

# Crime Detection In Credit Card Fraud

Ms. E.Anuradha

Associate Professor

Dept. Of Information Technology

Swami Vivekananda Institute Of Technology,Sec-Bad

Mr N S S R S Srikar

Btech Final Year Student

Dept. Of Information Technology Line Swami  
Vivekananda Institute Of Technology,Sec-Bad

Mr Aditya Sairam

Btech Final Year Student

Dept. Of Information Technology Line Swami  
Vivekananda Institute Of Technology,Sec-Bad

Mr P. Srikar

Btech Final Year Student

Dept. Of Information Technology Line Swami  
Vivekananda Institute Of Technology,Sec-Bad

## 1. Abstract

*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. Credit card fraud is an element of identity fraud. It can have far reaching effects, since the information on the card can be used to perpetrate other types of identity theft crimes. From using the signature on the back of a card that is stolen, to loaning a credit card to a friend or family member can cause someone to obtain what they need to open other credit card accounts or bank accounts in the victim's name. Credit applications are Internet or paper-based forms with written requests by potential customers for credit cards, mortgage loans, and personal loans. Credit application fraud is a specific case of identity crime, involving synthetic identity fraud and real identity theft. 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.*

## 2. Introduction

### 2.1 Problem Definition

The Existing System use 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 business rules and scorecards, and known fraud matching have limitations. Another existing is known as 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 applications are matched against the blacklist 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, and difficult to obtain, and have the potential of breaching privacy.

### 2.2 Problem Analysis

The Proposed System 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 white list-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.

## 3. System Analysis

### 3.1 EXISTING SYSTEM:

The Existing System use 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 business rules and scorecards, and known fraud matching have limitations. Another existing is known as 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 applications are matched against the blacklist 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, and difficult to obtain, and have the potential of breaching privacy.

### 3.2 PROPOSED SYSTEM:

The Proposed System 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 white list-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.

## 4. System Design

### OOAD OF THE SYSTEM:

An object oriented analysis and design language from the object management group. Many design methodologies for describing object-oriented systems were developed in the late 1980s. UML standardizes several diagramming methods, including Grady Booch's work at Rational Software .Rum Baugh's Object Modeling Technique and Ivan Jacobson's work on use cases.

### UML:

UML stands for Unified Modeling Language. UML is a standardized general-purpose modeling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group. The goal is for UML to become a common language for creating models of object oriented computer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modeling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of

software system, as well as for business modeling and other non-software systems. The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems. The UML is a very important part of developing objects oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

### GOALS:

The Primary goals in the design of the UML are as follows:

1. Provide users a ready-to-use, expressive visual modeling Language so that they can develop and exchange meaningful models.
2. Provide extendibility and specialization mechanisms to extend the core concepts.
3. Be independent of particular programming languages and development process.
4. Provide a formal basis for understanding the modeling language.
5. Encourage the growth of OO tools market.
6. Support higher level development concepts such as collaborations, frameworks, patterns and components.
7. Integrate best practices.

### TYPES

The following are the types of UML diagrams followed in this Project:

- Use Case Diagram.
- Activity Diagram.
- Class Diagram.
- Sequence Diagram.
- Collaboration Diagram.

## DETAILED DESIGN USE CASE DIAGRAM

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted. Use cases in the system are

1. Login.
2. Add\_Item.
3. Detect Fraud.
4. Payment.
5. Search Product.
6. Purchase

Actors in our system

1. Admin.
2. Customer.
3. Bank\_Admin

## ACTIVITY DIAGRAM

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.

## CLASS DIAGRAM

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information.

Class diagram of system consists of

1. Admin.
2. User.
3. Online\_Shopping.
4. Detection\_System.
5. Third\_Party.
6. Customer.

## SEQUENCE DIAGRAM

A sequence diagram in Unified Modeling Language (UML) is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams.

Sequence diagram of customer contain

- 1 Customer.
- 2 Online\_Shopping.
- 3 Detection\_System.

Sequence diagram of admin consists of

- 1 Admin.
- 2 Online\_shopping.
- 3 Bank\_Admin.
- 4 Detection\_System.

## 5. Testing

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub assemblies, assemblies and/or a finished product. It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner.

### 5.1 Test Cases

1. Check whether data is retrieved from the data set or not.
2. Is monthly transactions are viewed or not.
3. Can we select the particular customer or not.
4. Checking of monthly transactions of customers using Bar graph and weights.

5. whether the customer is able to select the item from the online shopping website or not.

## Conclusion

The main focus of this paper is Resilient Identity Crime Detection; in other words, the real-time search for patterns in a multilayered and principled fashion, to safeguard credit applications at the first stage of the credit life cycle. This paper describes an important domain that has many problems relevant to other data mining research. It has documented the development and evaluation in the data mining layers of defence for a real-time credit application fraud detection system. In doing so, this research produced three concepts (or “force multipliers”) which dramatically increase the detection system’s effectiveness (at the expense of some efficiency). These concepts are resilience (multilayer defence), adaptivity (accounts for changing fraud and legal behavior), and quality data (real-time removal of data errors). These concepts are fundamental to the design, implementation, and evaluation of all fraud detection, adversarial-related detection, and identity crime-related detection systems.

The implementation of CD and SD algorithms is practical because these algorithms are designed for actual use to complement the existing detection system. Nevertheless, there are limitations. The first limitation is effectiveness, as scalability issues, extreme imbalanced class, and time constraints dictated the use of rebalanced data in this paper. The counter-argument is that, in practice, the algorithms can search with a significantly larger moving window, number of link types in the whitelist, and number of attributes. The second limitation is in demonstrating the notion of adaptivity. While in the experiments, CD and SD are updated after every period, it is not a true evaluation as the fraudsters do not get a chance to react and change their strategy in response to CD and SD as would occur if they were deployed in real life (experiments were performed on historical data).

## 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](http://www.experian.com/products/pdf/experian_detect.pdf), 2008.