FraudAuditor A Visual Analytics Approach for Collusive Fraud in Health Insurance

ABSTRACT

Collusive fraud, in which multiple fraudsters collude to defraud health insurance funds, threatens the operation of the healthcare system. However, existing statistical and machine learning-based methods have limited ability to detect fraud in the scenario of health insurance due to the high similarity of fraudulent behaviors to normal medical visits and the lack of labeled data. To ensure the accuracy of the detection results, expert knowledge needs to be integrated with the fraud detection process. By working closely with health insurance audit experts, we propose FraudAuditor, a three-stage visual analytics approach to collusive fraud detection in health insurance. Specifically, we first allow users to interactively construct a co-visit network to holistically model the visit relationships of different patients. Second, an improved community detection algorithm that considers the strength of fraud likelihood is designed to detect suspicious fraudulent groups. Finally, through our visual interface, users can compare, investigate, and verify suspicious patient behavior with tailored visualizations that support different time scales. We conducted case studies in a real-world healthcare scenario, i.e., to help locate the actual fraud group and exclude the false positive group. The results and expert feedback proved the effectiveness and usability of the approach.

**EXISTING SYSTEM**

Akoglu et al. [2] extracted structural features, such as node degree or centrality, from the graph to find egonets. SpamCom [3] identified spammer communities on Twitter by using structure and attribute features such as Twitter content similarity, user topology, and user profile. In healthcare scenarios, Chen et al. [5] applied a spectrum analysis-based community detection method to detect patient referral fraud cases from a bipartite graph of physicians and specialists. Zhao et al. [13] generated a dynamic heterogeneous information network containing patients, hospitals, and diseases. Then, they identified anomalies that fit predefined fraud patterns (e.g., the high-cost single treatment) over fixedor variable periods. Statistics-based methods can produceinitial fraud candidates but may have erroneous results,requiring further validation by experts.

Xu et al. [7] propose GRC, a novel GNN model, that learns representations of different types of individuals and detects loan fraud by using attention mechanisms and imposing conditional random fields. However, these ML methods are supervised or semi-supervised and thus require fraud-labeled data, which is lacking in our health insurance scenario.

Niu et al. [4] use a node-link diagram to demonstrate the loan guarantee network, where each node belongs to a community defined by a random walk algorithm and is encoded with the corresponding color. In order to identify collective anomalies, Tao et al. [25] proposed a high-order correlation graph to support analysis processes starting with an abnormal node.Corresponding nodes that contribute to the anomaly can be easily identified through the high-order correlation graph. Our system incorporates graph and sequence visualization. To focus on collusive fraud in health insurance scenarios, our system provides richer contextual information, such as disease, drugs, and visit frequency.

Disadvantages

* + In the existing work, the system did not implement suspicious groups identification.
  + This system is less performance due to lack of Graph Neural Network.

**PROPOSED SYSTEM**

The system proposes a novel visual analytics approach to help health insurance audit experts identify suspicious groups, investigate the visit behavior of suspicious patients, and validate collusive fraud results. We propose a co-visit network to represent the relationship among patients. The weights of the edges are calculated by extracting the characteristics of collusive fraudsters, such as the time gap and number of visits. Suspicious groups with multiple simultaneous visits to the same location can then be identified by a weighted community detection algorithm.

The algorithm is integrated into a prototype system, Fraud Auditor, that supports experts in interactively browsing and improving model detection results. FraudAuditor can help experts quickly locate and examine fraud by observing co-visit links in visualizations of patient medical behavior.

Combined with contextual information such as disease, drug, and fee information, false positive groups can be verified and excluded. We provide case studies and expert interviews in real health insurance scenarios to validate the effectiveness of the proposed approach.

**Advantages**

A problem characterization that summarizes the requirements of collusive fraud detection in the scenario of health insurance.

A novel three-stage visual analytics approach to detect collusive fraud in health insurance that considers the visit pattern of fraud groups and expert knowledge.

An interactive prototype system, FraudAuditor, to facilitate the identification, examination, and validation of suspicious collusive fraud groups.

**SYSTEM REQUIREMENTS**

➢ **H/W System Configuration:-**

➢ Processor - Pentium –IV

➢ RAM - 4 GB (min)

➢ Hard Disk - 20 GB

➢ Key Board - Standard Windows Keyboard

➢ Mouse - Two or Three Button Mouse

➢ Monitor - SVGA

**SOFTWARE REQUIREMENTS:**

* **Operating system :** Windows 7 Ultimate.
* **Coding Language :** Python.
* **Front-End :** Python.
* **Back-End :** Django-ORM
* **Designing :** Html, css, javascript.
* **Data Base :** MySQL (WAMP Server).