Large Scale Medical Database Mining :Abuse and Fraud Detection

Clinical Research Informatics & Bio-Banks *

Hyunjung (Helen) Shin
Dept. of Industrial and Information Systems Engineering,
Ajou University, Korea
shin@ajou.ac.kr
http://www.alphaminers.net

March 26, 2010

Abstract

Medical abuse and fraud has incurred physical and fiscal cost to society, which has been worsened with aging population and advancing health technology. What the detection of abusive and fraudulent practice in healthcare makes difficult and unique is the major source of the problem lies in uncertainties involved in the practice of medicine, which result in variations in care processes. Therefore, the detection requires reviews by medical experts which can be quite time consuming and expensive. The advancement in information technology and digitization of healthcare information such as electronic medical records and electronic bills and claims opened a new venue for efficient and effective mechanisms of medical abuse and fraud detection. The objective of this talk is to introduce a novel model (in use since July 2009 by the Korean Health Insurance Review & Assessment Service, HIRA) that detects healthcare providers who show the pattern of abusive and fraudulent behavior in the provision of healthcare. The proposed model is designed to process large amount of information contained in healthcare insurance claims automatically and to output an index to be used for the decision whether a provider merits further investigation and intervention to change and prevent its abusive utilization behavior.

*2010 The Fourth Conference on Biomedical Informatics, Seoul, Korea

Large Scale Medical Database Mining:
Abuse and Fraud Detection

Hyunjung (Helen) Shin

Dept. of Industrial & Information Systems Engineering
Ajou University, Korea

shin@ajou.ac.kr
http://www.alphaminers.net
http://www.kyb.tuebingen.mpg.de/~shin

The authors would like to gratefully acknowledge support from the research grant from the **Health Insurance and Review Agency (HIRA 2007-82)**, National Research Foundation of the Korean Government **(2009-0065043)**, and **POST Brain Korea 21**.

Abstract

To detect fraudulent and abusive bill claims of medical care providers, a variety of indexes have developed and evaluated diverse aspects of bill claim pattern.

When taking all of indexes into consideration, however, it becomes confusing to find out **which index is of more importance than others**, and even more difficult if the respective results are significantly discordant.

To avoid the ambiguities, we propose **a method integrating the diverse degrees of anomaly** based on 2007 Korean HIRA data.

HIRA: Health Insurance Review and Assessment Service

Outline

Background	**Healthcare Fraud?**
	National Health System in Korea
	Difficulties in Review Process for Medical Bill Claims
	Problem Diagnosis
Method	**Scoring Model**
	Segmentation
Evaluation	**Test Run & Validation Results**
Conclusion	**Conclusion & Further work**
	References

Conclusion: Segmentation

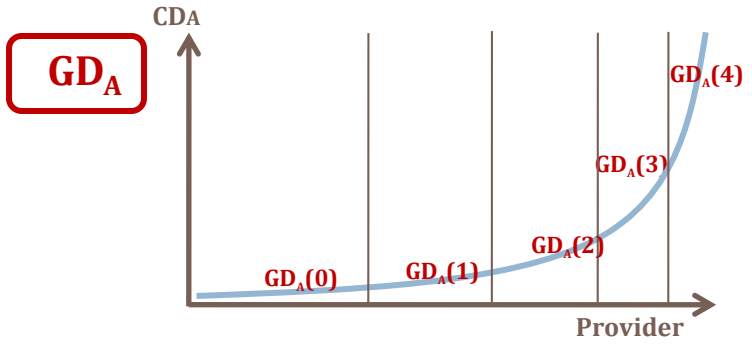
Scoring Indicator

- Measuring "**degree of anomaly**"
- **Integrating** most of the indicators

Segmentation

- **GD_A based on CD_A score**
- Gaining interpretability for the segments by re-classifying the GD_A groups using Decision Tree

