study. The interocurrence times of the activities performed by the people are typical because people usually perform the similar activities after similar time intervals. As a result, this study presents an approach that predicts the next preferences of user by using the $\widetilde{\mathrm{IOT}}$ values of the users’ performed activities and evaluating the normalized distances of the activities’ current time intervals to $\widetilde{\mathrm{IOT}}$ values of them. These distances show that how much a specific activity is close to its $\widetilde{\mathrm{IOT}}$. The closer normalized distance between the current IT of the activity and its $\widetilde{\mathrm{IOT}}$, the more likely the activity occurs. In other words, the activity that has the least normalized distance to its $\widetilde{\mathrm{IOT}}$ has the highest possibility to be performed by the user.

In order to compute the normalized distance between the current IT of the activity and its $\widetilde{\mathrm{IOT}}$, we use formula inspired from Mahalanobis Distance method (Mahalanobis, 1936). The Mahalanobis (r) Distance from test point (x) to mean μ is defined by using the sample standard deviation as:

$$r = \frac{|\mathrm{x} - \mu|}{\sigma}$$

In our proposed model, basically, we query context history to find entries matching to current context and then $\widetilde{\mathrm{IOT}}$ and $\widetilde{\mathrm{Var}}$ values are calculated for each found activity of the user. Then, the normalization of each activity's IOT values is performed by the below transformed formula:

$$\mathrm{D} = \frac{|\mathrm{IT} - \widetilde{\mathrm{IOT}}\,|}{\tilde{\sigma}}$$

IT is the interval time that was defined as the number of days between the last occurrence date of the activity and today's date. $\widetilde{IOT}$ is the mean of interoccurrence time values of the activity and $\tilde{\sigma}$ is the standard deviation of $\widetilde{\mathrm{Var}}$. Finally *D* is the normalized distance of the current interval time to the $\widetilde{\mathrm{IOT}}$ value of the activity.

D score for each activity indicates the distance of the current IT to the $\widetilde{\mathrm{IOT}}$ value. If the distance between the current IT and the $\widetilde{\mathrm{IOT}}$ value is small for an activity, then it might be concluded that it is highly probable that this activity occurs as the next activity of the user.

3.4. Ontology for Contexts

Ontology was defined as 'explicit specification of a conceptualization' (Gruber, 1993). Ontologies are used as explicit and formal representations that describe the concepts and their relationships for particular domains. They are used in various research areas. Pervasive computing is one of the popular areas that use ontology and a number of ontology based approaches including CoBra (Chen, Finin, & Joshi, 2003), SOCAM (Gu, Wang, Pung, &

Zhang, 2004) and Gaia (Ranganathan, McGrath, Campbell, & Mickunas, 2003) have been proposed for pervasive computing.

As seen in the literature, the contextual information can be clearly defined by ontologies. As a result, we have developed ontologies that define the structure of the contexts used in our study. This section introduces the collection of ontologies considered in this study. The proposed ontology models the basics concepts of context dimensions used in this study which are location, person, time frame, day and weather. There is also a proposed ontology that shows the available activities for certain locations.

Encountering new data and insufficient amount of data are the main problems in context aware prediction and recommender systems especially for the ones based on the user history. It is considerably possible to encounter a new context in such kind of systems. As an example, user may not have appeared in location X before or similarly user may have not been with person Y before. In such cases, there is no data in the context history of the user related with these new context data and so it is impossible to exactly match the current context information with the ones in user history.

In some cases the data that exactly matches with the current context of the user may be insufficient and making a proper prediction for the activity preference of the user may not be possible. The prediction based on large dataset returns the results that have higher accuracy. As a result, the deficiency of user context history causes appropriateness problems in the prediction based on user previous context information.

In order to prevent and overcome these problems, our study uses a set of ontologies that model the context information. These ontologies have many defined concepts and classes and enable the machines to relate information in one concept to the other one. As a result, the defined ontologies support the prediction by enriching the extracted data. Moreover, they also bring a solution for the situations of encountering new context data by using a more generalized concept which is related with the concept of context data.

The ontology hierarchies are given in the following sub sections. For all defined ontology structures, above entities correspond to more general concepts and below entities are more specific concepts.

3.4.1. Ontology for Location

This ontology describes the location of a user and has two levels. Sample location ontology is illustrated in Figure 1. It is divided into two main concepts which are “Shopping Center” and “Outdoor”. As an example, Cepa which is the instance of the “Shopping Center” is a shopping center.

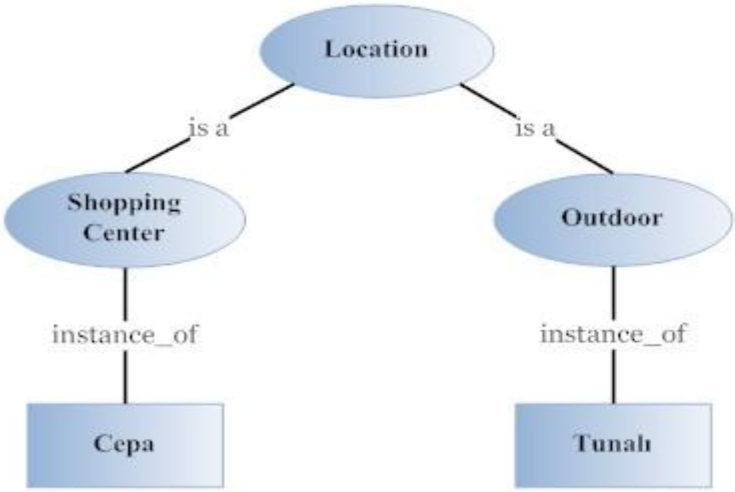

Figure 1: Sample Hierarchy for Location ontology

3.4.2. Ontology for Nearby Person of the User

Figure 2 shows a sample ontology hierarchy for the person accompanying the user. It is divided into six main concepts which are “Parents”, “Child”, “Partner/Spouse”, “Family Member”, “Friend” and “Colleague”. The instances of these sub concepts can be any person. As an example, according to Figure 2 Kate is one of the friends of the user and then she is labeled an instance under the sub concept of “Friend” for Mark.

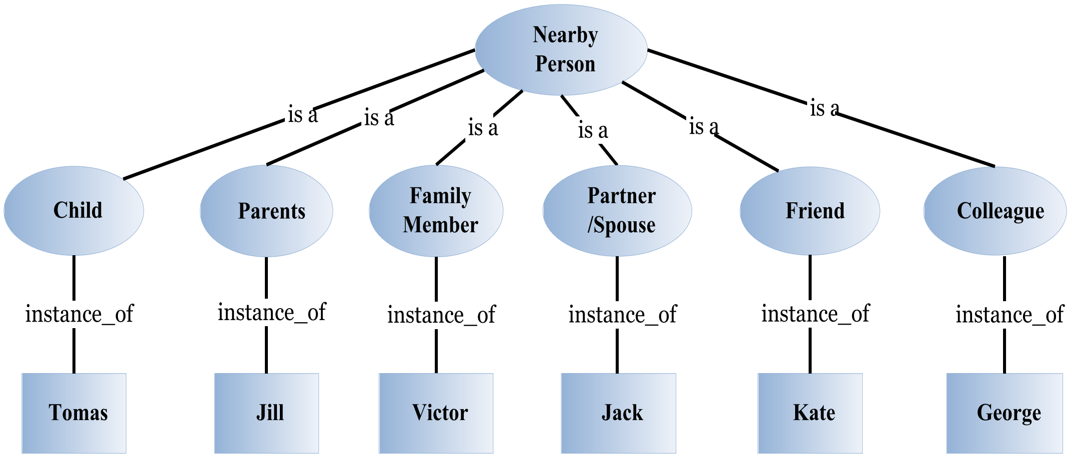

Figure 2: Sample Hierarchy for Person ontology

3.4.3. Ontology for Time Frame of the Day

Ontology hierarchy for the time ontology is given in Figure 3. It is divided into four main concepts which are the general time frames. These are “Morning”, “Afternoon”, “Evening” and “Night”. Since two hour time periods are used and the time between 08.00 and 02.00 are considered for the prediction in this study, the ontology for this context is developed accordingly.

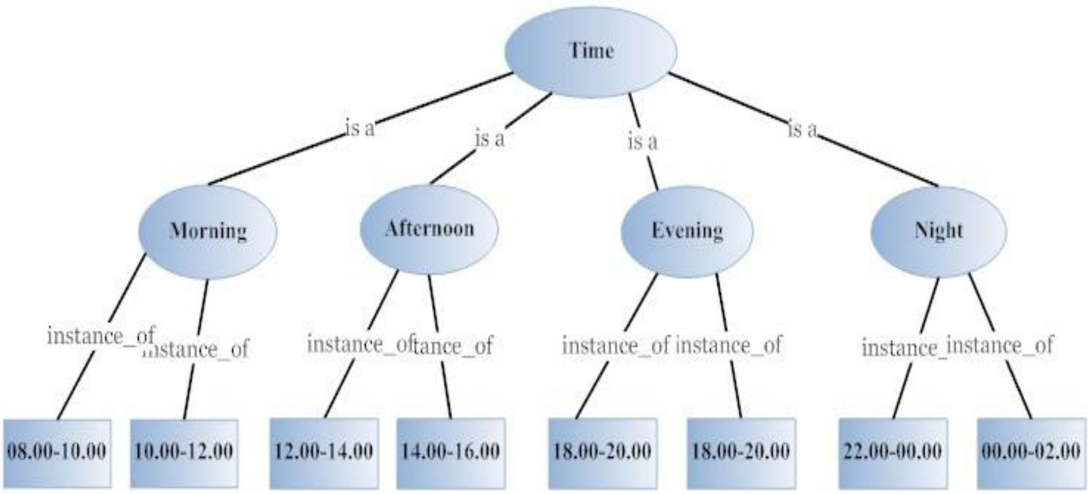

Figure 3: Hierarchy for Time frame ontology

3.4.4. Ontology for Day of Week

Ontology hierarchy for days is defined as given in Figure 4. It is divided into two main concepts: “Weekday” and “Weekend”. Users’ activity preferences can change according to days and people usually have similar preferences for the weekends that are generally holiday or for the weekdays that are routine working days for most of the users.

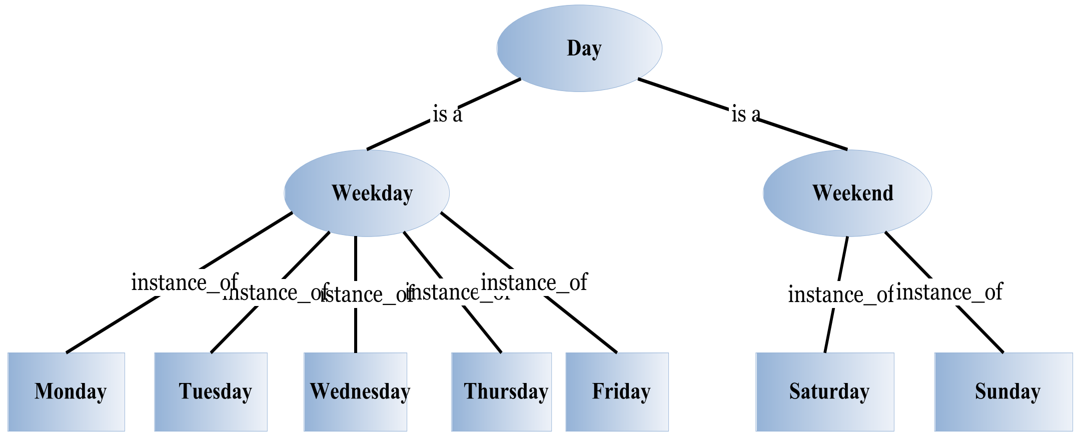

Figure 4: Hierarchy for Day ontology

3.4.5. Ontology Hierarchy for Weather

Ontology hierarchy for weather conditions is defined as Figure 5. Since the sunny and clear weather usually has similar effects on the people and similarly cloudy, rainy and snowy usually affect the people’s mood in similar way, the ontology specifies the weather context in two sub concepts as: “clear/sunny” and “cloudy/rainy/snowy”.

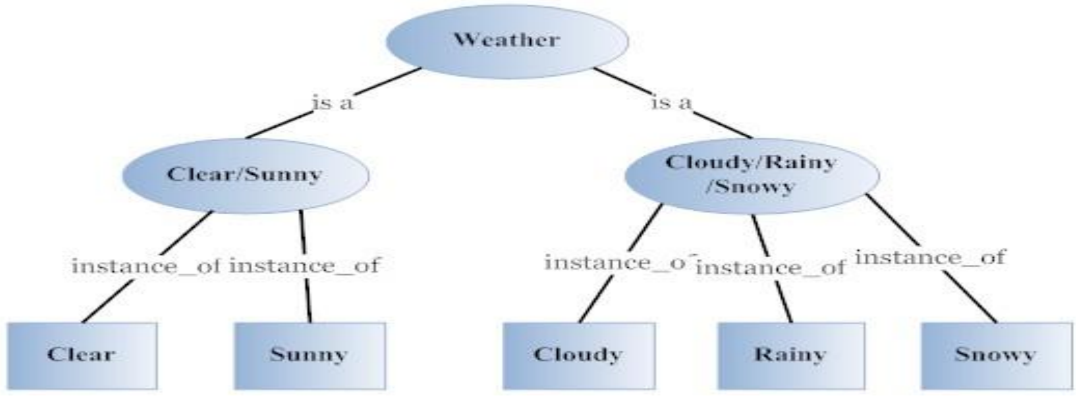

Figure 5: Hierarchy for Weather ontology

3.4.6. Ontology Hierarchy for Available Activities of Locations

Activity is the target context in this study. Each location may be appropriate for certain activities for the users. Therefore, the information of available activities in each location is kept and as an ontology. Thus, sample ontology hierarchy that shows the relations among the location context and activity context is defined as given in Figure 6. The available activities for each location is kept and provided by this ontology. As an example, the activity cinema is available in the locations of Cepa and Kentpark according to defined sample ontology in Figure 6.

