WWW.ICCSEEE2016.COM
5th International conference
On Computer Science , Electrical and Electronics Engineering
Malaysia – Kuala Lumpur
june 3-7, 2016

# تشخیص تقلب بیمه اتومبیل به‌وسیلهٔ الگوریتم ژنتیک و جستجوی پراکنده

**مژده معدّلی [۱]، حسن رشیدی [۲]**

[۱] دانشجوی کارشناسی ارشد دانشگاه آزاد اسلامی، واحد قزوین، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، قزوین، ایران
M.moaadeli@gmail.com

[۲] دانشیار دانشگاه آزاد اسلامی، واحد قزوین، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، قزوین، ایران hrashi@gmail.com

**چکیده**

با توجه به رشد روزافزون تقلب‌های صورت‌گرفته در صنعت بیمه در ایران، بیمه‌گذاران هرچه بیشتر درصدد یافتن روشی برای جلوگیری از این کلاهبرداری‌ها می‌باشند. مشکلی که صنعت بیمه در ایران با آن مواجه است عدم وجود پایگاه‌داده مناسب می‌باشد. در این مقاله ابتدا شاخص‌های جامع و کاربردی جدید مبتنی بر قوانین موجود در ایران برای تشخیص تقلب بیمه اتومبیل تعیین می‌شود. پس از آن، توسط تکنیک فرآیند تحلیل سلسله مراتبی، شاخص‌ها اولویت‌دهی می‌گردند و در الگوریتم پیشنهادی مورد استفاده قرار می‌گیرند. تاکنون الگوریتم‌های فراابتکاری کمتر مورد توجه محققان در این زمینه قرار گرفته‌است، لذا باتوجه به ویژگی‌های این الگوریتم‌ها، استفاده از ترکیب الگوریتم ژنتیک و جستجوی پراکنده برای تشخیص تقلب بیمه اتومبیل در این مقاله استفاده خواهد شد. در نهایت، نتایج به‌دست آمده با برخی از روش‌های داده‌کاوی که بیشترین سهم را در تحقیقات گذشته در تشخیص تقلب بیمه اتومبیل داشته‌اند، مورد مقایسه قرار خواهد گرفت. با توجه به ۵ معیار *Accuracy, Precision, Recall* و *Specificity, AUC* الگوریتم پیشنهادی این مقاله با دقت ۹۱/۴۲، ۹۷/۹۶ و ۹۵/۰۱ به ترتیب در معیارهای *Accuracy,Recall* و *AUC* برتری قابل ملاحظه ای نسبت به روش های مقایسه شده دارد.



**کلمات کلیدی:** الگوریتم ژنتیک، جستجوی پراکنده، تشخیص تقلب بیمه اتومبیل، تحلیل سلسله مراتبی، داده‌کاوی.

## ۱- مقدمه

تقلب‌های بیمه از مسائل مهم و خسارت‌زا برای شرکت‌های بیمه، در تمام رشته‌های بیمه است(فیروزی و شکوری،۱۳۹۰،ص۱۰۳). به‌دست آوردن الگویی برای کشف و جلوگیری از این تقلب‌ها دشوار و بسیار هزینه‌بر است. هم‌چنین زمان زیادی برای آن نیاز می‌باشد که حتی باعث می‌شود برخی از شرکت‌های بیمه ترجیح دهند که مبلغ را به کلاهبرداران پرداخت نمایند اما هزینه‌های کشف و ارزیابی تقلب‌ها را تقبل نکنند(دریگ،۲۰۰۲،ص۲۷۴).

شناسایی فعالیت‌های جعلی در اولین لحظه ممکن و تحت پیگیری قرار گرفتن متقلبان در سیستم امری ضروری است اما در تشخیص ادعای جعلی، شناسایی ویژگی‌هایی که آنها را از ادعاهای صحیح متمایز سازد، مشکل است؛ زیرا در پیکربندی امروز برای اجرای مراحل تشخیص- از طریق لیست گسترده‌ای از شاخص‌های تقلب (که در اختیار کارشناسان بیمه قرار دارد)- در لحظه‌ای که ادعا وارد سیستم می‌شود زمان کمی وجود دارد(ویان و همکاران،۲۰۰۵،ص۵۶۷).

در انواع بیمه‌نامه‌ها، بیمهٔ شخص ثالث دارای اهمیت بیشتری برای شرکت‌های بیمه است به این دلیل که اولاً این رشته از نظر سهم‌بری در پورتفوی حق‌بیمه دریافتی شرکت‌های بیمه در بسیاری از کشورها، از جمله ایران رتبه اول را دارد. ثانیاً، این رشته در کشور ما و بسیاری از کشورهای دیگر زیان‌ده است؛ بنابراین، مقابله با موارد تقلب بیمه‌ای که باعث کاهش خسارت‌های پرداختی و درنتیجه، کاهش زیان عملیاتی در بیمه شخص‌ثالث شود، اهمیت بیشتری پیدا می‌کند. ثالثاً، این رشته از نظر انواع تقلب ممکن در آن، شاید متنوع‌تر از دیگر رشته‌های بیمه باشد و در عمل بیشتر در معرض وقوع تقلب قرار داشته‌باشد(حجازی و همکاران،۲۰۱۳،ص۱).

نبودن شاخص‌های کاربردی جامع و مشترک بین شرکت‌های بیمه که برای تشخیص پرونده‌های مشکوک مورد استفاده قرار گیرد، از جمله تفاوت‌های مهمی است که ایران با کشورهای دیگر دارد. بدین معنی که هریک از شرکت‌های بیمه در ایران، علاوه بر چندین معیار کلی که مورد استفاده همهٔ آنهاست، معیارهای خاصی نیز برای تشخیص تقلب دارند که معمولاً دیگر شرکت‌ها از آن‌ها بی‌اطلاعند، که این موضوع، راه را برای متقلبان هموار می‌کند. بدین لحاظ انتخاب درست شاخص‌ها نه تنها منجر به بهبود فهم انسانی از دامنه مسئله می‌شود بلکه امکان راه‌حل‌های کارآتر و با هزینه کمتر را نیز فراهم می‌کند. هم‌چنین رفع افزونگی یا داده‌های بی‌ربط ممکن است به کارایی بیشتر منجر گردد.



باتوجه به محدودیت‌های ذکر شده اولین قدم برای حل مسئله تشخیص تقلب، تعیین شاخص‌های مورد نیاز برای تشخیص موارد مشکوک است. اضافه کردن هر شاخص منجر به افزایش بعد مسئله می‌گردد و در نتیجه یک مجموعه پراکنده تولید می‌کند که ممکن است دقت تشخیص را کاهش دهد؛ بنابراین راه‌حلی نیاز است که با محدودیت‌های ذکر شده سازگاری داشته‌باشد و نتایج قابل قبولی نیز در دقت تشخیص موارد تقلب ارائه کند. این مقاله هر دو مورد ذکر شده را پوشش خواهد داد.

