International Journal on Applications in Information and Communication Engineering
Volume 2: Issue 2: February 2016, pp 43-48. www.aetsjournal.com ISSN (Online) : 2394-6237

Efficient Approach For Resilient Detection Using Communal Method

M. Mohana , S. Aanjan Kumar

Abstract— Identity crime is well known, prevalent, and costly; and credit application fraud is a specific case of identity crime. The existing nondata mining detection system of business rules and scorecards, and known fraud matching have limitations. To address these limitations and combat identity crime in real time, this paper proposes a new multilayered detection system complemented with two additional layers: communal detection (CD) and spike detection (SD). CD finds real social relationships to reduce the suspicion score, and is tamper resistant to synthetic social relationships. It is the whitelist-oriented approach on a fixed set of attributes. SD finds spikes in duplicates to increase the suspicion score, and is probe-resistant for attributes. It is the attribute-oriented approach on a variable-size set of attributes. Together, CD and SD can detect more types of attacks, better account for changing legal behavior, and remove the redundant attributes. Experiments were carried out on CD and SD with several million real credit applications. Results on the data support the hypothesis that successful credit application fraud patterns are sudden and exhibit sharp spikes in duplicates. Although this research is specific to credit application fraud detection, the concept of resilience, together with adaptivity and quality data discussed in the paper, are general to the design, implementation, and evaluation of all detection systems.

Index Terms—Data mining-based fraud detection, security, data stream mining, anomaly detection.

I. INTRODUCTION

IDENTITY crime is defined as broadly as possible in this paper. At one extreme, synthetic identity fraud refers to the use of plausible but fictitious identities. These are effortless to create but more difficult to apply successfully. At the other extreme, real identity theft refers to illegal use of innocent people's complete identity details. These can be harder to obtain (although large volumes of some identity data are widely available) but easier to successfully apply. In reality, identity crime can be committed with a mix of both synthetic and real identity details.

Identity crime has become prominent because there is so much real identity data available on the Web, and confidential data accessible through unsecured mailboxes. It has also become easy for perpetrators to hide their true identities. This can happen in a myriad of insurance, credit, and telecommunications fraud, as well as other more serious crimes. In addition to this, identity crime is prevalent and costly in developed countries that do not have nationally registered identity numbers.

Data breaches which involve lost or stolen consumers' identity information can lead to other frauds such as tax returns, home equity, and payment card fraud. Consumerscan incur thousands of dollars in out-of-pocket expenses. The US law requires offending organizations to notify consumers, so that consumers can mitigate the harm. As a result, these organizations incur economic damage, such as notification costs, fines, and lost business

Credit applications are Internet or paper-based forms with written requests by potential customers for credit cards, mortgage loans, and personal loans. Credit applica-tion fraud is a specific case of identity crime, involving synthetic identity fraud and real identity theft.

As in identity crime, credit application fraud has reached a critical mass of fraudsters who are highly experienced, organized, and sophisticated Their visible patterns can be different to each other and onstantly change. They are persistent, due to the high financial rewards, and the risk and effort involved are minimal. Based on anecdotal observations of experienced credit application investigators, fraudsters can use software automation to manipulate particular values within an application and increase frequency of successful values.

Duplicates (or matches) refer to applications which share common values. There are two types of duplicates: exact (or identical) duplicates have the all same values; near (or approximate) duplicates have some same values (or characters), some similar values with slightly altered spellings, or both. This paper argues that each successful credit application fraud pattern is represented by a sudden and sharp spike in duplicates within a short time, relative to the established baseline level.

Duplicates are hard to avoid from fraudsters' point-of-view because duplicates increase their' success rate. The synthetic identity fraudster has low success rate, and is likely to reuse fictitious identities which have been successful before. The identity thief has limited time because innocent people can discover the fraud early and take action, and will quickly use the same real identities at different places.

