INFLUENCES OF SERENDIPITY ON CONSUMER MEDICAL INFORMATION PERSONALIZATION

Xiangyu Fan

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the Doctor of Philosophy degree in the School of Information and Library Science.

Chapel Hill
2015

Approved by:

Javed Mostafa

Bradley M. Hemminger

Robert Capra

Jane Greenberg

Luz Quiroga

ABSTRACT

XIANGYU FAN: Influences of Serendipity on Consumer Medical Information Personalization
(Under the direction of Javed Mostafa)

Serendipity is an important concept in the field of information science. It has the potential of enhancing information retrieval by promoting unexpected discovery. Serendipitous recommendation has been incorporated into the design of personalized systems to minimize blind spots in information delivery. Little evidence has been found to identify how serendipity influences personalization of consumer medical information delivery. This dissertation attempts to examine what roles serendipity plays in filtering consumer medical information and to understand how to incorporate serendipity in an effective manner. In addition, the study seeks to clarify users' attitudes associated with unexpected discovery of medical content during filtering as well as users' interest changes during this process.

To empirically analyze the influence of serendipity, a medical news filtering system named MedSDFilter was developed. The system personalizes the delivery of news articles based on users' interest profiles. In MedSDFilter, serendipitous recommendation was integrated into personalized filtering using three serendipity models, namely, a randomness-based, a knowledge-based, and an adaptive knowledge-based model. Using Medical News Today[1] site as information source, the three different system modalities were compared based

[1] Medical News Today (http://medicalnewstoday.com) is a market leader for medical news.

on a series of experimental sessions conducted with users. Thirty staff members were recruited to read and rate medical news delivered by one of the three system modalities.

The results of the user study indicated that serendipity has an important role in medical news content delivery. Regarding the question of how to incorporate serendipity, it was demonstrated that physicians' knowledge of medical topic association effectively enhanced serendipitous recommendation. In addition, the results suggested that the performance of recommendations were further improved after combing physicians' knowledge with a learning algorithm for more refined filtering. This study provided evidence of user-satisfaction associated with serendipitous recommendation. Finally, the study revealed several types of individual differences in seeking consumer medical information.

The results of this study yielded new insights and pointed out new means for avoiding potential drawbacks related to over-personalization in information delivery. This study enhanced our understanding of users' behavior regarding the consumption of medical information and generated new guidelines that can be used in developing consumer-centric information systems in the medical area.

This dissertation is dedicated to my parents, Kailiang Fan and Yangying Li.

ACKNOWLEDGEMENTS

It is highly challenging for me to pursue a Ph.D. degree in a field of which I knew little before. Without the support of family, mentors, and friends, I could not have achieved this goal.

My advisor Dr. Javed Mostafa introduced me to information science after I got MS degree in a different major. Under his directions, I had a chance to access a fantastic field - consumer medical information personalization. Dr. Mostafa is an excellent mentor to me. Though he was busy with multiple projects, he still spent a great deal of time discussing with me the details of the user study design. When I encountered various problems in my research work and personal life, he always encouraged me to keep an optimistic attitude. From him, I have learnt not only how to become a good researcher in my chosen academic field but also how to link the knowledge I have learnt in my research to business development.

My committee members have provided assistance during my dissertation study. Dr. Brad Hemminger has constantly provided me valuable advice. He has encouraged me to refine my experiments so that my research results would help resolve fundamental human-computer interaction issues. Dr. Robert Capra has consistently helped me improve the design of experiments to capture users' actual attitudes. He provided me detailed advice on how to improve the experiments so that the data I obtained can be closely linked to my research questions. Dr. Jane Greenberg has always encouraged me when I felt discouraged with slow progress of my research. She recommended good papers for me to study. After she moved to another University, she was still able to find time to advise me and serve on my committee. Dr. Luz Quiroga advised me to read some important papers in the information filtering field and also gave me good advice on the development of filtering system.

I would like to thank my parents. When I was pursing this degree, they always encouraged me to move forward. During the times when I felt exhausted with my research work and wanted to give up, I just went to see them and enjoyed delicious food they cooked for me. Despite staying far away from them, my parents tried their best to provide me with emotional and financial support. Without their continued support and love, I could not have accomplished this degree.

I also want to thank some of my colleagues, such as Fei Yu, my lab member, who helped me recruit subjects and conduct pilot study; Vincent Carrasco and Carlton R Moore, my lab members, who judged medical topic associations for my study; Dongqiuye Pu, my previous lab member, who helped me with literature collection; Xi Niu, my previous Ph.D. fellow who gave me good advice for my literature review and helped with editing my dissertation paper; Cassidy Sugimoto and Yan Zhang, my previous Ph.D. fellows, from whom I have learnt about user study designs from previous collaboration; and Weimao Ke, my previous Ph.D. fellow who taught me a great deal when I started to learn information technology. Without their kind help or collaboration, it would become much difficult for me to accomplish this degree.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.INTRODUCTION

In recent years, significant emphasis has been placed upon building personalized information retrieval (PIR) systems to help address the problem of information overload (Dumais et al., 2003; Radlinski & Dumais, 2006; Ma et al., 2007; Teevan et al., 2008). PIR-based systems offer users the advantage of retrieving better quality results in reduced time. However, they may not handle all types of information needs (e.g., browsing news articles in a casual environment) well, and information filtering (IF) systems work as an alternative option for personalized information delivery. Personalized information filtering (PIF) systems remove redundant or unwanted content from information stream and present relevant content to users based on their individual profiles. The presented content after filtering is "wanted" and expected to meet users' information needs. In certain situations, natural human information seeking may involve "a relatively unfocused sense of inquiry where the initial goal is not to find some particular answer or to fill some sort of reasonably anticipated informational gap" (Miksa, 1992). Serendipity in such situations helps inquirers learn their real information needs by accessing a wide range of information, and hence, guides the process of seeking relevant information. The concept of serendipity, or "the faculty of making happy and unexpected discoveries by accident" (Oxford English Dictionary Online, 2010), has a rich history in sociology and the acquisition of new knowledge (Merton and Barber, 2004). The survey of literature shows that serendipity is an important issue to address in the field of information science (Gup, 1998; Toms, 2000; Foster & Ford, 2003; Bruijn and Spence, 2008; Fan et al., 2012). In this dissertation concerning personalized filtering system, we define serendipity as the occurrence and development of unexpected discovery of information which is relevant to users' interest. This is different as compared to the traditional concept of serendipity which focuses on the exposure to diverse content in casual information seeking situation. In personalized filtering, serendipity is concerned with finding a way to actively assist users' unexpected discovery instead of relying on discoveries simply based on luck.

1.1 Problem Statement

Personalized information filtering involves a process of mapping users' profiles to the characteristics of information items. In filtering settings, capturing users' interest profiles is important in order to deliver relevant information to users, whose interest profiles are often described using vectors of topics. Each value in a vector represents how strongly users are interested in one particular topic. The simple way to acquire user interest profiles is through a dialog or interrogation. However, users may be unaware of or unable to articulate their actual information needs (Taylor, 1968; Belkin, 1980), and this situation tends to occur in the process of medical information seeking. Medical information includes consumer-level information and patient health records (PHR), which have different characteristics and therefore serve different purposes. Unlike PHR that are used for professional diagnosis and treatment, consumer-level medical information helps users improve their understanding of diseases, conditions, wellness issues, and eventually their health conditions. A large amount of consumer-level medical information is generated every day through various data sources, including a number of professional medical websites like MedlinePlus and WebMD, etc. In order to resolve the problem of medical information overload, various personalized filtering systems have been developed. However, it is challenging to acquire users' complete interests on medical topics. There are several barriers. First, some users would not like to reveal their personal preferences on all the medical topics of interest because there is concern that information about personal health conditions might leak, even though users may have interest in particular topics and the filtering systems adopt high security standards. Second, it may be the case that users are not certain about their exact interests in relation to specific medical topics, due to the lack of professional knowledge required to fully understand many medical topics. For instance, users' interests in medical topics are often related to their health conditions, and as a result, users may be incapable of linking their physical conditions to specific medical topics because they lack professional knowledge. Generally, when a situation arises whereby potentially relevant information is ignored or missed due to a gap or inaccuracy in the interest profile, the condition is called a blind spot. It may occur in many contexts and "Filter bubble" can be one of them. Some personalized search engines like Google analyze information using users' personal or local characteristics (e.g. location) and present information based on users' own culture or ideology ("Filter bubble," 2015). This phenomenon may cause blind spots if the system's assumption about user's preference is inaccurate and even wrong. In personalization system, whether the information is relevant or not is often judged in terms of the core topic of the retrieved information and users' interest in the topics. This means that blind spots can be evaluated from the angle of topics. In order to describe blind spots easily in this dissertation, we define them as the topics that are not presented to users in filtering sessions but users are actually interested in. To conclude, user profiles that exclusively rely on interrogation in filtering systems may be incomplete and lead to the occurrence of blind spots in medical information delivery.

To address the possibility of blind spot occurring and also to reduce blind spots, explicit incorporation of serendipity has been proved useful. Serendipitous recommendation has been incorporated into the design of many personalized systems in practice. LyricTime music system (Loeb, 1992) is one of the early information recommendation systems that accommodate serendipitous access to information by augmenting the information. The system occasionally adds randomly picked songs to the user's playlist. Similar techniques are now being used extensively by Amazon, Netflix, and other e-commerce systems. Despite successful applications in some commercial sites, little attention has been paid to developing information filtering systems that facilitate serendipitous discovery in medical domain.

To be more specific regarding serendipitous recommendation, little research has been conducted to investigate how to effectively implement it into filtering settings that focus on consumer-level medical information. This study focuses on the role of serendipity as well as the methods of incorporating serendipity, with its scope shown in Figure 1.1.

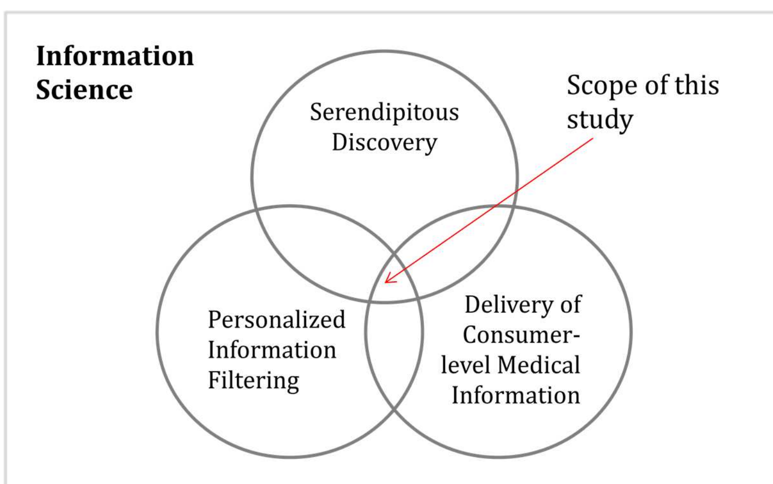


Figure 1.1: The Scope of This Study

Figure 1.1 indicates this study is closely related to three areas in information science: delivery of consumer-level medical information, personalized information filtering and serendipitous discovery. Based on previous research work in these areas, the study attempts to examine important factors associated with serendipitous discovery of consumer-level medical information in personalized filtering environment.

In order to make our problem statement clear, elaboration and explanation is necessary on the initialization phase of filtering systems. When users access a filtering system for the first time, they are often asked to select topics of interest from a list. These selected topics constitute the user's profile used in the whole filtering process by the system. Therefore, the topics are defined as system-profile (SP) topics. In contrast, the topics that are not selected by users are defined as non-system-profile (NSP) topics. In this study, incorporating serendipity is implemented in the system as presenting articles from NSP topics to users.

The traditional methods of serendipity incorporation usually adopt a random process of selecting NSP topics in presentation. This type of method is simple, but it may be not effective and stable due to its non-deterministic properties. Other types of strategies are likely to enhance serendipitous recommendation, with the incorporation of topic associations as an example. This method is based on the assumption that topics of user interests share certain inherent relationships. Based on this idea, it is inferred that NSP topics have a high possibility of being user interests if they are related to SP topics. This type of NSP topics should be considered as a priority in serendipitous recommendation. Compared with traditional methods which rely on "off-topic" serendipity, this method utilizes "near-topic" serendipity. In medical domain, users' interests in topics are often related to their health conditions. Therefore, user-related factors need to be carefully examined through empirical experiments before declaring the effectiveness of incorporating topic associations in enhancing serendipitous discovery in medical area.

In addition, the drawbacks of incorporating and presenting topic associations cannot be neglected. In serendipitous recommendation based on topic associations, NSP topics are presented around the SP topics with which they are associated. It is possible that one NSP topic is related to many SP topics and thus has a priority of presentation in multiple locations. If this NSP topic is not attractive to users, repeated presentation of the topic will make a negative impression on users in terms of information delivery performance. In order to resolve this problem, the method of adopting topic associations in serendipitous recommendation can be optimized by effectively integrating users' preferences for potentially relevant NSP topics. Based on users' interests, the frequency of presentation for each NSP topic is modified. For NSP topics presented to users over multiple times, the improved methods will exclude these topics in the subsequent sessions if users have low or no interest in old (or current) sessions. The methods also make sure that these NSP topics will continue to be presented in the subsequent sessions if users have strong interest in current sessions. By integrating the adaptive learning mechanism, this method incorporates serendipity which is closer to users' interest profile as compared with the method which only utilizes topic association.

Despite the potential usefulness of the strategies discussed above, little experimental data have been obtained to verify the effectiveness of selecting NSP topics based on topic associations and users' preferences. In addition, it is not clear whether users gain value from serendipitous discovery of medical content in filtering settings and if their interests are influenced by this type of discovery.

A previous study conducted by the researcher revealed that serendipity can play a positive role in medical information retrieval and should be given careful consideration in the design of personalized information retrieval systems (Fan et al., 2012). On the basis of findings in the study, this work further examines many important aspects of serendipity in the field of medical content delivery. Different from the previous study, this dissertation research puts an emphasis on how to effectively enhance serendipitous discovery. The incorporation of serendipity has the possibility of reducing blind spots through enhancing unexpected discovery in personalized information filtering. However, it remains an empirical research question as to how to make the serendipitous feature more effective in helping the occurrence of unexpected and useful discoveries in the medical domain.

1.2 Research Questions

The scope of this study is well-defined with a focus on filtering consumer-level medical information. All research questions are addressed within this scope. In this study, user interests are acquired through questionnaires for initializing the filtering environment. Medical news is used as the main information type because it is consumer-level, easy-to-read, and updated frequently. The main purpose of this study is to investigate the effectiveness and efficiency of incorporating serendipity in delivering relevant medical news to users in filtering environment. The main research question is whether and how incorporation of serendipity helps people find unexpected but relevant medical news. This study has four specific goals: 1) to examine whether serendipity helps people find unexpected but relevant medical news; 2) to introduce three serendipity models and determine their efficacy; 3) to investigate users' attitude related to unexpected discoveries; 4) to study the change of users' interests after incorporating serendipity. Based on these goals, four research questions are proposed and introduced in sequence in the following paragraphs. Questions 1 and 2 are proposed from the system's perspective, with the focus on serendipitous recommendation. Questions 3 and 4 are proposed from users' perspective, with the focus on unexpected discovery of relevant content. The four research questions (RQ) are described below:

RQ 1: Does incorporating serendipity help people find unexpected but relevant news?

The first research question focuses on the feasibility of incorporating serendipity to achieve unexpected but useful discoveries. The previous study (Fan et al, 2012) examined information seeking behaviors of users in a personalized information delivery system that ranks medical news articles based on the weight of users' interests. Experimental results demonstrated that serendipity can play a positive role in personalized medical content delivery. Based on the findings in previous work, this dissertation study attempts to examine whether incorporating serendipity influences unexpected discovery of relevant news content in a personalized filtering environment. The unexpected discovery is identified through comprehensively analyzing user interface actions (clicks and rating) and a survey of user interests. It is hoped that by comparing the results of this study with the results obtained in the previous experiments, we will clarify how serendipitous recommendation can affect personalization and effectiveness of medical content delivery.

RQ 2: How do the three well-known serendipity models deliver medical news content differently and in what ways?

This particular research question is the most important one in this study and concerns how to enhance unexpected discovery of relevant content in medical domain. Serendipity is typically incorporated into information filtering by inclusion of randomly selected documents into presentation. Beyond the random method (named as the RA method for serendipity Model 1 here), this study attempts to examine the feasibility of leveraging medical knowledge to strengthen unexpected discovery. In this updated method, the documents that help with unexpected discovery are not randomly chosen. Instead, they are chosen through analyzing the associations of medical topics (called KA methods for serendipity Model 2 here). For each SP topic, all NSP topics are sorted by their association with it. Based on this rank, these NSP topics are sequentially presented around the associated SP topic over many sessions. In the process of implementation, the NSP topic of the highest rank is shown first, followed by that of the second highest rank. Theoretically in both RA and KA methods, users have a chance to view the articles on all the NSP topics if the number of sessions is large enough. Since there exist hundreds of common medical topics in medical domain, it may take a long time to show all NSP topics in sessions. In addition, the interests of users may change over time. In order to improve user experience, it is significant to present NSP topics with a high possibility of being relevant to user's interests at the initial stage of serendipitous recommendation. KA method may perform better than RA in this aspect, based on additional data about medical topics. The associations among medical topics are determined based on the judgment of physicians.

As discussed in the section on research problems, KA methods without modification may cause repeated presentation of some irrelevant NSP topics that are closely associated with many SP topics. This weakens serendipitous recommendation of relevant NSP topics.

Adaptive KA methods (called KAA methods for serendipity Model 3 here) are developed to resolve the problem. In the implementation of KAA methods, all NSP topics are given a limited number of presentations. If users show interest in them in these presentations, the NSP topics continue to be shown to users. If users have low or no interest in them, the NSP topics are moved to the lowest position in the ranked list, implying no continuous presentations on them. More details of KAA method are given in the chapter of research methods. To clarify, KAA methods involve not only professional knowledge from physicians, but also the feedback from users.

The main difference between the three types of methods above is how to predict what NSP topics users would prefer to see. Figure 1.2 shows the distance of predicted NSP topics and users' interest profile.

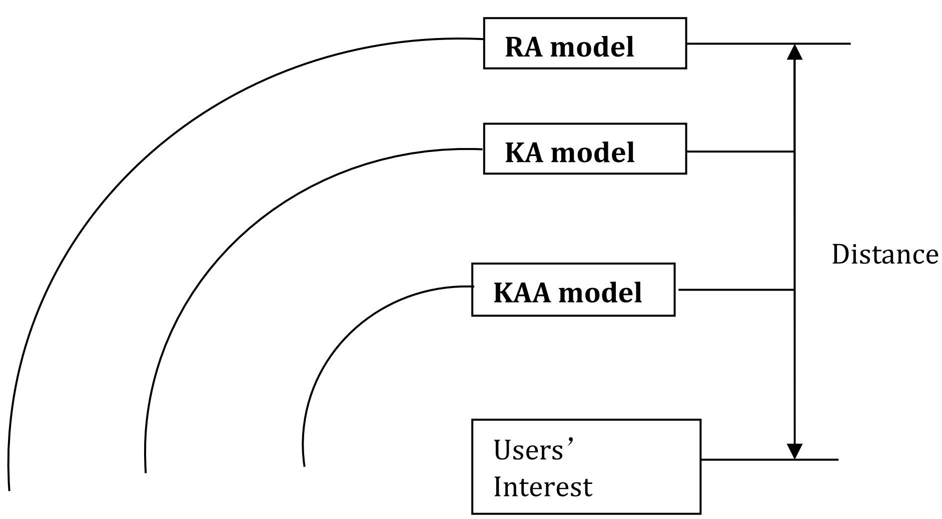


Figure 1.2: Distance between Serendipity Models and Users' Actual Interest Profile

Compared with RA methods, the serendipity topics are selected from NSP topics with shorter distance to users' interest profile when KA methods are implemented. The distance is further shortened when KAA methods are adopted. To clarify the difference, KA/KAA methods are developed to test "closely related" serendipity whereas RA methods are based on "random" serendipity. RQ2 is proposed to examine how "closely related" and "random" serendipity models perform.

The first two research questions cover the main aspects of incorporating serendipity from the perspective of system performance. In order to evaluate the serendipitous discovery comprehensively, another two research questions are proposed to analyze the role of serendipity from the perspective of user reactions.

RQ 3: Do users have a positive response to unexpected discovery of medical news?

This question is concerned with users' attitude related to unexpected discoveries of medical news in filtering settings. It is addressed through analyzing user's interests on clicked articles. This process involves two types of data associated with the articles on NSP topics. The first one is user rating of interest levels on article content over all sessions. The second one is the data from a post-study survey in which users directly describe what they think about the discovery of these unexpected articles and whether they would like to continue to use this feature. These data are comprehensively analyzed to clarify whether people like unexpected discoveries of medical news in filtering settings.

RQ 4: Does unexpected discovery causes a change in users' interest profile?

This question is answered from two aspects. The first one is whether there exists a change in users' interest profile and the second one is whether this change results from unexpected discovery. Since interest profile is a vector of interest strength on medical topics, the interest changes can occur in terms of dimension or strength. This study concentrates on interest dimensions more than strength. By comparing interest profiles acquired from surveys before and after filtering sessions, we can see whether users are interested in extra medical topics after accessing a wider range of medical content. If extra interest dimensions in post-study profiles exist, users will be asked to describe the factors associated with their interest changes. By qualitatively analyzing the comments of users, the connections between unexpected discovery and users' interest changes can be identified. This type of information from questionnaires provides direct evidence for answering RQ 4.

The ultimate objective is to develop reliable medical information filtering settings, which deliver more relevant information to users in an efficient manner, by reducing blind spots associated with personalization. The diversity and complexity of users and their interests make simulation study difficult to implement. As a result, this study adopts empirical research methods based on real users, and further details of the user study are introduced in the methods chapter.

1.3 Significance of the Study

This study contributes to resolving the problems associated with blind spots in medical information delivery in personalized filtering environment. The research examines the existence of blind spots as well as different methods to reduce blind spots in the context of medical domain. Based on literature review, we know serendipitous recommendation may have the potential to reduce blind spots by enhancing unexpected discovery. In this study, this potential is carefully investigated under different comparable strategies of incorporating serendipity. By comparing the unexpected but positive discoveries made by average users, the effects of incorporating serendipity in reducing blind spots in filtering settings can be understood in a more comprehensive manner. The results of this study can provide the designers of personalization systems with implications that can be utilized to reduce negative effects of over-personalization in information filtering.

By generating new evidence, this study seeks to further develop the concepts and theories of serendipity in the information seeking process. Despite the existence of successful applications in some commercial websites, the role of serendipity in IR is still somewhat controversial. Many scholars have noted the “threat” (Foster & Ford, 2003) to serendipity by particular electronic IR systems (Gup, 1997, 1998; Cooper & Prager 2000; Huwe, 1999). Although some scholars have described how serendipity can happen in IR systems, particularly in relation to browsing (Bruijn and Spence, 2008), IR systems have traditionally been created with the intention of enhancing the ability to conduct “known-item” or “known-problem” searches. Information filtering is similar to information retrieval in this respect since IF systems map known users’ profiles to information stream. Experiments conducted in this study help demonstrate the role of incorporating serendipity in personalized information delivery in the medical domain. The clarification of fundamental concepts and theories often results in technological innovations and revolutions. From the perspective of developing new serendipity-related concepts and applications, this study contributes to the field of information science. In addition, the serendipity-related features in assisting medical information seeking have not been carefully investigated. Specifically, there is little literature about serendipitous discovery in the field of consumer-level medical information. This study can make an important contribution to the research about serendipitous discovery of medical content.

This study also contributes to the understanding of users’ behaviors associated with the consumption of medical information. The heterogeneity in interest properties of users often results in variation of ways in which medical information is consumed, increasing the difficulty of designing robust filtering systems. It is important to see how users perform when seeking relevant information in filtering settings and how users' interface actions are related to their interest properties. This study generates some evidence to identify these internal connections. Additionally, users' intrinsic motivation for health information seeking can be identified from their comments, which helps health information provider better understand users' information needs and generate more appropriate information at consumer level.

This study involves different strategies to implement serendipitous recommendation, and the final impact of this work is the methodological contribution. From the results of literature review, the study is considered as an early work of applying topic association judged by human to serendipitous recommendation in the area of consumer-level medical information. The strategy developed in this study informs researchers about the characteristics of topic associations in medical domain, as well as how to utilize topic associations for presenting additional topics of potential relevance to users. In addition to the data of topic associations judged by physicians, user ratings, as one type of supportive data, are integrated for enhancing the performance of serendipitous recommendation in this study. Based on different types of supportive data, KA and KAA methods are formulated. Though these methods are implemented in a filtering system in this study, they involve a possibility of being utilized to enhance serendipitous discovery of medical content in other personalization systems (i.e. personalized medical search engines).

2. LITERATURE REVIEW

The area of focus for this dissertation is consumer medical information delivery, and the research examines how serendipity affects the delivery of consumer medical information in a personalized filtering system. Highlighting the problems associated with information delivery efficiency and coverage, this review aims to survey the literature in information overload and filtering technology, personalization in information delivery, and the role of serendipity in personalized information filtering.

Section 2.1 starts with a brief review of information overload ranging from its causes, effects on information seeking, and potential solutions. Then, Section 2.1 surveys the literature in the broad research area of filtering technology (one solution to information overload). The principles of fundamental theory as well as the approaches and types of information filtering are discussed. Section 2.2 reviews the theory and applications of personalization in information delivery, as well as the advantages and limitations. Based on the problems associated with personalization, the role of serendipity in information delivery is discussed and some recent findings in empirical studies are reviewed. Finally, Section 2.3 introduces the fundamental characteristics, delivery approaches, and information-seeking process of medical information to revisit the research problems proposed specifically for the medical domain.

2.1 Information Overload and Filtering

2.1.1 Information Overload

The rapid advances in computer and communication technologies have led to a generation where a deluge of data and information overload people every day. This overload has been apparent for a long time but has been exacerbated with broad adoption and use of the Internet. The term "information overload" conveys the simple notion of receiving too much information. The term and concept of information overload is a psychological phenomenon where information overload relates to sensory organs with an excess of incoming information. Toffler (1970) noted that individuals can have difficulty in understanding an issue and making decisions when they encounter too much information, and this difficulty can be considered information overload. Hanka (2000) explained the cause of information overload from the perspective of human interaction with knowledge expansion. He claimed that information overload results from the mismatch between the neural capacity of the human brain and the rate of expansion of human knowledge. In medical domain, overload on any given topic comes from the availability of hundreds of journals and guidelines, as well as large amounts of patient data and consumer medical information (e.g., news articles). Fuat et al. (2003) found that most clinicians were not familiar with existing guidelines, and 'guideline fatigue' is a common cause of stress when ascertaining a patient's diagnosis. This result is consistent with an older study conducted by Hibble et al. in 1998.

