



BIG DATA ANALYTICS: NEW WHISTLEBLOWER ON INSURANCE FRAUD

Sachin Pandhare

Abstract

As fraudsters get smarter and use technology advancements to their benefit, insurance companies face greater difficulties in preventing and detecting deception. Traditional methods of fraud detection often fail to identify these sophisticated attempts that can cost insurance providers millions of dollars. To stay a step ahead of these disruptive elements and prevent insurance crime, insurers must adopt new technologies in a big way. This paper discusses the role of big data analytics in detecting insurance fraud. It stresses the importance of building an IT environment to support the use of analytical insights obtained from myriads of data sources for successful fraud detection.



For all the tough measures, insurers lose millions of dollars to fraud every year. Apart from the carriers, the impact is also felt by policyholders as they end up paying higher premiums to compensate for the loss. According to the Federal Bureau of Investigation (FBI), the total cost of non-health insurance fraud is over \$40 billion per year, costing the average U.S. family between \$400 and \$700 per year in the form of increased premiums.¹

It is estimated that about 25% of claims contain some element of deception and account for around 10% of the total

insurance payout. Fraudulent claims can depict inflated losses or be based on staged mishaps to obtain the payout. Fraudsters can orchestrate accidents, undergo treatment for fictitious injuries, and claim large amounts as insurance payouts. Typically these include soft-tissue injuries, sprains and strains which are medically not easy to establish as false. Some drivers request their premiums to be reduced by deliberately manipulating details such as the number of miles driven, the use of the vehicle, etc. Arson, submitting duplicate claims – with slight

changes – for the same service, workers' compensation for bogus issues and unrelated injuries are also part of such attempts to deceitfully obtain insurance payouts.

With technology advancement, swindlers are adopting more sophisticated methods that can make frauds harder to detect. For instance, the anonymity provided by the Internet can at times allow professional criminals to take on fake identities and escape easily.

1. <http://www.fbi.gov/stats-services/publications/insurance-fraud>

2. <http://www.statsoft.com/Textbook/Fraud-Detection>

Traditional methods to counter new forms of fraud

Traditional methods are mostly rules-based systems where algorithms are used to search for anomalies, intrusions or unusual patterns in data or to detect patterns that are excessively regular and thus a cause for further investigation. Software solutions based on prediction scorecards and profiling use transaction and profile data to indicate the default propensity of a customer for every new transaction.

Often rigorous detailed checks and compilation of claimant and witness profiles are conducted for individual customers. Most such activities are performed manually. This includes verifying information presented by the claimant, referring suspicious claims to experts, investigating the situation, visiting the place of accident to collect evidence, and mapping of claims to check any link between the location of incidents and claimants. These manual checks are time consuming, people dependent, prone to slips, and not comprehensive.

Even when insurance companies use fraud detection tools, the focus of the solution is often on a specific area such as claim management and the entire framework is created around that area. This unbalanced approach can leave other areas in insurance vulnerable to deception. A holistic framework to detect fraud which integrates various aspects of insurance: claims, premiums, applications, employee information, and vendor details can offer all-around protection against scam attempts.

Human fraud detection techniques and lopsided technology solutions can be helpful to some extent in detecting fake claims at the individual customer level, but often fail to identify hoax activities and organized frauds which are a growing threat to the insurance industry today. In 2012, for instance, an auto insurance scam by the Russian mob claimed \$275 million in orchestrated car crashes and fake injuries in New York.³ Fortunately, insurers now have access to advanced technology and tools such as big data and predictive analytics to nip such crimes in the bud.



