Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 119 (2017) 234–244

Procedia Computer Science

www.elsevier.com/locate/procedia

6th International Young Scientists Conference in HPC and Simulation, YSC 2017, 1-3 November 2017, Kotka, Finland

Towards evolutionary discovery of typical clinical pathways in electronic health records

Anastasia A. Funkner[a*], Aleksey N. Yakovlev[a,b], Sergey V. Kovalchuk[a]

[a]*ITMO University, Saint Petersburg, Russia*
[b]*Almazov National Medical Research Centre*

Abstract

According to value-based health care, hospitals should deliver high quality and cost-effective medical treatments. Clinical pathways represent typical ways of treatment and indicate its impact on hospital resources. In this paper, we propose three approaches to discover and to cluster clinical pathways using the k-means method and genetic algorithms. A data set of 3434 electronic health records of patients with acute coronary syndrome is used to test proposed approaches. The approaches are compared with five clustering metrics. The best approach shows the best values for four of five metrics. In the future, we plan to embed the best approach of discovering clinical pathways in personalized Decision Support System of Almazov National Medical Research Centre (Saint Petersburg, Russia).

Peer-review under responsibility of the scientific committee of the 6th International Young Scientist conference in HPC and Simulation

Keywords: clinical pathways, genetic algorithm, process mining, clustering

1. Introduction

Value-based health care (VBHC) becomes priority policy in many clinics and hospitals all over the world. According to VBHC, clinical expertise and patient's' expectations have to be a base for medical decision-making [1]. Furthermore, healthcare staff should deliver high quality and cost-effective medical treatments by taking the

* Corresponding author. Tel.: +7-812-914-59-46.
E-mail address: funkner.anastasia@gmail.com

1877-0509
Peer-review under responsibility of the scientific committee of the 6th International Young Scientist conference in HPC and Simulation
10.1016/j.procs.2017.11.181

opinions of patients about the quality of life into account [1], [2]. One of the purposes of VBHC is to cure patients in an effective way without redundant therapy or screening which in turn saving hospital resources and patients' money [3]. So that, if patients' clinical pathway could be predicted in advance, it may help to plan allocation of resources (human and material) effectively and efficiently. Discovering clinical pathways can be a part of modeling of various hospital's aspects, e. g. discovered clinical pathways can be used to simulate clinical departments and patient flow through these departments of the hospital [4].

A clinical pathway (CP) means a common way how patients are treated in a hospital or its departments. Parameters of CPs are specially defined by a hospital, a country or experts who discover and use CPs in practice [5]. In this research, a CP is considered as the most common patient's movement through the departments and sub-departments (facilities with specialization within a department). The complexity of CP depends on the severity of a disease, presence of complications and methods of treatment. Because of patients' peculiarities, various hospitals' structures and individual experience of a doctor, different patients are treated in different ways and, consequently, there are many CPs which differ from medical guidelines. Moreover, CPs allow accessing the impact of treatment methods on resources of hospital and outcomes of treatments [6].

Generally, techniques for data and process mining are commonly used to discover CPs [7], [8]. In the case of discovering CPs for the whole set of thousands of patients, it is necessary to solve the problem of "Spaghetti" processes [9]. "Spaghetti" processes usually have many activities and it is a challenge to discover some patterns inside them. Patients with the same diagnoses can be treated in very different ways and, as a result, related to such treatment processes are usually too complex. Thus, we propose to divide patients into some groups and discover CPs inside each group.

The remaining part of the paper is organized as follows. Section 2 reviews existing techniques to discover clinical pathways. Section 3 describes our approaches of CPs discovering in detail. Section 4 demonstrates a comparative example of the proposed approaches' application by exploring clinical pathways of patients with Acute Coronary Syndrome (ACS) who were admitted to Almazov National Medical Research Centre (Saint Petersburg, Russia), one of the leading cardiological centers in Russia. Finally, the last section presents concluding remarks and future development of the proposed ideas.

2. Related Work

Discovery of clinical pathways is a nontrivial issue and does not have a unique solution. Data mining techniques include a frequent sequence mining technique, techniques based on optimization problems, a Bayesian networks technique etc. [10]–[12]. Liu et al develop an interactive recommender system called Pathway-Finder [12]. This system can automatically discover clinical pathways, visualize them and predict an outcome of a patient. Also, Perer et al develop Care Pathway Explorer which extract common sequences of medical events and explore how these sequences correlate with an outcome of a patient [10].

The most common process mining techniques are the Alpha-Algorithm [13], the Fuzzy Miner [14] and the Heuristic Miner [15]. They allow extracting typical CPs from a structured event log. With CPs it becomes possible to simulate the workflows through the hospitals and its departments [4], [16]. Furthermore, the unexpected CPs are also of greater importance for hospital managers and need to be explained by domain specialists [17], [18].

3. Methods and Algorithms

To solve "Spaghetti" process problem, we aim to divide patients into some groups and discover CPs for each group. Also, hospital data usually lack consistency, completeness, and correctness, as some events can be missing, have an erroneous time stamp or mistakes in event's descriptions [19]. Thus, proposed approach of CP's discovering also should be able to detect atypical CPs. In this section, three approaches of CP's clustering and discovering are described.

Preliminaries

This section introduces the main definitions used in the remainder of the paper.