It will be shown later in this paper that many fraudsters operate this way with these applications and that their characteristic pattern of behavior can be detected by the methods reported. In short, the new methods are based on white-listing and detecting spikes of similar applications. White-listing uses real social relationships on a fixed set of attributes. This reduces false positives by lowering some suspicion scores. Detecting spikes in duplicates, on a variable

M.Mohana , PG Student Shanmuganathan Engineering College (Email: M.Mohanamail@Gmail.Com)
S.Aanjan Kumar, Research Scholar Velammal College of Engineering and Technology (Email : Itsec1990@Gmail.Com)

set of attributes, increases true positives by adjusting suspicion scores appropriately.

Throughout this paper, data mining is defined as the real-time search for patterns in a principled (or systematic) fashion. These patterns can be highly indicative of early symptoms in identity crime, especially synthetic identity fraud

A. Main Challenges for Detection Systems

Resilience is the ability to degrade gracefully when under most real attacks. The basic question asked by all detection systems is whether they can achieve resilience. To do so, the detection system trades off a small degree of efficiency (degrades processing speed) for a much larger degree of effectiveness (improves security by detecting most real attacks). In fact, any form of security involves tradeoffs .

The detection system needs "defence-in-depth" with multiple, sequential, and independent layers of defence [25] to cover different types of attacks. These layers are needed to reduce false negatives. In other words, any successful attack has to pass every layer of defence without being detected.

The two greatest challenges for the data mining-based layers of defence are adaptivity and use of quality data. These challenges need to be addressed in order to reduce false positives.

Adaptivity accounts for morphing fraud behavior, as the attempt to observe fraud changes its behavior. But what is not obvious, yet equally important, is the need to also account for changing legal (or legitimate) behavior within a changing environment. In the credit application domain, changing legal behavior is exhibited by communal relation-ships (such as rising/falling numbers of siblings) and can be caused by external events (such as introduction of organizational marketing campaigns). This means legal behavior can be hard to distinguish from fraud behavior, but it will be shown later in this paper that they are indeed distinguishable from each other.

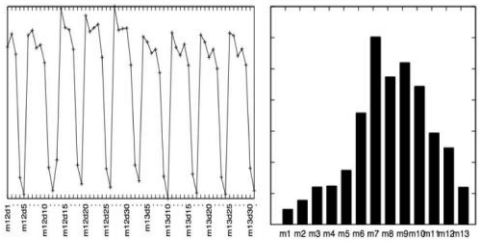
The detection system needs to exercise caution with applications which reflect communal relationships. It also needs to make allowance for certain external events.

Quality data are highly desirable for data mining and data quality can be improved through the real time removal of data errors (or noise). The detection system has to filter duplicates which have been reentered due to human error or for other reasons. It also needs to ignore redundant attributes which have many missing values, and other issues.

B. Existing Identity Crime Detection System

There are nondata mining layers of defence to protect against credit application fraud, each with its unique strengths and weaknesses.

The first existing defence is made up of business rules and scorecards. In Australia, one business rule is the hundred-point physical identity check test which requires the applicant to provide sufficient point-weighted identity documents face-to-face. They must add up to at least 100 points, where a passport is worth 70 points. Another business rule is to contact (or investigate) the applicant over the telephone or Internet. The above two business rules are highly effective, but human resource intensive. To rely less on human resources, a common business rule is to match an application's identity number, address, or phone number against external databases. This is convenient, but the public telephone and address directories, semipublic voters' register, and credit history data can have data quality issues of accuracy, completeness, and timeliness. In addition, scorecards for credit scoring can catch a small percentage of fraud which does not look creditworthy; but it also removes outlier applications which have a higher probability of being fraudulent.



(a) (b)

The second existing defence is known fraud matching. Here, known frauds are complete applications which were confirmed to have the intent to defraud and usually periodically recorded

into a blacklist. Subsequently, the current applications are matched against the blacklist. This has the benefit and clarity of hindsight because patterns often repeat themselves. However, there are two main problems in using known frauds. First, they are untimely due to long time delays, in days or months, for fraud to reveal itself, and be reported and recorded. This provides a window of opportunity for fraudsters. Second, recording of frauds is highly manual. This means known frauds can be incorrect expensive, difficult to obtain and have the potential of breaching privacy.