The effects of information overload on humans have been widely examined in accounting (Schick et al., 1990), management information systems (MIS) (Ackoff, 1967), organization science (Tushman & Nadler, 1978), and consumer research (Keller & Staelin, 1987). The findings are consistent in different disciplines: the performance of an individual in decision-making can be improved with the amount of information he or she receives. If more information is provided beyond a certain point, however, the performance of the individual may decline (Chewning & Harrell, 1990). The authors further explained that the information provided beyond this point causes the occurrence of information overload, which confuses the users and makes them fail to recall prior information they have received (O'Reilly, 1980, Schick et al., 1990). This curve of decision accuracy with information load is shown in Figure 2.1 (Eppler, 2004). It is consistent with the results of many studies on decision-making (Iselin, 1989; Hwang et al., 1999).

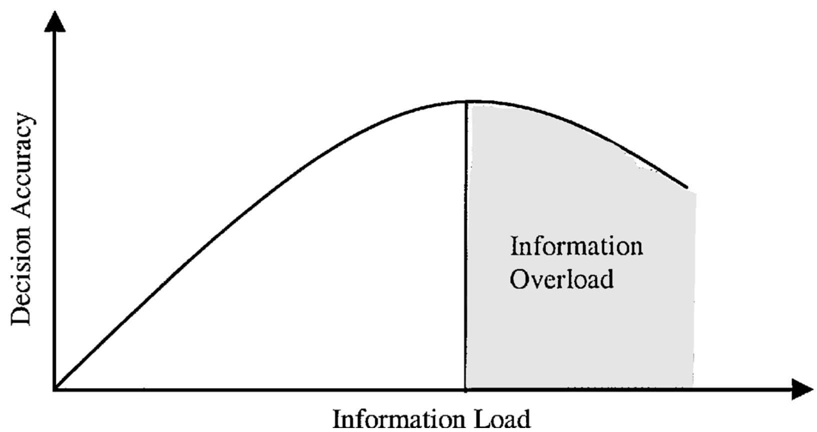


Figure 2.1: Decision Accuracy Change with Information Load (Eppler, 2004)

Hall and Walton (2004) overviewed various contexts of information overload within the health-care system and discussed the impacts overload from a clinic perspective. They note that health clinicians hold a widespread view that the effectiveness of their work is impaired by information overload. When faced with information overload, clinicians tend to have imprecise clinical judgments. The errors of clinicians may be because they have little time to preview and process data (Zeng et al., 2002).

The overview of the literature shows various countermeasures to information overload, ranging from general suggestions to very specific software tools. These approaches vary with the contexts and causes of information overload. Simpson and Prusak (1995) emphasized the quality of information and efficiency of information delivery and argued that the improvement of these two attributes is the main solution to information overload. Meyer (1998) proposed that information overload can be reduced by adopting the methods of visualization, compression, and aggregation. Many researchers are concerned about the information delivery performance and emphasize the importance of intelligent information systems that are capable of handling a large amount of information in an effective manner. They have developed various systems including decision-support systems (Wagholikar et al., 2012), automatic summarizers (Vu et al., 2001), information retrieval systems (Dumais et al., 2003; Radlinski & Dumais, 2006; Ma et al., 2007; Teevan et al., 2008), and information filtering systems (Mostafa et al., 1999; Mostafa & Lam, 2000). Extraction systems glean evidence from non-consumer medical information (e.g., electronic health records) to help disease diagnoses and treatment plans. Like information filtering systems, extraction systems attempt to garner useful information based on some criteria, from incoming texts (Sundheim, 1991). However, extraction systems are focused on extracting the facts hidden in the information rather than judging the relevance to user interests. Information retrieval (or search) is another important technique used for handling consumer medical information. Belkin and Croft (1992) proposed the interesting question of "whether information filtering and information retrieval can be viewed as two sides of the same coin," and systematically discuss the similarities and differences between these two techniques. Although their research compared the models of information filtering and retrieval in terms of the delivery performance for general information, the statements were still applicable to the medical domain.

Information filtering has the potential of enhancing the medical information seeking process in many situations. The general model in Figure 2.2 shows characteristics and application context of information filtering. It differentiates from a search in many ways and meets the information needs of users in a casual environment. First, information filtering is used to compare long-term users' profiles, while a search is used to compare one-time users' queries with text surrogates. As previously discussed, the process of seeking consumer medical information can be long-term because the information needs of people are always related to their health conditions over a period of time. Second, unlike a search, which emphasizes retrieval accuracy, information filtering is more concerned with removing irrelevant data in the information stream (to avoid the occurrence of information overload). This fact implies that it may be acceptable for users to view the information from a broader scope in filtering settings, which will be favorable to the users who have a desire to learn additional information about their disease of interest. Lastly, the information filtering process relies on user interests and is user-specific. This feature is important in medical information seeking because user interests can be, in practice, diverse and dynamic. More importantly, information filtering systems always present a list of medical topics for users to select from when building their interest profiles. Unlike typing queries, this method, which is based on selection, will be easy to use, especially when users have little knowledge in medical domain. To conclude, although different from a search, information filtering that stresses long-term users' interests can play an important role in enhancing the delivery of medical information, and due to its importance, this study selects information filtering as the main method for delivering consumer medical content to users.

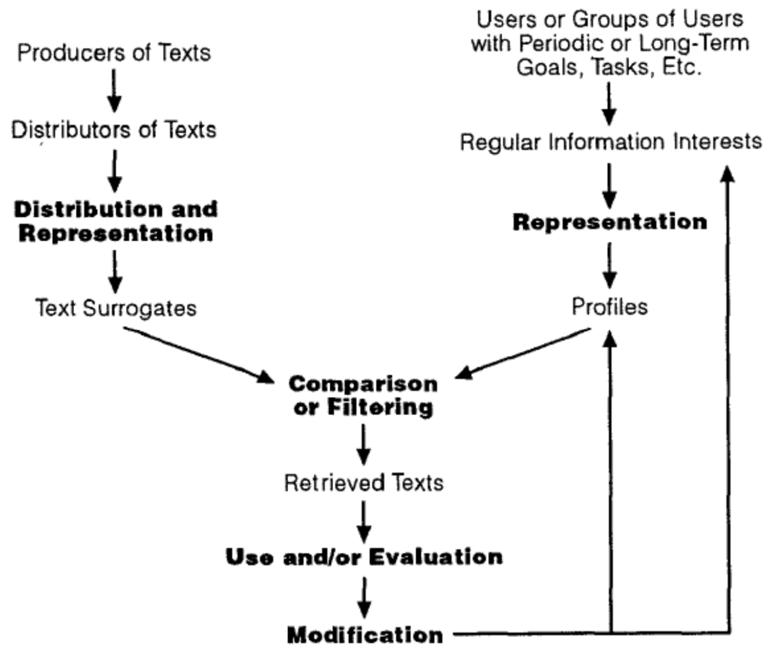


Figure 2.2: A General Model of Information Filtering (Belkin & Croft, 1992)

2.1.2 Types of Information Filtering

Information filtering systems have been used widely as a solution to information overload by increasing the semantic signal-to-noise ratio. Using computerized methods, this system removes redundant information from an information stream before users are exposed to relevant information. A typical information filtering system consists of four components: a data-analyzer component, a filtering component, a user-model component, and a learning component. These components work together to complete the filtering process. Information filtering can involve different initiatives of operation and adopt different presentation methods. In this section, the characteristics of different filtering types are discussed in an attempt to establish one likely to be most suitable for consumer medical information delivery.

1) Initiative of Operation

Initiative of operation distinguishes whether filtering systems are active or passive (Hanani et al., 2001). Active systems search for relevant data and actively present the data to users based on user profiles. Despite the fact that it reduces users' time spent on a search, push technology adopted in these active systems are always controversial. Edmunds and Morris (2000) gave an overview of the negative aspects of push technology and argued that push technology could cause information overload if the information pushed is unwanted information. Active systems make an effort to collect and combine data from different sources. As a result, sometimes the pushed data become spam to users due to the complexity of processing unstructured data. Active systems are often adopted in the areas where the information need is not very clear (e.g., online advertising). Unlike active systems, passive systems do not make an effort to collect the data items for users. Instead, they mainly process the data sent by other agents. Email and news feeds are two typical examples of passive systems. Professional medical websites (MedlinePlus and WebMD et al.) have a large amount of consumer-level medical information and many of them provide RSS feeds. Due to the high requirements of reliability in medical information, collecting content from non-professional sites may involve the risk of providing misleading information to users. It is thus important to get the data from carefully selected agents or sources in medical information delivery. From operation perspective, it appears passive systems may be more suitable for delivering consumer medical content to users than active systems. Passive systems perform as news feed subscription systems, which receive news from one or multiple sources.

2) Presentation Methods

Filtering systems may adopt different methods to present the information. Some filtering systems present part of the data after removing irrelevant information from the information stream, while others show all the data items ranked by relevancy. Each approach has advantages and shortcomings, and the selection of a particular approach is actually a trade-off situation. Less information presented to users can minimize the overload of information but may exclude some important information. Therefore, the selection of presentation method is dependent on whether the risk is affordable. For instance, a study by Oard (1997) indicated that Email users tend to use a filtering method that lists all messages because they are afraid of missing some messages of importance. Even if important messages are given a low rank position, most users consider this case acceptable because they can still scroll down to browse the messages. Compared to Emails, news can be more suitable for filtering systems that exclude irrelevant information from presentation, as the information in daily news is not as important to users as that in Email messages. Based on the analysis above, exclusion of irrelevant information is a good way of presenting the filtering results in a consumer medical information system.

2.1.3 Approaches of Information Filtering

Various filtering approaches have been developed and some have been carefully examined through prototypes. Early information filtering systems include the Video On Demand (VOD) service system (Raskutti et al., 1997), the CiteSeer system (Bollacker et al., 1999, 2000), and the LIBRA system (Mooney & Roy, 2000). Generally, filtering approaches can be categorized into two types: congnitive and sociological—with different interpretations of their scope. According to the definition of Malone et al. (1987), cognitive filtering characterizes "the contents of a message and the information needs of potential message recipients, and then us[es] these representations to intelligently match messages to receivers," while sociological filtering "works by supporting the personal and organizational interrelationships of individuals in a community." In much of the literature (Morita & Shinoda, 1994; Sheth, 1994; Shapira et al., 1997), cognitive and sociological filtering are interpreted as content-based and collaborative approaches, respectively. This dissertation is limited to the information seeking process of the individual user, and no collaborative approach is applied in the study. Therefore, this literature review focuses on content-based filtering techniques.

Content-based filtering techniques were used in the early recommendation systems (Basu et al., 1998; Mooney & Roy, 1999; Claypool et al., 1999). In addition, some filtering systems handling consumer medical documents also adopted content-based approaches (Mostafa et al., 1999; Mostafa & Lam, 2000). Content-based recommendation systems work by mapping a description of the item to a profile of the user's interests. The success of the content-based approach requires a good representation of target items, and it is important for the system to build representations of the items based on the features of the items. These items can be automatically extracted features (such as word frequency for text items) or human-edited features (such as genre of movies). Then, the similarity between the user's interest profile and the items in the database could be analyzed quantitatively, with a threshold established to determine whether the specific item should be presented to the user.

A variety of natural language processing (NLP) approaches (or text analytical methods) have been developed to analyze item descriptions. The common approach to working with free text is to convert it to a structured representation. One basic method is to treat every word as an attribute. This method uses a Boolean value (0 or 1) to indicate whether the word exists in the text or an integer value to indicate how frequently the words occur in the text. During this process, words are stemmed, which is a process that counts the root forms of words rather than words in analysis. Some words are not related to the task or the domain and need to be filtered out prior to applying NLP or statistical analysis. The words that are removed are called stop words.

A key issue in the representation of content is termed weighting. A variety of methods have been developed to reflect the significance of a word to a document in a collection or corpus. One of the popular weighting methods of words is term frequency—inverse document frequency, often called tf–idf (Salton and McGill, 1983). In tf–idf methods, the weight of a word is calculated by this formula:

$$W_{ij} = tf_{ij} * idf_i = tf_{ij} * \log(N/n_i) \tag{2.1}$$

where tf_{ij} is the frequency of term i over document j; where idf_i is the inverse document frequency of term i over a collection of N documents; and where n_i is the number of documents containing term i.

When representing item description, it is important to consider the concepts contained. Latent semantic indexing (LSI) technique uses a mathematical approach called singular value decomposition (SVD) to recognize the relationships between the terms (Deerwester et al., 1988). The assumption of LSI is that words used in the same contexts tend to have similar meanings. Based on this principle, LSI is expected to extract associations between those terms. This extra information can be used to help the representation of information filtering. Foltz (1990) evaluated how well LSI works for filtering articles. The subjects in the experiment rated Netnews articles as either relevant or not relevant, based on the subjects' interests. The ratings from the initial 80% of the articles read were used to predict the relevance of the remaining 20% of the articles. The results indicated an average of 13% improvement in prediction after adopting LSI, compared with the keyword matching method. The research also showed a 26% improvement in retrieval precision based on ranking. The positive effects of LSI were verified in another study (Foltz & Dumais, 1992). After using LSI, an improvement was observed for the mean ratings of users. This improvement was attributed to the fact that users were allowed to access relevant documents without words in their interest profiles.

Despite many successful applications, content-based recommendation systems may fail if the content does not contain enough information to distinguish items the user likes from items the user does not like. This can occur for the content where separated words cannot represent the content well. Poems and jokes are representative examples for this type of content. In order to resolve this problem, more content-specific information needs to be integrated. For instance, the introduction of poems (i.e., one of metadata of the poem) may be added into the representation of content in the recommendation. A representative commercial application is exemplified in the recommendation of movies. The movie recommendation systems not only use movie content to make recommendations but also the genres of movies, actors, and directors.

To conclude, information filtering is not a single technique but a combination of many information processing techniques. Information filtering is a process of mapping a description of information to a profile of the user's interests. This study adopted human (health professionals) judgments based on the genres of medical articles (information description) instead of classifying the articles through NLP techniques. This method greatly reduces the impact of text analysis on filtering, whereby inferring filtering performance is mainly determined based on user model (profile) and mapping methods. In the next section, the literature related to user model and mapping process is reviewed, with a focus on the role of personalization and serendipity in information filtering.

2.2 Personalization and Serendipity in Information Filtering

Information filtering not only involves information retrieval and representation but also the acquisition and representation of users' interests, which is challenging because user profiles are dynamic and sensitive in both time and context. If the system fails to keep up with the shift in user interests, the system would eventually fail which points out the importance of learning component in the information filtering process. In this section, the literature on user profiles and learning methods will be reviewed along with the challenges associated with personalizing based on user models.

2.2.1 User Model in Personalized Information Filtering

Traditional user models mainly include keywords, which represent user interests, while the enhanced user models keep high-level knowledge representation about users. The user profile is the most important part in the user model. An overview of literature shows that most of the existing filtering solutions adopt a single profile that is built from all user inputs. Therefore, the user model in this dissertation study was formulated through similar method. The user profile is a single monolithic data structure, which contains a list of preferences. Specifically in this study, the preferences are medical topics indicated by users based on their interests. Some early researchers employed the term "query" to refer to user models (Belkin & Croft, 1992). In this view, a user model is actually a saved query (or a set of saved queries). LyricTime (Loeb, 1992), a music recommendation system, adopted the mood specified by users as a user model. After users were logged in, the system would deliver a playlist based on the selected mood from users. Widyantoro et al. (2000) proposed a more complex user model with a two-fragment user profile, which contained the following components: a long-term interest profile that captures user's general interests and a short-term interest profile, which kept track of a user's more recent, faster changing interests. Traditional user models are utilized in commercial sites as well. Amazon employs "favorites" to represent user models and the data of "favorites" are generated from the preferred categories set by users (Brusilovsky et al., 2007).

Recent researchers have paid more attention to the users' location, activity, or other contextual information in building users' profiles. Loeb and Panagos (2011) described a mechanism by which a user profile can be constructed in real time from relevant sub-profiles. Each sub-profile corresponds to a specific context (time, location), mood, task, and social context. Moreover, the constructed profile in the researchers' methods can be updated based on events, feedback, context information, or explicit user updates. In addition to academic study, many commercial applications also add the contextual information in creating profiles. For instance, Amazon.com supports the creation of multiple account profile fragments. Amazon's system allows users to go back to their personal history and specify which of the items were purchased as gifts for others. In this way, the system can learn that these items specified as gifts do not represent the user's personal preference but, rather, their friends' tastes.

The methods of acquiring users' interests can be explicit or implicit, depending on the source of the data. If the data are generated from user interrogation, the methods are considered explicit. In contrast, implicit approaches often refer to handling the data on observed user behaviors. Modern information filtering (or recommendation) always involves the explicit and implicit approaches simultaneously and sets different weights on the signals of these approaches in the learning module.

User Models Built on Explicit Methods

User interrogation has been widely used and is always considered the more reliable approach in acquiring user knowledge compared to other means. The basis for this method is that conceptually users know themselves better than others who may know them. In the process of interrogation, users are required to provide information on their interest directly to the filtering system. Previous researchers have examined different types of interrogation with variable levels of flexibility granted to users. Some filtering systems present a predefined set of profiles, granting users the ability to choose one from these profiles (McCleary, 1994). This method makes users capable of creating profiles very quickly. To increase the freedom of choice, some systems provide users with a set of terms and ask users to construct their profiles based on these terms. Some systems allow users not only to pick terms but also set the importance of terms in building the profiles (McCleary, 1994). These aforementioned approaches generate user profiles, which are represented by a vector.

User Models Built on Implicit Methods

The implicit approach acquires knowledge from users by analyzing their recorded behaviors in using the systems. The core of the implicit method is to translate user behaviors into user tastes. Due to the influence of user's contextual environment and personal emotional status, this process is deemed highly challenging. For instance, users' browsing actions can be interrupted by some uncontrollable issues, such as phone calls. Despite these difficulties, researchers have achieved positive advances in understanding implicit user behavior. Previous studies indicated that the time users spend reading data items relates to users' interest in the data (Morita & Shinoda, 1994; Konstan et al., 1997). The work conducted by Morita and Shinoda examined the correlation between the usefulness scores of articles and the time users spent reading them. The study demonstrated that there is a relatively strong correlation between these two factors (correlation coefficient is 0.49), suggesting that users spend more time reading documents that they find relevant than they do reading documents that are irrelevant. Some other types of user behavior can be leveraged to acquire interest information implicitly. Goecks and Shavlik (2000) presented an approach that learned users' preferences through observing the hyperlinks clicked on and the users' activity, as tracked by their mouse and scrolling movements. The results were consistent with the surrogate measurements of user interests, indicating the well-designed implicit approach was capable of predicting a user interest profile. Calvi and De Bra (1997) conducted a similar study in which they developed an adaptive learning approach that analyzes users' past navigation history. This study performed well in the filtering module of educational hypermedia systems.

In addition to promising results reported by researchers, there exists evidence on the limitation of user behaviors in practical applications. Kelly and Belkin (2001) investigated whether an earlier finding on reading time (Morita & Shinoda, 1994) could be replicated in another information retrieval context. In their study, results obtained from practical experiments led to a conclusion that was the opposite of a previous study: no significant relationship exists between the length of time that a user spends viewing a document and the user's subsequent judgment on document relevance. This suggests that the theory concerning the relationship between reading time and relevance is scope-specific. In order to explain this clearly, Kelly and Belkin (2004) investigated the effects of tasks on the effectiveness of display time as implicit feedback. They analyzed online information-seeking behaviors of seven subjects during a fourteen-week period. In the study, subjects were asked to identify their tasks, classify the documents that they viewed according to these tasks, and evaluate the usefulness of the documents. The experimental results showed that tasks and the time spent on tasks vary with the settings of the individual studies, resulting in no general, direct relationship between viewing time and usefulness. The effects of tasks on viewing time partially explain the inconsistency in the conclusions of the previous studies.

2.2.2 Blind Spots in Personalized Information Filtering

This problem of "blind spots" was proposed in earlier studies on filtering systems that used the documents that were previously judged as relevant for mapping purposes. These techniques rely on the representation consisting of terms in relevant documents instead of a vector of user interest dimensions. The incoming document is compared to a relevant document representation, and similarities between the two are used to establish whether the incoming document is relevant (above threshold) or not (below threshold). One early system utilizing this method is Newsweeder (Lang, 1995). The method of matching to relevant document representation requires minimum user involvement, which may be effective for some users. However, the negative aspects of this method are apparent. If document representation built from a user's initial judgment does not include some topics that are of interest to the user, then the new documents from those topical areas would not be presented to users. In this specific case, the relevant document representation approach would never successfully acquire all of a user's interests through the learning algorithms. This problem of "blind spots" not only occurs for the relevant document-representation approach but also in the filtering techniques based on other types of user models.

Mostafa et al. (1997, 2003) attempted to resolve the problem of "blind spots" in the simulation study on information filtering. They developed the SIFTER system, which can filter information based on content and a user's specific interests. Beyond user interest profiles, which consist of many interest topics, the learning algorithm adopted another vector that represents the probability of the topic in the profile being selected as the most relevant topic. By modifying the module of generating the probability vector, all the topics in the incoming documents were given dynamic chances to be presented to users. This method reduced the probable occurrences of blind spots in the filtering process, which was verified in an experimental study (Mostafa, et al., 1997).

2.2.3 Serendipitous Recommendation

Serendipitous discovery has led to many medical breakthroughs in diagnosis and treatment techniques (Meyers, 2007). Despite this, little attention has been paid to developing information filtering systems that facilitate this type of discovery in health informatics. Although some scholars have described how serendipity can happen in IR systems, particularly in relation to browsing (Bruijn & Spence, 2008), serendipity is still considered as somewhat a "threat" (Foster, 2003) to particular IR systems (Gup, 1997, 1998; Cooper & Prager 2000; Huwe, 1999). This is because IR systems have traditionally been created for enhancing the ability to conduct "known-item" or "known-problem" searches. However, in some situations search systems involve the limitations in relation to natural human information seeking. Miksa (1992) gave a detailed description of these situations in his book:

> "Intellectual knowledge appears to be characterized by a relatively unfocused sense of inquiry where the initial goal is not to find some particular informational answer or to fill some sort of reasonably anticipated informational gap, but rather to bring order to (or to re-order) an ill-formed mass of ideas or to map some vaguely arranged area of knowledge. Information retrieval in such situations takes on the character, then, of helping an inquirer think about what he or she appears to be interested in, and might be better conceived as an exploratory and game-like mechanism rather than a precise response mechanism." (Miksa, 1992).

The aforementioned information-seeking situations highlight the importance of serendipity and its place in the field of personalized information delivery. To meet users' various information needs in natural situations, serendipity has been incorporated into many personalized recommendation systems. LyricTime music system (Loeb, 1992) is one of the early recommendation systems. The system accommodates serendipitous access to information through occasionally adding randomly picked songs to the user's playlists. The strategy involved in LyricTime is similar to that which is being used extensively by Amazon, Netflix, and other systems. Sieg et al. (2010) examined the effect of ontological information on the ability of the filtering in terms of increasing the coverage of the information filter. The experimental results showed that the semantic context broadened the set of recommended items, thus improving the diversity and serendipity of recommendations. For expanding users' interests and enabling the discovery of new items of interest in a fashion that is sensitive to both time and context, Loeb and Panagos (2011) presented a generalized model for the production of augmented and context sensitive profiles. With this model, users can be exposed to a potentially broader set of relevant information. The serendipitous discovery in search was examined in a previous study (André et al., 2009a). The paper showed how to measure serendipity by analyzing user ratings. For measuring serendipity, participants were asked to rate search results on two dimensions: relevance and interestingness. It hypothesized that search results that are interesting but not highly relevant indicate a potential for serendipity. The new method of evaluating serendipity provided some guidelines for the experimental design in this dissertation study.

Some general ideas can be obtained from previous studies in terms of the potential importance of serendipity in content delivery. Blandford and Buchanan (2003) described the role of serendipity from the perspective of the usability of digital libraries. They stated that "new, interesting, and possibly surprising material" will help create a positive experience for users and said that the importance of serendipity in search should not be underestimated. The preliminary results from a case study demonstrated that information selected by a particular individual user profile may not always be the most relevant (Loeb & Panagos, 2011), implying the importance of serendipitous mechanisms. Some studies (André et al., 2009b; Fan et al., 2012) have discussed the presence of serendipity in information recommendation and then demonstrated and explained the specific influence serendipity had on the performance of information delivery. Ziegler et al. (2005) investigated the impact of applying topic diversity algorithms on book recommendation systems. The study found that topic diversification lowered accuracy, but increased user satisfaction. In addition, serendipity is always related to browsing. Marchionini and Shneiderman (1988) defined Internet browsing as "an exploratory, information seeking strategy that depends upon serendipity."

Some general ideas can also be obtained from previous studies in terms of how to incorporate serendipity in content delivery. Mostafa et al. (1997) integrated the serendipitous mechanism into personalized recommendation system by randomizing the articles presented on the top of the presentation window. They considered this as an efficient solution to the problem of blind spots. Beale (2007) proposed a serendipitous environment for offering supplemental information to support users in making new discoveries while keeping them at the center of an interaction with data mining systems. Some researches show how serendipity can be incorporated through ontological information in filtering settings (Sieg et al., 2010). In these experiments, researchers examined the effect of ontological information on the ability of filtering to increase the coverage of the information filter. The experimental results showed that the semantic context broadened the set of recommended items, thus improving the diversity of recommendations.

2.2.4 Role of Serendipity in Personalized Information Filtering

Personalization systems usually generate user interest profiles based on users' interface actions, feedbacks, and browsing history. Since users' interests are often multi-dimensional and dynamic, some personalized systems learn from initial user feedbacks and then stabilize user interest profiles by not allowing any change. However, stabilized user interest profiles will jeopardize personalization systems by engaging the tunneling behavior, which dramatically reduces the possibility that users encounter relevant content serendipitously while they use personalized systems. As a result, some user interests are excluded from the user interest profile which was generated and stabilized by the system. This scenario most often occurs in over-personalized systems. For instance, in a typical information recommendation environment, systems are increasingly trained to provide users with only the information that they explicitly requested or the information similar to what they received before. There is major concern rising from these over-personalized systems because previous studies demonstrated the fact that information seekers were often unaware or unable to articulate their actual information needs (Taylor, 1968; Belkin, 1980). Therefore, it is important to study how the serendipity plays a role in the personalized filtering process.

Fan et al. (2012) conducted a user study to examine serendipity-related issues in the medical domain. Fan's study was an important empirical research conducted to investigate the role of serendipity in personalized delivery of consumer medical information. Since relevant research literature in the specific research area (i.e., serendipitous recommendation of consumer medical information) is very limited, it is of significance to review this previous study comprehensively. There are two general aims for discussing the past study: First, to demonstrate the role of serendipity in the medical information seeking process; Second, to clarify how people use a personalized system to seek information and how the serendipity feature can be implemented into personalization systems. Both types of findings in the past study contributed to this study in terms of formulating research questions and designing the user study.

