Using Trust-based Recommender Systems for Personalized Health Content

Morteza Rashidi Koochi
1 e-mail: rkmorteza2@live.utm.my

Ab. Razak Bin Che Hussin
2 e-mail: abrazak@utm.my

Author(s) Contact Details:
1, 2 Faculty of Computing, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

Abstract — Recommender Systems intend to discover personalized items for a user based on item’s descriptions and user’s preferences. In Health era, Recommender Systems can be used to evaluate and extract useful personalized health content through mass of unreliable information on the web. Unlike many other types of recommendations, health recommendation extremely depends on emotional, physical and psychological matters of the user. Therefore using Personal Health Records or health profiles of user for personalized recommendation is inevitable but involves considerations about trust. In this work, we try to have a review of recommendation solutions for health domain. Furthermore, we analyze them considering their trust evaluation models, and propose an initial health recommender framework. This paper describes concepts of recommendation, medicine 2.0, and trust, reviews some solutions in these areas and then analyzes three most important attributes of web 2.0 (including trust and reputation, ontology and semantics, and social network collaboration), finally, proposes a framework for harvesting personal and trustable health content through social media to recommend. As the result, some feature of web 2.0 and their role in reviewed solutions are analyzed. It displays some weaknesses in using ontologies and semantic analysis. Moreover using different methods of trust management displays the different views of researchers about the notion of trust. Developing appropriate standards and ontologies, using semantic profile and content analysis to derive trust scores, and using limited personal health information solve current drawbacks partially.

Keywords – recommender system; trust; social networks; semantic analysis; health; medicine 2.0

1. INTRODUCTION

Rapid development and growth of web 2.0 technologies influenced the common shape of utilities and facilities that the internet provides for us. Considering the difference between traditional systems with emerging collaboration-based systems shows that the transition to the new generation for so many traditional systems is inevitable. One of the most well-known faces of web 2.0 services appears in Social Networks. Social relationships can appear in social networks in addition to individual dependencies [1]. With this potential, we can use mass information sharing, study others’ opinion and apply the public or expert contribution to make appropriate decision or improve our service or production.

Recently, emerging professional virtual communities with participation of large group of experts are attracting attention of users and providing professional consultation and recommendation [2]. In health area, there is an strong tendency to use Health Social Networks (e.g., PatientsLikeMe1) and Health Recommender Systems to achieve personalized healthcare [3][4][5][6]. Intuitive notion of social networks is based on social collaboration and contribution for content provision and this notion has a strong contradictory with high sensitivity of health content. Nonetheless, most of people are going to refer to the web for primary health information and it needs to provide acceptable trust and reliability of contents for users.

The rest of the paper is arranged as follows: the second section describes Recommender Systems based on techniques and knowledge source. It also has a glance to newly emerged concept notion in web 2.0, called medicine 2.0. In third section, we study health recommender systems and continue with review of trust-based health recommender systems considering the web 2.0 attributes. In the last section we have conclusion and suggest some directions for future research.

2. BACKGROUND

There is a growing interest to use social networks for health information provision.[7]notices some new challenges in health information in social networks such as highly sensitivity and difficulty of monitor information by experts for long time. For health context, it needs to consider auxiliary parameters like health condition [6], semantics [5], and experts opinion [8], to provide reliable content through social network for user.

1http://www.patientslikeme.com
Intuitive notion of social networks is based on social collaboration and contribution for content provision and this notion has a strong contradictory with high sensitivity of health content as mentioned before. Nonetheless, most of people are going to refer to the web for primary health information and it needs to provide acceptable trust and reliability of contents for users. Recommender Systems, providing personalized content, meet these requirements.

A. Recommender Systems

Recommender Systems (RS) are services or applications which provide appropriate recommendations for user in various fields such as product purchase, service selection or any source usage in real or cyber world. For this aim, traditionally there are two major methods: content-based filtering and collaborative filtering. Content-based recommendation acts on content and item’s description whilst collaborative filtering uses similar users’ idea about the item [9].