3. <http://thomsonreuters.com/press-releases/092013/insurance-fraud-white-paper>

The data surge

The volume of data generated today in various spheres of life is astonishing. The increased use of technology and affordable devices populating global markets and households are creating new data constantly. Data is streaming through computers, phones, telematics devices, televisions, credit cards, wearable devices, cameras, and even sensors fitted in buildings, to name a few sources. Experts

believe the amount of data amassed in the past two years is greater than all the data generated since the dawn of human civilization.⁴

The massive phenomenon has impacted businesses with every enterprise accumulating huge amounts of data. Many organizations, unable to handle big data using existing relational databases and unprepared to invest in new technologies, discard most data retaining

only summaries or very short histories. Some outsource data storage to a vendor and receive limited reports and analytical capabilities. Such practices overlook the value of data and fail to turn the challenge of big data management into an opportunity to strengthen decision-making through powerful information and insights.

4. <http://harvardmagazine.com/2014/03/why-big-data-is-a-big-deal>



Using analytical insights

Big data technology combined with advanced analytics can extract the true value of data being amassed by businesses. However, many insurers remain unsure about the transformational power of analytics. A global survey conducted by EY in 2013 of more than 100 insurers indicated low levels of digital maturity and the need to act. Almost 80% of insurers were lagging behind in analytics and adoption of social media.

Insurers must take another look at the power of data being produced at an unimaginable pace to provide priceless insights to their business. Fraud detection is one of the areas in insurance where big data technology can bring about revolutionary change. This can be achieved by matching the sophistication of organized crime through advanced predictive analytics models that have proven precision in predicting probable events.

Let us examine a few key big data models that can help insurers strengthen their fraud detection capabilities.

Social network analysis

The phenomenon of social media, with its expanding embrace, is in the process of connecting the world in a way that never existed before. A billion social-media posts every two days is a massive amount of data that creates numerous social links.

The use of network theory to scrutinize these connections, analyze social networks and view social relationships as diagrammatic representations has found application in many areas of security including countering money laundering, tracking terrorist activities, identifying insider trading, and spotting attempts at fraud.

Social network analysis (SNA) allows companies to proactively look through large amounts of data to show relationships via links and nodes. Take a simple accident insurance case that may seem quite straightforward in the beginning. The same case can prove to be fraudulent when SNA reveals that the address given by one of the victims has been used for multiple claims or the vehicle has been involved in several other incidents. The ability to quickly gather this information saves insurers time and alerts them to the hidden factors involved in a potential fraud case before the claim is settled.

Predictive modeling

Predictive analytics combines a variety of statistical and analytical techniques, including text analysis and sentiment analysis, to scan big data and predict future events or behavior. Claim reports usually run into several pages, making it time consuming and often humanly impossible to check every detail. Predictive analytics solutions can examine the unstructured data and enable proactive fraud detection by identifying patterns and relationships between people and data. The objective of using predictive modeling is to spot suspect activities as early as possible to prevent payouts to fraudsters, which might prove difficult to recover later.

By developing predictive models based on both historical and real-time data on wages, medical claims, attorney costs, demographics, weather data, call center notes, and voice recordings, companies can equip themselves better to identify fraudulent claims early on in the claims lifecycle. For example, if a worker's compensation claim for a severe eye injury is investigated using predictive analytics, it may reveal that the claimant discussed on a social forum a movie he or she watched on the same day, soon after the injury occurred. This helps the investigators examine the possibility of the claim details being incorrect before the payment process starts.





The example of Massachusetts Health Insurance Exchange illustrates the point further. After moving away from the traditional pay-and-pursue model, Massachusetts Health Insurance Exchange recovered \$2 million in improper payments and avoided paying hundreds of thousands of dollars in fraudulent claims in six months of using predictive modeling to counter fraud in 2013, says technology research and assessment firm GCN.⁶

Social customer relationship management

Many insurers have become social creating Facebook and Twitter pages as part of their digital marketing strategy. Using the information they gather from this social network to understand behavior trends and customer demands and preferences

constitutes social CRM. By investing in sophisticated analytics tools, insurers can use the vast amount of social CRM data for link analysis to fight organized crime and to detect fraudulent claims. A claimant enquiring on social forums about an expensive heritage property for sale just weeks before submitting his or her large claim for fire insurance, for instance, might make investigators want to dig a little deeper before the carrier pays out.

