RNPS No. 2143
ISSN 2072-6260
Versión Digital

SERIE GRIS

REPORTE TÉCNICO
**Minería de Datos**

# Análisis de los métodos de detección de fraude en servicios de telecomunicaciones

Vitali Herrera-Semenets,
Mario Alfonso Prado-Romero y
Andrés Gago-Alonso

**RT_023** **febrero 2014**

CENATAV
Centro de Aplicaciones de Tecnologías de Avanzada
MINISTERIO DE LA INDUSTRIA BÁSICA

7ma. No. 21812 e/218 y 222,
Rpto. Siboney, Playa;
La Habana.
Cuba. C.P. 12200
**www.cenatav.co.cu**

RNPS No. 2143
ISSN 2072-6260
Versión Digital

SERIE GRIS

REPORTE TÉCNICO
**Minería de Datos**

# Análisis de los métodos de detección de fraude en servicios de telecomunicaciones

Vitali Herrera-Semenets,
Mario Alfonso Prado-Romero y
Andrés Gago-Alonso

**RT_023** **febrero 2014**

CENATAV
Centro de Aplicaciones de Tecnologías de Avanzada
MINISTERIO DE LA INDUSTRIA BÁSICA

7ma. No. 21812 e/218 y 222,
Rpto. Siboney, Playa;
La Habana.
Cuba. C.P. 12200
**www.cenatav.co.cu**

## Siglas y acrónimos

ABP: Asignaciones básicas de probabilidad.
ACP: Análisis de componentes principales.
BDS: Base de datos de llamadas sospechosas.
CDMA: División de código de acceso múltiple (en inglés *Code Division Multiple Access*).
CDR: Registro de detalles de llamadas (en inglés *Call Detail Record*).
CFCA: Asociación de control de fraude en las comunicaciones (en inglés *Communications Fraud Control Association*).
COI: Comunidad de interés (en inglés *Community of Interest*).
GSM: Sistema global para comunicaciones móviles (en inglés *Global System for Mobile Communications*).
ID: Identificador.
IRSF: Fraude internacional de coparticipación (en inglés *International Revenue Share Fraud*).
MDL: Longitud mínima de la descripción (en inglés *Minimum Description Length*).
MMS: Servicio de mensajes multimedia (en inglés *Multimedia Message Service*).
NASD: Asociación nacional de corredores de valores (en inglés *National Association of Securities Dealers*).
PBX: Centrales privadas (en inglés *Private Branch Exchange*).
PGA: Análisis de grupos semejantes(en inglés *Peer Group Analysis*).
PGM: Modelos gráficos probabilísticos (en inglés *Probabilistic Graphical Models*).
RPT: Árboles probabilísticos relacionales (en inglés *Relational Probabilistic Trees*).
PRS: Servicios de valor agregado (en inglés *Premium Rate Service*).
RBF: Funciones de base radial (en inglés *Radial Base Function*).
SIM: Módulo de identificación de abonado (en inglés *Subscriber Identity Module*).
SMS: Servicio de mensajes cortos (en inglés *Short Message Service*).
SOM: Mapas auto-organizados (en inglés *Self-Organizing Maps*).
SVD: Descomposición de valores singulares (en inglés *Singular Value Decomposition*).
TFF: Tabla de frecuencia de llamada de fraude.
TFL: Tabla de frecuencia de llamada legítima.
TIC: Tecnologías de la Informática y las Comunicaciones
USD: Dólares norteamericanos (en inglés *United State Dollars*).
VoIP: Voz sobre protocolo de internet (en inglés *Voice over IP*).
VoIPADE: Motor de detección de anomalías de VoIP (en inglés *VoIP Anomaly Detection Engine*).

## Notaciones

| | |
|---|---|
| $a, a_1, a_2$ | CDRs |
| $A$ | Conjunto de reglas |
| $b_1, b_2$ | Variables continuas |
| $\overline{b_i}, \overline{b_1}, \overline{b_2}$ | Variables cuantitativas |
| $\underline{b}$ | Variable aleatoria |
| $B, B_1$ | Números enteros no negativos |
| $c(x, d_1)$ | Contactos del subscriptor $x$ en el intervalo de tiempo $d_1$ |
| $C(R)$ | Conjunto de clientes detectados como fraudulentos por $R$ |
| $C(A)$ | Conjunto de clientes detectados como fraudulentos por las reglas del conjunto $A$ |
| $d, d_1, d_2, d_3, d_4$ | Intervalos de tiempo |
| $D_1, D_2$ | Eventos |
| $D_{d,y}$ | Evento que representa la ocurrencia de llamadas teniendo en cuenta un intervalo de tiempo $d$ y el tipo de llamada $y$ |
| $E$ | Conjunto de aristas de un grafo |
| $e_{ij}$ | Arista que conecta a $\hat{x_i}$ con $\check{x_j}$ |
| $\overline{f}(\check{a}_\gamma)$ | Frecuencia alta de llamadas de tipo $\check{a}_\gamma$ (toma valores de 0 o 1) |
| $\underline{f}(\check{a}_\gamma)$ | Frecuencia baja de llamadas de tipo $\check{a}_\gamma$ (toma valores de 0 o 1) |
| $\overline{F}$ | Hipótesis que implica que una llamada es fraudulenta |
| $G, G', \overline{G}$ | Grafos |
| $g_i, g_j$ | Subgrafos de algún grafo |
| $h_k$ | $k$-ésima característica de una llamada |
| $\overline{h_k}$ | Medida de la $k$-ésima característica de un vector |
| $H(V, d_1)$ | Contactos del conjunto $V$ en el intervalo de tiempo $d_1$ |
| $H(x, d_1)$ | Contactos del subscriptor $x$ en el intervalo de tiempo $d_1$ |
| $i, j, k$ | Subíndices (números enteros no negativos) |
| $\hat{I}$ | Conjunto de subscriptores de origen |
| $\check{I}$ | Conjunto de subscriptores terminales |
| $J[\psi_1]$ | Media de $\psi_1$ |
| $K_m$ | Matriz de correlación creada en el instante de tiempo $m$ |
| $K_m(i, j)$ | Celda de la matriz de correlación $K_m$ |
| $l$ | Tipo de fraude |
| $L$ | Hipótesis que implica que una llamada es legítima |
| $m$ | Instante de tiempo |
| $M_k$ | Matriz de la $k$-ésima característica |
| $n_i$ | Nivel de profundidad $i$-ésima de una COI |
| $N'$ | Historial de fraude genérico |
| $N(s_j)$ | Perfil histórico del cliente $s_j$ |
| $o(v_1, v_2)$ | Cambio métrico de $v_1$ con respecto a $v_2$ |
| $O_i$ | Salida de la neurona $i$-ésima |
| $p(\psi)$ | Función sigmoide evaluada en la variable real $\psi$ |
| $P(D_{d,y}\|F)$ | Probabilidad condicional de $D_{d,y}$ dada la hipótesis $F$ |
| $P(F)$ | Probabilidad de que se cumpla la hipótesis $F$ |
| $\dot{P}(h_k)$ | Probabilidad de ocurrencias de la $k$-ésima característica en el período de entrenamiento |
| $\ddot{P}(h_k)$ | Probabilidad de ocurrencias de la $k$-ésima característica en el período de prueba |

