

Using Analytics for Insurance Fraud Detection

3 innovative methods and a
10-step approach to kick
start your initiative

Ruchi Verma
Sathyan Ramakrishna Mani

If you've been used to thinking about analytics in terms of sales or marketing, think again. Today, analytics can reinvent your enterprise technologies — social networking, big data, CRM — to crack down on financial offenders. Giving you **more than an insight a day, to keep the fraud away.**

Digitization a new opportunity for fraud detection?

Digitization marked by a growing number of mobile devices and social media is changing the business landscape for all sectors — including insurance. The opportunities offered by this landscape for insurers are vast. Social networks and communities help insurers connect with their customers better, which in turn aids branding, customer acquisition, and retention. Insurance firms also receive a plethora of inputs from digital information in the form of feedback, which also can be used to come up with customized products and competitive pricing.

In addition to these opportunities, insurance companies are harnessing digitization — using data analytics for fraud detection. Handling fraud manually has always been costly for insurance companies, even if one or two low incidences of high-value fraud went undetected. In addition to this, the big data trend, (the growth in unstructured data) always leaves lot of room for a fraud going undetected if data is not analyzed thoroughly.

The big data trend, (the growth in unstructured data) always leaves lots of room for a fraud going undetected if data is not analyzed thoroughly

Traditionally, insurance companies use statistical models to identify fraudulent claims



Fraud detection by insurance companies

These models have their own disadvantages. First, they use sampling methods to analyze data, which leads to one or more frauds going undetected. There is a penalty for not analyzing all the data. Second, this method relies on the previously existing fraud cases, so every time a new fraud occurs, insurance companies have to bear the consequences of the first time. Finally, the traditional method works in silos and is not quite capable of handling the ever-growing sources of information from different channels and different functions in an integrated way.

Analytics addresses these challenges and plays a very crucial role in fraud detection for insurance companies. Some of the key benefits of using analytics in fraud detection are discussed below.



Identification of low-incidence events

Using sampling techniques comes with its own set of accepted errors. By using analytics, insurance companies can build systems that run through all critical data. This in turn helps detect low-incidence (0.001%) events. Techniques such as predictive modeling can be used to thoroughly analyze instances of fraud, filter obvious cases, and refer low-incidence fraud cases for further analysis.



Enterprise-wide solution

Analytics help in building a truly global perspective of the anti-fraud efforts throughout the enterprise. Such a perspective often leads to effective fraud detection by linking associated information within the organization. Fraud can occur at a number of source points: claims or surrender, premium, application, employee-related or third-party fraud. At the same time, insurance channel diversification is adding to the fragmentation of traceable data. Insurance-related activities can be done via mobile devices apart from the traditional online and face-to-face insurance. This can be viewed as an addition to information silos in the insurance industry. Given greater channel diversification and the increase in areas where fraud can occur, it is important for insurers to have accessible enterprise-level information about their business and customers.



Data integration

Analytics plays an important role in integrating data. Effective fraud detection capabilities can be built by combining data from various sources. Analytics also help in integrating internal data with third-party data that may have predictive value, such as public records. Data sources with derogatory attributes are all public records that can be integrated into a model. Examples include bankruptcies, liens, judgements, criminal records, foreclosures, or even address change velocity to indicate transient behavior. Other types of third-party data can be beneficial in enhancing efficiencies such as review of appraisal information to determine if damages match description or loss or injuries being claimed. One of the most under-utilized data sources is medical bill review data. This data, if used in a model properly, is a gold mine for companies investigating medical fraud. Uncovering anomalies, in billing and adding these to the other scoring engines or social network analysis will decrease the amount of time an investigator or analyst spends trying to pull all of the pieces together to identify fraudulent activity.



Harnessing unstructured data

Analytics helps in deriving the best value from unstructured data. Fraud can be soft fraud or hard fraud. This is based on whether it consists of a policyholder's exaggerated claims, or if it consists of a policy holder planning or inventing a loss. At a high level, fraud can occur during commission rebating, due to false documentation, collusion between parties or from mis-selling. Although lots of structured information is stored in a data warehouse as part of many applications, most of the crucial information about a fraud is in unstructured data, such as third party reports, which are hardly analyzed. In most insurance firms, information available in social media is not appropriately stored. A special-investigative-unit investigator will agree that unstructured data is very important for fraud analysis. Since textual data is not directly used for reporting, it does not find a place in most data warehouses. This is where text analytics can play a key role in reviewing this unstructured data and providing some valuable insights in fraud detection.