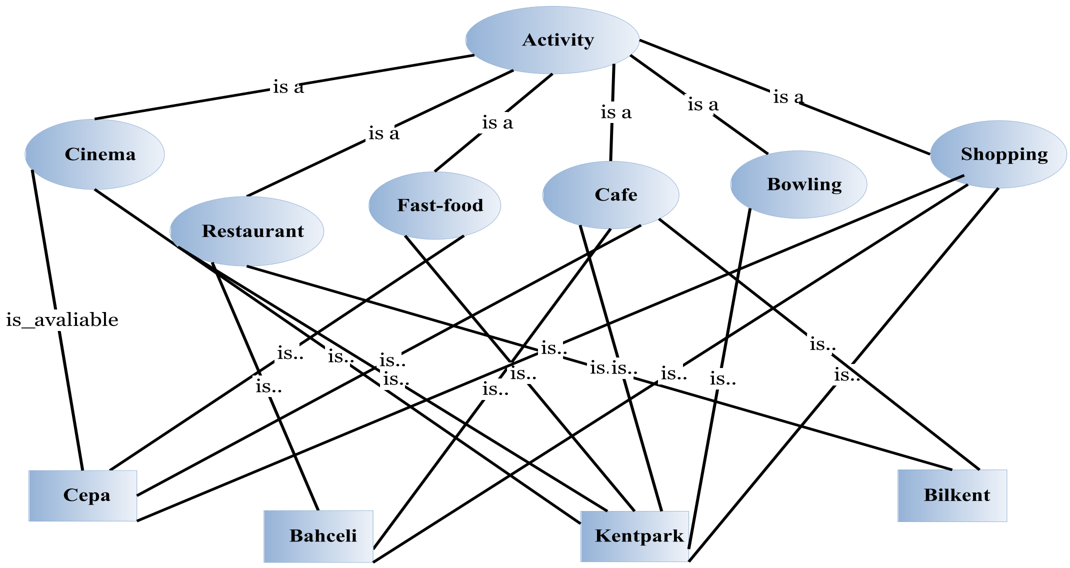

Figure 6: Sample Hierarchy for Available Activities of Locations

3.5. Activity Prediction Model and Related Algorithms

The algorithms used in our prediction model can be grouped under two main titles which are activity management and activity prediction. Table 2 lists the related concepts introduced in the previous sections and the corresponding acronyms used in this section.

Table 2: Notations Used in the Activity Prediction Model

Notations	Descriptions
IT	interval time
IOT	interoccurrence time
$\widetilde{\mathbf{IOT}}$	estimated interoccurrence time
$\widetilde{\mathbf{Var}}$	estimated variance
$\tilde{\boldsymbol{\sigma}}$	estimated standard deviation
$\tilde{\mathbf{D}}$	estimated D
CT	activity count threshold
PT	prediction threshold
AC	activity appearance count
DC	disable activity appearance count

3.5.1. Activity Management

The algorithms presented in this section are used to manage the statuses of each activity and forms the discard strategy for activity prediction. Changed or interrupted activities of any user are detected by the means of these algorithms and discarded activities are not included in the prediction process.

3.5.1.1. Management of Activity Status Algorithm

People perform the activities in specific time intervals and they can also interrupt or leave an activity preference in seasonal or periodic patterns. Therefore, adding these activities into the prediction process may cause inappropriate results in the prediction. Furthermore, if an activity occurs very frequently before it is interrupted, then the model may predict this activity incorrectly most of the time although the activity does not occur for a long period of time.

In order to solve this problem, our model proposes a status mechanism for the activities. With this mechanism, each activity has a status value that can be 'Enabled' or 'Disabled' and for every user, the status of each activity changes according to frequency of the activity's occurrence. 'Enabled' activities are included in the prediction process, whereas 'Disabled' are neglected in the prediction process until their statuses turn to 'Enabled' again. As a result, changing activity preferences of the users in time are recognized with this method.

This mechanism is based on the $\widetilde{IOT}$ and the $\widetilde{Var}$ of the IOT values of the activities. Our model keeps these values up to date and when an activity occurs, these values of this activity are updated according to equations that were explained in section 3.2.

'Enabled' activities are checked in daily manner and if any activity cannot fulfill the 'Enabled' condition, then the activity's status is changed to 'Disabled' and it is discarded from the prediction process until its status turns to 'Enabled' again.

The algorithm for occurred activity is used to update the $\widetilde{IOT}$ and $\widetilde{Var}$ values of occurred activity. Moreover, if the activity is 'Disabled', it checks include conditions to change the activity status to 'Enabled' and if they are fulfilled, changes the activity's status accordingly. The algorithm for occurred activity status update is given below.

```
BEGIN
   Calculate new IÕT and new Ṽar for IOT values
   IF activity is Disabled
      IF Current IOT < IÕT + σ̃ for IOT values
         SET activity status to Enabled
      ELSE IF disable activity count of the activity > CT
         IF IÕT + σ̃ < PT
             SET activity status to Enabled
         END IF
      ELSE
          SET DC to DC + 1
      END IF
   END IF
   UPDATE IÕT and Ṽar values
   SET AC to AC + 1
   SAVE data to the context history of the user
END
```

As it can be seen from the algorithm, when an activity occurs, new values of $\widetilde{IOT}$ and $\widetilde{Var}$ for the activity are recalculated and these values are updated.

When an activity's status is 'Enabled' but it does not occur today, discard conditions must be checked to change the activity status to 'Disabled'. The algorithm used to check 'Disabled' status is given below.

```
BEGIN
   IF activity is Enabled
       IF activity count > CT
          EXECUTE discard procedure
          IF the return is true
             SET the activity status to Disabled
          END IF
       END IF
   END IF
END
```

3.5.1.2. Activity Discard Procedure

$\widetilde{IOT}$ of each activity is recorded for the users and it is updated when the activity occurs. The difference between the current IT of the activity and $\widetilde{IOT}$ value of it indicates the distance of current IT to $\widetilde{IOT}$ of the activity. Furthermore, $\widetilde{\sigma}$ of IOT values for each activity shows the fluctuations of IOT values of that activity for the user. In order to detect the interrupted activities of any user, it is suitable to compare the difference between the current IT of the activity and $\widetilde{IOT}$ value of it and the $\widetilde{\sigma}$ of IOT values of each activity.

However, if an activity has not occurred for a specific user for a long time, the difference between the current IT of the activity and $\widetilde{IOT}$ value of it becomes considerably high. According to three-sigma rule (Pukelsheim, 1994), for a normally distributed data set, almost all (99.73%) of the samples lie within 3 standard deviations of the mean. If we customize this statement to our study, nearly all interocurrence times lie within the three $\widetilde{\sigma}$ of IOT values and so if the difference between the current IT of the activity and $\widetilde{IOT}$ value of it is higher than three $\widetilde{\sigma}$ of IOT value, then this activity is likely interrupted or left from the user. As a result, discard condition is defined as:

$$\mathrm{IT} - \widetilde{\mathrm{IOT}} > 3 * \tilde{\sigma} \text{ of IOT values}$$

Thus, the following is the corresponding discard procedure algorithm.

```
BEGIN
    IF (IT- IÕT) > 3 * σ̃ of IOT values
        Return true
    ELSE
        Return false
    END IF
END
```

As seen it can be seen, if difference between the current interval time of the activity and $\widetilde{\mathrm{IOT}}$ value of it exceeds three $\tilde{\sigma}$ of IOT values of the activity, then the activity is discarded and its status is changed to 'Disabled'.

For example, suppose that activity A is 'Enabled' activity for a specific user and has the following statics.

Last Occurrence Date	$\widetilde{\mathrm{IOT}}$ (days)	$\tilde{\sigma}$ of IOT values
15.07.2011	13.37	1.83

Suppose that today is 03.08.2011 and the algorithm checks each activity to manage the statuses of them. For activity A, the following values are obtained.

IT (days)	IT -$\widetilde{\mathrm{IOT}}$ (days)	$3 * \tilde{\sigma}$ of IOT values
19	5.63	5.49

Since difference between the IT and $\widetilde{\mathrm{IOT}}$ value of the activity exceeds three $\tilde{\sigma}$ of IOT values, then the activity A is discarded and its status is changed to 'Disabled'.

3.5.1.3. Activity Include Procedure

People can start doing some interrupted activities again, so it is essential to include these 'Disabled' activities into the prediction process. The most important issue in activity include procedure is to determine the time when 'Disabled' activity should be included in the prediction again.

The basic condition for the activities to change their status from 'Disabled' to 'Enabled' is the occurrence. In other words, if a 'Disabled' activity occurs, it has the chance to turn its status to 'Enabled'. The first time when the activity occurs after its status changed from 'Enabled' to 'Disabled', the intercurrence time of this occurrence of the activity is not added to $\widetilde{\mathrm{IOT}}$ of the activity. The reason behind this is that this current IOT value of the activity is too high and if it is added to $\widetilde{\mathrm{IOT}}$ value, it may change this value considerably. Actually this IOT value is the time between the interrupted and restarted dates of the activity. Because of this reason, the first IOT of the 'Disabled' activity is ignored and not added to $\widetilde{\mathrm{IOT}}$ value. As a result, only after the first time, if the 'Disabled' activity continues to occur, it may be included but there are some other conditions.

When the 'Disabled' activity continues to occur after its first occurrence and the current IOT value suits its previous 'Enabled' occurrence statistics, the first condition is applied. The first condition is that current IOT for the activity must be less than the sum of the $\widetilde{\mathrm{IOT}}$ and the $\tilde{\sigma}$ of IOT values. The idea behind this condition is that including the activities having similar occurrence manner with their previous 'Enabled' occurrence statistics in the prediction process as soon as possible. This condition is defined as the following equation:

$$\text{Current IOT} < \widetilde{\mathrm{IOT}} + \tilde{\sigma} \text{ of IOT values}$$

As shown in the equation, if current IOT is less than the sum of the $\widetilde{\mathrm{IOT}}$ and the $\tilde{\sigma}$ of IOT values, then the activity's status is changed to 'Enabled' and it is included in the next prediction processes.

As an example, assume that activity A is 'Disabled' for a specific user and it occurs after its status is changed to 'Disabled'. When this activity occurs again on 26.05.2011 and the related algorithm checks the status of activity A, if it can be changed to 'Enabled'. Activity A has the following statistics at that time.

Last Occurrence Date	$\widetilde{\mathrm{IOT}}$ (days)	$\tilde{\sigma}$ of IOT values (days)
06.05.2011	16.80	4.00

The current IOT of the activity can be calculated as;

(26 May)-(06 May) = 20 days

The current IOT of activity A is less than the sum of the $\widetilde{\mathrm{IOT}}$ and the $\tilde{\sigma}$ of IOT values of the same activity, therefore the activity's status is changed to 'Enabled' and it is included in the next prediction processes.

However, activity may not be in similar occurrence manner with the one when its status is 'Enabled', after its first occurrence while it is 'Disabled'. If the activity continues to occur, it means that the user restarts this activity but this time, the activity has a different occurrence manner as different from the one when it is 'Enabled'. In this case, its current IOT values naturally exceed the sum of the $\widetilde{\mathrm{IOT}}$ and the $\tilde{\sigma}$ of IOT values and so it is not possible for the activity to be included in the prediction process. To solve this problem, we propose a second condition that is if the disable activity count exceeds CT, the sum of new $\widetilde{\mathrm{IOT}}$ and new $\tilde{\sigma}$ of IOT values must be less than PT. This condition is defined as the following equation:

$$\text{new } \widetilde{\mathrm{IOT}} + \text{new } \tilde{\sigma} \text{ of IOT values} < \mathrm{PT}$$

As shown in the equation, if the sum of new $\widetilde{\mathrm{IOT}}$ and new $\tilde{\sigma}$ of IOT values is less than PT, then the activity's status is changed to 'Enabled' and it is included in the next prediction. Exceeding CT is defined as a precondition, since if the activity has a different occurrence manner with the one when its status is 'Enabled', it is expected that the activity must occur 'activity count threshold' (CT) times at least.

As an example, assume that activity A is 'Disabled' for a specific user and it occurs after its status is changed to 'Disabled'. When this activity occurs on 04.03.2011 and the related algorithm checks the status of activity A, if it can be changed to 'Enabled'. Assume that the first include condition is not fulfilled and disable activity count of the activity exceeds CT and let PT is 28 days for this scenario. Then, new $\widetilde{\mathrm{IOT}}$ and $\tilde{\sigma}$ of IOT values are calculated for the activity A and below values are obtained.

$\widetilde{\mathrm{IOT}}$ (days)	$\tilde{\sigma}$ of IOT values (days)	PT(days)
21.48	5.34	28

As seen in the above values, since the sum of the $\widetilde{IOT}$ and the $\widetilde{\sigma}$ of IOT values of the activity A is less than PT, then activity's status is changed to 'Enabled' and it is included in the next prediction processes.

3.5.1.4. Summary of Activity Management

An activity discard procedure is specified to disregard the activities of the user which have not been performed by the user for a long time. If difference between the current interval time of the activity and $\widetilde{IOT}$ value of it exceeds three $\widetilde{\sigma}$ of IOT values of the activity, then the activity is discarded and its status is changed to 'Disabled'. Hence, the 'Disabled' activity is not included the prediction until its status turns to 'Enabled' again.

We define two include conditions for the 'Disabled' activities to change their statuses to 'Enabled'. If the first condition is fulfilled by any activity, the second condition is not checked. If the first condition is not fulfilled by the activity, then the precondition of the second condition is checked and if it is fulfilled by the activity, then the second condition is checked. Finally, if a 'Disabled' activity suits any one of these conditions, its status is changed to 'Enabled', otherwise it remains 'Disabled' for the next prediction processes.

With this mechanism and algorithms, activities are managed and they are included into the prediction process according to their statuses. In summary, seasonal or periodic patterns and rarely and arbitrary performed activities are recognized for the users and their statuses are set as 'Disabled' and they are not included in the prediction process by this mechanism to improve the prediction quality.

3.5.2. Activity Prediction

This process is performed to predict the next activity preference of the user based on user's context history and Figure 7 illustrates this process.

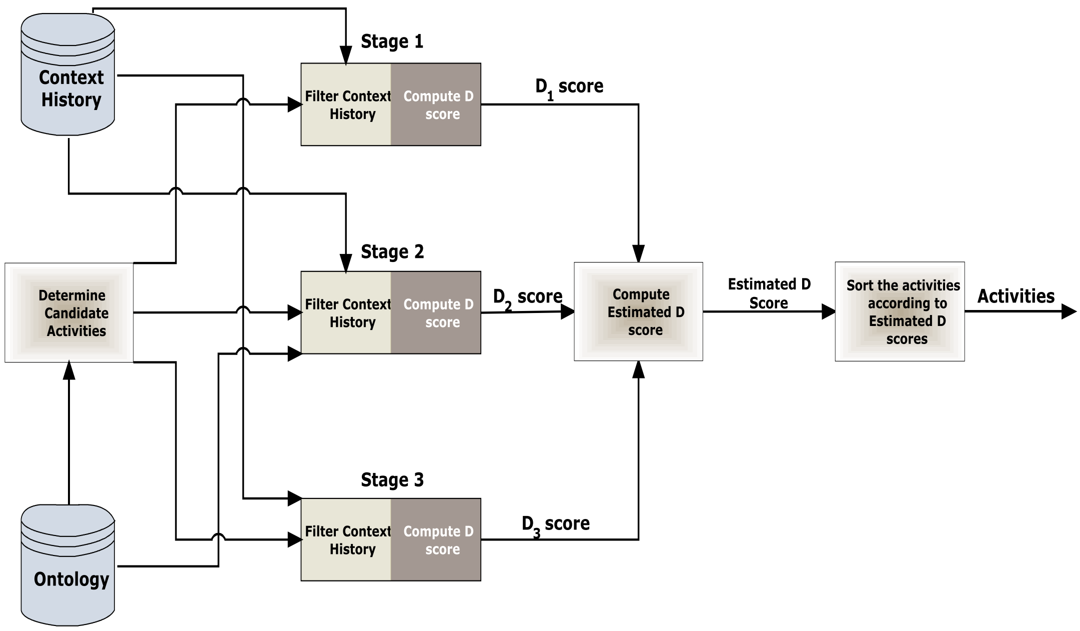

Figure 7: Activity Prediction Stages

At the beginning of the prediction process, the activities that are available in the current location context of the user are determined according to related ontology hierarchy. Moreover, the activities are selected from these specified activities which are 'Enabled' for that user and fulfill the conditions of PT and CT. Therefore, only those selected activities are included in the prediction.