The broad goal in Fan's study was to understand the influence of personalization and serendipity on medical content delivery. A system named MedSIFTER was developed as the experimental platform, which can personalize the presentation of news articles based on a system-generated interest profiles. A parameter "r" was adopted to control the serendipitous recommendation in the article list shown in Figure 2.3. A higher value of r indicates a higher level of personalization, whereas a lower value of r means a lower level of serendipity. Medical news articles from MedLinePlus (a public website maintained by the National Institute of Health) was utilized as presented medical information in MedSIFTER system.

1. Article on topic-X (serendipitous recommendation)
2. Article on the topic of the highest interest strength
3. Article on the topic of the second highest interest strength
4. Article on the topic of the third highest interest strength
5. Article on the topic of the fourth highest interest strength
6. Article on the topic of the fifth highest interest strength

……………

Figure 2.3: Top Articles Presented to Users in MedSIFTER System

Thirty participants from the School of Library and Information Science (SLIS) program at Indiana University Bloomington were recruited to participate in a four-week user study. Study participants were randomly assigned into one of three groups (ten participants per group). Each group was assigned to an instance of MedSIFTER with a different personalization level. At the beginning of this study, each user was presented with twenty-six topics and asked to identify the strength of their interests in the topics. Then, the classes with the high interest strength were selected as the scope of the user profile. The clicks on the articles from the topics that were not included in the profile were off-topic clicks.

Figure 2.4 indicates the proportion of off-topic clicks (POC) per session for every group. No significant difference existed in users' off-topic clicks between the three groups ($p>0.5$). Each user implemented approximately three off-topic clicks per session for all three groups on average. The POC for all sessions was fairly high in Group 3 (r=0.8), which indicates that the participants showed interests in off-topic classes even in a highly personalized environment.

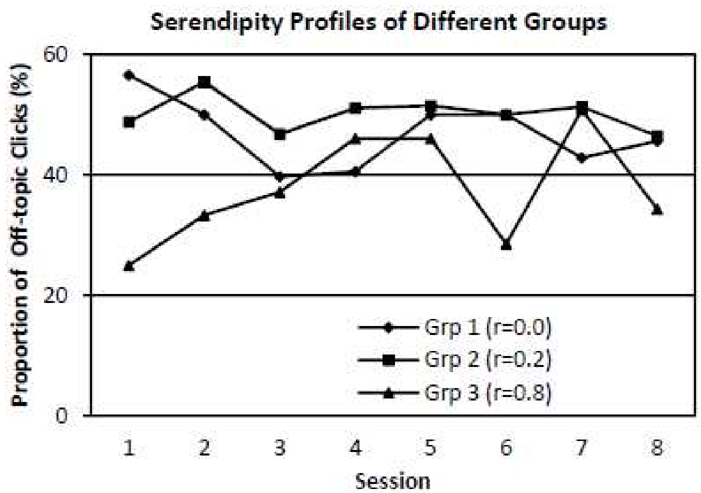


Figure 2.4: The Proportion of Off-topic Clicks Per Session (Fan et al., 2012)

For off-topic clicks, users' feedback on their satisfaction level was examined, with the results shown in Figure 2.5.

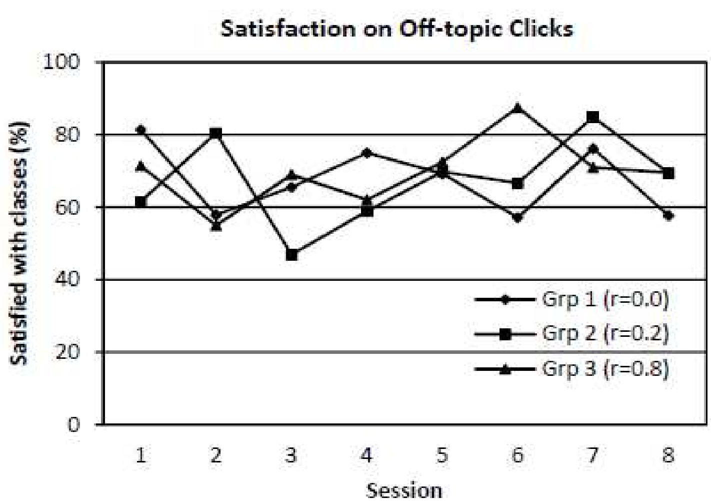


Figure 2.5: The Satisfaction on Off-topic Clicks (Fan et al., 2012)

Statistical analysis indicated no significant difference existed in satisfaction between the three groups ($p > 0.5$). The proportion of satisfaction with clicked news items per session (PSCS) remained fairly high across the study although there was fluctuation with each session. The off-topic clicks with satisfaction were aggregated across all participants and compared with total off-topic clicks. The results showed that users were satisfied with 67.5% of news items related to off-topic clicks.

In summary, the results of Fan's study demonstrated that serendipity can play a positive role in information retrieval. The study investigated the interaction between serendipity and personalization in an information retrieval setting. There were two main findings in the study. First, users had fairly high off-topic clicks, regardless of the personalization level to which they were assigned. Second, users showed uniform satisfaction with the outcome of serendipitous behavior across the three groups. The results from the usability analysis indicated that user responses were likely rooted in information-seeking characteristics that were dynamic or evolving in nature. In addition to the findings mentioned above, experimental system utilized in Fan's study is another concern for us and the experience learnt in the implementation of MedSIFTER system is useful to the design of our experiment system in this dissertation research.

2.3 Medical Information Seeking

Medical information is one domain where a fairly common scenario occurs: users' needs for information change suddenly as they deal with a personal health condition or the health condition of someone for whom they provide care. Therefore, it is important for individuals to learn the characteristics of medical information and the context of medical information consumption before they adopt appropriate medical information seeking strategies. This chapter first discusses the important properties of medical information and then introduces a variety of techniques and popular tools that have been recently developed for processing medical information.

2.3.1 Introduction of Medical Information

Information is a delicate commodity. One definition describes information as the data and knowledge that intelligent systems (human and artificial) use to support their decisions. Medical information systems assist doctors with their decisions and actions and also improve patient outcomes by making better use of information—creating more efficient means for patient data and medical knowledge to be captured, processed, communicated, and applied. Medical information mainly resides in electronic medical records systems deployed in inpatient hospitals, outpatient clinics, public health institutions, etc. Health information is different from the traditional format of information representation, which poses many challenges for intelligent analysis.

General health information includes consumer-level information and patient health records (PHR). The scope of this study is limited to handling consumer-level information and not medical information, which is mainly targeted exclusively to health providers. There exist numerous health-oriented websites and centralized databases that collect a large amount of consumer-level information (medical news, magazine articles, and some journal publications). The large amount and diversity of health information result in an overload of health information for consumers (including members of the general public and healthcare professionals). This study is conducted from the angle of enhancing the delivery of relevant consumer-level medical content to deal with information overload in medical domain.

2.3.2 Characteristics of Medical Information Seeking

Health information seeking is important for both members of the general public and healthcare professionals, such as physicians. The general public also desires to obtain information related to healthcare, especially after they have developed certain illnesses. Physicians strive for the newest medical information in order to update their own medical knowledge. An increasing emphasis on the use of evidence-based medicine, which focuses on the use of best evidence from the scientific literature in clinical decision making, also increases the importance of medical information seeking at the point of care. Due to the vast and ever increasing amount of medical information being made available on a daily basis, information seeking in the medical domain, in the absence of domain-specific tools, becomes unmanageable. Therefore, useful information delivery tools are becoming increasingly important in the medical domain, due to the explosion of clinically relevant evidence in published literature in particular.

Search is an important approach to helping people seek medical information. However, due to the complex nature of medical knowledge, simply limiting the target documents to those in the medical domain and indexing with a standard search engine is not sufficient. The range of sources of medical information, such as primary sources, secondary research sources, web pages, and popular publications, along with the range of end users, such as members of the general public, general practitioners, specialists, and researchers, lead to intricate requirements for search features.

Regarding the behavior of information seekers, most medical information searchers share two important sets of characteristics:

First, medical information searchers generally have limited abilities to express themselves when performing information seeking tasks. This is primarily due to a lack of medical knowledge. Such users would expect to read and learn all the information and knowledge related to their disease of interest. However, most search engines, even domain-specific ones, focus primarily on the accuracy of the retrieved documents, and this lack-of-diversity problem could be further aggravated by the nature of medical web pages. When discussing medical topics, medical web pages tend to use similar, but not exactly identical, descriptions by paraphrasing the contents of medical textbooks and research papers. Therefore, there is plenty of semantic redundancy in the pool of retrieved documents, many of which offer little use to information seekers. To appropriately address this problem while balancing the accuracy of search results, user queries would need to be refined or expanded to broaden the search scope horizontally.

Second, medical information searchers tend to input longer query terms, usually a paragraph or even a short essay, to describe their current health problems. This could also result from the lack of sufficient medical terms to appropriately describe their situations. Thus, seekers are often unclear about the medical problem they are facing and are unaware of the related medical terminology. It would be difficult for them to choose a few accurate medical phrases as a starting point for their searches. Unfortunately, most search engines will not perform well when dealing with long queries. This is made worse by the fact that some search engines have limits on the length of query texts and will simply reject the requests in the form of long queries.

Healthline (http://www.healthline.com/) is a domain-specific search engine dedicated to health information. It has a database containing descriptions of over 1,000 diseases and conditions, as well as over 4,500 symptom choices. However, the controlled vocabulary of symptoms limits the users' ability to describe medical situations appropriately. Such limitations may not provide the best match for the patients' current ailments during a one-time input of text (i.e. the description of symptoms). Realizing the special characteristics of medical information seekers and the complexity of medical information itself, researchers and engineers have since developed several medical domain specific information delivery systems in an effort to accommodate the needs of information seekers. These systems have symptom checker which utilizes filtering features.

WebMD site (http://www.webmd.com) provides a symptom checker system with interactive interface (shown in Figure 2.6).

Figure 2.6: The Symptom Checker Available in WebMD Site

After users select one body area from a body map, a list of related symptoms are presented to them. Based on each user's choices of symptoms, the possible conditions are retrieved through matching the symptoms to different conditions.

A health-education-oriented site called Symcat (http://symcat.com/) provides a way to increase the accuracy of describing medical condition by allowing users to input a list of symptoms step by step (as shown in Figure 2.7). Initially, users enter one symptom, then the system recommends a list of related symptoms, and users can either select from this list or enter another symptom. After a few steps, the system acquires a final list of symptoms. When users think they have fully described their current conditions, all the symptoms are submitted and they are mapped to possible diseases. Finally, a set of potential diseases is shown to users (see the right column in Figure 2.7). General educational information and treatment options are also presented upon selection of diseases. This filtering process may result in better performance of information delivery as it permits entry and specification of more symptoms.

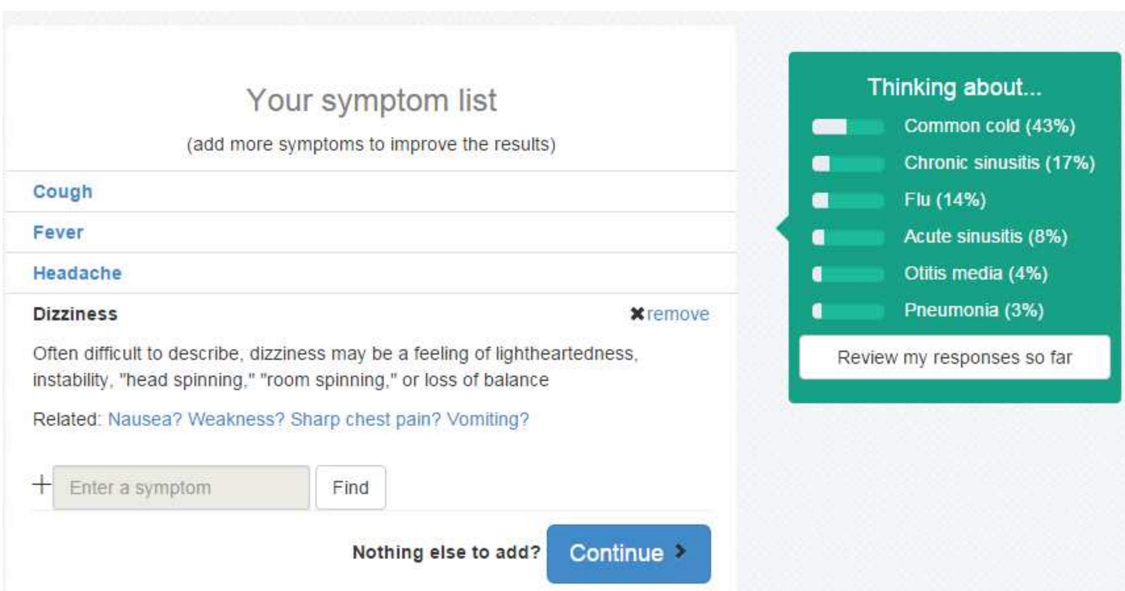


Figure 2.7: The Symptom Checker Available in Symcat Site

By reviewing current consumer medical information systems, we can see that search is an important solution when people have information needs in the medical area. Despite this, search systems may still pose barriers as query length is limited and domain knowledge may be required for formulating a good query. In IF systems, some of the challenges associated with search systems can be addressed or their negative impact can be minimized.

In summary, this chapter provides an overview of literature related to several important issues in the personalization of medical content delivery, thus clarifying the background of this study. It is clear from the literature review that the serendipity-related features in assisting medical information seeking have not been carefully investigated. Specifically, there is an extremely limited amount of literature about serendipitous discovery in the field of consumer medical information. Furthermore, the utility of IF systems for medical information delivery has not been fully explored. The gaps and potentials identified based on the literature review motivated this particular investigation and experimental study.

3.RESEARCH METHODS

The goal of this study is to understand whether and how incorporating serendipity can help people find unexpected but relevant news in the medical domain. For this purpose, the study adopted an empirical user evaluation to investigate the role of serendipity as well as different methods of incorporating serendipity in enhancing medical content delivery. This study falls under the broad research area of information seeking behaviors of users and the research methods were developed based on it. We start with an introduction of experimental system and user study procedures. After that, the critical categories of data and the methods for data collection are described. Finally, the chapter offers detailed discussion on the analytical methods of experimental results and approaches used for interpreting the potential findings. For this study, three filtering system modalities with different serendipitous features were developed to conduct user experiments. Based on these system modalities, three sets of experiments were carried out to answer the research questions, with each user group assigned to a specific modality. The details associated with experimental systems and the user study are presented respectively in the next several sections.

3.1 Setup of Experimental Environment

3.1.1 Personalized Filtering Environment

The personalized filtering environment was set up through implementing MedSDFilter system, which delivers medical news articles in a customized format. In every session with MedSDFilter, users are presented with a list of news articles (see Figure 3.1). A link to the full article is bound to each article title. In addition, an excerpt is added below each title to give a short description of the article content. All the articles are shown in the screen of browser. Users can see all the articles, without using a scroll bar. At the top-right corner of screen, there exists a “Go to Next Session” link. After reading the medical news in the current session, users can enter the next session to read more news by simply clicking the button. Users are not allowed to return to old sessions once they enter into a new session.

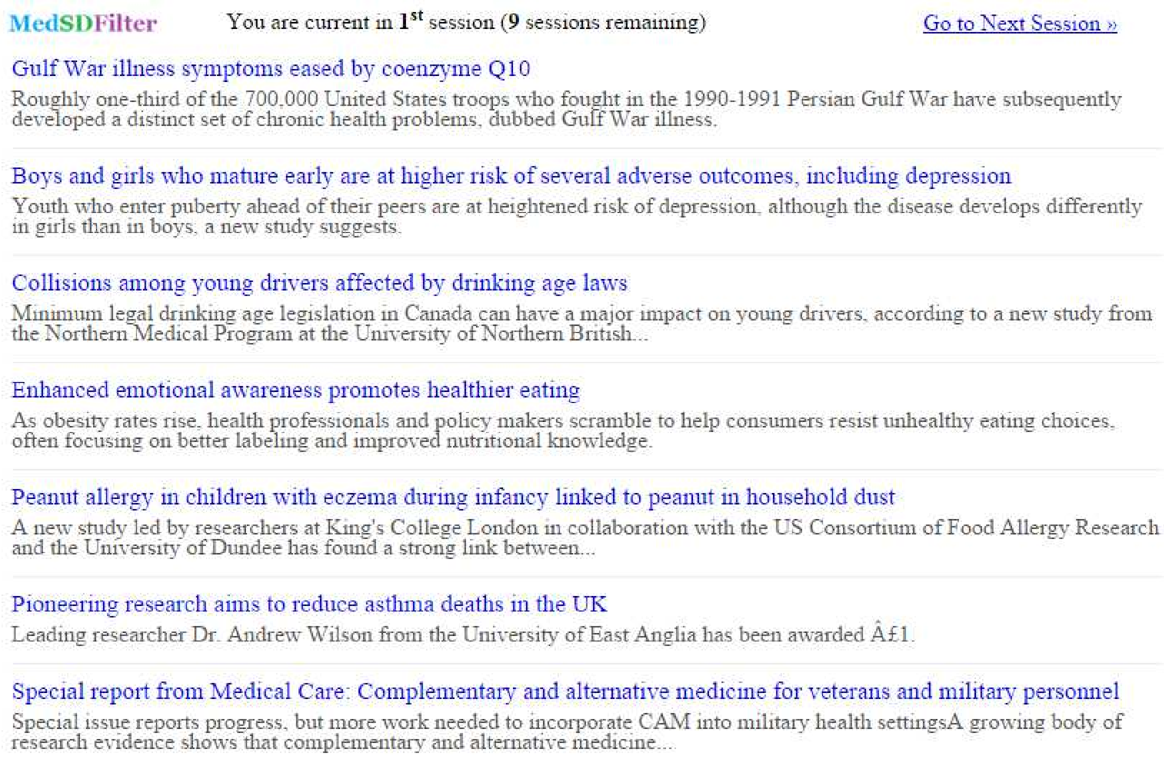


Figure 3.1: An Example of an Article List Shown in MedSDFilter

The Medical News Today (MNT) site was selected as the main data source for this study based on the site’s popularity. The site is a market leader for medical news, providing concise and accurate health information to general public and health professionals. In the MNT site, each medical news article is classified into one or multiple classes. News categories in the MNT site are alphabetically listed and shown in Figure 3.2. For this study, 30 classes with the high frequency of news update on the site were selected as medical topics of interest. They include diabetes, anxiety, women's/men's health, and heart disease, etc. (see Appendix A for the full set).

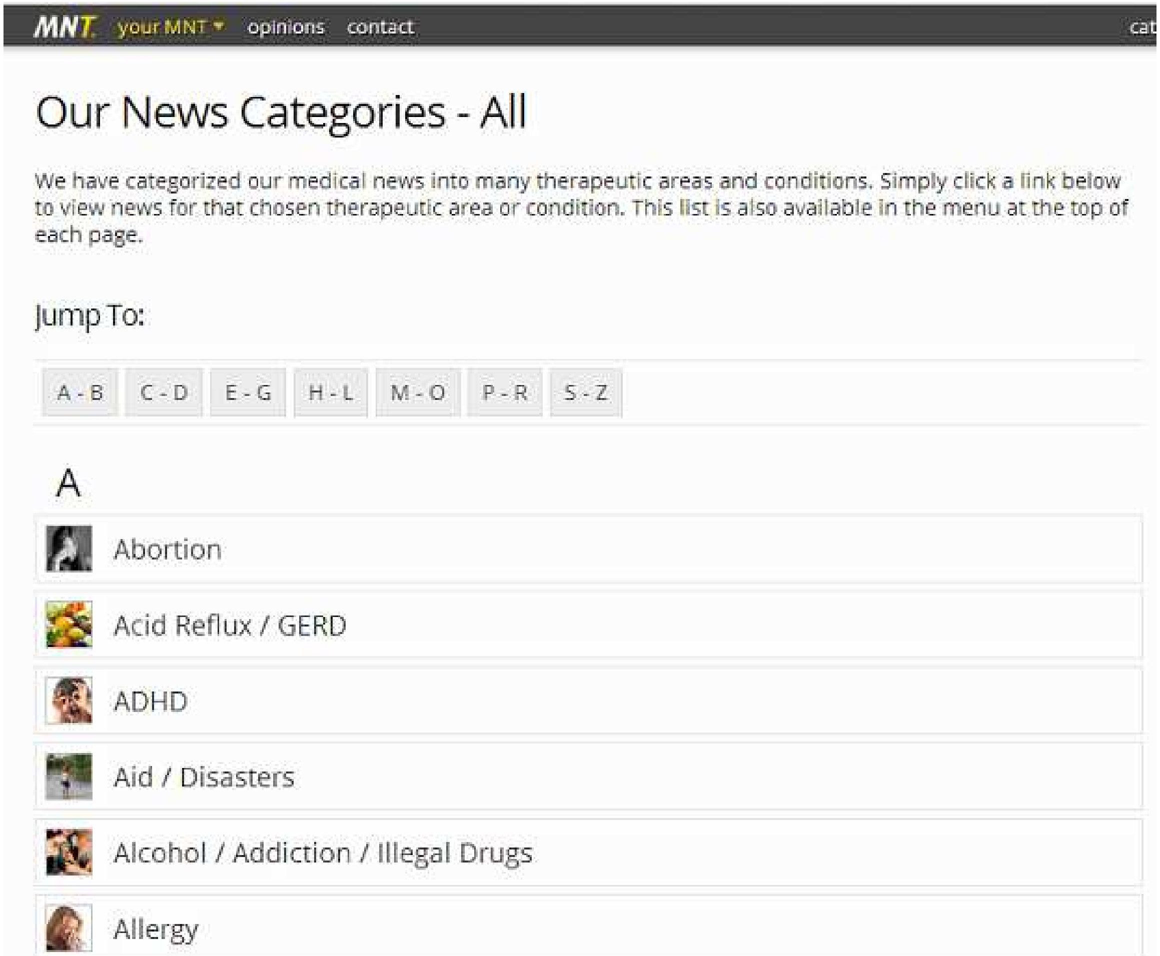


Figure 3.2: News Categories in the MNT Site

After obtaining approval from MNT, a crawler was run to retrieve recent news articles for each class before the study. As for the articles with multiple classes, they were only classified into the most important class. In this way, there were no duplicate articles when the data of all 30 classes were selected. Since the MNT site always ranks all the classes of one article by their importance, the top class in the list was considered as the most important one. As for 100 article pages crawled for each class, the article title, publication date and content were extracted. These data were sorted by publication date and stored into one table in the database. All 30 tables (each table for each class) were treated as the complete document set in the study. When filtering sessions start, the articles on each class are retrieved from the corresponding table in the order of ascending publication date. In order to avoid duplicate presentation of articles to one user, all articles in the database are labeled as “unused” before the first session starts. When one session completes, retrieved articles are re-labeled as “used” in the database and excluded from being retrieved in the following sessions.

3.1.2 Methods of Incorporating Serendipity

The section above introduces the setup process for the filtering environment. The next important issue in this study is how to incorporate serendipity in the filtering environment. In order to reduce the impact of ranking on serendipitous recommendation, we adopt a fixed position in the presented list to introduce serendipity. This means that our study does not test placement location. Placement is not an independent variable in this study.

Figure 3.3 shows how articles are presented to users in this study. In order to clarify the whole process of article retrieval and presentation, an example is given here. Let us assume a user selects 8 out of 30 medical topics to build system’s profile. These selected topics (SP topics) are ranked by the strength of user’s interest on each topic and are arranged in descending order. The other 22 medical topics (NSP topics), which are not selected by the user, are presented to the user by serendipitous recommendation. The filtering process involves the stages of article retrieval and presentation. In the retrieval stage, one article is selected from the dataset for each SP topic. As a result, 8 articles are retrieved from the experimental dataset. In the presentation stage, only the top 5 articles are listed according to the rank of their topics due to the limitation in total number of presented articles. Then, one article from a NSP topic is added below each article on SP topics, generating a total of 10 articles shown to the user in each session (see Figure 3.3).

1. Article on SP topic-1
2. Article on NSP topic-X (serendipity incorporated)
3. Article on SP topic-2
4. Article on NSP topic-X (serendipity incorporated)
5. Article on SP topic-3
6. Article on NSP topic-X (serendipity incorporated)
7. Article on SP topic-4
8. Article on NSP topic-X (serendipity incorporated)
9. Article on SP topic-5
10. Article on NSP topic-X (serendipity incorporated)

Figure 3.3: Article List Presented to Users in MedSDFilter System

From the description above, topic-X in Figure 3.3 is selected from all NSP topics. Presenting the articles on the NSP topics represents this study's strategy for incorporating serendipity. *In this context, how to pick topic-X from all NSP topics in each session becomes especially important.* Based on the article list described in Figure 3.3, *three methods* for picking topic-X are introduced respectively. In order to double check users' interests, topic-X is presented repeatedly for 2 times in consecutive sessions. In other words, users view 2 different articles on each topic-X. Below, we describe the three methods that introduced serendipity in this study.

1) RA Method (Based on Randomness)

RA method is the simplest solution of incorporating serendipity since it relies on randomness. It works as a baseline for comparing the other methods of incorporating serendipity. When RA method was implemented in this study, topic-X in every location was randomly selected from all NSP topics across all the sessions.

2) KA Method (Based on Knowledge of Topic Association)

In KA method, the topic-X is not randomly selected from all NSP topics. Instead, the topic-X is selected through analyzing its association (to be described later) with the SP topic above it in the presentation. For instance, as for the topic-X in the second article in Figure 3.3, all NSP topics are sorted by their association with topic-1. Based on this rank, these sorted topics are sequentially presented below topic-1 over all the 10 sessions. When KA method was implemented in this study, the NSP topic of the highest rank was selected as topic-X first, followed by that of the second highest rank.

For implementing the procedures above, it is important to establish topic associations. This study adopted manual methods instead of any technical solution in evaluating topic association because such an evaluation involves many contextual factors and should represent the real life clinical situations well. Two experienced physicians were involved in this study. Their long-term experience in the area of healthcare ensures that the data obtained about health topic association are valid. In the implementation, these two physicians were invited to fill out a survey that lists 435 topic pairs (based on 30 topics). As for each pair, the physicians were asked to select an integer value between 0 and 3 to best describe its association, based on their professional experience and knowledge. Here ‘3’ indicates the strongest association and ‘0’ means no association. In order to minimize judgment error across topics, the physicians were requested to adopt the same rating standards for all of these topic pairs when they conducted the evaluation. It is highly challenging to achieve this goal, considering the fact that a large number of judgments have to be made by each physician. As a result, a short range of topic associations (0-3, not 0-5 or more) was adopted to compress the range of selectable options for a judgment, making the workload of evaluation manageable. After the survey, the data obtained from the physicians were compared and analyzed through statistical methods. Based on average value of their judgments on topic association, the KA method was implemented for topic-X selection.