باوجود اینکه تجزیه و تحلیل اقتصادی یکی از مهم‌ترین بخش‌های تقلب بیمه می‌باشد اما راهکارهای کشف و تشخیص تقلب در بیمه بیشتر مورد توجه بوده و تاکنون راهکارهای متنوعی برای تشخیص تقلب انواع بیمه ارایه شده‌است که می‌توان آن‌ها را در مجموعه‌ای از تکنیک‌های داده‌کاوی یافت(ابراهیم حسن و آبراهام،۲۰۱۳،ص۳۴۲)؛ زیرا معمولاً تشخیص تقلب به‌عنوان یک مشکل داده‌کاوی دیده شده است(دومان و اوزلیک،۲۰۱۱،ص۱۳۰۵۸). تاکنون انواع روش‌های داده‌کاوی[1] برای تشخیص تقلب بیمه در اکثر کشورها مورد استفاده قرار گرفته‌است. دو ضعف عمده در روش‌های پیشین وجود دارد:

۱) به مجموعه داده برچسب خورده (اولیه) نیازمندند؛
۲) تنها برای مجموعه داده‌های بزرگتر و غنی‌تر مناسب هستند(سوبلج و همکاران،۲۰۱۱،ص۱۰۴۰).

در مدل‌های موجود تشخیص تقلب در بیمه‌ها دو نوع هدف وجود دارد:

۱) ارزیابی میزان تقلب صورت‌گرفته؛
۲) توسعه سیستم‌های هوشمند برای تشخیص تقلب؛

که مورد دوم، اساس این مقاله را شکل می‌دهد و باتوجه به ضعف عمده‌ای که پیشتر به آن اشاره گردید، این مقاله ترکیب الگوریتم ژنتیک و جستجوی پراکنده را برای تشخیص تقلب بیمه اتومبیل پیشنهاد می‌دهد.

[1] Data mining

## ۲- الگوریتم پیشنهادی

این بخش ابتدا الگوریتم ژنتیک و جستجوی پراکنده را با جزییات مورد نیاز تشریح می‌کند. سپس روش پیشنهادی مورد استفاده در مقاله، معرفی خواهد شد. باتوجه به ضعف عمده الگوریتم ژنتیک مبنی بر عدم درنظر گرفتن شرایط مسئله و انجام جستجوی تصادفی، این تحقیق ترکیب الگوریتم ژنتیک با جستجوی پراکنده را به‌منظور بهبود عملکرد الگوریتم ژنتیک در تشخیص تقلب بیمه اتومبیل پیشنهاد می‌کند.

### ۲-۱ الگوریتم ژنتیک

به‌طور کلی می‌توان روش را بدین صورت شرح داد که در نخستین مرحله الگوریتم ژنتیک، جمعیتی از کروموزوم‌ها به تعداد معین و به‌طور تصادفی تولید می‌شوند(ساستری و گولدبرگ،۲۰۰۵،ص۹۸). هرکروموزوم بیانگر یک جواب از فضای جستجو است و فرد نام دارد. مجموعه این افراد، جمعیت یا نسل فعلی نام دارند. به هر شخص، برازندگی بر اساس مقدار تعیین شده توسط تابع هدف تعلق می‌گیرد. پس از تعیین مقدار برازندگی اعضای جمعیت، می‌توان آن‌ها را با احتمالی متناظر با برازندگی نسبیشان انتخاب کرد و برای تولید نسل بعد ترکیب نمود.

در مرحله بعد نوبت به اعمال عملگرهای تقاطع و جهش می‌رسد و اشخاص مطابق با مقدار برازندگی برای جفت‌گیری انتخاب می‌شوند. فرایند به همین ترتیب تا تولید نسل بعد ادامه می‌یابد. الگوریتم ژنتیک هنگامی که برخی ضوابط مانند تعداد معینی تولید نسل و یا میانگین انحراف معیار عملکرد اشخاص جمعیت تأمین شود، به پایان می‌رسد.

### ۲-۲جستجوی پراکنده

این نسخه از جستجوی پراکنده برخی از روال‌های نسخه ۱۹۹۴ را ساده‌سازی کرده‌است(عبدالوهاب و همکاران،۲۰۰۶،ص۳۵۲). این الگو ۵ روال زیر را به خدمت گرفته است:

۱. روش تولید گوناگونی[1] برای تولید مجموعهٔ راه‌حل‌های متنوع، از راه‌حل آزمایشی اختیاری (رتبه‌بندی شده) به‌عنوان یک ورودی استفاده می‌کند.
۲. روش بهبود[2] به‌منظور انتقال راه‌حل آزمایشی به داخل یک یا چند راه‌حل‌های آزمایشی افزایش‌یافته است. اگر هیچ پیشرفتی از ورودی‌ها نتیجه گرفته‌نشود، راه‌حل‌های افزایش یافته همان راه‌حل‌های ورودی خواهند بود.
۳. روش به‌روز رسانی مجموعه مرجع[3] یک مجموعه مرجع که شامل بهترین راه‌حل‌ها است را ساخت و نگهداری می‌کند (مقدار آن معمولاً کوچک است، تقریباً ۲۰).
۴. روش تولید زیرمجموعه[4] روی مجموعه مرجع اجرا می‌گردد، بدین‌صورت که یک زیرمجموعه از راه‌حل‌ها را به‌عنوان پایه‌ای برای ایجاد راه‌حل‌های ترکیبی، تولید می‌کند.
۵. روش ترکیب راه‌حل[5] زیرمجموعه راه‌حل‌های تولیدشده توسط روش تولید زیرمجموعه را به داخل یک یا چند راه‌حل ترکیب‌شده انتقال می‌دهد.