Three innovative fraud detection methods

1. Social Network Analysis (SNA)

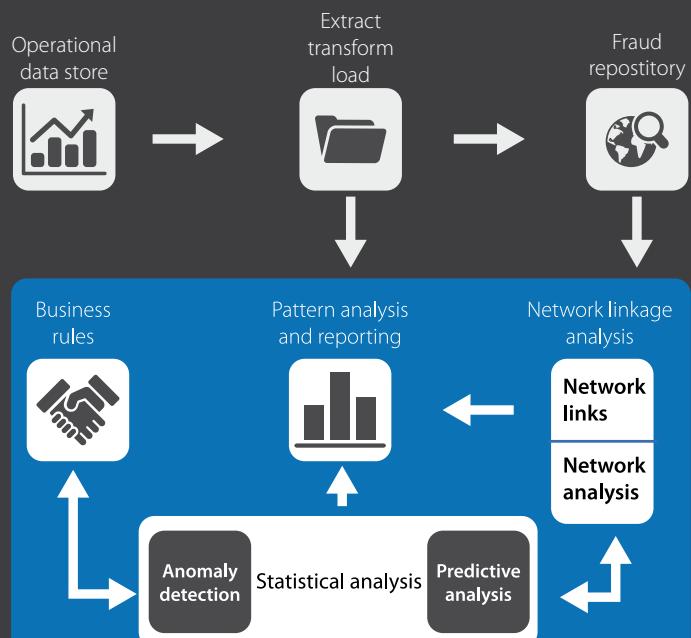
Let's take an example to explain the use of social network analysis (SNA). In a car accident, all people in the vehicle have exchanged addresses and phone numbers and provided them to the insurer. However, the address given by one of the accident victims may have many claims or the driven vehicle may have been involved in other claims. Having the ability to cull this information saves time and gives the insurer an insight into the parameters involved in the fraud case. SNA allows the company to proactively look through large amounts of data to show relationships via links and nodes.

The SNA tool combines a hybrid approach of analytical methods. The hybrid approach includes organizational business rules, statistical methods, pattern analysis, and network linkage analysis to really uncover the large amounts of data to show relationships via links. When one looks for fraud in a link analysis, one looks for clusters and how those clusters link to other clusters. Public records such as judgments, foreclosures, criminal records, address change frequency, and bankruptcies are all data sources that can be integrated into a model.

Using the hybrid approach, the insurer can rate these claims. If the rating is high, it indicates that the claim is fraudulent. This may be because of a known bad address or suspicious provider or vehicle in many accidents with multiple carriers.

SNA follows this path:

1. The data (structured and unstructured) from various sources is fed into the extract transform and load tool. It is then transformed and loaded into a data warehouse.
2. The analytics team uses information across a wide variety of sources and scores the risk of fraud and prioritizes the likelihood based on multiple factors. The information used can range anywhere from a prior conviction, a relationship in some manner to another individual with a prior case, multiple rejected claims, odd combinations of data, or even odd modifications to personal information.
3. Technologies such as text mining, sentiment analysis, content categorization and social network analysis are integrated into the fraud identification and predictive modeling process.
4. Depending on the score of the particular network, an alert is generated.
5. The investigators can then leverage this information and begin researching more on the fraudulent claim.
6. Finally, issues or frauds that are identified are added into the business use case system, which is a part of the hybrid framework.



Before implementing SNA, insurers should consider:

1. How fast data arrives
2. How clean the data is when it arrives
3. How deep the analysis must go to get the results
4. What type of user interface components need to be included in the SNA dashboard

Case study: GE Consumer & Industrial Home Services Division

Scenario

In GE Consumer & Industrial Home Services Division, claims typically came from technicians who repair consumer products that are under warranty. One of the biggest problems with their old process was that they could not identify patterns. With the amount of data available to them, no one could see unusual behavior emerging. Sometime back, GE got the perfect scenario to test an SNA solution from SAS, a developer of business analytics software. The company was tipped off to some service providers committing fraud. This situation made for an ideal pilot scenario. SAS was given the responsibility of analyzing the available data and identifying patterns in the data to find out who was committing the fraud.