3) KAA Method (Adaptive KA method)

If NSP topics are not relevant to users' interests but presented repeatedly, users may be distracted by noise introduced. Based on the KA method, the KAA method is developed to resolve the problem of over-presentation of irrelevant NSP topics. The KAA method is similar to KA but it incorporates user feedback representing their level of interest. In order to display more relevant NSP topics to users in a limited number of sessions, a NSP topic is excluded if it has been presented to users multiple times and users have low or no interest. In this study, two articles on topic-X (one NSP topic picked for presenting) were shown to users in consecutive sessions. If users were not interested in these articles, the topic-X was not shown to users again until all other NSP topics were presented.

To summarize, the 3 methods adopt different resources and strategies for incorporating serendipity in filtering settings. RA method uses a random mechanism. KA method employs physician's knowledge on topic associations. KAA method uses physician's knowledge on topic associations as well as user's feedback on topic relevance. Based on these methods, the influence of serendipity was analyzed in the study. Additional details of system implementation are discussed in the next section.

3.1.3 Example of Implementing Serendipitous Recommendation

In order to easily visualize the article list and explain how the list is generated, an example is given. This example assumes that one user chooses two topics from five available medical topics (not 30 in the real user experiments) to build the system's profile. These available topics include "anxiety," "diabetes," "sleep disorder," "breast cancer," and "hypertension." Let us assume the user is strongly interested in the topic "breast cancer" and also has a weaker interest in the topic "diabetes." From this selection, SP topics include "diabetes" and "breast cancer." In contrast, NSP topics include "sleep disorder," "breast cancer," and "hypertension." In filtering settings, the articles on "breast cancer" are retrieved from the dataset and shown at the top, followed by the articles on "diabetes" because the strength of user's interest on "breast cancer" is higher than that on "diabetes."

The next step involves obtaining the association of these five topics. In this study, the associations of these five topics were judged by the two physicians. Table 3.1 shows the results of physician judgments where '0' indicates no association and '3' means the strongest association. The normalized associations (*r*) of the five topics (see Table 3.2) are calculated from physician's judgment by this formula:

$$r = \frac{\mathrm{R1+R2}}{6} \tag{3.1}$$

where R1 and R2 are topic association strengths judged by two physicians respectively.

Table 3.1: The Strength of Topic Associations Judged by Two Physicians

Top-A	Top-B	R1	R2
Anxiety	Diabetes	1	1
Anxiety	Sleep disorder	3	1
Anxiety	Breast cancer	2	1
Anxiety	Hypertension	2	3
Diabetes	Sleep disorder	0	2
Diabetes	Breast cancer	0	1
Diabetes	Hypertension	2	3
Sleep disorder	Breast cancer	0	0
Sleep disorder	Hypertension	1	3
Breast Cancer	Hypertension	0	1

Table 3.2: The Normalized Association of Medical Topics

	Anxiety	Diabetes	Sleep Disorder	Breast Cancer	Hypertension
Anxiety	1.00	0.33	0.67	0.50	0.83
Diabetes	0.33	1.00	0.33	0.17	0.83
Sleep Disorder	0.67	0.33	1.00	0.0	0.67
Breast Cancer	0.50	0.17	0.0	1.00	0.17
Hypertension	0.83	0.83	0.67	0.17	1.00

From the results shown in Table 3.2, all the NSP topics are ranked based on their associations to the SP topics. If multiple NSP topics have the same association with one SP topic, they are ordered based on their rank in the complete topic list.

As for SP topic "breast cancer", NSP topics are ranked in the following way:

Anxiety (r=0.50) > Hypertension (r=0.17) > Sleep Disorder (r=0.0)

As for SP topic "diabetes," NSP topics are ranked based on:

Hypertension (r=0.83) > Anxiety (r=0.33) > Sleep Disorder (r=0.33)

According to the ranks of NSP topics, the article list is separately generated for different methods of incorporating serendipity, which is introduced respectively below.

When the RA method is used, there exists no priority of recommendation for the NSP topics. It is possible to see the articles list in Figure 3.4. In this case, the "sleep disorder" article occurs below the "breast cancer" article, even though "breast cancer" and "sleep disorder" have no apparent association according to physicians' judgments.

Session 1 and 2:	Session 3 and 4:
1. New "breast cancer" article	1. New "breast cancer" article
2. New "sleep disorder" article	2. New "anxiety" article
3. New "diabetes" article	3. New "heart disease" article
4. New "anxiety" article	4. New "sleep disorder" article

Figure 3.4: Article List Generated Using the RA Method

When the KA method is adopted, different NSP topics are presented around each SP topic in sequence, based on the rank of their associations (see Figure 3.5). "Anxiety" articles are next to "breast cancer" articles in the first two sessions, followed by "hypertension" articles in the second two sessions. Similar situation occurs in other "serendipity" locations of the article list. "Hypertension" articles are next to "diabetes" articles in the first and second sessions, followed by "anxiety" articles in the third and fourth sessions.

Session 1 and 2:	Session 3 and 4:
1. New "breast cancer" article	1. New "breast cancer" article
2. New "anxiety" article	2. New "hypertension" article
3. New "diabetes" article	3. New "diabetes" article
4. New "hypertension" article	4. New "anxiety" article

Figure 3.5: Article List Generated Using the KA Method

The KAA method takes into consideration the potential overload of NSP topics and makes adjustments accordingly by analyzing user's feedback. Based on the rule set in KA method, "hypertension" articles are supposed to be next to "breast cancer" articles in the third and fourth sessions because it is the second most highly associated topic with "breast cancer." When the KAA method is adopted, the "expected" results in the third and fourth sessions are modified according to user' interests in previous sessions. If users have low or no interest in all "hypertension" articles in the first and second sessions, "hypertension" articles are excluded from presentation next to "breast cancer" articles in the third and fourth sessions. "Sleep disorder" articles are shown instead because "sleep disorder" is the third most highly associated topic with "breast cancer." The constructed article list is shown in Figure 3.6. In contrast, "hypertension" is not excluded from presentation in the third and fourth sessions if users indicate interest in any "hypertension" article in the first two sessions.

Session 1 and 2:

1. New "breast cancer" article
2. New "anxiety" article
3. New "diabetes" article
4. New "hypertension" article

Session 3 and 4:

1. New "breast cancer" article
2. New "sleep disorder" article
3. New "diabetes" article
4. New "anxiety" article

Figure 3.6: Article List Generated Using the KAA Method

To summarize, three methods of incorporating serendipity are compared based on a simplified filtering environment. By analyzing the differences in the way each article list is generated, it is demonstrated how these methods performed in actual user experiments. Thirty topics involved in the user study generated much more complex combination of articles than the example described in this section. In the next section, the details are provided on the procedure of the user experiments to capture users' interests in the generated article combinations.

3.2 Setup of User Study

3.2.1 Design of User Experiments

Thirty subjects were recruited to participate in a user study through in-person contacts or via ads distributed in email lists. Information seeking behavior of users can be affected by many environmental factors in the medical domain. This makes conducting evaluations associated with serendipity a challenge. In order to reduce the impact of users' heterogeneity, the participants recruited in this study were college staff members because they came from a group of people who not only had similar experience but also shared common campus environments. The total time allocated to complete the study was roughly 1.5 hours. As compensation for the time spent, each participant received $40. The study was divided into two parts. The main part involved the presentation and interaction with news items. The other part consisted of the pre-session and post-session surveys. Considering the stringent requirement related to security of health-related information, this study adopted de-identification strategies. Each participant was given one of 30 unique identification codes (UIC) which are alphanumeric and 10-digital, which were generated and randomized prior to the study. Any of RA, KA and KAA methods can be enabled by using the 10 identification codes. This setting was adopted to ensure that each serendipity level had the same number of participants. Participants' data, such as results of questionnaires and logs of interface actions, were linked using the identification code. The MedSDFilter system and the database were installed in a laptop with password protection. It was brought to the testing sites convenient to the users on the UNC-CH campus. When the test was completed, the laptop was brought to the researcher's office with password-protected access, which is located in the Laboratory of Applied Informatics Research (LAIR), RM. 300 Manning hall on the UNC campus.

The study was conducted under a given scenario. At the beginning of the study, the participants were requested to read through the description of the scenario as shown below.

"For this study, we would like you to imagine that you are interested in learning about new medical information, especially as it relates to you, your family, and friends. Since there are a lot of medical-related news articles, you have decided to use a filtering system that will recommend articles of possible interest. The system allows you to indicate areas of interest and it will recommend articles that you can read during your free time."

The scenario of this study is based on typical filtering settings in which news articles are ranked according to the strength of users' interests on article topics. From the scenario description, users may know that filtering system delivers relevant information in a ranked order based on their interests in medical topics. This implies that users are not completely blind to the utility and application of topic ranking. A static ranking of interesting topics is likely to enhance the possibility that users find the patterns of how articles are sorted. In many situations, the designer of user study wants to prevent users from finding patterns through strategies like randomness. One concern of these designers is that users may apply the found pattern to guide their information seeking behaviors and that this changes the initial context of information seeking. However, it is not a problem in this study because the pattern of topic ranking in personalized information filtering may be expected by users, and we reinforce the utility of the profile by presenting content based on it.

3.2.2 Procedure of User Experiments

Before each experimental session, a consent form was presented to the subjects. After they reviewed and signed the form, a one-page instruction was shown. The instruction provided a description of the study scenario and the experiment procedures. The subjects were told that if they have any issue after reviewing the instruction, they can ask the researcher to clarify it. Then, the subjects accessed MedSDFilter system and entered unique identification code to start the experiment. Each user was asked to complete 10 filtering sessions. One short survey was given before and after the experimental session. The consent form, pre-session survey (2c, 2d and 2e in Table 3.3), in-session ratings, and post-session survey (2j, 3a and 3b in Table 3.3) were adapted from previous established user experiments (Fan et al., 2012). They were modified to fit the current study needs. The process flow for the experiments is presented in Table 3.3.

It is common that that people follow different paces in reading medical articles. Therefore, to make the study realistic, each session was limited to 10 minutes (i.e. an estimate based on average reading time from the pilot study). When a session times out, it is switched to the next one automatically. If a subject completed reading all the articles of interest ahead of time (i.e. <10 minutes), that individual could go to the next session by clicking on the "Go to next session" link at the top right corner of the window. In order to ensure users completed reading each article of interest, the system double-checks by generating a popup box to collect his or her affirmation. The session switch occurs only if users check the box to indicate that they have completed reading all the articles of interest. Participants were made aware of the time limit per session and the session switch feature during the introduction.

Table 3.3: Process Flow for User Experiment

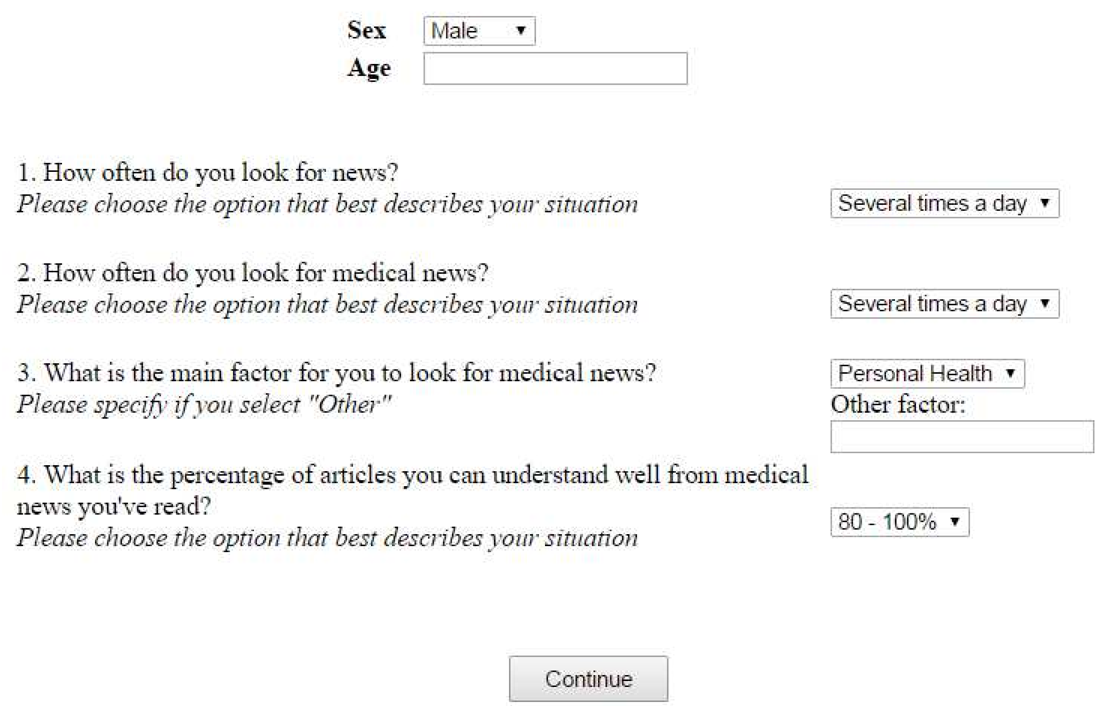
Before the experiment on MedSDFilter system	1a. Review and sign consent form 1b. Review study instruction (scenario description included)
	↓
Conducting the experiment on MedSDFilter system	2a. Launch MedSDFilter system (and its database) 2b. Enter UIC to load one serendipity model 2c. Questionnaire to obtain user's basic information 2d. Questionnaire to capture user's interest topics 2e. Questionnaire to capture user's interest strength 2f. Start one filtering session 2g. Read and rate articles based on user's interest 2h. Complete the filtering session 2i. Repeat [2f, 2h] for nine additional times 2j. Questionnaire to capture user's current interest 2k. Log out MedSDFilter system
	↓
After the experiment on MedSDFilter system	3a. Questionnaire to capture user's perceptions on serendipity feature 3b. Interview to clarify user's inputs in the questionnaire 3c. Compensation, receipt

3.3 Data Collection and Analysis

3.3.1 Collection of Experimental Data

The pre-session questionnaire has two parts. The first part of the survey was designed to collect data on participants' experiences of information seeking in their daily life. The screenshot of related questionnaire page is shown in Table 3.4.

Table 3.4: The Screenshot of Survey on Users' Information Seeking Experiences


Sex Male ▾
Age

1. How often do you look for news?
Please choose the option that best describes your situation Several times a day ▾

2. How often do you look for medical news?
Please choose the option that best describes your situation Several times a day ▾

3. What is the main factor for you to look for medical news? Personal Health ▾
Please specify if you select "Other" Other factor:

4. What is the percentage of articles you can understand well from medical news you've read?
Please choose the option that best describes your situation 80 - 100% ▾

Continue

The second part of the survey was designed to gather information on participants' interest on medical topics (including dimensions and strengths). Considering the MedSDFilter system is a filtering environment, a limitation was set on the maximum number of topics participants could select to build the system's profile. Instead of a given value, a range was adopted for the number of selected topics in the study since the number of interest dimensions of users may vary. Each participant was requested to select 5 to 10 topics of interest from 30 common medical topics (see Table 3.5). After that, the subjects were requested to select a value from a Likert-type scale (1-10) where '10' represents 'strongly interested' and '1' represents 'least interested,' as shown in Table 3.6. Users' selections on medical topics were then used to build the system's profile for the subsequent filtering sessions. In this study, content from the top 5 topics ranked by users' interest strength were presented to them.

Table 3.5: Screenshot of Online Survey to Capture Users' Interests

Please pick 5 to 10 health topics you are interested in

☐ Addiction	☐ Dentistry	☐ Mental Health
☐ Allergy	☐ Depression	☐ Neurology / Neuroscience
☐ Alternative Medicine	☐ Diabetes	☐ Nutrition / Diet
☐ Anxiety / Stress	☐ Eating Disorders	☐ Obesity / Weight Loss / Fitness
☐ Arthritis / Rheumatology	☐ Flu / Cold / SARS	☐ Pregnancy / Obstetrics
☐ Asthma	☐ Headache / Migraine	☐ Prostate / Prostate Cancer
☐ Cardiovascular / Cardiology	☐ Heart Disease	☐ Seniors / Aging
☐ Breast Cancer	☐ HIV / AIDS	☐ Sleep / Sleep Disorders
☐ Cholesterol	☐ Hypertension	☐ Stroke
☐ COPD	☐ Men's Health	☐ Women's Health

Continue

Table 3.6: The Screenshot of Survey Table to Capture Users' Interest Strength

Indicate the strength of your interests

Interest scale: 1 shows very little interest and 10 shows very strong interest.

Diabetes	◉ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Anxiety / Stress	◉ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Breast Cancer	◉ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Sleep / Sleep Disorders	◉ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10
Hypertension	◉ 1	○ 2	○ 3	○ 4	○ 5	○ 6	○ 7	○ 8	○ 9	○ 10

Ready to GO

During the filtering sessions, for each article the subjects clicked on, they were asked to indicate their interest in the article content. To operationalize this, a simple survey box (see Figure 3.7) was added in the full article window to collect an interest value regarding the article content. The MNT site (news source of this study) adopts a 5-star rating system. Similarly in this study, users were requested to identify their interest levels by selecting a value from a Likert scale (0-5), where ‘5’ represents ’strongly interested’ and ‘0’ represents ‘not interested.’ After users submitted their ratings, the full article window was closed. In the implementation of KAA method, users were considered to have low interest in the article if they selected a value smaller than 3 on the 0-5 scale. Beyond interest value, other types of data were also collected from interaction behaviors. When a subject clicked the title of the article, the article ID and the time stamp of the click to open an article were recorded. When users submitted their ratings on interest levels indicating they completed reading the article, the time stamps were recorded again. The difference between these two time points associated with an article was considered an important indicator of time spent reading the article. Reading time is used as indirect evidence in evaluating serendipitous discovery of relevant articles. In the pilot study, users usually clicked relevant article from top to bottom, implying that the order of users’ clicks was not strongly correlated with their interest strength in the context of this study. Therefore, we did not have particular concern over the order of users’ clicks in data collection.

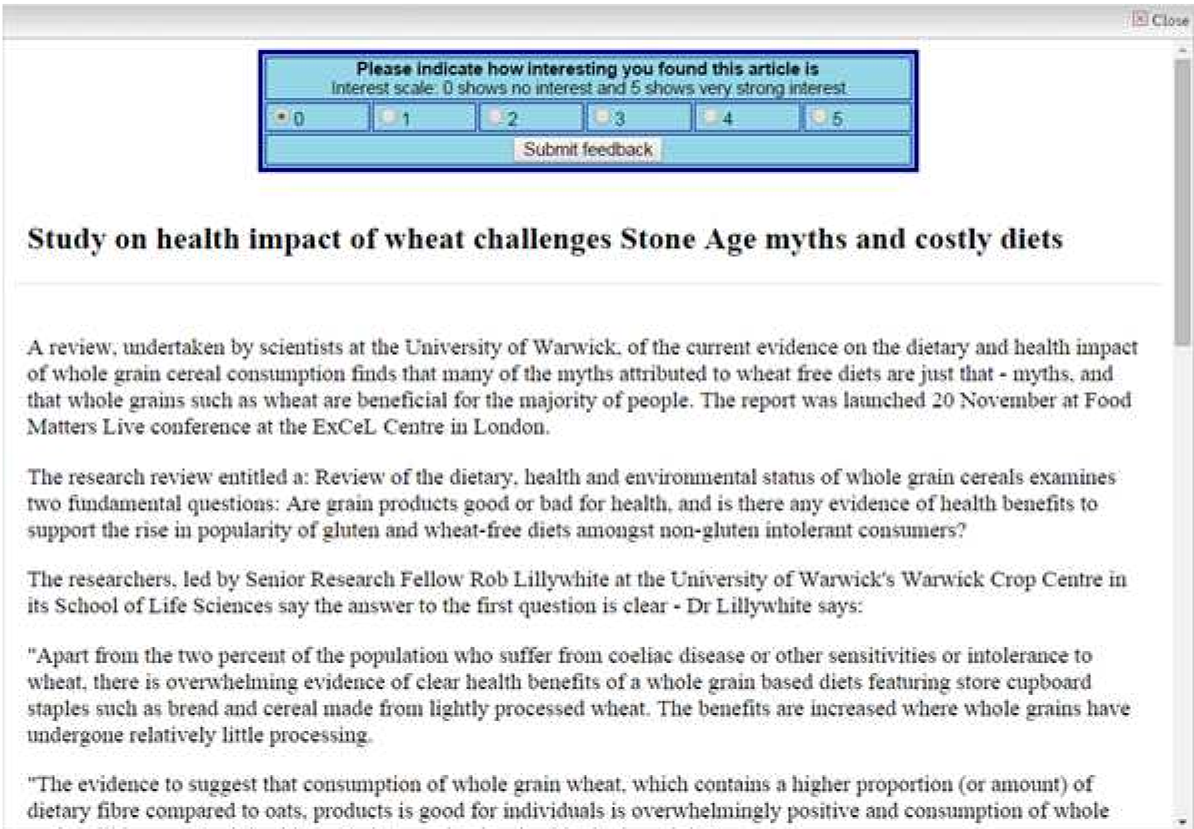


Figure 3.7: An Example of a Full Article Window

After completing 10 sessions, the post-session interest values were collected from each subject based on the survey as shown in Table 3.6. Users were presented with all 30 medical topics, and they were requested to make selections based on their current interests on these topics. By comparing the results of the pre-session and post-session surveys, the variations in users' interest profiles (whether in dimension or strength) were established. At the final stage of this study, users were asked to answer a few open questions. These questions (see Part 2 in Appendix D) were used to gather additional data to comprehensively evaluate users' perception of serendipity feature and system performance. After users completed this part of the study, the researcher reviewed their answers to the questions. If the information collected was not clear or insufficient, the researcher asked users to make some clarifications or provide additional information through a quick interview.

To summarize, this section describes the experimental data collection procedure in this study. In the next section, the methods for analyzing the collected data are introduced and discussed in detail.

3.3.2 Analysis of Quantitative Data

Both quantitative and qualitative data were collected. The first set of quantitative data originated from the answers to the questions in post-session survey (see Part 1 in Appendix D). By simple comparison of users' inputs or selections, the distributions of age, gender and information seeking experiences of subjects are established.

The second part of quantitative data concentrated on users' medical topic selection. The interest values collected before and after the 10 filtering sessions were main sources of data. By comparing pre-session and post-session data, it is possible to identify whether users' interest evolved and whether it expanded to include additional medical topics. The inclusion of extra interest topics would be due to many factors, and one of the most important factors may be the information in the articles users have read during filtering sessions. For instance, after reading some medical articles, users may have discovered that they were actually interested in the medical topics even though they originally did not include them in their profile. In order to clarify this, these extra interest topics are examined carefully. In pre-session questionnaires, users were asked to select a limited number (5-10 out of 30) of topics for building system's profile. It is possible that users had interest in unselected topics though their interest strengths were not as high as those on selected topics. For identifying the topics with definite interest change, our focus was placed on certain unselected medical topics that users rated highly in the post-session survey. By analyzing users' click and ratings on the articles from extra interest topics, additional evidence was generated to establish the correlation between users' interest changes and the content they have read. For better clarification, more direct evidence was retrieved from qualitative data. In the post-session survey, a set of questions were posed to users to clarify the factors associated with their interest changes (if extra interests were identified).

The third part of quantitative data is from user's ratings and clicks in using MedSDFilter system. The analysis helped establish users' actual interest based on both explicit and implicit methods. Two approaches are followed: a) Implicit approach – participants' mouse click actions and reading time at the article level were used to indirectly capture their probable interest in the medical topic; and b) Explicit approach – participants' feedback on interest strength at the article level was used to capture their actual interest in the medical topic. If users were interested in the article, they are considered to be at least partially interested in the article topic. The data captured using these two approaches were analyzed to elucidate key influences of serendipity on filtering. This part of the user study required judging whether a user's click on an article was an unexpected event. In order to reach this conclusion, the topic of the clicked article was compared to the system's profile. If the topic was not included in the system's profile, the click and opening of the article was considered an unexpected event. Since this type of clicks occurred on the articles from non-system-profile (NSP) topics, they were called NSP clicks. In contrast, the clicks on the articles from system-profile (SP) topics were called SP clicks. Interest strength indicated by participants was also analyzed for the corresponding clicked articles. This provided a more useful way of analyzing the influence of serendipity. By analyzing the behavior at the group level for the three experimental groups, additional evidence was obtained to establish if the different methods of incorporating serendipity assisted participants to find unexpected but relevant news.

The fourth type of quantitative data came from users' ratings of satisfaction on system's performance. It involved collecting users' judgment based on overall system's performance. The goal here was to establish whether there existed a significant difference in satisfaction levels between the three experimental groups. Since the main difference of these groups was in the methods of incorporating serendipity, the analysis provided a way to clarify the relative superiority of the serendipity methods. In addition, users' ratings of satisfaction on system's performance were compared to their ratings of interest in article contents.

After clarifying whether serendipity influences the way in which people seek information, the next important issue was to identify factors that participants thought as influential in serendipitous discovery of medical news. This question was answered by analyzing the qualitative data collected in the study.

3.3.3 Analysis of Qualitative Data

This phase of analysis was conducted to identify direct evidence of users' reaction and response to serendipity. The goal of this analysis was to learn how serendipity could be introduced in an effective manner and without causing distraction or blind spot. Upon completion of the last session, participants were asked to evaluate system performance. The evaluation was based on a list of questions (see Part 2 in Appendix D). The answers to these questions constituted qualitative data of this study. From users' comments and responses in the post-session survey, the dominant theme associated with serendipity was identified and compared across users.

In the post-session survey, Question 2.1 aimed to learn about the positive/negative aspects of filtering system through users' evaluation. In order to reach this goal, important points regarding system's performance were manually retrieved from participants' comments. By comparing these comments across groups, we could obtain general idea regarding how user's perception of the system differed across the three groups.

In the post-session survey, Question 2.2 dealt with the way the MedSDFilter system could be improved. Users were expected to describe the problems and potential solutions they have found when using MedSDFilter systems. These problems were not limited to serendipitous recommendation. They could have involved many aspects of the system in terms of the interface and functionality and speed. Similar analytical methods applied on Question 2.1 were adopted for handling the answers to this question. The problem description was manually extracted from the users' comments and then compared across users. Finally, possible causes and potential solutions of the problems were carefully analyzed, which provided useful insights into the design of a filtering system.

In the post-session survey, Question 2.3 directly asked whether and why participants would use the system. The negative and positive answers were counted and their rationale was analyzed. In practice, participants' willingness to use the system could be influenced by many factors such as the user interface, the content, and the operational speed. Considering serendipitous feature is the main focus in this study, the participants' opinion associated with content was extracted from answers and compared with related information in the answers to

Question 2.1 and 2.2. Through comprehensive analysis of all the responses in Question 2.1-2.3, we could obtain some useful evidence to explain users' information seeking behaviors in filtering sessions.

In the post-session survey, Question 2.4 and its sub questions focused on user's attitudes to unexpected discovery. In these questions, users were asked whether the system provided them with unexpected news items. They were also requested to comment on their experiences. By checking the themes in the answers for different groups, evidence for serendipity could be obtained in terms of its role in content delivery as well as the methods of effectively incorporating it into practice.