- Scalable
- Simple
- Pragmatic



Conclusion: Segmentation

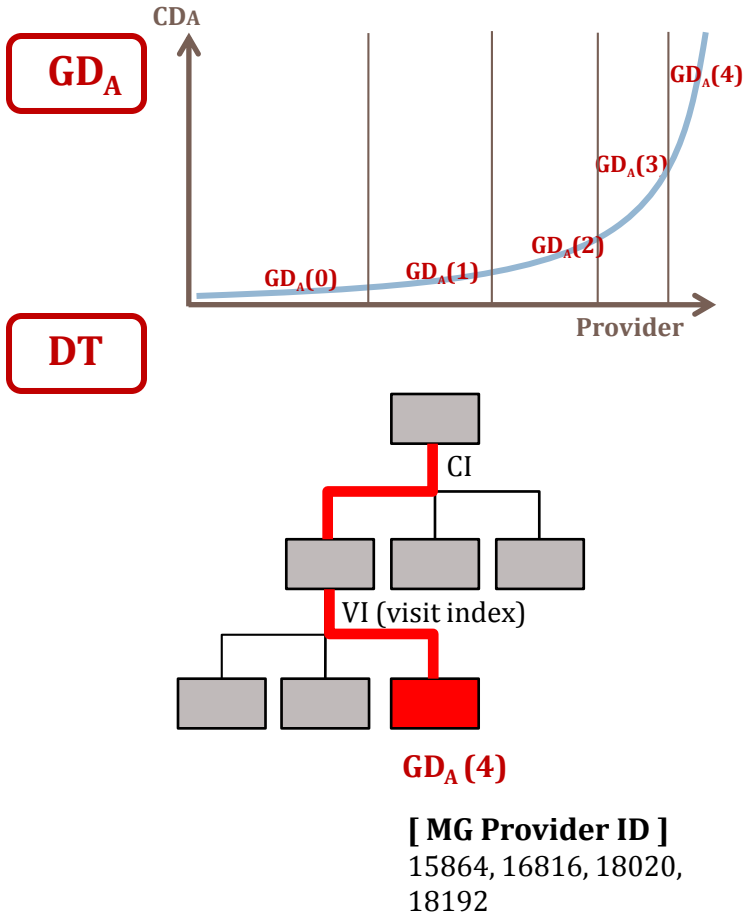
Scoring Indicator

- Measuring "**degree of anomaly**"
- **Integrating** most of the indicators

Segmentation

- GD_A based on CD_A score
- **Gaining interpretability for the segments** by re-classifying the GD_A groups using Decision Tree

- Scalable
- Simple
- Pragmatic



[MG Provider ID]
15864, 16816, 18020, 18192

Conclusion: Segmentation

Scoring Indicator

- Measuring "**degree of anomaly**"
- **Integrating** most of the indicators

Segmentation

- GD_A based on CD_A score
- Gaining interpretability for the segments by re-classifying the GD_A groups using Decision Tree

- Scalable
- Simple
- Pragmatic

Scailability

Simplicity

Pragmatic

YES!

Immediate application on the domain

Since July 2009, HIRA has taken immediate action on the proposed method, by introducing it to its hospital review and assessment system in order to prevent abusive and fraudulent medical-care bill claims.

Acknowledgement

The authors would like to gratefully acknowledge support from the research grant from the **Health Insurance and Review Agency (HIRA 2007-82)**, National Research Foundation of the Korean Government (2009-0065043), and POST Brain Korea 21.

References

References

- OECD. Health at a Glance 2007: OECD Indicators. OECD, 2007:85-89. Available from: http://www.sourceoecd.org/socialissues/9789264027329; accessed Aug 12, 2009.
- OECD. OECD Health Data 2009: Frequently Requested Data. OECD, 2009. Available from: http://www.oecd.org/document/16/0,3343,en_2649_34631_2085200_1_1_1_1,00.html; accessed Aug 12, 2009.
- Pontell HN, Jesilow PD, Geis G. Policing physicians: Practitioner fraud and abuse in a government medical program. Social Problems 1982;30(1):117-25.
- Feldman R. An economic explanation for Fraud and abuse in public medical care programs. J Legal Studies 2001;30(2):569-77.
- Rai AK. Health care fraud and abuse: A tale of behavior induced by payment structure. J Legal Studies 2001;30(2):579-87.
- Shane R. Detecting and preventing health care fraud and abuse-we've only just begun. Am J Health-Syst Pharm 2000;57:1078-80.
- Center for Medicare and Medicaid Services. Medicare Claims Review Programs: MR, NCCI Edits, MUEs, CERT, and RAC. CMS, 2008. Available from: http://www.cms.hhs.gov/MLNProducts/downloads/MCRP_Booklet.pdf; accessed Aug 12, 2009.
- The Problem of Health Care Fraud. Available from: http://www.nhcaa.org/eweb/DynamicPage.aspx?webcode=anti_fraud_resource_centr&wpscode=TheProblemOfHCFraud; accessed Aug 12, 2009.
- Center for Medicare and Medicaid Services. The Improper Medicare FFS Payments Report - May 2008. CMS, 2008. Available from: http://www.cms.hhs.gov/apps/er_report/preview_er_report_print.asp?from=public&which=long&reportID=9; accessed Aug 12, 2009.
- US Federal Bureau of Investigation. Financial Crimes Report to the Public: Fiscal Year 2007. Available from: http://www.fbi.gov/publications/financial/fcs_report2007/financial_crime_2007.htm#health; accessed Aug 12, 2009.