The analysis of claims, CRM and social CRM data can enable insurers to verify the validity of a claim through recent activities on social media sites when a claim is submitted. Social CRM tools are capable of searching millions of social media posts, including blogs, Facebook, Twitter, and other platforms to uncover information needed to investigate claims.

For example, in case of a motor accident claim, investigators can quickly answer questions such as: Is the claimant connected to the other person involved in the accident via a social networking site? Does third-party mobile phone information (GPS data) confirm where the claimant was at the time of the incident? The use of unstructured data on social media can allow claim adjusters to make more informed and quicker decisions on claims.

Telematics driving data analysis

The use of telematics as a means of recording vehicle data as well as driving behavior is catching on with insurers. Companies such as Tesco Bank in the UK, Progressive Corporation in the US, and Generali Group in Italy offer to fit telematic devices in policyholders' cars in order to

6. <http://gcn.com/Articles/2014/02/24/masshealth.aspx?Page=1>

provide usage-based products and to promote safe driving habits. However, the ability of telematics devices to provide vast amounts of real-time data on the vehicle also opens new possibilities for auto-insurers to verify claims and identify attempts at scam. Insurers now have access to facts about the vehicle and driving that can be analyzed to uncover the real story and catch fraudulent claims instead of paying out huge amounts in fake claims. For example, claims investigators can now verify whether the speed of the vehicle was within the prescribed limit, whether the traffic rules such as lane discipline were followed, if the brakes were applied appropriately, if the vehicle was in good condition and more. By harnessing the power of telematics analysis, insurers can now get clear answers in situations of ambiguity that could involve fraud.

In fact, telematics has evolved into a separate focus area due to the high potential for businesses to leverage the data generated by telematics devices in various areas such as insurance pricing, refining products and understanding customer behavior. Use of this data can help in fraud detection by supplementing expert judgement with plenty of minute details that were not available to investigators earlier.

