Healthcare Fraud Management using BigData Analytics

A Whitepaper by Trendwise Analytics

Venkata Reddy Konasani
M.Sc (Applied Statistics), IIT Bombay

Mukul Biswas
M.Tech (Computer Sc.) Jawaharlal Nehru University

Praveen Krishnan Keloth
M.Sc (Communication Engineering), Univ of Plymouth

Abstract:
Traditional Healthcare Fraud Management analytical solutions are based on enterprise data which is limited in many respects - variety of data, processing speed and analytical algorithms. Since BigData platform can process enterprise as well as social media data with ease, it is prudent to base such a solution on BigData platform. This paper provides one of many approaches to utilize fraud management solution to detect potential frauds. The solution is based on a high volume of historical data, predictive statistical models and social media analytics. It renders its services through client components like apps and web-services

Healthcare insurance suffering from Fraud

The healthcare domain has been an easy target for people who seek easy money by using fraudulent methods. Healthcare fraud is expected to continue to rise as people live longer. This increase will produce a greater demand for Medicare benefits. As a result, it is expected that the utilization of long and short term care facilities such as skilled nursing, assisted living, and hospice services will expand substantially in the future. Additionally, fraudulent billings and medically unnecessary services billed to health care insurers are prevalent throughout the world. These schemes are becoming increasingly complex and can be perpetrated by corporate-driven schemes and systematic abuse by certain provider types. The Figure on the left shows the fraud restitutions from 2001 to 2004 in Billions.

Fraud Restitution in Billions

Year	Value
2001	0.5
2002	0.3
2003	1.1
2004	1.1
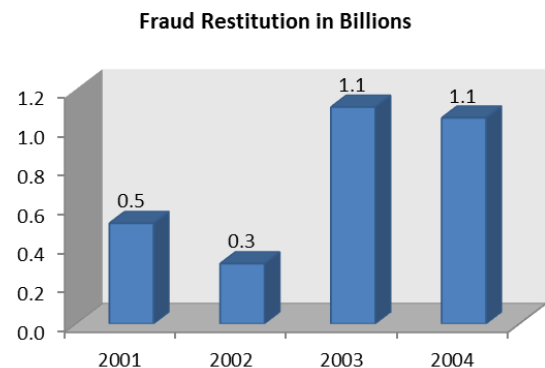
Appendix A provides a sample of the nature of fraud prevalent in Healthcare system.

Healthcare fraud prevention has resulted savings of nearly $4.1 billion

Approximately $4.1 billion stolen or otherwise improperly obtained from federal healthcare programs was recovered and returned to the Medicare Trust Funds, the Treasury and others in 2011(HHS 2012)

Health insurance fraud costs the industry an estimated \$70 billion to \$260 billion in the US and \$30 billion to \$100 billion in the EU each year(SAS 2012). Historical fraud detection methods only uncover about 10 percent of losses, and because of the post-payment nature of such methods and the resulting pay-and-chase recovery process, less than 5 percent of losses detected are ever recovered. Fueled by technology advancements that have made crimes such as identity theft and multiparty fraud schemes both easy to commit and hard to detect, healthcare fraud continues to grow. It holds particular appeal for organized crime syndicates, which account for a growing proportion of healthcare fraud, waste and abuse.

The Health Care Fraud and Abuse Control Program (HCFAC) has been awarded \$350 million to fight and also to provide additional tools and resources to help fight fraud.

Current form of Fraud Management is based on enterprise data and heuristics -

Fraud detection & prevention is mainly executed in two methods - fraud audit rules and fraud prediction score card. Fraud audit is the most widely used method compared to prediction scorecard.

More on audit rules and prediction scorecard can be found in Appendix A.