It was assumed that the three groups of users would have different views with regard to the unexpected news items they may have encountered. Unlike RA method, KA and KAA methods are developed to introduce NSP topics around SP topics based on their associations. In the article list built through KA or KAA method, medical topics of neighboring articles were likely to vary but in less disruptive fashion as compared with the RA method. Different from KA method, KAA method modifies the presentation of unexpected articles based on users' feedback. If an unexpected article and its topic were judged as relevant in the KAA group, new articles from the topic (i.e. an unexpected but relevant topic) were retrieved and presented. Therefore, it is likely that KAA method added more unexpected but relevant articles in the presented article list than KA method. In addition to difference in the article list built, users' evaluation on "unexpected" articles could be variable in the experiments. After a NSP topic was presented for a few times, users might think the topic was not "unexpected" to them. Based on the discussion above, it is possible that users' comments in the questionnaires could be not enough to differentiate three serendipity methods in terms of users' response to unexpected discovery. Therefore, users' interface actions were integrated for making a better evaluation.

Question 2.5 directly asked what factors caused the change of users' interests after 10 filtering sessions. This question was proposed only when a difference was detected between users' interest profiles in pre-session and post-session surveys. The analysis of these answers attempted to establish the influence of the role of serendipity in reducing potential blind spots during information filtering. By analyzing this part of data, we clarified some internal relations between the change in users' interest profile and unexpected news content delivery described in user's answers. Particularly, the analysis aimed to verify the possible identification of hidden interests by applying serendipity in medical information filtering systems.

In summary, this chapter starts with a detailed description of filtering environment setup and serendipity-incorporation methods. Then, it reports the procedures and data collection of user experiments. Last, analytical methods of quantitative and qualitative data are proposed, which are used to deal with experimental results described in next chapter.

4. RESULTS

4.1 Characteristics of the Subjects

4.1.1 Age and Gender

Thirty subjects were recruited, and all completed the study in its entirety. They were randomly assigned to one of three groups. The average and standard deviation of ages for each group were calculated respectively, with results listed in Table 4.1. The results show that the average age is 42.7 years. In terms of group comparison, the average age of Group 1 is slightly higher than that of Group 2 and 3. The results of standard deviation (STD) show that Group 2 and 3 have more apparent within-group age variance than Group 1.

Table 4.1: Average and Standard Deviation of Subjects' Ages in Three Groups

	Average	STD
Group 1	44.0	13.6
Group 2	42.4	18.0
Group 3	41.6	17.3
All Groups	42.7	15.9

Age is an important factor for medical information seeking. A previous study (Turk-Charles et al., 1997) reported that older and younger adults varied in terms of their information-seeking behavior. In order to examine if the differences in age among the three groups was significant, a one-way ANOVA test was conducted. The results indicated that there was no significant difference ($p=0.964$) in subject ages among three groups. Next, the distribution of subjects' age was analyzed, and the results are shown in Figure 4.1.

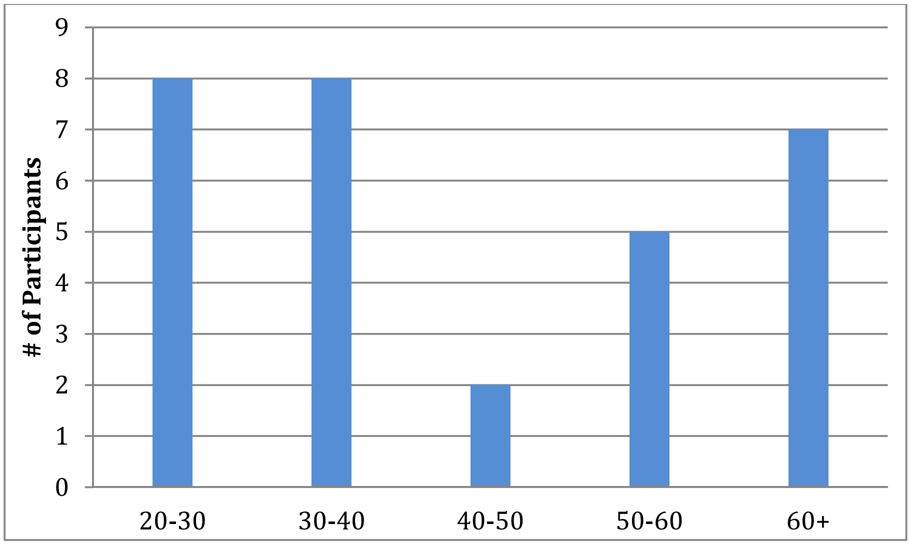


Figure 4.1: The Number of Subjects in Different Age Ranges

As shown in Figure 4.1, all participants are 20 years old or older. Twenty-two of them are 30 years or older, accounting for 73.3% of the total. This outcome was as expected, considering all the participants are UNC staff members who have worked for some years on campus. One of main features observed is that fewer subjects in the range of 40-60 years old participated, compared with other age ranges. Since recruitment can be affected by many contextual factors during the recruitment process, it is hard to accurately identify the cause of this observation. One possible cause may be that the staff members in the range of 40-60 years old were busy with their work-related responsibilities and thus did not actively respond to inquiries about participating in this study.

As for the gender of subjects, 13 of them were male and 17 of them were female, accounting for 43.3% and 56.7% respectively. This can be compared with the readers in the

MNT site, which was the source of the content used in the study. According to a Quantcast[2] report (see Figure 4.2), most of readers are female in the MNT site, matching the distribution of the subjects in this study.

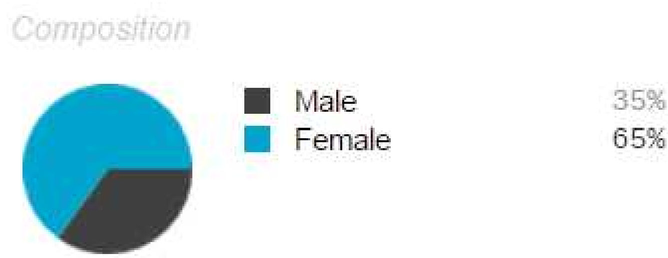


Figure 4.2: The Distribution of Subject's Gender in the MNT Site

Next, the distribution of subjects' gender across the three groups was analyzed. The results are shown in Figure 4.3. We can see that Group 1 and 2 have more females than males, while Group 3 has the same number of male and female subjects. Since the subjects were randomly assigned to different groups, the gender distribution in each group was unpredictable before the study.

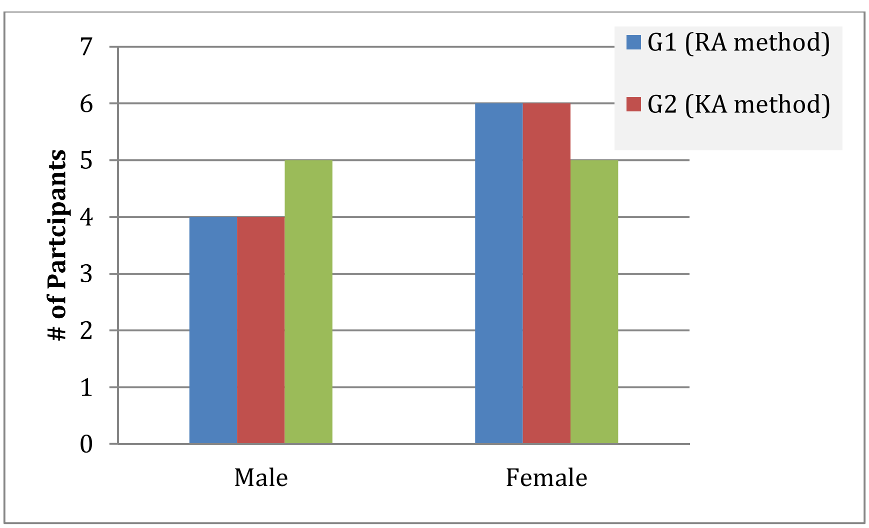


Figure 4.3: The Number of Male and Female Subjects in Different Groups

[2] Quantcast is one of leading tech companies which provides free, accurate and dependable audience insights for websites.

Besides the variables of age and gender, subjects' information seeking experiences may affect their performance. In order to obtain more background information to help judge participants' behavior, a list of related questions were utilized in the study. In the next section, the analysis conducted on the answers related to information seeking experience will be presented.

4.1.2 Information Seeking Experience

a) Frequency of Looking for News

Table 4.2 shows the frequency with which the subjects looked for news in general. It was found that all the subjects looked for news at least once a day. From this result, we can know that they often seek news in their daily life.

Table 4.2: The Frequency with Which the Subjects Looked for News

Frequency	Several times a day	Once a day	Once a week	Once a month	Once a year
% of Subjects	53.3	46.7	0.0	0.0	0.0

Subjects were asked about their experience with seeking medical news. Table 4.3 shows the frequency of seeking medical news for all the subjects in the study.

Table 4.3: The Frequency with Which the Subjects Looked for Medical News

Frequency	Several times a day	Once a day	Once a week	Once a month	Once a year
% of Subjects	0.0	36.7	23.3	23.3	16.7

From the table, we can see there exist wide variations in the frequency. This diversity was as expected because medical information seeking of subjects may be related to individual preferences and differences in their personal health. The majority of subjects indicated that they looked for medical news at frequencies ranging from once a day to once a week. As compared with the majority of subjects, a few subjects reported that they looked for medical news at very low frequency (i.e. once a year).

b) Factors of Why Participants Look for Medical News

Previously, it has been discussed that people's interests in medical topics can be affected by many contextual factors, such as personal health conditions. In order to clarify which factors play an important role during this process, subjects were asked to indicate the main factor that persuaded them to look for medical news. To capture other factors not included in the presented factor list, users have the option to type their answers directly. Table 4.4 shows the distribution across different factors.

Table 4.4: Factors that Influenced Seeking Medical News

Factors	Personal Health	Family Health	Friend's Health	Public Health	Other Factor
% of Subjects	60.0	16.7	6.7	13.3	3.3

The data in the table above clearly indicates that the majority of subjects looked for medical news because they were concerned about their personal health. Interestingly, friend's health as main factor involved a lower percentage than public health, given the fact that people always have a stronger relation with friends than with the general public. The low percentage may be related to the situation that personal health is private and not shared between friends in most cases. In addition, one subject who worked at UNC Hospital reported that she looked for medical news due to other factors (i.e. to make decisions in patient care). This indicates that, sometimes, the factor of why one looks for medical news can be related to one's work.

By comparing the data across three groups, we found no apparent difference regarding the factors of why participants look for medical news (Figure 4.4). This was as expected because subjects were randomly assigned to each group. In all groups, most subjects (6 out of 10) selected personal health as the main factor.

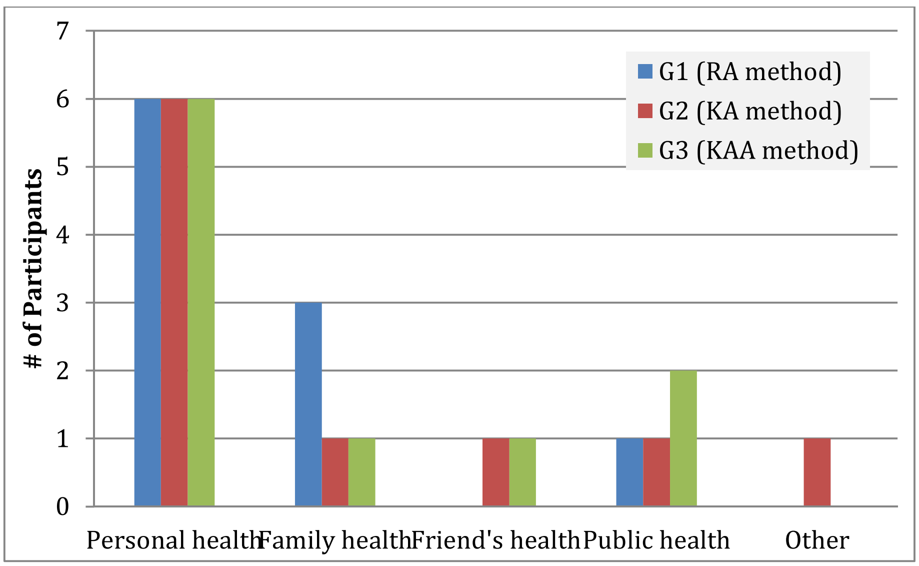


Figure 4.4: Preference of Medical News Seeking Factors in Different Groups

c) Levels of Understanding Medical News Content

During the experimental sessions, subjects were required to read medical news and rate their levels of interest. To ensure such ratings are valid, it is important that the subjects can comprehend the main points covered in medical news and make appropriate judgments based on the information. Table 4.5 shows the comprehensibility of medical news by the 30 subjects.

Table 4.5: Levels of Understanding Medical News Content

Understanding Percentage	80-100%	60-80%	40-60%	20-40%	0-20%
% of Subjects	43.3	33.3	23.3	0.0	0.0

Most subjects (76.6%) can understand 60% or more of medical news content well, and the others can understand 40-60% of medical content well. The results were not surprising. The content of medical news includes a large number of terminologies and jargon, and it is highly challenging for the general public to understand the content well due to the lack of professional knowledge. About 7 out of 30 subjects selected a comparatively low percentage (40-60%) when they were asked about how much information they can understand well. One thing to clarify here is that subjects' answers on the percentage were based on well-understood part of article content. This means that the subjects could still obtain some useful information from the content they could not understand well. This portion of content could play a positive role in helping the readers capture the main points of the article as a whole.

The data from the three groups were compared based on subjects' understanding levels of medical news, with the results shown in Figure 4.5. Group 1 and Group 3 have exactly the same pattern in participant's distribution regarding understanding levels. In contrast, more participants in Group 2 had high (80-100%) or medium level (60-80%) of understanding, compared with the participants in the Group 1 and Group 3.

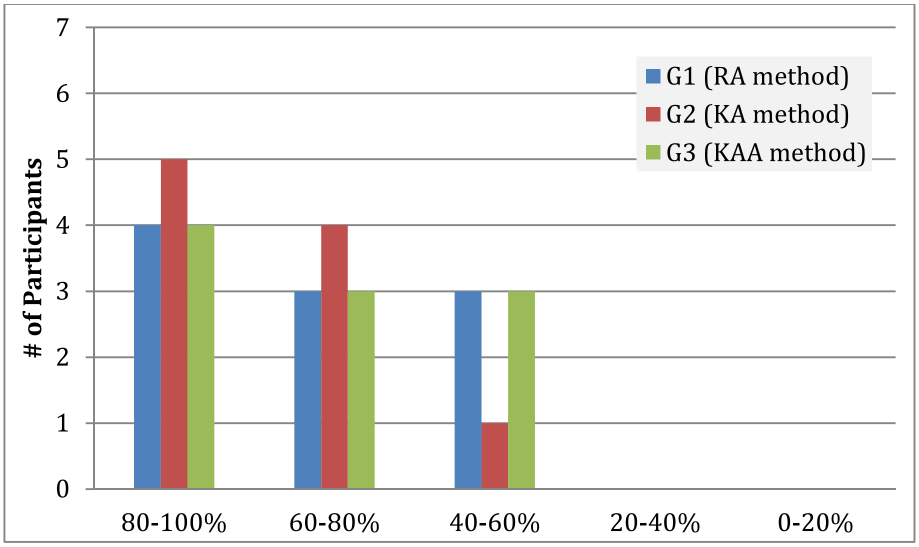


Figure 4.5: Levels of Understanding Medical News Content in Different Groups

4.1.3 Summary

This section presents the results obtained from the background questions (see Part 1 in Appendix D) in the pre-session questionnaires. These results were mainly used to clarify the characteristics of participants including age, gender, and information seeking experience. The important findings support the following observations. First, there was no significant difference in the age of subjects in different groups. Second, all participants looked for news once a day or more, and they looked for medical news much less frequently. Third, personal health was the main factor motivating medical news seeking for most participants. Fourth, the majority of participants had a high or medium level of understanding on medical news content. Beyond these characteristics of the subjects, medical topic association, established by physicians, is another important factor for this dissertation study, and it is discussed in the next section.

4.2 Topic Association Judged by Physicians

The evaluation of medical topic association is complex and may vary with a different angle of thinking regarding potential connections among topics. In order to obtain reliable data, two doctors of medicine (MD) who have worked as physicians for many years were invited to judge the association among thirty medical topics (see Appendix A). Admittedly, the evaluation from these two physicians was not perfect, but it provided useful assessments for the setup of an experimental platform, based on which the important research questions were examined.

4.2.1 Difference between Physicians' Judgments

As discussed in the chapter on research methods, two physicians were requested to select a value between 0-3 to indicate the strength of topic association. Figure 4.6 shows the distribution of physicians' judgments.

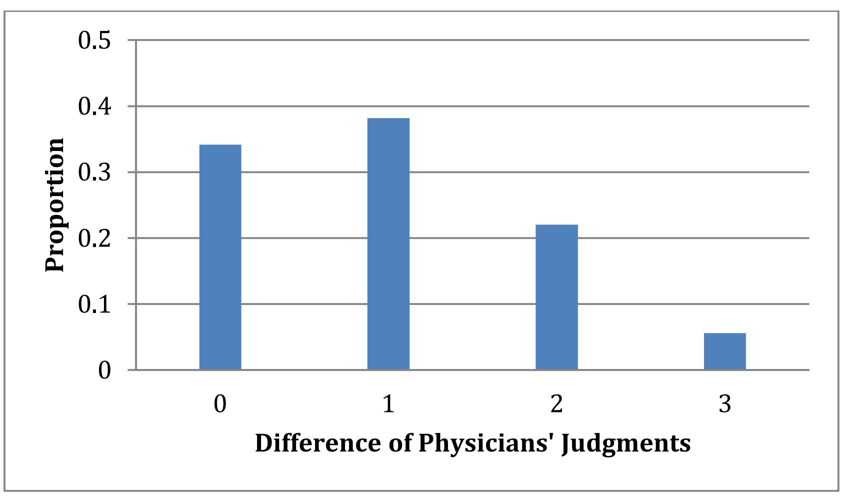


Figure 4.6: The Distribution of Difference between Physicians' Judgments

As shown in Figure 4.6, in about 72.4% of cases, the difference was rather small (i.e. the difference of 1) or no difference was detected. This finding indicated that the physicians had similar judgments on topic association in most cases. In addition, the difference in judgment between two physicians was distinct (i.e. the difference of 3) in 5.6% cases (25 of 435 judgments). In order to clarify the context in which such a difference between physicians' judgment occurred, the topics associated with these 25 cases were retrieved and then ranked by occurrence frequency. The 5 topics with the highest frequency are listed in Table 4.6.

Table 4.6: Topics with Distinct Difference in Physicians' Judgments

Topics	**Freq**
Women's Health	6
Sleep / Sleep Disorders	6
Men's Health	5
Alternative Medicine	4
Neurology / Neuroscience	4

The topics presented in Table 4.6 are related to general health issues such as women's and men's health. This implies that physicians are likely to make distinct judgments when they estimated topic pairs involving general health topics. This may be attributed to a wide variety of perspectives from which physicians may view general topics. In contrast, for specific topics such as diseases, it is likely that professional knowledge of physicians influences them to make similar judgments on medical topic association. In order to verify this, further analysis of physicians' judgments was conducted. Regarding any medical topic, each physician judged its association with the other 29 medical topics, which generated a vector on association strength for each topic. The correlation between the judgments of two physicians (i.e. two vectors on association strength) was calculated for each topic, with the results sorted in descending order. From the sorted topic list, we found that all top topics were concerned with specific medical areas. They included cholesterol, hypertension, addiction, cardiovascular, allergy, etc. This finding implied two physicians were more likely to make similar judgments on specific topics as compared with general topics. In order to evaluate the overall similarity and consistency of physicians' judgments, cosine-similarity and Pearson's correlation of two vectors of judgments (435) were calculated respectively. The similarity was 0.80 and the correlation coefficient was 0.59. Further analysis indicated that the inconsistency (the cause of correlation decrease) mainly occurred when physicians made judgments associated with general health topics.

4.2.2 Average Topic Association

Medical topic association is an important factor in serendipitous recommendation. For evaluating it, the normalized association strength (NAS) of one topic (i.e. Topic-X) with the other topics was calculated by using the formula 3.1 (see chapter 3). By averaging the value of NAS on 29 topics, the overall association between Topic X and all the other topics was comprehensively estimated, with the results shown in Table 4.7. As presented in the table, the five topics with strongest association include anxiety (2.03), seniors (2.02), obesity/weight loss/fitness (1.97), nutrition/diet (1.93), and women's health (1.93). In contrast, five topics with the weakest association include flu/cold/SARS (0.64), breast cancer (0.67), prostate/prostate cancer (0.69), allergy (0.72), and arthritis (0.88). Both KA and KAA method are dependent upon topic associations judged by physicians. This is because topic association can affect the presentation of topics, i.e. the NSP topic with strong association involves a high possibility of being presented to users in serendipitous recommendation. In this study, the topics such as anxiety and seniors were expected to have high frequency of presentations in MedSDFilter instances with KA and KAA methods.

Table 4.7: Overall Association between Topic-X and All Other Topics

Topic-A	**NAS Ave.**	**Topic-A**	**NAS Ave.**
Addiction	1.1	Headache / Migraine	1.29
Allergy	0.72	Heart Disease	1.69
Alternative Medicine	1.83	HIV / AIDS	0.97
Anxiety / Stress	2.03	Hypertension	1.64
Arthritis / Rheumatology	0.88	Men's Health	1.91
Asthma	0.91	Mental Health	1.59
Cardiovascular / Cardiology	1.76	Neurology / Neuroscience	1.47
Breast Cancer	0.67	Nutrition / Diet	1.93
Cholesterol	1.33	Obesity / Weight Loss / Fitness	1.97
COPD	1.21	Pregnancy / Obstetrics	1.24
Dentistry	0.9	Prostate / Prostate Cancer	0.69
Depression	1.78	Seniors / Aging	2.02
Diabetes	1.59	Sleep / Sleep Disorders	1.24
Eating Disorders	1.66	Stroke	1.47
Flu / Cold / SARS	0.64	Women's Health	1.93

4.3 Serendipitous Recommendation

In this dissertation, the first research question (RQ 1 in Chapter 1) is concerned with whether incorporating serendipity into personalized filtering system can aid discovery of unexpected but relevant news content. The analysis and discussion in this section is used to answer RQ 1. As described before, the relevancy of an article in this study is evaluated based on the relation of the article to users' interest. In personalized filtering environment formulated for the user study, whether a particular presented article is expected or not is judged by the topic of article. If the topic was not included in an interest profile, i.e., not selected by users in pre-session questionnaire, the article was classified as "unexpected." In addition, implicit and explicit approaches were adopted to capture users' interests in articles through analyzing their clicks and ratings. Article title and excerpt were shown as a way to aid users' judgments on whether to click, and the article content was presented to help users' ratings. As described in the section 3.3.2, the articles belonging to system-profile (SP) and non-system-profile (NSP) topics are SP and NSP articles respectively. Accordingly, the clicks on these two types of articles are called SP and NSP clicks. During the study, the subjects were instructed to click on the articles in which they were interested. If subjects were not interested in any article in the current session, they could enter into the next session before the current session timed out. This instruction ensured a strong correlation between subjects' clicks and their interests, which means that the clicks were interest-driven. Additionally, the subjects were asked to rate all the clicked articles in the instruction. As for the clicks on which rating data were missing, they were excluded from quantitative analysis in this study.

4.3.1 Users' NSP Clicks

User's interests were identified through analyzing their click events in previous studies (Fan et al., 2012). This method was applied to this study. NSP clicks were counted and aggregated across all participants in each session. Figure 4.7 shows the number of NSP clicks in each session.

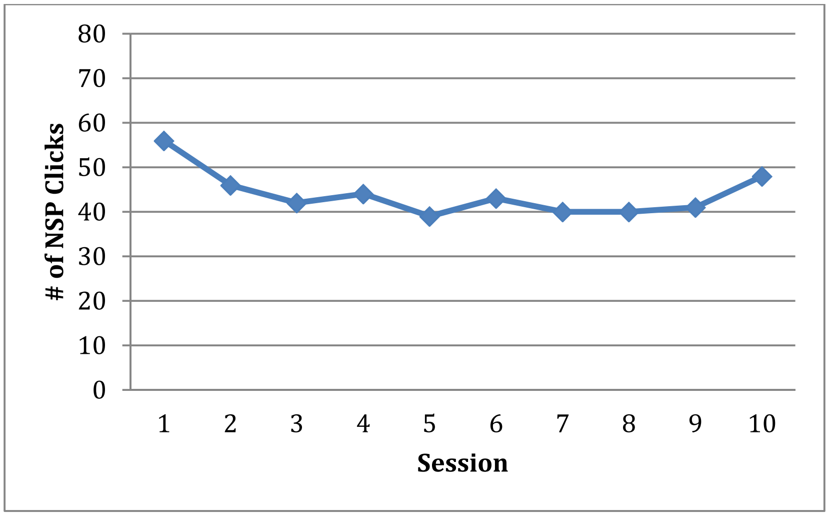


Figure 4.7: The Number of NSP Clicks in Each Session

From the figure, we can see that more than thirty NSP clicks were recorded in each session. These clicks were aggregated over all the sessions to obtain total NSP clicks in the study. It was found that thirty users executed a total of 439 NSP clicks, indicating that each user had an average of about fifteen NSP clicks.

As discussed in research methods, some retrieved articles had one main topic as well as multiple subtopics. In order to check the potential effect of this confounder, the subtopics of NSP articles were analyzed. The results indicated that, as for 17.6% of NSP articles presented to users, their subtopics included one or more SP topics. In addition, 24.1% of NSP articles clicked by users involved SP topics in their subtopics. Since there was no data about the importance of subtopics from our data source, it was hard to accurately evaluate the impact of this cofounder on users' clicks. Based on these findings, new research needs to be conducted to further investigate the effects associated with topic overlap in article content.

Since users' clicks were interest-driven in the context formulated in this study, clicks on articles implied the articles were relevant to users' interests. Therefore, the finding of NSP clicks indicated users were interested in some articles not from the topics in system' profile. An assumption is made here. If no serendipitous recommendation was implemented in MedSDFilter system, users could not have a chance to access these non-system-profile articles in which they had interest. This demonstrated that blind spots occurred in terms of content delivery in a personalized filtering environment without serendipity incorporation. In MedSDFilter, serendipitous recommendation was actually implemented and users also clicked on these non-system-profile articles. This indicated that serendipitous recommendation helped reduce blind spots by enabling users to discover articles of interest from the topics that were not included in the system's profile. In the next section, interest strength associated with these NSP clicks is analyzed, which provides additional evidence associated with blind spots.

SP clicks were also counted in each session to make a comparison with the results of NSP clicks. Here SP clicks refer to the clicks on the articles from system-profile topics. Figure 4.8 shows thirty users involved in more than sixty SP clicks in most sessions. These clicks were aggregated over all the sessions to obtain total clicks. It was found that, the users had a total of 683 SP clicks, meaning that each user involved an average of about twenty-three SP clicks.

