CONTEXTUAL PERSONALIZED RECOMMENDATION TECHNIQUE FOR ACADEMIC EVENT SELECTION

PARUKUTTY RAMAN

FACULTY OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2018
CONTEXTUAL PERSONALIZED RECOMMENDATION TECHNIQUE FOR ACADEMIC EVENT SELECTION

PARUKUTTY RAMAN

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER

FACULTY OF COMPUTER SCIENCE & INFORMATION TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2018
UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Parukutty Raman
Matric No: WGA120008
Name of Degree: Master of Computer Science
Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):
CONTEXTUAL PERSONALIZED RECOMMENDATION TECHNIQUE FOR
ACADEMIC EVENT SELECTION
Field of Study: INFORMATION SCIENCE

I do solemnly and sincerely declare that:
(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes, and any excerpt or extract from or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently, and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every right to the copyright of this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright, whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Signature of Candidate Date:

Subscribed and solemnly declared before,

Signature of Witness Date:

Name:
Designation:
ABSTRACT

A recommender system in the academic domain has significant importance to researchers. Incorporating contextual information in a recommender system is vital to ensure that the recommended information is relevant and in accordance to the user’s preferences. Consideration of “context” in the selection of an academic event is likely to have a profound effect on retrieving better recommendation results. Academic events can be classified as events related to any academic domain discipline, such as workshop, seminar or conference. A Recommendation System (RS), using a classical filtering approach, tends to fail when insufficient user preference information is available. The two main classical approaches are content based and collaborative. The objective of this study is to identify the most important contexts in selecting an academic event and develop a contextual personalized recommender technique for academic event selection. A survey is conducted to identify the most important contexts. Four important contexts of time, schedule, location and cost were identified and used in the technique development. Next, a tool was developed using context pre-filtering and collaborative searching techniques. The context will be pre-filtered and a search input that carries contextual data and keywords will be used to search for relevant events using a match matrix from the event database. The same events will then be checked again to see whether they have been attended by any neighbour of the user using the Top N weighted nearest neighbour technique. Contexts and keywords are explicitly given by the user. Average precision and mean average precision are used to evaluate the tool. The results will show that contextual personalized event selection techniques produce more relevant results than a technique that only uses classical approaches. This study proposes a context based personalized recommender technique to assist researchers in finding relevant academic events. The developed technique can also be used in other domains.
ABSTRAK

# TABLE OF CONTENTS

Abstract ............................................................................................................................................... iii

Abstrak ................................................................................................................................................ iv

Table of Contents ................................................................................................................................. v

List of Figures ........................................................................................................................................ ix

List of Tables .......................................................................................................................................... x

List of Symbols and Abbreviations .................................................................................................... xi

List of Appendices .............................................................................................................................. xii

## CHAPTER 1: INTRODUCTION ........................................................................................................ 13

1.1 Introduction .................................................................................................................................... 13

1.2 Research Motivation ....................................................................................................................... 14

1.3 Problem Statement ........................................................................................................................ 15

1.4 Research Objectives ...................................................................................................................... 16

1.5 Research Questions ....................................................................................................................... 16

1.6 Research Methodology .................................................................................................................. 16

1.7 Research Scope .............................................................................................................................. 18

1.8 Thesis Outline ............................................................................................................................... 18

## CHAPTER 2: LITERATURE REVIEW ............................................................................................ 20

2.1 Introduction .................................................................................................................................... 20

2.2 Academic Events .......................................................................................................................... 20

2.3 Recommender System .................................................................................................................. 21

   2.3.1 Content-Based Filtering Technique ....................................................................................... 23

   2.3.2 Collaborative Filtering Technique ......................................................................................... 24
2.3.3 Hybrid Filtering Technique .......................................................... 25
2.4 Context-Aware Recommender System .............................................. 26
  2.4.1 Contextual Taxonomy ................................................................. 27
  2.4.2 Contextual Filtering Techniques ................................................ 28
  2.4.3 Contextual RS in other Domains ............................................... 29
2.5 Academic Event Recommender Systems ......................................... 31
  2.5.1 Existing Tools and its Limitations .............................................. 32
2.6 Summary ....................................................................................... 34

CHAPTER 3: RESEARCH METHODOLOGY ................................................. 35
3.1 Introduction .................................................................................... 35
3.2 Requirement Planning Phase ......................................................... 35
  3.2.1 Information Gathering ............................................................... 36
    3.2.1.1 Sampling ........................................................................... 36
    3.2.1.2 Questionnaire ................................................................. 36
    3.2.1.3 Data collection ............................................................... 38
  3.2.2 Data Analysis ............................................................................ 38
3.3 User Design Phase .......................................................................... 40
  3.3.1 Requirements Analysis .............................................................. 40
  3.3.2 Identification of the technique .................................................. 41
  3.3.3 Development of Technique ...................................................... 42
3.4 Construction Phase .......................................................................... 43
  3.4.1 Development of Tool ............................................................... 43
3.5 Cut-Over – Testing Phase ................................................................. 43
  3.5.1 Evaluation of the Technique and Tool ...................................... 43
3.6 Summary........................................................................................................................................44

CHAPTER 4: TECHNIQUE AND TOOL DEVELOPMENT .................................................46
4.1 Introduction....................................................................................................................................46
4.2 Contextual Personalized Recommender Technique .................................................................46
4.3 Contextual Pre-Filtering Process ...............................................................................................48
4.4 Proposed Technique ....................................................................................................................50
4.5 Summary.......................................................................................................................................52

CHAPTER 5: EVALUATION OF THE TECHNIQUE .........................................................53
5.1 Introduction....................................................................................................................................53
5.2 Findings .......................................................................................................................................53
5.3 Experiment Result .......................................................................................................................54
5.4 Evaluation Results .......................................................................................................................57
5.5 Summary.......................................................................................................................................57

CHAPTER 6: CONCLUSION AND FUTURE WORK .......................................................59
6.1 Research Findings .......................................................................................................................59
6.2 Significance of the Study ............................................................................................................61
6.3 Limitations and Future Work ......................................................................................................61
6.4 Conclusion ...................................................................................................................................62
REFERENCES......................................................................................................................................63
BIBLIOGRAPHY.................................................................................................................................67
Appendix A: SURVEY QUESTIONNAIRE ...................................................................................78
Appendix B: SURVEY RESPONSE ...............................................................................................81
LIST OF FIGURES

Figure 1.1 Research Methodology Phases ................................................................. 18
Figure 2.1 Recommender System Filtering Technique Architecture ......................... 23
Figure 2.2 Contextual Information Hierarchy ............................................................. 26
Figure 3.1 Important Contexts in Choosing an Academic Event .................................. 39
Figure 3.2 Overall Architecture of the Contextual Personalized Tool ......................... 42
Figure 4.1 Contextualization Process using Contextual Pre-Filtering ......................... 48
Figure 4.2 Contexts Hierarchy .................................................................................... 48
Figure 4.3 Input, Process and Output for Proposed Technique .................................... 51
Figure 4.4 Pseudo Code of the Proposed Technique .................................................... 52
Figure 5.1 Pilot Study of Technique and Tool Evaluation ............................................ 54
LIST OF TABLES

Table 2.1 Contextual Taxonomy .............................................................................................. 27
Table 2.2 Comparison Chart of Existing Tools. .................................................................... 33
Table 3.1 Survey Questionnaire ............................................................................................ 37
Table 4.1 Example of User Preferences Distribution ............................................................... 49
Table 5.1 Experimental results for precision calculation ......................................................... 55
Table 5.2 Experimental results for recall calculation .............................................................. 56
### LIST OF SYMBOLS AND ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS</td>
<td>Recommender system</td>
</tr>
<tr>
<td>CARS</td>
<td>Context-aware recommender system</td>
</tr>
<tr>
<td>LCARS</td>
<td>Location context-aware recommender system</td>
</tr>
<tr>
<td>CF</td>
<td>Collaborative Filtering</td>
</tr>
<tr>
<td>CB</td>
<td>Content Based</td>
</tr>
<tr>
<td>KNN</td>
<td>K Nearest Neighbour</td>
</tr>
<tr>
<td>AP</td>
<td>Average Precision</td>
</tr>
<tr>
<td>RQ</td>
<td>Research Question</td>
</tr>
<tr>
<td>RO</td>
<td>Research Objective</td>
</tr>
<tr>
<td>CM</td>
<td>Context Modelling</td>
</tr>
<tr>
<td>Appendix</td>
<td>Pages</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------</td>
</tr>
<tr>
<td>Appendix A: Survey Questionnaire</td>
<td>68</td>
</tr>
<tr>
<td>Appendix B: Survey Response</td>
<td>71</td>
</tr>
<tr>
<td>Appendix C: Survey Response Charts</td>
<td>90</td>
</tr>
<tr>
<td>Appendix D: Programming Code</td>
<td>95</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION

1.1 Introduction

The current information system which has emerged over the years has to deal with a huge amount of data, thereby leading to the delivery of massive results. Thus, one solution to this problem is recommender system that can filter and personalize data according to the needs of users (Pommeranz, Broekens, Wiggers, Brinkman & Jonker, 2012). A recommender system (RS) is an information filtering system that deals with the problem of information overload in current information systems (Pham, Kovachev, Cao, Mbogos & Klamma, 2012). One can use RS technology to predict the preference ratings of items that no user has currently rated, and/or to output a personalised item ranking that will likely be interesting to the user (Ricci, Rokach, Shapira & Kantor, 2011). It uses three main classical filtering techniques, namely content-based, collaborative-based and hybrid techniques, to filter the information and to retrieve relevant information. A content-based filtering technique uses the past history of the user to recommend an item, while a collaborative-based technique looks for a neighbour that has similar interests with the user to recommend items. A hybrid filtering technique is a mixture of the two techniques mentioned above. Theoretically, a collaborative filtering technique uses two main methods, which are model-based and memory-based, to filter.

