Iranian Journal of Health Insurance Volume 1, Issue 4, Winter 2019

Research Article

Discovering Self-referral between Doctors and Pharmacists Based on Network Mining

Hassan Askarzadeh [1,*], Mohammad Jafar Tarokh [2]

[1] Department of Information Technology, Payame Noor University, Tehran, Iran
[2] Department of Information Technology, K.N. Toosi University of Technology, Tehran, Iran

*** Corresponding Author:** Hassan Askarzadeh, Department of Information Technology, Payame Noor University, Tehran, Iran. E-mail: haskarzadeh@pnu.ac.ir

Received: 31 Dec 2018 **Revised:** 17 Feb 2019 **Accepted:** 19 Feb 2019 **Online Published:** 15 Mar 2019

Abstract

Introduction: A significant amount of treatment cost is paid by health insurance organization. Insurance companies mostly use reliable people to audit documents, but due to the very high number of documents and the limitation of time and human resource, it is almost impossible to consider documents carefully and more importantly, some infringements are not identifiable according to only one document, but identifiable by accumulation of documents and intelligent analysis based on data mining. Detection of beneficial referral (self-referral and kickback) that a doctor refers a patient to a specific pharmacy that has benefits for him, is one of these things.

Methods: In the current study, data pool was prepared using Tehran Health Insurance data until 2017 and then after eliminating faulty data, according to network mining methods, actions were taken to detect anomalistic referrals on the network, data filtering, and weighing the edges of the network based on the views of reliable people. This method was implemented in the Knime environment and a short list was presented to monitoring department of the health insurance organization.

Results: In this research, according to the importance of detected interactions during network mining's process between doctors and pharmacies, and using visual tools in Knime, 73 doctors were detected that had meaningful relation with 26 pharmacies.

Conclusions: Inspectors of health insurance organization can have a more accurate and more effective examination with spending less time and human resource according to examination patterns based on network mining and visualization.

Keywords: Healthcare Insurance, Kickback, Self-Referral, Network Mining

Citation:
Askarzadeh H, Tarokh MJ. Discover self-referral between doctors and pharmacists based on network mining. Iran J Health Insur. 2019;1(4):153-158.

Copyright © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0) which permits copy and redistribute the material just in noncommercial usages, provided the original work is properly cite.

مقاله پژوهشی

کشف ارجاع منفعت‌طلبانه بین پزشک و داروخانه مبتنی بر شبکه‌کاوی

حسن عسکرزاده [۱،*]، محمدجعفر تارخ [۲]

[۱] عضو هیأت علمی گروه فناوری اطلاعات، گروه فناوری اطلاعات، دانشگاه پیام نور، تهران، ایران
[۲] گروه فناوری اطلاعات، دانشگاه خواجه نصیرالدین طوسی، تهران، ایران

* **نویسنده مسئول:** حسن عسکرزاده، عضو هیأت علمی گروه فناوری اطلاعات، گروه فناوری اطلاعات، دانشگاه پیام نور، تهران، ایران. ایمیل: haskarzadeh@pnu.ac.ir

تاریخ دریافت: ۱۳۹۷/۱۰/۱۰ **تاریخ تصحیح:** ۱۳۹۷/۱۱/۲۸ **تاریخ پذیرش:** ۱۳۹۷/۱۱/۳۰ **تاریخ انتشار آنلاین:** ۱۳۹۷/۱۲/۲۵

چکیده

مقدمه: سازمان بیمه سلامت، بخش مهمی از هزینه‌های درمان جامعه را پرداخت می‌کند. شرکت‌های بیمه عمدتاً از افراد خبره برای رسیدگی به اسناد بهره می‌برند که با توجه افزایش تعداد اسناد و محدودیت زمان و منابع انسانی، رسیدگی دقیق به اسناد تقریباً غیر ممکن است و از آن مهمتر بعضی از تخلفات از روی یک سند قابل شناسایی نیست، بلکه با تجمیع اسناد و تحلیل‌های هوشمند مبتنی بر داده‌کاوی قابل شناسایی است. کشف ارجاع منفعت‌طلبانه که در آن پزشک، بیمار را به داروخانه خاصی که در آن منافع دارد ارجاع می‌دهد، یکی از این موارد است.

روش بررسی: در این پژوهش با استفاده از داده‌های بیمه سلامت استان تهران تا سال ۱۳۹۶ هجری خورشیدی انباره‌داده تهیه شد و پس از حذف داده‌های پرخطا، بر اساس روش‌های شبکه‌کاوی نسبت به شناسایی ارجاعات غیرمتعارف در شبکه، فیلتر داده‌ها و وزن‌دهی یال‌های شبکه بر اساس نظر خبرگان اقدام شد. مدل فوق در محیط نرم‌افزار Knime پیاده‌سازی و لیستی از موارد مشکوک به بخش نظارت و رسیدگی ارایه گردید.

