

Identify suspicious collusion across service providers

Multi-channel fraud not only results in added expenses for bogus claims but also in the corruption of vital data on which underwriters rely. Certain tactics can help insurers to actively monitor the claim process for collusion and reduce costs

North American companies have developed sophisticated processes and technology to detect claims fraud. There are a variety of analytical techniques to use against claims data, such as automated red flags, predictive modeling, rules based analysis, data mining and others. At the most basic level, companies use decision rules to identify fraud at the claim level. Some companies process data at the provider level. They focus on the intermediaries who sit between the company and the customer. This is useful in cutting down fraud; however the weakness is that it assesses service provider behavior in isolation. A key gap in the arsenal is the capability to monitor the service provider network as a composite and assess pair-wise or group-wise culpability in fraud. This is the type of fraud we term "collusion." Let's discuss tactics to beat it.

What is collusion?

Collusion is a tacit agreement amongst two or more entities in the value chain between the insurer and the customer. The purpose of the agreement is to misrepresent or to inflate loss events and thus to defraud the insurer. A typical example of collusion fraud would be a third party adjuster approving fraudulent claims on soft tissue injuries as submitted by a complicit clinic. As with other type of fraud, collusion hurts the industry in two ways: First, there is the direct charge to the insurer for claims that are not legitimate. Second, the inaccurate or non-existent claims corrupt the data used by underwriters.

How rampant is collusion fraud?

Organizational inertia is one reason that insurers do not invest the time, resources, or the budget to explore multi-channel fraud. Companies need to take a long view and build a business case around the amount they are actually losing by not actively monitoring the claim process for this type of fraud. While we are not aware of an industry survey on the size of the problem, we highlight a couple of illustrative case studies. In 2008, the Manhattan District Attorney [1] indicted 11 persons who operated a fraudulent "medical mill" that had bilked more than \$6.2 million from insurance companies. Those charged included three medical doctors, a chiropractor, two acupuncturists, ten corporations, several 'runners,' and one mastermind behind the entire operation. The challenge with diagnosing such fraud is that when all the service providers are essentially validating each other, it becomes extremely hard to isolate inconsistencies at a single service provider level. In another example of collusion fraud, a single adjuster approved non-existent claims of net value of \$2.4M from a rogue tire dealer over a multi-year period [2]. The point to be made is that if an organization is not performing periodic audits across all service providers, then its claim process is at risk of subversion and could be hemorrhaging millions of dollars in profits. We next discuss the operational and the technical challenges in mitigating collusion fraud.

Privacy concerns and sharing data

One of the operational challenges in fighting claim fraud lies in gaining access and use of relevant data from multiple service entities. If third party service providers are involved, then contractual obligations should be enforced to ensure service data are made available. Service partners often hold a perception that service data cannot be shared with partners because of privacy legislation. This is fundamentally

IEF-004: This article appeared as a web exclusive in Claims Magazine (<http://claimsmag.com>) on Oct 11, 2009.

incorrect. Our point of view on this is not a legal opinion, but a clarification on the government stance about customer data privacy and security.

Government guidelines are clear about an organization's release of customer information to a third-party service provider. Customer confidentiality is enforced through a contractual agreement. Information that would identify customers through their name, address, date of birth, telephone number, social security identifier, or credit card number are not relevant and can be suppressed. The remaining service data can be shared and we particularly point to sub-section (e) in Section 6802 of the *Gramm-Leach-Bliley Act* [3], as being applicable to companies operating under U.S. law.

"Subsections (a) and (b) of this section shall not prohibit the disclosure of non-public personal information ... (3) (A) to protect the confidentiality or security of the financial institution's records pertaining to the consumer, the service or product, or the transaction therein; (B) to protect against or prevent actual or potential fraud, unauthorized transactions, claims, or other liability; (C) for required institutional risk control, or for resolving customer disputes or inquiries."

A similar bill passed by the Canadian Parliament set out the privacy preservation guidelines, explicitly stating that use of the relevant data is permissible for purposes such as statistical analysis. Guidelines similar to those listed above have also been listed in the seventh principle of the *UK Data Protection Act of 1998*. Data availability is core to fraud mitigation. It is in the best interests of the industry and the consumer that all parties involved provide transparency into their respective processes.

