# Multi-Agent Modeling of Risk-Aware and Privacy-Preserving Recommender Systems

By

Vishnu Srivastava

A thesis
Presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2017

# AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

## Abstract

Recent progress in the field of web recommender systems has led to the increase in the accuracy and better personalization of the recommendations [18]. These results are being achieved by gathering more user data and generating insights from it. However the privacy concerns of the user are often underestimated and ignored. In fact, many users are not sufficiently aware of the data that is collected or if such data is sold to the third party.

Research in the area of recommender system should strive towards not only achieving high accuracy of the generated recommendations but also maintain user privacy [2,4,5,6,11,12,14] and making recommender systems aware of the user's context [128] i.e. intentions of the user and the current situation of the user. Through research it has been established that a tradeoff is required between accuracy, privacy [130] and risk [7] in a recommender system and that it is highly unlikely to have recommender system satisfying all the requirements of being contextually aware and privacy preserving. Nonetheless, a significant attempt can be made to describe a novel modeling approach that supports designing a recommender system encompassing some of the mentioned requirements.

This thesis focuses on designing a multi-agent model of a recommender system by breaking it down into three different subsystems i.e. the data subsystem, the risk subsystem and the privacy subsystem. Within each subsystems, the operations are carried on by at least one agent, having a specific role in order to achieve a prescribed goal by performing some activity. Such a description of a system will be able to represent a small subset of recommender systems which can be classified as risk aware and privacy preserving web recommender system.

# Acknowledgements

I would like to thank Professor Paulo Alencar, for his patience, understanding, kindness and guidance. I learned a lot while working with him and I am proud to be his student. I am thankful to Professor Daniel Berry for serving as my co-supervisor. I also thank Professor Donald Cowan and Professor Gladimir Baranoski for agreeing to read my thesis and providing me valuable feedback.

# Table of Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

In recent times, Recommender Systems can take advantage of the semantic reasoning capabilities [128] to overcome common limitations and improve the recommendation quality. These systems uses domain properties, types and relationships to enhance the personalization. Current research in the area of recommender system has been focussed on the context aware recommender system [18]. A context-independent representation may lose predictive power because potentially useful information from multiple contexts is aggregated [128]. The ideal context-aware recommendation system would, therefore, be able reliably to label each user action with an appropriate context and effectively tailor the system output to the user in that given context. The majority of existing approaches to recommender systems focus on recommending the most relevant content to users using contextual information and do not take into account the risk of upsetting the user in by not providing accurate recommendations. However, in many applications, such as recommending personalized content, it is also important to consider the risk of upsetting the user by not being aware of the user's situation and intentions [7]. Therefore, the performance of the recommender system depends in part on the degree to which it has incorporated the risk into the recommendation process. The risk in recommender systems is the possibility to disturb or to upset the user which leads to a bad feedback from the user. With the advent of enormous amounts of personal data collection for the sake of personalization and improving recommendations quality, the focus of the current research on the recommender system has been shifting to Privacy Protection [129]. Personalization provides users with conveniences, and it can have a direct impact on marketing, sales, and profit. On the other hand, privacy, which is a serious concern for many users, is the price users have to pay for the convenience of recommender systems in a world with booming information. Users normally have no choice but to trust the service provider to keep their sensitive personal profile and information safe.

## 1.1 Research Issue

Since the majority of the focus in the area of recommender systems has been on the improvement of accuracy of the recommendations generated by the recommender system, there is a lack of focus on a modelling approach for the recommender systems which takes care of the aspect of upsetting the user by not having sufficient knowledge of the user's context and not taking enough measures to take into account the privacy of the users. A novel modeling for the recommender systems would make things much easier for domain experts to study and extend the research in the area of risk aware and privacy preserving recommender systems and thereby resulting in better design of such systems.

In the past few decades collaboration of multiple teams for a large software project has become a usual path for developing large scale software [15]. In spite of adoption of collaborative software development there is a lot of improvement to be done between what is needed and what has been provided today and because the software development landscape is changing rapidly. Multi-Agent software development is being utilized as a way to develop software by working on different aspect of a software system as separate agents, working in coherence to achieve the objective of the system. But a multi-agent approach to risk-aware and privacy protection is much scarce [18].

Since, most of the current research is focused on evaluating recommender system in terms of the accuracy of the results produced by such systems. There needs to be a discussion on a unifying universal a way to evaluate a model of a recommender system, not just on the accuracy of the underlying algorithm but also on the features it contains to enhance the user experience such as user's intention, context and also the privacy of the data that is used to produce recommendations.

## 1.2 Thesis Statement

The aim of this research is to provide a multi-agent based system model of a recommender system by introducing both privacy and risk-related abstractions into a traditional recommender system that supports designing these systems when privacy and contextual risk involving user data and information needs to be taken into account, followed by a case study of a job recommender system as a sample implementation of this approach.

## 1.3 Major Contributions

This research focusses the importance of the privacy and the risk aspect of the recommender systems i.e. how much a recommender system safeguards users' privacy and also how a recommender system utilizes the contextual data of a user in order to benefit the owner of the information as well as other users of the recommender system.

The approach utilizes multi-agent system description in a sense that the designers of the recommender systems can focus on individual units. This breakdown of the recommender system into small individual units enables fast and fault tolerant development of the system. It enables the designers of the recommender system to be aware of the each of the small objectives that must be accomplished by the each individual units in order to fulfil the objective of the entire system.

This approach combines the two existing research areas within the recommender systems i.e. risk and privacy into a system model. Towards the end of this thesis, a sample case study of this approach is also provides in the field of job recommender systems.

## 1.4 Thesis Organization

The thesis is divided into three parts. The first part is the introduction of the problem addressed in this paper, along with a survey of the Recommender Systems field which brings into light the risk and privacy issues, where this thesis is framed. The second part describes the related work in the recommender systems literature and provides an analysis of the design alternatives and statistical biases that may arise. It also provides a detailed discussion of the proposed approach to solve the issues with the existing models of the web recommender system. Towards the end of this part a brief case study of this model is show in the field of job recommender systems. The last part describes the work to be done in the future to extend this system model. In the Appendix a possible evaluation model for recommender systems based on model based approach has been discussed.

In more detail, the contents of this thesis are distributed as follows:

### Part I. Introduction

**Chapter 1** In this chapter, a brief description of the current focus in the area of recommender system has been provided followed by the description of the issues encountered with the current approach by the researcher and the domain experts regarding the recommender systems. The thesis statement is then provided to give an idea of what this paper is trying to achieve. This is followed by listing out some of the major contributions of this paper.

**Chapter 2** provides an overview of the state of the art in recommender systems, considering a classification of the main types of recommendation approaches. We also describe the weaknesses of the different recommendation techniques and present a broader class of hybrid recommenders that aim to overcome these limitations. We also discuss the risk and privacy issues in the recommender systems. And how these issues arises in the first place in such systems. The discussion is carried forward with the description of the some of the mitigating techniques that can be used to address this problem.

## Part II. The System Model

**Chapter 3** mentions some of the works in the field of recommender system that has contributed toward the conceptualization of the approach that is discussed in the paper.

**Chapter 4** presents the proposed approach. In this section, a detailed description of the system is provided along with the explanation of different aspects of this model.

**Chapter 5** presents the case study of the proposed approach in terms of the job recommender system. In this chapter a discussion of the previous approach to a job recommender system has been provided, followed by providing a sample instance of such a system, as part of the approach which is discussed in this paper. A discussion is also carried out along with the proposed enhancement(s) over the previous approach in order to make such recommender system privacy preserving and risk aware.

## Part III. Future work

**Chapter 6** Discusses some of the work(s) that can be carried out in the future to improve upon this approach of the instantiation of a multi-agent model for the recommender systems across different application areas.

**Appendix** This section discusses a possible model based approach for the evaluation of the recommender systems using the privacy preserving and the risk aware concepts.

# Chapter 2

# Recommender Systems

Recommender systems is a piece of software which produces a list of recommendations for its users by deploying generally two algorithms (i.e. collaborative filtering or content-based filtering) or a mix of these approaches into a hybrid approach. The approach used in Collaborative filtering utilizes the user's historic data (i.e. items purchased by the user, browsing/navigation history on the website or the feedback provided for the purchased item). The result of this approach is the list of recommendations produced by the system of interest to the user [22]. On the other hand, the Content-based filtering approaches employs some sort of attributes of an item in order to come up with a list of recommendations having item with similar attributes.[23] A Hybrid approach can be used as a combination of the previously discussed approaches in order to find the approach with the best accuracy of the recommendations.

## 2.1 Context Aware Recommender Systems

In paper [7] the authors have briefly discussed Context Aware Recommender systems. In order to make recommendations more accurate, the context at the time of generating recommendations is also an important factor. The contextual data can be added as a source of information for generating better recommendations or can help in filtering out non relevant recommendations from the list of resultant recommendations, generated by the system Therefore, from this utility of the context arises the Context Aware Recommender systems [21].

## 2.2 Privacy in Recommender Systems

There is a presence of large variety of information within the recommender system. Paper [19] discusses these diverse information types in detail. Some of this information can be confidential and should not be revealed to any other person/organization except for the owner of the information. On the user's end, there is always a trade-off between the amount of information to be provided to the recommender system and the accuracy of the recommendation. This aspect is represented in the paper [19] with the help of three dimension representation having duration of information storage, size of audience and the extent of usage as its three axis.

## 2.3 Privacy Protection

In order to alleviate the privacy concerns of the user to make the user provide more information to the system for better recommendations, some privacy protection techniques can be employed. One of the methods is Anonymization, which is removing any link in the data to a specific user while preserving the structure in the data. Paper [34] uses this approach using trust agents. The other methods to achieve the alleviation of the privacy concerns is to use Randomization and/or differential privacy servers.

### 2.3.1 User control

The paper [14] discusses two techniques to mitigate the concerns over the risk of privacy breach in the recommender systems by providing the user with the option to manage the release of information to the recommender system [41] or to provide appropriate reason for the requirement of release of information to the user [42]. These two methods helps in reducing concerns of privacy breach of the user.

## 2.4 Risk Aware Recommender Systems

The paper [7] discusses the Risk Aware Recommender Systems. In this variation of a recommender system an approach is used to calculate the trade-off of discovering contextual information and upsetting user with non-relevant recommendations. This trade-off factor is termed as risk and is calculated by using the multi-arm bandit optimization method. The techniques that are discussed in the paper are derived from the "variance of the cost" approach ([44, 43]), "the expected environment cost" approach [45, 46, and 47] and the hybrid approach ([48]).

# Chapter 3

# Related Work

## 3.1 Modelling Recommender system

In [1] a description of an ontology-driven model for usage mining in the context of agent-based Web recommender systems is provided. It first starts with a description of MADEM (Multi-Agent Domain Engineering Methodology) as a software development methodology for multi-agent domain engineering followed by the description of the modeling concepts, tasks and products for the development of a family of multi-agent systems in a problem domain.

## 3.2 Risk Aware Recommender Systems

After gaining understanding the concept of multi-agent system in context of recommender systems we now discuss the Dynamic risk aware recommender system, as described in [7].Dynamic Risk Aware Recommender System (DRARS) is essentially a context aware recommender system which uses the exploration-exploitation trade-off using Multi-Arm bandit optimization problem.

## 3.3 Privacy Preserving Recommender Systems

Paper [5] presents a collaborative privacy framework for preserving users' profile privacy in a social recommender service. It also provides an overview of EMCP components and the interaction sequence for a recommendations process in an IPTV content distribution scenario followed by the description of a novel two stage concealment process that offers to the users a complete privacy control over their ratings profiles. The concealment process utilizes hierarchical topology, where users will be organized in peer-groups, from which super-peers are elected based on their reputation.

Super-peers aggregate the preferences obtained from underlying users and then encapsulate them in a group profile and then send them to PRS. This paper also provides a test of performance of the proposed framework on a real dataset and the evaluation of how the overall accuracy of the recommendations depends on a number of users and requests. The experimental and analysis results showed that privacy increases under proposed middleware without hampering the accuracy of recommendations. Moreover the approach used in the paper has been shown to reduce privacy breaches on the concealed data without severely affecting the accuracy of recommendations based on collaborative filtering techniques by realizing that there are many challenges in building a collaborative privacy framework for preserving privacy in social recommender service. Paper [6] provides an evidence that the disclosure of user preferences in a recommender system seriously threatens users' personal privacy especially when service providers move their user data to an untrusted cloud. In this paper, a novel solution, called APPLET is presented, to address the significant challenges in privacy-preserving location aware recommender systems. For APPLET, an introduction of multiple cryptography methodologies has been provided in the paper in order highlight the aspect of protecting the privacy of the recommender system users without affecting the recommendation quality. Moreover, an evaluation has been provided that the effectiveness and performance of APPLET turns out to be well suited. In [12], the author(s) proposed a novel method for privacy preservation in collaborative filtering recommendation systems. The author(s) addressed the problem of protecting the users' privacy in the presence of an untrusted central server, where the server has access to users' profiles. To avoid privacy violation, a mechanism is proposed where users store locally an offline profile on their own side, hidden from the server, and an online profile on the server from which the server generates the recommendations. The online profiles of different users are frequently synchronized with their offline versions in an independent and distributed way. Using a graph theoretic approach, the author(s) developed a model where each user arbitrarily contacts other users over time, and modifies his own offline profile through a process known as aggregation. To evaluate the privacy of the system, this approach is then applied to the Netflix prize data set to investigate the privacy-accuracy trade-off for different aggregation types. Through experiments, it is concluded in the paper that such a mechanism can lead to a high level of privacy through a proper choice of aggregation functions, while having a marginal negative effect on the accuracy of the recommendation system. The results illustrated that similarity-based aggregation functions, where users receive items from other users proportional to the similarity between them, yield a considerable privacy level at a very low accuracy loss. The findings in [14] suggest that users' online information is multi-dimensional regarding privacy concerns, especially in a recommender context. Although this seems self-explanatory, it is often neglected in privacy research and recommender system design. Specifically, demographic information that is frequently required for online service registration can be divided into two categories i.e. unidentifiable information and identifiable information. Unidentifiable information consists of items describing one's personal attributes that cannot be used to uniquely pinpoint the individual, whereas identifiable information is more accurate in pointing to the individual's identity exclusively. It has been argued that the people are significantly more concerned about the recommender system accessing their identifiable information than their unidentifiable information. In a similar manner, product items can be broadly grouped into non-sensitive types and sensitive types. It is also shown that the users are significantly more worried about their previous purchases of sensitive products being accessed for personalized recommendations than they are about their previous purchases of non-sensitive products. These item-based analysis and categorizations provide a relative information-ranking system in terms of privacy concern in recommender systems, thus refining existing research on general privacy concern about user information. The categorization provided in the paper is extended prior research by extracting new factors, which can be used as a reference in future studies and designs. These new factors are shown to suggest that recommender system designers should treat users' information discriminatively and strategically based on their levels of sensitivity for pattern prediction and personalized recommendations. It is also concluded that the algorithm developers should be well aware of what information users are more hesitant to disclose, so as to adjust the degree of information tracking and use, as well as to provide appropriate coping strategies. In line with the "privacy-personalization Trade-off," unsolicited access to users' sensitive information may trigger severe privacy concerns that could affect users' overall experiences therefore, identifiable and sensitive data should be more cautiously handled in exchange for prediction accuracy. As a design suggestion, recommender systems should introduce user control or privacy assurance mechanisms to help alleviate users' privacy concerns. Also, user data with different sensitivity levels (e.g., identifiable vs. unidentifiable information) can be potentially protected with different levels of privacy remedies.

## 3.4 Privacy Preserving Methodologies for Recommender Systems

Traditional location-aware recommender systems are facing a significant challenge, namely, how to protect the location privacy of users while preserving the recommendation quality. There are several studies that have achieved location privacy, which are based on anonymity, differential privacy, and encryption schemes. The authors of [49–51] proposed some location-privacy preserving mechanisms (LPPMs) based on anonymity to protect the user's location privacy. Although these anonymity mechanisms are diversiform, each of them assumes the adversaries own specific prior knowledge. To solve the shortcomings of the above schemes, the authors of [52–54] introduced differential privacy mechanisms to protect the user's exact location independently from any side information that the adversary might possess. Notably, none of the work above can be directly used to protect the privacy in a recommender system, which also includes some other sensitive information.

# Chapter 4

# Proposed Approach

In this chapter we will discuss the proposed approach to tackle the challenges described in the previous sections. Let us start with a conceptual model of a recommender system as a system where the resultant recommendations are affected by the privacy factors (e.g. user controls, privacy settings etc.) and the contextual risk factors (e.g. location, social connections etc.). The privacy risk factors are can be understood as the parameters which are formulated by taking privacy instructions from the user and then filtering out the data to be considered for generating recommendations based on those privacy parameters set by the users. On the other hand, the contextual risk factors are the parameters that are obtained from the continuous/periodical stream of the user data followed by the filtering by the privacy parameters and which are used as one of the data source for generating the recommendations. Thus, in order to propose a model for the recommender system, we need to have a model which takes into account these two factors affecting the system.

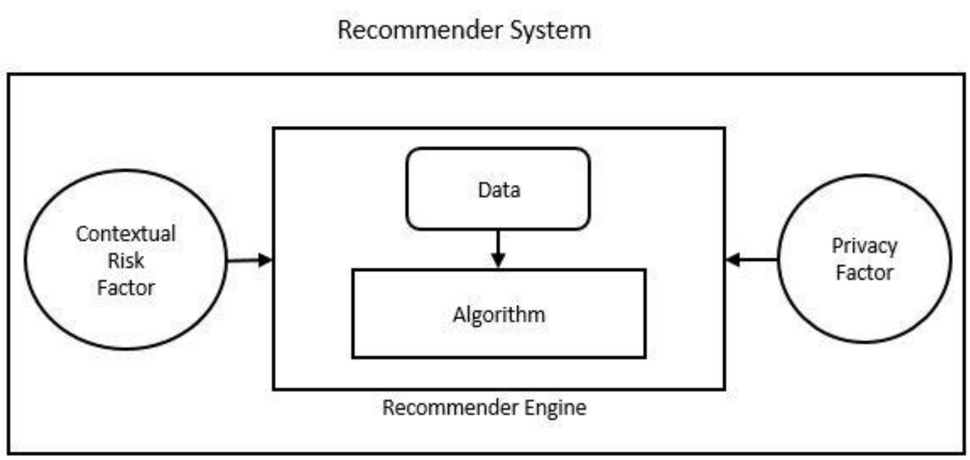


**Figure 1 Conceptualization of a Recommender System**

The proposed approach to describe a model for the web recommender system follows a sequence of steps in order to describe a model of the system. In the first step the conceptual system is broken down into 3 subsystems (i.e. the Data Subsystem, the Contextual Risk Subsystems and the Privacy Subsystem) to discuss the impact of the privacy and the risk factor on the overall objective of the system i.e. to produce the recommendations. This step also involves the introduction of an agent based approach where each subsystems is considered to be operated by one or more agents in order to accomplish the objective of that subsystem.