یافته‌ها: در این پژوهش بر اساس وزن تعاملات شناسایی شده در فرآیند شبکه‌کاوی بین پزشک و داروخانه و با استفاده از ابزار بصری ارایه شده در نرم‌افزار Knime تعداد ۷۳ پزشک که با ۲۶ داروخانه رابطه‌ی معناداری دارند شناسایی گردید.

نتیجه گیری: بازرسان سازمان بیمه سلامت می‌توانند بر اساس الگوهای رسیدگی مبتنی بر شبکه‌کاوی و بصری‌سازی با صرف زمان و منابع انسانی کمتر، رسیدگی دقیق و موثرتری را صورت دهند.

واژگان کلیدی: بیمه سلامت، ارجاع منفعت‌طلبانه، شبکه‌کاوی

مقدمه

هر چند مبالغ کشف تقلب در کشور ما اطلاع‌رسانی نمی‌شود اما می‌توان با آمار منتشر شده از سایر کشورها نسبتی معقول از حجم تقلب به دست آورد. Manyika و همکاران در گزارش موسسه McKinsey اظهار می‌دارند «از ۲۰۰۰ میلیارد دلار هزینه بهداشت و درمان در آمریکا، تقلب سهمی ۱۲۰ میلیارد دلاری دارد» [۱]. سهم تقلب در ایران با همین تناسب همه ساله بالغ بر ۵ درصد هزینه درمان برآورد می‌گردد و هزینه‌های تقلب سالانه بالغ بر ۵۰۰۰ میلیارد تومان برآورد می‌شود [۲]. یکی از موارد تقلب، ارجاع منفعت طلبانه است که به ارجاعاتی اطلاق می‌شود که در آن پزشک که مسئول تشخیص بیماری و هدایت فرآیند بهبودی بیمار است، بر اساس منافع مشترکی که با داروخانه یا آزمایشگاه یا مرکز درمانی دارد، اقدام به تجویز دارو یا آزمایش بعضاً غیر ضروری کرده و تاکید بر دریافت خدمات از داروخانه یا آزمایشگاه ذیربط داشته یا ادامه درمان بیمار را در یک مرکز درمانی که در آن منافعی دارند پیشنهاد می‌کند [۳]. همچنین ممکن است برای عدم خسارت به داروخانه، داروهایی که تاریخ انقضای آن نزدیک است توسط پزشک تجویز شود. شناسایی و رسیدگی به این نوع تقلب از هدر رفت منابع مالی نزد مردم و شرکت‌های بیمه جلوگیری می‌کند [۴].

نظام سلامت با توسعه سیستم‌ها و بخصوص سیستم نظارت بر عملکرد، می‌تواند بخشی مهمی از پشتیبانی از یکپارچگی نظام سلامت را ایجاد کند [۵]. برای اعمال نظارت بر عملکرد تعداد قابل ملاحظه‌ای پزشک و مراکز درمانی، چاره‌ای جز به کارگیری ابزارهای مبتنی بر داده کاوی نیست.

رفتار غیر متعارف پزشکان اعم از ویزیت تعداد زیادی از بیماران در زمانی کوتاه یا تجویز داروهای مغایر با وضعیت جنسی یا سنی بیمار به روش‌هایی مبتنی بر تحلیل اطلاعاتی یا شناسایی خوشه‌های غیر متعارف در روش‌های داده کاوی مبتنی بر خوشه‌بندی قابل شناسایی است [۶]؛ اما ارجاع منفعت‌طلبانه در روش‌های پیش گفته در خوشه‌های عادی طبقه‌بندی شده و قابل شناسایی نیست. علامت‌گذاری نسخ عادی بیمه‌شدگانی که در یک شبکه ارجاع منفعت‌طلبانه از قبل شناسایی شده‌اند باعث می‌شود شاخص جدیدی برای خوشه‌بندی پزشکان و متناسب با ریسک رفتار منفعت‌طلبانه آنان شناسایی شود که در واقع نشانگر احتمال وجود تقلب در نسخ عادی است. این گام در واقع مقدمه‌ای بر ایجاد شاخص جدید در خوشه‌بندی داده‌های درمان ایجاد می‌کند. ارجاع منفعت‌طلبانه، معرفی بیماران به ارائه‌دهنده

خدمات بهداشتی و درمانی خاص است. برای مثال، اگر یک پزشک منفعتی شخصی از یک کلینیک داشته باشد، نمی‌تواند هیچ بیماری را به آن کلینیک ارجاع دهد [۷]. در بعضی از کشورها ازجمله آمریکا قانونی برای مقابله با این امر وجود دارد. در ایران اشتراک منافع پزشکان با داروخانه‌ها و آزمایشگاه‌ها و بیمارستان‌ها به‌تناسب قرارداد سازمان‌های بیمه‌گر ممکن است با جرائمی همراه باشد.