| | |
|---|---|
| $q(X, Y)$ | Función para determinar si la media de la duración de las llamadas del conjunto $X$ es estadísticamente diferente a la de $Y$ |
| $Q_1, Q_2$ | Conjuntos de hipótesis |
| $\overline{r}(R)$ | Máxima correlación de la regla $R$ |
| $r(R_1, R_2)$ | Correlación entre reglas $R_1$ y $R_2$ |
| $R$ | Regla |
| $R_i$ | Regla $i$-ésima |
| $s, s_j$ | Clientes |
| $S$ | Hipótesis que implica que una llamada es sospechosa |
| $t$ | Total de tipos de fraude |
| $T$ | Conjunto de datos cuantitativos |
| $\overline{T}$ | Media aritmética de los valores de $T$ |
| $u_1(F), u_2(F)$ | Asignaciones básicas de probabilidad para la hipótesis $F$ |
| $U, U_1, U_2$ | Conjunto de vértices de un grafo |
| $v, v_1, v_2, v_i, v_j$ | Vectores |
| $v_1^\tau$ | Vector traspuesto de $v_1$ |
| $v_m$ | Vector de comportamiento de la matriz $K_m$ |
| $v_{m-1}$ | Vector de comportamiento típico de la matriz $K_m$ que se obtiene promediando a partir del instante de tiempo $m-1$, los últimos $W$ vectores de comportamiento |
| $V$ | Conjunto de subscriptores |
| $w_{ij}^*$ | Nuevo peso entre la neurona $i$-ésima y la neurona $j$-ésima |
| $w_{ij}$ | Peso entre la neurona $i$-ésima y la neurona $j$-ésima |
| $W$ | Ventana de tiempo |
| $x, x_1, x_2, x_k$ | Subscriptores |
| $y$ | Tipo de llamada |
| $\hat{x}_i$ | Subscriptor $i$-ésimo del conjunto $\hat{I}$ |
| $\check{x}_j$ | Subscriptor $j$-ésimo del conjunto $\check{I}$ |
| $z_1, z_2, z_k$ | Umbrales de probabilidad |
| $X, Y, Z$ | Conjuntos de CDRs |
| $\alpha(R)$ | Cobertura de fraude adicional de la regla $R$ |
| $\beta_\alpha$ | Umbral mínimo establecido para $\alpha(R)$ |
| $\beta_r$ | Umbral máximo aceptado para $\overline{r}(R)$ |
| $\gamma$ | Tipo de duración de una llamada (0 corta, 1 larga) |
| $\underline{\gamma}$ | Umbral máximo de tiempo de duración de llamadas |
| $\overline{\gamma}$ | Duración máxima de las llamadas de una ventana de tiempo |
| $\overrightarrow{\gamma}$ | Tiempo de duración de una llamada |
| $\delta$ | Parámetro para quitarle o darle influencia histórica a una COI |
| $\epsilon$ | Distancia Hellinger |
| $\epsilon_i$ | Distancia Hellinger en el $i$-ésimo intervalo de tiempo |
| $\epsilon_i'$ | Valor esperado de $\epsilon$ en el $i$-ésimo intervalo |
| $\tilde{\epsilon}_i$ | Desviación media en el $i$-ésimo intervalo |
| $\underline{\epsilon}$ | Umbral para determinar si $\epsilon$ representa a una anomalía |
| $\varepsilon_j$ | Error de la neurona $j$-ésima |
| $\theta_1, \theta_2$ | Umbrales de contactos |
| $\vartheta(R_i)$ | Cantidad de CDRs detectados como fraudulentos por la regla $i$-ésima |
| $\vartheta_l(R_i)$ | Cantidad de CDRs detectados como fraudes de tipo $k$ por la regla $i$-ésima |
| $\lambda$ | Variable que define la magnitud de la penalización en la técnica de *decaimiento de peso* |

| | |
|---|---|
| $\dot{\mu}(y)$ | Promedio de llamadas de tipo $y$ por día |
| $\ddot{\mu}(y)$ | Promedio de llamadas de tipo $y$ por mes |
| $\nu(X, Y)$ | Grado de variación de las llamadas del conjunto $X$ con respecto a las del conjunto $Y$ |
| $\xi_1, \xi_2$ | Coeficientes para hacer que $\epsilon'_i$ con la ayuda de $\tilde{\epsilon}_i$ converja lo más cerca posible a $\epsilon$ |
| $\xi_3$ | Coeficiente que define en el algoritmo que tan sensible debe ser $\underline{\epsilon}$ a los cambios |
| $\xi_4$ | Coeficiente que define en el algoritmo que tan adaptable debe ser $\underline{\epsilon}$ a los cambios de comportamientos en las llamadas |
| $\rho(v_1, v_2)$ | Función de coeficiente de correlación de Pearson de los vectores $v_1$ y $v_2$ |
| $\varrho(X)$ | Media de la duración de las llamadas del conjunto $X$ |
| $\sigma^2(X)$ | Varianza de las llamadas del conjunto $X$ |
| $\sigma(\psi_1, \psi_2)$ | Covarianza entre las variables reales $\psi_1$ y $\psi_2$ |
| $\sigma_F$ | Desviación de la frecuencia |
| $\varsigma(D_{d,y})$ | Número de ocurrencias de $D_{d,y}$ |
| $\varsigma(x, D_{d,a})$ | Número de ocurrencias de $D_{d,y}$ del subscriptor $x$ |
| $\phi$ | Término adicional de la función $w^*_{ij}$ |
| $\varphi$ | Variable que define el ritmo de aprendizaje, mientras menor sea su valor el aprendizaje se hace más lento |
| $\chi(N')$ | Cantidad de llamadas en el historial $N'$ |
| $\chi(x, N(s_j))$ | Cantidad de llamadas en el historial $N(s_j)$ del subscriptor $x$ |
| $\psi, \psi_1, \psi_2$ | Variables reales |
| $\aleph$ | Espacio muestral |
| $\mid A \mid$ | Cardinalidad del conjunto $A$ |
| $\mid \psi \mid$ | Valor absoluto del número real $\psi$ |
| $\cdot$ | Operador de multiplicación |
| $\setminus$ | Operador de diferencia de conjuntos |
| $\cup$ | Operador de unión de conjuntos |
| $\cap$ | Operador de intersección de conjuntos |
| $\in$ | Operador de pertenencia |
| $\oplus$ | Operador de suma entre dos grafos |
| $\emptyset$ | Conjunto vacío |
| $\Omega$ | Frecuencia de un tipo de llamada |
| $\underline{\Omega}$ | Umbral máximo de frecuencia de llamadas |
| $\bar{\Omega}_y$ | Máxima frecuencia de llamadas de tipo $y$ |
| $\check{\Omega}_\gamma$ | Frecuencia de llamada entrante de tipo de duración $\gamma$ |
| $\hat{\Omega}_\gamma$ | Frecuencia de llamada saliente de tipo de duración $\gamma$ |
| $\ddot{\Omega}$ | Por ciento que representa $\Omega$ de las otras llamadas |