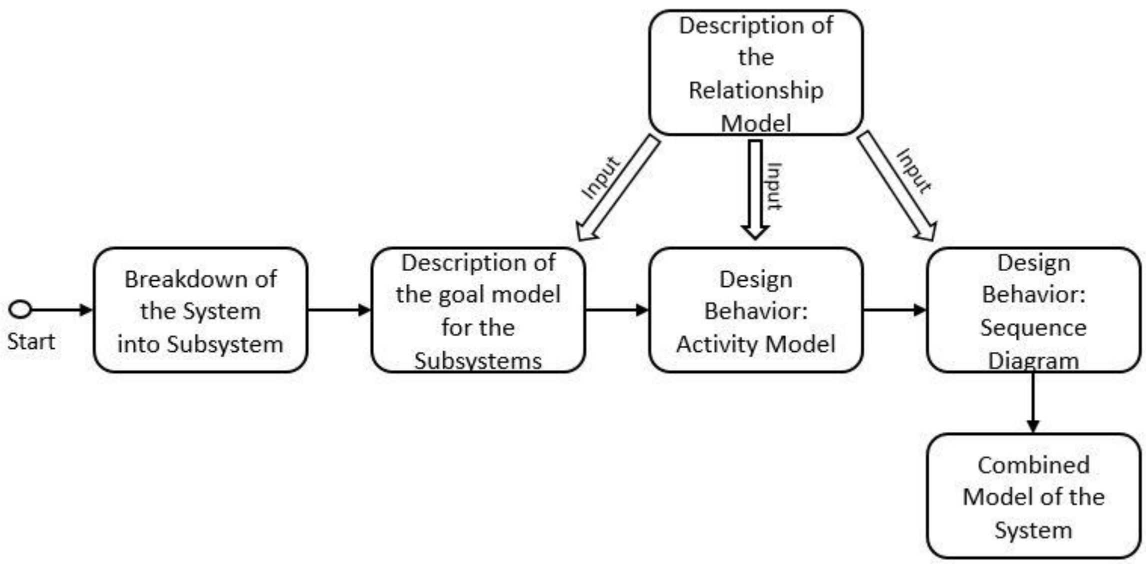


**Figure 2 Proposed Approach**

In the next step, we provide a goal model for each subsystems within our systems in order to specify the goal of these subsystems. The agents within these subsystems are described in terms of the roles they perform, the responsibilities they fulfill and the activities performed by these agents in order to achieve the objective of the subsystem. This is achieved partially by the introduction of the relationship model which provides a set of principle followed by each agent in order to accomplish its responsibilities within the subsystem.

We introduce two design behaviors for the next two subsequent steps. These design behaviors helps in understanding the system even better by providing the internal working of each subsystem. The first design behavior we discuss is the Activity model of the subsystems. It describes the working of the subsystem in context of the relationship model discussed earlier. The Activity model for each subsystems are then combined to form an activity model of the entire recommender system.

The second design behavior which is discussed is the Sequence diagram of the subsystems. The sequence diagram describes the sequence of events that occur within the subsystems. These sequence diagrams are then combined to form the sequence diagram of the recommender system. The working of the sequence diagram are based on the contextual information from the Relationship model of the subsystems.

The starting point of the description of the proposed approach is to discuss the UML modelling techniques. Before going further in the description of the system model, it is indispensable to describe the notations used in this approach. Various types of UML diagrams are used (activity diagram and sequence models) to provide the system description and to gain understanding of the workings of the subsystems ond the recommender systems as a whole. These diagrams are explained in the following section.

## 4.1 UML Diagrams

UML stands for Unified Modeling Language and is used in object oriented software engineering. Although typically used in software engineering, it is a rich language that can be used to model an application structures, behavior and even business processes. There are 14 UML diagram types but for the purpose of this thesis, we will be focusing only on the Activity Diagram and the Sequence Diagram.

### 4.1.1 Activity Diagram

The basic purposes of activity diagrams are similar to captures the dynamic behavior of the system by showing the message flow from one activity to another. Activity is a particular operation of the system. Activity diagrams are not only used for visualizing dynamic nature of a system but they are also used to construct the executable system by using forward and reverse engineering techniques. The only missing thing in activity diagram is the message part. It does not show any message flow from one activity to another. Activity diagram is some time considered as the flow chart. Although the diagrams looks like a flow chart but it is not. It shows different flow like parallel, branched, concurrent and single.

### 4.1.2 Sequence Diagrams

UML sequence diagrams are used to represent or model the flow of messages, events and actions between the objects or components of a system. Time is represented in the vertical direction showing the sequence of interactions of the header elements, which are displayed horizontally at the top of the diagram. Sequence Diagrams are used primarily to design, document and validate the architecture, interfaces and logic of the system by describing the sequence of actions that need to be performed to complete a task or scenario. UML sequence diagrams are useful design tools because they provide a dynamic view of the system behavior which can be difficult to extract from static diagrams or specifications. Although UML sequence diagrams are typically used to describe object-oriented software systems, they are also extremely useful as system engineering tools to design system architectures, in business process engineering as process flow diagrams, as message sequence charts and for protocol stack design and analysis.

## 4.2 Goal Model

Goal models for the recommender systems were introduced in [1]. In this thesis, the goal models are generalized to the subsystem model of the recommender system in order to describe the workings of the subsystems. This is an agent based model and the workings of the subsystems is represented diagrammatically with the help of the relationship model.

## 4.3 Multi-agent system Model and System Description

In this approach we will start with breaking-down the system into subsystems. Each subsystem will be responsible for accomplishing a pre-defined task. Each sub system will also consist of an agent. For the specification of the problem domain to be solved, we will focuses on modeling goals of the subsystem, roles of the agents, activities performed by the agents and finally the interactions of the agents. Agents possesses knowledge and that knowledge is used by it to exhibit autonomous behavior.

A subsystem is composed of agents having specific goals that establish what the subsystem intends to accomplish. The achievement of specific goals by the agents within a subsystem, allows reaching the general goal of the entire system when put together.

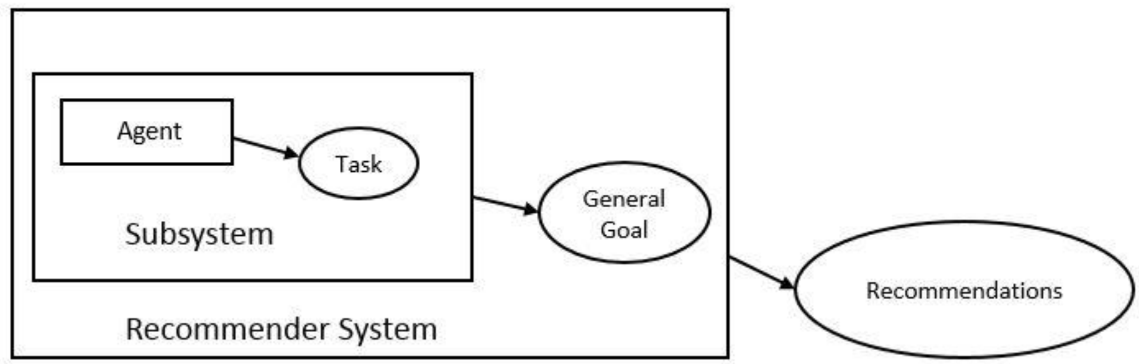


**Figure 3 Proposed Approach**

Specific goals by the agent within a subsystem are reached through the performance of responsibilities that agents have, by playing roles with a certain degree of autonomy. Responsibilities are exercised through the execution of activities by each individual agent within the subsystem. The set of activities associated with a responsibility are a functional decomposition of it. Roles have skills on one or a set of techniques that support the execution of responsibilities and activities in an effective way within the subsystem. Preconditions and post-conditions may need to be satisfied for/after the execution of an activity by each agent within the subsystem. Knowledge can be consumed and produced through the execution of an activity. Skills can be, for instance, the rules of the subsystem that agents know to access and structure its information sources. Sometimes, agents have to communicate with other agents to cooperate in the execution of an activity. This approach allows for such communication to take place between the agents within the subsystems.

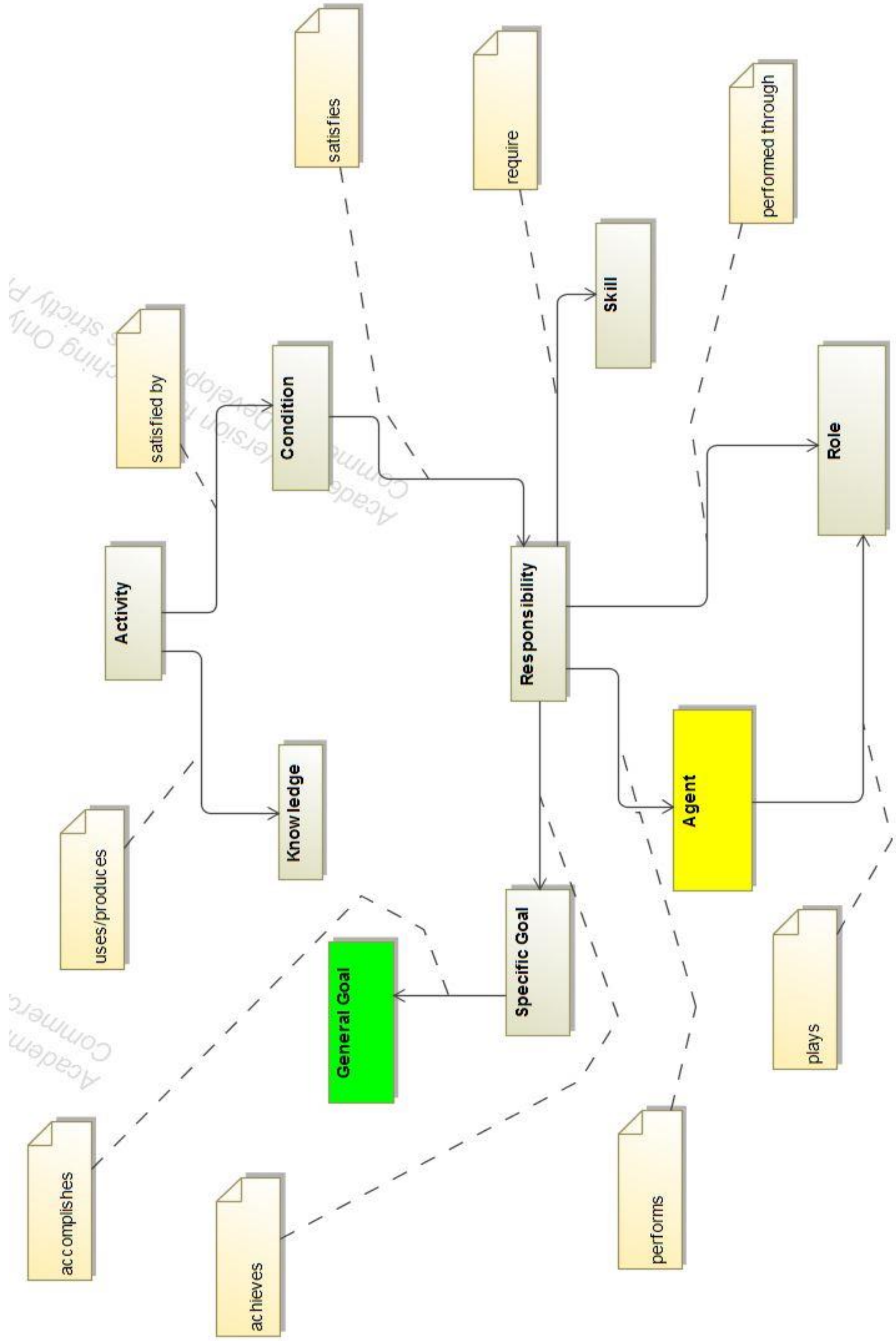


**Figure 4 Relationship model for subsystems**

## 4.4 Goal Models for the Subsystems

We will now discuss the goal models of the subsystems which makes up a risk aware and privacy preserving web recommender system and also explain the contribution of each subsystems and the agents involved in the respective subsystems.

### 4.4.1 Goal Model: Data Subsystem

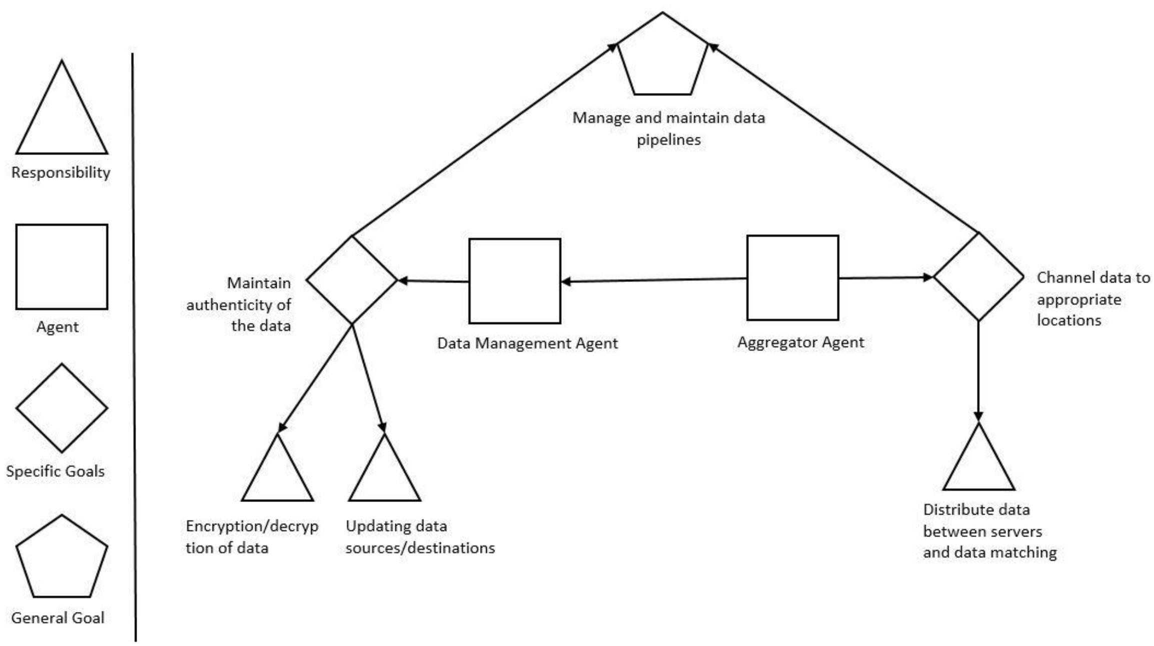


**Figure 5 Goal model for the data agent and the aggregator agent**

Let us first start with the data subsystem manages. This subsystem is responsible for managing the data inflow and outflow from the recommender system. This subsystem consists of two agent which are the Data Manager Agent and the Aggregator Agent.The goal of the data manager agent is to maintain the authenticity of the data by preventing it from getting corrupted and also to manage the piping of data from source to the desired destination. This goal for the data agent is achieved by fulfilling two responsibilities i.e. the responsibility of properly encrypting and decrypting the data from the source and the destination respectively and by updating the proper locations of source and destination for the data to be used by the system. . The main task of the Aggregator agent is to channel between the user interface and the various servers for computation, storage and generating recommendations. This specific goal is achieved by the proper distribution and redistribution of data within the system.

### 4.4.2 Goal Model: Privacy Subsystem

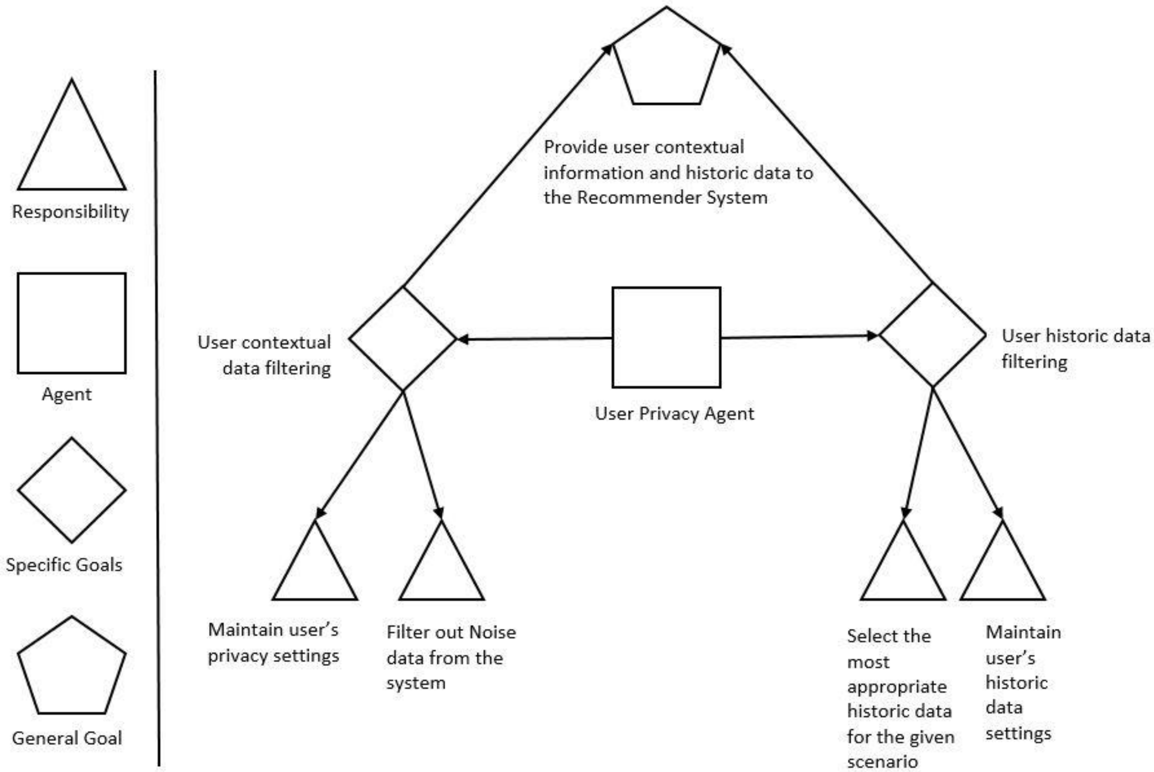


**Figure 6 Goal model for the user privacy agent**

The Privacy subsystem manages the privacy aspect of the web recommender system. This sub system consist of the User Privacy Agent to carry out its operations. The main role of this subsystem is to provide the user's contextual data and the historic data of the user to the computation server in order to generate recommendations for the users. The contextual information from the user can be in from of location, social information of the user, combined with the timing of the information. The user history data refers to the user's behavior while using the system that is being recorded for analysis.

To understand the role of the privacy subsystem within the recommender system mode, we need to look at the goals of the user privacy agent. The user privacy agent performs the task of maintaining user's privacy settings for the contextual data and also the responsibility of filtering out the noise from the contextual data being obtained from the user. These two responsibilities form the specific goal of filtering and maintaining user's contextual privacy information. On the other hand, the user privacy agent also fulfills the responsibility of maintaining the access to user's historic data based on the settings provided by the user and selecting the most appropriate data for generating the recommendations by filtering out the noise from the historic data.

### 4.4.3 Goal Model: Risk Subsystem

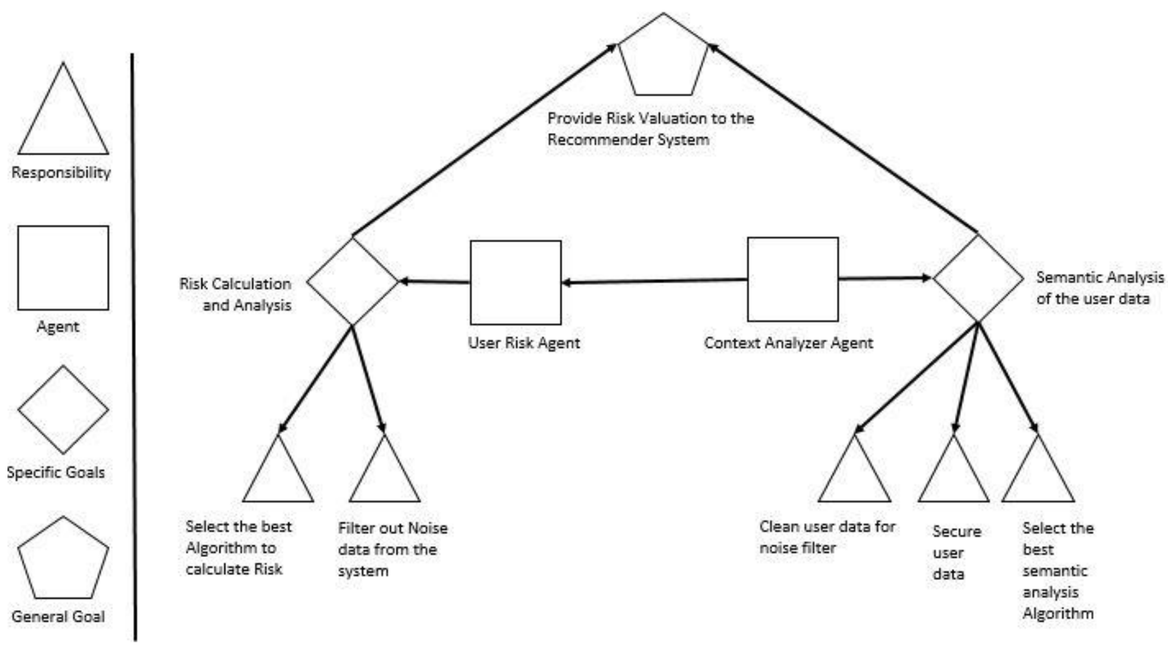


**Figure 7 Goal model for the User Risk & Context Analyzer agent**

This sub system handles the contextual risk by getting the contextual information i.e. time, location and social information from the user and then feeding this information to the recommender system. It consists of two agents the Context Analyzer Agent and the User Risk Agent.