به‌صورت خلاصه، قانون مقابله با ارجاع منفعت‌طلبانه زمانی نقض می‌شود که ارائه‌دهنده خدمات بهداشتی و درمانی بیماران را به جایی که ارتباط مالی با آن دارد ارجاع دهد. این معرفی‌ها توسط قوانین ضد ارجاع منفعت‌طلبانه ممنوع شده‌اند و در صورت رخ دادن تقلب محسوب می‌شوند [۸]. لذا این پژوهش با هدف کشف ارجاع منفعت‌طلبانه بین پزشک و داروخانه مبتنی بر شبکه‌کاوی در سال ۱۳۹۶ صورت پذیرفت.

روش بررسی

داده‌های پایه پژوهش

اطلاعات مرجع این پژوهش شامل اطلاعات نسخ درمانی ۹۲۴,۰۳۸ بیمار است که طی ۱,۰۵۴,۶۹۴ مراجعه به ۵۰,۰۳۳ پزشک حاصل‌شده است. داروهای تجویزشده از این نسخ از ۷۹۹ داروخانه دریافت شده‌اند. این اطلاعات طی فرآیندی پاک‌سازی و به یک انباره داده در محیط Oracleمنتقل و پس از ایجاد مدلی از پایگاه داده‌ای تحلیلی به نام Qubeبا محوریت پزشکان، خروجی جهت پردازش و مدل‌سازی به فرمت فایل Excel ذخیره گردیده است.

دانش پایه بکار رفته در مدل

با توجه به عدم وجود اطلاعات قبلی از این تخلف، دانش مدل، از تجربه خبرگان به دست آمد. بر این اساس از ۱۸ نفر افراد خبره در خصوص رسیدگی به اسناد سازمان‌های بیمه‌گر نظرسنجی و مصاحبه شد و اطلاعات زیر بدست آمد:

۱. به دلیل پایین بودن احتمال بهره‌مندی پزشک از منافع داروخانه دولتی نیازی به بررسی این داروخانه‌ها نیست.

۲. در صورتیکه درآمد سالانه داروخانه از ارجاعات یک پزشک کمتر از ۱,۰۰۰,۰۰۰ تومان در سال بود، احتمال وجود رابطه منفعت‌طلبانه ناچیز است.

۳. سهم هر پزشک از درآمد داروخانه یک ویژگی مهم در شناسایی ارجاع منفعت‌طلبانه است.

۴. درآمد داروخانه از هر پزشک نسبت به میانگین درآمد از همه پزشکان مرتبط با آن داروخانه، یک ویژگی مهم در شناسایی ارجاع منفعت‌طلبانه است.

ابزار مدل‌سازی، تحلیل و نمایش داده‌ها

نرم‌افزار مورداستفاده جهت مدل‌سازی، نرم‌افزار متن‌باز Knime است [۹]. مزیت استفاده از این نرم‌افزار در مقابل SPSS Modeler یا RapidMinerمتن‌باز بودن و عدم محدودیت در حجم داده‌های مورد پردازش است. ضمن اینکه برای تغییر در یک المان در مدل نیازی به محاسبه بسیار زمان‌بر همه مدل نیست. درواقع ویژگی منحصربه‌فرد Knimeنگهداری داده‌های پردازش‌شده از هر نود در مدل است که تا وقتی مبنای ورودی‌ها و تنظیمات آن تغییری نکند، دوباره محاسبه نخواهد شد. این ویژگی بخصوص در کار با پردازش حجم زیاد داده‌ها مزیت بزرگی در سرعت بخشیدن به تغییر مدل و مشاهده نتایج به دست می‌دهد.