A recent trend is the use of a context-aware RS as a component in a computing paradigm (Suvarna & Menezes, 2016). This use has been rapidly increasing in many domains over the years to improve the performance of the RS. According to Dey (2001), context is defined as any information that can be utilised in the characterisation of the situation of an entity, which can either be a place, person, or object. Furthermore, this characterisation is considered to possess a relevance to the interaction between an
application and the user, including the actual applications and the user themselves. Ranganathan and Campbell (2003) state that context refers to any information about the objects, circumstances, or conditions that surround a user and that is considered to have relevance to the interaction between the ubiquitous computing environment and the user. The inclusion of the context in the recommendation engine may improve the performance under certain circumstances (G Adomavicius & Tuzhilin, 2005).

Academic event refers to events that academicians academically organise not just to share ideas and knowledge but also to collaborate (Pham et al., 2012). It refers to workshops, seminars, and other educational events that have one or more subject matter experts sharing information mainly via discussion and lecture. Recommendations with regard to the selection of academic events play an important role in the retrieval of the most important events for students to attend. Furthermore, the context of students in the selection of academic events is important in ensuring that the most relevant events are recommended to students for them to attend. In this study, the important contexts concerning the student’s choice of an academic event were identified, followed by the development and evaluation of a tool incorporating contextual elements that are relevant to the evaluation matrix.

1.2 Research Motivation

Recommendations for academic events should be based on user preferences. This will ensure that students find the most relevant academic events according to their preferences, rather than making random selections. It is vital for students to find the most relevant events to gain the appropriate knowledge for their studies and to improve their skillset. While a substantial amount of research has already been performed in the area of recommender systems, most of the existing classical approaches focus on recommending
the most relevant items to users without taking into account any additional contextual information, such as time, location, and interest (Zhou, Luo, & Xu, 2012).

In order to solve some of the problems encountered in classical recommender techniques, a contextual RS has attracted the attention of academicians (Liu, Haohan, Zhang, Hui & He, 2015). Contextual data help to produce better recommendations by taking into count user preferences and other contexts that may affect a user’s choice of an item. In the selection of academic events, each student has his/her own preferences, and it is essential that these preferences (contexts), which affect the selection of academic events, be taken into considerations. The consideration of contexts in event selection will have a positive impact on the recommendation result.

1.3 Problem Statement

Each student may have their own preferences for attending an academic event, and this has to be identified accordingly. Nonetheless, students find it hard to determine the academic event that is the most relevant (Xia et al., 2013). When the wrong academic event is selected, the student will be not able to obtain the correct knowledge or skill that is relevant to their field of study and interest. The creation of recommendations for users in an academic domain to cater to their needs and tasks requires a deeper analysis of the contextual information that affects decision-making in this domain (Champiri, Shahamiri & Salim, 2015). Similarly, an RS that uses classical filtering approaches tends to fail when little knowledge about the user is known, and contextual factors in the academic domain have not been explored much over the years.
1.4 Research Objectives

The primary goal of this study was to retrieve more relevant academic events for students to attend by taking into consideration the contextual data that have an impact on the selection of academic events. The contextual personalized recommender technique was aimed at achieving the following research objectives:

1. To identify the user context for the selection of academic events.
2. To develop a contextual personalised recommender technique.
3. To evaluate the developed technique.

1.5 Research Questions

The research questions for this study were aimed at determining the relevance of the technique used to find academic events, as described below. The research questions were based on the research objectives mentioned below.

Objective 1: To identify the user context for the selection of academic events
Research Question: What are the important contexts for the selection of academic events?

Objective 2: To develop a contextual personalized recommender technique
Research Question: How can a system be developed using a contextual personalised technique?

Objective 3: To evaluate the developed technique
Research Question: How can the developed technique be evaluated?

1.6 Research Methodology

Three main methodological phases were involved in this study, namely, the identification of requirements, the development of a technique and tool, and the
evaluation phase. The technique and tool development process make use of the Rapid Application Development (RAD) methodology. This methodology is used to develop the tool with the use of its four phases - requirement planning phase, construction phase, user design phase, and cutover phase. The methodology helped this study to quickly adapt to changes and to develop the tool and technique at a faster pace.

In the first phase, the requirement planning focused on the gathering of information using online surveys, followed by data extraction to derive the relevant information for this study. In the first phase of the RAD, the information provided by the earlier phase was analysed to help in the development of the technique. In the user design phase, all the information was further compiled for integration into a design. The development of the technique was carried out in this phase.

In the construction phase, the developed technique was incorporated into the tool development. All the developments were completed in this phase, where the next phase, known as the cutover phase, involved the testing of the tool. The functionality of the tool and technique tested and evaluated in this study using precision and recall measurements. The evaluation and comparison of this tool against the classical approach were performed in this phase as well. Detailed evaluation is explained later in chapter 5.
1.7 **Research Scope**

This study focused mainly on students of the University of Malaya, and covered academic events that were conducted for them within or outside the university. Thus, it had certain limitations with regard to the primary goals of this study.

1. It only covered the selection of academic events by students within the University of Malaya
2. The retrieval of events was based on the academic events conducted at the University of Malaya.
3. Only students of the University of Malaya were chosen for the survey.

1.8 **Thesis Outline**

This study is comprised of six chapters. Chapter 1 presents an overall discussion of the recommender system (RS), the context-aware RS, and the academic event selection. It further goes through the research objectives, questions, scope and research motivation. Chapter 2 continues with the literature on previous work with regard to recommender systems, contextual recommender systems, and contextual recommender systems for the
selection of academic events. Furthermore, it also discusses the existing tools and explores their limitations in previous studies to support the proposed technique. Chapter 3 focuses fully on the methodology used to gather the information, and to develop and evaluate the technique. From there, Chapter 4 explains how the tool and technique were developed. The evaluation of the technique is clearly explained in Chapter 5, with a comparison of the experimental data and the results in order to discuss the technique measurements. The data from the experiment were examined and presented to analyse the contextualized technique. The last chapter, Chapter 6, presents the conclusion, a deliberation about the contribution of this study, and future work.
CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

Academic events are basically events organized by stakeholders in the academic field. The primary purpose of these events is to share knowledge and to collaborate with other students. Academic events can be seminars, workshops, technical talks, upskill programs, and other academic-related programs. Academic events play an important role in providing information about study strategies, tools or skills. Undoubtedly, they help students to enhance their skills academically and to apply the information in their academic domain. Furthermore, academic events also enhance academic profiles. For example, students joining thesis writing workshops tend to have an edge over those who do not because they learn skills about how to write an effective thesis. As this study focused more on a recommendation technique to get students to see which the more relevant academic events to attend are, this chapter discusses the previous work done with regard to academic events, recommender systems, and contextual recommender systems. Additionally, it also discusses the limitations of previous works and the existing tools that are available for finding academic events.

2.2 Academic Events

As mentioned above, academic events are crucial for academicians to share, learn and apply the knowledge gained to enhance, improve their skills and to collaborate effectively with other students (Xia et al., 2013). Thus, selecting an academic event is very important because each student has his/her own field of study and acquires a different skill set for self-improvement (Xia et al., 2013). In order to get the right knowledge, it is vital for students, especially postgraduates, to select the most accurate or relevant academic event according to their academic domain. Failure to do so will eventually become a drawback,
where they will be lacking in the skill set and strategies necessary for their studies. On the other hand, the large amount of data being recommended to students may cause them to miss the relevant academic events. Students find it difficult to do the selection manually by reviewing the events one by one to see if it fits their preference and also whether it is related to their academic domain.