As shown in Figure 7, this process consists of three stages for each activity and each stage consists of two steps. The first step in each stage is the filtering suitable context data from user context history and updating statistics of the selected activity for the returned entries. The second step is the computation of individual D score for the selected activity. The first stage applies to all historical context entries exactly matching with the current context for the selected activity. The second stage applies to all historical context entries that has similar context with the current context of the user for the selected activity. This stage uses the ontology for each context dimension to specify the similar contexts with the current context. The last stage applies to all historical data without considering the context information for the selected activity.

After the individual D scores are computed in each stage for the activity, an estimated D score is computed for the activity by using the D scores of each stage. The estimated D score is computed for each selected activity and then the activities are sorted according to their estimated D scores as explained in the following sections. Finally, the activity with the lowest score is predicted as the next most probable activity of the user.

3.5.2.1. Filtering User Context History

The first step in each stage is searching and filtering the data from the user context history matching to the current context of the user and the selected activity.

As mentioned previously matching consists of three stages. Firstly, the data exactly matching with the current context information of the user is extracted from the context history. However the extracted historical data with the exact matching may not be enough that could be used for the activity prediction or there is no extracted historical data at the end of the first stage. In such cases it is difficult to make good prediction with insufficient data. Therefore, ontology is used in the second stage. Ontology is used for each context dimension to incorporate more data in the prediction process. At the end of this stage the accuracy of the prediction is improved, since more data is collected from the context history than the first step. The ontology usage in the second stage will be described in the next sub section.

Finally, all the data that matches with the specified activities is extracted in the last stage. In this step, the current context data of the user is not used in the filtering. This step aims to get all the data that is independent of the context information for the specified activities. The reason behind this step is that some or all of the activities are independent of some of the context information for some users, because of that all the data is retrieved for the specified activities. Furthermore, all the extracted data is reprocessed during the filtering stage and this process will be explained as a sub section of this section.

3.5.2.1.1. Ontology Usage

Any current context may not match with the entries in the user's context history in the first stage. In that case, the exact matching is not possible as there is no data available for the prediction process. Moreover, extracted historical data matching with the current contexts of the user may not be enough to make proper prediction. As a result, ontology is used in the second stage to overcome these problems.

Ontology for each context dimension is applied by using a more generalized concept which is related to the concept of actual context data. As an example, assume that the person with the user is *Mark*. Other contexts of the user match with some data in the user history but there is no related data that matching with this context and so exact matching is not possible in the first step of the filtering stage. However, it is known that *Mark* is one of the friends of the user according to ontology structure of nearby person and generalization relations are searched to find a general concept that can be used in the matching process. Finally, as this instance belongs to "Friend" concept of the user and as a result of this; the data with the friends in user history is used in a more general way by the means of the ontology.

3.5.2.1.2. Data Preprocessing

The most important issue in our prediction model is the calculation of $\widetilde{\mathrm{IOT}}$ and $\widetilde{\mathrm{Var}}$ of IOT values for the filtered data of each activity in each stage. Since the prediction algorithm is based on the distance scores of each activity and also distance scores are calculated based on the $\widetilde{\mathrm{IOT}}$ and $\widetilde{\mathrm{Var}}$ of IOT values, these values are calculated for each selected activity of the user from the context history. Furthermore, they are updated after the first occurrence of the activity and this calculation is repeated for all selected the data of each activity that ordered according to occurrence date. Thus, each specified activity has an $\widetilde{\mathrm{IOT}}$ and an $\widetilde{\mathrm{Var}}$ of IOT values at the end of the filtering stage.

3.5.2.2. Prediction Algorithm

After filtering user context data from the user context history and calculating the $\widetilde{\mathrm{IOT}}$ and an $\widetilde{\mathrm{Var}}$ of IOT values for each specified activity in each stage, the IOT values are normalized and D score for that activity in that stage is computed. The formula for the computation of D score was given in the section 3.3. This process is repeated for each stage. Since context history records used in each stage is different, the computed D scores for each activity are also different at the end of each stage. As a result of this, in order to get a typical D score, we use an estimated D ($\widetilde{\mathrm{D}}$) score which is the weighted average of the all three D scores obtained in every stage. The formula for the estimated D score is:

$$\widetilde{\mathrm{D}} = (\beta * \mathrm{D}_1) + (\gamma * \mathrm{D}_2) + ((1 - \beta - \gamma) * \mathrm{D}_3)$$

In the formula, D_1 is the D score of the first stage one, D_2 is D score of second stage and D_3 is the D score of the last stage. The weighting factors (β) and (γ) are constant smoothing factors with values between 0 and 1 and also their sum must be between 0 and 1.

The presented formula is the general formula for $\widetilde{D}$ score for each activity and the formula can only be used, if more than one record is returned from filtering in each stage. Otherwise the stages that return less than two records are assigned to worst case D score of 3 and this score is multiplied by the corresponding weighting factor. For example, suppose that for the activity A only second and third stages have more than one records and their individual D scores are calculated. However, the first stage in which exact matching is done does not satisfy this condition, and then the D (D_1) score for that stage is set to 3 and $\widetilde{D}$ score for the activity A is calculated as:

$$\widetilde{D} = 3 * \beta + (\gamma * D_2) + ((1 - \beta - \gamma) * D_3)$$

Since D score of 3 is considered as the worst case in this study, the D scores in each stage that exceed 3are set to 3 in the calculation of $\widetilde{D}$ score for each activity.

$\widetilde{D}$ score is computed for each selected activity of the user and $\widetilde{D}$ scores of all the selected activities are listed in ascending order. Since $\widetilde{D}$ score indicates the distance of the current IT to the $\widetilde{IOT}$ value for each activity, low $\widetilde{D}$ scores are highly probable to occur as the user's next activity. In other words, lower $\widetilde{D}$ scores which show a close distance between current IT and the $\widetilde{IOT}$ value denote higher possibility of occurrence for the user. As a result of this, the activity that has the lowest $\widetilde{D}$ score is the highest possibility to be done by the user as a next activity and so it is recommended for the user's next activity.

3.5.2.3. Summary of Activity Prediction

The all activity prediction process can be summarized with the following algorithm.

```
BEGIN
Determine possible activities
FOR each possible activity
  FOR each stage
      Query the target input in the user context history
      WHILE data is extracted from the user context history that matches
      stage's criteria
          Calculate the new IÕT and Ṽar values of the selected activity
       END WHILE
           IF (the count of extracted data >= 2)
                Calculate IT
                Calculate the σ̃ of IOT values
                Compute D score
          ELSE
                SET D score to 3
          END IF
  END FOR
  Compute D̃ score for the activity using stages' D scores
END FOR
List the activities according to D̃ scores in ascending order
Return the sorted activities that may be the user's next activity
END
```

Before applying the prediction algorithm, the sum of $\widetilde{IOT}$ and the $\tilde{\sigma}$ of each activity are compared with the general PT and the activities whose sum exceeds PT are not included in the prediction process. Moreover, only the activities which are 'Enabled' and exceed CT are included in the prediction process. As a result of these, seasonal or periodic patterns and rarely and arbitrary performed activities are detected for the users and they are not included in the prediction. These controls in the algorithm improve the accuracy of the prediction and fulfill the users' expectancies.

Our activity prediction algorithm mainly consists of three stages and the idea behind using multiple stages is to use more data to improve prediction quality. $\tilde{D}$ score is the weighted average of obtained D scores in each stage and indicates distance of current IT of each activity to its $\widetilde{IOT}$ on that date. These scores show the occurrence possibility of each activity for the users. Less distance means that it is highly probable that this activity occurs for the next activity of the user. In conclusion, the users' next activity according to the current contexts of them can be high accurately predicted with our prediction method and algorithm.

3.6. Typical Scenarios

In this section, typical scenarios for our model are exemplified to show the application of the proposed algorithms in the model. The scenarios are presented in two sub sections. First, the scenarios related to the activity management are explained and then the scenarios that exemplify the activity prediction according to current contexts of the user are presented. For all scenarios of both activity management and activity prediction, the parameters given in Table 3 are used in the algorithms.

Table 3: The Values of the Variables in all the Scenarios

Variable	Value
CT	5
PT	35
α	0.3
λ	0.3
β	0.6
γ	0.25

3.6.1. Sample Scenarios for Activity Management

This section describes four basic scenarios that show how the system manages the activities and context history of the typical user.

For all four scenarios in this section, suppose that it is 12.45 on Monday and Gaye is eating his lunch at a fast food restaurant in Cepa shopping center and all the current context information for Gaye is given in the Table 4.

Table 4: Current Contexts for the Scenarios of Activity Management

Date	Day	Time	Location	Weather	Nearby Person	Activity
04.07.2011	Monday	12.00-14.00	Cepa	Sunny	Aslı	Fast-food

The variables of the activities that do not occur at that day are the same for all the scenarios and are given in Table 5.

Table 5: Variables of the Activities that does not occur for the Scenarios of Activity Management

Activity	Status	AC	DC	Last Occurrence	$\widetilde{IOT}$	$\widetilde{Var}$
Restaurant	Enabled	38	0	29.06.2011	6.601	1.880
Cafe	Enabled	21	0	25.06.2011	10.246	4.704
Cinema	Enabled	13	0	12.06.2011	12.328	8.404

Scenario 1

Suppose that the variables for the occurred activity are given in the Table 6 until that time.

Table 6: Occurred Activity Variables for the 1st Scenario of Activity Management

Activity	Status	AC	DC	Last Occurrence	$\widetilde{IOT}$	$\widetilde{Var}$
Fast-food	Enabled	29	0	27.06.2011	8.561	0.934

Since the occurred activity, fast-food is an Enabled activity, its new $\widetilde{IOT}$ and $\widetilde{Var}$ values are calculated and updated as:

Current IOT for the fast food activity = Current Occurrence Date – Last Occurrence Date

= 04.07.2011 - 27.06.2011 = 7 days

New $\widetilde{IOT} = (0.3 * \text{Current IOT}) + ((1-0.3) * \widetilde{IOT}_{n-1})$

= (0.3*7) + (0.7*8.561) = 8.093 days

New $\widetilde{Var} = ((|\text{Current IOT} - \text{Current } \widetilde{IOT}|^2) * 0.3) + (0.7 * \widetilde{Var}_{n-1})$

$= ((|7 - 8.093|^2) * 0.3) + (0.7 * 0.934) = 1.012$

Moreover Discard procedure is performed for the enabled activities of the user that does not occur on that day. If any activity that suits for the discard condition, then this activity is discarded and its status is changed to 'Disabled'.

For the activity of Restaurant:

Interval Time (IT) = Today's Date – Last Occurrence Date

= 04.07.2011 - 29.06.2011 = 5 days

The discard condition is: $(\text{IT}-\widetilde{IOT}) > 3 * \tilde{\sigma}$ of IOT values

$\text{IT}-\widetilde{IOT} = 5 - 6.601 = -1.601$

$3 * \tilde{\sigma}$ of IOT values $= 3* \sqrt{1.880} = 4.113$

Since -1.601 < 4.113, then the activity does not satisfy the discard condition and it is not discarded.

For the activity of Cafe:

IT = 04.07.2011 - 25.06.2011 = 9 days

IT– $\widetilde{IOT}$ = 9 – 10.246= -1.246

$3 * \tilde{\sigma}$ of IOT values $= 3*\sqrt{4.704} = 6.506$

Since -1.246< 6.506, then the activity does not satisfy the discard condition and it is not discarded.

For the activity of Cinema:

IT = 04.07.2011 - 12.06.2011 = 22 days

IT– $\widetilde{IOT}$ = 22 – 12.328= 9.672

$3 * \tilde{\sigma}$ of IOT values $= 3*\sqrt{8.404} = 8.697$

Since 9.672> 8.697, the activity satisfies the discard condition, and then activity is discarded. That is, its status is set to “Disabled”.

Finally, the context history of the user is updated by adding the context data of occurred activity and the activity variables of the user are updated as Table 7.

Table 7: Updated Activity Variables for the 1st Scenario of Activity Management

Activity	Status	AC	DC	Last Occurrence	$\widetilde{IOT}$	$\widetilde{Var}$
Fast-food	Enabled	**30**	0	**04.07.2011**	**8.093**	**1.012**
Restaurant	Enabled	38	0	29.06.2011	6.601	1.880
Cafe	Enabled	21	0	25.06.2011	10.246	4.704
Cinema	**Disabled**	13	0	12.06.2011	15.261	8.404

As seen in Table 7, changing variables are written in bold italics. Since fast-food activity occurs in the scenario, its variables change. Moreover, the status of cinema activity changes from Enabled to Disabled, because this activity is discarded at the end of the scenario.

Scenario 2

Suppose that the variables for the occurred activity are given in the Table 8 until that time.

Table 8: Occurred Activity Variables for the 2nd Scenario of Activity Management

Activity	Status	AC	DC	Last Occurrence	$\widetilde{IOT}$	$\widetilde{Var}$
Fast-food	Disabled	38	3	25.06.2011	8.764	2.409

Since the occurred activity, fast-food is a Disabled activity; it has the chance to turn its status to Enabled. The disable count (DC) of the activity does not exceed the general CT which is 5, and so the first condition of the include condition is checked for the activity.

The first condition is that current IOT for the activity must be less than the sum of $\widetilde{IOT}$ and $\tilde{\sigma}$ of IOT values.

Current IOT for the fast food activity = Current Occurrence Date – Last Occurrence Date

= 04.07.2011 - 25.06.2011 = 9 days

The include condition is: $\text{Current IOT} < \widetilde{IOT} + \tilde{\sigma}$ of IOT values

$\tilde{\sigma}$ of IOT values = $\sqrt{\widetilde{Var}} = \sqrt{2.409} = 1.552$

$\widetilde{IOT} + \tilde{\sigma}$ of IOT values = 8.764 + 1.552 = 10.316

Since 9 < 10.316, the activity satisfies the include condition, and then the activity's status is changed to Enabled.