تشریح مدل

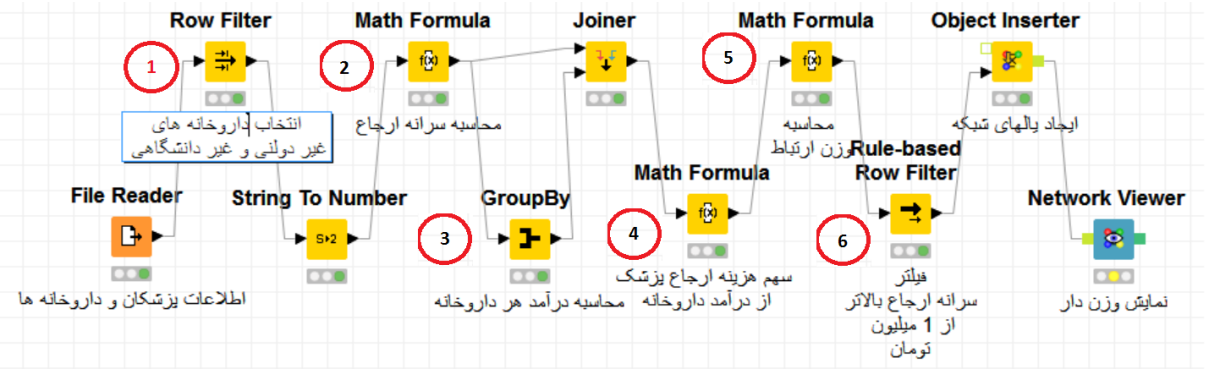
گام اول برای بررسی ارتباط ویژه بین پزشک و داروخانه آن است که بر اساس اطلاعات نسخ درمانی و روش‌های مبتنی بر کاوش در شبکه، از بین هزاران پزشک و صدها داروخانه، آن‌هایی که ارتباط معنی‌داری با هم دارند جداسازی شود.

برای تسهیل در امر پردازش داده‌ها، ابتدا با اتصال به پایگاه داده اسناد پزشکی سازمان بیمه سلامت در محیط اوراکل و ایجاد انباره داده و پس از ایجاد Qube پزشکان، با استفاده از نرم‌افزار Knime مدل شکل ۱ پیاده‌سازی گردید. فیلترهای اعمال شده بر داده‌ها و وزن‌های مرتبط با هر اتصال در شبکه بر اساس نظر خبرگان تعیین گردیده است.

ابتدا اطلاعات پزشکان و داروخانه‌ها خوانده و داروخانه‌های دانشگاهی و دولتی بر اساس نظر خبرگان از آن حذف گردید. در گام بعدی سرانه ارجاع بین پزشک و داروخانه محاسبه و به عنوان یک ویژگی به داده‌ها اضافه گردید. سپس درآمد هر داروخانه محاسبه گردید تا بتوان سهم هزینه ارجاع هر پزشک از درآمد داروخانه را محاسبه کرد. ویژگی محاسباتی فوق به سایر ویژگی‌ها اضافه گردید. در گام بعدی سهم هر پزشک از درآمد داروخانه محاسبه و سپس وزن هر یال در شبکه که از فرمول جمع هزینه ارجاع هر پزشک تقسیم بر سرانه ارجاع، محاسبه گردیده است به دست آمد. درواقع این محاسبه نشان می‌دهد ارتباط مالی هر پزشک و داروخانه چند برابر میانگین هزینه سایر پزشکان بوده است. در گام بعدی بر اساس نظر خبرگان ارجاعات با جمع مبلغ کمتر از ۱,۰۰۰,۰۰۰ تومان حذف و اعداد سرانه بالاتر از این مقدار فیلتر و به شبکه وارد شد. در شکل ۱ مدل پیاده‌سازی شده ارائه و هر گام در آن مستندسازی شده است.

یافته‌ها

فرضیه این پژوهش، امکان‌پذیر بودن شناسایی الگوهایی از تعاملات شبکه‌ای بین پزشک و داروخانه توسط مدل ارائه شده است که بیشترین احتمال وجود ارتباط منفعت‌طلبانه بین پزشک و داروخانه در آن وجود داشته باشد. بر اساس این پژوهش که مبتنی بر دانش اولیه تدوین‌شده توسط خبرگان است، به ورودی مدل پایگاه داده‌ای با ۷۹۹ داروخانه و ۵۰,۰۳۳ پزشک با ۴۱۸۷۴۸ ارتباط بین آنها، تزریق‌شده است که پس از اجرای مدل تنها شبکه‌ای متشکل از ۲۶ داروخانه و ۷۴ پزشک و ۷۴۴ ارتباط بین آن‌ها به‌عنوان شبکه‌ای با بیشترین احتمال وجود ارجاع متقلبانه شناسایی گردیده است. برای درک بهتر نتیجه پژوهش اگر فرض کنیم برای بررسی رابطه منفعت‌طلبانه مالی یک داروخانه با یک پزشک حداقل ۱۰ دقیقه زمان صرف شود؛ این پژوهش باعث صرفه‌جویی زمانی معادل ۶۹۶۶۷ ساعت فعالیت کارشناسی خواهد شد اگر هرروز کاری معادل ۸ ساعت و هر سال بدون احتساب روزهای جمعه و تعطیلات، ۲۸۸ روز کاری در نظر گرفته شود، ۳۰ سال زمان و حقوق یک کارشناس صرفه‌جویی خواهد شد.