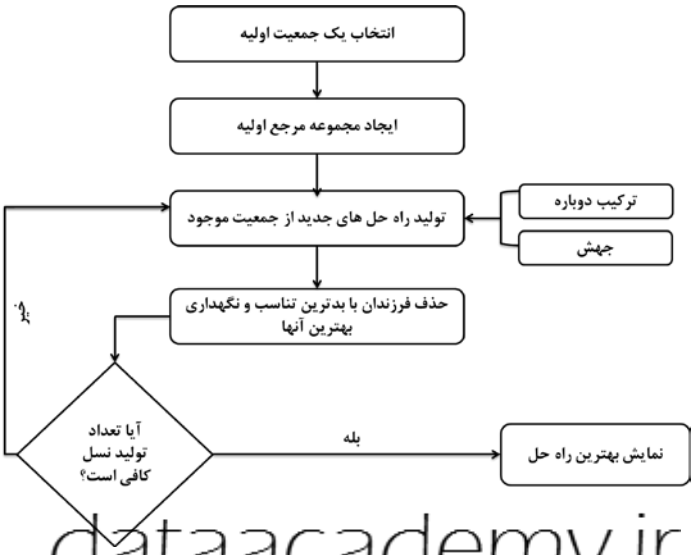


### ۲-۳ ترکیب الگوریتم‌های ژنتیک و جستجوی پراکنده

یکی از تفاوت‌های دو الگوریتم این است که درحالی‌که جمعیت اولیه الگوریتم ژنتیک بیشتر به صورت تصادفی تشکیل می شود، در جستجوی پراکنده ابتدا یک مجموعه از راه‌حل‌های تصادفی تولید می‌گردد و پس از آن، این راه‌حل‌ها با استفاده از روش بهبود محلی مانند جهش، بهبود یافته و برخی از این راه‌حل‌ها (اصلی و بهبود یافته) به صورت اولین مجموعه مرجع انتخاب می‌شوند که

[1] Diversification Generation
[2] Improvement
[3] Reference Set Update Methods
[4] subset generation method
[5] solution combination method

شامل راه‌حل‌های گوناگون و با کیفیت بالا هستند. الگوریتم پیشنهادی اساساً از مراحل الگوریتم ژنتیک پیروی می‌کند اما برخی از قطعات آن از مراحل جستجوی پراکنده استفاده خواهد شد(دومان و اوزلیک،۲۰۱۱،ص۱۳۰۵۹). در مقایسه با پیاده‌سازی الگوریتم ژنتیک معمولی اندازه کوچکتری از جمعیت نگه داشته می‌شود و اطمینان حاصل می‌شود که حداقل سطح از تنوع در هر نسل رعایت شده باشد. برای تولیدمثل از یک عملیات تولیدمثل استاندارد به‌جای عملگر تقاطع کلاسیک الگوریتم ژنتیک استفاده خواهدشد. هم‌چنین در پیاده‌سازی از عملگر جهش استفاده خواهدشد که برای هر دو الگوریتم ژنتیک و جستجوی پراکنده متداول است. شمای الگوریتم پیشنهادی در شکل (۱) نشان داده شده است.






شکل(۱) شمای الگوریتم پیشنهادی

مراحل الگوریتم پیشنهادی به ترتیب زیر است:

۱. اندازه مجموعه مرجع: تعداد راه‌حل‌هایی که برای نسل اولیه از والدین انتخاب می‌شود با توجه به اندازه مسئله تعیین می‌شود، یعنی جمعیت بزرگتر برای مسئله بزرگتر است.
۲. مجموعه مرجع بر اساس ارزش هر کروموزوم مرتب‌سازی می‌شود.
۳. در مرحله انتخاب زیرمجموعه، برای سهولت اجرا هر جفت ممکن از والدین ترکیب می‌گردد.
۴. در ترکیب راه‌حل‌ها، تنها نصف مجموعه مرجع تولید شده (بهترین‌ها) برای تولید مثل نگه داشته می‌شوند و نصف دیگر دور ریخته می‌گردد. راه‌حل‌های انتخاب شده در مجموعه‌ای که بهترین راه‌حل‌های والدین هم در آن قرار دارند، ریخته می‌شوند.
۵. عملگر جهش هم برای اطمینان از حداقل سطح تنوع اجرا می‌گردد.
۶. مراحل تولید نسل بهتر، تولیدمثل و انتخاب تا زمانی که نسل بهتر ۱۰ نسل تغییر نکند تکرار می‌شود.

## ۳- جمع آوری داده ها برای تعیین شاخص های تقلب

با راهنمایی‌هایی که از کارشناسان محترم خسارت شرکت‌های بیمه دریافت شد، ۶۰ شاخص تقلب که دارای بیشترین تأثیر در تشخیص تقلب در پرونده‌های خسارت بیمه اتومبیل وجود دارند استخراج گردید. اطلاعات موردنیاز در زمینه شناسایی شاخص‌های مؤثر در تشخیص تقلب بیمه از طریق تهیه پرسشنامه جمع‌آوری گردیده‌است. توزیع پرسشنامه به دو طریق سنتی (کاغذی) و اینترنتی انجام گرفته که پرسشنامه‌ها در بین کارشناسان خسارت، رؤسای شعب و کارمندان شرکتهای بیمه که دارای تحصیلات

تخصصی در رشته‌های بیمه بودند، توزیع شده‌است. مجموع تعداد پرسشنامه‌های توزیع‌شده ۶۵۳ و تعداد پرسشنامه‌های پاسخ داده‌شده ۲۶۸ عدد می‌باشد.

روایی پرسشنامه توسط چندین کارشناس خسارت تأیید گردیده‌است. پایایی پرسشنامه توسط آلفای کرونباخ مورد ارزیابی قرار گرفت و برای این آزمون مقدار ۰/۹۳۹ توسط نرم افزار ***SPSS*** به‌دست آمد. بدیهی است هرقدر شاخص آلفای کرونباخ به ۱ نزدیک‌تر باشد، همبستگی درونی بین سؤالات بیشتر و در نتیجه پرسش‌ها همگن‌تر خواهند بود. درنتیجه مقدار به‌دست آمده نشان از پایایی قابل قبول پرسشنامه تهیه‌شده، می‌باشد.

برای تعیین میزان اهمیت هر یک از شاخص‌ها یا به‌طور کلی تعیین وزن هر شاخص در مسئله تشخیص تقلب بیمه اتومبیل، تکنیکی مورد نیاز می‌باشد که از بین تعداد زیادی شاخص، پراهمیت‌ترین‌های آن‌ها را معرفی نماید. به‌همین منظور در این مقاله از فرآیند تحلیل سلسله مراتبی برای تعیین وزن هر شاخص استفاده می‌شود. وزن‌های به‌دست آمده برای هر شاخص در این بخش به‌صورت مستقیم در الگوریتم پیشنهادی این مقاله مورد استفاده قرار می‌گیرد.

### ۱-۳ استفاده از تکنیک فرآیند تحلیل سلسله مراتبی برای تعیین وزن هر شاخص

فرآیند تحلیل سلسله مراتبی با شناسایی و اولویت‌بندی عناصر تصمیم‌گیری شروع شده در فرآیند تحلیل، شامل محاسبهٔ وزن (ضرایب اهمیت) معیارها و زیرمعیارها، محاسبهٔ ضریب اهمیت (وزن) گزینه‌ها، محاسبهٔ نهایی گزینه‌ها و بررسی سازگاری منطقی قضاوت‌ها انجام می‌شود.