Fighting collusion fraud

Gathering information across all service provider entities involved in the claim process is not only a technical challenge but, as noted above, it can be difficult to engage multiple parties in the initiative. We emphasize that this is permissible, has precedent, and ultimately is in the consumer's interest because it helps the insurance industry keep costs down. Once the data are available, the data have to be linked across all the service providers on the claims they serviced.

To illustrate, once the data are captured and processed into a usable form, then it should be possible for an analyst to generate a profile on a closed claim that includes the date of the claim, the claimant(s) involved in the incident, the incident report, the name of the clinic(s) that appraised the claimant(s), the details on the treatment plan filed, the date of the filing, the adjuster who approved the treatment plan, the amount of the approval, the days of the treatment, the nature of the treatment, and so on. Such a data structure is known as an analytical data mart and is well within the scope of the technology capabilities of insurers. The decision support system that detects collusion patterns has a dependency on this data mart. The specific analytical technique underlying the decision support is called Association Analysis [4].

As the name suggests, Associations Analysis is used to detect associations among the entities that have serviced claims for an insurer. To run this analysis, the investigator has to first identify the targeted outcome. The targeted outcome is a suspicious event of interest. For example, it can be defined as the claims on incidents with a high loss amount, or as incidents with more than two claimants. Alternately, if the investigator has already identified a particular set of claims as being suspicious, then he or she would want to identify these as the targeted outcomes for the algorithm.

The next task for the investigator is to select the attributes of the claim to be mapped against the outcome. These attributes are defined per the data availability on the service providers. These can be any or all of the following: the broker who sold the claim, the clinic that appraised the claimant, the auto body shop that handled the repairs, the paralegal engaged by the claimant, the agency from which the vehicle was rented by the primary claimant, and so on.

Once the outcome and the attributes are defined, the analysis outputs the combination of claim attributes that are the strongest leading indicators for the targeted outcome. In technical terms, the output is an “if/then” rule whose “head” is the combination of attributes under consideration, and whose “tail” is the targeted outcome, the suspicious event of interest. Association Analysis outputs the if/then rules that are most statistically significant. A typical output rule would show that ‘Doctor A_ and B_ Rehabilitation Clinic’ (the head) have a strong correlation with suspicious claims (the tail).

Since there can be a very large number of if/then rules generated in the system, the algorithm filters these on the basis of two measures known as “support and confidence.” Essentially support and confidence are technical definitions on which more detail is available in standard textbooks on the subject [5]. In practical terms, if a rule has low support it occurs so infrequently that it may just as well be a chance occurrence and is thus not statistically significant. Rules with high confidence indicate a high reliability — and a strong correlation between the outcome and the attributes that comprise the rule “head.”

The effectiveness of this technology is in the intuitiveness of the results. In the example of the medical mill listed earlier [1], the combination of specific doctors, chiropractors, acupuncturists linking to the same set of unusual claims would be a dead giveaway that a medical mill was at work.

Next steps

While the analysis is powerful, it must be cautioned that the rules that are discovered do not necessarily imply a causal relationship between the “if” condition and the targeted outcome. The technique helps in formulating a hypothesis that a particular group of service providers is engaged in suspicious activity. The next step is to flag this provider set to the field investigators and have them initiate their evidence collection to prove or disprove the hypothesis.

References

- [1] New York County District Attorney’s Office news release, March 11, 2008.
- [2] Uniroyal Goodrich Tire Company v Mutual Trading Corporation, Nos 94-2915 & 94-3799.
- [3] Gramm-Leach-Bliley Act, 15 USC, Subchapter I, Sec. 6801-6809 - Disclosure of Non-public Personal Information.
- [4] Nearhos, J.; Rothman, M.; and Viveros, M. 1996. “Applying data mining techniques to a health insurance information system”. In Proc. of the 22nd Int’l Conference on Very Large Databases.
- [5] P-N Tan, M. Steinbach, V. Kumar, “Introduction to Data Mining”, ISBN-10: 0321321367, Addison-Wesley, 2006.

Biography

Varun Madhok heads the Client Services team for Infernotions Technologies’ ClaimsGator offering for mitigating claims abuse for the property and casualty industry. He may be reached at 416.516.3795, varun@claimsgator.com, <http://claimsgator.com> .