Artificial Intelligence in Medicine 65 (2015) 167–177

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

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

On local anomaly detection and analysis for clinical pathways

Zhengxing Huang[a,*], Wei Dong[b], Lei Ji[c], Liangying Yin[a], Huilong Duan[a]

[a] College of Biomedical Engineering and Instrument Science, Zhejiang University, Zhou Yiqing Building 510, Zheda Road 38#, 310008 Hangzhou, Zhejiang, China
[b] Department of Cardiology, Chinese PLA General Hospital, Fuxing Road 28#, 100853 Beijing, China
[c] IT Department, Chinese PLA General Hospital, Fuxing Road 28#, 100853 Beijing, China

ARTICLE INFO

Article history:
Received 4 December 2014
Received in revised form 2 July 2015
Accepted 2 September 2015

Keywords:
Local anomaly detection
Probabilistic topicmodel
Clinical event log
Clinical pathway analysis

ABSTRACT

Objective: Anomaly detection, as an imperative task for clinical pathway (CP) analysis and improvement, can provide useful and actionable knowledge of interest to clinical experts to be potentially exploited. Existing studies mainly focused on the detection of global anomalous inpatient traces of CPs using the similarity measures in a structured manner, which brings order in the chaos of CPs, may decline the accuracy of similarity measure between inpatient traces, and may distort the efficiency of anomaly detection. In addition, local anomalies that exist in some subsegments of events or behaviors in inpatient traces are easily overlooked by existing approaches since they are designed for detecting global or large anomalies.
Method: In this study, we employ a probabilistic topic model to discover underlying treatment patterns, and assume any significant unexplainable deviations from the normal behaviors surmised by the derived patterns are strongly correlated with anomalous behaviours. In this way, we can figure out the detailed local abnormal behaviors and the associations between these anomalies such that diagnostic information on local anomalies can be provided.
Results: The proposed approach is evaluated via a clinical data-set, including 2954 unstable angina patient traces and 483,349 clinical events, extracted from a Chinese hospital. Using the proposed method, local anomalies are detected from the log. In addition, the identified associations between the detected local anomalies are derived from the log, which lead to clinical concern on the reason resulting in these anomalies in CPs. The correctness of the proposed approach has been evaluated by three experience cardiologists of the hospital. For four types of local anomalies (i.e., unexpected events, early events, delay events, and absent events), the proposed approach achieves 94%, 71% 77%, and 93.2% in terms of recall. This is quite remarkable as we do not use a prior knowledge.
Conclusion: Substantial experimental results show that the proposed approach can effectively detect local anomalies in CPs, and also provide diagnostic information on the detected anomalies in an informative manner.

1. Introduction

Clinical pathways (CPs) have been widely adopted by healthcare organizations to standardize clinical workflow, normalize treatment behaviors, reduce variations, decrease costs and improve quality, etc. [1,2]. Unfortunately, patient traces, being affected by many factors and uncertainties, are unpredictable by nature and thus may not go well towards the expected direction, and they have deviations with the predefined CPs. In fact, the uncertainties, resulting from inter-observer variability, inaccurate evaluation of the patient and some deficiencies in grading scales [3,4], make the organization of clinical work corresponding challenging, and lead to breakdowns such as delays, waste of resources and cancelations on the day of surgery, etc. [1]. In such cases, anomalies happen inevitably.

As shown in Fig. 1, there are two types of anomalies in CPs. An anomalous patient trace of CPs, which is largely deviated from normal ones, called global or large anomaly. The others are local anomalies, which exist in some subsegments of clinical events or treatment behaviors in patient traces of CPs [5]. Local anomalies provide feedback on patient traces that do not conform to the normative CPs and quantify any treatment behavioral deviation. In general, local anomalies are extraneous in the global anomalous patient traces but are consistently absent from the normal patient traces. They may occur occasionally in regular patient traces, although in which normal treatment behaviors are

* Corresponding author. Tel.: +86 13819488552.
E-mail address: zhengxinghuang@zju.edu.cn (Z. Huang).

http://dx.doi.org/10.1016/j.artmed.2015.09.001
0933-3657/

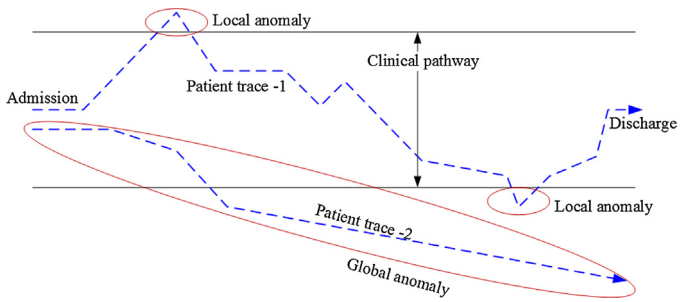

Fig. 1. Global and local anomalies in clinical pathways.

frequently and consistently present. For CP analysis, it is imperative to detect potential local anomalies and their causal dependencies from observations in a maximally-informative manner, to provide detailed feedback on non-compliance and support root cause analysis. Once these anomalies are detected, and their causal dependencies are derived, they can be leveraged as a base line to treatment compliance measurement for CPs, and a health-care organization can either update its CP specifications to cover the respective case, or it can improve new mechanisms to enforce best clinical practice execution. In this way, local anomaly detection and analysis are a central piece in the puzzle of advancing health-care organizations toward a higher degree of CP maturity.