2.3 Recommender System

A recommender system (RS) is one way to help users deal with the flood of information (Huynh, Luong & Hoang, 2012). With the large amount of data being distributed across the Internet for any field, especially the academic field, it is crucial to have an information filtering technique for a more organized retrieval of suitable information. As previously mentioned, RS helps manage vast amounts of data. It refers to a set of technologies that are utilised in sorting goods or information from a complete information source that a user might find interesting. For example, a well-known website like Amazon.com may recommend a huge number of items to users, where it would be impossible to show all of them at a glance to the user. So, when selecting a set of items to be shown to a user, it is very important to ensure that it is something that is preferred by or is of interest of the user. Different individuals have different preferences. Hence, the system must be able to identify specific or almost similar items that are of interest to the user and are more likely to be relevant to all users. A recommender system does just that. It creates a list of items that are likely to be of interest to each individual user. In summary, “any system that produces individualized recommendations as an output or has the effect of guiding the user in a personalized way” can be defined as an RS (Sieg, Mobasher & Burke, 2007).
RS is deemed as a system that is able to provide personalized recommendations through different types of algorithms (usually termed ‘filtering’). RS also can be described as a predicting system that work based on the interaction between a consumer (and business) with particular type of product and services (Bobadilla, Ortega, Hernando, Gutiérrez, 2013). It captures the taste of the specific customer based on their actions (comments, number of views, ratings) towards a particular product.

The main two units in the RS are the user and product (Madhusree, 2016). The users may share their view, opinion and experience about a product by utilizing the RS system well. Thus they able to receive recommendations about the new similar items based on their interest from the system as well. The recommendation is actually the unique preference of each customer towards the products in the system.

A successful recommender system has several conditions to be met. The primary element is requiring to contain with many items and offers under each domain. Next, the system requires to provide taste based choices to the customer and not limited for one solution. Subsequently, the system requires to equip interactive platform for the customer to review and express their view or experience about the product. Forth, the products under each domain must be homogenous so that all of the products can be covered or displayed under a taste.

An RS helps to retrieve relevant items using three main filtering approaches, namely collaborative filtering (CF), content-based (CB) and hybrid filtering. CF and CB are the classical approaches used in RS. CF performs recommendations based on the similar interests of other users, while the content-based approach uses the user’s history or the features of items (Klamma, Cuong, & Cao, 2009). Hybrid approaches achieve a better
recommendation result by combining CF and content based with few other techniques. The figure below shows the architecture of recommender system techniques and the collaborative filtering categories.

![Figure 2.1 Recommender System Filtering Technique Architecture](image)

2.3.1 Content-Based Filtering Technique

The concealed principle of the content-based (CB) filtering technique is where recommendation systems analyse item descriptions to identify those items that are of particular interest to the user (Pazzani & Billsus, 2007). A classical CB recommendation system recommends items to a user based completely on the past history of that particular user, and a profile is constructed by analysing the items that are of interest or have been purchased by this user. Examples of such systems are InfoFinder, NewsWeeder and many
more. A classical CB system has several specific drawbacks. The use of only the past history of the user may lead to cold start problems when the user does not have any past history to be analysed and to recommend items (Boehmer, Jung, & Wash, 2015).

2.3.2 Collaborative Filtering Technique

Collaborative filtering (CF) refers to a RS technique that is commonly utilised in commercial applications. Recommendations are provided by CF based on the previous preferences of the user and the opinions that other users who possess similar preferences have (Breese, 1998). User preferences can be obtained through implicit and explicit methods, where an implicit method interprets a user’s behaviour by using the user’s history of purchases or data browsing and so on. An explicit method uses the user’s rating for an item or the explicit information given by the user to recommend items. As shown in Figure 2.1 above, a CF filtering algorithm can be categorized as memory-based or model-based. Memory-based collaborative filtering algorithms generate the recommendations by operating on the entire user-item database; model-based collaborative filtering algorithms on the other hand make use of the user database for learning a model that it can then use for a recommendation. A user-based CF mainly uses the similarity between users (neighbours) to make recommendations, where it works entirely on a user-item-based matrix. This system uses a memory-based algorithm such as the KNN algorithm. According to Zohreh (2014), among the classical filtering methods used in RS, the usage of CF is high across the domains.

The CF approach is in contrast to the content-based (CB) technique, where it uses the similarity of the user’s likes to predict and recommend items. For this, each user will have a “nearest neighbour” with whom this user has the strongest similarity of purchasing history or search history. The strongest correlation helps to choose the neighbour. If there
is an unseen item score, it will use a combination of items to identify the nearest neighbour. It can be said that the classical collaborative method was solely developed on the basis of similarities to other users with the same interest. Examples of systems that are implementing this approach include the Bellcore video recommender, GroupLens, and many more. It has been observed that the classical CF method solves the drawbacks found in the CB method. The cold start problem can basically be identified, where the history of other users can be used to recommend items for those who do not have their own past history. However, the CF has its own shortcomings. It will be difficult for a newer item that has been added to the database to be recommended to a user until and unless someone has searched or purchased it, and then only will it appear to other users. Therefore, there may be sparseness issues if the number of users is relatively small compared to the volume of information within the system. Furthermore, a database that is very large or rapidly changing causes the coverage of ratings to become very sparse, which in turn scatters the gathering of recommendable items (Salman, 2016). Secondly, not all users have the same interest or like what others see. Hence, this will lead to poor recommendation results (Ricci, Bontcheva, Conlan, Lawless, 2015).

2.3.3 Hybrid Filtering Technique

The hybrid filtering technique combines both the content-based and collaborative techniques to provide recommendations (Liu, Haohan et al., 2015). A hybrid of the recommendation algorithm integrates other recommendation algorithms into one to recommend items to a user. Collaborative filtering (CF) and content-based (CB) technique algorithms are often combined to create a hybrid recommender system to solve the limitations found in both the CB and CF methods (Salman, 2016).
2.4  Context-Aware Recommender System

In dealing with the limitations of classical recommender methods, a context-aware recommender system (CARS) has captured the attention of academicians (Liu, Haoan et al., 2015). CARS is widely used in many domains and has been undertaken by many researchers over the years. Thus, the information for identifying the contexts used in CARS is indispensable and is key to understanding CARS in multi domains, which will eventually help in developing the technique. Papers on CARS have been identified, and the contexts used in each paper have been extracted to construct the contextual information. Using the review done by Champiri, Shahamiri & Salim (2015) as a reference, a hierarchy of contextual information was constructed. Most of the contexts used in CARS come under any of these categories.

![Contextual Information Hierarchy](image)

Figure 2.2 Contextual Information Hierarchy

26
2.4.1 Contextual Taxonomy

Past studies of CARS were reviewed across multiple domains to explore the contextual information affecting efficiency of recommendations. The contextual information was then sorted into each category, as shown in Figure 2.2, to construct a detailed contextual taxonomy. According to Table 2.1, the contextual conditions of the user and environment are mostly used in CARS. A common ‘context’ of interest in many recommendation systems is location. Other contexts, such as documents and citations, are also used in the academic domain to be incorporated in recommender systems to improve the results.