The information processed in this step is utilized by the recommender system to generate more contextually aware system by not only providing more relevant information to its users but also keeping itself aware of the risk associated with disturbing or negatively affecting the user with the bad recommendation. This tradeoff of providing relevant recommendations and the associated risk is the part of risk calculation through the exploration and exploitation problem.

The two agents have some specific goals and responsibilities. The responsibility of the user risk agent is to ensure that no noise remains in the data and to calculate the risk tradeoff for generating the recommendations and relevance of those recommendations to the user from user feedback for the previously generated recommendations. These two responsibilities helps in achieving the goal of carrying out risk calculation and analysis of the user data. The context analyzer agent is responsible for cleaning the data obtained from the risk calculation stage, selecting the best possible algorithm for the analysis and then by securing the generated data to be forwarded as recommendations to the user. This helps in achieving the task of semantic analysis of the user data and finally providing the analysis results as recommendations to the user of the system.

### 4.4.4 Combined Goal Model of the System

The combined Goal model of the web recommender system consists of the aggregation of the individual subsystems and the coherence of the agents working within each working subsystem to achieve the goals of the entire system.

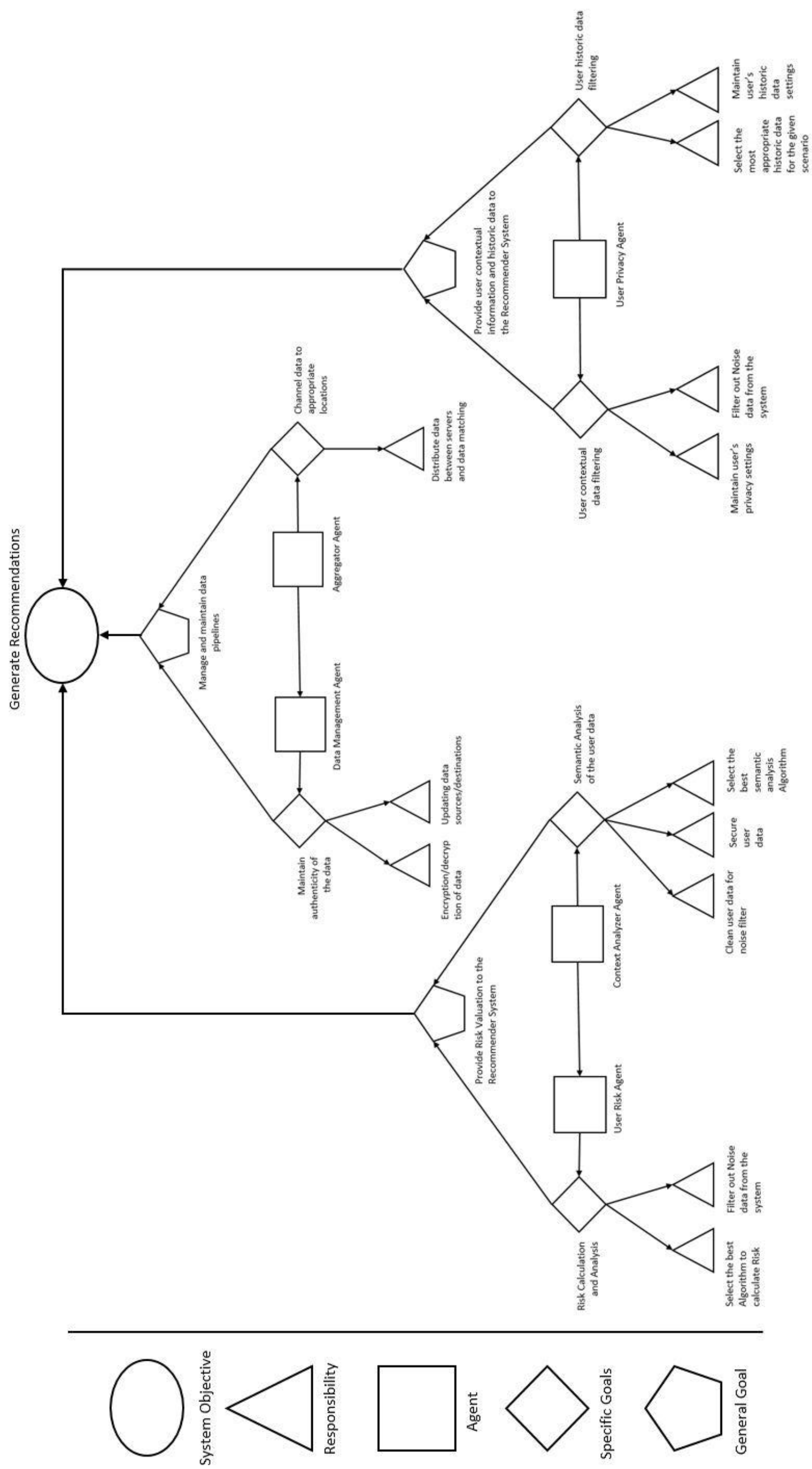


**Figure 8 System Goal Model**

## 4.5 Activity Models for the Subsystems

We will now discuss the Activity models of the subsystems which makes up a risk aware and privacy preserving web recommender system and also explain the contribution of each subsystems and the agents involved in the respective subsystems.

### 4.5.1 Activity Model: Data Subsystem

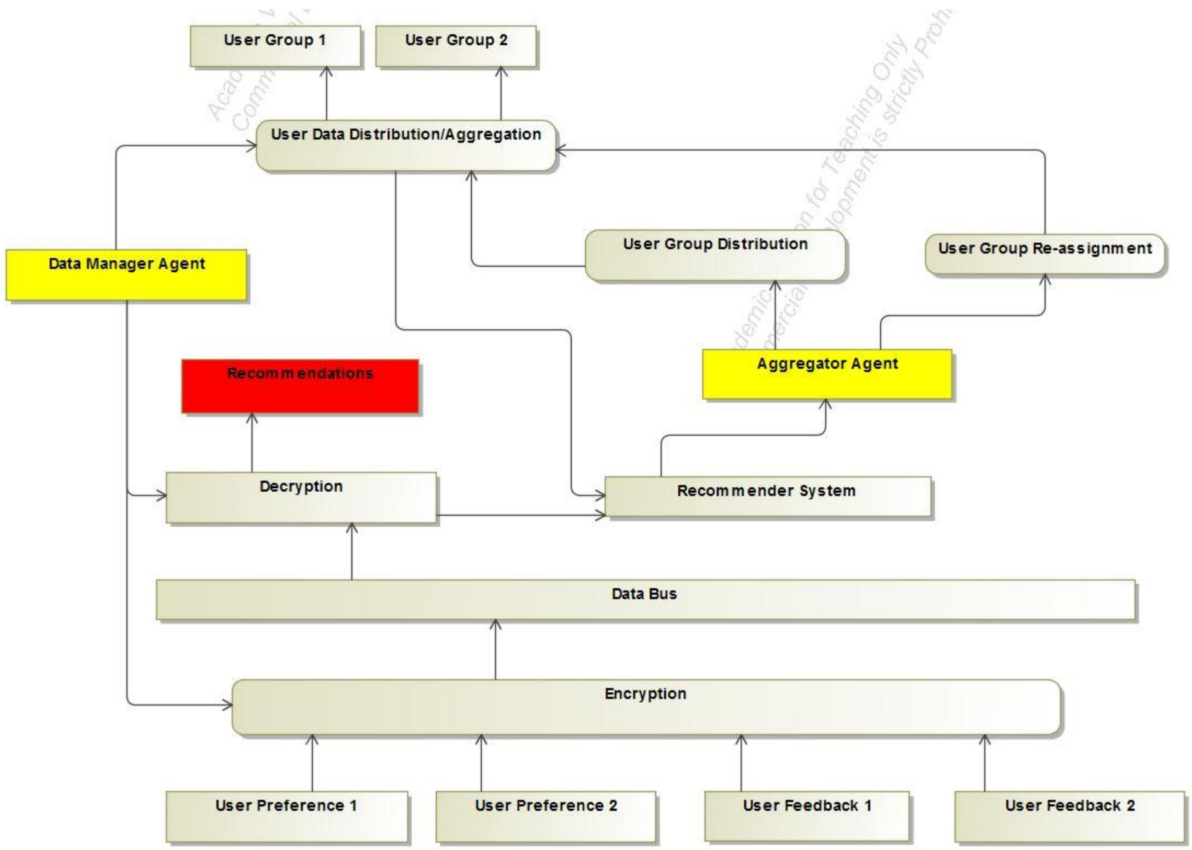


**Figure 9 Activity Diagram of Data Subsystem**

This subsystem absorbs data in form of User Preferences and User feedback. . It has multiple elements which performs the task that brings out the functioning of the data subsystems. The Data agent uses hashing, SHA, MD5 checking to ensure data authenticityThe best example of an Aggregator agent is a messaging brokers used in modern applications. Apache Kafa and RabbitMQ are two such message brokers. Together these two agents actives the objective of the data subsystem i.e. management and maintenance of the data pipelines within the system to enable the system.

### 4.5.2 Activity Model: Privacy Subsystem

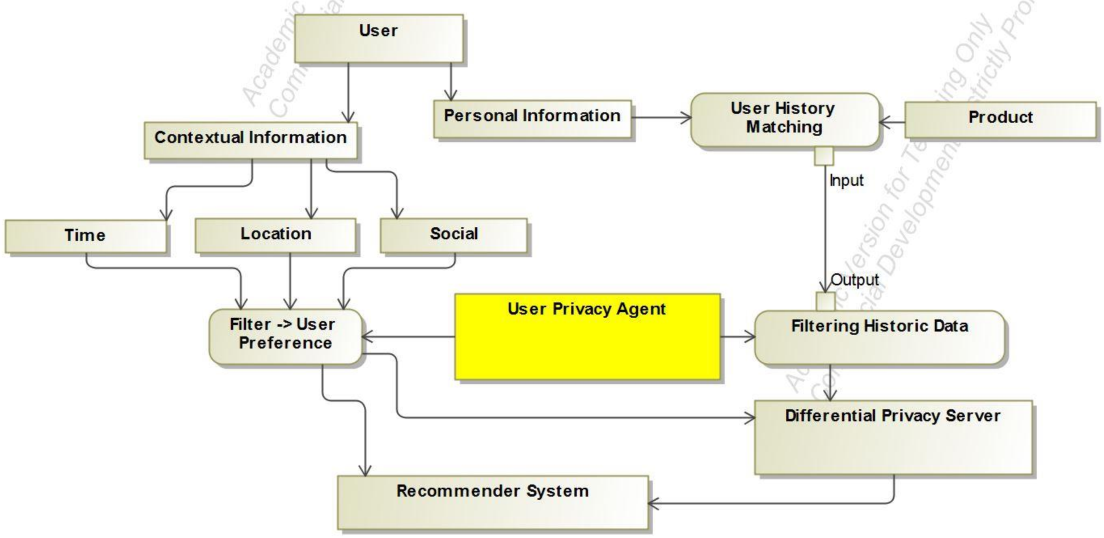


**Figure 10 Activity diagram for the Privacy Agent**

Within this subsystem the contextual and personal information is extracted from the user and fed into the recommender system. An addition differential privacy server is used to handle the differential privacy aspect of the subsystem. The contextual data from the user along with the historic data of the user provides valuable insights in order to provide quality recommendations to the user.

### 4.5.3 Activity Model: Risk Subsystem

The information processed in this step is utilized by the recommender system to generate more contextually aware system by not only providing more relevant information to its users but also keeping itself aware of the risk associated with disturbing or negatively affecting the user with the bad recommendation. This tradeoff of providing relevant recommendations and the associated risk is the part of risk calculation through the exploration and exploitation problem.

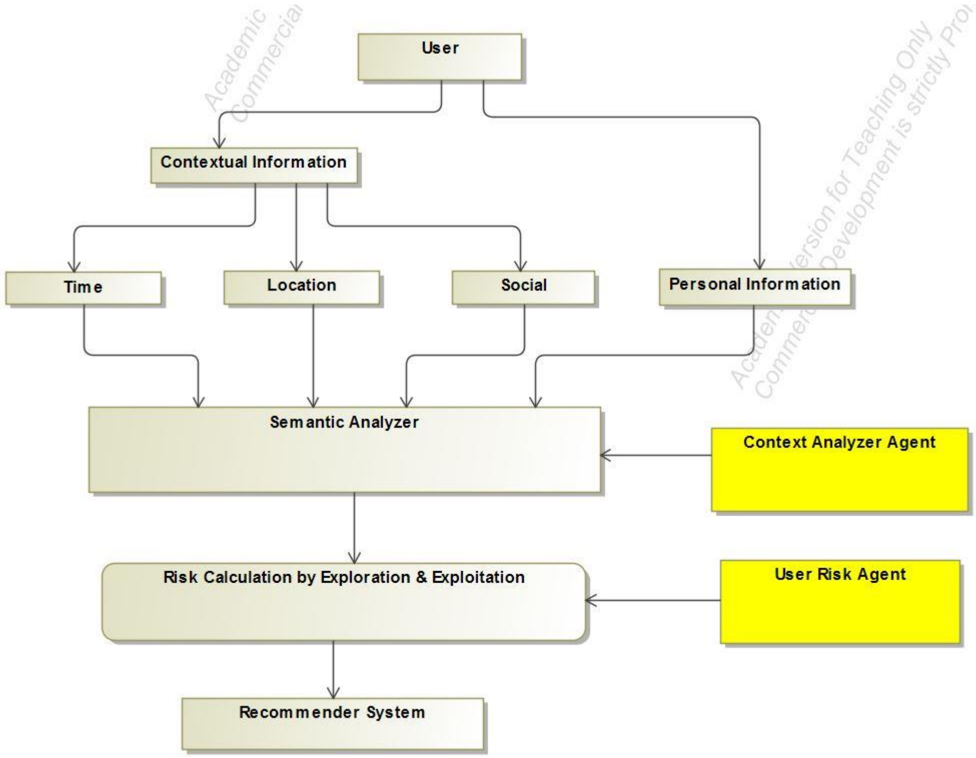


**Figure 11 Activity Diagram for the Risk Subsystem**

### 4.5.4 Combined Activity Model for the system

The combined Activity model of the recommender system consists of the aggregation of the individual subsystems and the coherence of the agents working within each working subsystem to achieve the goals of the entire system. The advantage of breaking down the web recommender system is to provide error detection and fault tolerance within the system. It also facilitates the understanding of the system in a clear sense. This model could be a better was of estimation of the value provided by the recommender system than the traditional way in the sense that it provides the domain experts with a better evaluation criteria.

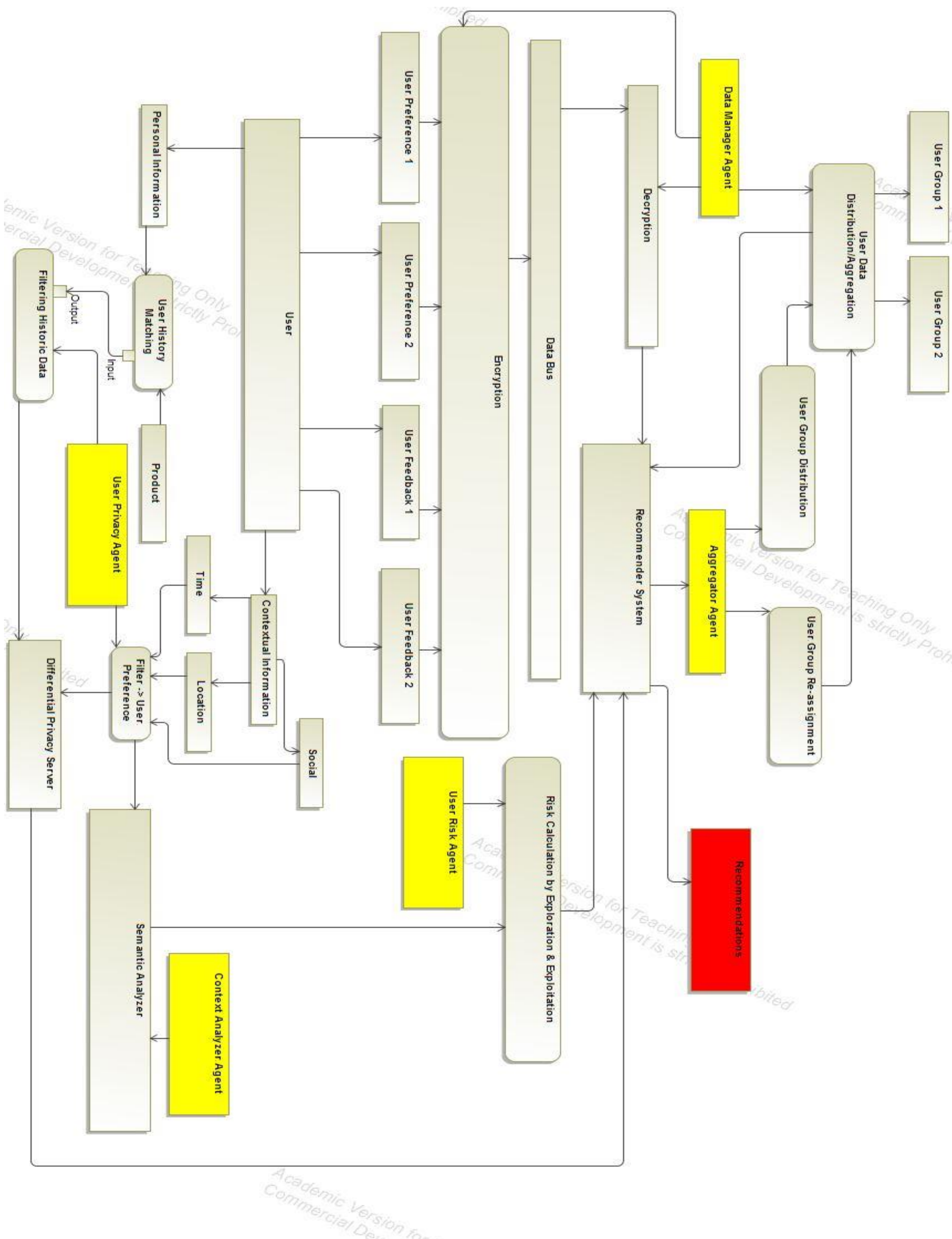


**Figure 12 Complete Activity model of the system**

### 4.6 Sequence Diagrams for the Subsystems

We will now discuss the Sequence Diagram of the subsystems which makes up a risk aware and privacy preserving web recommender system and also explain the sequence of actions that takes place within the subsystem.