In the real-time credit application fraud detection domain, this paper argues against the use of classification (or supervised) algorithms which use class labels. In addition to the problems of using known frauds, these algorithms, such as logistic regression, neural networks, or Support Vector Machines (SVM), cannot achieve scalability or handle the extreme imbalanced class [11] in credit application data streams. As fraud and legal behavior changes frequently, the classifiers will deteriorate rapidly and the supervised classification algorithms will need to be trained on the new data. But the training time is too long for real-time credit application fraud detection because the new training data have too many derived numerical attributes (converted from the original, sparse string attributes) and too few known frauds.

This paper acknowledges that in another domain, real-time credit card transactional fraud detection, there are the same issues of scalability, extremely imbalanced classes, and changing behavior. For example, FairIsaac—a company renown for their predictive fraud analytics—has been successfully applying supervised classification algorithms,

including neural networks and SVM.

C. New Data Mining-Based Layers of Defence

The main objective of this research is to achieve resilience by adding two new, real time, data mining-based layers to complement the two existing nondata mining layers discussed in the section. These new layers will improve

detection of fraudulent applications because the detection system can detect more types of attacks, better account for changing legal behavior, and remove the redundant attributes.

These new layers are not human resource intensive. They represent patterns in a score where the higher the score for an application, the higher the suspicion of fraud (or anomaly). In this way, only the highest scores require human intervention. These two new layers, communal and spike detection, do not use external databases, but only the credit application database per se. And crucially, these two layers are unsupervised algorithms which are not completely dependent on known frauds but use them only for evaluation.

The main contribution of this paper is the demonstration of resilience, with adaptivity and quality data in real-time data mining-based detection algorithms. The first new layer is Communal Detection (CD): the whitelist-oriented ap-proach on a fixed set of attributes. To complement and strengthen CD, the second new layer is Spike Detection (SD): the attribute-oriented approach on a variable-size set of attributes.

The second contribution is the significant extension of knowledge in credit application fraud detection because publications in this area are rare. In addition, this research uses the key ideas from other related domains to design the credit application fraud detection algorithms.

Finally, the last contribution is the recommendation of credit application fraud detection as one of the many solutions to identity crime. Being at the first stage of the credit life cycle, credit application fraud detection also prevents some credit transactional fraud.

Section 2 gives an overview of related work in credit application fraud detection and other domains. Section 3 presents the justifications and anatomy of the CD algorithm, followed by the SD algorithm. Before the analysis and interpretation of CD and SD results, Section 4 considers the legal and ethical responsibility of handling application data, and describes the data, evaluation measures, and experi-mental design. Section 5 concludes the paper.

II. BACKGROUND

Many individual data mining algorithms have been designed, implemented, and evaluated in fraud detection. Yet until now, to the best of the researchers' knowledge, resilience of data mining algorithms in a complete detection system has not been explicitly addressed.

Much work in credit application fraud detection remains proprietary and exact performance figures unpublished, therefore there is no way to compare the CD and SD algorithms against their leading industry methods and techniques. For example, [14] has ID Score-Risk which gives a combined view of each credit application's characteristics and their similarity to other industry-provided or Web identity's characteristics. In another example, [7] has Detect which provides four categories of policy rules to signal fraud, one of which is checking a new credit application against historical application data for consistency.

Case-based reasoning (CBR) is the only known prior publication in the screening of credit applications [29]. CBR

analyzes the hardest cases which have been misclassified by existing methods and techniques. Retrieval uses thresholded nearest neighbor matching. Diagnosis utilizes multiple selection criteria (probabilistic curve, best match, negative selection, density selection, and default) and resolution strategies (sequential resolution-default, best guess, and combined confidence) to analyze the retrieved cases. CBR has 20 percent higher true positive and true negative rates than common algorithms on credit applications.

