Information Processing and Management 50 (2014) 599–620

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/infoproman

Enriching queries with user preferences in healthcare

Tesfa Tegegne *, Th.P. (Theo) van der Weide

Institute for Computing and Information Sciences, Radboud University, Nijmegen, Netherlands

ARTICLE INFO

Article history:
Received 12 September 2012
Received in revised form 7 March 2014
Accepted 8 March 2014
Available online 23 April 2014

Keywords:
Query enrichment
User preference
Conditional preference

ABSTRACT

Query enrichment is a process of dynamically enhancing a user query based on her preferences and context in order to provide a personalized answer. The central idea is that different users may find different services relevant due to different preferences and contexts. In this paper, we present a preference model that combines user preferences, user context, domain knowledge to enrich the initial user query. We use CP-nets to rank the preferences using implicit and explicit user preferences and domain knowledge. We present some algorithms for preferential matching. We have implemented the proposed model as a prototype. The initial results look promising.

1. Introduction

Technological advances have brought tremendous progress in healthcare. Examples are: electronic medical records, mobile health and eHealth, videoconferencing, medical decision making, remote monitoring and many more. Besides these, healthcare now can provide personalized health decisions for patients and empower patients to participate in the decision making. Recently, patient empowerment and engagement has gotten a big boost. Patient-centeredness helps patients and families and/or doctors to make informed healthcare decisions during diagnosis and treatment (Arnold, 2013).

However, these benefits of technology in healthcare are not equitably available for developing countries with their restricted infrastructure and limited resources. Illiteracy, little medical knowledge and poor or ambiguous healthcare queries are factors that hinder the realization of patient empowerment. Low infrastructure, unavailability of technologies, costs and shortage and high turnover of clinicians are some of the obstacles to realize patient-centeredness in particular and eHealth in general. These challenges and opportunities have triggered the work in this paper. Especially we have focused on the Ethiopian situation (see (Berhan, 2008; Serra, Serneels, Lindelow, & Montalvo, 2010 in general and Tegegne & Weide, 2011; Tegegne, Kanagwa, & Weide, 2010) for a more detailed description).

In this paper we focus on a basic architecture to support health workers. This architecture supports the health worker to make a diagnosis, and then tries to find the best treatment for the patient at hand adapted to the environment of this patient. The dialogue support principles have been discussed in Tegegne and Weide (2013) and are outside the scope of this paper. In this paper we focus on the enrichment of the initial (medical) query, before matching it against a body of medical knowledge and making a prioritization of the possible treatments.

Rather than setting up an advanced expert system to cover for any medical case, we use a basic Information Retrieval approach to match a diagnosis with patterns defined for typical diseases. Then we use personal and environmental information to rank the associated treatments. The diagnosis can be seen as the initial query, that is enriched by personal and environmental information.

* Corresponding author. Tel.: +31 685177546.
E-mail addresses: t.tegegne@cs.ru.nl (T. Tegegne), tvdw@cs.ru.nl (Th.P. (Theo) van der Weide).

http://dx.doi.org/10.1016/j.ipm.2014.03.004
0306-4573/

Illiteracy and low level medical knowledge brings about poor query formulation. To this end, we incorporate personal profile and context to enrich the user initial query (which consists of signs and symptoms of the patient). This work also allows clinicians, specially community health workers (Health Extension Workers in Ethiopia context) with very limited training, to query diseases and treatments based on patients profile, symptoms and contexts.

The motive behind query enrichment is to enable low/semiliterate patients and clinicians (health workers) to create a medical query to search personalized healthcare information as well as to provide rich diagnosis and treatment options for both patients and clinicians. In general, enriching user queries using domain knowledge, personal profile and context can help to maximize cure and to minimize adverse effects, cost, and waiting time.

Queries may contain implicit preferences. This is for example the case in the query fragment *Martha is 6th months pregnant. She is on HIV/AIDS regimen; currently she shows some signs of malaria*. At this moment, the system should identify complications: antiretroviral may cause antimalarial resistance as well as antimalarial may affect pregnancy. Thus, while enriching the query, it is advisable to be cognizant of the impact of indirect preferences stated in the query. In the above query the woman might not have any knowledge about all the constraints resulting from her condition.

Other preferences may come from (if known) a user profile and a user context. The user profile contains interests and preferences, while the user context is considered as the actual state of the user current task. For the purpose of our paper, we see both user profile and user context as preferences for that user. In this paper we focus on handling user preferences in the context of healthcare application. Typically, the query (request) then is a consequence of a diagnosis obtained by the doctor or an automated diagnosis system. Consequently, we may assume that both user profile and user context are available upon entering the query (request).

The main purpose of this paper is to answer the following question: *How to develop a general preference model for enriching queries* in the context of healthcare application. This leads to the following sub-questions:

R-1 What is a general preference model that can handle medical preferences (medical preferences in this context is patients treatment or medication preference)?
R-2 How is preferential matching defined in this model?

The layout of this paper is as follows. Related work is presented in Section 2. In Section 3 we show the overall architecture of the proposed system. We shortly discuss the system components and discuss the overall quality of the system. The preference model is presented in Section 4. In Section 5 we discuss the preferential similarity model. Section 6 presents the application of preference prioritization in healthcare domain; and a medical case study is discussed to test our model in Section 7. Section 8 gives a brief description of the prototype. Finally, we conclude with some future research directions in Section 9.

2. Related work

The expansion of the Internet is providing a fertile ground to access a wealth of information. As the volume of heterogeneous web resources increases and the data become more varied, a massive response is issued to user queries. That makes it hard to distinguish relevant information from irrelevant or secondary information. Nowadays, various mechanisms are employed to enhance user queries, such as reformulation user queries, expanding query terms, adding a user profile and considering the user context.

Query enrichment is the process of transforming the initial query into a form that more adequately reflects the information need of the user. Query enrichment in this paper focuses on enhancing the initial user query by incorporating user profile, user context and medical knowledge. It is different from query reformulation and query expansion, which deals with rewriting the initial user query by adding or removing words, phrases or clauses. In some systems, expansion terms are added to the query by the user (interactive query expansion (Efthimiadis, 2000), like the Google search engine), while other systems the expansion terms are added by the search engine or system itself (automatic query expansion).

The query enrichment we propose here varies between query reformulation and query expansion. The enrichment is performed automatically by fetching the user profile, the user context and domain knowledge from the database and does not involve the user intervention.

Asfari, Doan, Bourda, and Sansonnet (2010) use user profiles and user tasks to improve user queries. They consider the state of user's task at a specified time. The drawback of this approach is that it does not take into account user's preference relations, preference ranking, etc. Query enrichment by integrating the user profile is addressed by Koutrika and Ioannidis (2010) and Kießling (2002).

Several authors propose a scheme for query enrichment by relevance feedback, beginning the search with an initial query and then modifying it by relevance judgments of the user (Bhogal, Macfarlane, & Smith, 2007). Then term expansion is used to reformulate the user query (Bai & Nie, 2008).

Different approaches are used for the process of query enrichment; for example Conesa, Storey, and Sugumaran (2008) try to formulate the user query by using semantic knowledge, while others use term relations to expand the user's initial query (Bai & Nie, 2008). Bhogal et al. (2007) and Sieg, Mobasher, and Burke (2007) propose ontology based query expansion. Semantic based query term relations in order to enhance queries is addressed by Benz, Krause, Kumar, Hotho, and Stumme (2009).