Functioning of the fraud detection system

Typically, there are some metrics and indicators on every claim that assist in identifying suspicious or fraudulent claims. GE's claims data is fed into the fraud detection software. There are 26 claim-level analyses, which are automatically calculated for each claim. There are some indicators like flags that are calculated based on various metrics and sent for auditing when they indicate that multiple elements in the claim fall out of the normal curve. Once these claims are flagged, the auditors at GE investigate these suspicious claims.

Outcome

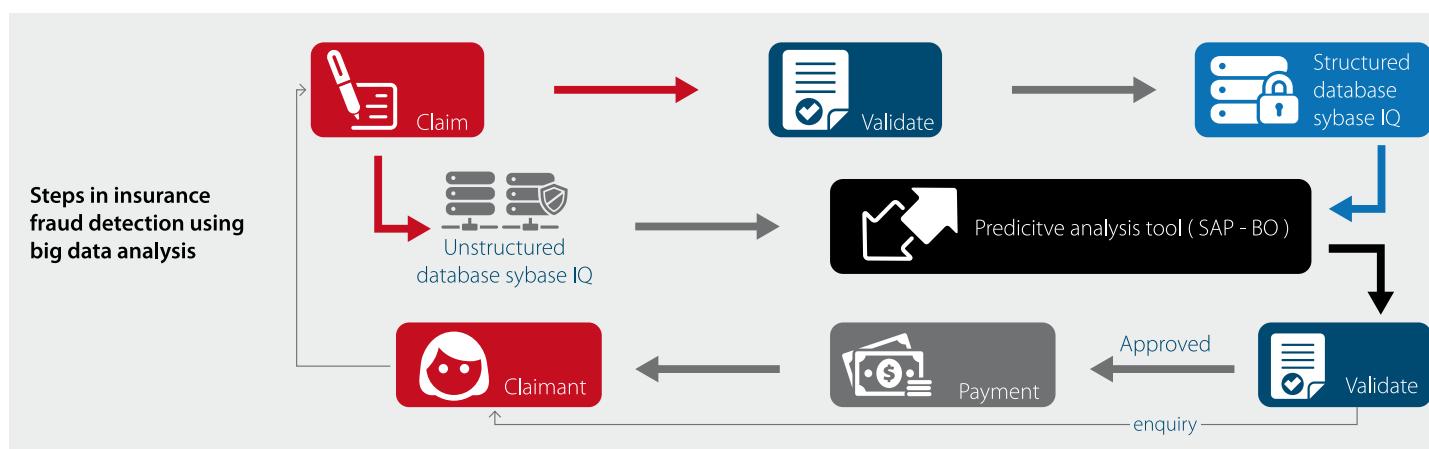
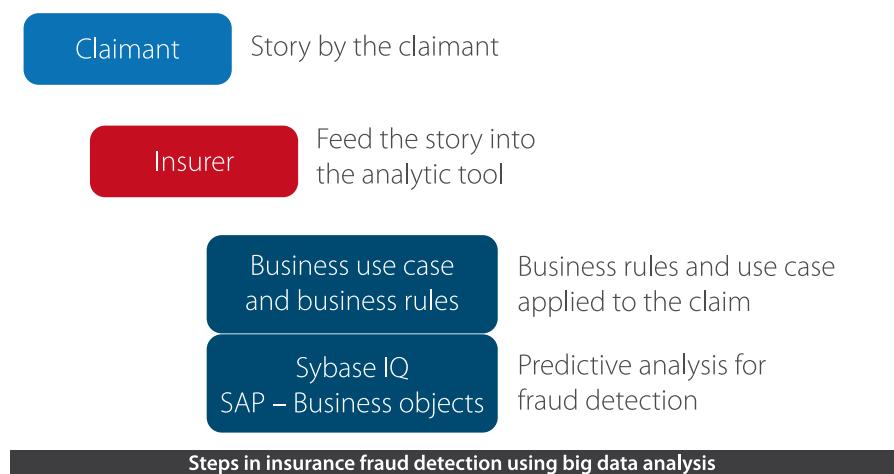
The GE Consumer & Industrial Home Services Division estimated that it saved about \$5.1 million in the first year of using SAS, to detect suspect claims.