Since fast-food activity occurs, its new $\widetilde{IOT}$ and $\tilde{\sigma}$ of IOT values are calculated and updated for the user as:

New $\widetilde{IOT} = (0.3 * \text{Current IOT}) + ((1 - 0.3) * \widetilde{IOT}_{n-1})$

= (0.3*9) + (0.7*8.764) = 8.835 days

New $\widetilde{Var} = ((|\text{Current IOT} - \text{Current } \widetilde{IOT}|^2) * 0.3) + (0.7 * \widetilde{Var}_{n-1})$

= ((|9 - 8.835 |^2) * 0.3) + (0.7 * 2.409) = 1.736

Since the variables of the activities that do not occur are the same, same calculations are made for these activities.

Thus, the context history of the user is updated by adding the context data of occurred activity and the activity variables of the user are updated as Table 9.

Table 9: Updated Activity Variables for the 2nd Scenario of Activity Management

Activity	Status	AC	DC	Last Occurrence	$\widetilde{\text{IOT}}$	$\widetilde{\text{Var}}$
Fast-food	**Enabled**	**39**	**0**	**04.07.2011**	**8.835**	**1.736**
Restaurant	Enabled	38	0	29.06.2011	6.601	1.880
Cafe	Enabled	21	0	25.06.2011	10.246	4.704
Cinema	**Disabled**	13	0	12.06.2011	15.261	8.404

As shown in Table 9, since fast-food activity occurs in the scenario, its variables change and also it status changes from Disabled to Enabled, because it suits the include condition. On the other hand, the status of cinema activity changes from Enabled to Disabled, because this activity is discarded at the end of the scenario.

Scenario 3

Suppose that the variables for the occurred activity are given in the Table 10 until that time.

Table 10: Occurred Activity Variables for the 3rd Scenario of Activity Management

Activity	Status	AC	DC	Last Occurrence	$\widetilde{\text{IOT}}$	$\widetilde{\text{Var}}$
Fast-food	Disabled	53	8	16.06.2011	9.2	2.085

Since the occurred activity, fast-food is a Disabled activity; it has the chance to turn its status to Enabled. Since the disable count (DC) of the activity exceeds the general CT which is 5, the activity has a different occurrence manner as different from the one when it is 'Enabled'. In that case, the second condition of the include condition is checked for the activity.

The second condition is that the sum of new $\widetilde{\text{IOT}}$ and new $\tilde{\sigma}$ of IOT values must be less than general PT.

Current IOT for the fast food activity = Current Occurrence Date – Last Occurrence Date

= 04.07.2011 - 16.06.2011 = 18 days

New $\widetilde{\text{IOT}} = (0.3 \ * \ \text{Current IOT}) + \ ((1 - 0.3) * \ \widetilde{\text{IOT}}_{n-1})$

= (0.3*18) + (0.7*9.2) = 11.84 days

New $\widetilde{\text{Var}} = ((|\ \text{Current IOT} - \text{Current } \widetilde{\text{IOT}}\ |^2) * 0.3) + (0.7 * \widetilde{\text{Var}}_{n-1})$

$$= ((\ |18 - 11.84|^2\ ) * 0.3\ ) + (\ 0.7 * 2.085) = 12.842$$

The include condition is: new $\widetilde{IOT}$ + new $\tilde{\sigma}$ of IOT values < PT

new $\tilde{\sigma}$ of IOT values $= \sqrt{\text{new } \widetilde{Var}} = \sqrt{12.842} = 3.854$

new $\widetilde{IOT}$ + new $\tilde{\sigma}$ of IOT values = 11.84 + 3.854 = 15.694

PT is identified as 35 days for this study and Since 15.694 < 35, the activity suits the include condition, and then the activity's status is changed to Enabled.

Since the variables of the activities that do not occur are the same, same calculations are made for the activities that do not occur.

Hence, the context history of the user is updated by adding the context data of occurred activity and the activity variables of the user are updated as Table 11.

Table 11: Updated Activity Variables for 3rd Scenario of Activity Management

Activity	Status	AC	DC	Last Occurrence	$\widetilde{IOT}$	$\widetilde{Var}$
Fast-food	**Enabled**	**54**	**0**	**04.07.2011**	**11.840**	**12.842**
Restaurant	Enabled	38	0	29.06.2011	6.601	1.880
Cafe	Enabled	21	0	25.06.2011	10.246	4.704
Cinema	**Disabled**	13	0	12.06.2011	15.261	8.404

As shown in Table 11 and like in the previous scenario, since fast-food activity occurs in the scenario, its variables change and also its status changes from Disabled to Enabled, because it suits the include condition. On the other hand, the status of cinema activity changes from Enabled to Disabled, because this activity is discarded at the end of the scenario.

Scenario 4

Suppose that the variables for the occurred activity are given in the Table 12 until that time.

Table 12: Occurred Activity Variables for the 4th Scenario of Activity Management

Activity	Status	AC	DC	Last Occurrence	$\widetilde{IOT}$	$\widetilde{Var}$
Fast-food	Disabled	47	6	11.05.2011	41.33	237.95

Since the occurred activity, fast-food is a Disabled activity; it has the chance to turn its status to Enabled. Since the disable count (DC) of the activity exceeds the general CT which is 5, the activity has a different occurrence manner as different from the one when it is 'Enabled'. In that case, the second condition of the include condition is checked for the activity.

The second include condition is that the sum of new $\widetilde{\mathrm{IOT}}$ and new $\tilde{\sigma}$ of IOT values must be less than general PT.

Current IOT for the fast food activity = Current Occurrence Date – Last Occurrence Date

= 04.07.2011 - 11.06.2011 = 54 days

New $\widetilde{\mathrm{IOT}}$ = $(0.3 * \text{Current IOT}) + ((1 - 0.3) * \widetilde{\mathrm{IOT}}_{n-1})$

= (0.3*54) + (0.7*41.33) = 45.13 days

New $\widetilde{\mathrm{Var}}$ = $((|\text{Current IOT} - \text{Current } \widetilde{\mathrm{IOT}}|^2) * 0.3) + (0.7 * \widetilde{\mathrm{Var}}_{n-1})$

= $((|54 - 45.13|^2) * 0.3) + (0.7 * 237.95) = 190.15$

The include condition is: new $\widetilde{\mathrm{IOT}}$ + new $\tilde{\sigma}$ of IOT values $<$ PT

new $\tilde{\sigma}$ of IOT values $= \sqrt{\text{new } \widetilde{\mathrm{Var}}} = \sqrt{190.15} = 13.79$

new $\widetilde{\mathrm{IOT}}$ + new $\tilde{\sigma}$ of IOT values = 45.13 + 13.79 = 58.82

PT is identified as 35 days for this study and Since 58.82 > 35, the activity does not suit the include condition, and the activity's status is not changed.

Since the variables of the activities that do not occur are the same, same calculations are made for the activities that do not occur.

Hence, the context history of the user is updated by adding the context data of occurred activity and the activity variables of the user are updated as Table 13.

Table 13: Updated Activity Variables for the 4th Scenario of Activity Management

Activity	Status	AC	DC	Last Occurrence	$\widetilde{IOT}$	$\widetilde{Var}$
Fast-food	Disabled	**48**	**7**	**04.07.2011**	**45.13**	**190.15**
Restaurant	Enabled	38	0	29.06.2011	6.601	1.880
Cafe	Enabled	21	0	25.06.2011	10.246	4.704
Cinema	**Disabled**	13	0	12.06.2011	15.261	8.404

As shown in Table 13, since fast-food activity occurs in the scenario, its variables change. The status of cinema activity also changes from Enabled to Disabled, because this activity is discarded at the end of the scenario.

3.6.2. Sample Scenario for Activity Prediction

This section demonstrates the activity prediction for a sample scenario. All the filtering stages of the prediction are detailed separately in this sample scenario.

Suppose that it is 16.30 on Saturday and Guven is in Cepa shopping center and he is free with his girlfriend. All the current context information for Guven is given in the Table 14.

Table 14: Current Contexts for the Activity Prediction Scenario

Date	Day	Time	Location	Weather	Nearby Person
30.04.2011	Saturday	16.00-18.00	Cepa	Rainy	Irmak

Sample context history of Guven is given as Table 15 until that time

Table 15: Sample from User Context History of Activity Prediction Scenario

Date	Day	Time	Location	Weather	Nearby Person	Activity
24.04.2011	Sunday	16.00-18.00	Cepa	Rainy	Irmak	Cafe
21.04.2011	Thursday	14.00-16.00	Bahceli	Cloudy	Serhat	Cafe
21.04.2011	Thursday	12.00-14.00	Bahceli	Cloudy	Serhat	Restaurant
….	….	….	….	….	….	….

As shown in Table 15, all activities that were performed by Guven are saved in this format and this context history for each user constitutes a basis for the prediction process.

Before prediction stages, the available activities are fast-food, restaurant, café, cinema and bowling in the location Cepa.

1st Filtering Stage

In the first stage, the context data from the user context history that matches with the current context of the user is filtered. Suppose that filtered data from context history of Guven according to current context information is obtained as shown in Table 16.

Table 16: Sample of Filtered data in the 1st Filtering Stage of Activity Prediction Scenario

Date	Day	Time	Location	Weather	Nearby Person	Activity
24.04.2011	Sunday	16.00-18.00	Cepa	Rainy	Irmak	Cafe
17.04.2011	Sunday	16.00-18.00	Cepa	Rainy	Irmak	Restaurant
02.04.2011	Sunday	16.00-18.00	Cepa	Rainy	Irmak	Cinema
….	….	….	….	….	….	….

Next, $\widetilde{\mathrm{IOT}}$ and $\widetilde{\mathrm{Var}}$ values are calculated for each activity and these values are updated for each data of the activity. Assume that data for the activity of cinema from the context history according to current context information of the user is as Table 17.

Table 17: Filtered data of cinema activity in the 1st Filtering Stage of Activity Prediction Scenario

Date	Day	Time	Location	Weather	Nearby Person	Activity
02.04.2011	Saturday	16.00-18.00	Cepa	Rainy	Irmak	Cinema
14.03.2011	Saturday	16.00-18.00	Cepa	Rainy	Irmak	Cinema
19.02.2011	Saturday	16.00-18.00	Cepa	Rainy	Irmak	Cinema
22.01.2011	Saturday	16.00-18.00	Cepa	Rainy	Irmak	Cinema

In order to calculate $\widetilde{IOT}$ and $\widetilde{Var}$ values for each activity of the selected data of the user, the data is sorted according to date and calculations are made in that order. According to Table 17, the data with date equal to 22.01.2011 is extracted first and since this is the first data, $\widetilde{IOT}$ and $\widetilde{Var}$ values cannot be calculated. Then, the data with date 19.02.2011 is extracted and $\widetilde{IOT}$ is set to current IOT which is as:

Current IOT for the cinema activity = Current Occurrence Date – Last Occurrence Date

= 19.02.2011 - 22.01.2011 = 28 days

With the extraction of second data $\widetilde{IOT}$ is set to 28 and the $\widetilde{Var}$ is set to 0. Next, the data whose date is 14.03.2011 is extracted and the values are calculated as:

Current IOT for the cinema activity = Current Occurrence Date – Last Occurrence Date

= 14.03.2011 - 19.02.2011 = 23 days

New $\widetilde{IOT} = (0.3 * \text{Current IOT}) + ((1 - 0.3) * \widetilde{IOT}_{n-1})$

= (0.3*23) + (0.7*28) = 26.5 days

New $\widetilde{Var} = ((|\text{Current IOT} - \text{Current } \widetilde{IOT}|^2) * 0.3) + (0.7 * \widetilde{Var}_{n-1})$

= (($|23 – 26.5|^2$) * 0.3) + (0.7 * 0) = 3.675

Finally, the last data whose date is 02.04.2011 is extracted and the values are calculated as:

Current IOT for the cinema activity = 02.04.2011 - 14.03.2011 = 19 days

New $\widetilde{IOT}$ = (0.3*19) + (0.7*26.5) = 24.25 days

New $\widetilde{Var}$ = (($|19 – 24.25|^2$) * 0.3) + (0.7 *3.675) = 10.84

Thus, $\widetilde{IOT}$ value is 24.25 and $\widetilde{Var}$ value is 10.84 for the activity of cinema.

This process is repeated for all data of the selected activities of the user. After this process is done for all the activities of the user, assume that the values are obtained for Guven as Table 18.

Table 18: Variables for the selected activities of the user in the 1st Filtering Stage of Activity Prediction Scenario

Activity	$\widetilde{IOT}$	$\widetilde{Var}$	$\tilde{\sigma}$	Last Occurrence Date
Fast-food	33.236	11.897	4.110	24.03.2011
Restaurant	15.491	2.709	1.646	21.04.2011
Cafe	7.756	2.912	1.706	24.04.2011
Cinema	24.25	10.840	3,292	02.04.2011
Bowling	Null	Null	Null	21.03.2011

As seen in Table 18, there is no calculated score for the activity of bowling, since the returned data from this filtering stage is less than two records. Therefore, D score for the activity of bowling is assigned to worst case D score of 3. D scores of other activities in this stage are computed according to the values given above.

For the activity of Fast-food:

Interval Time (IT) = Today's Date – Last Occurrence Date

= 30.04.2011 - 24.03.2011 = 37 days

$$D = \frac{|37-33.236|}{4.11} = 0.92$$

For the activity of Restaurant:

Interval Time (IT) = Today's Date – Last Occurrence Date

= 30.04.2011 - 21.04.2011 = 9 days

$$D = \frac{|9-15.491|}{1.646} = 3.94$$

Since D score of 3 is considered as the worst case in this study, the D scores that exceed 3are set to 3. Thus, D score for the activity of Restaurant is set to 3.

For the activity of Cafe:

Interval Time (IT) = Today's Date – Last Occurrence Date

= 30.04.2011 - 24.04.2011 = 6 days

$$D = \frac{|6-7.756|}{1.706} = 1.03$$

For the activity of Cinema:

Interval Time (IT) = Today's Date – Last Occurrence Date

= 30.04.2011 - 02.04.2011 = 28 days

$$D = \frac{|28-24.25|}{3.292} = 1.14$$

Hence, after the first filtering, the D scores of each selected activity for the user are computed and results are given in Table 19.

Table 19: D scores for the 1st Filtering Stage of Activity Prediction Scenario

Activity	D Score
Fast-food	0.92
Restaurant	3
Cafe	1.03
Cinema	1.14
Bowling	3

As shown in Table 19, since D score of Fast-food activity is the lowest, it has the highest possibility to be done by the user as the next activity according to the 1st filtering stage.

