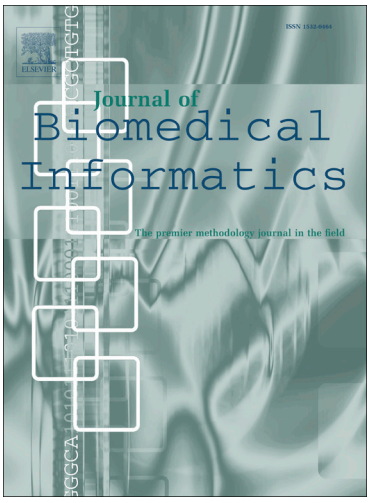

A cloud-based framework for large-scale traditional Chinese medical record retrieval

Lijun Liu, Li Liu, Xiaodong Fu, Qingsong Huang, Xianwen Zhang, Yin Zhang

PII: S1532-0464(17)30258-7
DOI: https://doi.org/10.1016/j.jbi.2017.11.013
Reference: YJBIN 2892

To appear in: *Journal of Biomedical Informatics*

Received Date: 12 April 2017
Revised Date: 2 November 2017
Accepted Date: 20 November 2017

Please cite this article as: Liu, L., Liu, L., Fu, X., Huang, Q., Zhang, X., Zhang, Y., A cloud-based framework for large-scale traditional Chinese medical record retrieval, *Journal of Biomedical Informatics* (2017), doi: https://doi.org/10.1016/j.jbi.2017.11.013

A cloud-based framework for large-scale traditional Chinese medical record retrieval

Lijun Liu[a], Li Liu[b], Xiaodong Fu[b], Qingsong Huang[b,*], Xianwen Zhang[c] , Yin Zhang[d]

[a] Department of Biomedical Engineering, Kunming University of Science and Technology, Kunming, 650500, China

[b] Department of Computer Science and Technology, Kunming University of Science and Technology, Kunming, 650500, China

[c] Medical Faculty, Kunming University of Science and Technology, 650051, Kunming, China

[d] Information Center, Kunming Municipal Hospital of Traditional Chinese Medicine, 650051, Kunming, China

* Corresponding author. Address: Room 403, Faculty of Information Engineering and Automation, No.727 South Jingming Rd., Chenggong District, 650500, Kunming, China.

Qingsong Huang E-mail : ynkmhqs@sina.com, cloneiq@126.com.

1. Introduction

Electronic medical records are increasingly common in medical practice. They represent a relatively novel and rich resource for clinical research. The secondary use of medical records provides great potential for improving clinical research efficiency [1, 2] and medical quality [3, 4]. However, the secondary use of medical records relies on the ability to retrieve the complete information about desired patient populations [5]. This fact makes medical search is becoming a critical technique for the rapid and effective access of patient information, which also provides great potential for facilitating research and improving quality in medical practice [5].

How to effectively and accurately retrieve relevant medical records is becoming a big challenge. The first difficulty lies in the huge amount of medical records. It is very time-consuming for healthcare providers to dig into the voluminous medical records of a patient to find the few that are indeed relevant to the patient's current problem [6]. Some works have been carried out on medical records retrieval to search relevant medical records, but most of them focus on centralized architecture [5, 7, 8], which cannot scale up well to large data volume. In addition, the latest relevant medical records are difficult to be indexed and retrieved in real-time due to the fact that creating and updating the large-scale indexes is very time-consuming. Second, the high quality medical search is a challenging task, in particular due to the inherent complexity and ambiguity of medical terminology [5]. The existing semantic approaches provide great potential to improve the quality of medical retrieval [5, 9], but they are not suitable to deal with Traditional Chinese Medical Records (TCMRs). A typical medical record contains a variety of structured data such as electronic prescriptions and laboratory values. A substantial portion of clinical data is also embedded in unstructured data in the form of narrative text notes [10]. The data usually has the variable and complex data types and formats [11, 12]. In this context, the unstructured data is lack of standardized description and friendly visualized interface. Therefore, the current medical record retrieval systems would be limited in terms of availability and universality.

In this paper, we propose an efficient and robust framework based on cloud for large-scale TCMRs retrieval called Cloud Medical Record Retrieval System (C-MRRS). The extensive experiments indicate that C-MRRS is specifically suitable to deal with large-scale TCMRs retrieval. It is more easily integrated with the existing clinical systems. The C-MRRS acts as a web service and provides a friendly web interface that can be used in various scenarios. It can be used as a medical decision support system for medical diagnosis, as a medical record integration platform for research or as a case-based learning system for education.

The remainder of this paper is organized as follows. Section 2 discusses the preliminary knowledge of medical