## Tabla de contenido



## Lista de figuras

## Lista de tablas

# Análisis de los métodos de detección de fraude en servicios de telecomunicaciones

Vitali Herrera-Semenets, Mario Alfonso Prado-Romero, y Andrés Gago-Alonso

Equipo de Investigaciones de Minería de Datos, Centro de Aplicaciones de Tecnologías de Avanzada (CENATAV), La Habana, Cuba
{vherrera,mprado,agago}@cenatav.co.cu

RT_023, Serie Gris, CENATAV
Aceptado: 21 de Enero de 2014

**Resumen.** El análisis de los métodos de detección de fraude en servicios de telecomunicaciones representa un punto de partida para comprender la forma en la cual se le hace frente a las diversas técnicas para cometer fraude que existen en la actualidad. En este trabajo son analizados métodos para la detección de fraude en servicios de telecomunicaciones, así como las técnicas para cometer fraudes. A partir del análisis desarrollado se proponen dos taxonomías, una para los métodos de detección y la otra para las técnicas de fraude. Para comprender de forma mas detallada el proceso de detección de fraude de los distintos enfoques reportados, para cada uno de ellos se elaboró un esquema de funcionamiento. También son analizadas otras técnicas para la detección de fraudes, pero aplicadas a otras esferas como en tarjetas de crédito, subastas online, mercado de valores, entre otras, que pudieran ser modificadas para ser aplicadas en servicios de telecomunicaciones. Finalmente son presentados los principales problemas y dificultades existentes en el area de la detección de fraudes en servicios de telecomunicaciones, los cuales brindan oportunidades para la investigación.
**Palabras clave:** técnicas de detección de fraude, técnicas para cometer fraude, servicios de telecomunicaciones.

**Abstract.** The review of fraud detection techniques in telecommunication services represents a starting point for understanding how to deal with the existing fraud techniques. In this paper, the reported methods for fraud detection in telecommunication services are analyzed, as well as fraud techniques. As a result of research two taxonomies are proposed: one for detection methods and other for fraud techniques. Each fraud detection technique was explained in details, including a scheme describing the detection process. Other fraud detection methods applied to other areas such as credit cards, online auctions, stock market, among others, are also presented. Such methods could be extended to be applied in telecommunication services. Finally, the main problems and difficulties detected in the area of fraud detection in telecommunication services are reported, providing new opportunities for future research.
**Keywords:** fraud detection techniques, fraud techniques, telecommunication services.

## 1. Introducción

El constante desarrollo de las Tecnologías de la Informática y las Comunicaciones (TIC) ha traído consigo la evolución de técnicas para cometer fraude en servicios de telecomunicaciones. Los fraudes representan pérdidas millonarias para las compañías que son afectadas directa e

indirectamente por esta actividad. Podemos encontrar ejemplos en las aplicaciones no confiables de teléfonos inteligentes (en inglés *smartphone*) que solicitan el envío de un mensaje de texto o realizar una llamada de voz a cierto número para completar su activación. Cuando el cliente ejecuta una de estas acciones se convierte en víctima ya que el número telefónico que contacta es usado por estafadores para cometer fraude. En muchos casos las aplicaciones no confiables realizan llamadas sin el consentimiento del cliente, las cuales son generalmente a números internacionales con altas tasas de pago.

El informe publicado por de la asociación de control de fraude en las comunicaciones (CFCA [1] por sus siglas en inglés) evaluó en 2011 las pérdidas globales por concepto de fraude en 40.1 billones de dólares norteamericanos (USD por sus siglas en inglés). Los tres tipos de fraude que mas pérdidas reportaron fueron:

1. Centrales privadas (PBX por sus siglas en inglés) comprometidas (4.96 billones de USD).
2. Robo de identidad y suscripción (4.32 billones de USD).
3. Fraude internacional de coparticipación (IRSF por sus siglas en inglés) (3.84 billones de USD).

Como resultado de muchas llamadas realizadas en períodos breves de tiempo, un gran volumen de información es generado. Por lo cual resulta prácticamente imposible para un analista poder procesar toda esa cantidad de datos sin hacer uso de métodos automatizados. Los analistas son las personas encargadas de detectar posibles fraudes en las compañías de telecomunicaciones. Los métodos automatizados deben detectar cuando se está en presencia de un fraude o cuando podría estar por cometerse alguno.

En este trabajo es presentado un estado del arte en el campo de la detección de fraudes en servicios de telecomunicaciones. El objetivo es analizar las soluciones que han sido propuestas y determinar los problemas existentes. Para tener una mejor idea de como tienen lugar los fraudes, se hace un análisis de las técnicas para cometer fraude. A partir de las técnicas reportadas son propuestas dos taxonomías: una para los métodos de detección y otra para las técnicas de comisión de fraude. El estudio de métodos de detección de fraude asociados a otras areas fuera de los servicios de telecomunicaciones también forma parte de este trabajo. A través de dicho estudio se puede determinar en cuales propuestas pueden tener lugar modificaciones para que puedan ser aplicadas en servicios de telecomunicaciones.

