**Resilient Identity Crime Detection**

**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.

**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.

**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.

**ALGORITHM USED:**

**Communal Detection:**

The CD algorithm works in real time by giving scores when there are exact or similar matches between categorical data 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.

**Spike Detection:**

CD has a fundamental weakness in its attribute threshold. Specifically, CD must match at least three values for our data set. With less than three matched values, our whitelist does not contain real social relationships because some values, such as given name and unit number, are not unique identifiers. The fraudster can duplicate one or two

important values which CD cannot detect. Before proceeding with a description of SD, it is necessary to reinforce that CD finds real social relationships to reduce the suspicion score, and is tamper resistant to synthetic social relationships. SD finds spikes to increase the suspicion score, and is probe resistant for attributes. Probe resistance reduces the chances a fraudster will discover attributes used in the SD score calculation. It is the attribute-oriented approach on a variable-size set of attributes

**MODULES:**

The main modules are

1. Admin module

2. User module

3. Identify Crime User

4.Online Shopping

**MODULES DESCRIPTION:**

**1. Admin module:**

In this module, the admin can add product details(product name, price, validity etc..) based on the category likes mobiles, computers, laptops etc.. and maintain the product details. The user enter their credit card details, the credit card is valid by Communal Detection and Spike Detection. If the card details is valid, the user can purchase their items else it report to the admin as “fraud transaction”.

**2. User module:**

The user can select purchasing products displayed in the home page or search the product using keyword or based on category. Then user can purchase the product using credit/debit card. To purchase, the user need to provide the following details like(credit card number, card holder name, date of birth, credit card provider). If the credit card is valid the user is allowed to purchase the product else it reports to the admin as “fraud transaction occur”

**3. Identify Crime User**

In the module we are using a real dataset “ <https://sites.google.com/site/cliftonphua/communal-fraud-scoring-data.zip>”. There are about 30 raw attributes such as personal names, addresses, telephone numbers, driver license numbers (or SSN), DoB, and other identity attributes (but no link attribute). Only 19 of the most important identity attributes are selected in table format as Training Dataset and Test Dataset. In the Dataset we identify the fraud transaction name by user name. The graph display the user fraud percentage across months and measure weight for Test dataset user.

4. **Online shopping**

In the module, we developed a website for online shopping. The user can purchase a products using credit card. If the fraud user uses their credit card to purchase item, the bank identify the fraud user using the weight and graph of the user in the bank database.

# SYSTEM CONFIGURATION:-

# HARDWARE REQUIREMENTS:-

# Processor -Pentium –III

* Speed - 1.1 Ghz
* RAM - 256 MB(min)
* Hard Disk - 20 GB
* Floppy Drive - 1.44 MB
* Key Board - Standard Windows Keyboard
* Mouse - Two or Three Button Mouse
* Monitor - SVGA

# SOFTWARE REQUIREMENTS:-

* Operating System : Windows95/98/2000/XP
* Application Server : Tomcat5.0/6.X
* Front End : Java, JSP
* Script : JavaScript.
* Server side Script : Java Server Pages.
* Database : MYSQL

**REFERENCE:**

Clifton Phua, Member, IEEE, Kate Smith-Miles, Senior Member, IEEE, Vincent Cheng-Siong Lee, and Ross Gayler, “**Resilient Identity Crime Detection**”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 24, NO. 3, MARCH 2012.