Mining Conflict Semantic from Drug Dataset for Detecting Drug Conflict

Shunxiang Zhang, Anhui University of Science and Technology, Huainan, China
Guangli Zhu, Anhui University of Science and Technology, Huainan, China
Haiyan Chen, East China University of Political Science and Law, Shanghai, China
Dayu Yang, Anhui University of Science and Technology, Huainan, China

ABSTRACT

The detecting of drug interactions hiding in the massive drug data, especially the conflict (i.e., some drugs react with each other) detecting, plays an important role in the medical information field. This kind of conflict detecting can not only relieve the cognitive burden for doctors, but also help some people (e.g., physicians and patients etc.) avoid the risk of reactions among drugs in some extend. This paper presents a Drug Conflict Detecting (DCD) algorithm to rapidly find reactions among several drugs according to the user’s query requirements. First, the user dictionary and waste words base are built according the data feature of medical data sources to effectively extract drug term including component and interaction terms. Then, all conflict semantics are mined to establish conflict knowledge base based on the results of drug term extraction. Finally, the DCD algorithm is proposed to provide rapid detection of drug conflict. The experimental results show that the proposed algorithm has high accuracy. It can effectively and rapidly implement the drug conflict detecting.

Keywords: Conflict Knowledge Base, Conflict Semantic, Drug Conflict Detecting, Drug Term Extraction, Incompatibility Conflict

1. INTRODUCTION

With the rapid developing of information technology, various forms of drugs data are also developing rapidly on the Web. Vast amounts of drugs data are stored in their specific organization way. It is difficult for users to find reasonable information in the large amounts of drugs data resource according to their requirements. For example, people want to know the information of drug interactions including conflict reaction to make a reasonable choice in selecting some drugs. Actually, detecting and finding this kind of conflict reaction information is very necessary. “Capital of Chinese Medicine” magazine has investigated some prescriptions from more than 50 hospitals whose grade is above second-class in Beijing. There are obvious conflicts in the three types of prescription accounted for 94% (Liu, 2006). This investigation shows that

DOI: 10.4018/IJCINI.2015070105
professional people such as doctors, physicians etc., also difficultly remember a large of the conflict reaction among drugs.

Therefore, it is necessary to develop a highly intelligent system to provide conflict knowledge/semantic search and detecting. Actually, some existing works can provide theoretical and technical support for this kind of intelligent system such as Inference algebra (IA) (Wang, 2012), cognitive computing (Wang, 2012), knowledge modeling and processing tool (Wang, et. al. 2011), semantic mining (Xu, 2014) and so on. At the same time, this kind of intelligent systems has widely developed for many fields such as the system of semantic news events (Xu, 2015), online social network (Caviglione, 2014; Chen, 2014), meeting transcripts (Liu, 2011) and video organizing (Xu, 2015). Obviously, this kind of conflict search and detecting can reduce the cognitive burden of some people such as doctors, physicians and patients. This work belongs to the new development branch of cognitive informatics (Wang, 2011; Wang, et. al. 2011). At the same time, a favorable algorithm can rapidly and exactly provide the conflict information. Thus, people can select drugs healthily and rationally. However, there are so many types of drugs and so many conflict reactions exist in these drugs. How to mine the conflict semantic from the large medical data is a difficult problem.

To solve this problem, this paper presents a Drug Conflict Detecting (DCD) algorithm to support the conflict information detection among several drugs for users. To support the drug conflict detection, the extraction of drug terms, mining of conflict semantic and the detecting algorithm of drug conflict are indispensable. Our work mainly includes three aspects. The first part is to build the user dictionary and waste words base, which are based on the data feature of the medical data sources. Chinese word segmentation technology (Wu, 2008) and our prior window-split idea (Zhang, 2015) can be used to extract component and interaction terms more accurately. The second part is mining the conflict semantic which is represented by conflict rule. The two elements of a conflict rule are component terms. And the conflict relation of a conflict rule is an interaction term. All conflict semantics are made up of conflict knowledge base. The third part is to design the DCD algorithm, which can be used to rapidly inquire drugs conflict from conflict knowledge base for some people such as physicians and patients etc.

The rest of this paper is organized as follows: Section 2 presents related work about term extraction and the construction of knowledge base. Section 3 introduces the drug data processing and term extraction. Section 4 introduces how to mine the conflict semantic and construct conflict knowledge base. Section 5 proposes the Drug Conflict Detecting (DCD) algorithm. Section 6 gives some experimental analysis in detail. Conclusions are given in Section 7.

2. RELATED WORK

In this section, two related aspects of the existed work are given. Some works about term extraction are introduced first. Then, the related work about the construction of knowledge base is given.

2.1. Term Extraction

Conrado et al. have proposed a machine learning approach to automatic term extraction using a rich feature set (Conrado, 2013). Nazar R. firstly proposed the statistical approach to term extraction, which has introduced the concept of mutual information to evaluate the combined capacity of the two words at the same time (Nazar, 2011). Liang et al proposed a hybrid strategy for high precision long term extraction (Liang, 2009). Wang has introduced the word automatically rich-keywords extracted from metadata and its positioning (Wang, 2005). Guang et al. presented a keyword extraction method based on the lexical chain for Chinese information processing
Some Remarks on the Concept of Approximations from the View of Knowledge Engineering
Tsau Young Lin, Rushin Barot and Shusaku Tsumoto (2012). Developments in Natural Intelligence Research and Knowledge Engineering: Advancing Applications (pp. 278-286).
www.igi-global.com/chapter/some-remarks-concept-approimations-view/66454?camid=4v1a

Robust Feature Vector Set Using Higher Order Autocorrelation Coefficients
www.igi-global.com/article/robust-feature-vector-set-using/49691?camid=4v1a