Este trabajo se divide en cuatro secciones como se describe a continuación. La primera sección consiste en la introducción al tema tratado. En la segunda sección se abordan algunos de los fraudes más comunes hoy en día. En la tercera sección se analizan los métodos de detección de fraude. En la cuarta sección, se abordan técnicas de detección de fraude en otras esferas. Finalmente en la quinta sección se presentan las conclusiones generales de la investigación.

## 2. Fraudes en telecomunicaciones: Definición y ejemplos

En la presente sección se abordan aspectos básicos sobre los fraudes, como su definición (sección 2.1) y algunos ejemplos de los más usados en la actualidad como el de bypass (sección 2.2), robo de identidad y suscripción (sección 2.3), IRSF (sección 2.4), las estafas (en inglés *Scam*) (sección 2.5) y el fraude superpuesto (sección 2.6).

En el trabajo de Laleh y Azgomi [1] se propone una taxonomía de distintos tipos de fraude no solo en telecomunicaciones, sino también en otras esferas como en tarjetas de crédito, en la

[1] CFCA: Communications Fraud Control Association, 2011 Global Fraud Loss Survey, www.cfca.org

web, seguros, entre otros. En su propuesta para el sector de las telecomunicaciones, solo aparecen representados dos tipos de fraude: subscripción y superpuesto. En nuestra revisión del estado del arte proponemos una taxonomía (figura 1) más ampliada para los tipos de fraude reportados en el sector de las telecomunicaciones.

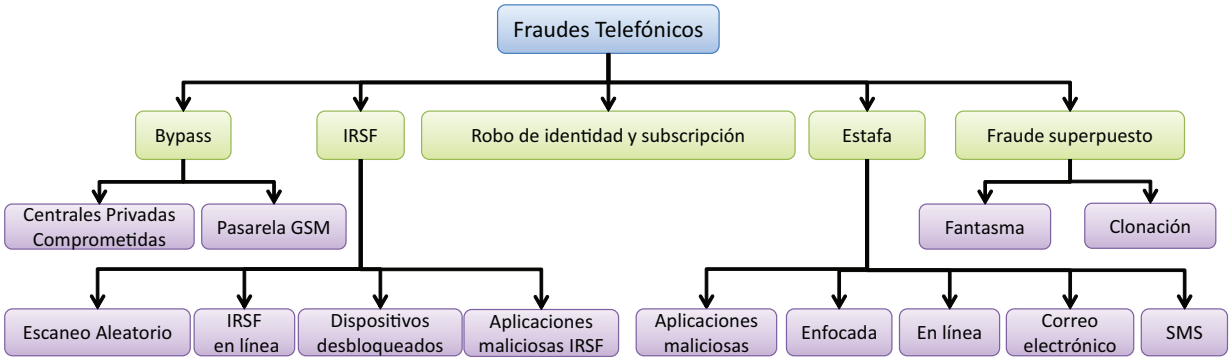


**Fig. 1.** Taxonomía de los tipos de fraudes reportados.

En la taxonomía presentada se observa como se descomponen los distintos tipos de fraude. Teniendo en cuenta las pérdidas reportadas por la CFCA (sección 1), cabe resaltar, como la técnica de PBX comprometidas que representa una variante de bypass, reportó más pérdidas que todas las variantes de IRSF juntas. Más adelante en este trabajo se explicará como funciona cada uno de los tipos de fraudes presentados en la taxonomía.

### 2.1. ¿Qué es un fraude en servicios de telecomunicaciones?

Un fraude consiste en utilizar los servicios de telecomunicaciones incumpliendo con las condiciones establecidas en los contratos y causando pérdidas para la compañía. Ejemplo de esto lo podemos encontrar en la reventa de comunicaciones al extranjero, el reenvío de llamadas a números de servicio de valor agregado (PRS por sus siglas en inglés) o a números internacionales con altas tasas de pago. Los fraudes pueden ser generados ya sea por personas involucradas conocidas como defraudadores o de manera automatizada. A continuación serán abordadas varias técnicas para cometer fraude.

### 2.2. Bypass

Dentro de los fraudes de tipo bypass[2] son incluidas algunas técnicas de terminación de llamadas de menor costo, como las PBX comprometidas. El bypass se caracteriza por evitar la interconexión legal de las llamadas, y desviar las llamadas internacionales de la red de sistema global para comunicaciones móviles (GSM por sus siglas en inglés) o la red de división de código de acceso múltiple (CDMA por sus siglas en inglés). En tal sentido, las conexiones de voz sobre el protocolo de internet (VoIP por sus siglas en inglés) son usadas como puerta de enlace, eludiendo así los impuestos de llamadas internacionales establecidos por las compañías.

Como se reporta en el informe de la CFCA, las PBX comprometidas representan el tipo de fraude que más pérdidas reportó a las compañías telefónicas. Generalmente es cometido contra

[2] Bypass: www.subex.com/pdf/bypass-fraud.pdf

los propietarios de las PBX. Los servicios privados proporcionan muchas funcionalidades, como correos de voz, asistentes automatizados, entre otros. Es por esa razón que los defraudadores utilizan a menudo dichas funcionalidades con el fin de obtener una señal para marcar. Luego, el defraudador coloca una llamada, y el costo es para el propietario de la PBX [2].

El fraude de tipo pasarelas GSM tiene lugar cuando es comprada una gran cantidad de tarjetas de módulo de identificación de abonado (SIM por sus siglas en inglés). El individuo u organización que poseen dichas tarjetas SIM, pasan a ofrecer llamadas gratuitas o muy económicas a celulares. Las tarjetas compradas se usan para realizar llamadas nacionales o internacionales evitando el conmutador internacional del operador y los cargos por terminación de llamada. De esta manera se entrega la llamada al destinatario como si fuese una llamada local, con la consecuente pérdida de ingresos para el proveedor.

### 2.3. Robo de identidad y suscripción

Teniendo en cuenta el informe de la CFCA, es el tipo de fraude que ocupó el segundo puesto entre los que más pérdidas reportan en el 2011. La técnica de robo de identidad y suscripción tiene lugar cuando alguien se suscribe a un servicio utilizando una identidad falsa para fraudulentamente obtener un contrato de servicio. El robo de identidad es un delito en el que una persona se apropia de la información privada de otra para cometer fraudes. Una forma de efectuar el robo de identidad es mediante la aplicación de técnicas de ingeniería social.