شکل ۱. مدل پیاده شده در محیط Knime

برای محاسبه یال‌های شبکه در مدل فوق در مرحله ۲، سرانه ارجاع پزشک به داروخانه از فرمول ۱ محاسبه شده است که در آن سرانه ارجاع پزشک i به داروخانه j و جمع هزینه نسخ بیماران ارجاع داده شده از پزشک i به داروخانه j و تعداد بیمارانی است که نسخ تجویزشده پزشک i به آنان در داروخانه j پیچیده شده است. سپس در مرحله ۳ درآمد هر داروخانه از فرمول ۲ محاسبه شده است که در آن درآمد داروخانه‌ی j از همه نسخ پزشکانی است که نسخ آنان توسط داروخانه jپیچیده شده است. سپس در مرحله ۴ سرانه سهم ارجاع هر پزشک به داروخانه از فرمول ۳ محاسبه شده و در انتها برای محاسبه وزن‌های شبکه در مرحله ۶ وزن ارتباطات شبکه‌ای هر داروخانه با پزشک در شبکه از فرمول ۴ استفاده می‌گردد.

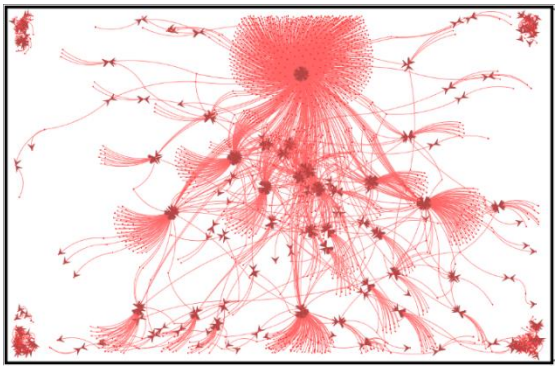
فرمول ۱ $$RPC_{j,i}=\frac{C_i}{N_i}$$

فرمول ۲ $$Pr_j=\sum_{i=1}^{N} C_i$$

فرمول ۳ $$Sprd_{j,i}=\frac{C_i}{Pr_j}$$

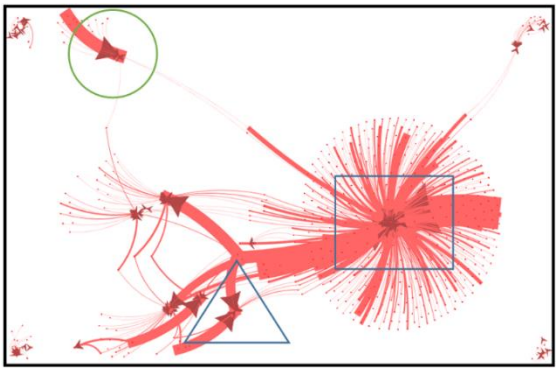
فرمول ۴ $$W_{j,i}=\frac{C_i}{Max(Sprd_{j,i})}$$

شکل ۲ گراف حاصل از اعمال داده‌های قبل از پردازش به شبکه است و چنانکه در این شکل نیز مشاهده می‌شود نمی‌توان موارد مشکوک را از میان هزاران گره و یال موجود در شبکه تمیز داد.



شکل ۲. شبکه ارتباطی پزشکان و داروخانه‌ها قبل از اعمال وزن‌ها و فیلترها

با اجرای مدل، شکل ۳ حاصل می‌شود که به‌وضوح یال‌های پررنگ متناسب با قطر آن نشان‌دهنده موارد مشکوک به ارجاع منفعت‌طلبانه هستند. در اشکال فوق، برای حفظ محرمانگی و اخلاق پژوهشی، برچسب گره‌ها که به نام پزشکان و داروخانه‌ها اشاره می‌کند همچنین برچسب یال‌های که به حجم تبادلات مالی اشاره می‌کند حذف شده است. قطر هر یال نشان‌دهنده حجم تعاملات مالی بین پزشک و داروخانه است.



شکل ۳. شبکه ارتباطی پزشکان و داروخانه‌ها پس از اعمال وزن‌ها و فیلتر یال‌های کم‌ارزش