Audit rules method involves tedious manual work; one needs to audit all the claims one by one to detect the fraud. The final judgment is taken by the auditor which leads to judgmental error & inconsistency between two auditor judgments. While audit rules do not rely on process automation, prediction scorecard is based on computer based statistical analysis; which unfortunately, suffers from lack of real time data. The models built on historical data, which tend to lose their prediction power beyond certain extent. Swindlers keep updating their fraudulent activities more frequently than the models are updated.
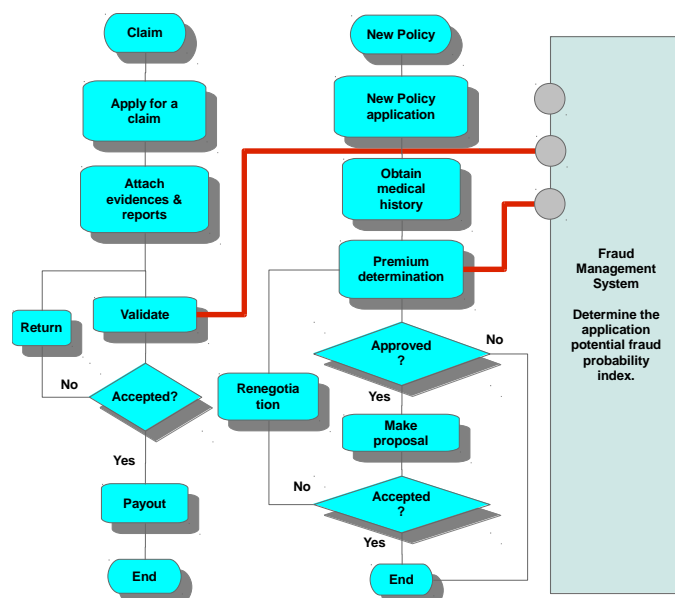

Figure: Current form of Healthcare Fraud Management

The quality of the outcome of the predictive analysis depends on the quality of historical data. A data set which covers a wide range of cases can made better predications. Also, analytical processing with more influencing factors (statistically speaking - explanatory variables) results with higher confidence. *Read more on explanatory variables in Appendix B.* In order to exploit both of these, one needs a more powerful computing platform. BigData platform is not only capable of processing terabytes or petabytes of data but also supports massively parallel processing.

Advent of BigData analytics makes Healthcare fraud detection more reliable and quicker

Healthcare claim system leverages the power of BigData platform by delegating its analytical needs. A BigData platform (as shown in the figure) has ability to sift through a huge amount of historical data in relatively shorter amount of time, so that the business transactions can use fraud detection on real time.

Typically, a BigData platform based FMS is hosted as a 3rd party web-service which is interacted on need basis. The FMS service provider manages the entire infrastructure with an assured quality of service (QoS). These services are, usually, available as software-as-a-service (SaaS) on pay as you go basis.
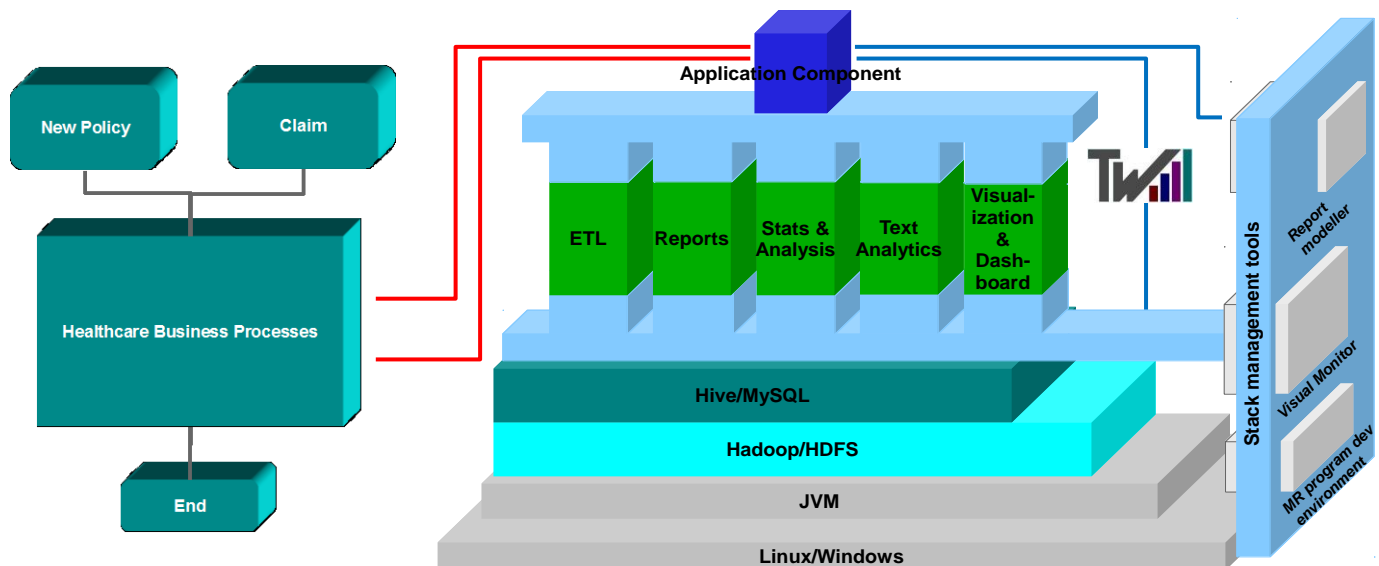