El fraude es realizado sin el conocimiento del suscriptor cuya identidad fue usada y puede repercutir, en un futuro, si la víctima quiere contratar algún servicio a la empresa afectada. Generalmente el fraude de suscripción queda oculto en las pérdidas de malos pagadores, es decir, los clientes que se retrasan en el pago. Por esa causa los analistas no lo identifican como fraude y la empresa nunca se recupera de las pérdidas.

### 2.4. Fraude internacional de coparticipación

El fraude internacional de coparticipación según el informe de la CFCA escaló hasta el tercer puesto en los fraudes que más pérdidas reportó en el 2011. En estos casos, el proveedor designa un conjunto de números de PRS, que a menudo tienen un precio mucho más alto que las llamadas legales que se realicen con destino a números ubicados en el mismo país que los PRS. En algunos países los números de PRS pueden tener tasas similares a las llamadas regulares. Al actuar como proveedores de contenido y atraer a las víctimas a llamar a números de PRS, los atacantes obtienen ingresos directos de las llamadas.

Los estafadores que están detrás del IRSF idean diversos métodos para solicitar a los clientes hacer tantas llamadas como sea posible a los números de fraude para maximizar sus ganancias. A pesar que las actividades de IRSF son comunes para los teléfonos fijos, se ha descubierto una serie de actividades IRSF que son únicos para las redes celulares [3], en los que se incluyen:

- **Aplicaciónes maliciosas (en inglés *Malware*)**: Es el caso de IRSF que más predomina representando el 13,2 % del total de las actividades de fraude. Estas aplicaciones móviles están diseñadas para ser capaces de iniciar llamadas involuntarias de los clientes a los números telefónicos utilizados por los estafadores para cometer fraude IRSF (en el resto del trabajo se les hará referencia como "*números IRSF*"). Los malware se pueden encontrar camuflajeados como una aplicación de juegos o como aplicaciones de útilidad. Una de las cosas más intere-

santes es que hay algunos malware que son capaces de conectarse con un servidor remoto y descargar listas de números IRSF periódicamente, por lo cual en muchas ocasiones se hace muy difícil encontrarlos y deshabilitarlos.

- **Escaneo aleatorio (en inglés *Random Scanning*)**: En este escenario, los estafadores utilizan un sistema automatizado para hacer llamadas a una serie de números de teléfono adyacentes sin dejar ningún mensaje. Esto hace que los clientes devuelvan la llamada a ese número.
- **IRSF en línea (en inglés *Online Media IRSF*)**: En este caso los estafadores ponen números IRSF en sitios muy populares en internet como Facebook, Twitter, entre otros. Un ejemplo es que publiquen un número de teléfono comentando que es de un artista famoso lo cual genera en un breve tiempo un número muy elevado de llamadas.
- **Dispositivos desbloqueados**: Los dispositivos desbloqueados son los teléfonos celulares que no están vinculados a una compañía específica. Este fraude comienza cuando un cliente lleva un celular a desbloquear. Tras ser desbloqueado, el estafador reconfigura el dispositivo y especifica números IRSF como puntos de acceso para el servicio de mensajes cortos (SMS por sus siglas en inglés) o para el servicio de mensajes multimedia (MMS por sus siglas en inglés). Esta acción hace que cada vez que el cliente envíe un SMS o un MMS se disparen llamadas a los números IRSF.

### 2.5. Estafa

En los escenarios de estafa existen dos etapas. En la primera etapa, los estafadores solicitan llamadas de los clientes; luego, en la segunda cuando los clientes devuelven la llamada, los estafadores aplican varias técnicas de ingeniería social. Estas técnicas son utilizadas para manipular a las personas con el fin de obtener información privada o confidencial. Existen varios tipos de estafas [3] dentro de los cuales se incluyen:

- **Enfocada (en inglés *Target Scam*)**: Este es el caso más común que representa el 55,4 % de todas las atividades de fraude. En este escenario el estafador selecciona las víctimas de las cuales ya tiene algún tipo de información personal como nombre, números de teléfono, direcciones, etc.
- **En línea (en inglés *Online Media Scam*)**: En este escenario, los estafadores ponen anuncios en foros online, los cuales son a menudo acerca de alquileres de casas de bajo costo y la venta de bienes raíces. Al final de los anuncios en muchas ocasiones se listan varios números de teléfonos para atrapar a las víctimas que son atraídas por los precios bajos.
- **Correo electrónico (en inglés *Email*)**: En este caso los estafadores piratean las cuentas de correo personal y envían correos a la lista de contactos haciéndose pasar por el propietario de la cuenta, reclamando una emergencia en un país extranjero y solicitando llamadas a las personas de la lista de contactos.
- **Aplicaciónes maliciosas (en inglés *Malware*)**: Este tipo de estafa se manifiesta en los ordenadores personales y atacan al sistema operativo Windows. Una vez infectado el ordenador, este queda bloqueado impidiéndole al usuario ejecutar un comando, incluso no permite entrar por modo seguro para eliminar el malware. Solo aparece un recuadro solicitando al usuario llamar a una lista de números internacionales para obtener el código para desbloquear el ordenador, y muchas veces los estafadores obligan a los usuarios a pagar para la restauración del sistema.

- **SMS**: Se caracterizan por ser SMS provenientes de empresas fantasmas que notifican a los clientes que son ganadores de increíbles premios y les solicitan a cambio un depósito para pagar los gastos administrativos.

### 2.6. Fraude superpuesto

Es la apropiación de un servicio sin contar con la autorización necesaria para hacer su uso. Esta situación puede ser fácilmente detectada mediante la identificación de llamadas no autorizadas que aparecen en la cuenta de teléfono. Este tipo de fraude se manifiesta de muchas formas, por ejemplo utilizando la tecnología Fantasma (en inglés *Ghosting*), que engaña a la red con el fin de obtener llamadas gratis. La Clonación es otra variante de fraude superpuesto, mediante la cual los defraudadores ganan acceso a todos los recursos de un móvil legítimo mediante la duplicación del mismo y el tráfico usado será facturado al propietario legal [4].

### 2.7. Conclusiones parciales

La existencia de una variedad tan amplia de técnicas para cometer fraude y la evolución continua de las tecnologías de la informática y las comunicaciones, propician el desarrollo de métodos y estrategias que afectan a las compañías de telecomunicaciones. Teniendo en cuenta lo anteriormente expuesto resulta prácticamente imposible tener una solución automatizada que detecte todos los tipos de fraude. El análisis de esta sección permite comprender como se realizan los distintos tipos de fraudes y el daño que pueden generar tanto a los clientes como a las compañías de telecomunicaciones. En la siguiente sección, se analizaran los métodos reportados en la literatura que permiten detectar algunos de estos fraudes.