The CD and SD algorithms, which monitor the significant increase or decrease in amount of something important (Section 3), are similar in concept to credit transactional fraud detection and bioterrorism detection. Peer group analysis [2] monitors interaccount behavior over time. It compares the cumulative mean weekly amount between a target account and other similar accounts (peer group) at subsequent time points. The suspicion score is a t-statistic which determines the standardized distance from the centroid of the peer group. On credit card accounts, the time window to calculate a peer group is 13 weeks, and the future time window is 4 weeks. Break point analysis [2] monitors intraaccount behavior over time. It detects rapid spending or sharp increases in weekly spending within a single account. Accounts are ranked by the t-test. The fixed-length moving transaction window contains 24 transactions: the first 20 for training and the next four for evaluation on credit card accounts. Bayesian networks [31] uncover simulated anthrax attacks from real emergency department data. Wong [30] surveys algorithms for finding suspicious activity in time for disease outbreaks. Goldenberg et al. [9] use time series analysis to track early symptoms of synthetic anthrax outbreaks from daily sales of retail medication (throat, cough, and nasal) and some grocery items (facial tissues, orange juice, and soup). Control-chart-based statis-tics, exponential weighted moving averages, and general-ized linear models were tested on the same bioterrorism detection data and alert rate [15].

The SD algorithm, which specifies how much the current prediction is influenced by past observations (Section 3.3), is related to Exponentially Weighted Moving Average (EWMA) in statistical process control research [23]. In particular, like EWMA, the SD algorithm performs linear forecasting on the smoothed time series, and their advan-tages include low implementation and computational com-plexity. In addition, the SD algorithm is similar to change point detection in biosurveillance research, which maintains a cumulative sum (CUSUM) of positive deviations from the mean [13]. Like CUSUM, the SD algorithm raises an alert when the score/CUSUM exceeds a threshold, and both detects change

points faster as they are sensitive to small shifts from the mean. Unlike CUSUM, the SD algorithm weighs and chooses string attributes, not numerical ones.

III. THE METHODS

This section is divided into four sections to systematically explain the CD algorithm (first two sections) and the SD algorithm (last two sections). Each section commences with a clearer discussion about its purposes.

A. Communal Detection

This section motivates the need for CD and its adaptive approach.

Suppose there were two credit card applications that provided the same postal address, home phone number, and date of birth, but one stated the applicant's name to be John Smith, and the other stated the applicant's name to be Joan Smith. These applications could be interpreted in three ways:

1. Either it is a fraudster attempting to obtain multiple credit cards using near duplicated data.
2. Possibly there are twins living in the same house who both are applying for a credit card.
3. Or it can be the same person applying twice, and there is a typographical error of one character in the first name.

With the CD layer, any two similar applications could be easily interpreted as (1) because this paper's detection methods use the similarity of the current application to all prior applications (not just known frauds) as the suspicion score. However, for this particular scenario, CD would also recognize these two applications as either (2) or (3) by lowering the suspicion score due to the higher possibility that they are legitimate.

To account for legal behavior and data errors, CD is the whitelist-oriented approach on a fixed set of attributes. The whitelist, a list of communal and self-relationships between applications, is crucial because it reduces the scores of these legal behaviors and false positives. Communal relationships are near duplicates which reflect the social relationships from tight familial bonds to casual acquaintances: family members, housemates, colleagues, neighbors, or friends [17]. The family member relationship can be further broken down into more detailed relationships such as husband-wife, parent-child, brother-sister, male-female cousin (or both male, or both female), as well as uncle-niece (or uncle-nephew, auntie-niece, auntie-nephew). Self-relationships highlight the same applicant as a result of legitimate behavior (for simplicity, self-relationships are regarded as communal relationships).

Broadly speaking, the whitelist is constructed by ranking link-types between applicants by volume. The larger the volume for a link-type, the higher the probability of a communal relationship. On when and how the whitelist is constructed, please refer to Section 3.2, Step 6 of the CD algorithm.