2nd Filtering Stage

In the second stage, to get more data for the prediction, related ontology is used for each context data and the matched data is extracted from context history of the user. Cepa as a location is a shopping center and so it is under the concept of shopping center according to ontology. Therefore all shopping centers can be used for the data in the prediction by the usage of ontology for location hierarchy. The day is Saturday and it is one of the weekend days according to ontology for day hierarchy. By the means of the ontology, the data with Sunday is also used for the prediction. Next, time '16.00-18.00' is in the time interval of evening according to time ontology. The weather is rainy and its usage is also generalized by the usage of weather ontology. As a result of these, a sample filtered data from context history of the user is like Table 20.

Table 20: Sample of Filtered data in the 2nd Filtering Stage of Activity Prediction Scenario

Date	Day	Time	Location	Weather	Nearby Person	Activity
24.04.2011	Sunday	16.00-18.00	Kentpark	Rainy	Irmak	Cafe
17.04.2011	Sunday	16.00-18.00	Cepa	Rainy	Irmak	Restaurant
16.04.2011	Saturday	18.00-20.00	Cepa	Cloudy	Irmak	Restaurant
….	….	….	….	….	….	….

Then the all the works in the first stage of filtering process are repeated and finally, D scores are calculated for all the selected activities of the user.

3rd Filtering Stage

In this stage, all the data is extracted that exactly matching with the specified activities. The current context data of the user is not used in the filtering. This stage aims to get all the data that is independent of the context information. Thus, a sample filtered data for activity of Cafe from context history of the user is like Table 21.

Table 21: Sample of Filtered data in the 3rd Filtering Stage of Activity Prediction Scenario

Date	Day	Time	Location	Weather	Nearby Person	Activity
24.04.2011	Sunday	16.00-18.00	Kentpark	Rainy	Irmak	Cafe
21.04.2011	Thursday	14.00-16.00	Bahceli	Cloudy	Serhat	Cafe
19.04.2011	Tuesday	20.00-22.00	Bilkent	Clear	Irmak	Cafe
….	….	….	….	….	….	….

Then the all the works in the first two stages of filtering process are repeated and finally, D scores are calculated for all the selected activities of the user.

Prediction

After the stages are executed the next activity preference of the user is predicted according to the estimated D ($\widetilde{D}$) scores of each activity. In order to compute $\widetilde{D}$ score of each selected activity, D scores of each stage for each activity are used. After D score calculation is done for all the selected activities in each stage, suppose that the D scores are as shown in Table 22.

Table 22: D scores for all filtering steps

Activity	D_1 Score for 1st stage	D_2 Score for 2nd stage	D_3 Score for 3rd stage
Fast-food	0.92	0.83	0.75
Restaurant	3.0	3,0	0.68
Cafe	1.03	0.98	0.71
Cinema	1.14	1.02	0.86
Bowling	3.00	2.46	2.86

The $\widetilde{D}$ score is calculated as:

$$\widetilde{D} = (\beta * D_1) + (\gamma * D_2) + ((1 - \alpha - \gamma) * D_3)$$

The weighting factors (β) and (γ) are defined for this scenario and their values are 0.6 and 0.25 respectively. Therefore, $\widetilde{D}$ score for the activity of Fast-food:

For the activity of Fast-food:

$\widetilde{D}$ = (0.6*0.92) + (0.25*0.83) + (0.15*0.75) = 0.87

The $\widetilde{D}$ scores are computed for each activity as the fast food activity. After calculating $\widetilde{D}$ scores of each selected activity for the user, the scores are obtained as shown in Table 23.

Table 23: Estimated D scores for selected activities

Activity	$\widetilde{D}$
Fast-food	0.87
Restaurant	2.65
Cafe	0.96
Cinema	1.07
Bowling	2.84

When the $\widetilde{D}$ scores of the activities are listed in ascending order, we have a list:

{Fast-food, Café, Cinema, Restaurant, Bowling}

The $\widetilde{\mathrm{D}}$ score indicates the distance of the current IT to $\mathrm{I\widetilde{O}T}$ value for each activity. As a result of this, the activity that has the lowest $\widetilde{\mathrm{D}}$ score is the highest possibility to be done by the user as a next activity.

When we look at the table, the activity of Fast-food is highly probable, since its distance between current IOT and $\mathrm{I\widetilde{O}T}$ value is close. On the other hand, this difference for the activities of Restaurant and Bowling is very high and so the possibilities of their occurrences are low. In conclusion, the activity of Fast-food has highest possibility for the next activity preference of the user café, cinema, restaurant and bowling follow this activity respectively.

CHAPTER 4

A SAMPLE PROTOTYPE IMPLEMENTATION

In this chapter, the prototype software implemented to show feasibility of the proposed model is introduced.

4.1. Aim and Scope

We developed a web based prototype application that implements the core features of the proposed model. It is a kind of a personalized predictor that predicts the activity preferences of the user according to provided current context. The prototype mainly aims to show the applicability of the proposed model and its algorithms. The prototype application focuses on the outside activity prediction and the context dimensions and the ontology models defined in the previous chapter are used in this prototype. Like in the proposed model, the prototype assumes that the high level context information is provided by the user. That is using raw sensor data and processing it to obtain high level context information is not performed by the prototype.

4.2. High Level Requirements

High level requirements of developed prototype are follows:

- System shall make use of all defined context dimensions.
- System shall accept the context history of each user as the input.
- System shall accept the current context information of any user as the input.
- System shall keep a context history for each user.

- System shall manage the activities of each user separately.
- System shall query context history of each user according to provided context information.
- System shall use the ontology models.
- System shall get current context information and predict and display the list of all possible activities for a user.

4.3. System Design

The prototype is implemented as a web based system on the ASP.NET platform and Microsoft SQL Server is used as the database management system.

The user is the only external entity interacting with the system and the interaction between the user and the system is illustrated in Figure 8. The system takes the context history of the user and while the data is retrieved from the user, it manages statuses of the activities and saves the data for the context history of the user. Moreover, the current context which is generated outside the system is entered to system as an input by the user and prediction results are returned to the user by the system.

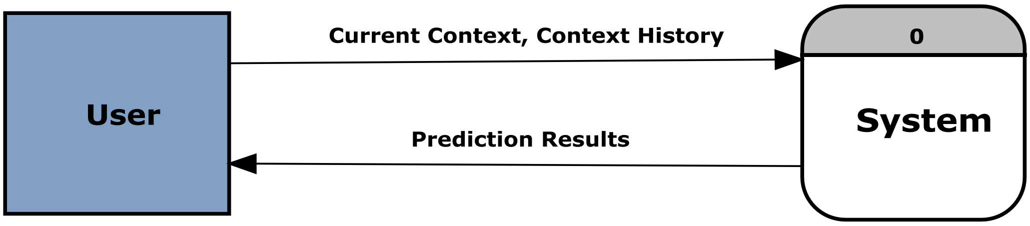

Figure 8: Level 0 DFD of the Prototype Implementation

Our web based system consists of two main modules. These are context history acquisition and activity prediction modules. The context history acquisition module takes the context data from the user by the *context history handler component*. Moreover, the activities of the user are managed by *activity manager* component. On the other hand, the activity prediction module takes the current context data from the user and performs the prediction by *context* *data handler*, *filtering agent*, *ontology provider*, *prediction engine* components. The modules of the system and constituent components are depicted in Figure 9.

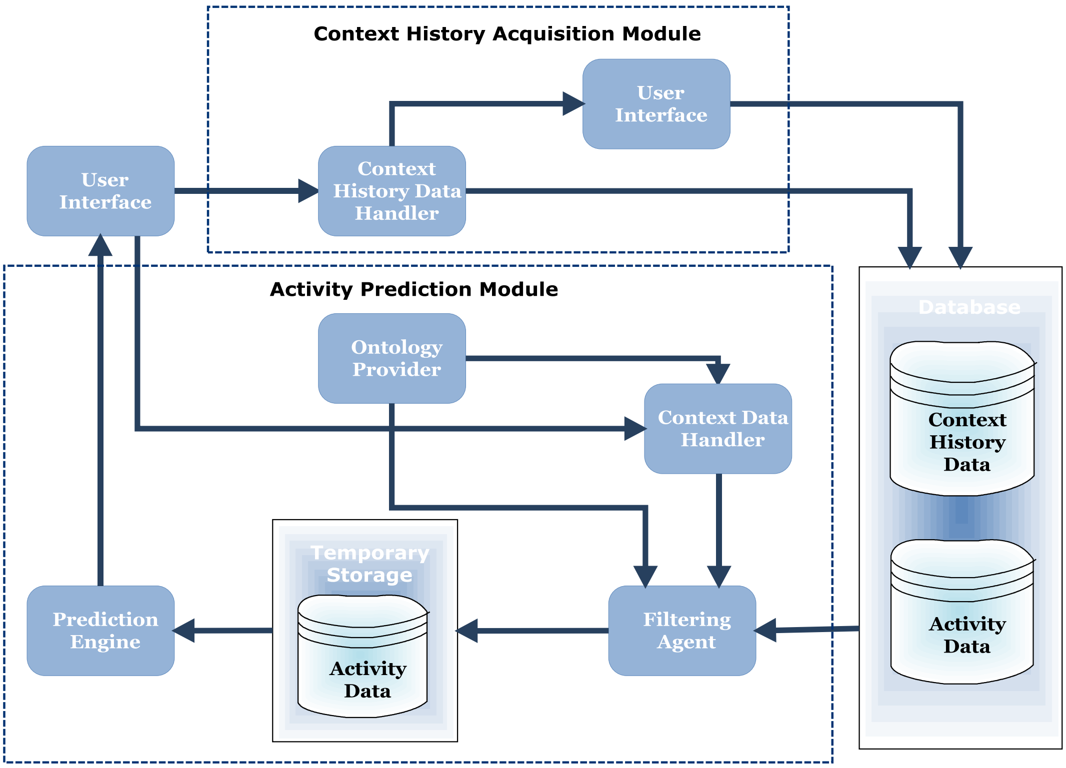

Figure 9: System Structure of the Prototype Implementation

4.3.1. Context History Acquisition Module

This module takes the context history from the user and manages the activities of the user according to retrieved data. This module monitors the user's activities and for each data entry all activities of the user are checked and related changes are done. This module consists of the *context history handler* and *activity manager* components.

4.3.1.1. Context History Handler

Context history data of the users are delivered first to this component in the system. This component stores the historical data for the occurred activities to the database. Moreover, context history handler provides related data to the *activity manager*.

4.3.1.2. Activity Manager

This component takes the data sent by the *context history handler* and manages the activities of the user. If the occurred activity is 'Disabled', this component checks the include conditions for the activity to change its status to 'Enabled'. For each of the 'Enabled' activities, it checks the discard conditions to change the activity status to 'Disabled'. If the activity fulfills the conditions, changes the activity's status correspondingly. Moreover, this component checks the activities that do not occur frequently. This component also updates $\widetilde{\mathrm{IOT}}$ and $\widetilde{\mathrm{Var}}$ values for each occurred activity. Finally, upon processing related records are updated in the database.

4.3.2. Activity Prediction Module

Activity prediction module takes the context data entered by the user and performs three stages of prediction. This module is composed of four main components and each component performs a specific step of the prediction. These components are: *context data handler*, *filtering agent, ontology provider* and *prediction engine.*

4.3.2.1. Context Data Handler

Context data handler component takes the current context from the user and delivers this data to the *filtering agent*. This component also uses *ontology provider* to specify the candidate activities for the current location.

4.3.2.2. Ontology Provider

Ontology provider has three important roles. Firstly, it keeps and manages the ontology structures for each context dimension. When new concepts occur for any context dimension, the related concept in the ontology structure of this context is updated by this component. Secondly, it keeps and provides the available activities for the current location to *context data handler*. Finally, *filtering agent* applies to the ontology provider to get the more generalized concept specifications of each context. Thus, ontology returns this information to *filtering agent* to use in the second stage of the prediction.

4.3.2.3. Filtering Agent

Filtering agent module takes the data sent by the *context data handler* and *ontology provider* and performs the matching processes. This agent performs three different matching types. Firstly, it matches the current context of the user with the user context history. Secondly it matches the more generalized concepts taken from the *ontology provider*. Lastly, this agent filters the all data of the activities that are specified at the beginning of the prediction process. Filtering agent also calculates $\widetilde{IOT}$ and $\widetilde{Var}$ values for each activity in each filtering stage and it updates these values for each retrieved data from the user context history. The computed values for each activity are temporarily stored in order to be used in the prediction process.

4.3.2.4. Prediction Engine

This is the actual component that performs the prediction algorithm. This engine takes the $\widetilde{IOT}$ and $\widetilde{Var}$ values of each activity for each filtering stage from the temporary storage and computes firstly $\tilde{\sigma}$ of IOT values. Next, it computes D scores for each activity in each filtering stage. Then it calculates $\tilde{D}$ score for each activity by combining D scores obtained from each filtering stage. Finally, prediction engine orders the activities according to ascending order of $\tilde{D}$ scores and displays them.

4.3.2.5. Temporary Storage

Temporary storage stores $\widetilde{IOT}$ and $\widetilde{Var}$ values of each activity for each filtering stage. This data is supplied by the filtering agent and the prediction engine retrieves the data and computes D scores to calculate $\tilde{D}$ score for each activity.

4.3.3. Database

The prototype primarily stores and maintains two main types of data. First one is context history data that is used for the future activity prediction of the users. Context history consists of the context data for each occurred activity of the user. This data is entered by the user using the “data history upload” module.

Second one is the activity data for each user and it is used for the management of activities. This data consists of $\widetilde{\mathrm{IOT}}$ and $\widetilde{\mathrm{Var}}$ values, status of the activity, and the last occurrence date of the activity. This data is updated upon uploading the user context history or adding new context history.

4.4. User Interface

The prototype requires authentication and any user must log in the system to use it. After the user logs into the system, s/he gets the home page and the main menu which has the choices of data history upload and activity prediction.

In order to upload user context history, data history upload option is selected and the upload screen shown in Figure 10 is displayed. As seen in the figure, the stored context history of the user is listed and two options are offered: The context history of the user can completely be changed or new historical records can be added to stored context history.

If update option is selected, currently stored context history is deleted and the uploaded context history becomes the new history. On the other hand, if add option is selected, uploaded context history is added to stored current history. Uploading process accepts Microsoft Excel files and the expected format of the file is specified in the screen. For each entry of the uploaded data set, the activities of the user are managed by the related algorithms.