Table 2.1 Contextual Taxonomy

<table>
<thead>
<tr>
<th>Contextual Type</th>
<th>Features</th>
<th>Conditions</th>
<th>References</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profile</td>
<td>Name, Age, Gender, Email, Occupation, Contact No., Identity No., Fax No., Address, Marital Status, Education, Income, Faculty, etc.</td>
<td>(Guo &amp; Lu, 2015; Herzog &amp; Wolfgang, 2016; Kanetkar, Nayak, Swamy &amp; Bhatta, 2014; M. Li, Sagl, Mbura &amp; Fan, 2016; Miyazawa, Yamamoto &amp; Kawabe, 2013; Park, Yoo &amp; Cho, 2006; Uddin, Banerjee &amp; Lee, 2016; Xia et al., 2013; Yin et al., 2013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preferences</td>
<td>Session Preferences, Domain Preferences, Speaker Preferences, Topic Preferences, Genre Preferences, Short/Long time Preferences, Budget Preferences, etc.</td>
<td>(Etaati &amp; Sundaram, 2015; Gupta &amp; Singh, 2013; Herzog &amp; Wolfgang, 2016; Kanetkar et al., 2014; Uddin et al., 2016; Xia et al., 2013; Yin et al., 2013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>USER Behaviour</td>
<td>Very keen, just seeking information, potential buyer, dispute customer, mental state, assumptions, learning personalities, health situation, user searching behaviour, user browsing behaviour, etc.</td>
<td>(Bourkoukou, Bachari &amp; Adnani, 2016; Dehghani, Afshar, Jamali &amp; Nematyakhsh, 2011; Etaati &amp; Sundaram, 2015; Guo &amp; Lu, 2015; Li et al., 2016; Mobasher, Dai, Luo &amp; Nakagawa, 2002; Sarkaleh, 2012; Sieg et al., 2007; Uddin et al., 2016; Yin et al., 2013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Types</td>
<td>Student, researcher, postgraduate, graduate, auditor, user, purchaser, traveller, tour agent, musician, listener, etc.</td>
<td>(Dehghani et al., 2011; Etaati &amp; Sundaram, 2015; Herzog &amp; Wolfgang, 2016; Miyazawa et al., 2013; Sieg et al., 2007; Xia et al., 2013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge</td>
<td>Very Basic, Basic, Proficiency, Specialty, Expertise</td>
<td>(Dehghani et al., 2011; Li et al., 2014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Relationship</td>
<td>Collaboration with other users, work team, co-author relationships &amp; connections between scholars, colleagues' work, Google Wave, Blog information/posts, friends, companionship, family, etc.</td>
<td>(Dehghani et al., 2011; Herzog &amp; Wolfgang, 2016; Li et al., 2014; M. C. Pham et al., 2012; X. H. Pham, Nguyen, Jung, &amp; Nguyen, 2014; Singh, Shubhankar, &amp; Pudi, 2012; Uddin et al., 2016; Zheng, 2016)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| ENVIRONMENT Location | User location, item location, services location, longitude, latitude | (Etaati & Sundaram, 2015; Gupta & Singh, 2013; Herzog & Wolfgang, 2016; Kanetkar et al., 2014; Madadiipouya, 2015; Narayan & Cherukuri, 2016; Pham et al., 2012; X. H. Pham, Jung, Vu
### Contextual Filtering Techniques

Adomavicius & Tuzhilin (2008; 2011) described three types of contextual representational approaches, namely contextual pre-filtering, post-filtering and contextual modelling methods as follows:

A. Contextual pre-filtering is when the contextual data is pre-processed or filtered by removing the irrelevant ratings before they are used for computing recommendations with standard (non-contextual) methods.

B. Contextual post-filtering uses a post-contextual filtering process after non-contextual recommendation methods have been applied to the data.

C. Contextual modelling (CM) assumes that the contextual information is used within the recommendation-generating algorithms together with the user and item data.
2.4.3 Contextual RS in other Domains

As described above, contextualization is widely used in many domains such as movie recommendations, e-commerce, tourism, and book recommendations. Each of these domains works differently in contextualization because each has its unique contexts based on the user, environment and other factors. The contextual data used in movie recommendations cannot be used in book recommendations and vice versa. This is due to the fact that the decision factor in the selection of a movie and in the selection of a book may vary in terms of the domain. The preferences of the users also vary individually under the same domain. For example, if Ali and Amin are looking for movie recommendations, it does not mean that both of them have the same interest. Ali may like horror movies and Amin might prefer comedies. At the same time, both also might have the same preference to watch movies at night. Hence, contextualized filtering techniques are used here to gather the preferences of these people separately and, at the same time, to match their similar preferences to provide better recommendation results.

The emergence of the RS concept of “context” in computer science as well as other related fields has a significant influence on the development of RS. Slowly but gradually, contextual information is gaining the attention of recommender systems. Ranking models can be estimated and these models may be constructed using context as they include marginal and additional data about defined items and users. Contextual data are considered to be a vital source of precision in RSs (Baltrunas, 2008). Furthermore, several studies have shown that the performance of RS can be improved by utilising contextual data (Adomavicius & Tuzhilin, 2005; Gediminas Adomavicius & Tuzhilin, 2008; Panniello & Gorgoglione, 2010).
In this study, CARS from a previous work was studied to understand how contexts work across domains. The output was discussed in a previous session. According to Veras, Prota, Bispo, Prudencio, Ferraz (2015), the study focused on a combined cross-domain context-aware model for the recommendation of television shows and books. They used contextual pre- and post-filtering techniques to recommend the contextualized data. The data were clustered according to similarity across the domain before the contextualization took place. The study was limited to only one context, namely, location. Later, Dehghani (2011) did a study on a multi-layer contextual model for digital libraries which uses multiple contexts without considering the location. Using too many contexts also creates the problem of sparseness in an RS. Other than that, a location-aware RS for movies was developed in 2015 by Madadipouya, where a CF Pearson Correlation algorithm was incorporated with contextualized data to provide more effective movie recommendations to users. Location is a context that can be used across many domains, including the academic domain. Tourism is also one of the domains where contexts are being widely used, mainly because location is one of the most important contexts for tourism recommendation systems. A tourism RS based on location, mobile devices and user features was developed by Sarkaleh in 2012 to give better recommendation results. A content-based image retrieval RS was explored by Miyazawa in 2013 using implicit attributes. Another content-based music RS developed in 2006 using a Fuzzy Bayesian technique also uses mood as its context to recommend music but it is limited to content-based and also implicit attributes. It uses a dynamic context like mood and only a content-based approach for recommendations, which may lead to cold start issues.
2.5 Academic Event Recommender Systems

Quite a number of studies have been done in the academic domain, such as a scientific paper recommender system using a layered approach (Manouselis & Verbert, 2013). In addition, a social awareness recommender system using strong social ties between the presenter and the participant to recommend good presentation sessions was developed in 2014 by Asabere. It is a contextualized venue recommender system that can be used in academic event recommenders. However, there are at times good sessions occurring at two different venues at the same time, thereby causing issues with regard to which one to select. An RS using disjointed user/item sets was introduced by Hornick in 2012. However, this study derived limited recommendations when the registration survey was very shallow. Even though there are recommender systems that are being explored in the academic domain, a systematic review of scholar context-aware recommender systems in 2014 by Champiri confirmed that there is less incorporation of contextual data in recommender systems in the academic domain compared to other domains like e-commerce, movie recommendations, tourism and so on.

CB and CF techniques are considered as classical approaches to recommender systems, where the past history of the user or similarities in the user community are used to recommend items without taking into consideration user preferences and also the context of the environment of a user. According to Qiao, Zhang, Cao, Zhou & Guo (2014) in their study to improve collaborative recommender systems, those recommender systems that use classical approaches tend to fail when little knowledge about the user is known or no one has a similar interest to the user.
2.5.1 Existing Tools and its Limitations

There are a number of academic event recommendation tools available. Tools and methodologies developed for academic event management still faces issues (Klamma et al., 2009). On the other hand, this study focused on a personalized academic event recommender for postgraduate students from the UM, so that it concentrated on those events that are academic-based and that are also according to user preferences. There are many existing tools for recommending an event, however the most related tools are Acadevent (AcadEvent, 2017), Eventbrite and Allconferences (AllConferences, 2000).

One of the existing tools is Acadevent (AcadEvent, 2017). Acadevent uses classical (content and collaborative) filtering techniques to recommend events, and it does not consider user preferences and other environmental contexts. Users who register will be asked to provide only basic information, and they will able to use the tool to search for events. There are chances that the events recommended are not those that are preferred by the user since the user only uses a keyword to search. All the events that match the keyword will be retrieved regardless of whether it is based on the user’s preferences. This also causes the tool to retrieve a huge amount of data, and the user will have to manually review each event before choosing the most relevant one.

Another tool is Eventbrite (Eventbrite, 2006), which carries out some filtering with regard to the location and type of event. However, this tool does not take into consideration the preferences of different users. The filtered categories are general, and a large number of events are produced. This may lead to sparseness, where the user may need to spend extra hours to retrieve the event one by one before deciding if it is relevant or irrelevant according to his/her preferences.
Another existing academic event recommendation tool is Allconferences (AllConferences, 2000). Allconferences is a tool to search for conferences all over the world. However, the recommendations in this tool is very wide. Even though filtering can be done in this tool, it does not take into consideration of the user contexts. Thus will produce a vast results and led to sparseness. It produces a long list of event according to keyword and locations, but user still need to go and filter the results manually or check one by one, to see the more relevant academic events for them. This tool also does not include workshops, seminars and other small scaled academic events and only focus on big conferences all over the world. Table 2.2 is the comparison summary of all three tools explained above.

Table 2.2 Comparison Chart of Existing Tools.