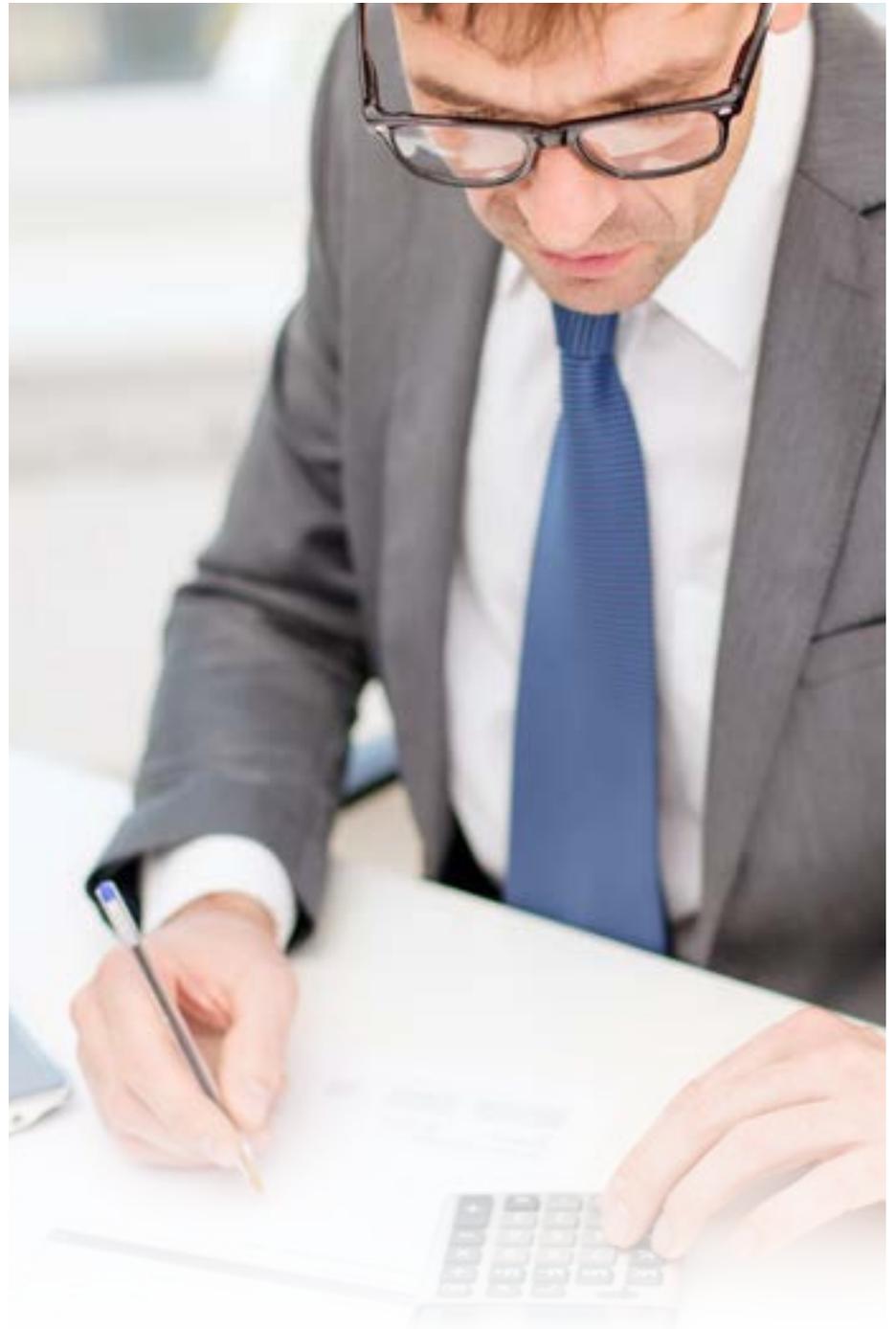
Data generated by different sources such as various social media platforms, claim reports, CRM, and telematics can be cross-leveraged to detect fraud more effectively. The fraudster is less likely to be consistent in providing the same false information on these varied forums. Analysis of this data can help companies find gaps and inconsistencies in information and lead to more effective detection and prevention.

Aware of the great potential of big data to prevent fraud, many insurers are looking to build adequate IT infrastructure to support big data analytics and move toward their objective of fraud-free payments.

Conclusion

Although the conservative nature of the industry has held insurers back from adopting new technologies, they need to review this attitude in order to compete and thrive in today's digital economy. This is critical – especially in areas such as fraud detection and prevention – at a time when tech-savvy crooks are often more than a step ahead of the law. Big data analytics in particular has tremendous potential

to spot clandestine fraudulent activities before it is too late. After implementing the web, mobile and social media aspects of their business strategy, insurers must look at big data analytics, predictive modeling, social CRM and telematics to fight sophisticated organized crime and smart fraudsters that have been siphoning off insurance profits quietly for years.



About the Author



Sachin Pandhare has over 22 years of IT and industry experience. He leads the innovation lab for Insurance, Cards and Payments vertical, in Infosys to drive various innovation initiatives for the unit. He focuses on delivering innovative technology solutions, leveraging the concepts of cloud, mobility, and analytics. Sachin has extensive experience in product development, delivering practice leadership, program management, and consulting to customers across the globe in banking and financial services, insurance, communications, energy, and other key industries.

For more information, contact askus@infosys.com



© 2018 Infosys Limited, Bengaluru, India. All Rights Reserved. Infosys believes the information in this document is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of other companies to the trademarks, product names and such other intellectual property rights mentioned in this document. Except as expressly permitted, neither this documentation nor any part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, printing, photocopying, recording or otherwise, without the prior permission of Infosys Limited and/ or any named intellectual property rights holders under this document.