References

- Eisenberg JM. Physician utilization: The state of research about physicians' practice patterns. Med Care 2002;40(11):1016-35.
- Hager G, Upton C, Graycarek R, Knowles V, McNees E, Perry J. Information Systems Can Help Prevent, but Not Eliminate, Health Care Fraud and Abuse. Kentucky Legislative Research Commission, 2006. Available from: http://www.lrc.ky.gov/lrcpubs/RR%20333_forweb.pdf; accessed Aug 12, 2009.
- Li J, Huang KY, Jin J, Shi J. A survey on statistical methods for health care fraud detection. Health Care Manage Sci 2008;11:275-87.
- Health Insurance Review & Assessment Service and National Health Insurance Corporation. 2007 National Health Insurance Statistical Yearbook. Health Insurance Review & Assessment Service and National Health Insurance Corporation, 2008. Available from: http://www.hira.or.kr/common/dummy.jsp?pgmid=HIRAF010303000000; accessed Aug 12, 2009.
- Phua C, Lee V, Smith K, Gayler R. A comprehensive survey of data mining-based fraud detection research. Artificial Intelligence Review 2005:1-14.
- He H, Hawkins S, Graco W, Yao X. Application of genetic algorithms and k-nearest neighbour method in real world medical fraud detection problem. J Advanced Computational Intelligence and Intelligent Informatics 2000;4(2):130-7.
- Major J, Riedinger D. EFD: A hybrid knowledge/statistical-based system for the detection of fraud. J Risk and Insurance 2002;69(3):309-24.
- Bonchi F, Giannotti F, Mainetto G, Pedreschi D. A classification-based methodology for planning auditing strategies in fraud detection. In: Proceedings of the 5th ACM SIGKDD international conference on knowledge discovery and data mining; San Diego, CA. New York: ACM Press; 1999. 175-84.
- Cox E. A fuzzy system for detecting anomalous behaviors in healthcare provider claims. In: Goonatilake S and Treleaven P (eds). Intelligent Systems for Finance and Business. New York John Wiley and Sons; 1995. 114-34.
- Brockett P, Xia X, Derrig RA. Using Kohonen's self-organizing feature map to uncover automobile bodily injury claims fraud. J Risk and Insurance 1998;65(2):245-74.

References

- Viaene S, Richard A, Dedene D. A case study of applying boosting Naive Bayes to claim fraud diagnosis. IEEE Transactions on Knowledge and Data Engineering 2005;16(5):612-20.
- Kob HC, Tan G. Data mining applications in healthcare. J Healthcare Inform Manage 2005;19(2):64-72.
- Williams G, Huang Z. Mining the knowledge mine: The Hot Spots methodology for mining large real world databases. Lecture Notes Computation Science 1997;1342:340-48.
- Long J, Irani E, Slagle J. Automating the discovery of causal relationships in a medical records database. In: Piatestsky-Shapiro G, Frawley W (eds). Knowledge discovery in database. Menlo Park, CA: AAAI Press; 1991.
- Bartlett P, Ben-David S, Kulkarni S. Learning changing concepts by exploiting the structure of change. Machine Learning 2000;41:153-74.
- Yamanishi K, Takeuchi J, Williams G, Milne P. On-line unsupervised outlier detection using finite mixtures with discounting learning algorithms. Data Mining and Knowledge Discovery 2004;8:275-300.
- He H, Graco W, Yao X. Application of genetic algorithms and k-nearest neighbour method in medical fraud detection. In: Simulated Evolution and Learning. Berlin: Springer; 1999. 74-81.
- He H, Wang J, Graco W, Hawkins S. Application of neural networks to detection of medical fraud. Expert Systems with Applications 1997;13(4):329-36.
- Williams G. Evolutionary hot spots data mining: An architecture for exploring for interesting discoveries. In: Proceedings of PAKDD-99; Beijing, China. Berlin: Springer; 1999. 184-95.
- Yang WS, Hwang SY. A process-mining framework for the detection of healthcare fraud and abuse. Expert Systems with Applications 2006;31:56-68.
- Chan C, Lan C. A data mining technique combining fuzzy sets theory and bayesian classifier-an application of auditing the health insurance fee. In: Proceedings of the International Conference on Artificial Intelligence; Las Vegas, NV; 2001. 402-8.
- Ortega PA, Figueroa CJ, Ruz GA. A medical claim fraud/abuse detection system based on data mining: A case study in Chile. In: Proceedings of the 2006 International Conference on Data Mining; Las Vegas, NV. CSREA Press; 2006. 224-31.