<table>
<thead>
<tr>
<th></th>
<th>EventBrite</th>
<th>AcadEvent</th>
<th>AllConferences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Search</strong></td>
<td>Keywords and venue</td>
<td>Tagging and Keywords</td>
<td>Keywords and location</td>
</tr>
<tr>
<td>User Context</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Web Enable</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Search by Domain / Categories</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Search by User Preferences</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Notification of Upcoming event</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Search by Location</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Search by Time</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Limitation</strong></td>
<td>Sparseness and did not give consideration to user preference</td>
<td>Did not give consideration to user preference</td>
<td>Sparseness</td>
</tr>
</tbody>
</table>
2.6 Summary

In summary, based on previous studies, it can be clearly seen that context-aware recommendations have not been explored much in academic event recommender systems compared to other domains and classical approaches that tend to fail without having the contexts of the user and the environment to provide an effective recommendation. Hence, this study focused on a contextual personalized academic event recommender technique by taking into consideration the contexts of user preferences and location. Based on literature review, the user context and location play a vital role in deriving more relevant recommendations. Hence, later in the tool development, user contexts and location are given more important compared to other contextual information to get more relevant result.

Moreover, the collaborative filtering technique is chosen as the filtering technique as this study more focused on user and item based recommendation and the most relevant filtering is Collaborative filtering (CF). Nevertheless, CF alone is not good due to the limitations of CF where not all users have the same interest or like what others see and this may lead to poor recommendations. This is where contextual data proposed to be incorporated with CF technique to produce better result of recommendation and the evaluation of this study of classical approach (CF technique) and contextual approach (CF + context) will be producing the result of the comparison.
CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction

This chapter discusses the methodology of the study, which was implemented using rapid application development (RAD) methodology. Basically, RAD methodology is used for fast system development, and it is a tighter fit between user requirements and system specifications. It consists of four main phases, namely requirement planning phase, user design phase, construction phase, and cutover phase. RAD is known to provide the ability to swiftly develop an application or technique, and it is flexible in terms of modifications or changes to requirements without leading to a repetitive development environment. Dennis et al (2000) and Martin (1991) defined RAD as a development life cycle that is meant to offer higher quality results and faster development compared to the traditional life cycle. In the fourth phase of this study, called the cut-over phase, the technique and tool are evaluated and the results are constructed.

3.2 Requirement Planning Phase

The requirement planning phase is a phase in which the requirements of this study are gathered and pre-processed prior to the development of the technique and tool. In this phase, the requirements are identified. Since the main objective of this study was to develop a contextual technique for academic events, it was essential to identify the most important contexts of academic events that should be contextualized and used for the technique development. It was crucial to identify the contexts as these would carry their own weightage during the filtering technique, and have an impact on the resulting technique and tool.
3.2.1 Information Gathering

Information gathering can be done in many ways such as by observation, behaviour tracking, focus groups, self-assessments, interviews, surveys, case studies and existing data. Academic events are events that are mostly attended by postgraduate students to develop certain skills and to collaborate with other participants (Pham, Kovachev, Cao, Mbogos, & Klamma, 2012). Hence, postgraduates were selected as the target group for this study.

3.2.1.1 Sampling

As mentioned in the above section, academic events are more related to postgraduates where they need to collaborate ideas with other researchers for better research outcome. Therefore, homogeneous sampling is chosen where the people from similar background and experience is chosen as a sampling for this study to reduce variation (Patton, 2001). This is sampling widely used when we use a focus group, in this case the focus group are the postgraduate of University Malaya, as the research scope is academic event conducted by University of Malaya and it is limited to students who study in University of Malaya. A questionnaire (survey) was distributed to the postgraduates. The questions were mainly based on the criteria for the selection of academic events and the most important elements that postgraduates look for when selecting an academic event. A set of questions to better understand the academic event selection pattern were listed out for the students to answer.

3.2.1.2 Questionnaire

The aim of this survey was to understand the most important elements (contexts) affecting a student’s decision in choosing an academic event. This survey questionnaire consisted of 11 questions in five sections, as given below (Refer Appendix A):
Table 3.1 Survey Questionnaire

<table>
<thead>
<tr>
<th>Section</th>
<th>Question</th>
<th>Total number of questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Identity details (level of study and faculty)</td>
<td>2 questions</td>
</tr>
<tr>
<td>B</td>
<td>Academic events attended and what makes them attend these events</td>
<td>3 questions</td>
</tr>
<tr>
<td>C</td>
<td>What are the preferred criteria when selecting an academic event?</td>
<td>1 question</td>
</tr>
<tr>
<td>D</td>
<td>What is the preferred choice (weekend/weekday, free event/with cost, morning/afternoon/night events, within campus/outside/anywhere convenient)?</td>
<td>4 questions</td>
</tr>
<tr>
<td>E</td>
<td>Do the above criteria affect the student’s choice of an academic event?</td>
<td>1 question</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>11 questions</strong></td>
</tr>
</tbody>
</table>

The survey was conducted online by means of a Google form. This research survey used the quantitative method that focused on the respondent’s answers to the following questions:

1. What were the criteria that affected a student’s choice of past academic events?
2. What are the expected criteria in choosing academic events?
3. What are the preferred choices for the criteria that have been listed above?
4. Do the above criteria affect a student’s choice of academic events?

In section C, the criteria is distributed and not limited to only particular criteria. Students are not limited to choose the criteria, if there is a criteria out of the listed ones, the also can write that down. This is to ensure other contexts which affect the students decision choosing an academic event captured accordingly. However in section D, this question more focused on time and location based questions. Location is said one of the most dominated context in the context aware recommender system and time together can produce a more innovative and efficient tool (Xia et al., 2013). Based on this the question
is developed to know if the time and location context influence the decision making of the students.

3.2.1.3 Data collection

This survey received a very good response from the postgraduates. These postgraduates were from different fields of study who had attended academic events prior to participating in this study. From the results, it was encouraging to see that 99% of the respondents were postgraduates because this study was mainly focused on postgraduates to identify the expected criteria for choosing an academic event. Furthermore, all the respondents completed the survey, where all 12 questions were answered by everyone. Hence, the survey results were used as the deciding factor to identify the contexts that affect a student’s selection of an academic event. Moreover, the reason for recruiting respondents from different fields of study was to ensure that the criteria for attending academic events were similar across the faculties. For example, the students can be from either the accounting or business faculty, but from the results obtained, both chose similar contexts when selecting an academic event. Using this pattern of answers, it can be stated for certain that the academic event selection criteria (contexts) were common among all students. On the other hand, this study was all about developing a personalized recommender tool for academic events, and therefore, the scope of the users was also focused on postgraduates alone.

3.2.2 Data Analysis

Generally, data extraction is used to provide a description of the study in general, to obtain findings from every study, and help in the interpretation of the findings and the requirement analysis phase. From the survey, four main important contexts (cost, time, location and schedule) for the selection of academic events were identified. These four
main contexts carried a weightage of 80% and above, which meant that more than three quarters of the survey participants agreed that these were the important elements in the selection of an academic event.

![Chart showing important contexts in choosing an academic event]

**Figure 3.1 Important Contexts in Choosing an Academic Event**

Subsequent to the survey, the most to least affecting contexts of choosing academic events captured and plotted in a chart. Based on the chart, cost was the most important context, with 92 students selecting it (86.8%), followed by 91 (85.8%) students selecting time as another important context. The other two important contexts were location and schedule, which were selected by 86 students (81.1%) and 76 students (71.7%), respectively. Four contexts (cost, time, location and schedule) were identified and denoted as C1 (cost), C2 (time), C3 (location) and C4 (schedule) accordingly in the next phase.
3.3 User Design Phase

In this phase, interactions with users with the developed technique are analysed, and the input and output from and to the users are illustrated in architectural diagrams. The detailed development of technique is explained in Chapter 4, where it covers the contextual personalised recommender technique, contextual pre-filtering technique and the proposed technique which incorporate contextual technique with collaborative filtering technique.

3.3.1 Requirements Analysis

In this phase, gathered requirements in Section 3.2 were analysed to illustrate the proposed design. The four important contexts that were identified were used to translate the requirements into procedural specifications. Basically, in the recommender system, the User (U) is recommended with an Item (I) and a Rating (R). Hence, it is represented as

\[ R: \text{User} \times \text{Item} \rightarrow \text{R} \]

The requirement of incorporating contexts into the classical recommendation technique will be represented as

\[ R: \text{User} \times \text{Item} \times \text{Context} \rightarrow \text{R} \]

Using a contextual technique, the rating function R is represented as

\[ R = D_1 \times D_2 \times D_3 \times \ldots \times D_n \rightarrow \text{Ratings} \]
Here, the contextual information gathered explicitly by the users of the system will be processed and used as the contextualized data to filter the events and to recommend the most relevant academic event to the user.

3.3.2 Identification of the technique

In this study, a classical technique algorithm was chosen to incorporate the contextual information so as to provide more relevant academic events to the student. Hence, the classical algorithms were reviewed, and the most suitable algorithm to incorporate the contextual information was finalized. Two main classical approaches, namely, the content-based and collaborative filtering approaches, use some algorithms to predict items to users.