بحث

اهمیت مداوم و در حال رشد اطلاعات با کمک توسعه سریع سیستم‌هایی که حجم زیادی از اطلاعات را جمع‌آوری و پردازش می‌کنند، از نظر پردازش و تجزیه‌وتحلیل اطلاعات، یک مشکل بزرگ است. پاسخ به نیازهای فعلی و آینده بازار، یک انبار داده است که از طریق فرایند استخراج داده‌ها از سامانه‌های مختلف پشتیبانی می‌شود. فرآیند انتقال داده‌ها از سامانه‌های مختلف ETL (Extract, Transform, Load)نامیده می‌شود که شامل استخراج، پاک‌سازی و تبدیل و بارگذاری در انبار داده می‌گردد. پس از استقرار داده‌ها در انباره داده امکان پردازش بدون وقفه یا انجام تجزیه‌وتحلیل مهیا می‌گردد [۱۰]. شبکه کاوی مبتنی بر نظریه گراف است که شاخه‌ای از ریاضیات بوده و درباره گراف‌ها بحث می‌کند. این مبحث درواقع شاخه‌ای از توپولوژی است که با جبر و نظریه ماتریس‌ها پیوند مستحکم و تنگاتنگی دارد. نظریه گراف برخلاف شاخه‌های دیگر ریاضیات نقطه آغاز مشخصی دارد و آن انتشار مقاله‌ای از لئونارد اویلر، ریاضیدان سوئیسی، برای حل مسئله پل‌های کونیگسبرگ در سال ۱۷۳۶ میلادی است به‌کارگیری ریاضیات مرتبط با گراف در داده‌کاوی، بخش جدیدی از تحلیل‌های مبتنی بر داده‌های بزرگ را به نام شبکه کاوی به وجود آورده است [۱۱]. یک گراف که مشتمل بر تعدادی رأس یا گره و یال است می‌تواند روابط بین عناصر مختلف در یک شبکه را نمایش دهد.

یال‌ها می‌تواند جهت‌دار یا بدون جهت باشد، همچنین یال‌ها می‌تواند حاوی اطلاعاتی نظیر حجم تبادلات دو گره یا فقط نشان‌دهنده ارتباط دو گره باشد. با گسترش شبکه‌های اجتماعی، اهمیت تحلیل ارتباط تأثیرگذار از طریق آنالیزهای مرتبط با شبکه دوچندان شده است. Akoglu و Faloutsos در سال ۲۰۱۳ میلادی روشی را برای کشف ناهنجاری و تقلب در شبکه‌های بزرگ داده‌ای ارائه دادند. روش آن‌ها تلفیقی از یادگیری ماشین و داده‌کاوی با نظریه گراف بود. عمده‌ترین ضعف این روش نیازمندی آن به وجود موارد از قبل شناسایی‌شده برای اعمال در بخش یادگیرنده بود [۱۲]. Liu و Bier در سال ۲۰۱۶ میلادی با استفاده از تحلیل گراف در یک نرم‌افزار تجاری تولیدشده توسط شرکت Xerox توانستند تعاملات بین پزشکان، بیماران و داروسازانیکه مجوز تولید دارای مواد مخدر هستند را شناسایی کنند [۱۳]. وجود دو ویژگی مهم باعث موفقیت استفاده از الگوی فوق گردید. اولین ویژگی وجود فاصله مکانی بین بیمار، داروساز و پزشک در پایگاه داده مورداستفاده و دومین ویژگی وجود موارد متعدد تقلب گزارش‌شده بود. به دلیل عدم وجود دو ویژگی مهم فوق امکان به‌کارگیری الگوهای مورد استفاده از تحقیق فوق وجود ندارد. در آمریکا دو بسته بیمه‌ای مهم به نام مدیکیر و مدیکید که اولی برای افراد سالمند و دومی برای افراد کم‌بضاعت تدارک شده است اهدافی مانند طرح تحول سلامت در ایران را دنبال می‌کند [۱۴]. کمپانی Oracle نیز در سال ۲۰۱۶ میلادی جعبه‌ابزاری برای تحلیل تقلب در دو بسته‌ی بیمه‌ای فوق را ارائه کرده است [۱۵]. این بسته به نام Oracle PGX نیز بر اساس تئوری گراف و شبکه کاوی نسبت به شناسایی موارد مشکوک با روش‌های بدون نظارت اقدام می‌کند. بسته‌های بیمه‌ای مدیکیر و مدیکید همانند نظام ارجاع از طریق پزشک خانواده، بر اساس روال مشخصی تدوین‌شده که آزادی عمل در دریافت خدمات بین بیمار و پزشک و مرکز درمانی وجود ندارد؛ لذا کشف تقلب محدود به عوامل محدودتری است. با توجه به اینکه در نظام فعلی سلامت ایران، بیماران خود پزشک، داروساز، آزمایشگاه را انتخاب و پزشکان بدون دسترسی به سوابق آزمایش‌ها و داروهای بیماران، می‌توانند آزمایش‌های تکراری یا داروهای غیر مؤثر تجویز کنند، پایگاه داده‌ای مرتبط با پزشک و دارو و مراکز درمانی با آنتروپی یا بی‌نظمی بسیار زیاد همراه است و همین امر باعث می‌شود در فرآیند کشف الگو، منابع حافظه‌ای و پردازشی بسیار زیادی به کار گرفته شود. اجرای مدل ارائه‌شده در این پژوهش از شروع تا خاتمه کمتر از یک ساعت به طول انجامیده که برای این حجم از داده بسیار مطلوب به نظر می‌رسد.