## 3. Métodos de detección de fraude en servicios de telecomunicaciones

Debido a que es muy grande el volumen de información a ser procesado por los analistas, se hace necesario el empleo de técnicas automatizadas que faciliten su trabajo. Los sistemas de detección de fraude pueden ser de dos tipos: de análisis absoluto o de análisis diferencial. Las técnicas reportadas en este trabajo se incluyen dentro del conjunto de análisis diferencial, con excepción de las técnicas basadas en reglas que son de análisis absoluto.

Las técnicas de análisis absoluto determinan los comportamientos fraudulentos basados en reglas, cotas o umbrales, que son definidas teniendo en cuenta el comportamiento previamente conocido de cada fraude. El hecho que las técnicas de análisis absoluto sean basadas en reglas, las hace eficientes sobre los fraudes que ya se tiene conocimiento de su existencia. Aunque dichas técnicas fallan en la detección de cambios más sutiles o los nuevos tipos de fraude que aparecen y no se tienen definidas reglas o umbrales para ellos.

Las técnicas de análisis diferencial permiten detectar cambios del comportamiento de los clientes en el uso de los servicios o llamadas. Esta técnica funciona almacenando por cada cliente un perfil a largo plazo (historial del cliente) y uno a corto plazo (perfil actual del cliente). Una vez calculado los perfiles, algunas comparaciones son realizadas con el objetivo de detectar cambios en el comportamiento del cliente.

Por lo general, el conjunto de datos sobre el cual se realiza el análisis de los métodos son los registros de detalles de llamadas (CDR por sus siglas en inglés) [5], que incluye suficiente información como la fecha, tiempo de duración, número de origen y destino, entre otros. Los CDR son generados automáticamente en la compañía telefónica cada vez que se efectúa una llamada y pueden ser almacenados durante largos períodos de tiempo, llegando a billones de registros.

En Becker et al. [6] se describen algunos componentes claves a tener en cuenta en un sistema de detección de fraude. Esos componentes son: una fuente continua de CDR, una base de datos, algoritmos de detección, analistas, y herramientas de visualización para ayudar a los analistas a hacer diagnósticos. En la figura 2 se presenta la taxonomía de las técnicas de detección de fraude reportadas en este trabajo.

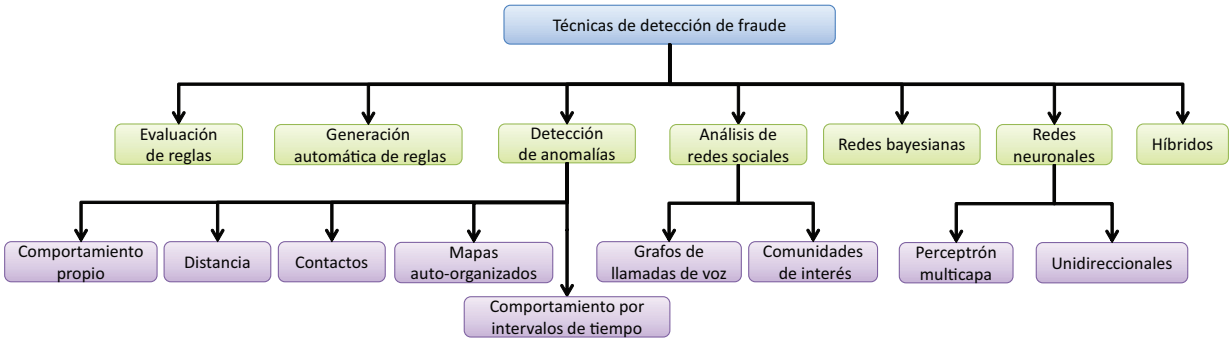


**Fig. 2.** Taxonomía de las técnicas de detección de fraude analizadas en servicios de telecomunicaciones.

En la taxonomía se representa como están organizadas las técnicas de detección de fraude. Dentro de dichas técnicas, las basadas en anomalías son las que más variantes poseen. Cada uno de los métodos que conforman la taxonomía así como trabajos relacionados a estos serán expuestos a continuación.

### 3.1. Evaluación de Reglas

Existen técnicas propuestas basadas en reglas para detectar fraudes en redes de VoIP [7], así como en redes de telefonía móvil [8], donde el conjunto de reglas a utilizar se define con anterioridad por un grupo de analistas. En dichas técnicas intervienen tres elementos importantes. El primero es un listado de datos de las cuentas telefónicas. Otro consiste en un repositorio donde están las reglas que definen el fraude. Por último el motor de reglas el cual determina si es descrito en alguna regla y lanza alertas en tales casos (ver figura 3).

Las técnicas basadas en reglas [9,10,8] son muy eficientes, pero muy difíciles de administrar, se requiere de mucho trabajo para especificar reglas para cada caso de fraude posible. Además estos métodos necesitan ser actualizados con frecuencia para poder detectar los nuevos fraudes que aparezcan. Existen también sistemas como el Nikira[3] que son basados en reglas para detectar fraudes.

Para crear las reglas además de tener en cuenta la información generada por los CDRs también se puede utilizar la información de las cuentas de los clientes. Entre los datos de las cuentas de

[3] Nikira fraud management system: http://pinpoint.microsoft.com/en-in/applications/nikira-fraudmanagement-system-12884906185

los clientes se puede encontrar la edad del cliente, modelo de teléfono, plan de tarifa, crédito y estado civil.