فرآیند تحلیل سلسله مراتبی، نیازمند شکستن مسئله تصمیم با چندین شاخص به سلسله مراتبی از سطوح است. سطح اول شامل هدف کلی از تصمیم‌گیری می‌باشد. در سطح دوم معیارهای کلی قرار دارند که تصمیم‌گیری بر اساس آن‌ها صورت می‌گیرد و در آخرین سطح نیز گزینه‌های تصمیم که در اینجا شاخص‌های تشخیص تقلب بیمه اتومبیل هستند، مطرح می‌شوند. تمامی فرآیندهای تحلیل سلسله مراتبی مورد استفاده در این تحقیق توسط نرم افزار ***Expert Choice 11*** انجام گرفته است.



در این تحقیق از معیارهای معرفی شده در تحقیق "بلهاجی و همکاران، ۲۰۰۰" استفاده شده است(بلهاجی و همکاران، ۲۰۰۰،ص۵۲۱). شش معیار "سوابق ادعا"، "بیمه‌نامه"، "هزینه"، "مدارک و شواهد"، "حوادث خاص" و "دیگر عوامل" از معیارهای مهم برای تعیین میزان اهمیت هر یک از شاخص‌ها می‌باشد (هدف). در سطح سوم شاخص‌هایی که از طریق پرسشنامه به‌دست آمده اند، وجود دارد.

نتایج حاصل از مقایسه زوجی معیارها در جدول (۱) ارائه شده است. ارجحیت هر یک از گزینه‌ها (۶۰ شاخص به‌دست آمده از پرسشنامه) در ارتباط باهریک از معیارها مورد قضاوت قرار گرفته‌است.

**جدول (۱) ضرایب ماتریس‌های ارزیابی و بردار وزنی معیارها**

| معیار | وزن نسبی | معیار | وزن نسبی |
|---|---|---|---|
| سوابق ادعا | ۰/۱۸۴ | هزینه | ۰/۱۳۷ |
| بیمه‌نامه | ۰/۱۴۷ | مدارک و شواهد | ۰/۳۶۱ |
| حوادث خاص | ۰/۰۹۰ | دیگر عوامل | ۰/۰۸۱ |

نتیجه محاسبات، میزان سازگاری معیار و گزینه‌ها را کمتر از ده درصد نشان داده است و در نتیجه اولویت‌بندی مقایسه زوجی ماتریس‌ها مورد قبول می‌باشد. قضاوت معیارها و گزینه‌ها برمبنای مقایسه دودویی و براساس مقیاس ۹ کمیتی عددی صورت پذیرفته است. پس از مشخص شدن وزن نسبی معیارها و گزینه‌ها، با ادغام وزن‌های نسبی به‌دست آمده، وزن نهایی در جدول (۲) ارائه شده‌است.

امتیازهای نهایی از ۰/۰۰۷ تا ۰/۰۲۹ تعیین شده‌است. ۵ شاخص اول جدول با بالاترین امتیاز نهایی دارای بیشترین تأثیر و ۷ شاخص انتهایی جدول با کمترین امتیازهای نهایی دارای کمترین تأثیر در تشخیص تقلب بیمه اتومبیل خواهند بود. به‌طور کلی با

کاهش امتیاز نهایی هر شاخص از میزان تأثیر آن در تشخیص تقلب بیمه اتومبیل کاسته می‌شود. تمامی شاخص‌هـا در جـدول (۲) به اختصار $Q$ بیان شده است که این شاخص‌ها در پیوست به‌صورت کامل ارائه شده‌است.

**جدول (۲) نتیجه اولویت بندی نهایی گزینه‌ها با توجه به امتیازات مربوطه**

| شاخص‌ها | امتیاز نهایی | شاخص‌ها | امتیاز نهایی | شاخص-ها | امتیاز نهایی |
|---|---|---|---|---|---|
| $Q_{29}$ | ۰/۰۲۹ | $Q_{4}$ | ۰/۰۱۸ | $Q_{10}$ | ۰/۰۱۲ |
| $Q_{42}$ | ۰/۰۲۹ | $Q_{7}$ | ۰/۰۱۸ | $Q_{55}$ | ۰/۰۱۲ |
| $Q_{50}$ | ۰/۰۲۹ | $Q_{13}$ | ۰/۰۱۸ | $Q_{56}$ | ۰/۰۱۲ |
| $Q_{51}$ | ۰/۰۲۹ | $Q_{22}$ | ۰/۰۱۸ | $Q_{60}$ | ۰/۰۱۲ |
| $Q_{52}$ | ۰/۰۲۹ | $Q_{3}$ | ۰/۰۱۷ | $Q_{2}$ | ۰/۰۱۱ |
| $Q_{26}$ | ۰/۰۲۸ | $Q_{6}$ | ۰/۰۱۷ | $Q_{15}$ | ۰/۰۱۱ |
| $Q_{14}$ | ۰/۰۲۷ | $Q_{1}$ | ۰/۰۱۶ | $Q_{9}$ | ۰/۰۱۰ |
| $Q_{23}$ | ۰/۰۲۶ | $Q_{5}$ | ۰/۰۱۶ | $Q_{28}$ | ۰/۰۱۰ |
| $Q_{24}$ | ۰/۰۲۶ | $Q_{8}$ | ۰/۰۱۵ | $Q_{30}$ | ۰/۰۱۰ |
| $Q_{12}$ | ۰/۰۲۵ | $Q_{18}$ | ۰/۰۱۵ | $Q_{48}$ | ۰/۰۱۰ |
| $Q_{35}$ | ۰/۰۲۱ | $Q_{19}$ | ۰/۰۱۵ | $Q_{49}$ | ۰/۰۱۰ |
| $Q_{36}$ | ۰/۰۲۱ | $Q_{20}$ | ۰/۰۱۵ | $Q_{58}$ | ۰/۰۱۰ |
| $Q_{37}$ | ۰/۰۲۱ | $Q_{21}$ | ۰/۰۱۵ | $Q_{59}$ | ۰/۰۱۰ |
| $Q_{40}$ | ۰/۰۲۱ | $Q_{34}$ | ۰/۰۱۵ | $Q_{16}$ | ۰/۰۰۹ |
| $Q_{45}$ | ۰/۰۲۱ | $Q_{38}$ | ۰/۰۱۵ | $Q_{27}$ | ۰/۰۰۹ |
| $Q_{47}$ | ۰/۰۲۱ | $Q_{39}$ | ۰/۰۱۵ | $Q_{31}$ | ۰/۰۰۹ |
| $Q_{53}$ | ۰/۰۲۱ | $Q_{41}$ | ۰/۰۱۵ | $Q_{32}$ | ۰/۰۰۹ |
| $Q_{57}$ | ۰/۰۲۱ | $Q_{44}$ | ۰/۰۱۵ | $Q_{43}$ | ۰/۰۰۹ |
| $Q_{17}$ | ۰/۰۱۹ | $Q_{11}$ | ۰/۰۱۳ | $Q_{54}$ | ۰/۰۰۹ |
| $Q_{25}$ | ۰/۰۱۹ | $Q_{33}$ | ۰/۰۱۳ | $Q_{46}$ | ۰/۰۰۷ |