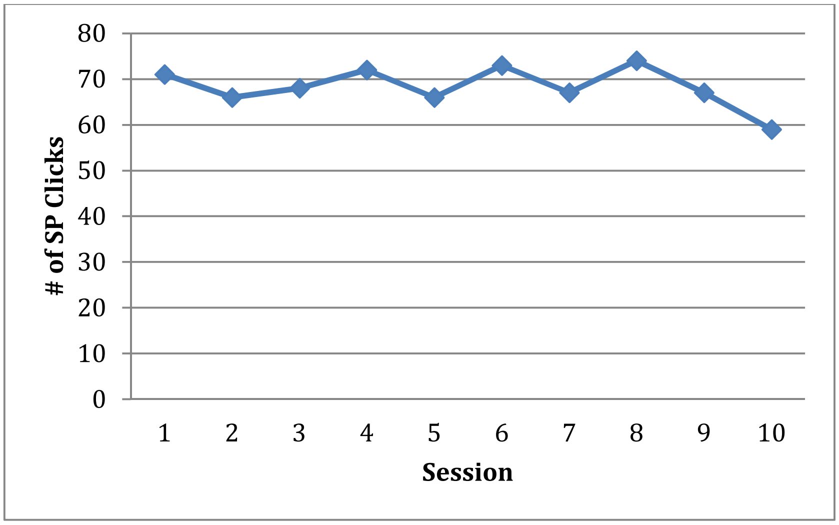


Figure 4.8: The Number of SP Clicks in Each Session

According to the experimental setup, each user was presented with the same number (i.e. 50) of NSP and SP articles in the filtering sessions. The proportion of NSP and SP articles on which users clicked was calculated and compared. The results showed that users clicked on 29.3% NSP and 45.5% SP articles that were presented to them. The higher proportion for SP articles was as expected because users indicated their interests in all SP topics in the pre-session questionnaire. Considering users excluded all NSP topics in building their profile for filtering, the proportion of clicks on NSP articles still generated good evidence for demonstrating unexpected discovery. The clicks on almost 30% NSP articles may be attributed to the adoption of KA and KAA methods for serendipitous recommendation. Twenty out of 30 users were assigned to the groups that utilized these two methods. The methods employed physicians' knowledge and users' feedback to help serendipitous recommendation. Therefore, it is possible that recommendation of relevant articles was enhanced in this situation. In addition, the SP topics were static in the study, meaning that no information about user's feedback was used to update these topics. Even if users did not like the articles on some SP topics, the topics were still presented to users in each session. During the study, the static presentation of SP topics may have reduced their advantage over NSP topics in attracting users' clicks.

4.3.2 Users' NSP Clicks with Strong Interest

Different than clicks, which implicitly show users' interests, ratings on articles explicitly indicate users' interests. In this study, users were requested to rate their interest strength by selecting a value from a Likert scale (0-5), where '5' represents 'strongly interested' and '0' represents 'not interested.' Users were considered to have a strong interest in an article if they selected a value equal or greater than 3 on the 0-5 scale to describe their interest strength. In this study, NSP and SP clicks with strong interest (SI) were called NSP-SI and SP-SI clicks respectively. Figure 4.9 shows the number of NSP-SI clicks in each session.

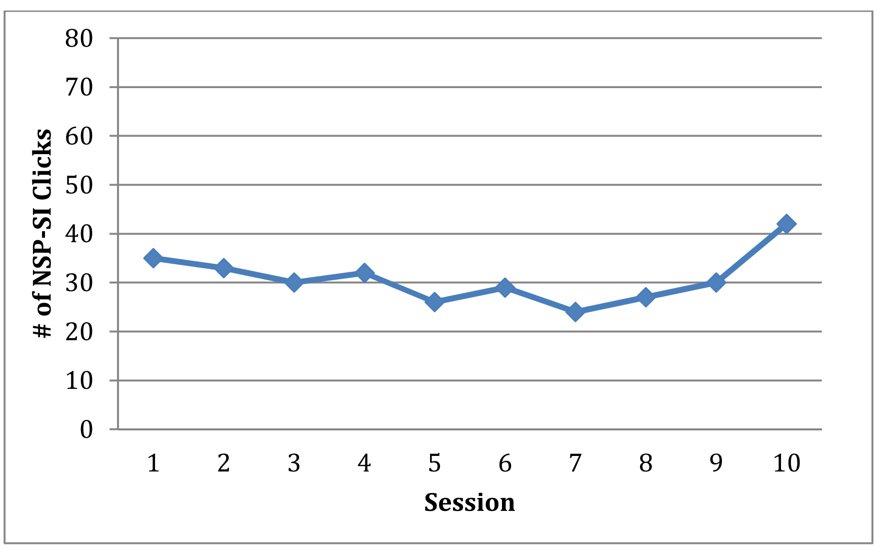


Figure 4.9: The Number of NSP-SI Clicks in Each Session

NSP-SI clicks were added up for each session, and the result shows 308 NSP-SI clicks recorded in the study for all the users. On average, each user had about 10 NSP-SI clicks in the study. SP-SI clicks were also counted in each session to make a comparison with the results of NSP-SI clicks. It was found that 584 SP-SI clicks recorded in the study for all the users. On average, about 19 SP-SI clicks were detected per user. All clicks with strong interest were calculated by summing the total number of NSP-SI and SP-SI clicks. NSP-SI clicks accounted for 34.7% of total clicks with strong interest. Therefore, a comparatively high proportion of NSP-SI clicks (35%) clearly indicated that users discovered many NSP articles that were unexpected but relevant to their interests. Additionally, the result showed that users would miss about one third of the articles of strong interest if serendipitous recommendation was not implemented in MedSDFilter system. This demonstrated the occurrence of blind spot in personalized medical content delivery as well as the potential of serendipity in reducing blind spot.

4.3.3 Summary

In this study, implicit data recorded users' clicks on the articles from non-system-profile topics (i.e. unexpected events for users). In addition, the analysis of explicit data (i.e. users' ratings) identified users' interests on these clicked articles. In this section, both data were integrated to clarify the role of serendipitous recommendation in delivering medical news content relevant to users' interests. The findings of users' NSP clicks and interests on clicked articles demonstrated that incorporating serendipity helped users discover unexpected news content relevant to their interests, which answered the first research question proposed in this dissertation.

4.4 Difference between RA, KA, and KAA Methods

In this dissertation, the second research question (RQ 2 in Chapter 1) was concerned with how the three serendipity models deliver medical news content differently and in what ways. The three serendipity models refer to the model built on RA, KA, and KAA methods. RA method relies on randomness. KA method makes use of topic associations judged by physicians. The KAA method, as adaptive KA method, utilizes both the judgment from physicians and the ratings collected from users. As was stated before, the analysis in this section is based on the users' clicks and ratings.

4.4.1 Users' NSP Clicks

Thirty participants were randomly assigned to three groups, with 10 participants in each group. Group 1, 2, and 3 involved RA, KA, and KAA methods for implementing serendipitous recommendation, respectively. The clicks on NSP articles for each group were counted and aggregated across all users in each session. On average, users in Group 1, 2, and 3 clicked on almost 11, 14, and 20 articles belonging to NSP topics per session, respectively. Further, these aggregated NSP clicks were compared with the total clicks per session of each group to calculate the proportion of NSP clicks. The metric, proportion of NSP clicks in relation to total clicks, is referred to as PNC. Figure 4.10 shows PNC in each session for the three groups.

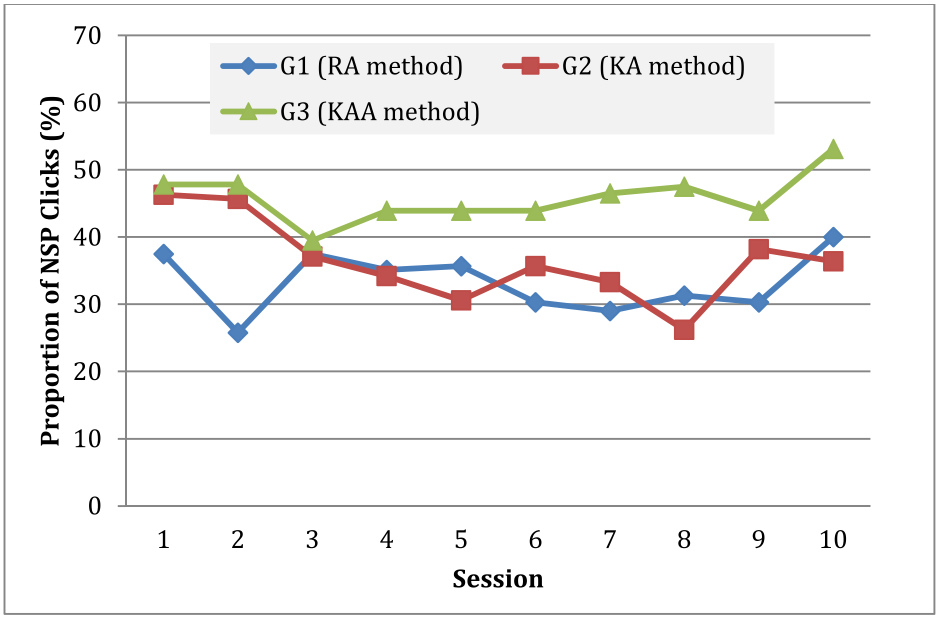


Figure 4.10: The Proportion of NSP Clicks for Different Groups

From Figure 4.10, the PNC for all sessions was higher than 25% in three groups. On average, the PNC values across all ten sessions are 33.3, 36.4, and 45.8% for Group 1, 2, and 3, respectively. This indicates that the users in Group 3 (KAA method) involved higher proportion of NSP clicks than those in Group 1 and 2. It also shows that the average PNC in Group 2 was slightly higher than that of Group 1. The one-way ANOVA test showed significant difference in PNC between three groups ($p<0.0001$). Additional analyses were conducted to determine what is driving the significant difference between the three groups. Tukey's post-hoc method was used to compute the pair-wise comparisons between groups, with the results shown in Table 4.8.

Table 4.8: Pair-wise Comparisons on Proportion of NSP Clicks for Three Groups

Group Pair	Group 1 and 2	Group 1 and 3	Group 2 and 3
P value	0.342	0.00001	0.00056

The results above show that the difference between Group 3 and 1 is significant in PNC. In addition, the difference in PNC between Group 2 and 3 is significant as well. This was as expected because KAA method involved additional information (i.e. user's feedback) to modify serendipitous recommendation, compared with RA and KA methods. In this study, users were presented with the articles from different NSP topics across the filtering sessions. Since users' interest strength varied with NSP topics, their ratings on presented NSP articles were likely to change during this period. Theoretically, KAA method excluded NSP topics of no or low interest from presentation based on users' feedback. Therefore, the impact of interest variation on serendipitous recommendation could be weakened by the adaptive learning method, implying a comparatively stable recommendation performance. The result of theoretical analysis was consistent with our observation. From Figure 4.10, we can see that the PNC value in KAA method was comparatively stable, compared with the PNC variation in KA method.

Though the PNC average across all sessions in Group 2 is higher than that in Group 1, the difference between them is not significant. This result may be partially explained through our findings. Despite high fluctuation of PNC across the sessions, the variation trend of PNC was still apparent. In the first two sessions, Group 2 and 3 have almost the same values of PNC, which is apparently higher than that of Group 1. From the third sessions on, the PNC value in Group 2 begins to decrease and approach that of Group 1. The decrease of PNC in Group 2 may be attributed to the KA method adopted. KA method shows different NSP topics below SP topics in each session in descending order based on their associations. In the study, the presented NSP topics were strongly associated to SP topics in the first few sessions. If users are interested in the SP topics, it is likely that they liked these NSP topics as well. After a few sessions, the associations between NSP and SP topics became weak, implying reduced impact of topic associations on users' clicks.

4.4.2 Users' NSP Clicks with Strong Interest

In the study, subjects were requested to rate clicked articles to indicate their interest on article content. In each group, the clicks on NSP articles with strong interest (i.e. NSP-SI clicks) were counted and aggregated across all users in each session. On average, users in Group 1, 2, and 3 clicked and indicated strong interest on approximately 6, 11, and 14 articles belonging to NSP topics per session respectively. Further, these aggregated NSP-SI clicks were compared with the total clicks with strong interest to calculate the proportion of NSP-SI clicks. We refer to this metric, proportion of NSP-SI clicks in relation to total clicks with strong interest, as PNSC. Figure 4.11 shows PNSC in each session for three groups.

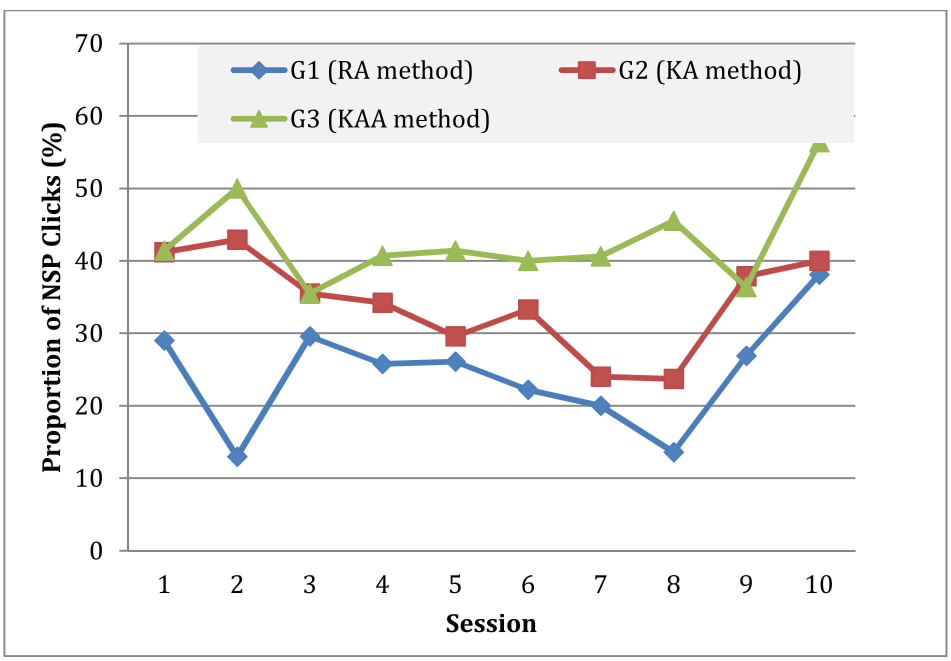


Figure 4.11: The Proportion of NSP-SI Clicks for Different Groups

The PNSC for all sessions was equal or higher than 13.0, 23.7 and 35.5% in Group 1, 2, and 3 respectively. On average, the PNSC values across all ten sessions were 24.4, 34.2, and 42.8% for Group 1, 2, and 3 respectively. These results indicated that the users in Group 3 (KAA method) involved higher proportion of NSP-SI clicks than those in Group 2, which had a higher proportion of NSP-SI clicks than those in Group 1. The one-way ANOVA test showed significant difference in PNSC between the three groups ($p<0.0001$). In order to determine what is driving the difference between the three groups, Tukey's post-hoc method was used for the pair-wise comparisons between groups.

Table 4.9: Pair-wise Comparisons on Proportion of NSP-SI Clicks for Three Groups

Group Pair	Group 1 and 2	Group 1 and 3	Group 2 and 3
P value	0.010	0.000007	0.026

Table 4.9 shows the results of the statistical analysis. Significant difference in PNSC was identified between group pairs. The data in Figure 4.10 showed that the proportion of NSP clicks (PNC) in Group 1 and 2 were close in most sessions (i.e. from the third session). The proportion of NSP-SI clicks (PNSC) in Group 2 was significantly higher than that in Group 1. This difference implied that, compared with users in Group 1, the users in Group 2 were more likely to indicate strong interest once they clicked articles. In order to verify this, the proportion of indicating strong interest (PISI) in NSP clicks was calculated through a simple method; i.e. for clicked NSP articles, indications of strong interest were counted and compared to total ratings. Based on this comparison, we found that PISI values for NSP clicks in Group 1 (RA method) and 2 (KA method) were 0.59 and 0.80 respectively. The

lower PISI value for NSP clicks in Group 1 was as expected; that is, users indicated strong interest in the clicked non-system-profile articles at a comparatively low proportion when the RA method was used. In the instructions of this study, users were asked to find and view the articles they were interested in. In this context, if users could not find articles of strong interest, they may have lowered the standard of judging and clicked on some articles in which they in fact had weak interest. The phenomenon requires further investigation.

4.4.3 Relationship between Subjects' Characteristics and NSP Clicks

In order to examine the relationship between users' age, gender and previous experience with their interface actions, a group of correlations were estimated, with the results shown in Table 4.10.

Table 4.10: Coefficients of Correlation between Users' Characteristics and NSP Clicks

	PNC	PNSC
Age	-0.203 (Pearson)	-0.209 (Pearson)
Gender (Male: 1, Female: 2)	0.127 (Pearson)	0.405 (Pearson)
Frequency of looking for news	-0.217 (Spearman)	-0.112 (Spearman)
Frequency of looking for health news	0.050 (Spearman)	0.071 (Spearman)
Level of understanding health news	0.102 (Spearman)	0.067 (Spearman)

The negative correlation coefficient of age and PNC (or PNSC) indicated that younger users were more interested in non-system-profile articles than older users. In addition, the negative coefficient of gender and PNC (or PNSC) showed that female users were more interested in NSP articles than the male. Positive correlations were also found between NSP clicks and the frequency of looking for health news as well as the level of understanding health news. In contrast, there existed a negative correlation between NSP clicks and the frequency of looking for news. Admittedly, the findings above may be biased due to the evaluation of correlations based on the data associated with a small number of users in this study. New research involving larger sample size should be conducted before reaching clear conclusions regarding the relationship between users' characteristics and NSP clicks.

4.4.4 Summary

This study attempted to clarify the efficacy of three serendipity models utilized for recommendation. To achieve this goal, users' clicks and ratings were analyzed. It was found that KAA method generated significantly higher proportion of clicks on NSP articles than both RA and KA methods. The value of serendipitous recommendation is directly reflected on rating provided by users, meaning that the evidence from ratings is more direct and thus persuasive than the evidence associated with clicks. Therefore, the proportion of NSP-SI clicks (PNSC) was used as the main indicator when comparing different methods of incorporating serendipity. By comparing the average PNSC in the three groups, we can conclude that, in terms of delivering relevant articles in serendipitous recommendation, KAA method outperformed than KA method and KA method did better than RA method. The results and analysis of users' clicks and ratings in this section answered the RQ 2, the second research question of this study.

4.5 Users' Response to Unexpected Discovery

The third research question (RQ 3 in Chapter 1) is concerned with determining users' response to unexpected discovery of relevant medical news. The analysis in this section is used to answer the research question. Users' response was evaluated based on quantitative and qualitative data collected from the experiments. The quantitative evaluation focused on the strength of users' interests in unexpected articles on which they clicked during the filtering sessions. The qualitative analysis attempted to identify users' opinions associated with unexpected discovery from the comments they provided in the post-session questionnaires.

4.5.1 Users' Interest in Clicked NSP Articles

As described before, an article is considered "unexpected" if its main topic is not part of the system's profile. In other words, the unexpected articles are actually the articles from non-system-profile (NSP) topics, i.e. NSP articles. In this study, users' preference associated with unexpected articles was evaluated by analyzing their interests in the NSP articles. Figure 4.12 shows the distribution of users' interest strength in the NSP articles they clicked on.

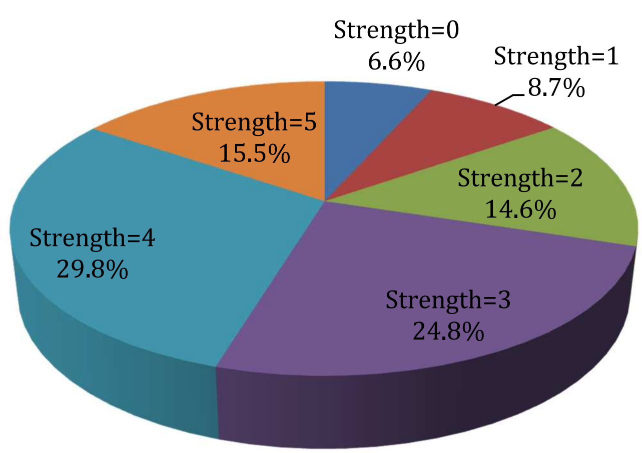


Figure 4.12: Distribution of Interest Strength in Clicked NSP Articles

As shown in the figure, users indicated strong interest (strength>=3 on a scale of 0-5) in 70.1% of total rating cases. In addition, at 15.5% of total rating cases, users selected the highest value of 5 to describe their interest strength. In contrast, users indicated no interest in 6.6% of totally rating cases. The results above indicated that users generally demonstrated strong interests in their unexpected discovery of articles. The ratings of interest strength on all clicked NSP articles were aggregated across all participants in each group. The one-way ANOVA test showed there was significant difference in the ratings between the three groups ($p=0.00043$). In order to evaluate differences of group pairs, Tukey's method was adopted for multiple comparisons. The results showed that the difference between Group 1 and 2, as compared to Group 1 and 3, was significant ($p=0.0003$ and $p=0.0095$ respectively), while the difference between Group 2 and 3 was not significant ($p=0.36$).

Users' interest strength could have been influenced by several factors associated with delivered articles, such as the quality of article content. In order to clarify the findings associated with unexpected discovery, user's interest strength in unexpected articles was compared with that in expected articles. In personalized filtering environment, expected articles are from system-profile (SP) topics and they are actually SP articles. Figure 4.13 shows the results of statistical analysis of users' interest strength in clicked SP articles.

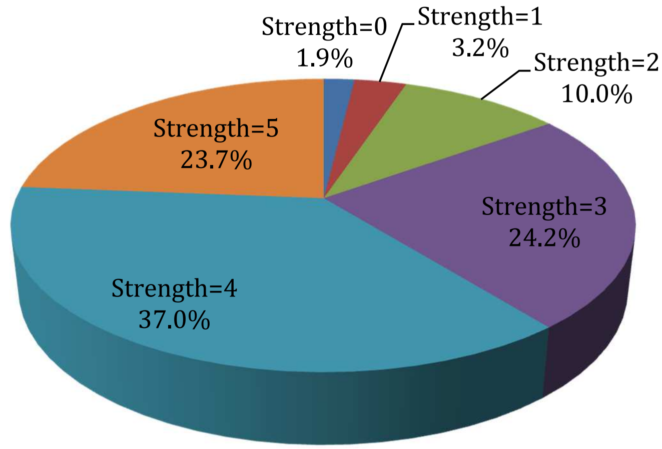


Figure 4.13: Distribution of Interest Strength in Clicked SP Articles

For 1.9 and 84.9% of total rating cases, users indicated no and strong interest respectively. As this finding indicated, users were likely to select higher interest strength when they clicked on SP articles as compared to NSP articles. The higher interest strength in clicked SP articles was expected, considering that users indicated strong interest in most of SP topics in pre-session questionnaires. Though users did not select NSP topics to build the system's profile, their interest in clicked NSP articles remained comparatively strong. In order to further examine users' actual interest in the articles associated with unexpected discovery, users' comments in post-session questionnaires were also analyzed. Reflection on the analysis is presented in the next subsection.

4.5.2 Users' Opinions on Unexpected Discovery

Upon execution of the last filtering session, users were requested to complete a questionnaire to investigate their opinion associated with unexpected discovery. The results from each of the questions in the questionnaire (see Part 2 in Appendix D) are presented here, segregated into experimental groups. In order to avoid confusion in the analysis below, we first clarify the definition of "unexpected." Whether an article is unexpected is evaluated by comparing the topic of the article to the interest topics selected by users in the pre-session survey (i.e. system's profile). However, users may think some articles from non-system-profile topics are not "unexpected" in practice. This can be attributed to the difference between presentations of articles and topics. When articles are presented, the terms associated with the interest profile are not highlighted. As a result, users may not know the existence of the core topics, particularly when the article covers diverse topics.

When asked whether the system ever presented unexpected news items (Question 2.4 in Appendix D), all the users in Group 1, 2, and 3 responded that they found some articles to be unexpected. Users were then requested to provide the associated keywords as examples. The keywords provided by many users were not exactly the terms of medical topics (see Appendix A). For instance, one user gave the term "Alzheimer's" in her answer. It was likely that unexpected articles she noticed were from the topic "Seniors / Aging," which made the user think that the main topic of the article was different from the core topics in the interest profile. Furthermore, users were required to answer two extra questions regarding unexpected discovery. The questions were concerned with whether users click and like unexpected articles. Users were asked to explain their answers in detail.

(a) Clicks on unexpected articles

The majority of users in Group 1 (RA method) stated that they clicked on unexpected articles. Regarding the reasons for clicking, one user commented that he "just want[s] to check it out what it is." Two users stated, "the title seemed interesting" and "the headline caught my attention." Additionally, some users explained they clicked on some unexpected articles because the articles were related to their personal or family health. The responses in

Group 2 (KA method) and Group 3 (KAA method) were somewhat similar with those in Group 1. From the collected answers, all but one user from Groups 2 and 3 mentioned that they clicked on unexpected items. The reasons of clicking were described as “the title looks informative” and “the article on the topic is related to my family health.”

The number of clicks on unexpected articles varied with users. One user in Group 2 noted, “one or two” articles while another one in the same group stated, “I clicked on at least 3-4 articles per session.” Additionally, users did not click on all the unexpected articles. One user in Group 1 stated that she clicked on 50% of unexpected articles and further explained her standard for deciding to click. She stated, “If the title seemed interesting I would read the article, otherwise I skipped it.” Other users did not indicate what proportion of the unexpected articles they clicked on. They commented that they clicked on unexpected articles if the articles were interesting. In summary, the majority of users in the three groups indicated they clicked on some unexpected articles and the reasons behind their clicks included attractive article title/headlines and relation to personal or family health. There was no apparent difference in users’ responses among the three groups. The motivation for users to click on the unexpected articles is that the articles looked interesting to them. This is consistent with the analysis in section 4.3, meaning that the users’ clicks on NSP articles are interest-driven.

(b) Preference on clicked articles

Regarding the preference for unexpected articles on which users clicked, almost all the users in Group 2 (KA method) and Group 3 (KAA method) indicated that they liked at least some of these articles. One user stated, “the surprise ones were actually some of my favorites.” These users voiced positive opinions, stating they liked the articles because they learned something they had not previously known. For instance, they acquired new knowledge in other areas by reading these unexpected articles. Their comments were: "presented new information on a general level" or "presented new ideas" useful to one's personal or family health. Additionally, one user liked unexpected articles simply because of curiosity. Users in Group 1 (RA method) had more negative responses compared with those in Groups 2 and 3. In the RA group, one user reported no clicks on unexpected articles, implying no or little preference on them. Therefore, the responses of nine users were collected and analyzed. Six of them indicated that they favored the unexpected articles on which they clicked. The other three users noted that they did not like some unexpected articles they had read. One user stated the article content was not as interesting as he thought. Another one stated he was not interested in research articles and wanted "the more clinically oriented articles." The third user did not like the articles related to studies funded by commercial companies.