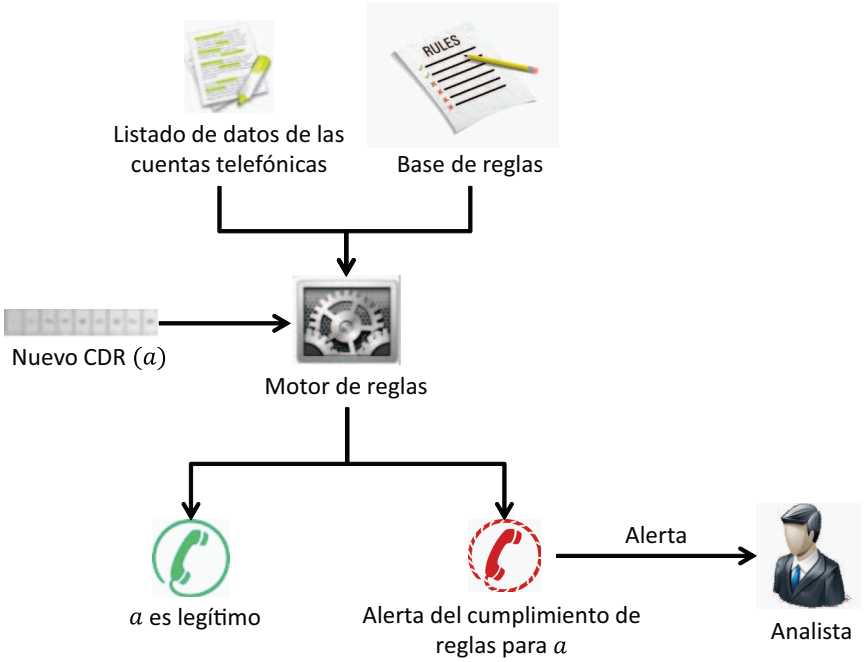


**Fig. 3.** Arquitectura del método basado en evaluación de reglas.

### 3.2. Generación automática de reglas

Las técnicas basadas en generación automática de reglas parten de datos etiquetados como fraudulentos y no fraudulentos. La base de datos está formada por un listado de CDRs históricos y un listado que incluye atributos de las cuentas de los clientes. Los listados son procesados por un algoritmo que permite encontrar relaciones entre los atributos de ambos listados obteniéndose como resultado reglas. En la propuesta de Rosset et al. [11], los criterios empleados para seleccionar las reglas son los siquientes:

- Las reglas deben ser diferentes en el sentido de que deben identificar distintos tipos de fraude.
- El conjunto de reglas debe ser lo suficientemente pequeño para que sea procesado por un analista.
- Alta precisión (la mayoría de los casos detectados deben ser realmente fraudulentos).
- Alta cobertura (la mayoría de los casos de fraude deben ser detectados).

A partir de los criterios anteriores, es implementado un algoritmo para la selección de reglas, que es explicado más adelante.

En el trabajo de Rosset et al. [11], se propone un modelo de dos etapas para detectar el fraude de subscripción y el fraude superpuesto. Una etapa es la fase de generación de reglas y la otra es la fase de selección del conjunto de reglas que va a ser utilizado. El esquema de este método se muestra en la figura 4.

La base de datos de este modelo está compuesta por dos componentes. El primero es un listado de CDRs históricos etiquetados en fraudulentos y no fraudulentos. De dicho listado se obtiene un resumen de las llamadas de cada cliente, teniendo en cuenta atributos como el total de llamadas, llamadas locales, llamadas a móviles y llamadas internacionales. El segundo es un listado de datos de las cuentas de los clientes, que tiene atributos como el tipo de cliente, área de residencia, crédito, entre otros.


**Fig. 4.** Esquema de generación automática de reglas.

Para la generación de reglas se procesan los datos usando una modificación del algoritmo C4.5 [12] que permite encontrar relaciones entre los dos listados de datos. Las relaciones obtenidas constituyen reglas. Un ejemplo de una relación que define un fraude es la siguiente: un cliente que vive en determinada área de residencia y hace muchas llamadas internacionales, es fraudulento. Este mismo procedimiento para la generación de reglas se presenta en el trabajo de Rajani y Padmavathamma [13].

Para seleccionar una nueva regla se utiliza un algoritmo voraz (en inglés *Greedy*) que se divide en dos subprocesos. El primero ordena las reglas candidatas teniendo en cuenta su cobertura y precisión. El segundo es quien realiza la selección, escaneando secuencialmente todas las reglas ordenadas comenzando por la mejor regla de acuerdo a su cobertura y precisión. La reglas son seleccionadas teniendo en cuenta los criterios definidos en términos de umbrales, donde la diferencia entre dos reglas viene dada por su correlación.

El proceso de selección de una nueva regla $R$ comienza calculando su cobertura de fraude adicional $\alpha(R)$ (ver ecuación 1) y su máxima correlación $\overline{r}(R)$ (ver ecuación 3) con las reglas en el conjunto $A$ de reglas seleccionadas. La variable $C(R)$ representa el conjunto de clientes detectados como fraudulentos por la regla $R$ y $C(A)$ denota el conjunto de clientes detectados como fraudulentos por las reglas del conjunto $A$. Los conjuntos $C(R)$ y $C(A)$ son definidos como cobertura de fraude de la regla $R$ y el conjunto $A$ respectivamente. Para seleccionar la regla $R$ el analista define dos umbrales. El umbral $\beta_\alpha$ es el valor mínimo aceptado para $\alpha(R)$ y el umbral $\beta_r$ que es el valor máximo aceptado para $\overline{r}(R)$. La función $r(R_1, R_2)$ (ver ecuación 2) obtiene la correlación entre dos reglas $R_1$ y $R_2$ del conjunto $A$, se define:

$$\alpha(R) = |C(R) \backslash C(A)|, \tag{1}$$

$$r(R_1, R_2) = \frac{\sum_{l=1}^{t}(\vartheta_l(R_1) - \vartheta(R_1)) \cdot (\vartheta_l(R_2) - \vartheta(R_2))}{\sqrt{\sum_{l=1}^{t}(\vartheta_l(R_1) - \vartheta(R_1))^2} \cdot \sqrt{\sum_{l=1}^{t}(\vartheta_l(R_2) - \vartheta(R_2))^2}}. \tag{2}$$

Donde el valor de $t$ representa la cantidad de tipos de fraude, el número entero $\vartheta(R)$ representa el total de CDRs detectados como fraudulentos por la regla $R$, en la base de datos, y el número entero $\vartheta_l(R)$ representa la cantidad de CDRs detectados como fraudes de tipo $l$ por $R$. En el caso de esta propuesta el fraude de subscripción es representado por $l = 1$ y el fraude superpuesto por $l = 2$.

$$\overline{r}(R) = \max_{R_i \in A}\{r(R, R_i)\}, \tag{3}$$

si $\alpha(R) \geq \beta_\alpha$ y $\overline{r}(R) \leq \beta_r$, la regla $R$ es seleccionada.

Este modelo proporciona un mayor entendimiento a los analistas, puesto que la salida del sistema son reglas, a diferencia de otros que funcionan como cajas negras. En la etapa de generación se obtienen muchas reglas, ya que se trata de no excluir ninguna. Por tanto, en la etapa de selección se hace necesario filtrar las reglas generadas, eliminando aquellas que no cumplan los criterios establecidos.