Content-based filtering is a model of filtering based on evaluations and information that retrieves from the content of sources. However because of the fast growing of content on the web and different professional and non-professional websites, big amount of content of the web is overlooked [10]. One of the problems in content-based filtering method is related to information retrieval, because common keyword-based search mechanisms fetches mass amount of information commonly not related to the subject. For this aim semantic relations can be used for jointing similar information together [11]. However complementary researches show that because of the information system limitations, both keyword-based and semantic content retrieval methods have limitations for retrieving appropriate content based on user’s stated need [12].

Collaborative filtering is a model of filtering that uses the opinions and ratings of similar users with similar interests to recommend based on opinion aggregation. This model is more known and used in recommender systems. Whereas Burke[13] states: "Collaborative filtering is also insensitive to preference changes of individual users”. Hence, another method of filtering so called “user-based collaborative filtering” emerges to utilize previously success recommendation for new users with similar interests [14]. With introductory statements mentioned above, we can redefine recommender systems as tools that predict utility of a specific item for suggesting to a specific user. However the common and primary part of every sort of recommender system is the knowledge source of recommendation (figure 2). The knowledge source can be the user generated data such as ratings and comments or can be system generated database about the domain. These knowledge sources can be used for rating prediction of users or provide top N-recommendations for users. According to the type of knowledge source of recommender system, Burke[15] divides recommender systems into different subdivisions (see figure 1):

- Content-based: content-based recommendations use user’s history of rating to find similar items with items that user has liked at the past, believing that everyone will like items similar to ones that liked before [16].
- Collaborative: collaborative recommendations use other users’ ratings, find out users with similar taste to current user and recommend him/her new items that other similar users have used or liked [17].
- Demographic: demographic recommendations use demographic items of user’s profile to recommend appropriate items suitable for him based on his profile’s attributes [18].
- Knowledge-based: knowledge-based recommendations use an explicit knowledge source (usually in an specific domain) inferring user’s needs or preferences [19], [20].

![FIGURE 1: Recommendation Techniques [15]](image)

Each of the above mentioned techniques has some drawbacks in implementation; for example, a controversial problem so called cold-start, takes place at the first entrance of new user or item in learning-based techniques, when there is not enough rating history to use in providing recommendations (content-based, collaborative, and demographic), or gray sheep problem when user’s preferences are dissimilar to majority (collaborative). Another type of problem occurs at the converse situation when the user’s preferences change, systems cannot adapt until new ratings with new preferences update the
scales. Among the above mentioned techniques knowledge-based recommendations have fewer problems, because they
do’t use historical preferences of users or don’t need retraining at the case of preference’s change [15]. Actually, real
recommender systems merge two or more techniques to tackle the weaknesses and gain better performance.

![Knowledge Sources in Recommender Systems](image)

**FIGURE 2: Knowledge Sources in Recommender Systems [20]**

**B. Medicine 2.0**

Eysenbach[21] used “Medicine 2.0” term for the first time to introduce the web-based health services using Web 2.0 and
semantic web technologies to facilitate “social networking, participation, apomediation (in place of intermediation, high-
quality information provision by web 2.0 tools), collaboration, and openness. It is considerable that joining between new
emerged social network systems like Facebook or a professional one like PatientsLikeMe[22], with electronic medical
systems such as “Personal Health Records” give users this opportunity to share their information with others.

Supposing this aim that medical information records contribute in recommendation in social networks or professional
communities we can design several scenarios for taking advantages. First of all, we can introduce entities such as: patient,
related physician, (virtual) community of experts, pharmacists, and nurses. For example, in case of uncertainty in diagnosis
the physician can use discussion facilities and share patient’s disease documents. After that, all of mentioned entities can
contribute in aggregated consultation, recommendation even diagnosis.

**3. HEALTH CONTENT RECOMMENDATION**

**A. Health Recommender Systems**

There are several works that attempt to use semantic web infrastructure to supply appropriate health information for
users. Wiesner, Rotter and Pfeifer[23] present a solution that semantic query used in information retrieval techniques
provides health information for users. They suggest semantic networks as formal relation among concepts to use for
categorizing health concepts like symptoms, diseases, and psychological entities. In spite of these requirements, health
category needs additional consideration such as emotional, physical and psychological matters of user [24].