So far, the analysis has been mainly concerned with the unexpected articles that were clicked by users. Users' satisfaction on unexpected discovery also involved their opinions on unexpected articles they did not click on during the experimental sessions. In the next subsection, a comprehensive evaluation is conducted based on evidence from the other parts of the post-session questionnaires.

4.5.3 Users' Satisfaction on Unexpected Discovery

In this study, the main difference in system's performances among three groups was that unexpected articles were presented to users in different manners. This difference implied that the difference in users' satisfaction on system performance was likely to be related to their attitude to unexpected discovery. Therefore, users' satisfaction regarding system's performance was analyzed. In post-session questionnaires, users' overall satisfaction on the system's performance was collected. Different satisfaction levels were summarized in each group, with the results shown in Figure 4.14.

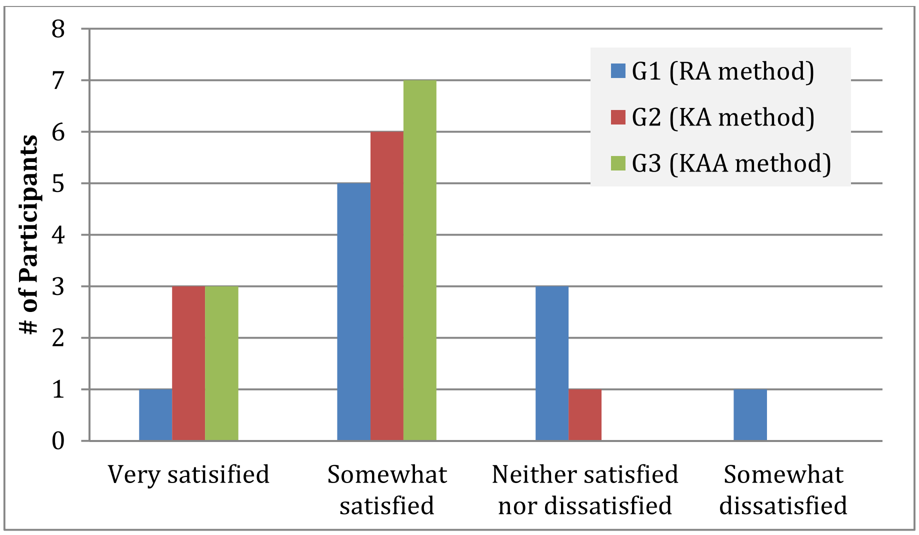


Figure 4.14: Distribution of Users' Satisfaction on System's Performance

From the distribution of satisfaction levels, we can see that all users in Group 3 (KAA method) were satisfied with the performance of the system, and three users were very satisfied. Similarly, the majority (9 of 10) of users in Group 2 (KA method) indicated satisfaction with the performance of system, with one user declaring "neither satisfied nor dissatisfied." In contrast, satisfaction on the system's performance was mixed for the Group 1 users (RA method). Six of 10 users in the group indicated they were satisfied while the other users did not note satisfaction. One user even stated, "somewhat dissatisfied." As for the users who were very satisfied with the system's performance, Group 1 has just one, compared with three in each of other groups. It is concluded that in terms of the system's performance, Group 3 involves slightly higher satisfaction than Group 2, and both groups have apparently higher satisfaction than Group 1. The results are consistent with the findings regarding users' NSP clicks with strong interest as shown in Figure 4.11. Across Group 1, 2, and 3, the proportion of NSP-SI clicks increased gradually.

Users' satisfaction can be affected by many factors, and unexpected discovery is one of these factors. Users' ratings of satisfaction on system's performance were compared to their ratings of interest strength on NSP articles during filtering sessions. As described in the section 4.4.2, users' ratings on NSP articles in Group 2 or 3 was significantly higher than that in Group 1 while the difference between Group 2 and 3 was not significant. The results were consistent with users' ratings of satisfaction on system's performance, i.e. users in Group 2 and 3 involved similar ratings of satisfaction and they were more satisfied with system's performance than users in Group 1.

In order to evaluate the impact of unexpected discovery on users' opinions on system's performance, users' comments associated with a few post-session questions (see 2.1-2.3 in Appendix D) were analyzed.

When asked whether to use a system like this, 4 of 10 of users in Group 1 clearly indicated that they would not like to use the system to access health information. All of these users indicated "neither satisfied nor dissatisfied" or "somewhat dissatisfied" with the system's performance. The main reason they provided is that some of the recommended articles were not relevant to their interests. This statement implied that users' dissatisfaction with the system's performance was strongly correlated with the occurrence of irrelevant articles. Some users indicated that they would select not to use the system in some situations. For instance, one user stated that he would not use it if he "was researching on specific topic" because he would be "distracted by the unrelated articles." In contrast, users in Group 2 and 3 gave much more positive opinions on whether to use this type of system. Eight users in Group 2 indicated that they would use the system. Another two users in the group stated they were not sure about whether to use it because they wanted to use a system that was capable of seeking health information on a specific topic. All the users in Group 3 indicated that they would use the system (or use it conditionally). One user stated he was interested in a system like this because the stories delivered seemed relevant and interesting.

When asked about likes and dislikes on the system and suggestions to improve the system, one user in Group 1 mentioned that he liked "a nice variety" in article recommendation. In contrast, five users in the same group considered that some of the presented articles were not relevant and stated that the system should be improved through recommending more relevant articles to the user. In contrast, there were far fewer complaints about irrelevant articles in Group 2 and 3. By analyzing users' comments, we found that the majority in these two groups did not mention the occurrence of irrelevant articles. Instead, these users emphasized that articles recommended to them were relevant and that they liked the articles. For instance, one user stated that the system "seemed to successfully predict articles that I found interesting." Another user noted that, besides very relevant articles, the other articles were still "catchy" and on target.

4.5.4 Summary

Through comprehensive analysis of users' interface actions (clicks and ratings) as well as related comments in post-session questionnaires, it can be concluded that most users in the three groups had a positive response to the unexpected articles they clicked. In terms of unexpected discovery, the opinions of users in Group 2 and 3 were more positive than those in Group 1. Additionally, a fairly high portion of users in Group 1 indicated negative opinions on unexpected discovery, with the main reason being that some of the unexpected articles were not relevant to their interests.

4.6 Impact of Unexpected Discovery on Users' Interest

The fourth research question (RQ 4 in Chapter 1) is concerned with whether users' interest change is caused by unexpected discovery. In this section, we mainly focused on the change of users' interest, which was correlated with unexpected discovery in the filtering sessions. Through the analysis of quantitative data (interest profile change) and qualitative data (factors of interest change), we attempted to clarify the impact of unexpected discovery on user's interest.

4.6.1 Quantitative Analysis of Newly-Added Topics

The quantitative analysis was concerned with newly-added profile (NAP) topics. These topics were identified by comparing interest profiles acquired from questionnaires filled out before and after the filtering sessions. One thing to clarify is that quantitative analysis in this subsection was not targeted on all NAP topics because some NAP topics were not related to users' interest change. Here, an example is provided to explain the possible cause. For instance, a user was interested in 12 medical topics. MedSDFilter system only allowed him or her to select at most 10 topics for filtering purpose. In this situation, two interest topics were not included in pre-session profile. During the sessions, the user did not have interest change in these two topics. Post-session questionnaires did not set a limitation in the number of profile topics because the profile was not utilized for filtering. Therefore, the user added these two topics into his or her post-session profile. In this specific scenario, the two topics, as newly-added profile topics, did not get involved in the user's interest change.

NAP topics with definitive interest change are the newly-added profile topics of concern in this study (named NAP-C topics). In this study, we detected NAP-C topics through a simple method. It could be assumed that users' interest strength in NAP topics in pre-session stage was equal to or smaller than the minimum interest strength of pre-session profile topics (MISPPT) because they did not select NAP topics for pre-session profile. Therefore, NAP-C could be detected through simply comparing interest strength in post-session profile with MISPPT. As for NAP topics, users' interest change was definitive during the filtering sessions if post-session interest strength of them were higher than MISPPT. This is because MISPPT was higher than pre-session interest strength of NAP topics.

It was found that the average number of NAP-C topics per user in Group 1, 2, and 3 are 2.5, 3.2, and 3.7, respectively. One-way ANOVA test showed no significant difference ($p>0.5$) in the number of NAP-C topics between three groups because of high variance between users in each group. The number of NAP-C topics involved a similar pattern with the proportion of NSP clicks with strong interest in each group (see Figure 4.12), i.e. Group 3 involved a higher value than Group 2, and both groups had a higher value than Group 1. This indicated a possibility of association between users' preference on NSP article content and newly-added topics in profile.

In order to check the potential relationship between NAP-C topics and users' unexpected discovery in the filtering sessions, the clicks and ratings on the articles from

NAP-C topics were analyzed. It was found that each user clicked about 5 articles from NAP-C topics on average in this study and 75% of ratings on these articles were positive, implying users' interest change in NAP-C topics was associated with their positive opinions on the articles from NAP-C topics. In addition, the relationship between NAP topics and users' NSP clicks was examined. It was found that approximately 55% NSP clicks occurred on the articles from NAP topics. Besides, about 60% NSP clicks with strong interest were executed on the articles from NAP topics. These results indicated an association between NSP articles on which users clicked in the filtering sessions and new topics in which users were interested after the filtering sessions. In order to further clarify the factors associated with users' interest change, qualitative analysis of users' response in post-session questionnaires was conducted, with the detail discussed in the next subsection.

4.6.2 Qualitative Analysis of Users' Interest Change

If newly-added topics were found in the post-session profile, as compared to the pre-session profile, users were asked to provide their opinions on what might have caused the change. In Group 1, one user did not think that there was a change in her interest profile, and other users showed various reasons behind their interest change. Two users noted that they selected more health topics because these topics were related to their personal health condition or family health. One user stated that the reason for interest change was, "I like to read about a lot of different health issues." Additionally, six users mentioned that their interest change was related to the articles they read in the sessions. The response from one user is representative. She stated, "after reading the articles, I found that I was more interested in some things than I thought." Similarly, other users mentioned the causes of interest change such as “the medical topics mentioned in article content” and “what I read.” From some articles, she learnt about the connection between what she deemed as important health issues to those things she worried less about, implying she became interested in more topics due to new discoveries.

The greatest diversity of responses was found in Group 2. From the collected answers, 7 of 10 users in Group 2 indicated that their interest change was related to the articles presented to them. One user stated that she found some interesting articles outside of her chosen topics and further explained that she might be more interested in those topics than she originally considered. Another user noted that the pre-session profile was consisted of major interest topics and the post-session profile included extra topics from interesting articles she read during the sessions. In addition, two users mentioned that the limited number of interest topics in pre-session profile caused the difference between pre-session and post-session profiles. One of them stated, “In the pre-session I was asked to choose only 10 topics, and in the post-session I could rate all of the topics.” In the comment section, the user also indicated that her interest was affected by the articles she read in the sessions. Unlike most users who focused on article content, one user responded with comments on article titles. She mentioned that the cause of her interest change was related to article title, which revealed her interests more than just keywords of the listed topics.

In comparison to Group 1 and 2, one dominant theme emerged from Group 3. All the users in the group stated that their interest change was related to the articles they read during the sessions. One user considered that her interest change was related to unexpected discovery of relevant articles by stating, “I have seen some articles which were not in my interest areas. Their titles were attractive. After I clicked and read the content of these articles, I found their topics were interesting." This statement is consistent with the comments of one user who changed his interest profile after reading the articles out of his interest area. He noted, "the system seemed to add in articles that I was not asking for or really expecting, but after reading a few of the articles those choices became more salient to me and I began reading more of those." When users mentioned their interest change, they mostly discussed their increased interest strength in some topics. In contrast, one user mainly described how her interest strength was decreased in her comment. Based on her description, she had no idea about what to get when she selected interests in the pre-session questionnaires. After she selected interest topics "heart disease" and "women's health," she found the content from these two categories were not exactly what she wanted, so she indicated weak interest in these two topics in post-session questionnaires.

4.6.3 Summary

In summary, the results of newly-added profile (NAP) topics demonstrated that users' interests were expanded after the filtering sessions. Based on their comments, interest changes were potentially caused by the unexpected articles they read, which answered the fourth research question (RQ 4) in this dissertation. The majority of users indicated that they became interested in more topics because some presented articles out of their selected topics were interesting for reading. Considering these articles were a product of serendipitous recommendation, it can be inferred that incorporating serendipity had a strong correlation with an increase in the number of indicated interest topics by users. The existence of newly-added profile topics in the study implied the positive impacts of system-imposed serendipity on the expansion of users' interest.

5. CONCLUSIONS AND DISCUSSION

5.1 Conclusions

This research attempted to clarify the impact of serendipity on filtering consumer medical information and how to incorporate serendipity in an effective manner. In addition, the study strived to clarify user attitudes associated with unexpected discovery of medical content in filtering settings as well as users' interest change during this process. To achieve these goals, a medical news filtering system named MedSDFilter was developed. Serendipitous recommendation was integrated into personalized filtering environment based on one of the three serendipity models (random, knowledge-based, and adaptive knowledge-based).

The first significant finding of this study is that incorporating serendipity into personalized filtering systems can help users discover unexpected but relevant medical news articles. Each user executed a comparatively high number of clicks on articles from non-system-profile topics, and users indicated strong interest in these clicked articles in most cases. We observed occurrence of blind spots in personalization systems. Importantly, it was demonstrated that serendipity helped reduce blind spots by enabling users to discover relevant articles from the topics that were not included in the system's profile.

The second significant finding is that using physician knowledge effectively enhanced serendipitous recommendation, which was further improved after an adaptive learning algorithm was adopted. The performance of serendipitous recommendation was evaluated through clicks and ratings on non-system-profile articles. The results demonstrated statistically significant increase in clicks associated with strong interest after utilizing supportive data in serendipitous recommendation. Two factors were particularly important in attracting clicks with strong interest: incorporation of physicians' knowledge and leveraging users' feedback.

This study also clarified users' attitudes associated with unexpected discovery. The opinions of users in the groups with knowledge-based and adaptive knowledge-based methods were more positive than those in the group with random-based method. A fairly high portion of users in the group with random-based method showed negative opinions on unexpected discovery. They stated that many unexpected articles were not relevant to their interests. In contrast, users in the other two groups were more concerned with the positive impact of unexpected discovery, i.e. the articles they discovered were interesting to them.

Finally, the study demonstrated that users' interest changed during filtering sessions. A distinctive dimension of the interest change was that users' interest expanded to incorporate additional topics. After reading some medical articles in topics they originally did not select, users discovered that they were actually interested in these topics. These unexpected articles were recommended through serendipity incorporation models, implying that incorporating serendipity had a marked impact on selections of additional interest topics by users.

Overall, the results of this study provided the system developers with a few features which may help avoid potential drawbacks related to over-personalization in information filtering, i.e. blind spot. This study enhanced the understanding of users' behavior and attitude regarding the consumption of medical information and generated new guidelines, which take into consideration constraints, capabilities, domain knowledge and human factors in developing information systems in the medical area.

5.2 Implications

The findings in this study compelled the researcher to explore a few critical issues for advancing the area of personalization. Some of the main implications are discussed below.

The first one is the meaning of “unexpected.” In quantitative analysis, whether the discovery of an article is unexpected to one user is determined by the relationship between the main topic of the article and interest topics indicated by the user in pre-session questionnaires. If a topic is not included in the interest topics of the user, the article is considered “unexpected.” However, in filtering sessions, users may not think these articles are unexpected to them. Several factors may contribute to this result; for example, the article may contain diverse information.

The second important issue is how to evaluate medical topic association effectively. As discussed in the research method chapter, people think about the association of medical topic pair from different perspectives, which implies that the judgments of topic association are always subjective. Also, a large number of such judgments on topic associations are not manageable in practice. For example, the evaluation of pair wise associations of one hundred topics requires at least four thousand judgments. It is highly challenging to make such large number of judgments in a consistently accurate manner, particularly if judges are operating under time limitations. In this study, topic association judgments were used to enhance serendipitous recommendation. Typically, it may then imply that physicians should be involved in evaluating recommendation performance. However, the filtering system was designed as a consumer-oriented information system and therefore it was important to conduct its evaluation based on end-users. Since the purpose of recommendation is to meet users' personal information need, the evaluation of topic association from users' perspective may be more useful than that from a physician's perspective.

In addition to physicians' judgment, medical topic association can be evaluated by analyzing users' click histories. There exists a large amount of log data in various health information sites. The idea of how to analyze the data is inspired by the "item-to-item collaborative filtering" (Linden, 2003) method implemented in Amazon. In this method, the correlation between different items (i.e. products) is identified through analyzing co-occurrence of these items in users' purchasing histories. Utilizing similar strategies, we may obtain the correlation of topics by analyzing their co-occurrence frequency. For example, if users always click the articles from topic A and B in the same browsing sessions, it is likely that these two topics have inherent connections. Based on this assumption, topic correlation (or association) of medical topic pairs can be calculated for later use in serendipitous recommendation. In addition, the data of medical topic association can be obtained by natural language processing techniques. For instance, we can retrieve the description text of medical topics from various professional resources such as MedlinePlus. After keyword recognition and weighing, the text can be represented with a vector of terms from medical vocabulary. The calculation of the vector similarity will generate the topic association. The MetaMap tools developed by National Library of Medicine (NLM) can be used to map medical text to medical vocabulary in this process. From the analysis above, we can see that there are some available methods for topic association evaluation. Despite their potential use, these methods need to be examined in terms of their practical efficacy before being considered for future research.

5.3 Comparison with Previous Studies

The researcher expects to clarify how serendipitous recommendation affects personalization and effectiveness of medical content delivery, by comparing the results of this study with the results from previous experiments (Fan et al, 2012). Although it is hard to make a valid comparison between studies since audience, collections, presentation, and methods of serendipity incorporation are different, it is still possible to have an improved understanding of the influence of serendipity on personalized medical content delivery "qualitatively." The previous study showed the proportion of off-topic clicks for all sessions was fairly high in three groups, which indicated that the participants showed interests in off-topic classes. Similar results were also found in this study. On average, the proportion of non-system-profile clicks in all sessions was higher than 0.3, even in the group with random-based method. Regarding off-topic clicks, user ratings were compared between two studies: the previous study showed that users were satisfied with 67.5% of news items related to off-topic clicks; the result from the previous study is consistent with the findings of this study in which users indicated strong interest at 70.1% of total rating cases for non-system-profile clicks. The results from both studies demonstrated that most users have a positive response to the unexpected articles they clicked on.

As for between-group difference, two user experiments involved distinct results. Statistical analysis in the previous study indicated no significant difference existed in either users' off-topic clicks or users' satisfaction between the three groups with different serendipity levels. In this study, we identified significant difference among the three groups in terms of non-system-profile clicks and ratings. The difference in statistical analysis between this study and the previous work is likely to be related to the methods used for serendipitous recommendation. In the previous study, different variants of random-based method were adopted for all three groups. The difference observed in serendipity levels in the previous study among the three groups was small because only one article was from serendipitous recommendation in presentation. In this study, we employed physicians' knowledge and users' ratings to generate serendipitous recommendation. We also increased the number of positions for serendipitous recommendation to five. These changes probably magnified the difference among the groups.

5.4 Individual Difference

The individual difference is one of the important features in this study. This is expected because the subjects in this study were only limited to staff members currently employed at UNC Chapel Hill.

As for users' background and knowledge profiles, based on the answers provided by participants in the pre-session questionnaires, we found them to be diverse, indicating the existence of substantive individual differences. The apparent difference (see Table 4.3) occurred when subjects were asked about their experience with medical news seeking and consumption. Users' understanding levels of medical content is distinct. This may be partially a result of the fact that the participants have different levels of medical knowledge. For instance, some participants are from UNC hospital system; thus they have very good understanding of article contents in medical domain. Other participants did not have occupations in the medical field and thus may have faced the difficulties in understanding medical article contents due to a large number of different terminologies and jargon.

Individual difference was further demonstrated in quantitative and qualitative analysis of user interface action and post-session questionnaire. In this study, the number of participants' screen clicks varied widely, even if they were assigned to the same method group. For instance, one participant from Group 1 clicked on more than five articles each session on average. In contrast, another participant from the same group only clicked on about two articles each session on average. It was observed that one type of participants clicked more articles in the first few sessions and clicked less in following sessions. Another type of participants clicked less at the early stage, while a group of users had comparatively stable number of clicks across all the sessions.

Additionally, the responses to unexpected discovery are different among the users in the same group. One user in Group 1 mentioned that he liked "a nice variety" in article recommendation and five users in the same group indicated that they did not like irrelevant articles presented to them. Finally, the pattern of interest change across users, even in the same group, was not consistent. One participant indicated ten new interest topics in the post-session survey. In contrast, one user, from the same group, indicated no interest change. In addition, most users indicated general satisfaction with the system and the pattern of rating varied across users.

In summary, all the findings above suggest that users' medical information seeking experience, habits, and behavior are very personal, and thus individual differences have to be carefully handled in order to develop robust serendipitous recommendation system.

5.5 Limitations

The major limitation lies in the diversity of users' interest types, which may result in the failure of serendipity models introduced in this study. For instance, the internal connections of users' interest topics can pose an important impact on their information seeking behavior as well as the performance of serendipitous recommendation. In order to explain this easily, let us assume there are two users with different types of interest. Interest profiles of the first user consist of strongly related and coherent topics such as "headache" and "sleep disorder," while a second user has disparate interest dimensions such as "headache" and "breast cancer". Both users indicate that they are interested in the topic "headache." From a physician's perspective, "headache" may have strong link to "sleep disorder" and no apparent links to "breast cancer." If the physician-association method is followed, "sleep disorder" will be granted a priority in presentation to both users. Considering the actual interest profiles of the two users, this recommendation is good for the first user, but not for the second user. From this simple logical analysis, we can see that association-based method may not always lead to useful serendipitous discovery.

Another limitation exists for the filtering environment utilized in this dissertation research. In most personalized filtering systems, users typically select only a subset of the available topics to build their profile for filtering purpose. Such a selection ensures that only part of topics can be tracked to detect interest change. Accurate identification of interest strength change related to unselected topics therefore become impossible.

In addition, clicks and ratings at the article level were utilized as two metrics for measuring users' preferences associated with medical topics. However, when medical news articles contain mixed topics, their link to such articles is weak. Finally, users' ratings on articles may be influenced by several confounding factors such as the quality of article content.

5.6 Suggestions for Future Work

In this study, the size of research sample was not large. To generalize to the whole research population, future research is required involving a larger size of subject pool. Additionally, the participants in this study have a variety of education background, which may have implication for different levels of medical knowledge possessed. Further, we know that age is an influential factor for medical information seeking (Turk-Charles et al., 1997). Subjects in this study had a wide age range. The future research should strive for less diversity in subjects' education background and age.

This post-session questionnaire collected a large amount of data about participants' responses in terms of what they liked and disliked in the MedSDFilter system. These data were not directly related to the research questions and thus were not exhaustively analyzed in this study. However, the data could be useful toward design of a new system for a future study. For example, some participants suggested clustering the articles in presentation to help them view similar articles easily. Some participants suggested that the system should filter out excessively technical articles before they are presented. Some participants wanted to see more images in the presentation. Moreover, many participants provided positive feedback regarding the way in which article titles and short excerpts were presented together. This feature should probably be maintained. These are good design guidelines requiring serious consideration.

The evaluation of medical topic association based on human opinion is a complex task, which may vary due to different perspectives about the association among medical concepts. In order to obtain reliable association data, more doctors of medicine (MD) should be invited to judge the associations among medical topics. It should be also noted that this research did not investigate physicians' thinking process as well as the factors associated with making judgments.

In conclusion, this research attempted to achieve an improved understanding of the influence of serendipity on consumer oriented medical information delivery. The results of this study provide practical directives for system designers to reduce blind spot in personalized content delivery by enhancing the potential for serendipitous discovery.

APPENDIX A. MEDICAL TOPICS

1. Addiction
2. Allergy
3. Alternative Medicine
4. Anxiety / Stress
5. Arthritis / Rheumatology
6. Asthma
7. Cardiovascular / Cardiology
8. Breast Cancer
9. Cholesterol
10. Chronic Obstructive Pulmonary Disease (COPD)
11. Dentistry
12. Depression
13. Diabetes
14. Eating Disorders
15. Flu / Cold / SARS
16. Headache / Migraine
17. Heart Disease
18. HIV / AIDS
19. Hypertension
20. Men’s Health
21. Mental Health
22. Neurology / Neuroscience
23. Nutrition / Diet
24. Obesity / Weight Loss / Fitness
25. Pregnancy / Obstetrics
26. Prostate / Prostate Cancer
27. Seniors / Aging
28. Sleep / Sleep Disorders
29. Stroke
30. Women’s Health

APPENDIX B. MEDICAL TOPIC ASSOCIATIONS

Here PID is created on the IDs of topic pair for each judgment and *r* is the normalized association calculated through the formula 3(1). Topic ID refers to Appendix A.