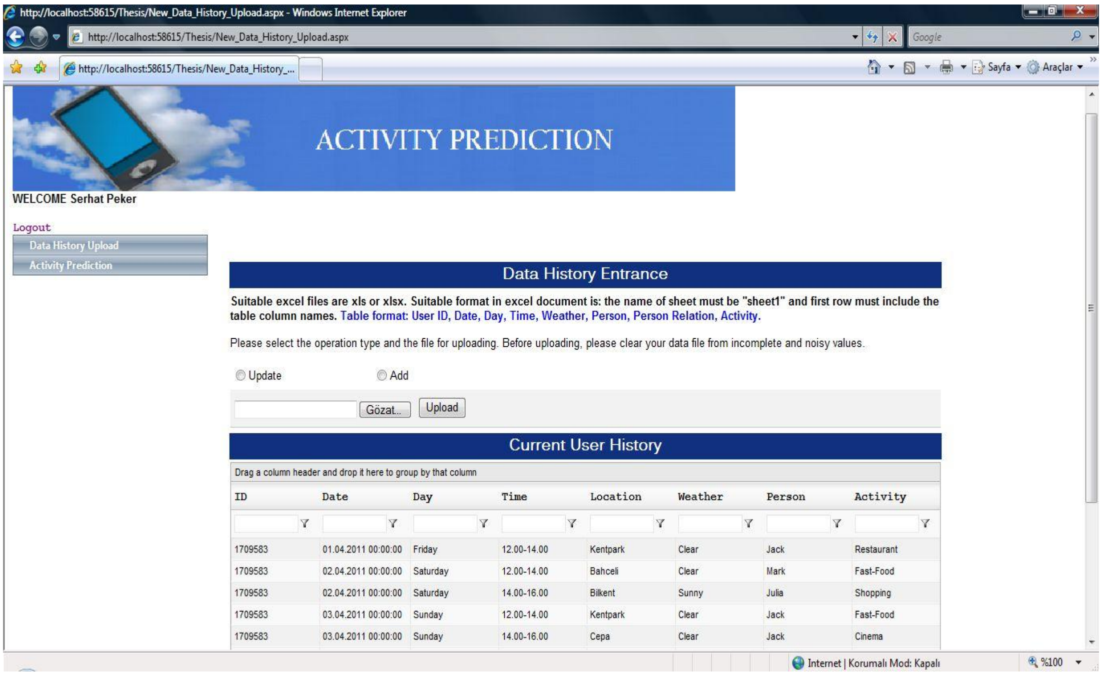

Figure 10: User Context History Uploading Screen

Activity prediction option in the main menu is selected to provide a prediction for the entered current context. The screen for activity prediction is given in Figure 11. As seen in the figure, this screen has a form input for the necessary fields for the current context.

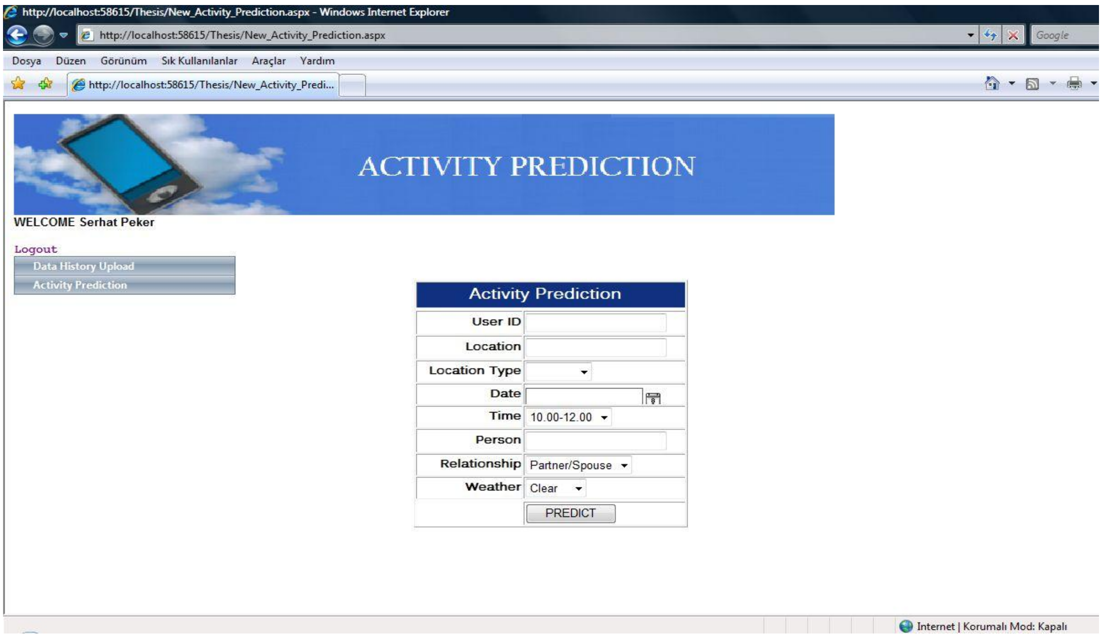

Figure 11: User Activity Prediction Screen

When all the information is entered and 'predict' button is clicked, prediction process algorithms are run. Then a list of possible activities is displayed for the next activity preference of the user on the screen as Figure 12. The list is sorted with the most probable activity is shown on the top.

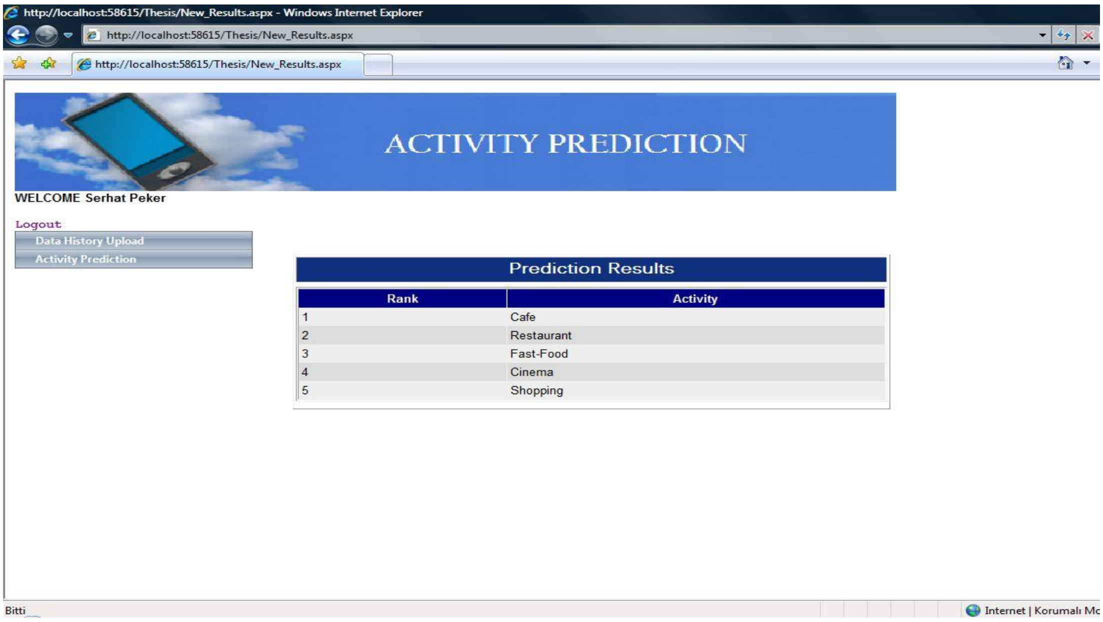

Figure 12: Results Screen

4.5. Evaluation and Test Case

The evaluation of our implemented prototype is performed through the use of synthetic context history. We didn't perform an evaluation with real data of a real user as we do not have such a database and collecting real user data require a very long time and it would not be possible to have enough data within the time span of this thesis study. Therefore, we generate context history for activity preferences of a user in different contexts over a period of four months. The generated context history is given in the Appendix A. This context history is generated in a logical manner to demonstrate that our model gives reasonable results with the consistent context history. As an example, cinema preferences of the user in the generated context history are given in Table 24.

Table 24: Cinema Preferences of the User

Date	Day	Time	Location	Weather	Nearby Person	Relationship	Activity
08.01.2011	Saturday	14.00-16.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cinema
22.01.2011	Saturday	14.00-16.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cinema
05.02.2011	Saturday	16.00-18.00	Kentpark	Snowy	Gaye	Partner/Spouse	Cinema
19.02.2011	Saturday	14.00-16.00	Kentpark	Rainy	Özge	Friend	Cinema
26.02.2011	Saturday	16.00-18.00	Bilkent	Snowy	Gaye	Partner/Spouse	Cinema
12.03.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cinema
24.03.2011	Thursday	20.00-22.00	Kentpark	Rainy	Davut	Friend	Cinema
09.04.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cinema
17.04.2011	Sunday	14.00-16.00	Kentpark	Sunny	Gaye	Partner/Spouse	Cinema

As seen in Table 24, we assume that the user mostly goes to cinema in the "Kentpark" mostly on Saturdays. It is also assumed that most of the time the user prefers to go cinema with his partner at rainy or cloudy weathers. Moreover, we assumed that the user goes to cinema once in a two week period. The $\widetilde{\mathrm{IOT}}$ and $\widetilde{\mathrm{Var}}$ values for the activity of cinema are 11.8 and 7.7 respectively.

There are also similar considerations for other activities in our data set. As a result, the synthetic data that we generated is consistent and we expect out model give meaningful results with this synthetic context history.

For the test cases, the parameters given in Table 3 are used in the activity management and activity prediction algorithms. The statistics for the activities obtained from the context history are as shown in Table 25. We analyzed three cases. In the first case the prediction is done for a given context. Then we explore other cases by inserting new entries into the initial context history and observing the changes in the prediction results. For each case, the results obtained are discussed.

Table 25: Statistics for the Activities before the Prediction

Activity	Last Occurence	Status	$\widetilde{\mathrm{IOT}}$	$\widetilde{\mathrm{Var}}$	AC
Restaurant	26.04.2011	Enabled	4.51	1.43	22
Fast-Food	19.04.2011	Enabled	9.08	4.15	14
Cafe	28.04.2011	Enabled	2.51	1.86	42
Cinema	17.04.2011	Enabled	11.82	7.71	9
Bowling	10.04.2011	Enabled	32.2	428.65	3

Activity Prediction – Case 1

Suppose that it is 16.15 on Saturday and the user is with his partner Gaye in Kentpark. All the current context information for the user is given in Table 26. This context information is entered to the prototype system for the activity prediction.

Table 26: Current Context of the User for all Cases

Date	Day	Time	Location	Weather	Nearby Person	Relationship
30.04.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse

The historical context data that exactly matches with the current context of the user is given in Table 27.

Table 27: Entries Returned in the 1st Prediction Stage

Date	Day	Time	Location	Weather	Nearby Person	Relationship	Activity
08.01.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cafe
22.01.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cafe
12.03.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cinema
09.04.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cinema

The historical context data for similar context with the current context of the user is given in Table 28.

Table 28: Entries Returned in the 2nd Prediction Stage

Date	Day	Time	Location	Weather	Nearby Person	Relationship	Activity
08.01.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cafe
15.01.2011	Saturday	16.00-18.00	Kentpark	Snowy	Gaye	Partner/Spouse	Cafe
22.01.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cafe
05.02.2011	Saturday	16.00-18.00	Kentpark	Snowy	Gaye	Partner/Spouse	Cinema
12.02.2011	Saturday	16.00-18.00	Kentpark	Cloudy	Gaye	Partner/Spouse	Cafe
05.03.2011	Saturday	16.00-18.00	Kentpark	Snowy	Gaye	Partner/Spouse	Cafe
12.03.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cinema
09.04.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cinema

As shown in Table 27 and Table 28, only the activities of Café and Cinema were performed in the past in the same and similar contexts by the user. The prediction results produced by the prototype for the current context of the user are given in Table 29.

Table 29: Results for Case 1

Rank	Activity	D_1	D_2	D_3	$\widetilde{D}$
1	Cinema	1.82	1.64	0.42	1.57
2	Cafe	1.82	1.75	0.37	1.59
3	Restaurant	3	3	0.43	2.76
4	Fast-Food	3	3	0.94	2.84

In Table 29 for each activity, D_1 is the distance score of historical context entries exactly matching with the current context of the user, D_2 is the distance score of all historical context entries that has similar context with the current context of the user and D_3 is the distance score of all historical data of the user without considering the context information.

As Table 29 shows, the prototype application does not return score for “bowling” activity as it is not included in the prediction process. The reason behind this is that “bowling” activity’s AC does not exceed defined CT for the case (3<5). Moreover, D_1 and D_2 scores of the activities of Restaurant and Fast-Food are taken as 3 as there is no matching entry in the first and second stages of prediction.

However D_3 score of Café is less than the score of the Cinema and the $\widetilde{D}$ score of Cinema is less than the score of Café. The reason behind this is that D_2 score of the Cinema is less than that of Café’s and weight of second stage in the overall score is larger than that of third stage in our parameter set.

Cinema and Café are highly probable in this scenario as their $\widetilde{D}$ scores are very small. On the other hand, $\widetilde{D}$ scores of Restaurant and Fast-food are large and so the possibilities of their occurrences are low. In conclusion, the cinema has highest possibility for the next activity preference of the user; cafe, restaurant and fast-food follow this activity respectively.

As seen in Table 27 and Table 28 the activities of Restaurant and Fast-food have not been performed in the same or similar contexts by the user before. As a result, they are least likely to be the next activity of the user. Moreover, Café and Cinema were performed in the same or similar contexts by the user several times in the past. Thus, they turn out to be the first two activities that are highly possible as the next activity of the user. Cinema is the most likely to be the next activity of the user as this activity has the lowest $\widetilde{D}$ score. Time line for the activity Cinema is depicted in Figure 13.

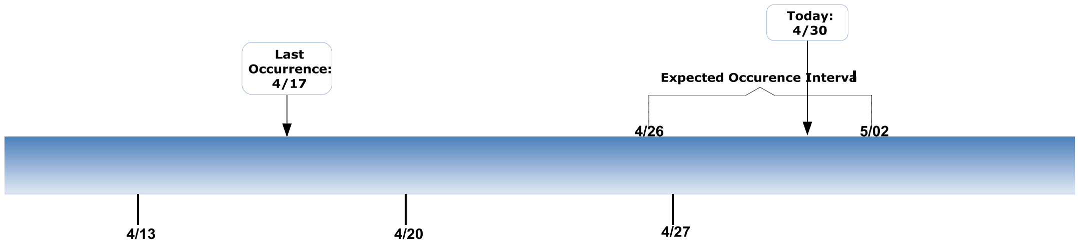

Figure 13: Time Line for Cinema Activity in Case 1

According to Table 25, its $\widetilde{IOT}$ is 11.8 and $\widetilde{\sigma}$ is calculated as 2.7. As Figure 13 shows the last occurrence date of the cinema is 17.04.2011. Let's define expected occurrence interval as the one standard deviation neighborhood of the expected next occurrence date of an activity. That is:

The expected occurrence interval = ($\widetilde{IOT}$ +/- $\widetilde{\sigma}$) days after the last occurrence date of the activity

The expected occurrence interval for activity Cinema can be found as 11.8 +/- 2.7 days after the last occurrence date of the Cinema activity. Thus, expected occurrence interval is between 26.04.2011 and 02.05.2011. Since today is 30.04.2011 and it is in the expected occurrence interval of the activity of cinema, cinema is highly probable as a next activity of the user.