### ۳-۲ داده‌های مورد استفاده در این مقاله

داده‌های مورد استفاده در این تحقیق از ۶ شرکت بیمه در شهر کرمانشاه جمع‌آوری شده‌است. مجموعه داده‌ها به دو دسته پرونده-های تقلبی و پرونده‌های صحیح طبقه‌بندی می‌شود. معیار تقلب و یا صحیح بودن پرونده‌ها، تایید کارشناسان محتـرم شـرکت‌هـای مزبور می‌باشد. داده‌های جمع‌آوری شده مربوط به بازه زمانی سال ۸۹ تا بهار ۹۴ می‌باشد. تعداد پرونده‌های تقلبی در دسـترس کـه تقلب آن‌ها اثبات شده باشد، باتوجه به مشکلاتی که قبلاً به آن‌ها اشاره شد به‌طور قابل ملاحظه‌ای کم اسـت. بـرای مقابلـه بـا ایـن مشکل، باتوجه به صلاحدید کارشناسان محترم خسارت، آن دسته از پرونده‌هایی که تقلب آن برای کارشناسان محرز بوده اما بنا بـه هر دلیلی امکان اثبات آن وجود نداشته است نیز به داده‌ها اضافه گردیده‌است.

مجموع پرونده‌هایی که به‌عنوان داده‌های این مقاله استفاده خواهد شد و به تایید کارشناسان محترم خسارت شرکت‌های مذکور رسیده‌است ۳۹۴ عدد می‌باشد؛ که از این تعداد ۳۵ پرونده تقلبی و ۳۵۹ پرونده صحیح می‌باشد (تعداد پرونده‌های تقلب تقریباً ۱۰ درصد پرونده‌های صحیح می‌باشد). تعداد پرونده‌های صحیح در سال‌های مورد بررسی این تحقیق بسیار زیاد می‌باشد؛ اما تنها پرونده‌هایی در این مقاله استفاده شده‌است که حداقل یکی از شاخص‌های تقلب در آن وجود داشته است؛ به‌عبارت دیگر پرونده‌ای مشکوک به تقلب بوده است اما صحت این پرونده برای شرکت‌های بیمه اثبات شده می‌باشد.

## ۴- روند اجرای الگوریتم پیشنهادی

در این بخش روند اجرایی الگوریتم که توسط زبان برنامه‌نویسی سی شارپ[1] صورت گرفته است، شرح داده خواهد شد. از آنجا که عملگرهای استفاده شده در الگوریتم پیشنهادی به‌طور کلی عملگرهای استاندارد الگوریتم‌های ژنتیک و جستجوی پراکنده بوده‌است، لذا در این بخش از تکرار آن‌ها صرف‌نظر شده و تنها به معرفی و تشریح تابع تناسب الگوریتم پرداخته خواهد شد.

ابتدا مسئله تشخیص تقلب در ساده‌ترین حالت ممکن بررسی می‌گردد. این مسئله از تعدادی شاخص تشکیل شده است (۶۰ عدد) که هر یک از این شاخص‌ها با ارزش‌های متفاوت (اهمیت متفاوت) در تشخیص تقلب تأثیرگذار هستند. به‌منظور تعیین اهمیت هر یک از شاخص‌ها از فرآیند تحلیل سلسله مراتبی استفاده گردید که در بخش قبل به‌طور مفصل بیان شد. هم‌چنین تعداد تکرار هر یک از این شاخص‌ها نیز در تشخیص تقلب اهمیت دارد؛ زیرا تجربه کارشناسان خسارت ثابت کرده‌است اگرچه برخی از این شاخص‌ها دارای ارزش کمتری می‌باشند اما باتوجه به تعداد تکرار آن‌ها در پرونده‌های مختلف تأثیرگذار تلقی می‌شوند و گاه ممکن است منجر به تشخیص تقلب شوند. پس منطقی به‌نظر می‌رسد که تعداد این تکرارها نیز در الگوریتم به‌کار برده شوند. حال مسئله‌ای پیش رو است که تعدادی شاخص با ارزش‌های متفاوت دارد و هریک از این شاخص‌ها دارای تکرار متمایز و هستند. این مسئله با محدودیت "وجود ۱۰ شاخص" در هر مسئله نیز مواجه است. پس می‌توان نتیجه گرفت با توجه به آنچه ذکر شد استفاده از تابع هدف کوله‌پشتی صفر و یک برای تابع ارزیاب الگوریتم پیشنهادی دور از ذهن نخواهد بود. در ادامه ویژگی‌های بیشتری از این دو مطرح می‌گردد.



به‌طور کلی، مسئله کوله‌پشتی صفر و یک، روند انتخاب از اقلام مختلف به داخل کوله‌پشتی با برخی از محدودیت‌ها می‌باشد(کونگ و همکاران،۲۰۱۵،ص۵۳۳۷). هدف به‌حداکثر رساندن سود از اقلام بسته‌بندی شده در کوله پشتی است که تحت شرایطی باشد که حجم کل مربوطه کمتر یا مساوی ظرفیت حجم داده شده باشد. از آنجا که متغیرهای تصمیم به صفر یا یک محدود می‌شوند، فضای متغیر از مجموعه‌ای از نقاط گسسته متناهی که در آن راه‌حل بهینه واقع شده است تشکیل می‌گردد.

برای استفاده از تابع هدف کوله‌پشتی در مسئله تشخیص تقلب بیمه اتومبیل ابتدا باید شباهت‌های آن‌ها را مورد بررسی قرار داد. اطلاعاتی که از پرونده‌های بیمه اتومبیل در دسترس می‌باشد تعداد ۶۰ شاخص تشخیص تقلب است که در هر پرونده می‌تواند وجود داشته باشد یا نداشته باشد. هر یک از شاخص‌ها متغیری باینری است که تنها مقدار صفر یا یک می‌پذیرد. وجود هریک از این متغیرها در پرونده تقلب دارای ارزش متفاوتی در تشخیص تقلب بیمه اتومبیل خواهد بود که این ارزش‌ها با استفاده از تکنیک فرآیند تحلیل سلسله مراتبی مرتب‌سازی گردیدند.