پس از اجرای مدل پیشنهادی روی داده‌ها می‌توان ۳ الگو را مشاهده کرد که نیازمند تحلیل مجزا هستند. الگوی اول داروخانه‌ای که برای جمع نسخه‌های بالاتر از یک‌میلیون تومان از چندین پزشک ارجاع دارد اما سهم ارجاع از یک پزشک بسیار بالاتر از سایر پزشکان است و احتمال وجود ارجاع منفعت‌طلبانه بین آنان وجود دارد. دلیل دیگر این حجم از تعاملات می‌تواند نزدیکی موقعیت مکانی پزشک و داروخانه باشد. داروخانه متمایز شده در دایره سبز در شکل ۳ نمونه‌ای از این الگو است. الگوی دوم شامل داروخانه‌هایی است که از ۲ یا ۳ پزشک ارجاع ویژه دارند. این الگو می‌تواند حاصل ترکیبی از ارجاع منفعت‌طلبانه و نزدیکی موقعیت پزشک و داروخانه باشد. داروخانه متمایز شده توسط مثلث آبی در شکل ۳ نمونه‌ای از این الگو است. الگوی سوم شامل داروخانه‌هایی است که از چندین پزشک ارجاعات با هزینه داروی بالا دارند. دلیل ارجاع در این داروخانه‌ها علاوه بر ارجاع منفعت‌طلبانه و نزدیکی موقعیت مکانی داروخانه و پزشک، می‌تواند تخصصی بودن داروخانه نیز باشد. به‌عنوان مثال در صورتیکه داروخانه‌ای به‌صورت تخصصی داروهای هموفیلی ارائه کند، مقصد دریافت نسخ موردنظر از پزشکان متخصص زیادی است و با توجه به اینکه پزشکان خاصی نیز می‌توانند این داروها را برای بیماران تجویز کنند؛ بنابراین یال‌های پررنگی از تعداد محدودی پزشک به یک داروخانه به وجود می‌آید. نمونه متمایز شده در مربع آبی‌رنگ از این الگو پیروی می‌کند. با بررسی الگوهای مشابه، اغلب این موارد شامل داروخانه‌های تخصصی است. ممیزین و بازرسان با استفاده از این لیست کوتاه شده و با بررسی سایر استاد مرتبط موارد تخلف دیگری را شناسایی نمودند. به عنوان مثال بازرسان مواردی از پزشکان فاقد تخصص لازم در صدور داروی خاص را مشاهده نمودند.

با توجه به اینکه وقتی داروخانه و پزشکی در موردی خطاکار بود، احتمال بروز تقلب در سایر موارد نیز وجود دارد، وزن ارتباطات غیرعادی بین داروخانه و پزشک که از فرآیند شبکه کاوی حاصل‌شده به‌عنوان ویژگی جدید در فرآیند خوشه‌بندی اطلاعات سلامت می‌تواند مورداستفاده قرار گیرد. با اضافه شدن این ویژگی، امکان استخراج نسخی که در روش‌های معمول خوشه‌بندی عادی شناسایی می‌شوند، به‌عنوان موارد مشکوک قابل‌شناسایی می‌گردد. در پژوهش با عنوان "ارائه مدل کشف تقلب در بیمه سلامت" از این شاخص به‌عنوان نشانه اسناد مشکوک در تعامل بین داروخانه و پزشک استفاده شد.

نتیجه‌گیری

با توجه به حجم بالای اطلاعات داروهایی که توسط پزشک برای بیمار تجویزشده که داروی آن توسط داروخانه‌ها تأمین می‌شود، بررسی ارتباط داروخانه و پزشک بدون به‌کارگیری روش‌های پردازش اطلاعات حجیم غیرممکن است و همین امر باعث ایجاد زمینه سوءاستفاده پزشک و داروخانه می‌گردد. روش‌های مبتنی بر تحلیل و شبکه‌کاوی به ارزیابان بیمه این امکان را می‌دهد که به جای بررسی هزاران ارتباط بین پزشک و داروخانه، تنها مواردی را که به‌احتمال بیشتر همراه با تقلب است را بررسی نمایند. نتایج این پژوهش نشان می‌دهد، با بهره‌برداری از روش‌های مبتنی بر شبکه‌کاوی می‌توان نسبت به ارزیابی حجم زیادی از داده‌ها درزمانی کوتاه و با صرف منابع انسانی محدود اقدام نمود.