Figure: Using BigData Analytics in Fraud Management

Following are the benefits of a BigData based FMS over a traditional process –
1. Can process a large volume of historical data using complex algorithms
2. Business process automation can run almost real-time analysis before proceeding to the next step
3. Analysis is based on both enterprise and unstructured information from the Net.
4. Service is available online as a web-service which does not mandate software installation on the client premise - "Pay as you use".
5. FMS undergoes a continuous self-learning on the basis of transaction made on an ongoing basis.

Appendix A:

1. Doctors, who treated whopping, say 50+ patients in a day.
2. Providers administering far higher rates of tests than others.
3. Providers costing far more, per patient basis, than others.
4. Providers with high ratio of distance patients.
5. Providers prescribing certain drugs at higher rate than others.
6. Distance between claimant's home address and medical provider
7. Multiple medical opinions/providers
8. Changing providers for the same treatment (possibly correlated with other claim activity)
9. High number of treatments for type of injury
10. Abnormally long treatment time off for the type of injury
11. Accident severity does not correlate with severity of injury
12. Providers billing for services not provided.
13. Providers administering (more) tests and treatments or providing equipments that are not medically necessary.
14. Providers administering more expensive tests and equipments (up-coding).
15. Providers multiple-billing for services rendered.
16. Providers unbundling or billing separately for laboratory tests performed together to get higher reimbursements.
17. Providers charging more than peers for the same services.
18. Providers conducting medically unrelated procedures and services.
19. Policy holders traveling long distance for treatment which may be available nearby. (Possibly scams by bogus providers.)
20. Policy holders letting others use their healthcare cards.

Appendix B:

Fraud Audit Rules: The audit rules or general check points are designed purely based on previous experience and intuition. Audit rules are used to check the genuineness fields like Total amount billed, Total number of patients, Total number of patient visits, Per-patient average billing amounts, abnormally long treatment time off for the type of injury, Providers charging more than peers for the same services and so on. Each application is checked against these rules, if a claimant fails to satisfy one or more rules his/her claim will be rejected or subjected to further scrutiny.

Fraud Prediction Scorecard: In this method we predict the probability of fraud by using a mathematical predictive model. The model is built based on historical data which involves fraud or non-fraud indicator along with other explanatory elements like billed amount, number of patients, Reporting Lags, Treatment Characteristics, Years of Experience of doctor, Number of Prior Incidents from doctor etc.,. We use advanced statistical techniques to find the patterns in the historical data and come up with a predictive model. The model takes above explanatory fields (found in claim application) as input and gives a predicted fraud score as output.

References
HHS 2012: http://www.hhs.gov/news/press/2012pres/02/20120214a.html
SAS 2012: http://www.sas.com/industry/healthcare/insurer/fraud-detection.html
FBI 2005:http://www.fbi.gov/stats-services/publications/fcs_report2005#HEALTH%20CARE%20FRAUD

Trendwise Analytics
109 Regent Prime
48 Whitefield Road
Bangalore 560 066 INDIA

info@trendwiseanalytics.com
http://trendwiseanalytics.com