Future Generation Computer Systems 33 (2014) 21–31

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

Similarity-based behavior and process mining of medical practices

Shusaku Tsumoto [a,*], Haruko Iwata [b], Shoji Hirano [a], Yuko Tsumoto [c]

[a] Department of Medical Informatics, School of Medicine, Faculty of Medicine, Shimane University, 89-1 Enya-cho Izumo 693-8501, Japan
[b] Division of Nursing, Shimane University Hospital, 89-1 Enya-cho Izumo 693-8501, Japan
[c] Department of Fundamental Nursing, School of Nursing, Faculty of Medicine, Shimane University, 89-1 Enya-cho Izumo 693-8501, Japan

HIGHLIGHTS

- Temporal data mining process for hospital management was proposed.
- The chronological overview of hospital activities was visualized.
- Clustering was applied for showing temporal characteristics of divisions.
- Trajectories mining was applied for capturing characteristics of divisions.
- Similarity-based analysis were applied for characterizing nursing orders.

ARTICLE INFO

Article history:
Received 5 April 2013
Received in revised form
28 September 2013
Accepted 4 October 2013
Available online 11 November 2013

Keywords:
Temporal data mining
Visualization
Clustering
Hospital information system

ABSTRACT

This paper presents data mining results in which the temporal behavior of global hospital activities is visualized. The results show that the reuse of stored data will provide a powerful tool for hospital management and lead to improvement of hospital services.

1. Introduction

Twenty years have passed since clinical data were stored electronically as a hospital information system [1,2]. Stored data provide all the histories of clinical activities in a hospital, including accounting information, laboratory data and electronic patient records. Due to the traceability of all the information, a hospital cannot function without an information system.

However, reuse of the stored data has not yet been discussed in detail, except for laboratory data and accounting information to which OLAP methodologies are applied. The data mining methods started getting attention in hospital applications just ten years ago [3,4] and the area is still waiting for a comprehensive research [5–7].

In this paper, we first propose a scheme for the innovation of hospital services based on data mining. Then, based on this scheme, we applied data mining techniques to the data extracted from hospital information systems. The results included several interesting findings, which suggest that the reuse of stored data will provide a powerful tool to improve the quality of hospital services.

The paper is organized as follows. Section 2 proposes a general framework on the innovation of hospital services based on data mining. Section 3 briefly explains how a hospital information system works, as the background to this study. Section 4 explains on data preparation and the mining process. Section 5 shows the results of the visualization of hospital activities by using HIS data. Section 6 shows a clustering-based analysis of similarities between divisions. Section 7 applies a trajectories mining technique to temporal analysis of the number of orders. Finally, Section 9 concludes this paper.

2. Our goal: data-mining based hospital services

Fig. 1 shows our goal for hospital services, which consists of the following three layers of hospital management: services for

* Corresponding author. Tel.: +81 853202171.
E-mail addresses: tsumoto@computer.org, tsumoto@med.shimane-u.ac.jp (S. Tsumoto), haruko23@med.shimane-u.ac.jp (H. Iwata), hirano@med.shimane-u.ac.jp (S. Hirano), tsumotoy@med.shimane-u.ac.jp (Y. Tsumoto).

0167-739X/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.future.2013.10.014

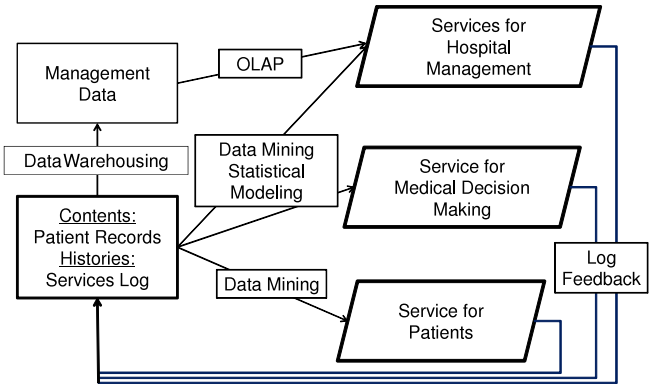

Fig. 1. Service-oriented hospital management.

hospital management, devices for the medical staff and services for patients. Data mining in a hospital information system plays a central role in achieving these layers.