PID	*r*	PID	*r*	PID	*r*	PID	*r*	PID	*r*
1-2	0.00	4-8	0.50	7-23	1.00	11-28	0.83	17-22	0.50
1-3	0.33	4-9	0.17	7-24	1.00	11-29	0.00	17-23	1.00
1-4	1.00	4-10	0.50	7-25	0.33	11-30	0.67	17-24	1.00
1-5	0.33	4-11	0.50	7-26	0.00	12-13	0.33	17-25	0.33
1-6	0.00	4-12	1.00	7-27	1.00	12-14	1.00	17-26	0.17
1-7	0.17	4-13	0.33	7-28	0.50	12-15	0.17	17-27	0.83
1-8	0.17	4-14	0.83	7-29	1.00	12-16	0.83	17-28	0.50
1-9	0.00	4-15	0.00	7-30	0.83	12-17	0.33	17-29	0.83
1-10	0.17	4-16	0.83	8-9	0.00	12-18	0.33	17-30	0.83
1-11	0.17	4-17	0.83	8-10	0.17	12-19	0.33	18-19	0.17
1-12	1.00	4-18	0.67	8-11	0.00	12-20	0.67	18-20	0.83
1-13	0.00	4-19	0.83	8-12	0.67	12-21	1.00	18-21	0.67
1-14	1.00	4-20	0.83	8-13	0.17	12-22	0.83	18-22	0.33
1-15	0.00	4-21	1.00	8-14	0.00	12-23	0.50	18-23	0.33
1-16	0.67	4-22	1.00	8-15	0.00	12-24	0.83	18-24	0.33
1-17	0.17	4-23	0.67	8-16	0.00	12-25	0.83	18-25	0.50
1-18	0.50	4-24	1.00	8-17	0.17	12-26	0.83	18-26	0.17
1-19	0.00	4-25	0.67	8-18	0.17	12-27	0.83	18-27	0.33
1-20	0.50	4-26	0.67	8-19	0.17	12-28	0.83	18-28	0.17
1-21	1.00	4-27	0.83	8-20	0.17	12-29	0.83	18-29	0.33
1-22	0.83	4-28	0.67	8-21	0.67	12-30	0.67	18-30	0.67
1-23	0.17	4-29	0.50	8-22	0.17	13-14	0.83	19-20	0.83
1-24	0.50	4-30	0.83	8-23	0.50	13-15	0.33	19-21	0.33
1-25	0.33	5-6	0.17	8-24	0.33	13-16	0.33	19-22	0.67
1-26	0.00	5-7	0.17	8-25	0.17	13-17	1.00	19-23	0.83
1-27	0.33	5-8	0.17	8-26	0.00	13-18	0.50	19-24	0.83
1-28	0.67	5-9	0.00	8-27	0.67	13-19	0.83	19-25	0.67
1-29	0.17	5-10	0.17	8-28	0.00	13-20	0.83	19-26	0.17
1-30	0.50	5-11	0.00	8-29	0.00	13-21	0.33	19-27	0.83
2-3	0.50	5-12	0.50	8-30	1.00	13-22	0.50	19-28	0.67
2-4	0.33	5-13	0.17	9-10	0.33	13-23	1.00	19-29	0.83
2-5	0.50	5-14	0.17	9-11	0.00	13-24	1.00	19-30	0.67
2-6	1.00	5-15	0.00	9-12	0.00	13-25	0.67	20-21	0.67
2-7	0.17	5-16	0.17	9-13	1.00	13-26	0.17	20-22	0.50
2-8	0.00	5-17	0.33	9-14	0.83	13-27	0.67	20-23	0.83
2-9	0.00	5-18	0.00	9-15	0.00	13-28	0.33	20-24	0.83
2-10	0.67	5-19	0.17	9-16	0.33	13-29	0.83	20-25	0.00
2-11	0.17	5-20	0.67	9-17	1.00	13-30	0.67	20-26	0.83
2-12	0.00	5-21	0.17	9-18	0.17	14-15	0.00	20-27	0.83

2-13	0.17	5-22	0.17	9-19	1.00	14-16	0.33	20-28	0.50
2-14	0.17	5-23	0.50	9-20	1.00	14-17	0.83	20-29	0.67
2-15	0.50	5-24	0.67	9-21	0.00	14-18	0.33	20-30	0.50
2-16	0.33	5-25	0.17	9-22	0.50	14-19	0.33	21-22	0.83
2-17	0.00	5-26	0.00	9-23	1.00	14-20	0.50	21-23	0.67
2-18	0.00	5-27	1.00	9-24	1.00	14-21	1.00	21-24	0.67
2-19	0.17	5-28	0.17	9-25	0.17	14-22	0.67	21-25	0.67
2-20	0.50	5-29	0.17	9-26	0.00	14-23	1.00	21-26	0.33
2-21	0.00	5-30	0.67	9-27	0.83	14-24	1.00	21-27	0.67
2-22	0.17	6-7	0.33	9-28	0.00	14-25	0.67	21-28	0.67
2-23	0.67	6-8	0.00	9-29	1.00	14-26	0.33	21-29	0.67
2-24	0.17	6-9	0.00	9-30	0.83	14-27	0.33	21-30	0.83
2-25	0.17	6-10	0.83	10-11	0.17	14-28	0.50	22-23	0.33
2-26	0.00	6-11	0.00	10-12	0.33	14-29	0.33	22-24	0.50
2-27	0.17	6-12	0.33	10-13	0.33	14-30	0.67	22-25	0.17
2-28	0.00	6-13	0.17	10-14	0.17	15-16	0.00	22-26	0.00
2-29	0.00	6-14	0.33	10-15	0.83	15-17	0.00	22-27	0.83
2-30	0.50	6-15	0.67	10-16	0.17	15-18	0.33	22-28	0.50
3-4	0.67	6-16	0.00	10-17	0.67	15-19	0.17	22-29	1.00
3-5	0.67	6-17	0.33	10-18	0.17	15-20	0.33	22-30	0.50
3-6	0.67	6-18	0.00	10-19	0.50	15-21	0.17	23-24	1.00
3-7	0.67	6-19	0.17	10-20	0.67	15-22	0.17	23-25	0.83
3-8	0.50	6-20	0.33	10-21	0.33	15-23	0.33	23-26	0.50
3-9	0.67	6-21	0.00	10-22	0.17	15-24	0.00	23-27	0.83
3-10	0.50	6-22	0.00	10-23	0.17	15-25	0.17	23-28	0.17
3-11	0.50	6-23	0.67	10-24	0.50	15-26	0.00	23-29	0.50
3-12	0.67	6-24	0.50	10-25	0.33	15-27	0.67	23-30	0.83
3-13	0.67	6-25	0.67	10-26	0.00	15-28	0.00	24-25	0.83
3-14	0.67	6-26	0.00	10-27	0.83	15-29	0.00	24-26	0.17
3-15	0.50	6-27	0.50	10-28	0.33	15-30	0.33	24-27	0.67
3-16	0.67	6-28	0.17	10-29	0.33	16-17	0.17	24-28	0.83
3-17	0.67	6-29	0.00	10-30	0.50	16-18	0.17	24-29	0.67
3-18	0.67	6-30	0.33	11-12	0.17	16-19	0.83	24-30	0.83
3-19	0.67	7-8	0.00	11-13	0.17	16-20	0.50	25-26	0.00
3-20	0.67	7-9	1.00	11-14	0.50	16-21	0.83	25-27	0.00
3-21	0.50	7-10	0.83	11-15	0.17	16-22	1.00	25-28	0.33
3-22	0.50	7-11	0.33	11-16	0.67	16-23	0.33	25-29	0.50
3-23	0.83	7-12	0.50	11-17	0.67	16-24	0.33	25-30	1.00
3-24	0.83	7-13	1.00	11-18	0.17	16-25	0.33	26-27	0.83
3-25	0.50	7-14	0.67	11-19	0.17	16-26	0.33	26-28	0.17
3-26	0.67	7-15	0.33	11-20	0.67	16-27	0.50	26-29	0.33
3-27	0.67	7-16	0.17	11-21	0.17	16-28	0.50	26-30	0.00
3-28	0.67	7-17	1.00	11-22	0.17	16-29	0.67	27-28	0.50

APPENDIX C. CONSENT FORM

University of North Carolina at Chapel Hill
Consent to Participate in a Research Study
Adult Participants

Consent Form Version Date: 10/17/2014
IRB Study # 14-1549
Title of Study: Evaluating the influences of serendipity on personalized medical news delivery
Principal Investigator: Xiangyu Fan
Principal Investigator Department: School of Information and Library Science
Principal Investigator Phone number: 9192594246
Principal Investigator Email Address: xyfan@email.unc.edu
Faculty Advisor: Javed Mostafa
Faculty Advisor Contact Information: 919-962-2182

What are some general things you should know about research studies?
You are being asked to take part in a research study. To join the study is voluntary. You may refuse to join, or you may withdraw your consent to be in the study, for any reason, without penalty.

Research studies are designed to obtain new knowledge. This new information may help people in the future. You may not receive any direct benefit from being in the research study. There also may be risks to being in research studies.

Details about this study are discussed below. It is important that you understand this information so that you can make an informed choice about being in this research study.

You will be given a copy of this consent form. You should ask the researchers named above, or staff members who may assist them, any questions you have about this study at any time.

What is the purpose of this study?
The purpose of this research study is to identify the impacts of serendipity on personalized medical news content delivery.

Are there any reasons you should not be in this study?
You should not be in this study if you are not healthy English speaking adults (18 years of age and older).

How many people will take part in this study?
There will be approximately 30 people in this research study.

How long will your part in this study last?
We expect the total duration of your participation to be about 2 hours.

What will happen if you take part in the study?
Users will be introduced to a medical news filtering system (i.e. MedSDFilter) and then asked to complete 10 filtering sessions using MedSDFilter system. During each session, users will be presented with a list of articles. For each clicked article, users are asked to show whether they are interested in the article content. At the end of the test, they will be asked questions about their experience to help interpret system's effectiveness in delivering relevant news to users.

What are the possible benefits from being in this study?
Research is designed to benefit society by gaining new knowledge. You will not benefit personally from being in this research study.

What are the possible risks or discomforts involved from being in this study?
There are no known risks to participating in this study.

What if we learn about new findings or information during the study?
You will be given any new information gained during the course of the study that might affect your willingness to continue your participation.

How will information about you be protected?
Subject privacy and confidentiality will be maintained in several ways. The user test results will be coded with unique identifiers. Linkage between the unique identifiers and the users' names will be kept in a password-protected database and retained separately from the user test results. The identifiers and test results will be destroyed after the data analysis is complete.

Participants will not be identified in any report or publication about this study. Although every effort will be made to keep research records private, there may be times when federal or state law requires the disclosure of such records, including personal information. This is very unlikely, but if disclosure is ever required, UNC-Chapel Hill will take steps allowable by law to protect the privacy of personal information. In some cases, your information in this research study could be reviewed by representatives of the University, research sponsors, or government agencies (for example, the FDA) for purposes such as quality control or safety.

What if you want to stop before your part in the study is complete?
You can withdraw from this study at any time, without penalty. The investigators also have the right to stop your participation at any time. This could be because you have had an unexpected reaction, or have failed to follow instructions, or because the entire study has been stopped.

Will you receive anything for being in this study?
You will be receiving compensation of $40 for taking part in this study. At any point in time you may withdraw from the study; if you do you will receive a prorated portion of the $40

stipend.

<u>Will it cost you anything to be in this study?</u>
It will not cost you anything to be in this study.

<u>What if you are a UNC student?</u>
You may choose not to be in the study or to stop being in the study before it is over at any time. This will not affect your class standing or grades at UNC-Chapel Hill. You will not be offered or receive any special consideration if you take part in this research.

<u>What if you are a UNC employee?</u>
Taking part in this research is not a part of your University duties, and refusing will not affect your job. You will not be offered or receive any special job-related consideration if you take part in this research.

<u>What if you have questions about this study?</u>
You have the right to ask, and have answered, any questions you may have about this research. If you have questions about the study (including payments), complaints, concerns, or if a research-related injury occurs, you should contact the researchers listed on the first page of this form.

<u>What if you have questions about your rights as a research participant?</u>
All research on human volunteers is reviewed by a committee that works to protect your rights and welfare. If you have questions or concerns about your rights as a research subject, or if you would like to obtain information or offer input, you may contact the Institutional Review Board at 919-966-3113 or by email to IRB_subjects@unc.edu.

<u>Participant's Agreement</u>:

I have read the information provided above. I have asked all the questions I have at this time. I voluntarily agree to participate in this research study.

____ ________________

Signature of Research Participant Date

____

Printed Name of Research Participant

____ ________________

Signature of Research Team Member Obtaining Consent Date

____

Printed Name of Research Team Member Obtaining Consent

APPENDIX D. QUESTIONNAIRES

PART 1: PRE-SESSION QUESTIONS

1.1 What is your age?

1.2 What is your gender?

1.3 How often do you look for news?

- ○ Several times a day
- ○ Once a day
- ○ Once a week
- ○ Once a month

1.4 How often do you look for medical news?

- ○ Several times a day
- ○ Once a day
- ○ Once a week
- ○ Once a month

1.5 What is the main factor for you to look for medical news?

- ○ Personal Health
- ○ Family Health
- ○ Friend's Health
- ○ Public Health
- ○ Other (Please specify if you select "Other")

1.6 What is the percentage of articles you can understand well from medical news you've read? (Please choose the option that best describes your situation)

- 80 - 100%
- 60 - 80%
- 40 - 60%
- 20 - 40%
- 0 - 20%

PART 2: POST-SESSION QUESTIONS

2.1 What did you like or dislike about the system?

2.2 Suggest one or two ways in which this system can be improved?

2.3 Do you think you will use a system like this to access health information? Why or why not?

2.4 Did the system provide you with news items you didn't expect? Provide an example (keywords are sufficient)

(a) Did you read or click on these articles? Why or why not?

(b) Did you like these items? Why or why not?

2.5 The system detected a difference between your interest profiles in pre-session and post-session surveys, please explain what caused your interest changes if possible?

* This question is optional, please ask researcher whether you need answer it.

2.6 Please provide an overall rating of the system's performance?

- Strongly satisfied
- Somewhat satisfied
- Neither satisfied nor dissatisfied
- Somewhat dissatisfied
- Strongly dissatisfied

APPENDIX E. RECRUITING EMAIL

Subject: Participate in a user study to earn $40 for reading medical news

Body:
Dear UNC staff members,

Do you like reading news based on your interests? If so, would you like to participate in a study which investigates how to effectively recommend medical news to you every day?

This research study is investigating the impacts of serendipity on personalized medical news delivery. If you choose to participate in this study, you will meet with a researcher individually in a campus location convenient to you and be asked to use a medical news filtering system, and then answer a few questions. The whole study takes about [x] hour.

****

Participants who complete the study will receive $40 as compensation for their time.

****

This is an IRB approved study (IRB#: 14-1549, approved date: 11/13/2014). Oversight for this study is provided by the UNC Chapel Hill Institutional Review Board (IRB) for Social and Behavioral Research. If you have questions or concerns about this study please contact the IRB at 919-966-3113 or by email at IRB_subjects@unc.edu.

For more information or to take part in this study please contact Xiangyu Fan at xyfan@email.unc.edu, or call 919-259-4246.

Sincerely,

Xiangyu Fan

APPENDIX F. RECEIPTS

IRB Study #14-1549

I acknowledge receipt of [$40] for participating in this research study.

______________________________________ __________________

Signature of Research Participant Date

____

Printed Name of Research Participant

____

Signature of Researcher Date

____

Printed Name of Researcher

Xiangyu Fan, Principal Investigator
CB #3360
School of Information & Library Science, UNC-CH
Chapel Hill NC, 27599-3360
xyfan@email.unc.edu

REFERENCES

Ackoff, R. L. (1967). Management misinformation systems. Management science, 14(4), B-147.

André, P., Teevan, J., & Dumais, S. T. (2009a). From x-rays to silly putty via Uranus: serendipity and its role in web search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2033-2036). ACM.

André, P., Teevan, J., & Dumais, S. T. (2009b). Discovery is never by chance: designing for (un) serendipity. In *Proceedings of the seventh ACM Conference on Creativity and Cognition* (pp. 305-314). ACM.

Basu, C., Hirsh, H., & Cohen W. (1998). Recommendation as classification: using social and content-based Information in recommendation. In *Proceedings of the 15th National Conference on Artificial Intelligence*, Madison, WI, USA, 714-720.

Beale, R. (2007). Supporting serendipity in data mining and information foraging. *International Journal of Human Computer Studies*.

Belkin, N. J. (1980). Anomalous states of knowledge as a basis for information-retrieval. *Canadian Journal of Information Science-Revue Canadienne Des Sciences De L Information*, 5(MAY), 133-143.

Belkin, N. J., & Croft, B. (1992). Information Filtering and Information Retrieval: Two Sides of the Same Coin? *Communications of the ACM*, 35(12), 29-38.

Blandford, A., & Buchanan, G. (2003). Usability of digital libraries: a source of creative tensions with technical developments. *IEEE Technical Committee on Digital Libraries Bulletin*, 1(1).

Bollacker, K. D., Lawrence, S., & Giles, C.L. (1999). A system for automatic personalized tracking of scientific literature on the web, In *Proceedings of the fourth ACM Conference on Digital Libraries*, August 11-14, 1999, Berkeley, CA, USA, 105-113.

Bollacker, K. D., Lawrence, S., & Giles, C. L. (2000). Discovering relevant scientific literature on the web. *Intelligent Systems and their Applications*, IEEE, 15(2), 42-47.

Bruijn, O., & Spence, R. (2008). A new framework for theory-based interaction design applied to serendipitous information retrieval. *ACM Transactions on Computer-Human Interaction*, 15(5), 5-38.

Brusilovsky, P., Kobsa, A., & Nejdl, W. (Eds.). (2007). *The adaptive web: methods and strategies of web personalization* (Vol. 4321). Springer Science & Business Media.

Calvi, L., & De Bra, P. (1997). Proficiency-adapted information browsing and filtering in hypermedia educational systems. *User Modeling and User-Adapted Interaction*, 7(4), 257-27.

Chewning, E. G., & Harrell, A. M. (1990). The effect of information load on decision makers' cue utilization levels and decision quality in a financial distress decision task. Accounting, Organizations and Society, 15(6), 527-542.

Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D., & Sartin, M. (1999, August). Combining content-based and collaborative filters in an online newspaper. In *Proceedings of ACM SIGIR workshop on recommender systems* (Vol. 60).

Cooper, J.W., & Prager, J.M. (2000). Anti-serendipity: finding useless documents and similar documents. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences* (Piscataway, NJ: IEEE Computer Society).

Deerwester, S. (1988). Improving Information Retrieval with Latent Semantic Indexing, In *Proceedings of the 51st Annual Meeting of the American Society for Information Science*, 25, 36-40.

Dumais, S., Cutrell, E., Cadiz, J. J., Jancke, G., Sarin, R., & Robbins D.C. (2003). Stuff I've Seen: A system for personal information retrieval and re-use. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, Toronto, Canada, 72-79.

Edmunds, A., & Morris, A. (2000). The problem of information overload in business organisations: a review of the literature. *International journal of information management*, 20(1), 17-28.

Eppler, M. J., & Mengis, J. (2004). The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. *The Information Society*, 20, 325-344.

Fan, X., Mostafa, J., Mane, K., & Sugimoto, C. (2012). Personalization is not a Panacea: Balancing Serendipity and Personalization in Medical News Content Delivery. *ACM SIGHIT International Health Informatics Symposium*, 709-714.

Filter bubble. (n.d.). In Wikipedia. Retrieved April 1, 2015, from http://en.wikipedia.org/wiki/Filter_bubble.

Foltz, P. W. (1990). Using latent semantic indexing for information filtering. In *ACM SIGOIS Bulletin*, 11(2-3), 40-47.

Foltz, P. W., & Dumais, S. T. (1992). Personalized information delivery: an analysis of information filtering methods. Communications of the ACM, 35(12), 51-60

Foster, A., & Ford, N. (2003). Serendipity and information seeking: An empirical study. *Journal of Documentation*, 59(3), 321-340.

Fuat, A., Hungin, A. P. S., & Murphy, J. J. (2003). Barriers to accurate diagnosis and effective management of heart failure in primary care: qualitative study. *Bmj*, 326, 196-198.

Goecks, J., & Shavlik, J. (2000). Learning users' interests by unobtrusively observing their normal behavior. In *Proceedings of the 5th international conference on Intelligent user interfaces*, 129-132.

Gup, T. (1997). Technology and the end of serendipity. *The Chronicle of Higher Education*, 44(21), A52.

Gup, T. (1998). Technology and the end of serendipity. *The Education Digest*, 6, 48-50.

Hall, A., Walton, G. (2004). Information overload within the health care system: a literature review. *Health Info Libr J.*, 21(2), 102-8.

Hanani, U., Shapira, B., & Shoval, P. (2001). Information filtering: Overview of issues, research and systems. *User Modeling and User-Adapted Interaction*, 11(3), 203-259.

Hanka, R., & Fuka, K. (2000). Information overload and 'just-in-time' knowledge. *The Electronic Library*, 18, 279-284.

Hibble, A., Kanka, D., Pencheon, D., & Pooles, F. (1998). Guidelines in general practice: the new Tower of Babel? *British Medical Journal*, 317, 862-863.

Huwe, T. K. (1999). New search tools for multidisciplinary digital libraries. *Online*, 23(2), 67-70, 72-74.

Hwang, M. I., & Lin, J. W. (1999). Information dimension, information overload and decision quality. *Journal of Information Science*, 25, 213-218.

Iselin, E. (1989). The impact of information diversity on information overload effects in unstructured managerial decision making. *Journal of Information Science*, 15, 163-173.

Keller, K. L., & Staelin, R. (1987). Effects of quality and quantity of information on decision effectiveness. *Journal of consumer research*, 200-213.

Kelly, D., & Belkin, N. J. (2001). Reading time, scrolling and interaction: exploring implicit sources of user preferences for relevance feedback. In *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*, 408-409.

Kelly, D., & Belkin, N. J. (2004). Display time as implicit feedback: understanding task effects. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval,* 377-384.

Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., & Riedl, J. (1997). GroupLens: applying collaborative filtering to Usenet news. *Communications of the ACM*, 40(3), 77-87.

Lang, K. (1995). Newsweeder: Learning to filter netnews. In *Proceedings of the 12th international conference on machine learning*, 331-339.

Linden, G., Smith, B., & York, J. (2003). Amazon. com recommendations: Item-to-item collaborative filtering. *Internet Computing, IEEE,* 7(1), 76-80.

Loeb, S. (1992). Architecting Personalized Delivery of Multimedia Information. *Communications of the ACM*, 35(12), 39-48.

Loeb, S., & Panagos E. (2011). Information Filtering and Personalization: Context, Serendipity and Group Profile Effects, *The 8th* Annual *IEEE Consumer Communications and Networking Conference*, 393-398.

Ma, Z., Pant, G., & Liu Sheng, O.R. (2007). Interest-based personalized search. *ACM Transactions on Information Systems*, 25(1), 5.

Malone, T. W., Grant, K. R., Turbak, F. A., Brobst, S. A., & Cohen, M. D. (1987). Intelligent information-sharing systems. *Communications of the ACM*, 30(5), 390-402.

Marchionini, G., & Shneiderman, B. (1988). Finding facts vs. browsing knowledge in hypertext systems. *IEEE Computer*, 21(1), 70-80.

McCleary, H. (1994). Filtered information services: a revolutionary new product or a new marketing strategy? *Online*, 18(4), 33-42.

Merton, R.K., & Barber, E. (2006). *The travels and adventures of serendipity: A study in sociological semantics and the sociology of science*. Princeton University Press.

Meyer, J. A. (1998). Information overload in marketing management. *Marketing Intelligence & Planning*, 16(3), 200-209.

Meyers, M. A. (2007). *Happy accidents: Serendipity in modern medical breakthroughs*. Arcade Publishing.

Miksa, F. L. (1992). Library and information science: two paradigms. *Conceptions of library and information science: historical, empirical and theoretical perspectives. Londres, Los Angeles: Taylor Graham*, 229-252.

Mooney, R. J., & Roy L. (1999). Content-based book recommendation using learning for text categorization. In *Proceedings of the Workshop on Recommender Systems: Algorithms and Evaluation*, Berkeley, California, USA.

Morita, M., & Shinoda, Y. (1994). Information filtering based on user behavior analysis and best match text retrieval. In *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, 272-281.

Mostafa, J., & Lam W. (2000). Automatic classification using supervised learning in a medical document filtering application. *Information Processing & Management*, 36(3), 415-444.

Mostafa, J., Mukhopadhyay, S., Lam, W., & Palakal, M. (1997). A multilevel approach for intelligent information filtering: model, system and evaluation. *ACM Transactions on Information Systems*, 15(4), 368-399.

Mostafa, J., Mukhopadhyay, S., & Palakal, M. (2003). Simulation studies of different dimensions of users' interests and their impact on user modeling and information filtering. *Information Retrieval*, 6(2), 199-223.

Mostafa, J., Quiroga, L.M., & Palakal, M. (1999). Filtering medical documents using automated and human classification methods. *Journal of the American Society for Information Science*, 49(4), 1304-1318.

Oard, D. W. (1997). The state of the art in text filtering. *User modeling and user-adapted interaction*, 7(3), 141-178.

O'Reilly, C. A. (1980). Individuals and information overload in organizations: Is more necessarily better? *Academy of management journal*, 23(4), 684-696.

Oxford English Dictionary Online. Retrieved on January 23, 2010 from http://dictionary.oed.com/entrance.dtl

Radlinski, F., & Dumais, S. (2006). Improving personalized web search using result diversification. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, Seattle, Washington, USA, 691-692.

Raskutti, B., Beitz, A., & Ward, B. (1997). A feature-based approach to recommending selections based on past preferences. *User Modeling and User-Adapted Interaction* 7(3), 179-218.

Salton, G., & McGill, M. J. (1983). *Introduction to modern information retrieval*, McGraw-Hill.

Schick, A. G., Gordon, L. A., & Haka, S. (1990). Information overload: A temporal approach. *Accounting, Organizations and Society*, 15(3), 199-220.

Shapira, B., Shoval, P., & Hanani, U. (1997). Stereotypes in information filtering systems. *Information Processing & Management*, 33(3), 273-287.

Sheth, B. D. (1994). A learning approach to personalized information filtering (Doctoral dissertation, Massachusetts Institute of Technology).

Sieg, A., Mobasher, B., & Burke, R. (2010). Ontology-based collaborative recommendation, In *Proceedings Of the 8th Workshop on intelligent Techniques for Web Personalization and Recommender Systems*, 20-31.

Simpson, C. W., & Prusak, L. (1995). Troubles with information overload—moving from quantity to quality in information provision. *International Journal of Information Management*, 15(6), 413-425.

Sundheim, B. Ed. (1991). *Proceedings of the third message understanding evaluation and conference*. Morgan Kaufmann, Los Altos, Calif.

Taylor, R. S. (1968). Question-negotiation and information seeking in libraries. *College & Research Libraries*, 29, 178-194.

Teevan, J., Dumais, S.T., & Liebling, D.J. (2008). To personalize or not to personalize: modeling queries with variation in user intent. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, Singapore, Singapore, 163-170.

Toffler, A. (1970). Future Shock, Amereon Ltd., New York, 350.

Toms, E. G. (2000). Serendipitous Information Retrieval. In *DELOS Workshop: Information Seeking, Searching and Querying in Digital Libraries*.

Turk-Charles, S., Meyerowitz, B. E., & Gatz, M. (1997). Age differences in information-seeking among cancer patients. *The International Journal of Aging and Human Development*, 45(2), 85-98.

Tushman, M. L., & Nadler, D. A. (1978). Information Processing as an Integrating Concept in Organizational Design. *Academy of management review*, 3(3), 613-624.

Vu, S., Bader, C., & Purdy, D. (2001). U.S. Patent Application 09/908, 443.

Wagholikar, K. B., MacLaughlin, K. L., Henry, M. R., Greenes, R. A., Hankey, R. A., Liu, H., & Chaudhry, R. (2012). Clinical decision support with automated text processing for cervical cancer screening. *Journal of the American Medical Informatics Association*, 19(5), 833-839.

Widyantoro, D.H, Ioerger, T.R., & Yen, J. (2000). Learning user interest dynamics with a three-descriptor representation. *Journal of the American Society for Information Science*, 52(3), 212-225.

Zeng, Q., Cimino, J. J., & Zou, K. H. (2002). Providing concept orientated views for clinical data using a knowledge-based system. *Journal of the American Medical Informatics Association*, 9, 294-305.

Ziegler, C. N., McNee, S. M., Konstan, J. A., & Lausen, G. (2005). Improving recommendation lists through topic diversification. In *Proceedings of the 14th international conference on World Wide Web*, 22-32, ACM.