در هر پرونده بیشتر از ۱۰ شاخص از ۶۰ شاخص معرفی شده نمی‌تواند به‌طور همزمان وجود داشته باشد. این مسئله بدیهی است که متقلبان برای رسیدن به مقصود خود -گرفتن خسارت غیرواقعی از شرکت‌های بیمه- بسیار زیرکانه عمل می‌کنند و این موضوع که بیشتر از ۱۰ شاخص در پرونده درخواست خسارت آن‌ها وجود داشته باشد امری محال است. به‌عبارت دیگر در صورت وجود بیشتر از ۱۰ شاخص از شاخص‌های معرفی شده، پرونده بدون بررسی بیشتر، تقلب شناخته شده و بدون پرداخت خسارت به مدعی، مختومه اعلام می‌گردد. هم‌چنین برخی از این شاخص‌ها در پرونده‌های مشکوک به تقلب بیشتر دیده می‌شوند. به‌عبارت دیگر تکرار بیشتر هر یک از آن‌ها در پرونده‌های پیشین و وجود آن‌ها در پرونده خسارت جاری برای کارشناسان این الزام را به‌وجود می‌آورد که پرونده را بیشتر از پرونده‌های عادی (که هیچ‌یک از این شاخص‌ها در آن وجود ندارد) مورد بررسی قرار دهند. با توجه

[1] C#

به جدول (۳) اگر ارزش هر متغیر $v_i$ ، وزن هر متغیر $w_i$ و ظرفیت کوله‌پشتی $W$ نامیده شود صورت ریاضی مسئله کوله‌پشتی صفر و یک طبق رابطه (۱) محاسبه می‌گردد.

$$\begin{aligned} &Maxf(x) = \sum_{i=1}^{n} v_i x_i \\ &\sum_{i=1}^{n} w_i x_i \le W_{\max} \\ &i = 1,2,\ldots,n \\ &xi \in \{0,1\} \end{aligned} \qquad (۱)$$

**جدول (۳) معادل‌سازی اصطلاحات تشخیص تقلب بیمه با کوله پشتی صفر و یک**

| ویژگی‌ها | مسئله کوله‌پشتی صفر و یک | تشخیص تقلب بیمه اتومبیل |
|---|---|---|
| ارزش متغیرها | ارزش هر شی | هر متغیر دارای ارزش خاصی است که از تکنیک فرایند تحلیل سلسله مراتبی به دست آمده است |
| وزن هر متغیر | وزن هر شی | وزن هر متغیر بر اساس تعداد تکرار آن در پرونده‌های خسارت |
| محدودیت | جمع وزن اشیا باید به‌گونه‌ای باشد که از ظرفیت کوله‌پشتی تجاوز نکند | در هر پرونده بیشتر از ۱۰ شاخص باهم نمی‌تواند وجود داشته باشد |
| وضعیت هر متغیر | یا برداشته می‌شود یا خیر (صفر یا یک) | یا در پرونده وجود دارد یا خیر (صفر یا یک) |



تفاوت مسئله کوله‌پشتی با تشخیص تقلب در این است که در مسئله کوله‌پشتی هدف بیشینه کردن سود است، یعنی چه اشیایی برداشته شود که سود به حداکثر برسد؛ اما در مسئله تشخیص تقلب، حداکثرسازی هدف نخواهد بود و تشخیص درست موارد تقلب و موارد صحیح، هدف الگوریتم می‌باشد.

برای حل این تفاوت در الگوریتم پیشنهادی درصدی به‌عنوان درصد تقلب درنظر گرفته شده است؛ یعنی الگوریتم تمامی حالات تقلب را با توجه به پارامترهای ارزش و وزن به‌دست می‌آورد (تا این قسمت مسئله کوله‌پشتی واضح است) اما از آنجا که تمامی حالات تقلب بسیار زیاد بوده و عملاً ربطی به مسئله تشخیص تقلب ندارد، محدودیت‌هایی بر الگوریتم اجرا شده‌است. این محدودیت، محاسبه درصد تعلق ورودی به آخرین نسل از تقلب صورت گرفته می‌باشد. این نسل دارای ارزش‌های متفاوتی خواهد بود. هرچه درصد محاسبه شده تعلق ورودی به ارزش‌های نسل آخر بیشتر باشد احتمال وقوع تقلب بیشتر است و هرچه این درصد کمتر باشد احتمال صحیح بودن ورودی تشدید می‌شود.

## ۵- نتایج الگوریتم پیشنهادی و مقایسه با سایر رویکردها

روش‌های انتخاب شده برای مقایسه با الگوریتم پیشنهادی، از بین پرتکرارترین روش‌های تشخیص تقلب بیمه که تاکنون مورد بررسی قرار گرفته‌اند انتخاب شده است. به‌طور کلی الگوریتم‌هایی که در ادامه معرفی خواهند شد از محبوب‌ترین روش‌های مورد استفاده در تشخیص تقلب می‌باشند که در حال حاضر در بسیاری از کشورها مورد استفاده کاربردی قرار گرفته‌اند. تمامی الگوریتم‌های داده‌کاوی در این تحقیق با استفاده از نرم‌افزار رپیدماینر[1] نسخه ۵ انجام گرفته است. در اجرای این الگوریتم‌ها ۷۰ درصد داده‌ها به‌عنوان دادهٔ آموزشی و ۳۰ درصد داده‌ها به‌عنوان دادهٔ آزمایشی مورد استفاده قرار گرفته‌اند. تمامی روش‌های مورد استفاده در هریک از الگوریتم‌ها، روش‌های پیش‌فرض و استاندارد نرم افزار رپیدماینر می‌باشد. نتایج حاصل از ارزیابی الگوریتم پیشنهادی با چندین روش شناخته‌شده داده‌کاوی مقایسه شد که جدول (۴) نتایج آن‌ها را نشان می‌دهد.

[1] RapidMiner 5

همان‌طور که ملاحظه می‌شود الگوریتم پیشنهادی با دقت تشخیص ۹۱/۴۲ درصد از موارد تقلب بالاترین میزان دقت را در میان روش‌های مورد مقایسه دارد (معیار ***recall***). هم‌چنین معیار ***Accuracy*** برای سنجش دقت کلی روش‌ها استفاده شده است که دقت ۹۷/۹۶ درصدی هم برتری قابل ملاحظه الگوریتم پیشنهادی را نسبت به روش‌های مورد مقایسه نشان می‌دهد.