2. Predictive analytics for big data

Consider a scenario when a person raises a claim saying that his car caught fire, but the story that was narrated by him indicates that he took most of the valuable items out prior to the incident. That might indicate the car was torched on purpose. Predictive analytics include the use of text analytics and sentiment analysis to look at big data for fraud detection. Claim reports span across multiple pages, leaving very little room for text analytics to detect the scam easily. Big data analytics helps in sifting through unstructured data, which wasn't possible earlier and helps in proactively detecting frauds. There has been an increase in the use of predictive analytics technology, which is a part of big data analytics concept, to spot potentially fraudulent claims and speed the payment of legitimate ones. In the past, predictive analytics were used to analyze statistical information stored in the structured databases, but now it is branching out into the big data realm. The potential fraud present in the written report above is spotted using text analytics and sentiment analysis.

Here's how the text analytics technology works:

- Claim adjusters write long reports when they investigate the claims
- Clues are normally hidden in the reports, which the claims adjuster would not have noticed
- However, the computing system, which is based on business rules, can spot evidence of possible fraud
- The most important point to observe is that people who usually commit fraud alter their story over time. The fraud detection system can spot these discrepancies



Case study: Infinity Insurance Co.

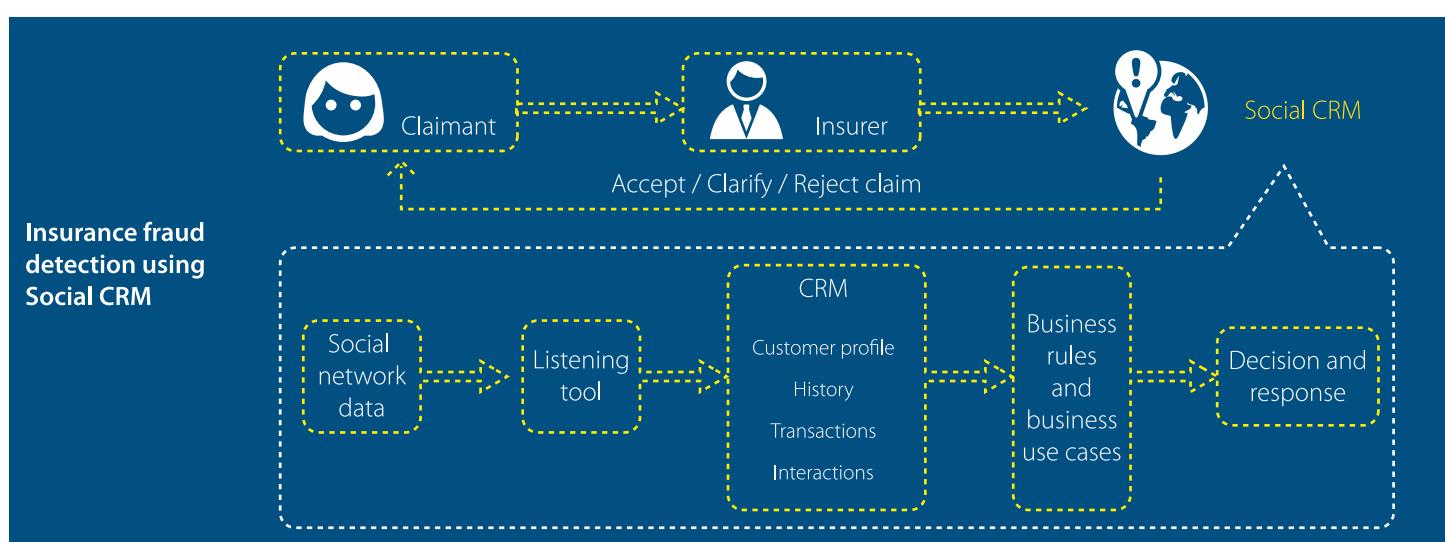
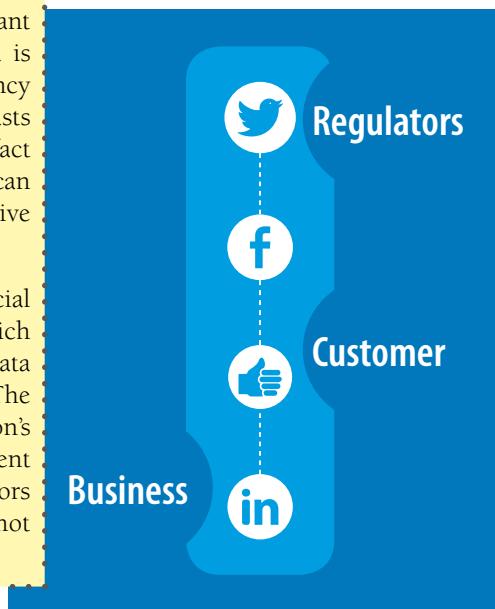
Infinity, a property and casualty company, came up with the idea of ‘scoring’ insurance claims from customers to look for signs of fraud. Its target market is mainly drivers who have higher than normal risks and pay high rates compared to others. With the kind of exposure Infinity has, spotting insurance fraud, either while raising the claim or while calculating the premium to be paid, is even more important than it is to other insurance companies. Infinity uses a predictive analytics technology to spot potentially fraudulent claims and speed the payment of legitimate ones.