Activity Prediction – Case 2

In this case, we again used the same context history except one record addition. We suppose that the user went to cinema on 24.04.2011 and the context record given in Table 30 is added to context history of the user.

Table 30: Added Historical Context Record for Case 2

Date	Day	Time	Location	Weather	Nearby Person	Relationship	Activity
24.04.2011	Sunday	16.00-18.00	Kentpark	Cloudy	Davut	Friend	Cinema

This time, the statistics for the cinema activity will change and new statistics are given in Table 31.

Table 31: Updated Statistics for Cinema Activity in Case 2

Activity	Last Occurrence	Status	$\widetilde{IOT}$	$\widetilde{Var}$	AC
Cinema	24.04.2011	Enabled	10.37	8.81	10

According to these statistics, the time line for the activity cinema will be like the one given Figure 14. As Figure 14 shows the expected occurrence interval is between 02.05.2011 and 08.05.2011. Since today is 30.04.2011 and it is not in the expected occurrence interval of the activity of cinema, it is less probable as a next activity of the user than one in the case 1.

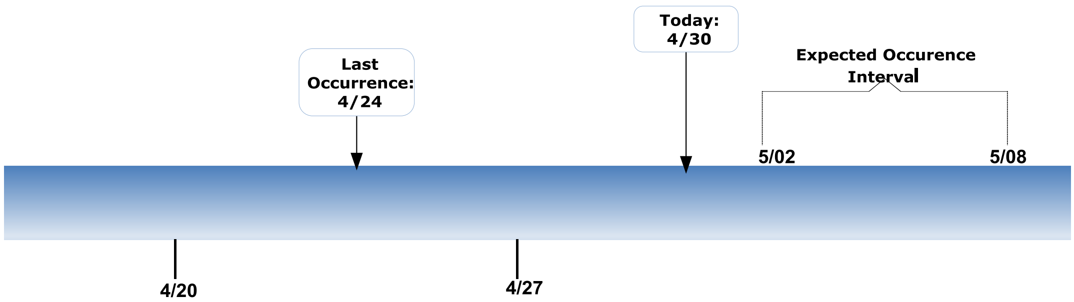

Figure 14: Time Line for Cinema Activity in Case 2

Then, the results for the prediction according to current context information of the user in Table 26 are generated by the prototype as in Table 32.

Table 32: Results for Case 2

Rank	Activity	D_1	D_2	D_3	$\widetilde{D}$
1	Cafe	1.82	1.75	0.37	1.59
2	Cinema	1.82	1.64	1.48	1.73
3	Restaurant	3	3	0.43	2.76
4	Fast-Food	3	3	0.94	2.84

As seen in Table 32, D_3 score of the cinema is the only change when the results are compared to the results in case 1. As we inserted a recent Cinema activity the last occurrence date of the cinema changes and therefore today falls outside the expected occurrence interval as shown in Figure 14. Therefore D_3 score for the activity of cinema increases.

According to Table 32, Cafe has highest possibility to be selected as the next activity of the user. Cinema, restaurant and fast-food follow this activity respectively. Time line for the activity cafe is given in Figure 15.

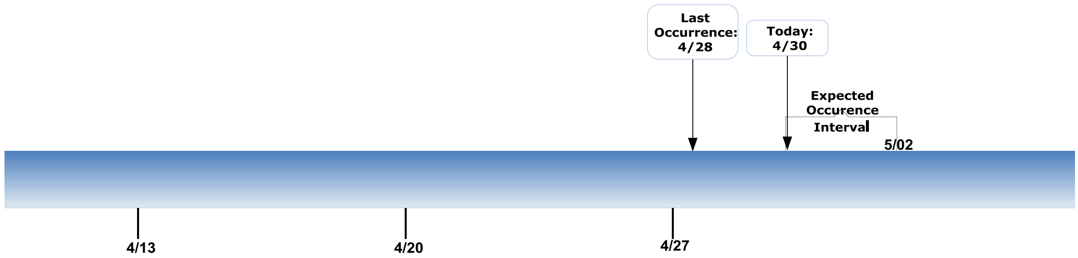

Figure 15: Time Line for Cafe Activity in Case 2

As the figure illustrates the expected occurrence interval of cafe is between 30.04.2011 and 02.05.2011. Since today is 30.04.2011 and it is in the expected occurrence interval of the activity of cafe, cafe is also highly probable as a next activity of the user.

Activity Prediction – Case 3

Now we add the historical context data given in Table 33 to the context history used in Case 2. According to inserted record the user went to café on the same day of the activity prediction.

Table 33: Added Historical Context Record for Case 3

Date	Day	Time	Location	Weather	Nearby Person	Relationship	Activity
30.04.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cinema

The statistics for the cafe activity change and new values are shown in Table 34.

Table 34: Updated Statistics of Cafe Activity in Case 3

Activity	Last Occurrence	Status	$\widetilde{IOT}$	$\widetilde{Var}$	AC
Cafe	30.04.2011	Enabled	2.35	1.34	43

As seen in Table 34, $\widetilde{IOT}$ is 2.35 and $\tilde{\sigma}$ is calculated as 1.16 and the last occurrence date is also updated as 30.04.2011 for the user.

Thus, expected occurrence interval for café is now between 02.05.2011 and 04.05.2011. Since today is 30.04.2011 and it is not in the expected occurrence interval of cafe, it is less probable as the next activity of the user as compared to the case 2. The time line for the activity cafe is given in Figure 16.

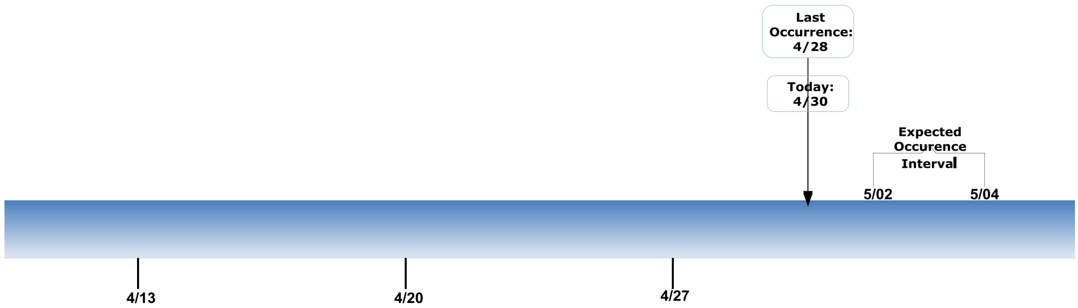

Figure 16: Time Line for Café Activity in Case 3

As a result the prediction according to current context information of the user is generated by the prototype as shown in Table 35.

Table 35: Results for Case 3

Rank	Activity	D_1	D_2	D_3	$\tilde{D}$
1	Cinema	1.82	1.64	1.48	1,73
2	Cafe	1.82	1.75	2.03	1.84
3	Restaurant	3	3	0.43	2.76
4	Fast-Food	3	3	0.94	2.84

A seen in Table 35, D_3 score of the cafe is the only change from the results of case 2. Since the last occurrence data of the cafe changes and today falls outside the expected occurrence interval of café as shown in Figure 16. Thus, D_3 score for the activity of café increases.

According to Table 35, Cinema has highest possibility to be selected as the next activity by the user, since this activity has the lowest $\widetilde{IOT}$ score. Cinema, restaurant and fast-food follow this activity.

CHAPTER 5

CONCLUSION AND FUTURE WORK

There is a growing need to properly predict the next activities of the individuals to get the most valuable and appropriate recommendations from recommender systems. However, existing context aware recommender systems usually do not exactly meet the users' actual needs in heterogeneous environments offering different activity alternatives.

In this thesis, we propose a model to properly predict the activity preferences of the individuals to improve the service quality of the context aware recommender systems and satisfaction of the users. Our activity prediction model is based on collecting the past activity preferences of the individuals in certain contexts and activity preferences in similar contexts using historical data. That is, we use the context history of the individuals for the prediction. The next activity of the user in the current context is predicted using the interoccurrence times of the past activities of the user in similar contexts.

Typically, recommender systems suffer from cold start problem. That is, there may not be enough historical data to use for the activity prediction for an individual user. In such cases, it is difficult to achieve high prediction accuracy with insufficient data. Therefore, in this thesis, ontology hierarchies are defined for each context dimension and related ontology is used to consider activities performed in similar contexts. Our model use three filtering stages and in one of these stages, we consider similar contexts to incorporate more data in the prediction process. As a result, the accuracy of the prediction is improved.

The individuals usually perform the activities in specific time periods. Therefore our prediction approach makes use of the interoccurrence times of the past activities. In the prediction model, for each activity interoccurrence time statistics are calculated and prediction is performed on these statistics. We defined and used $\widetilde{\mathrm{D}}$ score which is the weighted average of obtained D scores in each filtering stage and indicates distance of current IT of the activity to its $\mathrm{I}\widetilde{\mathrm{OT}}$ on that date. $\widetilde{\mathrm{D}}$ score shows the occurrence possibility of each activity for the users. The activity with least $\widetilde{\mathrm{D}}$ score is determined as the next activity of the user with the highest possibility to occur.

To minimize inappropriate predictions using stale data, our model also proposes an activity management mechanism which manages the activity status. The algorithms in this mechanism are proposed to realize the seasonal or periodic patterns in activity occurrence dates and exceptions, changed or interrupted activities of the user, so the prediction adapts to the change of user's preferences in time. With this mechanism activities are included into or excluded from the prediction process according to their statuses. For example, the activities that do not occur for a long time are discarded and they are not included in the prediction process. Thus, this mechanism contributes the improvement of the prediction precision.

Some activities rarely occur for some users. To overcome this problem, activity count and prediction thresholds are defined and used in our model to recognize the exceptions. As a result, an activity is included in the prediction process if and only if it fulfills the defined threshold values. Hence the rarely occurring activities are not included in the prediction process to avoid mostly incorrect predictions.

In order to assess the applicability of the proposed model, a sample prototype software application is implemented. This prototype allows user to import his/her activity/context history and predicts the user's next activity preference according to the entered current context of the user by running the prediction algorithms of the proposed model. The results obtained from the use of the prototype suggest that the proposed model can be used to improve the service quality of the context aware recommender systems and satisfaction of users in ubiquitous computing environments.

The proposed model in this thesis contributes the context aware recommender systems by predicting the next activity preferences of the individuals. This model does not require large sets of inputs from the user to provide recommendations and the explicit user interaction will be kept as low as possible. Hence, CARS work seamlessly and provide more personalized services for the users, so the users would not be disturbed with irrelevant recommendations at inappropriate moments.

However, some improvements can still be made on this model. First of all, five main context dimensions which are most widely used and popular are used in this thesis. In the future, other context dimensions might be considered in prediction process so using more information will improve the accuracy of the context prediction.

This study does not consider the special dates for the users and also the holidays in the years. In these dates, the user's preferences and habits may be different. Therefore such special days might be considered in prediction.

Two level ontology structures are used for each context dimension but in the future more levels can be introduced for each ontology structure. Moreover, upper, more generalized concepts are used for each context dimension in the second step of the filtering stage. However, cross matching in the second step of the filtering might be used as a future work. With this improvement, a specific context can be filtered with the more generalized concepts of other remaining contexts.

Furthermore, a web based prototype is developed to show the applicability of the model and the generated synthetic data is used in the tests instead of real time sensor data due to the time limitations. In the future, a mobile application might be developed and so the real time sensor data might be used by the usage of related technologies in the evaluation of the prototype.

Privacy concerns related to such recommendation systems might be considered in a future work. A policy might be defined for the user to control the sharing of their private information and so the proposed model checks this policy if it is permitted to acquire the required information.

The proposed model considers the domain of outside activities in this study. It might be specified for a context recommender system for an activity such as activity of cinema. The habits of individuals for a specific activity might be predicted by the usage of this model in the future. If the related changes are made in the model, it might be also applied for the different domains such as ageing people and infant caring. Thus, the needs of ageing people or infants might be predicted and satisfied.

REFERENCES

Abbar, S., Bouzeghoub, M., & Lopez, S. (2009). Context-Aware Recommender Systems: A Service Oriented Approach. *Proceedings of the 3rd International Workshop on PersDB, In conjunction with VLDB'09. Lyon, France* (pp. 31–36).

Adomavicius, G., & Tuzhilin, A. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering*, *17*(6), 734-749.

Adomavicius, G., Sankaranarayanan, R., Sen, S., & Tuzhilin, A. (2005). Incorporating Contextual Information in Recommender Systems Using a Multidimensional Approach. *ACM Transactions on Information Systems*, *23*(1), 103-145.

Anderson, C. (2006). The Long Tail: How Endless Choice is Creating Unlimited Demand.

Balabanović, M., & Shoham, Y. (1997). Fab: Content-Based, Collaborative Recommendation. *Communications of the ACM*, *40*(3), 66-72.

Byun, H. E., & Cheverst, K. (2004). Utilizing Context History to Provide Dynamic Adaptations. *Applied Artificial Intelligence*, *18*, 533-548.

Chen, H., Finin, T., & Joshi, A. (2003). An Ontology for Context-Aware Pervasive Computing Environments. *The Knowledge Engineering Review*, *18*(03), 197–207.

Davison, B. D., & Hirsh, H. (1998). Predicting Sequences of User Actions. *Predicting the Future: AI Approaches to Time-Series Problems* (pp. 5–12).

Dey, A. K., & Abowd, G. D. (2000). Towards A Better Understanding of Context and Context-Awareness. *CHI 2000 workshop on the what, who, where, when, and how of context-awareness* (Vol. 4, pp. 1–6).

Gruber, T. R. (1993). A Translation Approach to Portable Ontology Specifications. *Knowledge acquisition*, *5*(2), 199–220.

Gu, T., Wang, X. H., Pung, H. K., & Zhang, D. Q. (2004). An Ontology-Based Context Model in Intelligent Environments. *Proceedings of communication networks and distributed systems modeling and simulation conference* (Vol. 2004, pp. 270–275).

Hong, J., Suh, E. H., Kim, J., & Kim, S. Y. (2009). Context-Aware System for Proactive Personalized Service Based on Context History. *Expert Systems with Applications*, *36*(4), 7448–7457.

Jacobson, V. (1988). Congestion Avoidance and Control. *ACM SIGCOMM Computer Communication Review* (Vol. 18, pp. 314–329).

Kaowthumrong, K., Lebsack, J., & Han, R. (2002). Automated Selection of The Active Device in Interactive Multi-Device Smart Spaces. *Workshop at UbiComp'02: Supporting Spontaneous Interaction in Ubiquitous Computing Settings*.

Laasonen, K., Raento, M., & Toivonen, H. (2004). Adaptive On-Device Location Recognition. *Pervasive Computing* (Vol. 3001, pp. 287-304). Berlin, Heidelberg: Springer

Lee, K. C., Cho, H., & Lee, S. (2010). U-BASE: General Bayesian Network-Driven Context Prediction for Decision Support. *Advances in Information Technology*, 63–72.