#### 4.6.1 Sequence Diagrams: Data Subsystem

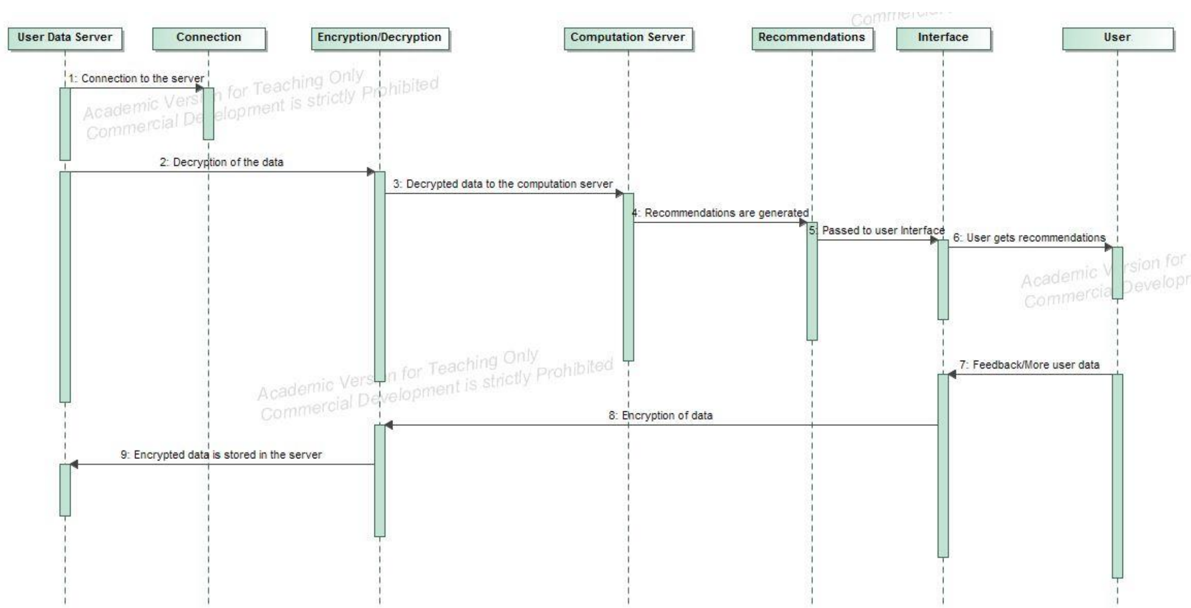


**Figure 13 Data Subsystem sequence diagram**

The sequence diagram of the data subsystem has been provided. In this diagram, a recommendation generation process starts when a connection is established between the user-data database and the computation server where the data to be used is decrypted. This data is then piped to the computation server. After the processing at the communication server, the recommendations are generated and are then forwarded to the user through an interface. Based on the quality of recommendation, the user provides a feedback which is stored in the user-data database. The transfer of data between the servers including the encryption and the decryption process is carried out within the data subsystem. These tasks can be assumed to be carried out by the data agent and the aggregator agent within the data subsystem, the outline of which has been provided in the previous section

### 4.6.2 Sequence Diagram: Privacy Subsystem

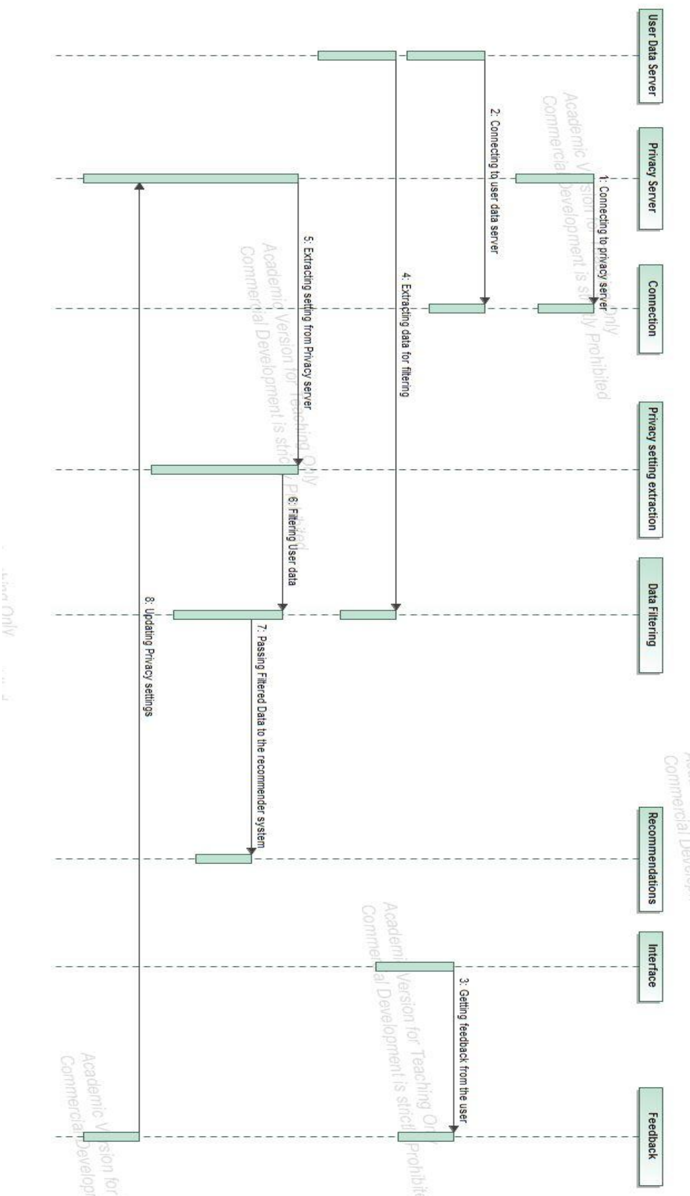


**Figure 14 Privacy subsystem sequence diagram**

In order to gain the understanding of the privacy subsystem is to get the knowledge of flow of control that happens within the subsystem. The first step involves establishing a connection with the user-data server and with the privacy server. This is followed extracting the user data and user privacy settings from the system. Once the data has been extracted from the server, it is filtered against the user settings. The user data includes the contextual data i.e. location, time and social data as well as the user's previous behavior pattern obtained while the user interacted in the system. The user is made aware of the data through the user controls and asking permission from the user to utilize the data for generating the recommendations.

Once the data has been filtered of the noise and against the user settings, it is piped through the computation server to generate the recommendations to the user. Once the recommendations has been generated, they are forwarded to the user via interface.

Based on the quality of the recommendations, the user provides a feedback or exhibits certain behavior pattern (clicks, navigation, dismiss) which indicates the user's perspective on the quality of the generated recommendations. This feedback data is then encrypted and stored in the user-data database to serve as an input future for the future computations for generating recommendations.

### 4.6.3 Sequence Diagram: Risk Subsystem

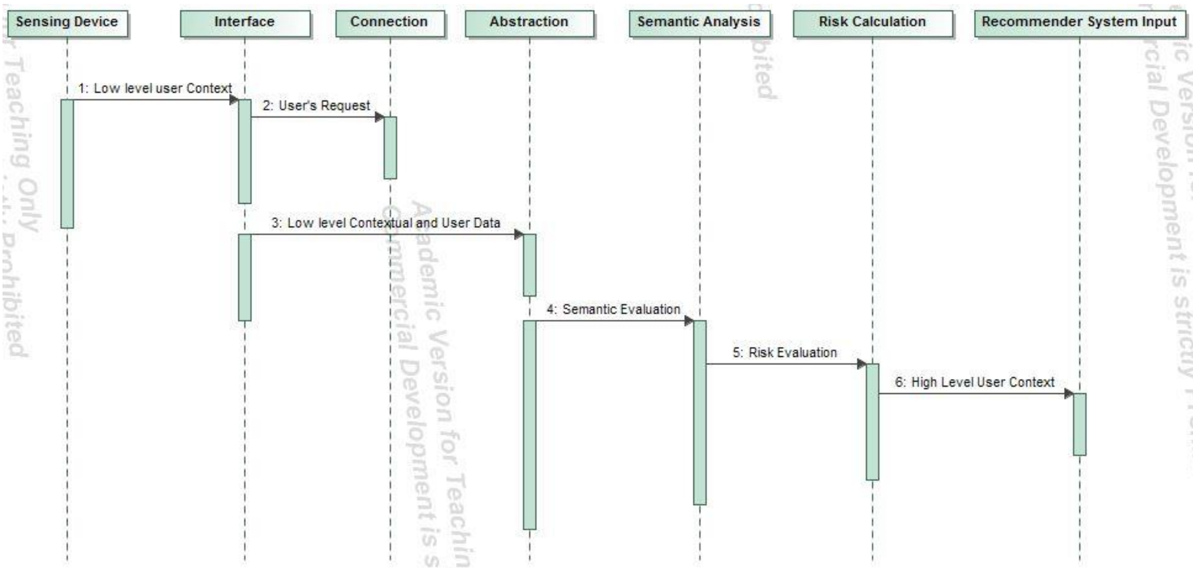


**Figure 15 Contextual Risk Subsystem Sequence Diagram**

The sequence diagram helps in understanding the steps that take place within the contextual risk subsystem. First, a connection is established with a sensing device at the user's end, through an interface. This step is followed by the low level abstraction of the user's data and feeding it to the servers running the semantic analysis. As a result of this, the risk is calculated and based on the value of this parameter the recommendations are forwarded to the user.

### 4.6.4 Combined Sequence Diagram

The Sequence diagram of the web recommender system consists of the aggregation actions taking place within the system to achieve the goals of the entire system.


**Figure 16 Combined Sequence Diagram**

# Chapter 5

# Case Study: Job Recommender System

Generally a recommendation system suggests personalized choices from a large set of possible options with the objective of reducing complex decision making. The last decade has witnessed the emergence of lots of job portals offering such services to their users. Generally a recommendation system works on information filtering technique and provides information which is of the interest to the concerned user. Typically, a recommendation engine, which employs a set of algorithms, compares the user's profile to some reference characteristics collected from the job description across multiple jobs posted on the job portal or the user's social environment, and seeks to predict a set of suitable jobs for the user.

## 5.1 Problem Description

We will now describe a recommender system proposed by [9] in 2013 and [10] in 2016. Paper [9] describes a hybrid recommender system for job seeking and recruiting websites. The described hybrid recommender system exploits the job and user profiles and the actions undertaken by users in order to generate personalized recommendations of candidates and jobs. The data collected from the website is modeled using a directed, weighted, and multi-relational graph, and the 3A ranking algorithm [16] is exploited to rank items according to their relevance to the target user. This paper also provides a preliminary evaluation based on simulated data and production data from a job hunting website in Switzerland. The approach in the paper consisting of modelling the entity-interaction based relations in the followed by the formation of a graph consisting of these entities and computation of ranking from this graph.

**Table 1 Interaction Entities proposed in [9]**

| User / Object | Candidate | Employer | Job |
|---|---|---|---|
| **Candidate** | Similar | Visit, Like Match, Favor, Apply | Visit, Like Match, Favor, Apply |
| **Employer** | Visist, Favourite Match | Similar,Visist | Post, Visit |
| **Job** | Match | Posted | Similar |

The technique used in the paper involves interaction based relations. The first of these relations is the 'POST' relation which is described as a bidirectional relation between the employer and its jobs which comes into play while comparing two similar jobs posted by different employers. The next relationship that is described in the paper is the 'APPLY' which signals that a candidate is interested in the job. This signal leads the candidate to other jobs similar to the ones he/she applied for. The next relationship that is described in the paper is 'FAVOURITE', using which, a user can add an entity into his/her 'favorite list'. This is also a strong and explicit signal of interest. Similar to the previous relationship, the 'LIKE' relationship with a difference that a user may not revisit the items they liked. In the paper, the 'LIKE' relationship is considered as an explicit feedback but weaker than 'APPLY' or 'favorite'. The final relationship signal that is discussed in the paper is 'VISIT' which is an implicit feedback for user's interest.

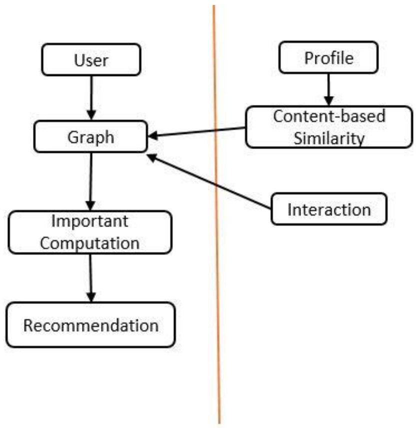


**Figure 17 Graph Framework described in [9]**

A pipelined hybrid recommendation approach is described and implemented in [9] along with providing the result of content-based similarity which is fed into a relation-based algorithm as an additional relation after normalization. The above figure shows the recommendation framework described in the paper for generating personalized job recommendations to the users.

On the other hand, paper [10] describes a resume matching system which intelligently extracts the qualifications and experience of a job seeker directly from his/her résumé, and relevant information about the qualifications and experience requirements of job postings. Using a novel statistical similarity index the resume matching system returns results that are more relevant to the job seekers experience, academic, and technical qualifications, with minimal active user input.

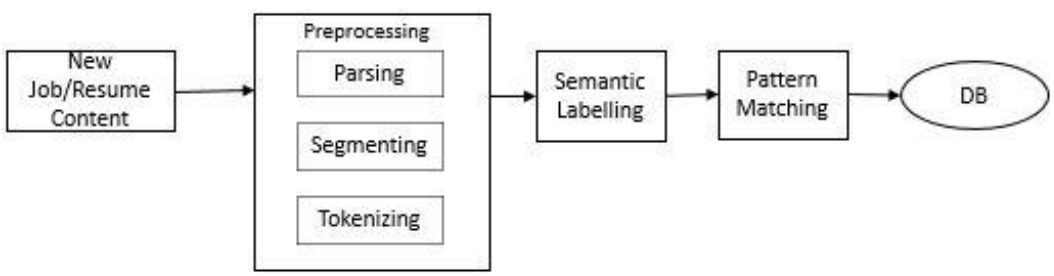


**Figure 18 Resume matcher System as described in [10]**

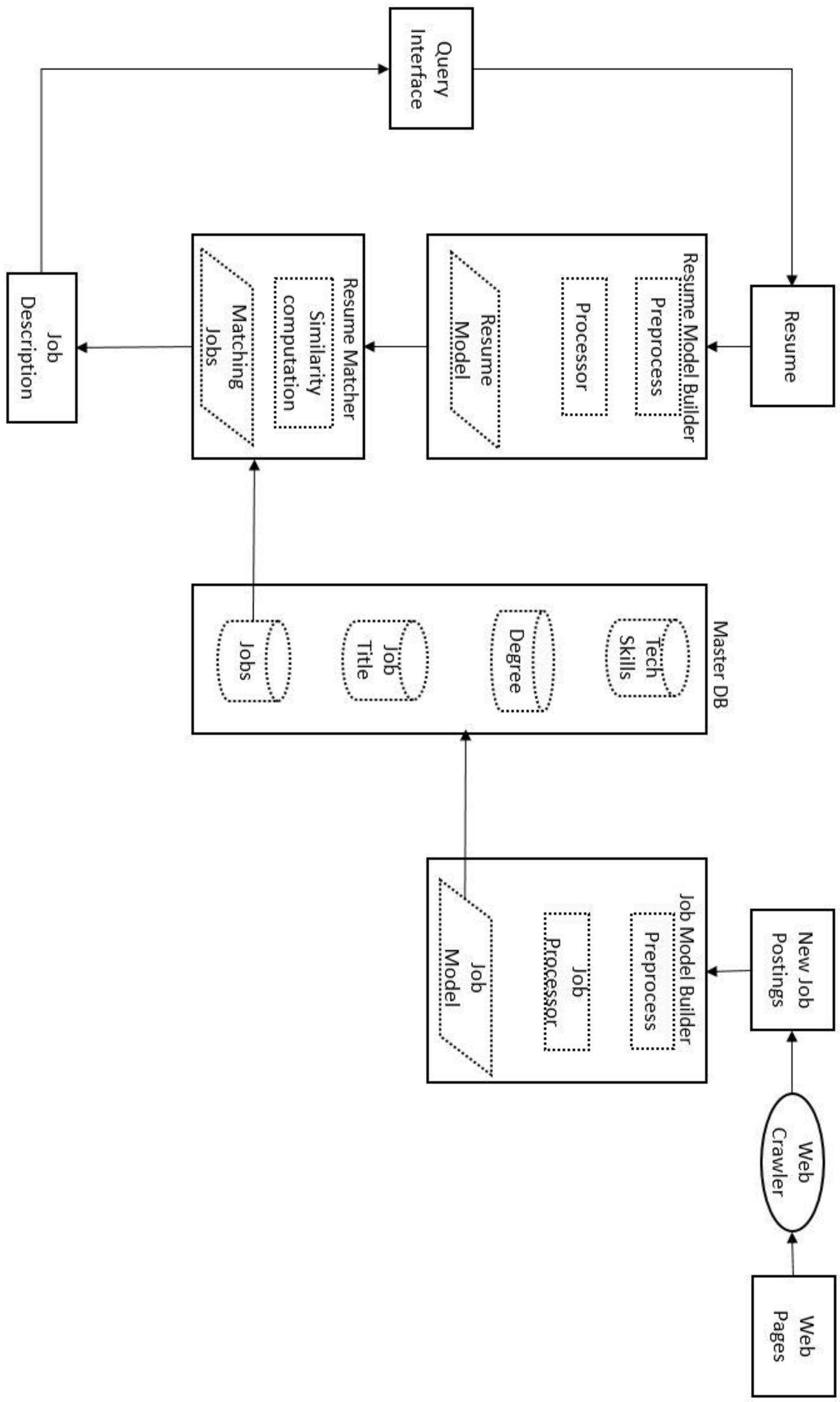


**Figure 19 Information Processing Pipeline described in [10]**

## 5.2 Approach

Representation of the two job recommender systems that are described in the previous section and combining them in a model of a job recommender system which can be efficiently described by the approach used in this thesis will be discussed in this section. The first step towards this description is to determine the features of the described recommender systems and then laying out those features in terms of the discussed approach. This involves breaking down the recommender system and focusing on the multi-agent aspect of the system and separating different components of the system into different subsystems and finding out a way to integrate the subsystems into one compact unit.

## 5.3 Goal Models of the subsystems

This subsystem has two responsibilities. The first responsibility is to encrypt the data obtained from the employers and/or the candidate and store it in the database and also making it available it for use by fetching it from the system and decrypting it. The second responsibility is to not only maintain the pipelines of candidate's data and employer's data within the system but also help in anonymization of the data by piping it through the differential privacy servers. These responsibilities gives rise to two goals of the system i.e. maintain the authenticity of the data and the channelling of the data through the system while protecting it as well. These tasks are performed by the data management agent and the aggregator agent. The end goal of this subsystem is to manage and maintain the data pipelines thought the subsystem. As before there are multiple software libraries used within this subsystem. Some examples of such software are messaging brokers like Apache KAFKA, RabbitMQ etc. working within a distributed system framework like Hadoop Distributed File Fystem.