After using predictive analysis, the claims fraud system increased the success rate in pursuing fraudulent claims from 50–88 % and reduced the time required to refer questionable claims for investigation by as much at 95%.

3. Social customer relationship management (CRM)

Social CRM is neither a platform nor a technology, but rather, a process. It is important that insurance companies link social media to their CRM. When social media is integrated within multiple layers of the organization, it enables greater transparency with customers. Mutually beneficial transparency indicates that the company trusts its customers and vice versa. This customer-centric ecosystem reinforces the fact that increasingly the customer is in control. This customer-centric ecosystem can be beneficial to the business as well, if the business is able to leverage the collective intelligence of its customer base.

Social CRM uses a company’s existing CRM and gathers data from various social media platforms. It uses a ‘listening’ tool to extract data from social chatter, which acts as reference data for the existing data in the current CRM. The reference data along with information stored in the CRM is fed into a case management system. The case management system then analyzes the information based on the organization’s business rules and sends a response. The response from the claim management system as to whether the claim is fraudulent or not, is then confirmed by investigators independently, since the output of social analytics is just an indicator and should not be taken as the final reason to reject a claim.



Case study: AXA OYAK, Turkey

AXA OYAK is a Turkish insurance company that has been using the SAS Social CRM solution to manage risk and prevent fraud. AXA OYAK built an intelligent enterprise around social CRM in such a way that it integrates all customer-related information into a single and coordinated corporate vision.

Using its social CRM, AXA was able to clean up their customer portfolio data. This helped them find and correct inconsistencies in this data, which enables AXA to link two slightly different records to the same customer. With cleaner data, AXA can run more accurate customer analysis and investigate fraudulent claims more efficiently. Using SAS, AXA OYAK was quickly able to find the relationships between customer behavior and fraudulent claims. With the SAS data warehouse, AXA is able to segment their customer data based on flags that are generated while analyzing certain relationships between data sets.

A 10-step approach to implement analytics for fraud detection

1

Perform SWOT

Insurance companies are realizing the importance of analytics in the fraud detection space and hurriedly opting for expensive fraud solutions that are not aligned to the company's weakness and strengths. In order to leverage analytics solutions to the fullest, insurance companies should first do a SWOT analysis of existing fraud detection frameworks and processes to identify gaps.

2

Build a dedicated fraud management team

Usually, in a traditional insurance company, no specific team or person is proactively accountable for fraud detection. When fraud is detected internally, people point fingers, raise alarms and take measures to fight it. It is important that a dedicated team is identified and made accountable for fraud detection. The team should report to senior management for necessary buy in.

3

Whether to build or buy

Once the SWOT is complete and a team of dedicated people for fraud detection have been identified, insurance companies should review how they want to implement analytics and what data sources they want to analyze. Insurance firms need to be honest in answering whether the skill set for building analytics solutions are available in-house or whether there is a need to buy an analytical fraud detection solution from an external vendor. If there is a need to buy the analytics solution, insurance firms should evaluate different analytics vendors in the market to find a solution that best fits the company's requirements. Key parameters to judge an external vendor are cost, user interface, scalability, ease of integration and ability to add new data sources.