The first layer is called services for patients. It supports the improvement of healthcare service delivery for patients. This is a fundamental level of healthcare services in which medical staff directly provide medical services to the patients. Patient records and other results of clinical examinations support the quality of this service. The second layer is called services for medical staff. This layer supports the decision making of the medical practitioner. Patient histories and clinical data are applied to data mining techniques which gives useful patterns for medical practice. Specifically, patient risk detection, such as adverse drug reactions or the temporal status of chronic diseases will improve the qualities of medical services. The top layer is called services for hospital management. This level is achieved by capturing the global behavior of a hospital: the bridge between the microscopic behavior of medical staff and the macroscopic behavior of the hospital is very important in deploying medical staff in an optimal way for improving the performance of the hospital.

Here, data mining plays a central role in bridging between data storage (service from human to computer) and data-oriented computerized service (from computer to human). Especially, it is expected that data mining will be a useful tool for the refinement of computerized hospital service.

3. Background

3.1. Hospital information system: cyberspace in the hospital

Clinical information has been stored electronically as a hospital information system (HIS). The database stores all the data related with medical actions, including accounting information, laboratory examinations, and patient records described by medical staff. Incident or accident reports are not exceptions; they are also stored in HIS as clinical databases. For example, Fig. 2 shows the structure of the HIS in Shimane University Hospital, where all the clinical inputs are shared through the network service in which medical staff can retrieve their information from their terminals [8,9].

Since all the clinical data are distributed, stored, and connected as a large-scale network, HIS can be viewed as a cyberspace in the hospital: all the results of clinical actions are stored as "histories". It is expected that techniques in web mining or network mining can be applied to the data. Dealing with cyberspace in a hospital will provide a new and challenging problem in hospital management in which the spatiotemporal data mining, social network analysis and other new data mining methods may play central roles [10].

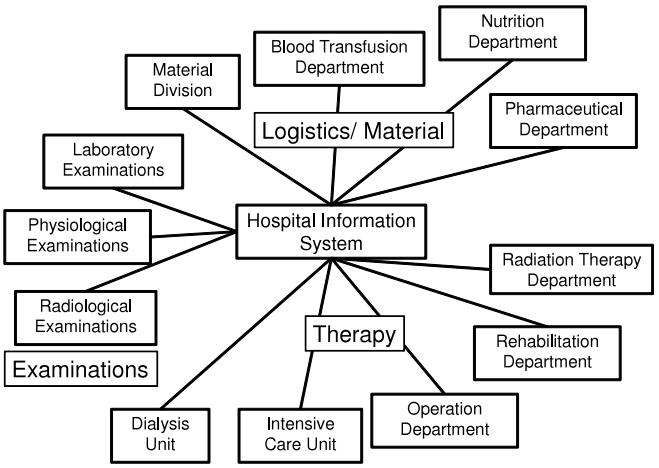

Fig. 2. Hospital information system in Shimane university.

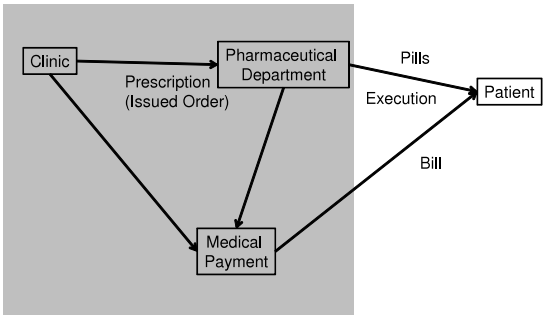

Fig. 3. Workflow of prescription order.

3.2. Basic unit in HIS: order

The basic unit in HIS is an "order", which is a kind of document or message, conveying an order from a medical practitioner to others. For example, a prescription can be viewed as an order from a doctor to a pharmacist. The prescription order is executed as follows:

1. outpatient Clinic;
2. a prescription given from a doctor to a patient;
3. the patient brings the order to the medical payment department;
4. the patient brings the order to the pharmaceutical department;
5. execution of the order in a pharmacist's office;
6. delivery of the prescribed medication;
7. payment.

The second to fourth steps can be viewed as information propagation: thus, if we transmit the prescription through the network, all the departments involved in this order can easily share the ordered information and execute the order immediately. This also means that all the results of the prescription process are stored in HIS.

Fig. 3 depicts the workflow of prescriptions between doctors, pharmacologists, patients and the pay desk. For comparison, Fig. 4 shows the workflow of an injection.

These sharing and storing processes, including histories of orders and their results, are automatically collected as a database: HIS can also be viewed as a cyberspace of medical orders.

3.3. Related work

Although it is well recognized that hospital information system stores "big data", application of data mining to big hospital data