### 5.3.1 Goal Model: Data Subsystem

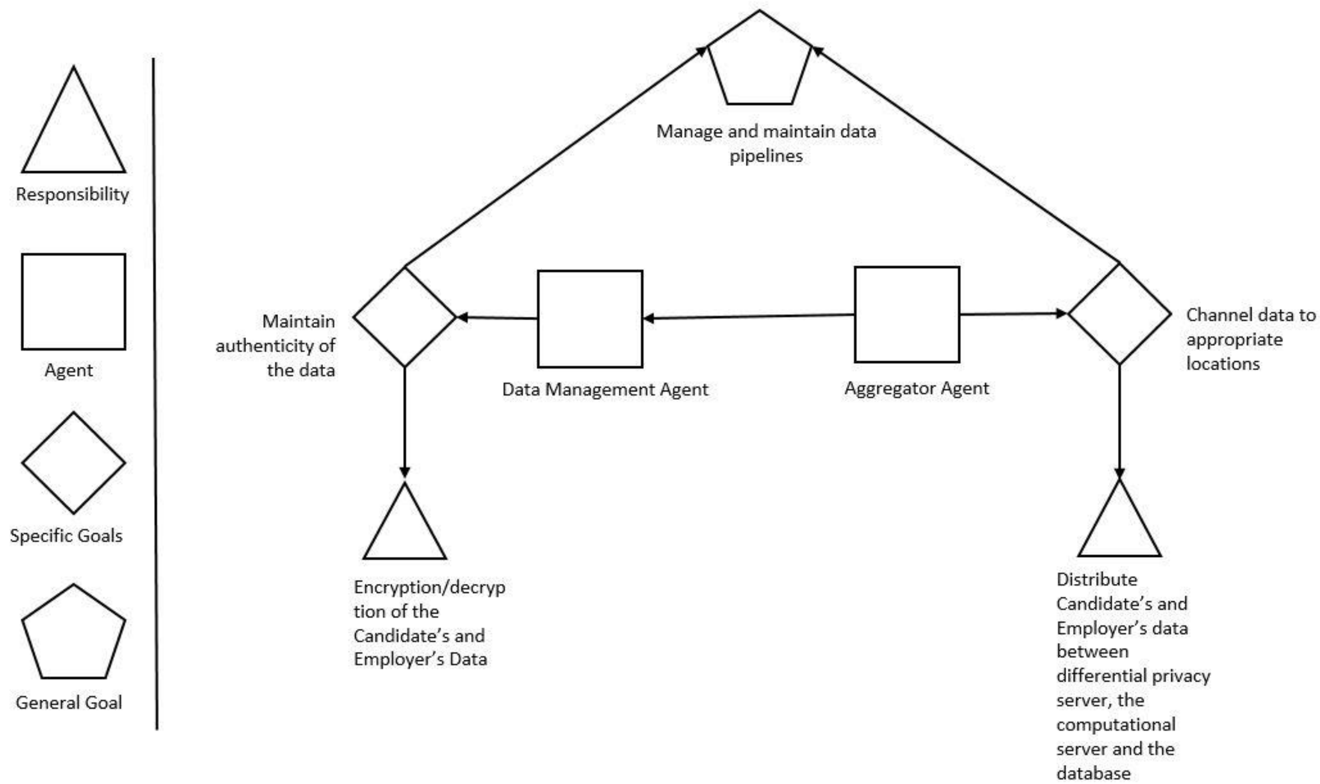


**Figure 20 Goal Model: Data Subsystem**

### 5.3.2 Goal Model: Privacy Subsystem

The privacy subsystem has a user privacy agent. The goals of the privacy agent is to filter the contextual data and the historic data. But the filtering has to be carried out by fulfilling the responsibilities. The first responsibility is to maintain the privacy settings of the employers and the candidate's data in the system. This is followed by the responsibility of filtering out contextual data based on the privacy settings. This is followed by maintaining the historic data setting for both type of the users and then filtering out the historic data based on those settings. These goals and responsibilities helps in achieving the goal of the privacy subsystem i.e. provide filtered user contextual data and user historic data to the Recommender system.

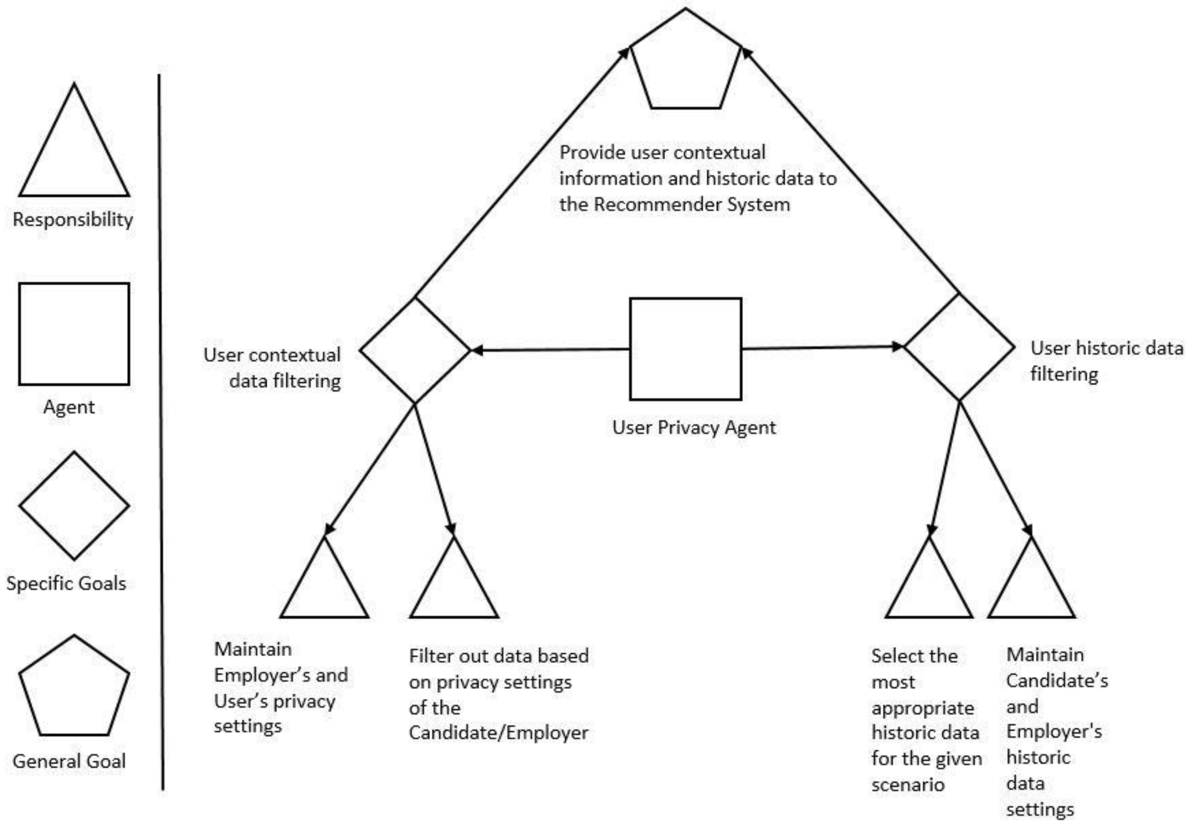


**Figure 21 Goal Model: Privacy Subsystem**

### 5.3.3 Goal Model: Risk Subsystem

T Risk subsystem has two agents, the user risk agent and the Context Analyzer Agent. The goal of the user risk agent is to calculate the risk factor for contextual data. The goal of the context analyzer is to carry out the semantic analysis of the user data. These goals help in fulfilling responsibilities. The responsibilities are calculating risk using candidate's profile information, employer's job description, extract relavent information from the candidate's profile and the job description and then use a anlysis /matching algorithm for the current scenario. The overall objective of the subsystem is to provide risk valuation to the recommender system.

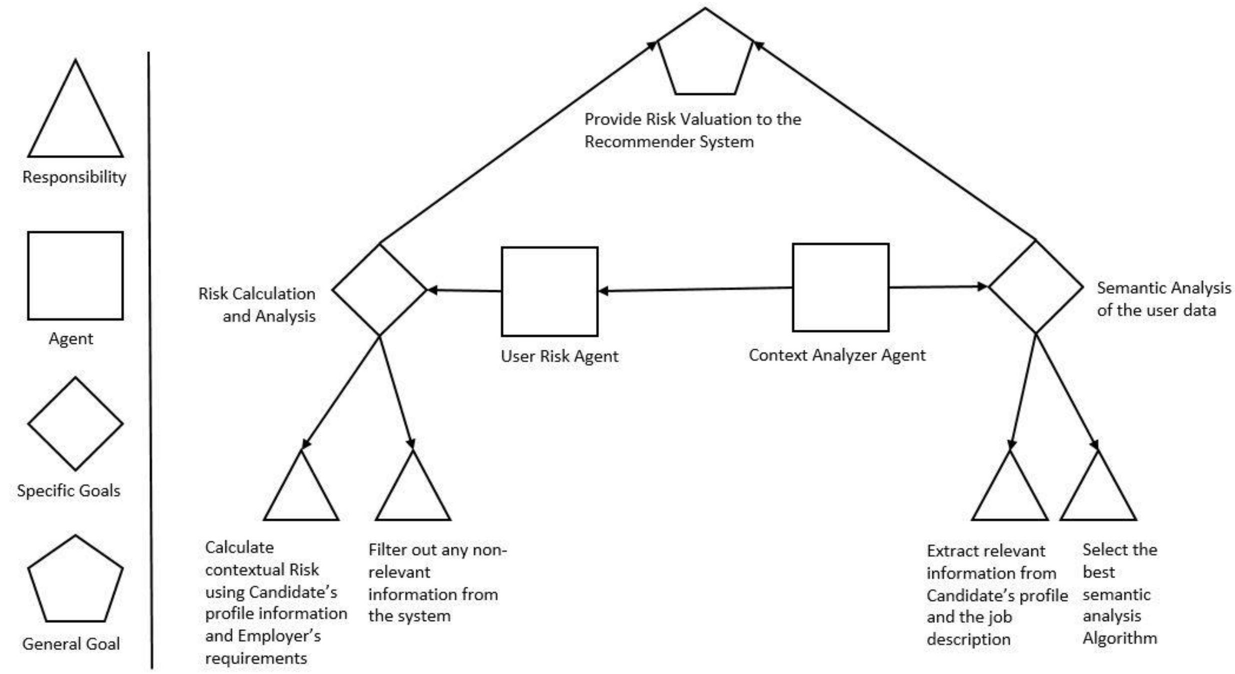


**Figure 22 Goal Model: Risk Subsystem**

### 5.3.4 Combined Goal Model of the System

The combined Goal model of the web recommender system consists of the aggregation of the individual subsystems and the coherence of the agents working within each working subsystem to achieve the goals of the entire system. The risk subsystem, the data subsystem and the privacy subsystem in the recommender system accomplishes the goal of the entire system by generating recommendations.

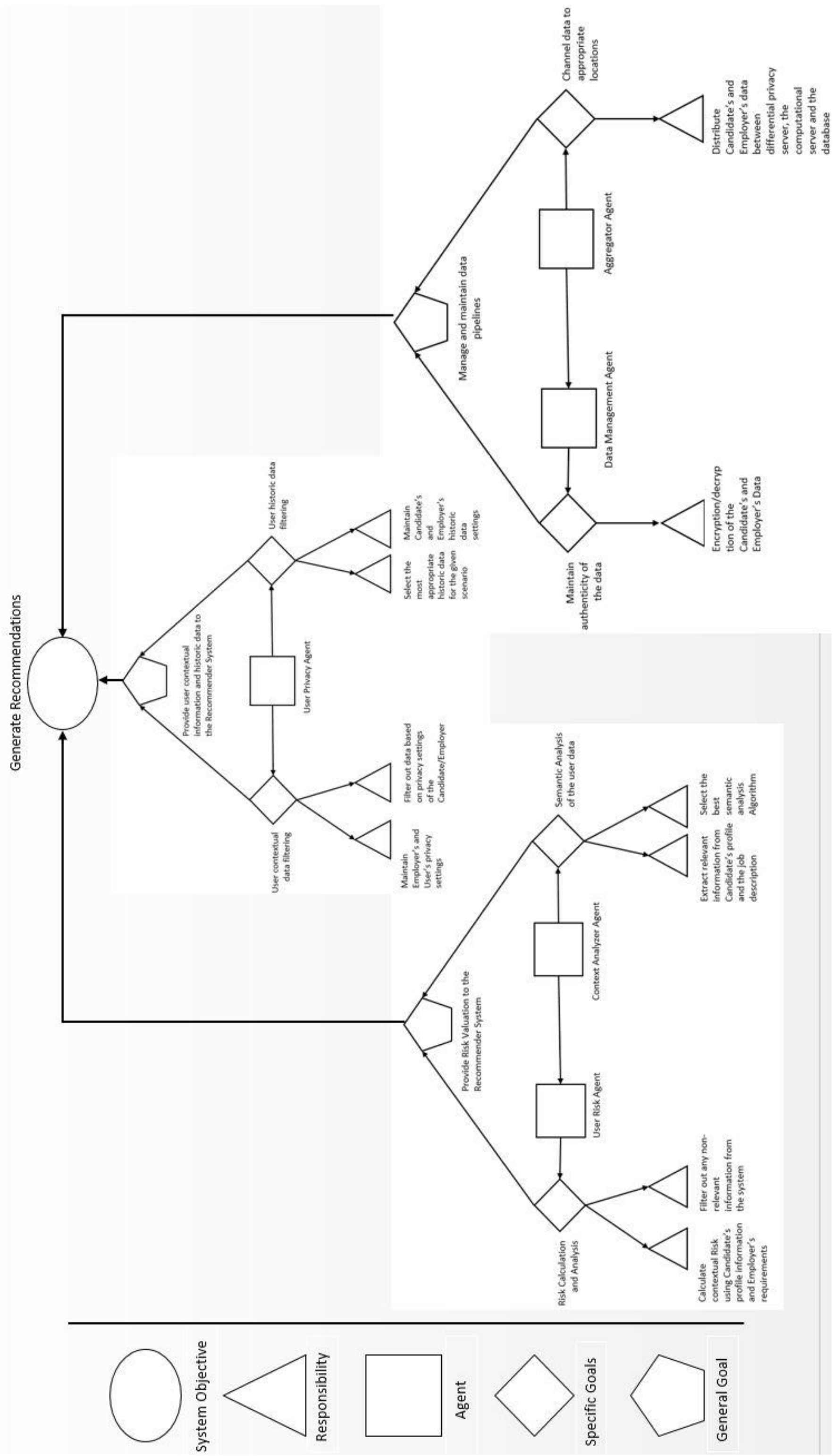


**Figure 23 Combined Goal Model of the Job Recommender System**

## 5.4 Activity Models of the subsystems

### 5.4.1 Activity Model: Data Subsystem

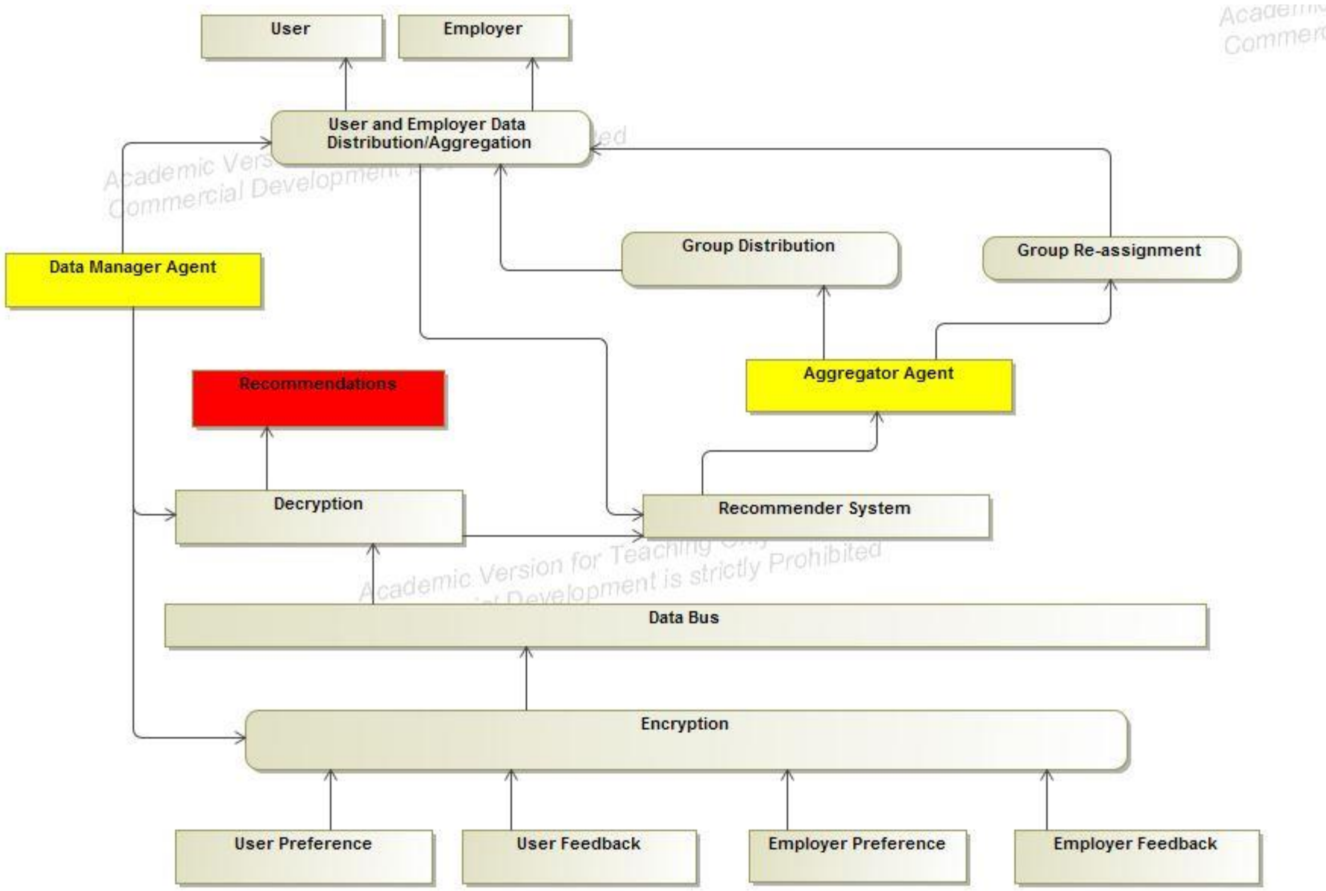


**Figure 24 Job Data Agent**

The Data subsystem manages the data flow within the recommender system. It manages the data from the candidate and the employer and the subsequent distribution of that data between different channels and filters such as noise filters which is usually followed by encryption/decryption. This is one of the most important subsystems and probably serves as the backbone of the entire system.

### 5.4.2 Activity Model: Risk Subsystem

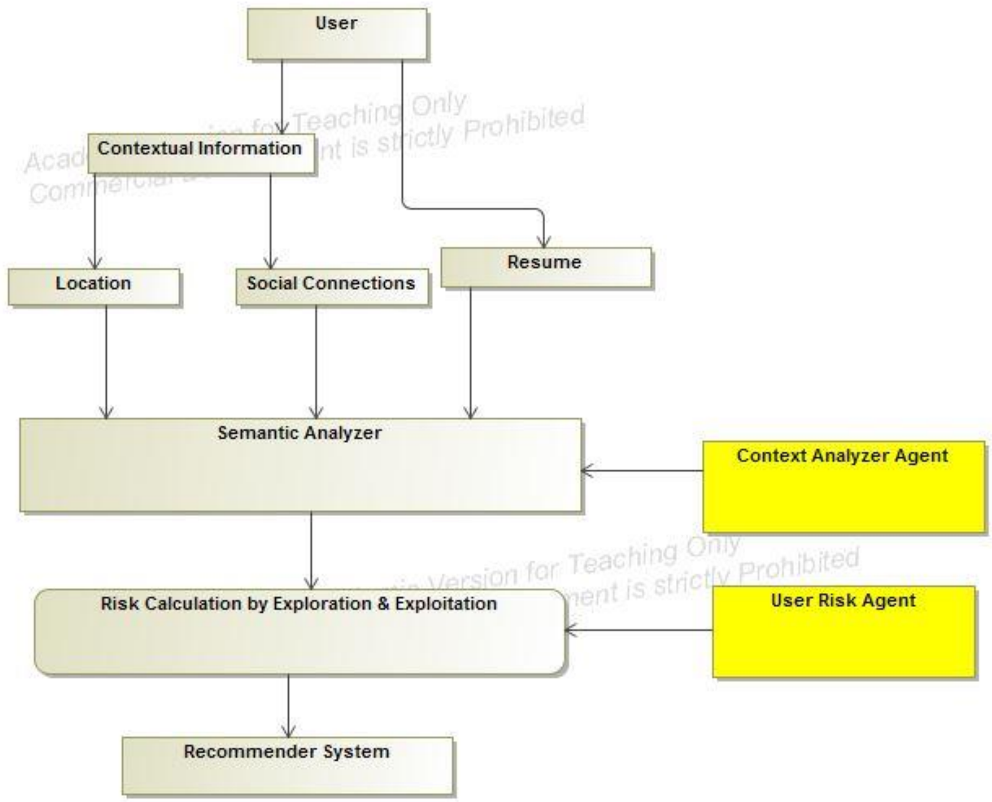


**Figure 25 Risk Agent for job recommender**

The contextual risk subsystem, as described earlier provides the risk calculation in order to generate suitable recommendations by the recommender system. The contextual information in the job recommender system is the location of the candidate and the employer and the social connections of the candidate. As described in the earlier sessions, this system consists of two agents the Context Analyzer Agent and the User Risk Agent. The information processed in this step is utilized by the recommender system to generate more contextually aware system by not only providing more relevant information to its users but also keeping itself aware of the risk associated with disturbing or negatively affecting the user with the bad recommendation. This tradeoff of providing relevant recommendations and the associated risk is the part of risk calculation through the exploration and exploitation problem.