4

Clean data

Integrate siloed databases and remove inefficiencies from processes and redundancies from data sources.

5

Come up with relevant business rules

Insurance companies should leverage existing domain expertise and experienced resources to come up with business rules. Certain types of fraud are very specific to the industry and, in some cases, certain companies. Without inputs from in-house capabilities, it will be difficult for any internal or external team to build a robust fraud detection solution.

6 Come up with pre-determined anomaly detection thresholds

Whether the analytics framework is built in-house or by using a third-party vendor, insurance companies should provide inputs for threshold values for different anomalies. The number of claims received for life insurance is different from the number of claims received in non-life insurance. Key performance indicators associated with tasks or events are baselined and thresholds are set using anomaly detection. Setting the threshold is a major decision in anomaly detection. If thresholds are set too high, too many fraudulent claims could slip through the system. When thresholds are set too low, there can be risks of wasting time, alienating members and providers, and can result in late-payment penalties. Certain statistical analyses take an empirical value by determining ‘normal’ ranges for predetermined metrics.

7 Use predictive modeling

An important fraud detection method is one that utilizes data mining tools to build models that produce fraud propensity scores linked to unidentified metrics. Claims are automatically scored to look for any indication of a discrepancy or fraud. After this, the results are made available for review and further analysis.

8 Use of SNA

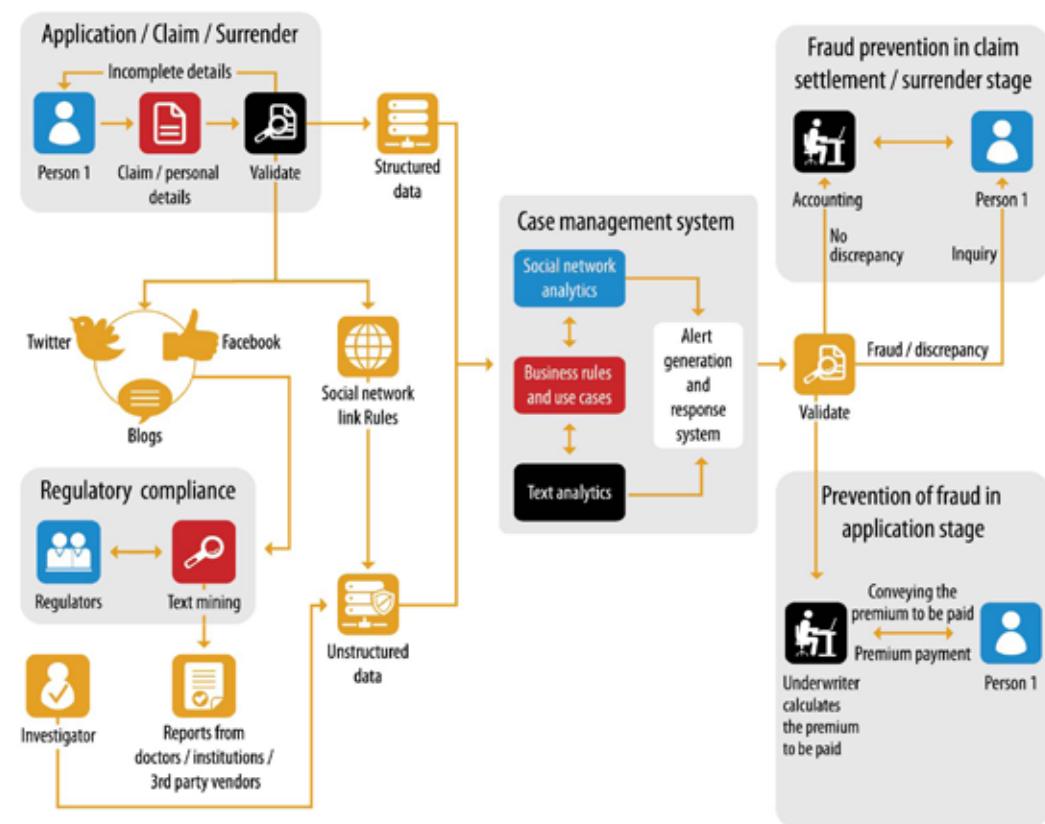
SNA has proven effective in identifying organized fraud activities by modeling relationships between various entities involved in the claim. Entities can range anywhere between locations to telephone numbers. The number of linkages between certain types of entities may be found to be much greater than the average number of connections expected based on statistical analyses of other ‘networks’ of entities.

