Towards On-Line ADR Detection: An Innovative Data Cube-based Approach

Wen-Yang Lin, Jhih-Wei Du

Department of Computer Science and Information Engineering, National University of Kaohsiung, Taiwan

Abstract and Objective

The problem of Adverse Drug Reactions (ADRs) detection is a key issue in the pharmacovigilance community. Although many related researches have been proposed in recent years, majority of them were concerned about the accuracy of the detected adverse drug reactions, very few paying attention to the performance issue of how to quickly accomplish the detection process. In this work, we propose the concept of multi-valued-dimension contingency cube (MDC-cube) to facilitate multidimensional, fast and online detection of ADRs, especially those caused by drug interactions. Experiments conducted on the FDA AERS data set show that our MDC-cube based method is an order of magnitude faster than the state-of-the-art ABCM-MS method.

Keywords:

Adverse drug reaction, Data cube, Drug safety, Pharmacovigilance

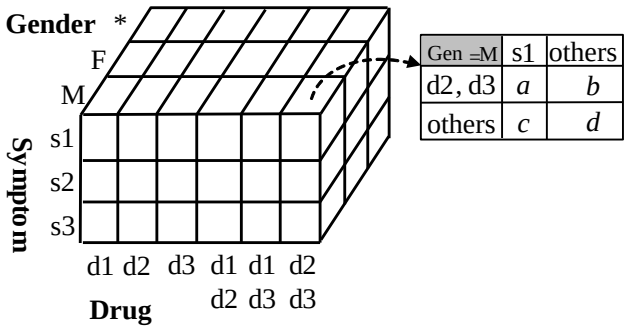
Introduction

Adverse Drug Reaction (ADR) refers to uncomfortable and/or harmful side effects yielded by normal doses of drug usage. Since many ADRs are only seen in long term, post-marketing surveillance of drug usage, spontaneous reporting systems (SRSs) have been established to collect suspected adverse drug events for further analysis, e.g., the USA FDA Adverse Event Reporting System (FAERS).

The constant and fast increasing in the amount of reports accumulated by SRSs has precluded the possibility for human analysis to detect ADR signals, motivating the development of many statistical- or data-mining-based algorithms for ADR detection. Unfortunately, contemporary methods for ADR detection were developed aiming at improving the signal quality, very few paying attention to the performance aspect. This not only lengthens the process of ADR detection, but also hinders the analysts from performing interactive, stratified analysis, i.e., inspecting from many different aspects and demographic conditions, such as gender, age, etc.

Methods

Inspired by the success of data cube technique to facilitate OLAP analysis widely used in business intelligence, we propose a new type of data cube, namely multivalued-dimension contingency cube (MDC-cube), which embodies the concept of multivalued dimension, i.e., the member of a dimension can be a set of items (multivalues) and stores in each cube cell a contingency table essential for measuring the disproportionality (e.g., PRR, ROR) of an ADR signal. The following figure illustrates a MDC-cube composed of three dimensions, Gender, Symptom, and Drug, where the cube cell [M, {d2, d3}, s1] stores the corresponding contingency table used for computing the measurement of the ADR signal: Gender = Male, Drug = {d2, d3} → Symptom = s1, e.g., PRR = $a(c + d) / [c(a + b)]$. In this way, the computation of each signal can be generated on the fly. Note that the number of dimensions (variables) involved in a MDC-cube can be more than three, yet all dimensions other than Drug and Symptom are optional.



Results

We have compared our MDC-cube based method with an associative classification based approach, ABCM-MS, which employs a tree structure similar to the FP-tree structure used by the leading FP-Growth algorithm for associations mining. We tested both methods over the 2007 FAERS dataset, which contains 140183 reports, composed of 9644 drugs and 7689 symptoms. The evaluation was inspected from two aspects, the number of predicates, indicating the query complexity, and the data size. Specifically, we consider two extremes of query complexity, one with only one predicate (demographic) attribute and the other with all predicate attributes. The results show that our MDC method is an order of magnitude faster than ABCM-MS, with speedup ranging from 10 ~ 30.

Conclusion

In this paper, we have presented our proposed new type of data cube, MDC-cube, which can facilitate an OLAP-like online detection of ADR signals from huge volume of SRS dataset. This method has been implemented to enhance the performance of our developed iADRs system, a web-based interactive system for ADR detection. We currently are pursuing ways to extend this cube model to cope with the uncertainty issue, e.g., missing values.