However, there are two problems with the whitelist. First, there can be focused attacks on the whitelist by fraudsters when they submit applications with synthetic communal relationships. Although it is difficult to make definitive statements that fraudsters will attempt this, it is also wrong to assume that this will not happen. The solution proposed in this paper is to make the contents of the whitelist become less predictable. The values of some parameters (different from an application's identity value) are automatically changed such that it also changes the whitelist's link types. In general, tampering is not limited to hardware, but can also refer to manipulating software such as code. For our domain, tamper resistance refers to making

it more difficult for fraudsters to manipulate or circumvent data mining by providing false data.

Second, the volume and ranks of the whitelist's real communal relationships change over time. To make the whitelist exercise caution with (or more adaptive to) changing legal behavior, the whitelist is continually being reconstructed.

B. CD Algorithm Design

This section explains how the CD algorithm works in real time by giving scores when there are exact or similar matches between categorical data; and in terms of its nine inputs, three outputs, and six steps.

This research focuses on one rapid and continuous data stream [19] of applications. For clarity, let G represent the overall stream which contains multiple and consecutive f. . . ; g_{x_2}; g_{x_1}; g_x; $g_{xþ1}$; $g_{xþ2}$; . . .g Minidiscrete streams.

. g_x: current Minidiscrete stream which containsmultiple and consecutive $fu_{x;1}$; $u_{x;2}$;. . .; $u_{x;p}$g micro-discrete streams.

. x: fixed interval of the current month, fortnight, orweek in the year.

. p: variable number of microdiscrete streams in aMinidiscrete stream.

Also, let $u_{x;y}$ represent the current microdiscrete stream which contains multiple and consecutive $fv_{x;y;1}$; $v_{x;y;2}$;. . .;$v_{x;y;q}$g applications. The current application's links arerestricted to previous applications within a moving window, and this window can be larger than the number of applications within the current microdiscrete stream.

. y: fixed interval of the current day, hour, minute,or second.

. q: variable number of applications in a microdis-crete stream.

Here, it is necessary to describe a single and contin-uous stream of applications as being made up of separate chunks: a Minidiscrete stream is long-term (for example, a month of applications); while a microdiscrete stream is short-term (for example, a day of applications). They help to specify precisely when and how the detection system will automatically change its configurations. For example, the CD algorithm reconstructs its whitelist at the end of the month and resets its parameter values at the end of the day; the SD algorithm does attribute selection and updates CD attribute weights at the end of the month. Also, for example, long-term previous average score, long-term previous average links, and average density of each attribute are calculated from data in a Minidiscrete stream; short-term current average score and short-term current

average links are calculated from data in a microdiscrete stream.

With this data stream perspective in mind, the CD algorithm matches the current application against a moving window of previous applications. It accounts for attribute weights which reflect the degree of importance in attributes. The CD algorithm matches all links against the whitelist to find communal relationships and reduce their link score. It then calculates the current application's score using every link score and previous application score. At the end of the current microdiscrete data stream, the CD algorithm determines the SoA and updates one random parameter's value such that it trades off effectiveness with efficiency, or vice versa. At the end of the current Minidiscrete data stream, it constructs the new whitelist.

Table 1 shows the data input, six most influential parameters, and two adaptive parameters.

TABLE 1
Overview of Communal Detection Algorithm

Inputs
v_i (current application)
W number of v_j (moving window)
$\Re_{x,link-type}$ (link-types in current whitelist)
$T_{similarity}$ (string similarity threshold)
$T_{attribute}$ (attribute threshold)
η (exact duplicate filter)
α (exponential smoothing factor)
T_{input} (input size threshold)
SoA (State-of-Alert)

Outputs
$S(v_i)$ (suspicion score)
Same or new parameter value
New whitelist

CD algorithm

Step 1: Multi-attribute link [match v_i against W number of v_j to determine if a single attribute exceeds $T_{similarity}$; and create multi-attribute links if near duplicates' similarity exceeds $T_{attribute}$ or an exact duplicates' time difference exceeds η]