MORMED [25] is a platform based on social networking facilities to supply collaboration among researchers, medical
doctors, and patients trying to cover features of Medicine 2.0 that Eysenbach declares in his well-known essay[21]. Whereas
it is a medical version of their previous project for knowledge management platform so-called OrganiK[26]. They try to
improve the quality of this platform by using semantic technologies, social networking and multilingual tools.

HARE [6] is a Health-Aware Recommender System that considers health, location, and context awareness. López-Nores
et al. after counting demographic, content-based, user-based, and item-based collaborative as different filtering strategies,
introduce property-based collaborative filtering as a new approach. This approach divides items into clusters based on similar
properties and makes a 3-dimentional dynamic matrix (item, property, user), equipped with ontology “a set of class
hierarchies and semantic attributes” and semantic reasoning to model the knowledge among the users and the items.

HRS (Health Recommender System) introduced by Wiesner and Pfeifer[5] uses a semantic graph extracted from
Wikipedia. They mention the benefits of using ontologies and semantic networks to leverage the quality and relativity of
health recommendations to requirements of users, but because of some drawbacks (e.g. being time consuming, costly, and
difficult to keep up-to-date), they use Wikipedia as ontology. HRS also uses Personal Health Record (PHR) system for more
tailored and personalized recommendation [23].
Dragu and Gomoi[27] suggest using “Topic Maps” as an standard for knowledge representation for structured and unstructured data and rules. They also mention subjectivism of the knowledge containing topic maps and cost of creating and maintaining as two major drawbacks.

### TABLE 1: Survey of Health Recommender Systems

<table>
<thead>
<tr>
<th>Author</th>
<th>Proposed Model</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[6]</td>
<td>Property-Based Collaborative Filtering for Health-Aware Recommender</td>
<td>Filtering strategy that characterizes the items and the users for recommendations relies on an ontology. (a set of class hierarchies and semantic attributes)</td>
</tr>
</tbody>
</table>

#### B. Trust-based Health Recommender Systems

There are mass of unreliable, incorrect and overlooked information on the websites that makes it hard to use internet sources for correct recommendation and decision making. It is necessary to present a solution that user can “trust” to information and knowledge that retrieve from recommender systems. Trust-aware recommender systems use various trust mechanisms to provide sufficient trustworthiness based on their trust factors to satisfy recommendation users. For this aim, recommender systems use trust factor of agents to show the authority of recommendation mechanism. Different trust evaluation models are used to estimate the trustworthiness of users by measuring individual or global trust score based on user’s behavior or reputation in the system (see table 2).

There are two distresses using information in social networks or internet-based solutions for personalized recommendation or consultation; the first problem is related to security and privacy concerns. For example, using Personalized Health Records in web-based applications are exposed to security attacks. Although, security is not in scope of this work, but it worth mentioning that this challenge is not enough to block the opportunity of using web 2.0 facilities in health context; Song and Marsh[29] propose a health condition indexing method for measuring the similarity in social networks anonymously. They introduce “d-words” (discriminant words) as words that show the health condition of user anonymously, instead of using personal health records to meet security concerns. They explain “Linear Variance Analysis” method for selecting d-words. The second problem is related to the trustworthiness of information or knowledge that we retrieve from the system. This concept of trust means how much people or systems can rely on achieved information[30]. In this paper we argue on the second concept of trust and discuss that how we can understand about the reliability of the information, what kind of methods we can use for that and what limitations we have for this aim.

Recently, a new approach called trust-aware recommender system proposed by Massa and Avesani[31] try to improve Collaborative Filtering method with search in web of trust for trustable users instead of similar users. This is one of the first solutions which exploit relationship among users to provide more accurate recommendation for the user. Trust-aware recommendation follows the natural intuition that “recommendation from trusted users is more acceptable for people”. According to this notion of trust relation, Huang et al.[11] claim clinicians trust more on suggestions from known experts and suggest “collaboration-based medical knowledge recommendation” that collaborate among clinicians. This method works in two phases; generating trust profile for clinician and generating recommendation. Clinicians search knowledge items using tags, recommender system proposes knowledge items with collaborative-tagging process considering working experience of clinicians to improve the trustworthiness of recommendations. Consumers can influence the priority with rating and tagging on items.