The collaborative technique (CF) technique was chosen to be incorporated with the contextual data after an analysis of the advantages and disadvantages of both the classical approaches. In an academic event recommender, it is vital to take into consideration the history of other students from the same faculty or working under the same research domain who are attending an event. Thus, the CF has an edge over the CB method in recommending better results to students. However, the issue of differences in individual interests was identified using user preference contexts, where the first level of filtering was done based on user preferences before looking out for the nearest neighbour. Although the CF uses many algorithms to find the nearest neighbour, this study focused on only the user-KNN algorithm, which will be explained further in Chapter 4.

Overall, the architecture was constructed to understand how the tool and technique worked from the point of view of both the user and the organizer. Explicit inputs from the user are captured for the contexts, and the inputs are contextualized (pre-filtered) prior
to the user’s search. When a user searches for events, the contextualized data together with the keyword will be presented as the search input to the databases for the events listed. These events will be based on the keyword and will be filtered using the contextualized data given. Next, using the user-KNN algorithm, the nearest neighbour of the user is identified, and the events will be filtered using the KNN algorithm prior to being displayed as the results to the user.

![Figure 3.2 Overall Architecture of the Contextual Personalized Tool](image)

### 3.3.3 Development of Technique

In this sub-section, the well-known KNN algorithm was incorporated with the context data, which had previously been explicitly given by the user. Hence, context filtering, followed by collaborative filtering, were used to develop the technique.
3.4 Construction Phase

During the construction phase, the technique and tool were combined to make a complete contextual personalized recommender tool. The requirements of the future end-users are used as the basis for the detailed definition of the design of every completed function.

3.4.1 Development of Tool

The overall architecture of the tool, as shown in Figure 3.2, was drafted to help in the development of the tool. Three main entities were developed, namely, the D1-User, D2-Organizer and D3-Event. The input given explicitly by D1 on preferences was contextualized. D2 used the same preferences to register an event being organized by them. The databases consisted of contextualized users, ratings, feedback, and events. Contextualized users here refer to users who had already registered and saved their preferences into the system. These contextualized users would later be used by the KNN-User to find the nearest neighbour.

3.5 Cut-Over – Testing Phase

The cutover phase is more on developing the test data needed to verify the system’s operational capacity and also to evaluate the system relevancy to the target audience. In this study, this phase is used to evaluate the technique and tool developed and evaluation result are explained in detail in chapter 5.

3.5.1 Evaluation of the Technique and Tool

One of the objectives of this study is to evaluate the tool and technique of contextualization. Thus, a precision and recall evaluation is used to evaluate whether the relevancy matrix of the retrieved events was according to the user’s preferences. Firstly,
the baseline for comparing the system against was chosen, namely, Acadevent.com. Acadevent.com was chosen as the baseline system because it uses classical approaches to recommend events. Since this study was aimed at showing that the contextualized approach is better than the classical approach, Acadevent.com considered the best to be chosen as the benchmark for this study.

Precision was chosen in performing the evaluation towards the technique and tool because it measures exactness and determines the fraction of the relevant retrieved items out of all the items retrieved. For instance, precision is used in measuring the movie recommender systems based on the proportion of recommended movies that are good in actuality. Relevant items gain more usefulness when they are found earlier in the recommendation list. Precision takes that into account as well. Whilst recall is to measure the completeness of the relevant items retrieved where it does take into count the fraction of relevant items which is not able to be retrieved in the results. Using both the precision and recall method, the technique and tool evaluated of exactness and completeness of the relevant events compared to classical approach.

3.6 Summary

RAD was the key on developing this entire chapter. Each phases has been given equal importance from the requirement analysis phase up to the cut-over phase. This is to ensure the technique and tool able to fulfill student academic event recommendation objective and at the same time produce a more relevant results compared to the classical approaches. The technique and tool development was only briefly explained in Chapter 3, but it is the Chapter 4, which holds the detailed process of the technique and tool. Chapter 4 constructed to be more precise on the development of technique and tool in
order to able to posturize the incorporation of the contextual information into collaborative filtering technique
CHAPTER 4: TECHNIQUE AND TOOL DEVELOPMENT

4.1 Introduction

This chapter explores the technique and tool development in detail, and shows the contextual personalized recommender technique in architectural figures.

4.2 Contextual Personalized Recommender Technique

The contextual personalized technique proposed in this study was a combination of context filtering followed by classical approach filtering to provide a better recommender. The contextualized data derived from the contextual filtering were further processed with the classical approach filtering to recommend items to the user.

The reduction-based approach explained by Adomavicius & Tuzhilin (2005) is an example of pre-filtering where, prior to the application of a content-based method to recommend items, the data are first computed and pre-processed with contextual data to give rise to contextualized data for further filtering. In this study, contextual pre-filtering was chosen to pre-filter the user preferences into a set of data, and these data were contextualized prior to their application in the collaborative filtering technique (user-KNN). One of the primary reason contextual pre-filtering selected is due to contextual information uses data or data constructed prior to running the recommendation engine. To be precise, information of the context, c is contextualized before the data input processed by the recommender engine. (Adomavicius and Tuzhilin, 2008).

Whilst in contextual post filtering the contextual data is intially ignored and only conextextualized after the recommendation engine recommends data from the 2D recommender. This will led to inability to fullfill the the research objective of this study as the contextual data is the input for contextual personalised recommender technique and
proposed to contextualized prior to performing the filter. On the other hand, the contextual modelling technique is where the data is used directly to the system and use a model based approach led to more complex model. This study does not use model based technique, and it adopts memory based technique and contextual modelling technique is not suitable.

The pre-filtering approach in a contextual RS uses the data in a few phases. The data and the explicit input of contexts are combined to generate the contextualized data. The user-KNN algorithm will then be used on this contextualized data to match students from the same faculty, and the events search mechanism will start to look for the best relevant event based on the student’s preferences, and those that have been attended by the nearest neighbour will also be top-listed in the recommendation results. Thus, the user can easily identify the most relevant event that matches his/her preferences and that has been registered by his/her neighbour, who could be from the same faculty or working on the same domain.
4.3 Contextual Pre-Filtering Process

In earlier chapters, the four most important contexts were finalized for incorporation into the contextualized RS. Therefore, the contexts were denoted as C1 (cost), C2 (time), C3 (location) and C4 (schedule). Each context had its own weightage, and all of the above were contextualized based on the weightage provided. Specifically, in reference to Adomavicius (2008), one can define contextual information with a set of contextual dimensions $C$, with every contextual dimension $C$ in $C$ being definable using a set of $q$ attributes $K = (K_1, \ldots, K_n)$. This context possesses a hierarchical structure that is similar to majority of the context-aware profiling and recommender systems. From this, the user selection can be of any set from the hierarchy, as shown in the table below. Each user there will have a set of their own preferences.

![Figure 4.2 Contexts Hierarchy](image)

An example of context conditions is described below, where six random users were chosen from different faculties, and each one of their preferences were captured. These
preferences were an illustration of the distribution of preferences between users. Hence, a set of contextual data processing for each user was created during the pre-filtering process.

Table 4.1 Example of User Preferences Distribution

<table>
<thead>
<tr>
<th>Contextual Conditions</th>
<th>Ali</th>
<th>Abu</th>
<th>Kumar</th>
<th>Nisa</th>
<th>Ah Meng</th>
<th>Razak</th>
</tr>
</thead>
<tbody>
<tr>
<td>no cost</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with cost</td>
<td>yes</td>
<td></td>
<td>yes</td>
<td>Yes</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>morning</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>afternoon</td>
<td></td>
<td></td>
<td>Yes</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>night</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>within campus</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>outside</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>Yes</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>weekday</td>
<td>yes</td>
<td></td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>weekend</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

According to the example of the table above, each user (Ali, Abu, Kumar, Nisa, Ah Meng and Razak) has very different preferences. Ali and Kumar prefer events without any cost, but, at the same time, both have different preferences with regard to the time of the event. Similarly, Nisa, Ah Meng and Razak do not mind attending an event with cost, but only Ah Meng and Razak prefer afternoon events, while Nisa prefers night events. This shows that each user has their own unique preferences. The contexts were denoted as C1, C2 and so on, whereby each condition of the context was denoted as below.