**جدول (۴) خلاصه ای از مهم ترین نتایج به دست آمده در این فصل**

| | *Recall* | *Accuracy* | *AUC* | *Specificity* | *precision* |
|---|---|---|---|---|---|
| الگوریتم پیشنهادی | ۹۱/۴۲ | ۹۷/۹۶ | ۹۵/۰۱ | ۹۸/۶۰ | ۸۶/۴۹ |
| روش بیزین | ۸۸/۸۹ | ۹۷/۴۶ | ۹۳/۵۳ | ۹۸/۱۷ | ۸۰/۰۰ |
| شبکه عصبی | ۸۰/۰۰ | ۹۷/۴۵ | ۹۰/۰۰ | ۹۹/۰۷ | ۸۸/۸۹ |
| لجستیک ساده | ۷۷/۷۸ | ۹۶/۶۱ | ۸۷/۹۷ | ۹۸/۱۷ | ۷۷/۷۸ |
| درخت تصمیم | ۷۷/۷۸ | ۹۷/۴۶ | ۸۷/۵۱ | ۹۷/۲۵ | ۷۰/۰۰ |

## ۶- بحث و نتیجه‌گیری

رشد تقلب، خطر بزرگی برای حیات شرکت‌های بیمه است؛ زیرا در صورت عدم تشخیص به‌موقع تقلب، هزینه‌های هنگفتی از منابع شرکت‌های بیمه به حساب متقلبان واریز می‌شود که در صورت تکرار آن بیمه‌ها را با خطر ورشکستگی مواجه می‌کند. عدم وجود لیست یکپارچه از شاخص‌های تشخیص تقلب بیمه و نبود روش تشخیص تقلب بیمه اتومبیل مبتنی بر سیستم‌های کامپیوتری در ایران، دو دلیل مهم است که تاکنون راه را برای متقلبان باز گذاشته است. در این مقاله هر دو زمینه تعیین شاخص‌های کاربردی تقلب و هم ارائه سیستم تشخیص تقلب مبتنی بر کامپیوتر پوشش داده شده‌است. استفاده از ترکیب الگوریتم ژنتیک و جستجوی پراکنده به‌منظور تشخیص تقلب بیمه اتومبیل برای اولین بار مورد استفاده قرار گرفته‌است. به‌منظور بهبود محدودیت الگوریتم ژنتیک مبنی بر تصادفی بودن جمعیت اولیه، از جستجوی پراکنده برای تولید جمعیت اولیه استفاده گردید. هم‌چنین عملگر تقاطع نیز از الگوریتم ژنتیک با عملگر ترکیب دوباره از جستجوی پراکنده جایگزین شد. نتایج حاصل از الگوریتم پیشنهادی برتری قابل ملاحظه‌ای نسبت به انواع روش‌های دادکاوی داشت. از آنجا که الگوریتم‌های فرابتکاری کمتر در این زمینه مورد استفاده قرار گرفته‌اند، استفاده از دیگر الگوریتم‌ها می‌تواند به‌عنوان پیشنهادی برای تحقیقات آینده مطرح گردد.



## ۷- فهرست منابع

[۱] فیروزی، مهدی؛ شکوری، مرتضی؛ کاظمی، لیلا؛ زاهدی، سحر.(۱۳۹۰)؛ «شناسایی تقلب در بیمهٔ اتومبیل با استفاده از روش‌های داده‌کاوی»، پژوهشنامهٔ بیمه (صنعت بیمهٔ سابق)، سال ۲۶، شماره ۳، صفحات ۱۰۳ تا ۱۲۸.

[2] Abdule-Wahab, Rasha.S, (2006), "A Scatter Search Algorithm for the Automatic Clustering Problem", *Springer-Verlag Berlin Heidelberg*, pp. 350-364.

[3] Belhadji, El.Bachir; Dionne, Georges; Tarkhani, Faouzi,(2000), "A Model for the Detection of Insurance Fraud", *The Geneva Papers on Risk and Insurance Vol. 25, No.4*, pp. 517-538.

[4] Derrig , Richard, (2002),"Insurance Fraud", *The Journal of Risk and Insurance*, pp. 271-287.

[5] Duman, Ekrem; Ozcelik, M.Hamdi,(2011), "Detecting credit card fraud by genetic algorithm and scatter search", *Expert Systems with Applications* 38, pp. 13057–13063.

[6] Hedjazi, Arya et al. (2013),"A report of five cases of self-mutilation for the purpose of insurance fraud",*Journal of Forensic and Legal Medicine*, pp. 1-4.

[7] Ibrahim Hassan, Amira. Kamil; Abraham, Ajith, (2013),"Computational Intelligence Models for Insurance Fraud Detection: A Review of a Decade of Research", *Journal of Network and Innovative Computing*, pp. 341-347.

[8] Kong, Xiangyong et al. (2015), "A Simplified Binary Harmony Search Algorithm for Large Scale 0–1 Knapsack Problems", *Expert Systems with Applications 42*, pp. 5337–5355.

[9] Sastry,Kumara; Goldberg,David,(2005), "Genetic Algorithm", *Search Methodologies*, pp. 97-125.

[10] Šubelj, Lovro; Furlan, Štefan; Bajec, Marko,(2011), "An expert system for detecting automobile insurance fraud using social network analysis", *Expert Systems with Applications 38*, pp. 1039–1052.

[11] Viaene, Stijn et al.(2007), "Strategies for detecting fraudulent claims in the automobile insurance industry", *European Journal of Operational Research 176*, pp. 565–583.

## پیوست

شاخص‌های مورد استفاده در مقاله که در بخش روش تحقیق در جدول (۲) معرفی شده‌اند در این قسمت به صورت کامل تشریح می‌گردند.

Q.1 فاصله‌زمانی وقوع حادثه با تاریخ شروع بیمه‌نامه و یا اتمام آن کم باشد.

Q.2 مقصر حادثه بلافاصله پس از صدور الحاقیه ادعای خود را مطرح کند.

Q.3 پس از یک وقفه طولانی از انقضای بیمه‌نامه سال قبل، بیمه‌نامه تمدیدی مقصر حادثه به صورت کوتاه مدت صادر شده‌باشد.

Q.4 مقصر حادثه در گذشته ادعای مشکوکی را مطرح کرده باشد.( قبلا پرونده مشابهی تشکیل داده باشد.)

Q.5 مقصر حادثه بیشتر از دوبار در سال ادعا مطرح کرده باشد.

Q.6 نام مقصر حادثه و یا زیان‌دیده در لیست سیاه بیمه مرکزی وجود داشته باشد.

Q.7 ادعای خسارت در مدت زمان قانونی به شرکت بیمه گزارش نشده باشد.



Q.8 مبلغی که زیان دیده مطالبه کند مبلغ قابل توجهی باشد.

Q.9 قبل از وقوع حادثه بیمه‌گذار سوالات زیادی از چگونگی پوشش بیمه بپرسد که مشابه آنچه اتفاق افتاده است، باشد.

Q.10 بیمه‌گذار اطلاعات کافی از چگونگی بروز حادثه نداشته باشد و نتواند به یاد بیاورد حادثه کجا رخ داده است.