To the best of our knowledge, most of existing studies are usually developed for detecting global anomalous patient traces in a structured manner, e.g., sequence alignment [6]. In clinical environments, the diversity of treatment behaviors in CPs is far higher than that of common business processes. CPs, as typical human-centric processes, always take place in a loosely structured manner. Bringing order in the chaos of CPs probably requires different anomaly detection strategies rather than existing methods. In addition, a patient trace may be quite long and complicated; however, anomalies usually occur in local subsegments. Therefore, it is understandable that local anomalies cannot be targeted by existing methods, which are designed for the whole patient trace, and may ignore the local anomalies existing in some subsegments of clinical events or treatment behaviors in CPs. To this end, we propose in this paper a probabilistic approach for anomaly detection and analysis in CPs. In detail, we employ a probabilistic topic model, i.e., clinical pathway model (CPM) [7], to disclose essential features of CPs from clinical event logs, and thus provide an accurate description of CPs by combining different classes of distributions. More specifically, the derived CPM recognizes patient traces as a probabilistic combination of underlying treatment patterns, and describes treatment patterns as a probabilistic combination of various clinical events [7], which aids to detect local anomalies. Furthermore, we present a diagnostic method to detect the causal dependencies between local anomalies so as to provide detailed feedback on anomalies and support root cause analysis in a maximally informative manner. The proposed approach is evaluated via a real-world data set collected from Chinese PLA General hospital.

The paper is structured as follows. A brief overview of relevant related works is provided in Section 2. Section 3 presents preliminaries for our investigations. In Section 4, we present our approach of local anomaly detection and analysis in CPs. In detail, we, firstly, introduce how to generate a CPM from the collected clinical event log. And then, we explain how the generated CPM can be used for local anomaly detection. Furthermore, we introduce an approach of deriving diagnostic information on detected anomalies. Section 5 presents findings from our validation, for which we implemented a prototype and tested it on a real-world clinical event log. Section 6 discusses contributions, limitations and possible extensions of our approach. Finally, Section 7 concludes the paper and identifies topics for future work.

2. Related work

Many efforts have been expended on the CP anomaly detection and analysis. CP anomalies are usually conceptualized in terms of clinical outcomes [8,9], and a well-defined body of knowledge has developed that takes this perspective [8,10,11]. In clinical practice, for example, length of stay (LOS), mortality, and infection rate, etc., are commonly used measures of CP anomalies [12–14]. As valuable as these measures are, they restrict the attention to the analysis of CP anomalies from the external perspective. In medical informatics, there is an emerging attention to use data mining and machine learning technologies to acquire insights into CPs. These techniques are also called process mining [15–17], which has been widely studied in the domain of business process management. Process mining techniques attempt to extract non-trivial and useful information from event logs. Based on process execution data, with its logic and reasoning ability, process mining guarantees integrity, objectivity and universality of the discovered process patterns and knowledge.

As a clearly important subject to the development of more efficient and effective business process management, recent researches in the process mining domain have been addressing the problem of identifying abnormal behaviors in business processes. For example, Aalst and Medeiros [18] present a α-algorithm-based anomaly detection method that has been implemented in a well-known process mining tool, i.e., ProM [19]. In [20], Bezerra and Wainer systematically present how ProM tool can support anomaly detection in event logs for business process management. In particular, they discussed four algorithms for detecting anomalies in logs of process aware systems [21], and validated their approach with a real case.

The idea of using process mining to detect and analyze CP anomalies is based on an assumption that we can collect clinical event logs from various hospital information systems, which regularly record normal and abnormal treatment behaviors in CPs [6]. The approaches of process mining-based CP anomaly detection usually first extract one or several consensus models taking place in patient traces, classify a new patient trace to one of the discovered CP models, and then automatically detect if it is anomalous with respect to the general characteristics of its membership model. For example, Bouarfa and Dankelman proposed a process mining algorithm to derive a consensus model from a clinical event log, based on which CP anomalies can be detected automatically and without prior knowledge from clinical experts [6].

Although most process mining algorithms can detect anomalous process traces in a structured manner [22,9], the assumption that the processes take place in a structured fashion is not valid for CPs. In clinical settings, the complexity and diversity of treatment behaviors in CPs are far higher than that of common business processes [17,23]. In fact, many treatment behaviors can occur arbitrarily without a particular order in clinical practice. Detecting local anomalies for loosely-structured CPs probably requires a different mining strategy rather than existing process mining algorithms.

3. Preliminary

Our approach for local CP anomaly detection and analysis is based on clinical event logs extracted from hospital information systems. First, we clarify some concepts and their data representation, including clinical event, patient trace, and clinical event log.

Definition 1 *(Clinical event)*. Let A be the treatment activity domain, and T the time domain. A clinical event e is represented as $e = (a, t)$, where a is the clinical activity type of $e (a \in A)$, and t is