C1 (Cost) = C1A (no cost), C1B (with cost)
C2 (Time) = C2A (morning), C2B (afternoon), C2C (night)
C3 (Location) = C3A (within campus), C3B (outside campus)
C4 (Schedule) = C4A (weekday), C4B (weekend)
Thereinafter, the user will have a set of preferences as below.

\[
\begin{align*}
\text{Ali} &= \{C1A, C2A, C3A, C4A\} & \text{Nisa} &= \{C1B, C2C, C3B, C4A\} \\
\text{Abu} &= \{C1B, C2C, C3B, C4B\} & \text{Ah Meng} &= \{C1B, C2B, C3B, C4B\} \\
\text{Kumar} &= \{C1A, C2C, C3B, C4A\} & \text{Razak} &= \{C1B, C2B, C3B, C4B\}
\end{align*}
\]

Using this set of contextual data, the contextual pre-filtering will be done prior to the incorporation with the CF method.

4.4 Proposed Technique

The proposed technique, as shown in the flowchart in Figure 4.3, used three entities, namely User, Organizer and Event. Thus, these entities either provided the input or the output. A user, who is known as a system user, will register himself/herself in order to use the tool. He/she will then need to fill in the basic information to be eligible to use the tool. Moreover, it is compulsory for the student (user) to key in his/her preferences on entering the system for the very first time. The preferences can later be updated by the students based on their change of preferences. The contextual pre-filtering mechanism will use the latest updated preferences to generate a set of contextualized data for the particular user, and these preferences will be saved in the database for future use. If changes have been done to the preferences, the database will be updated the next time the student looks out for an event.

Organizers, who are considered to be another important entity, play a role in updating the events organized by them. Organizers will also need to go through the process of registration, and will have to key in the details of the events. The information or data on the event that is given by the organizers will be saved in a different database instance. Events, contextualized users and ratings were the three databases used in this study.
At the processing level, the user context and event information provided by the organizer will be matched. If there are events that match the user preferences, they will be checked for further filtering. Once they are confirmed to be in accordance with the preferences of the user, the next search technique using the KNN algorithm will look for the nearest neighbour for this user. Databases containing the basic and contextual information of the user will be searched for faculty members or researchers from the same domain to identify the list of neighbours. The identified neighbours will be searched for their attendance at events or registration for events. The events retrieved using this technique will be displayed as the recommendation results. The display will show which events were attended by which neighbours, listing the recommendations from the highest to the least according to the relevancy to the user. The pseudo code for the proposed technique is illustrated in Figure 4.4. That are the process flow used in the contextualized technique to retrieve the most relevant events.

![Figure 4.3 Input, Process and Output for Proposed Technique](image-url)
In summary, this chapter explains in detail the development of the technique and tool. The incorporation of contextual elements in classical approaches is to be evaluated to prove that the proposed technique and tool has an edge over the classical approaches. Moreover as described in literature review, context is not taking into consideration into making recommendations will lead to sparsely produced data and can cause manual filtering by the user to find the most relevant data. On the other hand, using contextual data in many RS domain has showed positive and better result of the recommendation (Adomavicius & Tuzhilin, 2005; Gediminas Adomavicius & Tuzhilin, 2008; Panniello & Gorgoglione, 2010). Hence, as stated in research problem that the contextual data is not widely explored in academic domain is the main and focal reason this technique and tool is proposed to understand how contextual data can be incorporated in academic domain to produce more relevant results.

**Figure 4.4 Pseudo Code of the Proposed Technique**

4.5 **Summary**

In summary, this chapter explains in detail the development of the technique and tool. The incorporation of contextual elements in classical approaches is to be evaluated to prove that the proposed technique and tool has an edge over the classical approaches. Moreover as described in literature review, context is not taking into consideration into making recommendations will lead to sparsely produced data and can cause manual filtering by the user to find the most relevant data. On the other hand, using contextual data in many RS domain has showed positive and better result of the recommendation (Adomavicius & Tuzhilin, 2005; Gediminas Adomavicius & Tuzhilin, 2008; Panniello & Gorgoglione, 2010). Hence, as stated in research problem that the contextual data is not widely explored in academic domain is the main and focal reason this technique and tool is proposed to understand how contextual data can be incorporated in academic domain to produce more relevant results.

```
**Input:** Set $KC$ which contain keyword and context info, $Ua$ as a user with profile and $ed$ as event database.

**Output:** Set $LHN$ asset of recommender events results.

1. **Begin**
2. **For** $Ua$ search query input for keyword and context, $KC \in ed$
3. **For** each keyword or context, $KC$ in event database, $ed$
4. **Search** with $KC$ being registered by neighbour in $ed$
5. **For** all highest neighbour ($HN$) being registered in $ed$,
6. **If** ($HN$ date > current date)
7. **Add** in the list of highest neighbour, $LHN \leq 20$
8. **Recommended** event based on contexts and neighbour, $LHN$
```
CHAPTER 5: EVALUATION OF THE TECHNIQUE

5.1 Introduction

In this chapter, the contextualized personalized recommender technique developed for academic events was evaluated and the results were analysed using the precision measurement method. The overall results of the technique evaluation were further explored to determine if the technique met the objectives of this study.

5.2 Findings

A pilot study was conducted for the evaluation, as shown in Figure 5.1, to compare the classical approach (Acadevent.com) to the contextualized approach (proposed technique for this study). A set of retrieved items, AP@4, were then checked for relevancy according to the user’s preferences. The recommended events were denoted as E1, E2, and so on. The relevant items were denoted as ‘R’ and the irrelevant items as ‘N’. If the first prediction was relevant, $1/1=1$ was given as the score, while the second relevant item in the set (E.g.: In the first set, it was the third event retrieved that was relevant) was given a score of $2/3=0.66$. The third, and fourth relevant items were scored using the same method.

Once all items had been checked with regard to their relevancy, the average precision was calculated using the formula for good items. Average precision (AP) is a ranked precision metric that emphasises highly ranked correct predictions (hits).

**Precision = good items recommended / all recommendations

**Average Precision = total score of the precision / number of recommended items.
Figure 5.1 Pilot Study of Technique and Tool Evaluation

5.3 Experiment Result

The table of comparison below shows the evaluation that was carried out on 20 students using both the classical approach and the contextualized approach, where the scores for each approach were calculated based on the relevancy of the recommended items. Using the pilot study, it was shown that the contextualized approach had an edge over the classical approach, where the top-listed items that were recommended were relevant compared to the classical approach.

Data were retrieved from 20 users to evaluate both the classical approach and the contextualized approach with two different keyword searches. The average precision for each user for 10 retrieved events (AP@10) was calculated, and all the average precisions for all 20 users were summed up to calculate the average.
By using the above formula, the average overall precision was calculated for each keyword using the classical approach and the contextualized approach.

Table 5.1 Experimental results for precision calculation

<table>
<thead>
<tr>
<th>Respondent</th>
<th>Classical Approach</th>
<th>Contextualized Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision AP@10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Keyword1</td>
</tr>
<tr>
<td>1</td>
<td>0.66</td>
<td>0.831</td>
</tr>
<tr>
<td>2</td>
<td>0.75</td>
<td>0.755</td>
</tr>
<tr>
<td>3</td>
<td>0.66</td>
<td>0.733</td>
</tr>
<tr>
<td>4</td>
<td>0.525</td>
<td>0.79</td>
</tr>
<tr>
<td>5</td>
<td>0.75</td>
<td>0.66</td>
</tr>
<tr>
<td>6</td>
<td>0.755</td>
<td>0.525</td>
</tr>
<tr>
<td>7</td>
<td>0.5</td>
<td>0.785</td>
</tr>
<tr>
<td>8</td>
<td>0.66</td>
<td>0.831</td>
</tr>
<tr>
<td>9</td>
<td>0.66</td>
<td>0.609</td>
</tr>
<tr>
<td>10</td>
<td>0.75</td>
<td>0.755</td>
</tr>
<tr>
<td>11</td>
<td>0.66</td>
<td>0.5</td>
</tr>
<tr>
<td>12</td>
<td>0.525</td>
<td>0.66</td>
</tr>
<tr>
<td>13</td>
<td>0.66</td>
<td>0.79</td>
</tr>
<tr>
<td>14</td>
<td>0.525</td>
<td>0.733</td>
</tr>
<tr>
<td>15</td>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td>16</td>
<td>0.66</td>
<td>0.79</td>
</tr>
<tr>
<td>17</td>
<td>0.609</td>
<td>0.733</td>
</tr>
<tr>
<td>18</td>
<td>0.755</td>
<td>0.75</td>
</tr>
<tr>
<td>19</td>
<td>0.5</td>
<td>0.66</td>
</tr>
<tr>
<td>20</td>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td>Average Precision AP@10</td>
<td>0.6442</td>
<td>0.7105</td>
</tr>
</tbody>
</table>
Another method of evaluation used is recall to measure the relevancy of the event recommended out of all relevant events. Recall is a measure of completeness, determines the fraction of relevant items retrieved out of all relevant items. Below is the recall formula to measure the relevancy based on the recommended events to the users. Out of 15 relevant event overall, table constructed of how many relevant events retrieved with the sample of 20 users. Data constructed to show the comparison between classical approach and contextualised approach.