### 5.4.3 Activity Model: Privacy Subsystem

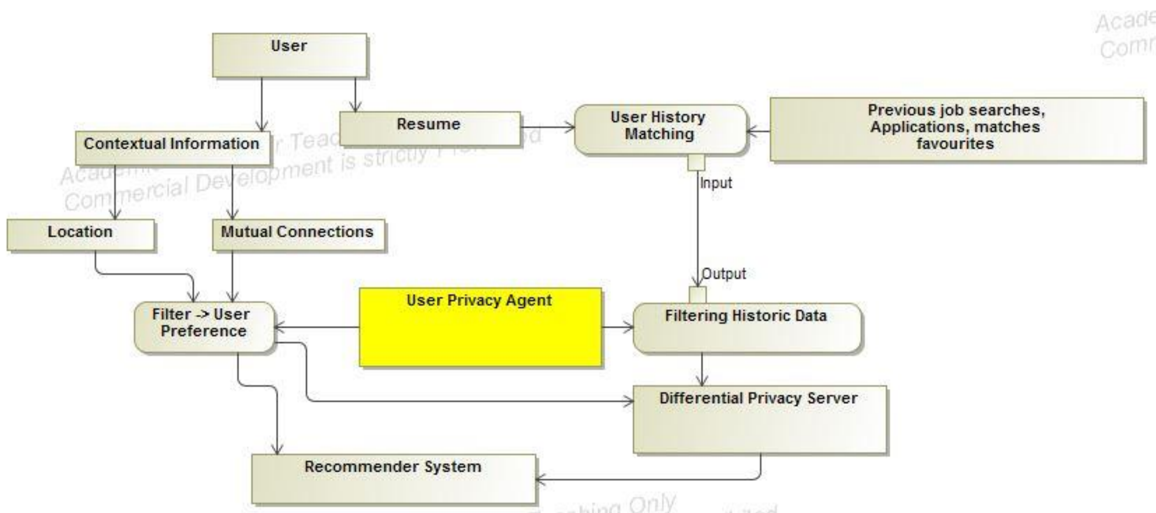


**Figure 26 Privacy Agent for job recommender**

The above diagram is the activity model of the privacy subsystem of the web recommender system. Within this subsystem the contextual and Resume Information in form of resume information is extracted from the user and fed into the recommender system. A differential privacy server manages the anonymity of data within this subsystems by implementing privacy differential algorithms. The main role of this subsystem is to provide the user's contextual data, personal information and the historic data i.e. favorites, visits and applications, of the user to the computation server in order to generate recommendations for the users. The user history data refers to the user's behavior while using the system that is being recorded for analysis. The contextual data from the user along with the historic data of the user presents valuable insights in order to provide quality recommendations to the user.

### 5.4.4 Combined Activity Model of the system

The combined Activity model of the job recommender system consists of the aggregation of the individual subsystems and the coherence of the agents working within each working subsystem to achieve the goals of the entire system. The advantage of breaking down the web recommender system is to provide error detection and fault tolerance within the system. It also facilitates the understanding of the system in a clear sense. This model could be a better was of estimation of the value provided by the recommender system than the traditional way in the sense that it provides the domain experts with a better evaluation criteria.

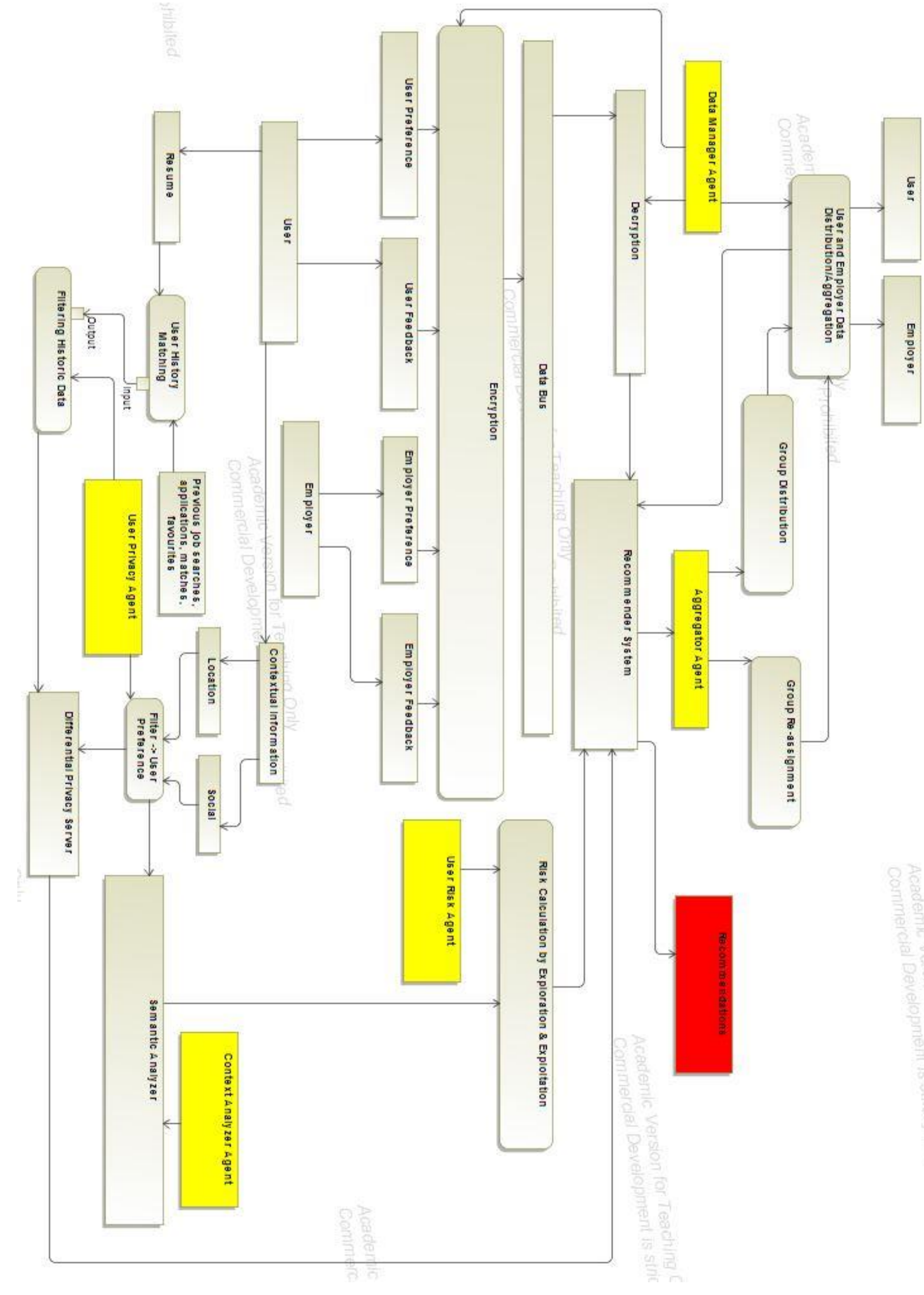


**Figure 27 Job recommender system model**

## 5.5 Sequence Diagram for the Subsystems

### 5.5.1 Sequence Diagram: Data Subsystem

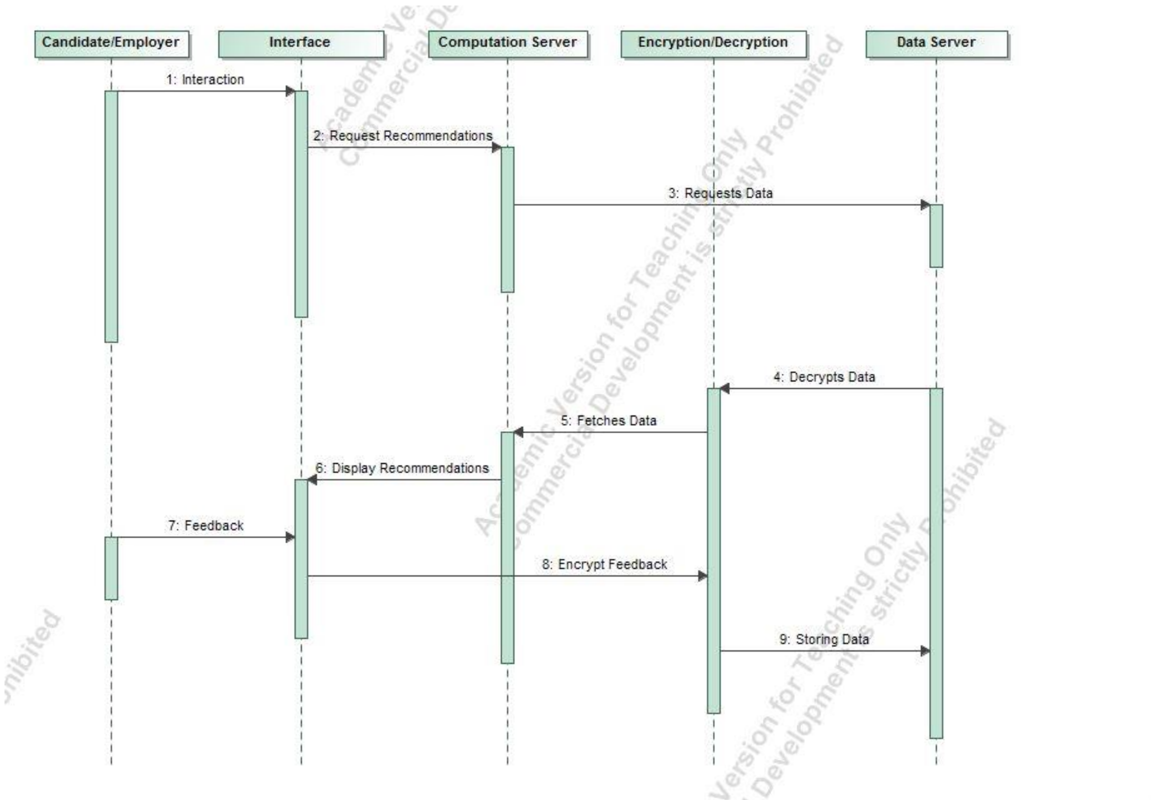


**Figure 28 Sequence Diagram: Data Subsystem**

The sequence diagram of the data subsystem has been provided. The process within the data subsystem is initiated when the candidate interacts with the interface. The interface can be the website or a mobile device. The data from the interface is sent to the computation server from where the recommendations are generated. The data is then encrypted and stored in the data server. The recommendations are forwarded to the interface and the feedback is obtained in order to enhance the recommendations. This data is again stored in the database.

### 5.5.2 Sequence Diagram: Risk Subsystem

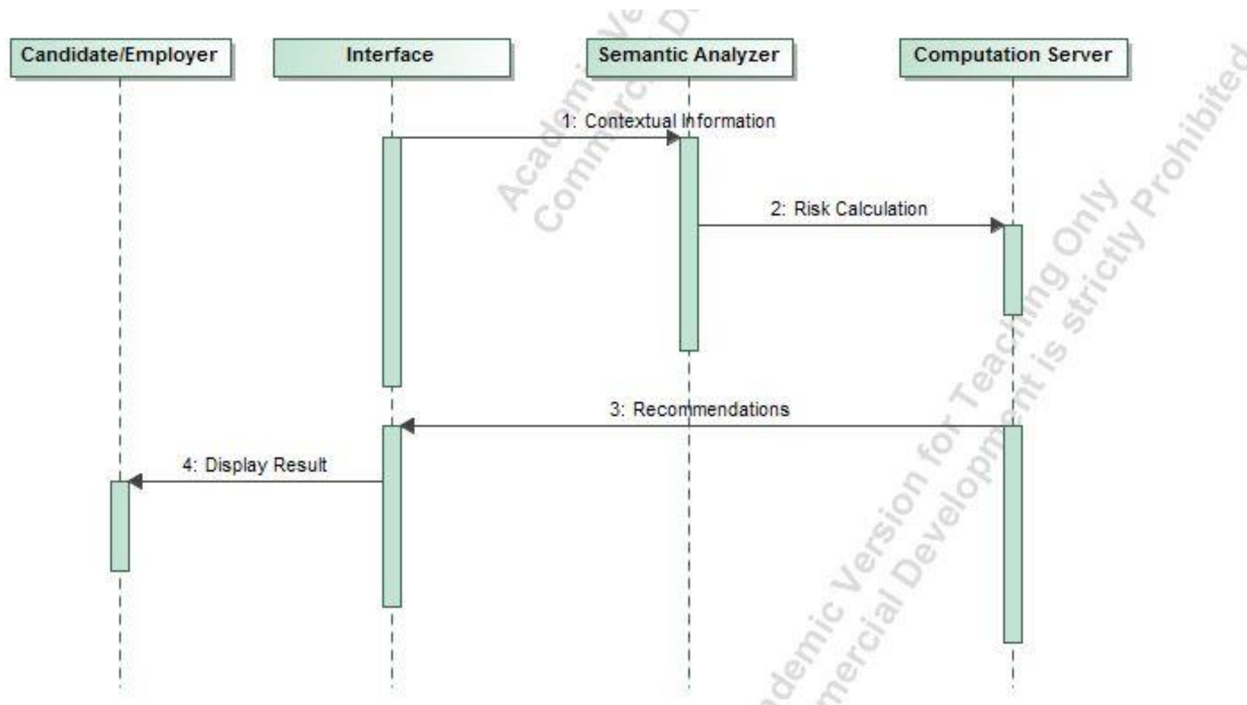


**Figure 29 Sequence Diagram: Risk Subsystem**

The sequence diagram of the risk subsystem is shown above. The contextual information is fed into the computation server through the interface, which is passed through a semantic analyzer. Based on the algorithms on the computation server, the recommendations are generated and forwarded to the interface and displayed to the user.

### 5.5.3 Sequence Diagram: Privacy Subsystem

The Sequence diagram of the privacy subsystem represents the sequence of processes within the system. The contextual data is first passed through a privacy filter before travelling to the database or the server. The filtered data is recovered from the database for the purpose of generating recommendations. It is passed through a differential privacy server to introduce anonymity. This data is then processed by the recommendation server to generate recommendations and pass it to the user through an interface.

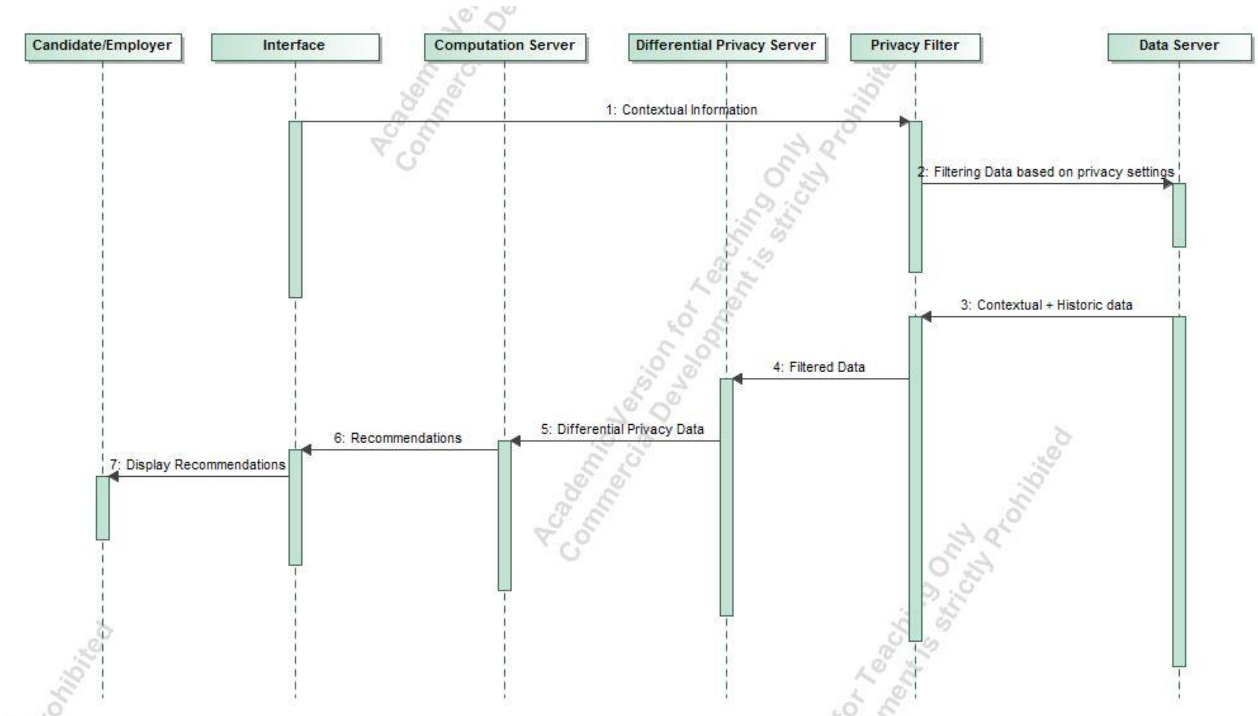


**Figure 30 Sequence Diagram: Privacy Subsystem**

### 5.5.4 Combined Sequence Diagram of the System

The Sequence diagram of the web recommender system consists of the aggregation actions taking place within the system to achieve the goals of the entire system.

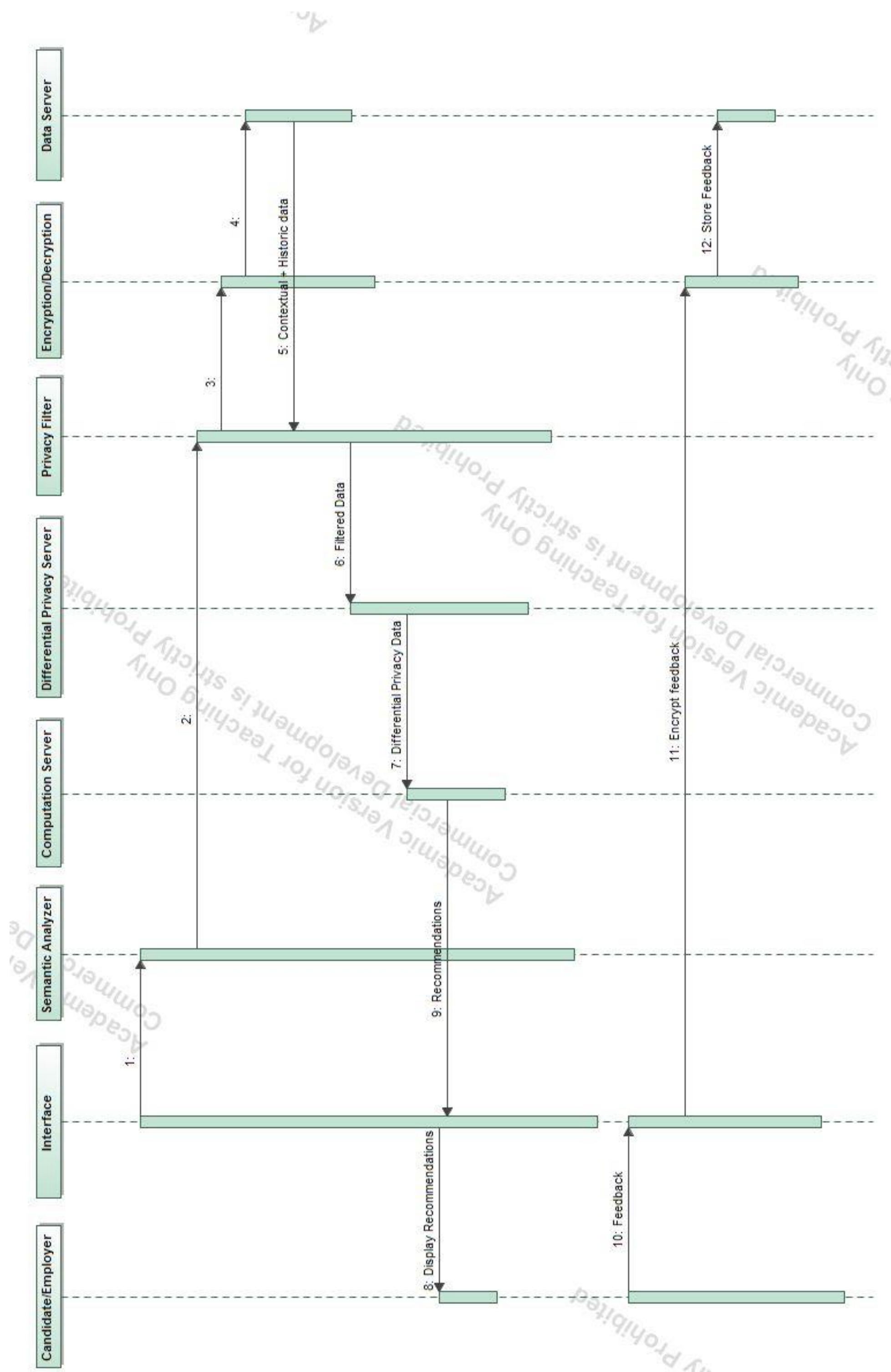


**Figure 31 Combined Sequence diagram of the Job Recommender System**

# Chapter 6

# Conclusion & Future Work

## 6.1 Conclusion

We have now show how the proposed approach provides a broader aspect to the system description and a prospect for evaluation of a recommender system across multiple application areas.  We have also shown that this approach utilizes multi-agent system description in a sense that the designers of the recommender systems can focus on individual units by breaking the recommender system into small individual units enables fast and fault tolerant development of the system. It also enables the designers of the recommender system to be aware of the each of the small objectives that must be accomplished by the each individual units in order to fulfil the objective of the entire system. This high level approach to describe the system is helpful for domain experts to gain valuable knowledge of a recommender system, operational in a particular application area in a short period of time.