Step 2: Single-link score [calculate single-link score by matching Step 1's multi-attribute links against $\Re_{x,link-type}$]

Step 3: Single-link average previous score [calculate average previous scores from Step 1's linked previous applications]

Step 4: Multiple-links score [calculate $S(v_i)$ based on weighted average (using α) of Step 2's link scores and Step 3's average previous scores]

Step 5: Parameter's value change [determine same or new parameter value through SoA (for example, by comparing input size against T_{input}) at end of $u_{x,y}$]

Step 6: Whitelist change [determine new whitelist at end of g_x]

. v_i: unscored current application. N is its number ofattributes. $a_{i;k}$ is the value of the kth attribute in application v_i.

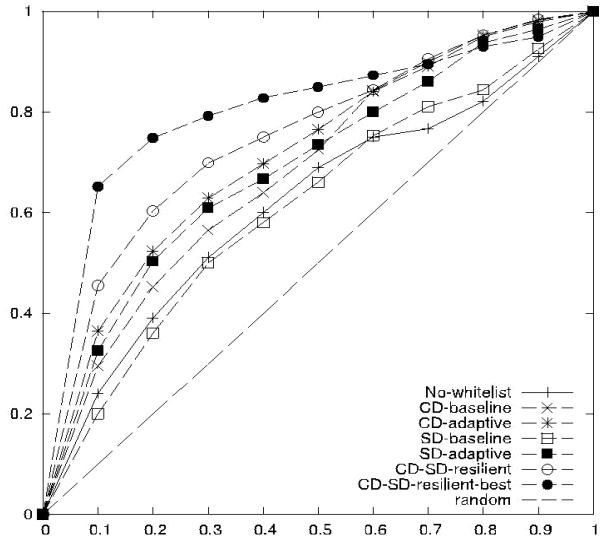
. W: moving (or sliding) window of previous applications. It determines the short time search space for the current application. CD utilizes an application-based window (such as the previous 10,000 applica-tions). v_j is the scored previous application. $a_{j;k}$ is the value of the kth attribute in application v_j.

. $<_{x;link_type}$is a set of unique and sorted link-types (indescending order by number of links), in the link-type attribute of the current whitelist. M is the number of link-types.

. $T_{similarity}$: string similarity threshold between twovalues.

. $T_{attribute}$: attribute threshold which requires a mini-mum number of matched attributes to link two applications.

. _: exact duplicate filter at the link level. It removes links of exact duplicates from the same organization



REFERENCES

[1] A. Bifet and R. Kirkby Massive Online Analysis, Technical Manual, Univ. of Waikato, 2009.

[2] R. Bolton and D. Hand, "Unsupervised Profiling Methods for Fraud Detection," Statistical Science, vol. 17, no. 3, pp. 235-255, 2001.

[3] P. Brockett, R. Derrig, L. Golden, A. Levine, and M. Alpert, "Fraud Classification Using Principal Component Analysis of RIDITs," The J. Risk and Insurance, vol. 69, no. 3, pp. 341-371, 2002, doi: 10.1111/1539-6975.00027.

[4] R. Caruana and A. Niculescu-Mizil, "Data Mining in Metric Space: An Empirical Analysis of Supervised Learning Performance Criteria," Proc. 10th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD '04), 2004, doi: 10.1145/1014052.1014063.

[5] P. Christen and K. Goiser, "Quality and Complexity Measures for Data Linkage and Deduplication," Quality Measures in Data Mining, F. Guillet and H. Hamilton, eds., vol. 43, Springer, 2007, doi: 10.1007/978-3-540-44918-8.

[6] C. Cortes, D. Pregibon, and C. Volinsky, "Computational Methods for Dynamic Graphs," J. Computational and Graphical Statistics, vol. 12, no. 4, pp. 950-970, 2003, doi: 10.1198/1061860032742.