\[
\text{Recall} = \frac{\text{good items recommended}}{\text{good item recommended} + \text{good items not recommended}}
\]

Table 5.2 Experimental results for recall calculation

<table>
<thead>
<tr>
<th>Respondent</th>
<th>Classical Approach</th>
<th>Contextualized Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recall</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Keyword1</td>
<td>Keyword2</td>
</tr>
<tr>
<td>1</td>
<td>0.4</td>
<td>0.66</td>
</tr>
<tr>
<td>2</td>
<td>0.46</td>
<td>0.66</td>
</tr>
<tr>
<td>3</td>
<td>0.3</td>
<td>0.66</td>
</tr>
<tr>
<td>4</td>
<td>0.5</td>
<td>0.46</td>
</tr>
<tr>
<td>5</td>
<td>0.66</td>
<td>0.46</td>
</tr>
<tr>
<td>6</td>
<td>0.2</td>
<td>0.66</td>
</tr>
<tr>
<td>7</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>8</td>
<td>0.46</td>
<td>0.5</td>
</tr>
<tr>
<td>9</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>10</td>
<td>0.3</td>
<td>0.46</td>
</tr>
<tr>
<td>11</td>
<td>0.4</td>
<td>0.66</td>
</tr>
<tr>
<td>12</td>
<td>0.46</td>
<td>0.3</td>
</tr>
<tr>
<td>13</td>
<td>0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>14</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>15</td>
<td>0.3</td>
<td>0.66</td>
</tr>
<tr>
<td>16</td>
<td>0.46</td>
<td>0.46</td>
</tr>
<tr>
<td>17</td>
<td>0.2</td>
<td>0.46</td>
</tr>
<tr>
<td>18</td>
<td>0.4</td>
<td>0.66</td>
</tr>
<tr>
<td>19</td>
<td>0.46</td>
<td>0.46</td>
</tr>
<tr>
<td>20</td>
<td>0.3</td>
<td>0.46</td>
</tr>
<tr>
<td>Average Recall</td>
<td>0.403</td>
<td>0.502</td>
</tr>
</tbody>
</table>
5.4 Evaluation Results

The evaluation results showed that for the first keyword, the classical approach attained an average precision of 0.6442 over the 10 retrieved items, while the same keyword search in the contextualized approach attained a slightly higher average precision of 0.77085.

The second keyword search was measured and the contextualized approach scored an average precision of 0.76965, which was higher than the average precision of 0.7105 that was scored by the classical approach. The evaluation results further confirmed that the contextualized approach was better than the classical approach. This did not mean that the classical approach did not recommend good items, but the contextualized approach recommended items that were more relevant at the top of the list compared to the classical approach, where the recommendations were distributed among the relevant and non-relevant items.

As for the recall an average of 0.6935 achieved for keyword 1 is gained for contextual approach compared to classical approach is 0.403. Contextual approach has a higher fraction of relevancy compared to classical approach. The same goes to the second keyword search using contextual approach which register a higher fraction of relevant item which is 0.7045 compared to classical approach of 0.502.

5.5 Summary

The contextualized approach obtained the relevant items based on the preferences explicitly given by the user. Thus, when the user saw the recommended items, the top listed ones were the items that matched their preferences. Hence, the first few items that were retrieved were considered to be good recommendations based on user preferences.
Classical approaches which only uses pure filtering technique derived good results. But yet, the contextual data of four element incorporation has shown a higher fraction of exactness and completeness in the recommendation results. Location, time, cost and schedule has influenced the result in general. Students explicitly given their preferences of the four context finds the recommendation result is more relevant to what they are searching compare to the classical search only by keyword. Even though classical approach does filtration to produce result, it still can be sparse as it is not what the student may be really want to be recommended on. Thus, having contextual data incorporated will help the classical approach to get more efficient and relevant results based on the user preferences.

In summary, the evaluation results showed a positive outcome from the contextualized approach, where this experiment can be extended to more users to produce the mean average for precision and recall
CHAPTER 6: CONCLUSION AND FUTURE WORK

This study investigated the preferences (contexts) of students in choosing an academic event. Numerous tools are available for the recommendation of events to students. However, the existing tools still encounter some limitations when it comes to recommending more relevant events to students. All the research questions set out earlier were answered and discussed in Chapters 2, 3, 4 and 5. This chapter discusses the findings of this study, the research contribution, its limitations and future work. According to this study, the objectives and research questions (RQs) were determined as:

Objective 1: To identify the user context for the selection of academic events.

Research Question 1: What are the important contexts in the selection of academic events?

Objective 2: To develop a contextual personalized recommender technique.

Research Question 2: How can a system be developed using a contextual personalised technique?

Objective 3: To evaluate the developed technique.

Research Question 3: How can the developed technique be evaluated?

6.1 Research Findings

The RQs and objectives for this study were revisited. The findings for each RQ are discussed below.

Objective 1: To identify the user context for the selection of academic events
Research Question 1: What are the important contexts in the selection of academic events?

Many elements affect students in their choice of academic events. This was discussed in detail in Section 3.2.1, and based on the results obtained from the questionnaire, the four most important contexts were found and were further used to develop the technique. This answered question 1. The objective 1 is achieved.

Objective 2: To develop a contextual personalized recommender technique

Research Question 2: How can a system be developed using a contextual personalised technique?

As described in Chapter 4, the contexts derived explicitly from the users were incorporated into the selected classical filtering user-KNN algorithm to process the input and retrieve recommendations based on user preferences. Contextual pre-filtering was used at the beginning of the process to contextualise the preferences into a set of inputs from each user prior to searching the inputs using the keyword. The personalization occurred when the keyword was combined with the contextual data used to search for the events. This answered question 2. The objective 2 is achieved.

Objective 3: To evaluate the developed technique

Research Question 3: How can the developed technique be evaluated?

An experiment was carried out to evaluate the proposed technique using precision, where each retrieved item was checked as to its relevance according to the preferences of the user as discussed in section 3.2.2. The results obtained were compared with the existing
tools available to confirm that the contextual recommender technique was better than the classical technique. This answered question 3. The objective 3 is achieved.

6.2 Significance of the Study

This work was aimed at helping students from the academic field to find the most suitable events to participate in and to collaborate with other students. The contribution of this work was the development of a contextual personalized academic event selection recommender technique using collaborative filtering, where the user’s preferences and location are taken into account to recommend more relevant academic events. Based on the literature it was proven in other RS domain that contextual data has immensely shown efficient and innovative results. Hence, this study proposed to learn and incorporate contextual data into academic event recommender to produce better result. The experimental result of this study clearly shows that incorporating contextual data will improve the recommendation result. The exactness and completeness of the relevant events retrieved were a result of contextual information incorporation into classical approach. This work will help future researchers to further explore context-aware recommender systems in the academic domain.

6.3 Limitations and Future Work

A few limitations of this study were identified, which can be further explored:

- This study only focused on the KNN algorithm. Thus, this study can be further extended to incorporate data on contexts using different CF algorithms and comparing the results.
- Sample size is small. The use of more respondents will give better average precision results for the contextual technique.
• The database for the events was limited to events happening at the University of Malaya. Importing more events will have a good impact on the results.

6.4 Conclusion

The primary focus of this study was to identify the preferences of students that affect their decision to attend academic events. The results of the investigation showed that students face difficulties in finding the relevant academic events based on their preferences due to the limitations of existing tools. Thus, this study focused on developing a contextual personalized recommendation technique that incorporates contextual information in its retrieval process. The contextual pre-filtering technique was chosen due to its suitability in helping to contextualize data prior to the searching process.

This study found that contextual information, which can also be defined as user preferences, can be mapped into recommender systems. The main significance of this study is that user preferences can be identified and incorporated into the classical recommendation technique.

A classical user-based recommendation algorithm, the user-KNN, and a context pre-filtering technique were embedded in the technique development. The contextual data were incorporated into the K-nearest neighbour (KNN) algorithm to recommend events based on user preferences.

Finally, from the experimental results obtained, it is believed that all the objectives set out at the beginning of this study have been successfully achieved.
REFERENCES


AllConference (2000), AllConference. Retrieved from AllConferences.com


Jung, J. J. (2011). Ubiquitous conference management system for mobile recommendation services based on mobilizing social networks: A case study of u-


Conference on Information and Computer Science, NICS 2015, 141–146. https://doi.org/10.1109/NICS.2015.7302179