## 6.2 Future Work

The multi-agent approach can be extended in the future to include other domains of recommender system such as news recommender or restraint recommender system. The system can be represented with other UML models like a state model or class diagram. Frameworks can be implemented to use domain specific language to generate class structure for the recommender system or can be used to generate code. But most importantly, a model verification method, possibly through experimentation can be used to enhance the model.

# Appendix

This section discusses a model based approach for the purpose of evaluation of the recommender systems. Traditionally, recommender systems are evaluated based on the accuracy of the results produced by the system but, using this approach, the recommender systems can be evaluated based on the features they possess.

## Privacy scope of a system

We introduce a coordinate system to describe the state of a web recommender system in terms of the privacy it offers to the user. It is a 3 dimensional representation with each of the mutually independent axis representing the state of the recommender system. On one of the axis we have a feature which states the size of the audience to which recommendations will be disclosed using data of a participant in the system i.e. if a user allows the system to use his/her data, then how many people other than the user, will be able to receive the recommendation based on that user's data in a collaborative environment. The extent of usage axis refers to the amount of information that is extracted from each participant in the system. The third and the final axis represents the duration for which the data remains in the system.

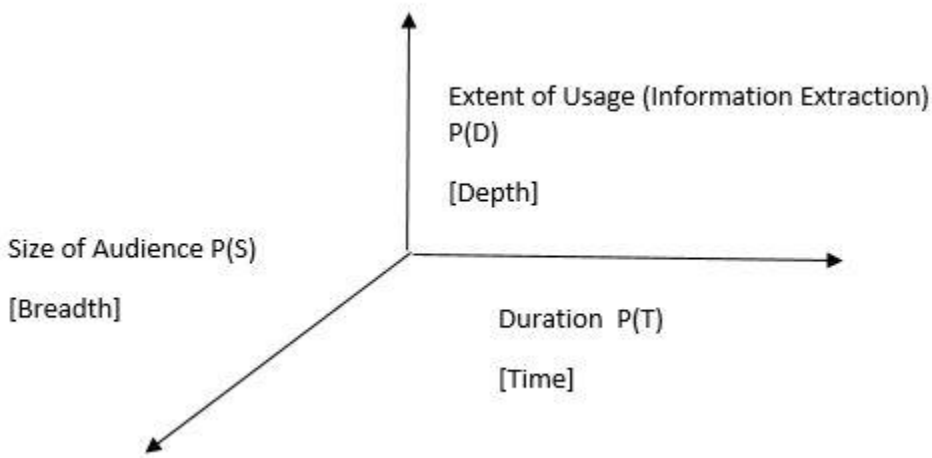


**Figure 32 Privacy Scope**

## Contextual risk scope

This section describe the contextual risk scope of a recommender system. Similar to the description of the privacy scope, the contextual scope is also a three dimensional representation for the purpose of characterizing recommender systems. The three axis of the contextual risk scope are mutually independent. The first axis is the similarity axis denoted of the R(s) notation. It is defined as the extent of similarity between the user and the user group into which the user is placed. The second axis denoted by R(C) is the axis of intention and is described as the extent of awareness of the user's intention by the system. This axis is conceptual i.e. the valuation provided by the recommender system based on this metric is highly based on experimentation results. The third and the last axis is the axis of duration and is the measure of how long the contextual data will be stored by the system.

This axis is represented by the notation R(T) and may also represent the period of data used by the recommender system for the purpose of generating recommendations.

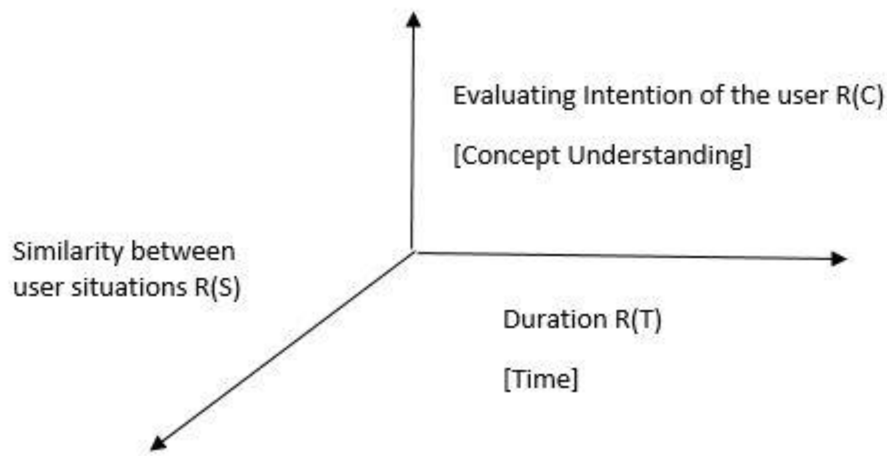


**Figure 33 Contextual Risk Scope**

### Explanation of a multidimensional system diagram

We are now in position to describe a web recommender system in a five dimensional representation. Parallel coordinates is a visualization technique used to plot individual data elements across many dimensions. Each of the dimensions corresponds to a vertical axis and each data element is displayed as a series of connected points along the dimensions/axes. Thus, a recommender system can be described as a series of connected points along the diagram, intersecting at these axis/dimensions.

Since it is a start of a research for visualizing a recommender system by these axis, there is the dearth of data for exactly calculating the exact value of a particular recommendation. Hence, through this thesis, an approximate representation is used for visualizing a recommender system by using this method.

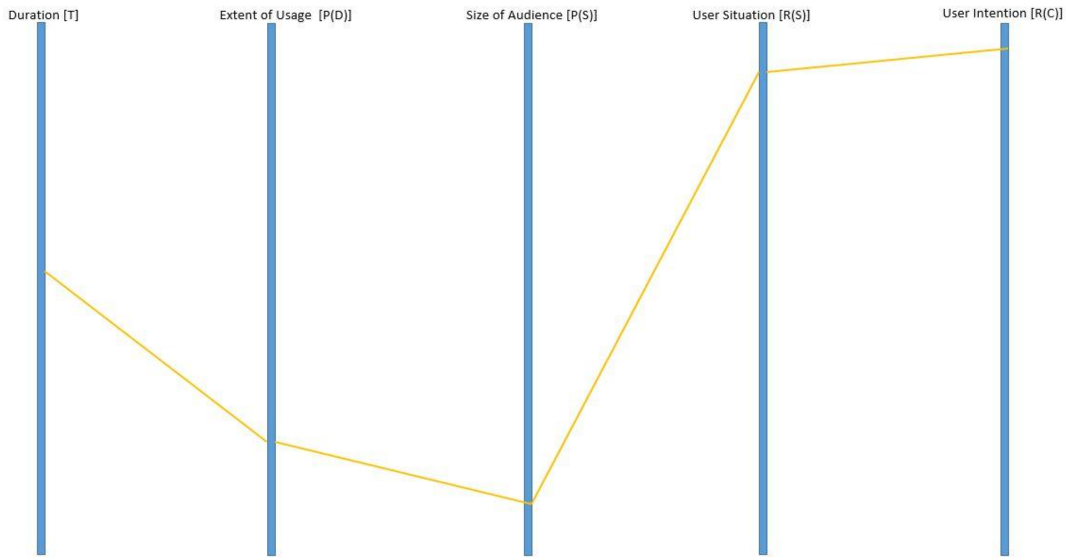


**Figure 34 Dimensional Plot of a recommender System**

**Table 2 General Dimensional Analysis of various Algorithms**

| | Duration | Extent of Usage | Size of Audience | User Situation | User Intention | Dimensions |
|---|---|---|---|---|---|---|
| Collaborative filtering | o | o | Nil | Nil | O | 3 |
| Demographic | o | o | Nil | o | Nil | 3 |
| Context-aware | O | O | o | O | o | 5 |
| Hybrid | O | O | o | O | o | 5 |
| Social | o | O | O | o | O | 5 |

In the above table an approximate idea has been provided about the possible dimensions that can be used by the recommender system using different methodologies, architecture and algorithms. The conclusion listed above is a result of review of current text in the area of web recommender systems. The extent of utilization/valuation of different metrics on the five dimensional recommendations is being distinguished in the table by using two different notations. The ‘O’ symbol is used where the utilization/valuation on a particular dimension is conceptually high and the ‘o’ notation is used for those system whose utilization/valuation on a particular dimension is relatively low.

## Extension of the evaluation method to the case study

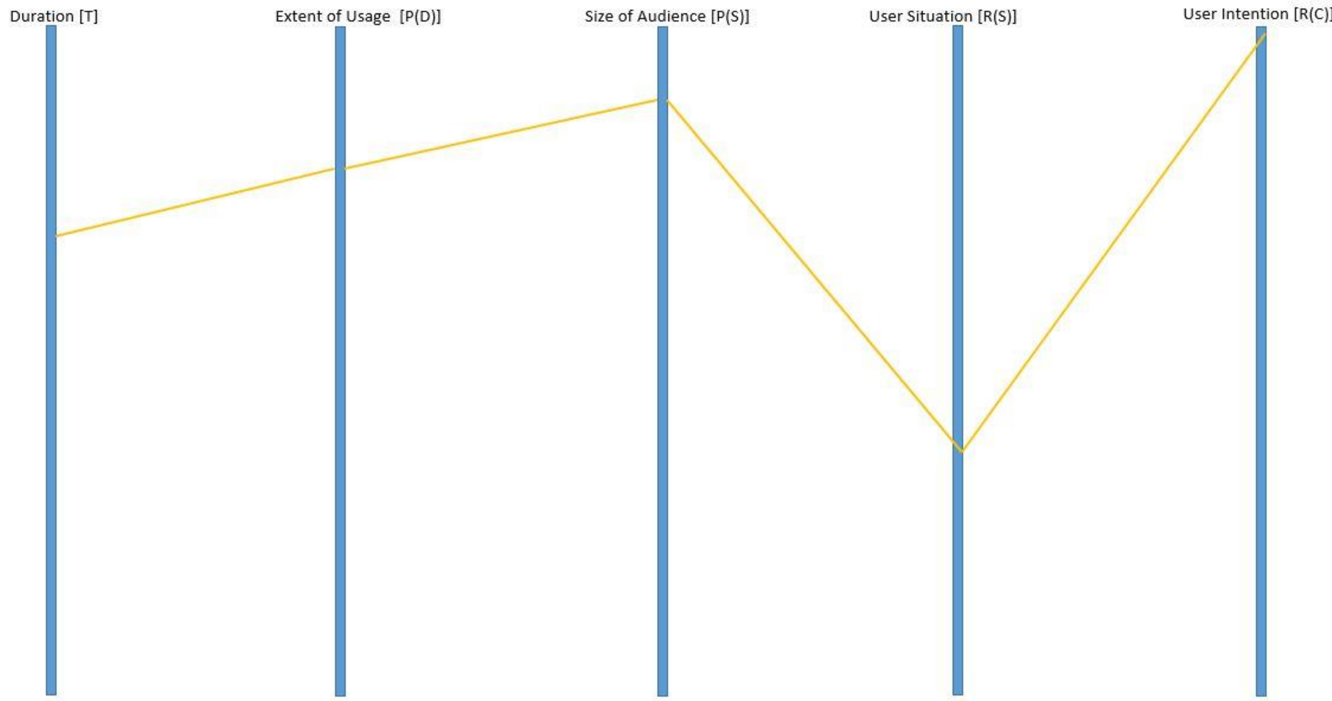


**Figure 35 Multidimensional description of the Job Recommender system**

The five dimensional representation of the recommender system, described in the previous section is now provided. The duration dimension is described as the period of time for which the job data and the resume were kept in the system and duration of chunk of historic data being used for generating recommendation. It is evident from the papers that this factor is on the higher side.

The next factor to consider is the extent of usage of user data by the recommender system. Since, large extent of user's personal data is available to the system in form of resume and user's actions (like, favorite, apply etc.) were being recorded by the system, the extent of data usage is supposed to be at high levels.

The size of audience in this scenario is also a big factor on the higher side. It can be considered to be higher than the valuation/utilization of the two previously discussed dimension because the data is available to many organizations and users that are accessing the system for their job search and getting recommendations from the system.

Since, most of the user data that is obtained, stored and utilized by the system is in static form involving personal information of both the job applicant and the employers, the value of user situation awareness by the recommender system is on the lower side.

Finally, the user intention factor of the system is at a high level in the graph because of the fact that the main objective of the system is to obtain meaningful job recommendation to the user and being aware of the user's intention to display better results.

The future work can be focused on either reducing the number of dimensions from five, in order to better represent the system by finding relationship or equations between the existing dimensions. More dimensions can be added into the system by figuring out more parameter for the evaluation of the recommender system across multiple platforms.

Quantitative analysis can be performed over the recommender systems across multiple dimension in order to find the optimal values for each of the existing dimension that the recommender system must satisfy. These optimal value can be served as the threshold values for these dimensions and the recommender systems can be characterized based on these threshold values. The characterization of the recommender system could lead to a standard for evaluation for these systems contrary to the existent metric i.e. the accuracy and predictability.

This metric of recommender system evaluation will be more efficient and fair because sample used for predicting the accuracy can be colluded or the testing and validation set for determine the accuracy of the recommender system might not be applicable in the real world applications.

# Bibliography


[1] Girardi, Rosario, and Leandro Balby Marinho. "A model of Web recommender systems based on usage mining and collaborative filtering." *Requirements Engineering* 12.1 (2007): 23-40.

[2] Rasmussen, Curtis, and Rozita Dara. "Recommender Systems for Privacy Management: A Framework." *High-Assurance Systems Engineering (HASE), 2014 IEEE 15th International Symposium on*. IEEE, 2014.

[3] Sankar, C. Prem, R. Vidyaraj, and K. Satheesh Kumar. "Trust Based Stock Recommendation System–A Social Network Analysis Approach." *Procedia Computer Science* 46 (2015): 299-305.

[4] Zhao, Vicky Na, Melody Moh, and Teng-Sheng Moh. "Contextual-Aware Hybrid Recommender System for Mixed Cold-Start Problems in Privacy Protection." *Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS), 2016 IEEE 2nd International Conference on*. IEEE, 2016.

[5] Elmisery, Ahmed M., Seungmin Rho, and Dmitri Botvich. "Collaborative privacy framework for minimizing privacy risks in an IPTV social recommender service." *Multimedia Tools and Applications* 75.22 (2016): 14927-14957.

[6] Ma, Xindi, et al. "APPLET: a privacy-preserving framework for location-aware recommender system." *Science China Information Sciences* 60.9 (2017): 092101.

[7] Bouneffouf, Djallel. *DRARS, A Dynamic Risk-Aware Recommender System*. Diss. Institut National des Télécommunications, 2013.

[8] Wang, Zhibo, et al. "Friendbook: a semantic-based friend recommendation system for social networks." *IEEE Transactions on Mobile Computing* 14.3 (2015): 538-551.

[9] Guo, Shiqiang, Folami Alamudun, and Tracy Hammond. "RésuMatcher: A personalized résumé-job matching system." *Expert Systems with Applications* 60 (2016): 169-182.

[10] Lu, Yao, Sandy El Helou, and Denis Gillet. "A recommender system for job seeking and recruiting website." *Proceedings of the 22nd International Conference on World Wide Web*. ACM, 2013.

[11] McSherry, Frank, and Ilya Mironov. "Differentially private recommender systems: building privacy into the net." *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009.

[12] Shokri, Reza, et al. "Preserving privacy in collaborative filtering through distributed aggregation of offline profiles." *Proceedings of the third ACM conference on Recommender systems*. ACM, 2009.

[13] Shang, Shang, et al. "Beyond personalization and anonymity: Towards a group-based recommender system." *Proceedings of the 29th Annual ACM Symposium on Applied Computing*. ACM, 2014.

[14] Zhang, Bo, Na Wang, and Hongxia Jin. "Privacy concerns in online recommender systems: influences of control and user data input." *Symposium on Usable Privacy and Security (SOUPS)*. 2014.

[15] Zimmermann, Thomas, and Christian Bird. "Collaborative software development in ten years: Diversity, tools, and remix culture." *Proceedings of the Workshop on The Future of Collaborative Software Development*. 2012.

[16] El Helou, Sandy, et al. "The 3A contextual ranking system: simultaneously recommending actors, assets, and group activities." *Proceedings of the third ACM conference on Recommender systems*. ACM, 2009.

[17] Adomavicius, Gediminas, and Alexander Tuzhilin. "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions." *IEEE transactions on knowledge and data engineering* 17.6 (2005): 734-749.

[18] Beel, Joeran, et al. "Research-paper recommender systems: a literature survey." *International Journal on Digital Libraries* 17.4 (2016): 305-338.

[19] Jeckmans, Arjan JP, et al. "Privacy in recommender systems." *Social media retrieval*. Springer London, 2013. 263-281.

[20] Jafarkarimi, Hosein, Alex Tze Hiang Sim, and Robab Saadatdoost. "A naive recommendation model for large databases." *International Journal of Information and Education Technology* 2.3 (2012): 216.

[21] Dey, Anind K. "Understanding and using context." *Personal and ubiquitous computing* 5.1 (2001): 4-7.

[22] Melville, Prem, and Vikas Sindhwani. "Recommender systems." *Encyclopedia of machine learning*. Springer US, 2011. 829-838.

[23] Mooney, Raymond J., and Loriene Roy. "Content-based book recommending using learning for text categorization." *Proceedings of the fifth ACM conference on Digital libraries*. ACM, 2000.

[24] Grudin, Jonathan. "Partitioning digital worlds: focal and peripheral awareness in multiple monitor use." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2001.

[25] Kang, Jerry. "Information privacy in cyberspace transactions." *Stanford Law Review* (1998): 1193-1294.

[26] Palen, Leysia, and Paul Dourish. "Unpacking privacy for a networked world." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2003.

[27] Palen, Leysia, and Paul Dourish. "Unpacking privacy for a networked world." *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2003.