In a noticeable review on capacities and also challenges of exploiting recommender system to supply personalized health content, Fernandez-luque, Karlsen, and Vognild[32] compare Recommender Systems with Computer-Tailoring Health Education Systems and believe Social Networks’ capacities and Semantic Technologies can be used as complementary of Recommender Systems in health context. Social Networks can be used for discovering new interests of people and harvesting knowledge about contents and users, and Semantic Technologies can be used for integration of different
components and creation of semantic-enabled solutions. They also mention semantic health standards such as HL7\(^2\) and OpenEHR\(^3\) can accelerate achieving this aims. They also present HealthTrust\([33]\) as a ranking algorithm based on well-known algorithm HITS, to discover and score videos and channels on YouTube network about diabetes.

With the fast growing up of health related knowledge content like many other areas, collaboration and sharing among experts with similar interest and expertise is undeniable. To this aim, medical modules \([34]\) containing knowledge and expertise for collaboration among physicians is developed. It is a sort of MediGrid\([35]\) project that capsule pieces of knowledge in a set of so-called modules containing some attributes such as scientific publications, input and output data and description and capability to search with adding users’ tag and comment. The authors of \([34]\) claim inefficiency of traditional information retrieval methods in distributed environments and provide a reputation mechanism to supply accessibility to more reputable modules in case of knowledge overload.

<table>
<thead>
<tr>
<th>Author</th>
<th>Proposed Model</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[11]</td>
<td>Collaboration-based medical knowledge recommendation</td>
<td>Clinician trust profile based on the measure of trust factors combine the clinicians’ own experiences with scientific evidences from the medical community</td>
</tr>
<tr>
<td>[34]</td>
<td>Reputation Based Trust Management Supporting Collaboration in a Medical Application</td>
<td>Information modules created by clinicians, rated by users. Reputation of a module is based on popularity and reputation of the author.</td>
</tr>
<tr>
<td>[33]</td>
<td>HealthTrust: Trust-based Retrieval of YouTube’s Diabetes Channels</td>
<td>Analyzes the reliability of information in diabetes online community(YouTube’s Channels). Algorithm: Hyperlink-Induced Topic Search (HITS) ranking authoritative sources.</td>
</tr>
<tr>
<td>[37]</td>
<td>Ranking Medical Claims based on Community Knowledge</td>
<td>Predicts the trustworthiness of a medical claim based on experiences shared by users rank sites based on aggregating the trust scores of claims from the site.</td>
</tr>
<tr>
<td>[38]</td>
<td>UserRank: Influential Users in an Healthcare Social Network</td>
<td>Extract Social Network structure from MedHelp forum. Assess user’s influence based on distance and content similarity.</td>
</tr>
</tbody>
</table>

4. COMPARATIVE ANALYSIS AND FRAMEWORK SUGGESTION

User-generated contents and social relations and interactions are two most important of web 2.0 have emanated in social networks. These features have both threats and opportunities; contents which are generated by user contain emerging topics of interest, but, usually are non-professional, ambiguous, and with vulgar terminology. Moreover, people whom user interact with, may have malicious or biased aims. Therefore, to overcome the above mentioned threads, we will analysis two technologic domains respectively: Semantics and Trust in social relation and collaboration.

Moturu, Liu, and Johnson\([39]\) declare that for trusting to a content there are many parameters to study from different points of view including contributors, references, objectivity and evaluation. According to \([39]\) “The biggest challenge for Health 2.0 research might be posed by the lack of standards among websites in the same application domain.” They offer using text mining and (natural) language processing techniques for semantic analysis, classification and clustering for data processing and statistics and machine learning approaches for link mining, graph mining and social network analysis.

According to table 3, the mentioned researches have various points of view to calculate trust and reputation. Huang et al. \([11]\) generate trust profile with collaboration of others clinicians, Sp´anek et al. \([34]\) use users’ rating, Weitzel et al. \([36]\) that propose their solution based on twitter, consider higher retweets for higher trust, and Tang and Yang \([38]\) give more point of trust for more influential users with similar content.