9 Build an integrated case management system leveraging social media

Integrated case management capabilities allow investigators to capture all key findings that are relevant to an investigation, including claims data, network diagrams, adjuster notes, and social media, which can contain structured or unstructured data. Metrics are the key indicators of fraud or abuse and can be automatically tabulated for comparison at the individual entity or network level (using the anomaly threshold or SNA). Case workflow enables a full and complete assessment of investigative workload, efficiency, and return on investment.

10 Forward-looking analytics solutions

Insurance companies should keep looking for additional sources of data and integrate those with existing fraud detection solutions, for building the most efficient fraud detection system possible to address a variety of new frauds that may emerge in the future.



The proposed system can

- Rapidly organize and analyze the unstructured data present in the claims submitted by the claimant, notes of the claim adjuster and third-party reports
- Examine the sentiments of the claimant to help drill down to the specific concerns that bother at-risk customers
- Synthesize complex fraudulent patterns that contain the presence of multiple red flag indicators
- Detect and provide early warning of potential issues before they become problems
- Uncover early patterns in fraudulent activity

The way forward

Insurance firms always hesitate in implementing analytics because of the initial time investment needed for analytics solutions. However, it has been seen that analytics goes a long way in detecting fraud proactively and earlier in the insurance lifecycle. It culminates in reducing the overall cost of fraud detection and improving the overall ROI of insurance fraud solutions.

Insurers must now exploit the existing data in any form (structured or unstructured) by using analytics to effectively detect, manage, and report frauds. The earlier the fraud is detected in the insurance lifecycle, the lesser it costs to manage it. Analytics can play a very important role in identifying fraud early in the insurance lifecycle, and failing to act on this opportunity could quite literally equate to a gargantuan loss.

About the Authors



Ruchi Verma
Senior Consultant, Financial Services and Insurance Unit

She has around eight and half years of experience in Infosys in varied roles across multiple accounts. Her areas of interest includes emerging trends and regulations in the financial services and insurance domain.

She can be reached at ruchi_verma@infosys.com



Sathyan Ramakrishna Mani
Senior Associate Consultant, Financial Services and Insurance Unit

He has close to three years of experience in varied roles across multiple accounts. His interests are in the area of capital markets. He is also a keen follower of macroeconomic events that take place around the world.

He can be reach at sathyan_mani@infosys.com

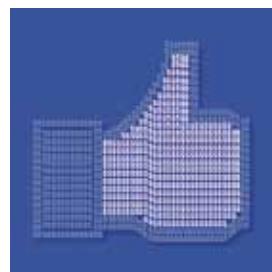
Other articles in this edition of FINsights



Serve them all,
serve them well



Multi-channel user
experience – a banking
perspective



Banking on customer
satisfaction in a
digital world



Enhancing customer
experience with
immersive
correspondence

Like this
article?
Share it
with your
network



Bgff[Y S` S_ Wfa
SXSW



? a` W[Sf[a` , ^aa] [Y
SffZWgfgd\laX
eWZewh[UWTS'] [Y



More the merrier:
the hacker's motto



Financial transactions
get personalized and
secure with biometrics



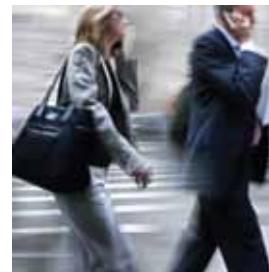
Go digital, reduce fraud



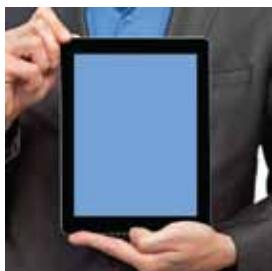
Financial institutions
reduce fraud risk with
social media



Reinventing bank
marketing with mobility



Delivering 'on the go'
services for employees
and customers



Just what the insurer
wanted: a 'tabletized'
future



The social makeover of
the financial services
industry



Digital transformation
framework



Straight-through
processing



Read previous
article



Read next
article