گام بعدی در این پژوهش می‌تواند با درج اطلاعات موقعیت مکانی پزشک و داروخانه در اطلاعات‌پایه و تأثیر دادن آن در تحلیل شبکه صورت پذیرد. همچنین در صورت ثبت و در اختیار قرار گرفتن نتایج بررسی‌های صورت گرفته روی نتایج این تحقیق، می‌توان از مدل‌های یادگیرنده یا نیمه نظارت‌شده بهره برده و مدل را بهبود بخشید.

پیشنهادها

پیشنهاد می‌شود سازمان بیمه سلامت یا سازمان تأمین اجتماعی با استقرار سامانه فوق در بخش رسیدگی به اسناد یا واحد بازرسی و نظارت سازمان و ارائه آموزش‌های لازم به ممیزین اسناد، به‌تدریج فرآیند رسیدگی به اسناد را به روش‌های مبتنی بر دانش تغییر دهد.

سپاسگزاری

به دلیل محدودیت‌های موجود در ارتباط محقق با بازرسان و تأخیر در دریافت نظر خبرگان سازمان بیمه سلامت، در این پژوهش از کمک کارشناسان رسیدگی اسناد پزشکی سایر سازمان‌های بیمه‌گر نیز استفاده شده است که بدین‌وسیله از آنان قدردانی می‌گردد. این مقاله بر اساس پژوهش سفارشی سازمان بیمه سلامت به شماره ۹۵/۸۷۲۶۰ مصوب ۹۵/۲/۲۷ تدوین گردیده است.

References

1. Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, et al. Big data: The next frontier for innovation, competition, and productivity. New York: McKinsey and Company; 2011.
2. Askarzadeh H, Tarokh MJ, editors. Discovery of health insurance fraud based on data mining. International Conference of Modern Research in Management, Economics and Tourism Industry Capability in Development; 2017; Mashad, Iran.
3. STARK LAW - Information on penalties, legal practices, latest news and advice ENVISAT 2018 [updated 2019; cited 2018]. Available from: http://starklaw.org.
4. Thornton D, Brinkhuis M, Amrit C, Aly R, editors. Categorizing and Describing the Types of Fraud in Healthcare ,Conference on ENTERprise Information Systems. International Conference on Project Management on Health and Social Care Information Systems and Technologies; 2015.
5. Seyedamini B, Riahi L, Mahmoudi-Majdabadifarahani M, Tabibi SJ, Masoudi-Asl I. Functional-Informational Factors Affecting Health System Integration. Health Inf Manage. 2018;15(3):99-105.
6. van Capelleveen G, Poel M, Mueller RM, Thornton D, van Hillegersberg J. Outlier detection in healthcare fraud: A case study in the Medicaid dental domain. Int J Account Info Sys. 2016;21:18-31. doi: 10.1016/j.accinf.2016.04.001
7. Oude Nijeweme-d'Hollosy W, van Velsen L, Poel M, Groothuis-Oudshoorn CGM, Soer R, Hermens H. Evaluation of three machine learning models for self-referral decision support on low back pain in primary care. Int J Med Inform. 2018;110:31-41. doi: 10.1016/j.ijmedinf.2017.11.010 pmid: 29331253
8. A Roadmap for New Physicians Fraud & Abuse Laws 2018 [updated 2019; cited 2018]. Available from: https://oig.hhs.gov/compliance/physician-education/01laws.asp.
9. Knime Software Features 2018 [updated 2019; cited 2018]. Available from: https://www.knime.com/knime-for-data-scientists.
10. Gorawski M, Gorawska A. Research on the Stream ETL Process. Beyond Databases, Architectures, and Structures Communications in Computer and Information Science. 4242014. p. 61-71.
11. Mashaghi AR, Ramezanpour A, Karimipour V. Investigation of a protein complex network. Europe Physic J. 2004;41(1):113-21. doi: 10.1140/epjb/e2004-00301-0
12. Akoglu L, Faloutsos C, editors. Anomaly, Event, and Fraud Detection in Large Network, Datasets. Proceedings of the sixth ACM international conference on Web search and data mining; 2013; Rome, Italy.
13. Liu J, Bier E, Wilson A, Guerra-Gomez JA, Honda T, Sricharan K, et al. Graph analysis for detecting fraud, waste, and abuse in healthcare data. AI Magazine. 2016;37(2):33-46.
14. The official U.S. government site for Medicare 2018. Available from: https://www.medicare.gov.
15. Oracle Labs, PGX 2, 3, 2 Documentation 2018 [updated 2019; cited 2018]. Available from: https://docs.oracle.com/cd/E56133_01/2.3.1/use-cases/healthcare-fraud/index.htm.