---

\(^2\) HL7, http://www.hl7.org

\(^3\) openEHR, http://www.openehr.org
From the point of view of collaboration, some researches contemplate the professional communities and use collaboration between experts [11][34], some consider common social networks [38][36], and even contribute between other entities like movie channels [33].

### TABLE 3: Web 2.0 attributes in Health Recommender Systems

<table>
<thead>
<tr>
<th>Specifications</th>
<th>Proposed Model</th>
<th>Trust/Reputation</th>
<th>Ontology/Semantic</th>
<th>Collaboration/Social Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reputation Based Trust Management System Supporting Collaboration in a Medical Application [34]</td>
<td>Reputation Based Trust Management System Supporting Collaboration in a Medical Application [34]</td>
<td>Reputation Calculation Based on users’ rating</td>
<td>---</td>
<td>collaboration between experts in Biomedicine</td>
</tr>
<tr>
<td>HealthTrust: Trust-based Retrieval of YouTube’s Diabetes Channels [33]</td>
<td>HealthTrust: Trust-based Retrieval of YouTube’s Diabetes Channels [33]</td>
<td>Adaptation of HITS score: &lt;authority,hub&gt;</td>
<td>---</td>
<td>Channels as Nodes and Links between them as Edges</td>
</tr>
<tr>
<td>Ranking Medical Claims based on Community Knowledge [37]</td>
<td>Ranking Medical Claims based on Community Knowledge [37]</td>
<td>Trustworthiness of medical claim based on textual features</td>
<td>Redirect link-graph of Wikipedia</td>
<td>---</td>
</tr>
<tr>
<td>UserRank: Influential Users in an Online Healthcare Social Network [38]</td>
<td>UserRank: Influential Users in an Online Healthcare Social Network [38]</td>
<td>Influential users based on distance and content similarity</td>
<td>---</td>
<td>Social Network from Thread Analysis</td>
</tr>
</tbody>
</table>

Using ontology and applying semantic analysis to improve the accuracy and reliability of recommendations for targeted users, seems to be a crucial phase for trust-based recommender systems. Referring to respective column of table 3 shows that a few of solutions use ontology and semantic analysis. Huang et al.[11] use latent semantic indexing for semantic search and Vydiswaran, Zhai, and Roth [37] use graph of Wikipedia as a public ontology. It shows that there are serious limitations using and maintaining domain ontologies also developing solutions to analysis the content semantically.

As a proposed solution, A framework is presented in figure 2 containing the main elements of harvesting content through social media, estimating trust level from local source (web of trust) and global source (influential users), and finally, clarification and annotation process on content using auxiliary resources such as domain ontologies (formal knowledge for a certain domain).

![FIGURE 2: A generic framework of trust-based and personalized health recommender](image)

User’s web-of-trust helps to find suitable content which satisfies personalization and reliability concerns. Furthermore, because of high-sensitivity of health content and need to validate by experts, exploiting influential users of the social media

---

\(^4\) Latent Semantic Indexing
and implicit verification of the content helps to provide reliable and trustworthy content to recommend. Thus, for this aim, we verify both content’s trustworthy and author’s reliability (reputation). Expert medical terminologies are often unclear for common people and contrary, vulgar terms used by people are obscure for physicians or researchers. Therefore, health domain ontologies need to be equipped with both terminologies to transform the content to the readable version for user. Finally, recommender engine, with the help of user’s health profile supply personalized and trustable content to recommend to the user.

5. CONCLUSION

Successful outcomes using recommender systems in electronic commerce causes more attention from other fields such as health and medicine. In health area because of high importance of information used for diagnosis and treatment, trustworthiness of recommendation is more remarkable. Discordance between several mentioned proposed models shows that researchers look at trust concept from different view of point and this plurality will increase when trust is used in new areas. People use social networks and recommender systems for their common daily issues and share their personal information to retrieve appropriate and personalized results. Trust is a necessary requirement in both faces: first privacy and security and second reliability. Linking health information records to public networks is a solution but arise trust concerns. Nevertheless making suitable standard and ontology and using implicit personal information satisfies the security concerns of the solutions. Personal recommendation needs personal information and personal information is subject of dissemination. There is no complete solution solving the entire problem. However, the best solutions maybe are based on equal attention to double side of the problem; security and personalization.

REFERENCES