Lee, S. Y., & Tran, D. L. (2010). Users' Habits-based Context Prediction applied to Smart Buildings. *Lecture Notes in Engineering and Computer Science, 2186*.

Lee, S., Lee, K. C., & Cho, H. (2010). A Dynamic Bayesian Network Approach to Location Prediction in Ubiquitous Computing Environments. *Advances in Information Technology*, 73–82.

Lombardi, S., Anand, S. S., & Gorgoglione, M. (2009). Context and Customer Behavior in Recommendation. *RecSys09 Workshop on Context-aware Recommender Systems*.

Mahalanobis, P. C. (1936). On the Generalized Distance in Statistics. *Proceedings of the National Institute of Science, Calcutta* (Vol. 12, p 49).

Mayrhofer, R. (2005). An architecture for context prediction. Doctoral dissertation, Linz University.

McCarthy, J. F. (2001). Active Environments: Sensing and Responding to Groups of People. *Personal and Ubiquitous Computing, 5*(1), 75–77.

Mozer, M. C. (1998). The Neural Network House: An Environment hat Adapts to its Inhabitants. *Proc. AAAI Spring Symp. Intelligent Environments*.

Oku, K., Nakajima, S., Miyazaki, J., & Uemura, S. (2006). Context-Aware SVM for Context-Dependent Information Recommendation. *Mobile Data Management*, *IEEE International Conference* (Vol. 0, p 109). Los Alamitos, CA, USA: IEEE Computer Society.

Patterson, D. J., Liao, L., Fox, D., & Kautz, H. (2003). Inferring High-Level Behavior from Low-Level Sensors. *UbiComp 2003: Ubiquitous Computing* (Vol. 2864, pp 73-89). Berlin, Heidelberg: Springer Berlin Heidelberg.

Petzold, J., Bagci, F., Trumler, W., & Ungerer, T. (2003). Global and Local State Context Prediction. *Artificial Intelligence in Mobile Systems*.

Petzold, J., Bagci, F., Trumler, W., Ungerer, T., & Vintan, L. (2004). Global State Context Prediction Techniques Applied to A Smart Office Building. *The Communication Networks and Distributed Systems Modeling and Simulation Conference*.

Poslad, S. (2009). Ubiquitous Computing: Smart Devices, Environments and Interactions. John Wiley & Sons Inc.

Prahalad, C. K. (2004). Beyond CRM: CK Prahalad Predicts Customer Context is the next Big Thing. *American Management Association MwWorld*.

Pukelsheim, F. (1994). The three sigma rule. *American statistician*, 88–91.

Ranganathan, A., McGrath, R. E., Campbell, R. H., & Mickunas, M. D. (2003). Use of Ontologies in A Pervasive Computing Environment. *The Knowledge Engineering Review*, *18*(03), 209–220.

Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, *40*(3), 56-58.

Sigg, S., Haseloff, S., & David, K. (2006). A Novel Approach to Context Prediction in UBICOMP Environments. *2006 IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications* (pp. 1-5).

Sigg, S., Haseloff, S., & David, K. (2010). An Alignment Approach for Context Prediction Tasks in UbiComp Environments. *IEEE Pervasive Computing*, *9*(4), 90-97.

Tennenhouse, D. (2000). Proactive computing. *Communications of the ACM*, *43*(5), 43–50.

Tuzhilin, A., & Adomavicius, G. (2008). Context-aware recommender systems. *Proceedings of the 2008 ACM conference on Recommender systems RecSys 08*, *16*(RecSys), 335.

Van Setten, M., Pokraev, S., & Koolwaaij, J. (2004). Context-Aware Recommendations in the Mobile Tourist Application COMPASS. *Adaptive Hypermedia and Adaptive Web-Based Systems* (pp. 515–548).

Weiser, M. (1991). The Computer for the 21st Century. *Scientific American*, *265*(3), 94–104.

Yujie, Z., & Licai, W. (2010). Some Challenges for Context-Aware Recommender Systems. *2010 5th International Conference on Computer Science and Education (ICCSE)* (pp. 362-365).

Chen, Z., Shao, Z., Xie, Z., & Huang, X. (2010). An Attribute-Based Scheme for Service Recommendation Using Association Rules and Ant Colony Algorithm. *Wireless Telecommunications Symposium (WTS), 2010* (pp. 1-6).

APPENDICIES

APPENDIX A

SAMPLE USER CONTEXT HISTORY USED IN TESTS

Date	Day	Time	Location	Weather	Companion	Relationship	Activity
04.01.2011	Tuesday	12.00-14.00	Kentpark	Snowy	Süleyman	Collegue	Restaurant
08.01.2011	Saturday	12.00-14.00	Kentpark	Rainy	Gaye	Partner/Spouse	Restaurant
08.01.2011	Saturday	14.00-16.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cinema
08.01.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cafe
09.01.2011	Sunday	14.00-16.00	Bahceli	Snowy	Gaye	Partner/Spouse	Cafe
09.01.2011	Sunday	16.00-18.00	Bahceli	Snowy	Özge	Friend	Cafe
09.01.2011	Sunday	18.00-20.00	Bahceli	Snowy	Özge	Friend	Fast-Food
13.01.2011	Thursday	12.00-14.00	Kentpark	Cloudy	Oguzhan	Collegue	Restaurant
14.01.2011	Friday	18.00-20.00	Ankuva	Cloudy	Davut	Friend	Fast-Food
14.01.2011	Friday	20.00-22.00	Ankuva	Cloudy	Davut	Friend	Bowling
15.01.2011	Saturday	14.00-16.00	Kentpark	Snowy	Gaye	Partner/Spouse	Restaurant
15.01.2011	Saturday	16.00-18.00	Kentpark	Snowy	Gaye	Partner/Spouse	Cafe
16.01.2011	Sunday	14.00-16.00	Bahceli	Cloudy	Özge	Friend	Cafe
16.01.2011	Sunday	16.00-18.00	Bahceli	Cloudy	Özge	Friend	Cafe
20.01.2011	Thursday	18.00-20.00	Cukurambar	Snowy	Gaye	Partner/Spouse	Restaurant
20.01.2011	Thursday	20.00-22.00	Bilkent	Snowy	Gaye	Partner/Spouse	Cafe
22.01.2011	Saturday	14.00-16.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cinema
22.01.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cafe
23.01.2011	Sunday	16.00-18.00	Bahceli	Rainy	Davut	Friend	Cafe
23.01.2011	Sunday	18.00-20.00	Bahceli	Rainy	Davut	Friend	Fast-Food
26.01.2011	Wednesday	12.00-	Cepa	Cloudy	Süleyman	Collegue	Restaurant

		14.00					
29.01.2011	Saturday	12.00-14.00	Kentpark	Snowy	Gaye	Partner/Spouse	Restaurant
29.01.2011	Saturday	14.00-16.00	Cukurambar	Snowy	Gaye	Partner/Spouse	Cafe
29.01.2011	Saturday	16.00-18.00	Cukurambar	Snowy	Gaye	Partner/Spouse	Cafe
30.01.2011	Sunday	14.00-16.00	Bilkent	Snowy	Gökcen	Friend	Cafe
30.01.2011	Sunday	16.00-18.00	Ankuva	Snowy	Gökcen	Friend	Bowling
01.02.2011	Tuesday	12.00-14.00	Cepa	Cloudy	Süleyman	Collegue	Fast-Food
05.02.2011	Saturday	14.00-16.00	Cepa	Snowy	Gaye	Partner/Spouse	Restaurant
05.02.2011	Saturday	14.00-16.00	Cepa	Snowy	Gaye	Partner/Spouse	Cafe
05.02.2011	Saturday	16.00-18.00	Kentpark	Snowy	Gaye	Partner/Spouse	Cinema
06.02.2011	Sunday	14.00-16.00	Bahceli	Snowy	Gaye	Partner/Spouse	Cafe
06.02.2011	Sunday	16.00-18.00	Bahceli	Snowy	Gaye	Partner/Spouse	Cafe
11.02.2011	Friday	12.00-14.00	Kentpark	Cloudy	Oguzhan	Collegue	Fast-Food
12.02.2011	Saturday	14.00-16.00	Kentpark	Cloudy	Gaye	Partner/Spouse	Restaurant
12.02.2011	Saturday	16.00-18.00	Kentpark	Cloudy	Gaye	Partner/Spouse	Cafe
14.02.2011	Tuesday	18.00-20.00	Cukurambar	Snowy	Gaye	Partner/Spouse	Restaurant
14.02.2011	Tuesday	20.00-22.00	Cukurambar	Snowy	Gaye	Partner/Spouse	Cafe
19.02.2011	Saturday	12.00-14.00	Cepa	Rainy	Özge	Friend	Fast-Food
19.02.2011	Saturday	14.00-16.00	Kentpark	Rainy	Özge	Friend	Cinema
19.02.2011	Saturday	16.00-18.00	Bilkent	Rainy	Özge	Friend	Cafe
20.02.2011	Sunday	14.00-16.00	Bahceli	Rainy	Gaye	Partner/Spouse	Restaurant
20.02.2011	Sunday	16.00-18.00	Bahceli	Rainy	Gaye	Partner/Spouse	Cafe
23.02.2011	Wednesday	12.00-14.00	Cepa	Cloudy	Süleyman	Collegue	Fast-Food
24.02.2011	Thursday	18.00-20.00	Bahceli	Rainy	Davut	Friend	Fast-Food
24.02.2011	Thursday	20.00-22.00	Bahceli	Rainy	Davut	Friend	Cafe
26.02.2011	Saturday	12.00-14.00	Kentpark	Snowy	Gaye	Partner/Spouse	Restaurant
26.02.2011	Saturday	14.00-16.00	Kentpark	Snowy	Gaye	Partner/Spouse	Cafe
26.02.2011	Saturday	16.00-18.00	Bilkent	Snowy	Gaye	Partner/Spouse	Cinema
01.03.2011	Tuesday	16.00-18.00	Bahceli	Cloudy	Erhan	Friend	Cafe
01.03.2011	Tuesday	18.00-20.00	Bahceli	Cloudy	Erhan	Friend	Restaurant
05.03.2011	Saturday	14.00-16.00	Kentpark	Snowy	Gaye	Partner/Spouse	Fast-Food
05.03.2011	Saturday	16.00-18.00	Kentpark	Snowy	Gaye	Partner/Spouse	Cafe

06.03.2011	Sunday	14.00-16.00	Bahceli	Snowy	Gaye	Partner/Spouse	Cafe
06.03.2011	Sunday	16.00-18.00	Bahceli	Snowy	Gaye	Partner/Spouse	Cafe
09.03.2011	Wednesday	12.00-14.00	Kentpark	Cloudy	Süleyman	Collegue	Restaurant
12.03.2011	Saturday	14.00-16.00	Kentpark	Rainy	Gaye	Partner/Spouse	Restaurant
12.03.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cinema
13.03.2011	Sunday	14.00-16.00	Bahceli	Sunny	Özge	Friend	Fast-Food
13.03.2011	Sunday	16.00-18.00	Bahceli	Sunny	Özge	Friend	Cafe
15.03.2011	Tuesday	16.00-18.00	Bahceli	Sunny	Gaye	Partner/Spouse	Cafe
19.03.2011	Saturday	12.00-14.00	ODTU	Sunny	Gaye	Partner/Spouse	Restaurant
19.03.2011	Saturday	14.00-16.00	Bilkent	Sunny	Gaye	Partner/Spouse	Cafe
19.03.2011	Saturday	16.00-18.00	Bilkent	Sunny	Gaye	Partner/Spouse	Cafe
21.03.2011	Monday	12.00-14.00	Cepa	Cloudy	Oguzhan	Collegue	Fast-Food
24.03.2011	Thursday	18.00-20.00	Kentpark	Rainy	Davut	Friend	Restaurant
24.03.2011	Thursday	20.00-22.00	Kentpark	Rainy	Davut	Friend	Cinema
27.03.2011	Saturday	14.00-16.00	Bilkent	Sunny	Gaye	Partner/Spouse	Cafe
27.03.2011	Saturday	16.00-18.00	Bilkent	Sunny	Gaye	Partner/Spouse	Cafe
28.03.2011	Sunday	14.00-16.00	Bahceli	Sunny	Özge	Friend	Fast-Food
28.03.2011	Sunday	16.00-18.00	Bahceli	Sunny	Özge	Friend	Cafe
04.04.2011	Wednesday	18.00-20.00	Cukurambar	Clear	Gaye	Partner/Spouse	Restaurant
04.04.2011	Wednesday	20.00-22.00	Bilkent	Clear	Gaye	Partner/Spouse	Cafe
08.04.2011	Friday	12.00-14.00	Cepa	Cloudy	Süleyman	Collegue	Fast-Food
09.04.2011	Saturday	14.00-16.00	Kentpark	Rainy	Gaye	Partner/Spouse	Restaurant
09.04.2011	Saturday	16.00-18.00	Kentpark	Rainy	Gaye	Partner/Spouse	Cinema
10.04.2011	Sunday	14.00-16.00	Bilkent	Cloudy	Gökcen	Friend	Cafe
10.04.2011	Sunday	16.00-18.00	Ankuva	Rainy	Gökcen	Friend	Bowling
14.04.2011	Thursday	12.00-14.00	Kentpark	Rainy	Süleyman	Collegue	Restaurant
16.04.2011	Saturday	14.00-16.00	Bilkent	Sunny	Gaye	Partner/Spouse	Cafe
16.04.2011	Saturday	16.00-18.00	Bilkent	Sunny	Gaye	Partner/Spouse	Cafe
17.04.2011	Sunday	10.00-12.00	ODTU	Sunny	Gaye	Partner/Spouse	Cafe
17.04.2011	Sunday	12.00-14.00	ODTU	Sunny	Gaye	Partner/Spouse	Cafe
17.04.2011	Sunday	14.00-16.00	Kentpark	Sunny	Gaye	Partner/Spouse	Cinema
19.04.2011	Tuesday	12.00-	Cepa	Cloudy	Oguzhan	Collegue	Fast-Food

		14.00					
20.04.2011	Wednesday	16.00-18.00	ODTU	Sunny	Özge	Friend	Cafe
23.04.2011	Saturday	14.00-16.00	ODTU	Sunny	Gaye	Partner/Spouse	Restaurant
23.04.2011	Saturday	16.00-18.00	ODTU	Sunny	Gaye	Partner/Spouse	Cafe
24.04.2011	Sunday	10.00-12.00	ODTU	Sunny	Davut	Friend	Cafe
26.04.2011	Tuesday	12.00-14.00	Kentpark	Cloudy	Süleyman	Collegue	Restaurant
28.04.2011	Thursday	18.00-20.00	ODTU	Sunny	Davut	Friend	Cafe