Q.11 اطلاعات خودرو بیمه‌گذار با اطلاعات خودرو حادثه آفرین همخوانی نداشته باشد.

Q.12 بیمه گذار و یا زیان دیده قبل از بررسی کارشناس آثار خسارت را از بین ببرند.

Q.13 طرفین حادثه فاقد مدارک و شاهدانی باشند که برای اثبات کافی باشد.

Q.14 جزئیات گزارش درمانی طرفین حادثه مبهم باشد و با اطلاعات مدارک همخوانی نداشته باشد.

Q.15 امضای بیمه گذار مقصر حادثه در فرم پیشنهاد بیمه نامه با امضای فرد مدعی در زمان خسارت همخوانی نداشته باشد.

Q.16 گزارش بیمارستانی وجود نداشته باشد.

Q.17 زیان دیده صدمات غیر مرتبط با تصادف را در پزشکی قانونی مطرح سازد.

Q.18 یک تصادف خفیف منجر به هزینه های گزافی شود.

Q.19 حاثه بین یک اتومبیل مدل پایین و در شرایط اسقاط و یک اتومبیل لوکس و گرانبها رخ داده باشد.

Q.20 در خسارت‌های جرحی مدعی مشتاق باشد خسارت قبل از مهلت اتمام دادرسی پرداخت گردد.

Q.21 در خسارت‌های مالی مدعی مشتاق باشد، خسارت بدون انجام بازدید مجدد کارشناس بیمه پرداخت گردد.

Q.22 فاصله زمانی بین حادثه، کروکی، پرونده بالینی و رای دادگاه از حد معمول کمتر و یا بیشتر باشد.

Q.23 توسط مقصر حادثه، به کارشناس اجازه ی بازدید داده نشود.

Q.24 گزارش محلی درباره وضعیت، متناقض با ادعای مطرح شده باشد.

Q.25 طرفین حادثه توضیحات شاهدان را انکار کنند.

Q.26 یکی از طرفین حادثه و یا هر دو، بیکار باشند و یا در یک منطقه‌ی فقیرنشین زندگی کنند.

Q.27 اطلاعات بیمه‌نامه طرفین حادثه کامل نباشد.

Q.28 طرفین حادثه خیلی پرخاشگر و دعوایی باشند و یا در شرکت بیمه دعوای ساختگی ایجاد کنند.

Q.29 طرفین حادثه از پاسخ به سوالات کارشناس بیمه در مورد حادثه طفره بروند.

Q.30 در طول مدت کارشناسی توسط بیمه‌گر، طرفین حادثه حالت عصبی و دستپاچه داشته باشند.



Q.31 مالکیت خودروها به تازگی تغییر کرده باشد.

Q.32 بیمه‌نامه صادره طرفین حادثه، دارای تخفیفات بیش از حد باشد.

Q.33 شماره سریال بیمه‌نامه صادره بر روی لاشه بیمه‌نامه با کوپن‌های خسارت یکی نباشد.

Q.34 استعلام بیمه نامه مقصر حادثه در سایت بیمه مرکزی وجود نداشته باشد.

Q.35 امکان بازدید از شماره شاسی حک شده بر روی خودروها برای کارشناس مقدور نباشد.

Q.36 در زمان تشکیل پرونده کپی مدارک به جای اصل آنها، ارائه شود.

Q.37 طرفین حادثه در زمان های مختلف خودروهای خود را جهت بازدید کارشناس بیمه به شرکت معرفی نمایند و حاضر نباشند همزمان بیایند.

Q.38 زیان دیده حاضر به دریافت خسارت کمتر از واقع باشد به شرط آنکه به راهنمایی و رانندگی جهت ترسیم کروکی مراجعه ننماید.

Q.39 مقصر حادثه به علت بی مبالاتی و با وجود تماس با ۱۱۰،جریمه نشده‌باشد (برگ جریمه نداشته باشد).

Q.40 باوجود خسارت جرحی هیچ‌کدام از خودروها به پارکینگ منتقل نشده باشند.

Q.41 طرفین حادثه نسبت به ارائه قبض پارکینگ بی اطلاع باشند.

Q.42 راننده مقصر حادثه در گزارش‌های اولیه و گزارشات تکمیلی متفاوت باشد.

Q.43 برای بیمه نامه مقصر حادثه، المثنی صادر شده باشد.

Q.44 خسارت خودروی مقصر حادثه کمتر از خسارت خودروی زیان دیده بوده و گزارش پلیس ارائه نگردد.

Q.45 بدون اطلاع مقامات انتظامی، شکواییه تنظیم و توسط مقامات قضایی، پرونده بررسی گردد.(کروکی فرضی توسط شورای حل اختلاف و کارشناس رسمی دادگستری ترسیم شده باشد.)

Q.46 شماره سریال کروکی بلافاصله پس از ترسیم کروکی در سیستم راهور ثبت شده باشد.

Q.47 مدارک خودرو، اعم از بیمه نامه، کارت ماشین و سند خودرو زیان دیده مفقود شده باشد.

Q.48 اکثر جراحت‌های بدنی از کمر به پایین و یا بیشتر در ناحیه صورت (و بیشتر بینی) باشد.

Q.49 در واژگونی خودرو، افراد مصدوم نسبت خانوادگی نداشته باشند.

Q.50 راننده مقصر حادثه در صحنه حادثه شناسایی نشود.

Q.51 مواضع آسیب دیده خودروهای مقصر حادثه و زیان دیده با همدیگر مطابقت نداشته باشد.

Q.52 اظهارات مقصر حادثه و زیان‌دیده در طول روند دادرسی تغییر کرده باشد.

Q.53 خسارت‌های جرحی فاقد صورت جلسه اولیه مامورین کلانتری باشد.

Q.54 بیمه گذار و مقصر حادثه و مالک خودرو اشخاص متفاوتی باشند.

Q.55 خودروی مقصر و یا زیان دیده در حالت سقوط در پرتگاه و در انتهای پرتگاه قرار داشته باشد.

Q.56 خودروی مقصر حادثه و یا زیان دیده دچار حریق گردیده باشد.

Q.57 گزارش افسر راهور فاقد مهر پایگاه راهور باشد.

Q.58 مقصر حادثه و یا زیان دیده دارای شغل مرتبط با بیمه نیست اما اطلاعات زیادی نسبت به مراحل پرداخت و اصطلاحات بیمه و قوانین بیمه نامه شخص ثالث دارد.

Q.59 مقصر حادثه و یا زیان¬دیده دارای ظاهر نامتعارف باشد.(اعم از اعتیاد به مواد مخدر، الکل و یا عدم تعادل روانی)

Q.60 مقصر و یا زیان دیده درخواست پرداخت رشوه داشته باشد.