[28] Gross, Ralph, and Alessandro Acquisti. "Information revelation and privacy in online social networks." *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*. ACM, 2005.

[29] Tufekci, Zeynep. "Can you see me now? Audience and disclosure regulation in online social network sites." *Bulletin of Science, Technology & Society* 28.1 (2008): 20-36.

[30] Shyong, K., Dan Frankowski, and John Riedl. "Do you trust your recommendations? An exploration of security and privacy issues in recommender systems." *Emerging Trends in Information and Communication Security*. Springer Berlin Heidelberg, 2006. 14-29.

[31] Konstas, Ioannis, Vassilios Stathopoulos, and Joemon M. Jose. "On social networks and collaborative recommendation." *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2009.

[32] Ramakrishnan, Naren, et al. "Privacy risks in recommender systems." *IEEE Internet Computing* 5.6 (2001): 54.

[33] Rosenblum, David. "What anyone can know: The privacy risks of social networking sites." *IEEE Security & Privacy* 5.3 (2007).

[34] Cissée, Richard, and Sahin Albayrak. "An agent-based approach for privacy-preserving recommender systems." *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*. ACM, 2007.

[35] Polat, Huseyin, and Wenliang Du. "SVD-based collaborative filtering with privacy." *Proceedings of the 2005 ACM symposium on Applied computing*. ACM, 2005.

[36] Polat, Huseyin, and Wenliang Du. "Privacy-preserving top-N recommendation on distributed data." *Journal of the American Society for Information Science and Technology* 59.7 (2008): 1093-1108.

[37] Berkovsky, Shlomo, et al. "Enhancing privacy and preserving accuracy of a distributed collaborative filtering." *Proceedings of the 2007 ACM conference on Recommender systems*. ACM, 2007.

[38] Shokri, Reza, et al. "Preserving privacy in collaborative filtering through distributed aggregation of offline profiles." *Proceedings of the third ACM conference on Recommender systems*. ACM, 2009.

[39] Bugliesi, Michele, et al., eds. *Automata, Languages and Programming: 33rd International Colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006, Proceedings*. Vol. 4051. Springer, 2006.

[40] McSherry, Frank, and Ilya Mironov. "Differentially private recommender systems: building privacy into the net." *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009.

[41] Lampe, Cliff, Nicole B. Ellison, and Charles Steinfield. "Changes in use and perception of Facebook." *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. ACM, 2008.

[42] Knijnenburg, Bart P., and Alfred Kobsa. "Making decisions about privacy: information disclosure in context-aware recommender systems." *ACM Transactions on Interactive Intelligent Systems (TiiS)* 3.3 (2013): 20.

[43] Filar, Jerzy A., Dmitry Krass, and Kirsten W. Ross. "Percentile performance criteria for limiting average Markov decision processes." *IEEE Transactions on Automatic Control* 40.1 (1995): 2-10.

[44] Marcus, Steven I., et al. "Risk sensitive Markov decision processes." *Systems and control in the twenty-first century*. Birkhäuser Boston, 1997. 263-279.

[45] Geibel, Peter, and Fritz Wysotzki. "Risk-sensitive reinforcement learning applied to control under constraints." *J. Artif. Intell. Res.(JAIR)* 24 (2005): 81-108.

[46] Yu, Xiang, et al. "An autonomous robust fault tolerant control system." *Information Acquisition, 2006 IEEE International Conference on*. IEEE, 2006.

[47] Hans, Alexander, et al. "Safe exploration for reinforcement learning." *ESANN*. 2008.

[48] Castro, Dotan D., Aviv Tamar, and Shie Mannor. "Policy gradients with variance related risk criteria." *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*. 2012.

[49] Gao, Sheng, et al. "TrPF: A trajectory privacy-preserving framework for participatory sensing." *IEEE Transactions on Information Forensics and Security* 8.6 (2013): 874-887.

[50] Niu, Ben, et al. "Enhancing privacy through caching in location-based services." *Computer Communications (INFOCOM), 2015 IEEE Conference on*. IEEE, 2015.

[51] Cicek, A. Ercument, Mehmet Ercan Nergiz, and Yucel Saygin. "Ensuring location diversity in privacy-preserving spatio-temporal data publishing." *The VLDB Journal* 23.4 (2014): 609-625.

[52] Andrés, Miguel E., et al. "Geo-indistinguishability: Differential privacy for location-based systems." *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*. ACM, 2013.

[53] Xiao, Yonghui, and Li Xiong. "Protecting locations with differential privacy under temporal correlations." *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2015.

[54] To, Hien, Gabriel Ghinita, and Cyrus Shahabi. "A framework for protecting worker location privacy in spatial crowdsourcing." *Proceedings of the VLDB Endowment* 7.10 (2014): 919-930.

[55] Shao, Jun, Rongxing Lu, and Xiaodong Lin. "Fine: A fine-grained privacy-preserving location-based service framework for mobile devices." *INFOCOM, 2014 Proceedings IEEE*. IEEE, 2014.

[56] Popa, Raluca Ada, et al. "CryptDB: processing queries on an encrypted database." *Communications of the ACM* 55.9 (2012): 103-111.

[57] Calandrino, Joseph A., et al. "" You Might Also Like:" Privacy Risks of Collaborative Filtering." *Security and Privacy (SP), 2011 IEEE Symposium on*. IEEE, 2011.

[58] Bhagat, Smriti, et al. "Recommending with an agenda: Active learning of private attributes using matrix factorization." *Proceedings of the 8th ACM Conference on Recommender systems*. ACM, 2014.

[59] Staff, C. A. C. M. "Recommendation algorithms, online privacy, and more." *Communications of the ACM* 52.5 (2009): 10-11.

[60] Celdrán, Alberto Huertas, et al. "PRECISE: Privacy-aware recommender based on context information for cloud service environments." *IEEE Communications Magazine* 52.8 (2014): 90-96.

[61] Zhu, Jieming, et al. "A privacy-preserving qos prediction framework for web service recommendation." *Web Services (ICWS), 2015 IEEE International Conference on*. IEEE, 2015.

[62] Jorgensen, Zach, and Ting Yu. "A Privacy-Preserving Framework for Personalized, Social Recommendations." *EDBT*. 2014.

[63] Guerraoui, Rachid, et al. "D 2 P: distance-based differential privacy in recommenders." *Proceedings of the VLDB Endowment* 8.8 (2015): 862-873.

[64] Shen, Yilin, and Hongxia Jin. "Privacy-preserving personalized recommendation: An instance-based approach via differential privacy." *Data Mining (ICDM), 2014 IEEE International Conference on*. IEEE, 2014.

[65] Liu, Bisheng, and Urs Hengartner. "pTwitterRec: a privacy-preserving personalized tweet recommendation framework." *Proceedings of the 9th ACM symposium on Information, computer and communications security*. ACM, 2014.

[66] Samanthula, Bharath K., et al. "Privacy-Preserving and Efficient Friend Recommendation in Online Social Networks." *Trans. Data Privacy* 8.2 (2015): 141-171.

[67] Gong, Yanmin, Yuanxiong Guo, and Yuguang Fang. "A privacy-preserving task recommendation framework for mobile crowdsourcing." *Global Communications Conference (GLOBECOM), 2014 IEEE*. IEEE, 2014.

[68] Hoens, T. Ryan, et al. "Reliable medical recommendation systems with patient privacy." *ACM Transactions on Intelligent Systems and Technology (TIST)* 4.4 (2013): 67.

[69] Guo, Linke, Chi Zhang, and Yuguang Fang. "A trust-based privacy-preserving friend recommendation scheme for online social networks." *IEEE Transactions on Dependable and Secure Computing* 12.4 (2015): 413-427.

[70] Xin, Yu, and Tommi Jaakkola. "Controlling privacy in recommender systems." *Advances in Neural Information Processing Systems*. 2014.

[71] Tinghuai, M. A., et al. "Social network and tag sources based augmenting collaborative recommender system." *IEICE transactions on Information and Systems* 98.4 (2015): 902-910.

[72] Aïmeur, Esma, et al. "Alambic: a privacy-preserving recommender system for electronic commerce." *International Journal of Information Security* 7.5 (2008): 307-334.

[73] Zhu, Hengshu, et al. "Mobile app recommendations with security and privacy awareness." *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2014.

[74] N. Lao, “Efficient Random Walk Inference with Knowledge Bases,” PhD Thesis. The Carnegie Mellon University, 2012.

[75] N. Lao and W. W. Cohen, “Personalized Reading Recommendations for Saccharomyces Genome Database,” *Unpublished Paper, http://www.cs.cmu.edu/ nlao/publication/2012/2012.dils.pdf*, pp. 1–15, 2012.

[76] N. Lao and W. W. Cohen, “Personalized Reading Recommendations for Saccharomyces Genome Database,” *Unpublished Poster, http://www.cs.cmu.edu/ nlao/publication/2012/2012.dils.poster.portrat.pdf*, 2012.

[77] N. Lao and W. W. Cohen, “Contextual Recommendation with Path Constrained Random Walks,” *Unpublished, http://www.cs.cmu.edu/ nlao/doc/2011.cikm.pdf*, pp. 1–9, 2011.

[78] P. Lakkaraju, S. Gauch, and M. Speretta, “Document similarity based on concept tree distance,” in *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*, 2008, pp. 127–132.

[79] N. Lao and W. W. Cohen, “Relational retrieval using a combination of path-constrained random walks,” *Machine learning*, vol. 81, no. 1, pp. 53–67, 2010.

[80] K. D. B. S. Lawrence, “A System For Automatic Personalized Tracking of Scientific Literature on the Web,” in *Proceedings of the fourth ACM conference on Digital libraries*, 1999, pp. 105–113.

[81] S. R. Lawrence, K. D. Bollacker, and C. L. Giles, “Autonomous citation indexing and literature browsing using citation context,” U.S. Patent US 6,738,780 B2Summer-2004.

[82] S. R. Lawrence, C. L. Giles, and K. D. Bollacker, “Autonomous citation indexing and literature browsing using citation context,” U.S. Patent US 6,289,342 B1Nov-2001.

[83] H. Li, I. Councill, W.-C. Lee, and C. L. Giles, “CiteSeerx: an architecture and web service design for an academic document search engine,” in *Proceedings of the 15th international conference on World Wide Web*, 2006, pp. 883–884.

[84] Y. Liang, Q. Li, and T. Qian, “Finding relevant papers based on citation relations,” in *Proceedings of the 12th international conference on Web-age information management*, 2011, pp. 403–414.

[85] J. Lin and W. J. Wilbur, “PubMed Related Articles: a Probabilistic Topic-based Model for Content Similarity,” *BMC Bioinformatics*, vol. 8, no. 1, pp. 423–436, 2007.

[86] Y. Lu, J. He, D. Shan, and H. Yan, “Recommending citations with translation model,” in *Proceedings of the 20th ACM international conference on Information and knowledge management*, 2011, pp. 2017–2020.

[87] S. M. McNee, N. Kapoor, and J. A. Konstan, “Don’t look stupid: avoiding pitfalls when recommending research papers,” in *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, 2006, pp. 171–180.

[88] S. E. Middleton, H. Alani, and D. C. De Roure, “Exploiting synergy between ontologies and recommender systems,” in *Proceedings of the Semantic Web Workshop*, 2002, pp. 1–10.

[89] S. E. Middleton, D. De Roure, and N. R. Shadbolt, “Ontology-based recommender systems,” in *Handbook on Ontologies*, Springer, 2009, pp. 779–796.

[90] S. E. Middleton, D. C. De Roure, and N. R. Shadbolt, “Foxtrot recommender system: User profiling, ontologies and the World Wide Web,” in *Proceedings of the WWW Conference*, 2002, pp. 1–3.

[91] S. E. Middleton, D. C. De Roure, and N. R. Shadbolt, “Capturing knowledge of user preferences: ontologies in recommender systems,” in *Proceedings of the 1st international conference on Knowledge capture*, 2001, pp. 100–107.

[92] M. Mönnich and M. Spiering, “Adding value to the library catalog by implementing a recommendation system,” *D-Lib Magazine*, vol. 14, no. 5, pp. 4–11, 2008.

[93] S. M. McNee, I. Albert, D. Cosley, P. Gopalkrishnan, S. K. Lam, A. M. Rashid, J. A. Konstan, and J. Riedl, “On the Recommending of Citations for Research Papers,” in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, 2002, pp. 116–125.

[94] S. E. Middleton, N. R. Shadbolt, and D. C. De Roure, “Ontological user profiling in recommender systems,” *ACM Transactions on Information Systems (TOIS)*, vol. 22, no. 1, pp. 54–88, 2004.

[95] M. Monnich and M. Spiering, “Einsatz von BibTip als Recommendersystem im Bibliothekskatalog,” *Bibliotheksdienst*, vol. 42, no. 1, pp. 54–54, 2008.

[96] A. Naak, “Papyres: un système de gestion et de recommandation d’articles de recherche,” Master Thesis. Université de Montréal, 2009.

[97] A. W. Neumann, “Recommender Systems for Information Providers,” Springer, 2009, pp. 91–119.

[98] A. Naak, H. Hage, and E. Aimeur, “A multi-criteria collaborative filtering approach for research paper recommendation in papyres,” in *Proceedings of the 4th International Conference MCETECH*, 2009, pp. 25–39.

[99] A. Naak, H. Hage, and E. Aimeur, “Papyres: A Research Paper Management System,” in *Proceedings of the 10th E-Commerce Technology Conference on Enterprise Computing, E-Commerce and E-Services*, 2008, pp. 201–208.

[100] R. M. Nallapati, A. Ahmed, E. P. Xing, and W. W. Cohen, “Joint latent topic models for text and citations,” in *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2008, pp. 542–550.

[101] C. Nascimento, A. H. Laender, A. S. da Silva, and M. A. Gonçalves, “A source independent framework for research paper recommendation,” in *Proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries*, 2011, pp. 297–306.

[102] T. Ozono, S. Goto, N. Fujimaki, and T. Shintani, “P2p based knowledge source discovery on research support system papits,” in *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 1*, 2002, pp. 49–50.

[103] T. Ozono and T. Shintani, “P2P based Information Retrieval on Research Support System Papits,” in *Proceedngs of the IASTED International Conference on Artificial and Computational Intelligence*, 2002, pp. 136–141.

[104] T. Ozono and T. Shintani, “Paper classification for recommendation on research support system papits,” *IJCSNS International Journal of Computer Science and Network Security*, vol. 6, pp. 17–23, 2006.

[105] T. Ozono, T. Shintani, T. Ito, and T. Hasegawa, “A feature selection for text categorization on research support system Papits,” in *Proceedings of the 8th Pacific Rim International Conference on Artificial Intelligence*, 2004, pp. 524–533.

[106] D. M. Pennock, E. Horvitz, S. Lawrence, and C. L. Giles, “Collaborative filtering by personality diagnosis: A hybrid memory-and model-based approach,” in *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*, 2000, pp. 473–480.

[107] Y. Petinot, C. L. Giles, V. Bhatnagar, P. B. Teregowda, and H. Han, “Enabling interoperability for autonomous digital libraries: an API to citeseer services,” in *Digital Libraries, 2004. Proceedings of the 2004 Joint ACM/IEEE Conference on*, 2004, pp. 372–373.

[108] Y. Petinot, C. L. Giles, V. Bhatnagar, P. B. Teregowda, H. Han, and I. Councill, “A service-oriented architecture for digital libraries,” in *Proceedings of the 2nd international conference on Service oriented computing*, 2004, pp. 263–268.

[109] S. Pohl, “Using Access Data for Paper Recommendations on ArXiv. org,” Master Thesis. Technical University of Darmstadt, 2007.

[110] S. Pohl, F. Radlinski, and T. Joachims, “Recommending related papers based on digital library access records,” in *Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries*, 2007, pp. 417–418.

[111] T. Researchgate, “Researchgate Recommender,” *http://www.researchgate.net/directory/publications/*, 2011.

[112] L. Rokach, P. Mitra, S. Kataria, W. Huang, and L. Giles, “A Supervised Learning Method for Context-Aware Citation Recommendation in a Large Corpus,” in *Proceedings of the Large-Scale and Distributed Systems for Information Retrieval Workshop (LSDS-IR)*, 2013, pp. 17–22.

[113] Sarkanto, “About the Sarkanto Recommender Demo,” *http://lab.cisti-icist.nrc-cnrc.gc.ca/Sarkanto/about.jsp*. 2013.

[114] T. Strohman, W. B. Croft, and D. Jensen, “Recommending citations for academic papers,” in *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, 2007, pp. 705–706.

[115] K. Sugiyama and M.-Y. Kan, “Scholarly paper recommendation via user’s recent research interests,” in *Proceedings of the 10th ACM/IEEE Annual Joint Conference on Digital Libraries (JCDL)*, 2010, pp. 29–38.

[116] D. Thomas, A. Greenberg, and P. Calarco, “Scholarly Usage Based Recommendations: Evaluating bX for a Consortium,” *Presentation, http://igelu.org/wp-content/uploads/2011/09/bx_igelu_presentation_updated_september-13.pdf*. 2011.

[117] R. Torres, S. M. McNee, M. Abel, J. A. Konstan, and J. Riedl, “Enhancing digital libraries with TechLens+,” in *Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*, 2004, pp. 228–236.

[118] K. Uchiyama, H. Nanba, A. Aizawa, and T. Sagara, “OSUSUME: cross-lingual recommender system for research papers,” in *Proceedings of the 2011 Workshop on Context-awareness in Retrieval and Recommendation*, 2011, pp. 39–42.

[119] A. Vellino, “A comparison between usage-based and citation-based methods for recommending scholarly research articles,” in *Proceedings of the American Society for Information Science and Technology*, 2010, vol. 47, no. 1, pp. 1–2.

[120] A. Vellino and D. Zeber, “A hybrid, multi-dimensional recommender for journal articles in a scientific digital library,” in *Proceedings of the 2007 IEEE/WIC/ACM International Conference on Web Intelligence*, 2007, pp. 111–114.

[121] Y. Wang, E. Zhai, J. Hu, and Z. Chen, “Claper: Recommend classical papers to beginners,” in *Seventh International Conference on Fuzzy Systems and Knowledge Discovery*, 2010, vol. 6, pp. 2777–2781.

[122] S. Watanabe, T. Ito, T. Ozono, and T. Shintani, “A paper recommendation mechanism for the research support system papits,” in *Proceedings of the International Workshop on Data Engineering Issues in E-Commerce*, 2005, pp. 71–80.

[123] A. Woodruff, R. Gossweiler, J. Pitkow, E. H. Chi, and S. K. Card, “Enhancing a digital book with a reading recommender,” in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2000, pp. 153–160.

[124] C. Yang, B. Wei, J. Wu, Y. Zhang, and L. Zhang, “CARES: a ranking-oriented CADAL recommender system,” in *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries*, 2009, pp. 203–212.

[125] F. Zarrinkalam and M. Kahani, “SemCiR - A citation recommendation system based on a novel semantic distance measure,” *Program: electronic library and information systems*, vol. 47, no. 1, pp. 92–112, 2013.

[126] F. Zarrinkalam and M. Kahani, “A New Metric for Measuring Relatedness of Scientific Papers Based on Non-Textual Features,” *Intelligent Information Management*, vol. 4, no. 4, pp. 99–107, 2012.

[127] D. Zhou, S. Zhu, K. Yu, X. Song, B. L. Tseng, H. Zha, and C. L. Giles, “Learning multiple graphs for document recommendations,” in *Proceedings of the 17th international conference on World Wide Web*, 2008, pp. 141–150.

[128] Adomavicius, Gediminas, and Alexander Tuzhilin. "Context-aware recommender systems." *Recommender systems handbook*. Springer US, 2015. 191-226.

[129] Zhan, Justin, et al. "Privacy-preserving collaborative recommender systems." *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 40.4 (2010): 472-476.

[130] Zhou, Tao, et al. "Solving the apparent diversity-accuracy dilemma of recommender systems." *Proceedings of the National Academy of Sciences* 107.10 (2010): 4511-4515.