[7] Experian. Experian Detect: Application Fraud Prevention System, Whitepaper, http://www.experian.com/products/pdf/ experian_detect.pdf, 2008.

[8] T. Fawcett, "An Introduction to ROC Analysis," Pattern Recognition Letters, vol. 27, pp. 861-874, 2006, doi: 10.1016/j.patrec. 2005.10.010.

[9] A. Goldenberg, G. Shmueli, R. Caruana, and S. Fienberg, "Early Statistical Detection of Anthrax Outbreaks by Tracking Over-the-Counter Medication Sales," Proc. Nat'l Academy of Sciences USA (PNAS '02), vol. 99, no. 8, pp. 5237-5240, 2002.

[10] G. Gordon, D. Rebovich, K. Choo, and J. Gordon, "Identity Fraud Trends and Patterns: Building a Data-Based Foundation for Proactive Enforcement," Center for Identity Management and Information Protection, Utica College, 2007.

[11] D. Hand, "Classifier Technology and the Illusion of Progress," Statistical Science, vol. 21, no. 1, pp. 1-15, 2006, doi: 10.1214/ 088342306000000060.

[12] B. Head, "Biometrics Gets in the Picture," Information Age, p.p 10-11, Aug.-Sept. 2006.
[13] L. Hutwagner, W. Thompson, G. Seeman, and T. Treadwell, "The Bioterrorism Preparedness and Response Early Aberration Re-porting System (EARS)," J. Urban Health, vol. 80, pp. 89-96, 2006.
[14] IDAnalytics, "ID Score-Risk: Gain Greater Visibility into Indivi-dual Identity Risk," Unpublished, 2008.
[15] M. Jackson, A. Baer, I. Painter, and J. Duchin, "A Simulation Study Comparing Aberration Detection Algorithms for Syndro-mic Surveillance," BMC Medical Informatics and Decision Making, vol. 7, no. 6, 2007, doi: 10.1186/1472-6947-7-6.
[16] J. Jonas, "Non-Obvious Relationship Awareness (NORA)," Proc. Identity Mashup, 2006.
[17] M. Kantarcioglu, W. Jiang, and B. Malin, "A Privacy-Preserving Framework for Integrating Person-Specific Databases," Proc. UNESCO Chair in Data Privacy Int'l Conf. Privacy in Statistical Databases (PSD '08), pp. 298-314, 2008, doi: 10.1007/978-3-540-87471-3_25.
[18] J. Kleinberg, "Temporal Dynamics of On-Line Information Streams," Data Stream Management: Processing High-Speed Data Streams, M. Garofalakis, J. Gehrke, and R. Rastogi, eds., Springer, 2005.
[19] O. Kursun, A. Koufakou, B. Chen, M. Georgiopoulos, K. Reynolds, and R. Eaglin, "A Dictionary-Based Approach to Fast and Accurate Name Matching in Large Law Enforcement Databases," Proc. IEEE Int'l Conf. Intelligence and Security Informatics (ISI '06), 72-82, 2006, doi: 10.1007/11760146.
[20] J. Neville, O. Simsek, D. Jensen, J. Komoroske, K. Palmer, and H. Goldberg, "Using Relational Knowledge Discovery to Prevent Securities Fraud," Proc. 11th ACM SIGKDD Int'l Conf. Knowledge Discovery in Data Mining (KDD '05), 2005, doi: 10.1145/ 1081870.1081922.
[21] T. Oscherwitz, "Synthetic Identity Fraud: Unseen Identity Challenge," Bank Security News, vol. 3, p. 7, 2005.
[22] S. Roberts, "Control-Charts-Tests Based on Geometric Moving Averages," Technometrics, vol. 1, pp. 239-250, 1959.
[23] S. Romanosky, R. Sharp, and A. Acquisti, "Data Breaches and Identity Theft: When Is Mandatory Disclosure Optimal?," Proc. Ninth