### 3.3. Detección de anomalías

Los trabajos analizados para la detección de fraudes basados en anomalías parten de listados de CDRs históricos no etiquetados. De los CDRs se extraen las características con las cuales se define el comportamiento de los subscriptores. Luego, empleando distintas técnicas, se buscan anomalías en los comportamientos de los subscriptores y en caso de detectarse alguno se reporta.

#### *3.3.1. Basado en comportamiento propio*

En este tipo de enfoque se parte de obtener perfiles de los subscriptores a partir de características recopiladas en un período de tiempo determinado, por ejemplo: las llamadas recibidas, llamadas realizadas, la media de estas llamadas, el total de subscriptores con los que interactuó. Teniendo en cuenta que las variables $i$, $j$ y $k$ son números enteros no negativos. Para cada subscriptor $x_k$, existe una matriz $M_k$ que en la fila $i$-ésima correspondiente al $i$-ésimo instante de tiempo (el tiempo es representado por días), y la columna $j$-ésima correspondiente a la $j$-ésima característica contiene un celda $M_k(i, j)$ con el valor de la característica (ver figura 5).

La dimensión de la matriz $M_k$ depende del período de tiempo de los datos guardados y la cantidad de características. Para cada subscriptor $x_k$, el analista define una misma ventana de tiempo $W$. De la matriz $M_k$ para cada $j$-ésima característica se extrae su vector $v_j$ correspondiente de dimensión $W$. Luego para cada par de vectores se calcula el coeficiente de correlación de Pearson definido como:

$$\rho(v_1, v_2) = \frac{\sigma(v_1, v_2)}{\sigma(v_1)\sigma(v_2)}, \tag{4}$$

donde $v_1$ y $v_2$ son los vectores de dimensión $W$, $\sigma(v_1, v_2)$ es la covarianza de dichos vectores. Los valores de $\sigma(v_1)$ y $\sigma(v_2)$ representan la desviación típica de $v_1$ y $v_2$ respectivamente. A partir de los coeficientes calculados es creada una matriz de correlación $K_m$ de los vectores de características como se explica a continuación, donde $m$ es el instante de tiempo en el que fue creada la matriz. En la matriz $K_m$ la fila $i$-ésima corresponde a la $i$-ésima característica y la columna $j$-ésima corresponde a la $j$-ésima característica. Posteriormente se desplaza la ventana de tiempo y se obtiene otra matriz de correlación $K_{m+1}$. Para cada ventana de tiempo construyen una matriz de

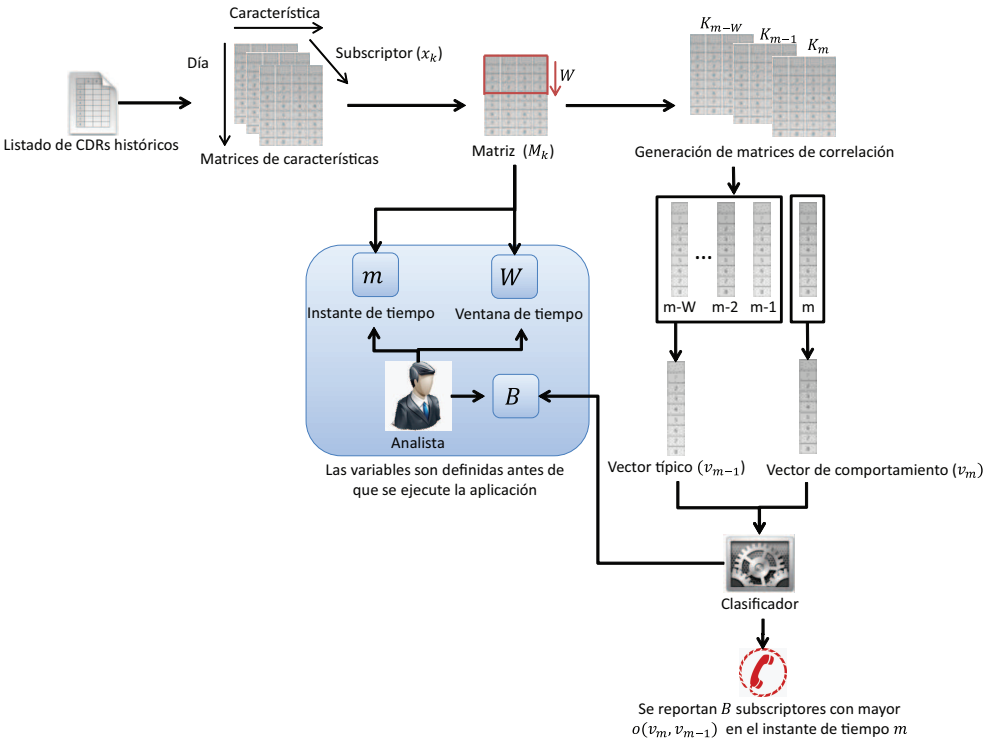


**Fig. 5.** Esquema de detección de anomalías basado en comportamiento propio.

correlación $K_m$, donde $K_m(i,j) = \rho(v_i, v_j)$ sobre la ventana de tiempo $W$. El proceso se repite hasta que se llega al final de los datos.

De cada $K_m$ generada, son extraídos los vectores propios que son llamados vectores de comportamiento. Para el vector de comportamiento calculado en el instante de tiempo $m$, definido como $v_m$, se calcula un vector típico $v_{m-1}$ que promedia a partir del instante de tiempo $m-1$, los últimos $W$ vectores de comportamiento obtenidos (ver figura 5).

El vector típico $v_{m-1}$ es comparado con el último vector de comportamiento obtenido $v_m$. Para ello se define el cambio métrico entre dos vectores $v_1$ y $v_2$ como $o(v_1, v_2) = 1 - v_1^\tau \cdot v_2$, donde $v_1^\tau$ representa el vector traspuesto de $v_1$. Aplicando el cambio métrico a los vectores de comportamiento y típico $o(v_m, v_{m-1})$, se calcula el producto escalar de los dos vectores. Cuando ambos vectores son idénticos el producto escalar es 1 por tanto $o(v_m, v_{m-1}) = 0$. El analista define una cantidad $B$ de subscriptores para ser reportados por la aplicación. Se reportan los $B$ subscriptores con el valor más elevado de $o(v_m, v_{m-1})$ en un instante de tiempo $m$ predefinido, como subscriptores con comportamiento anómalo en $m$.

En Akoglu y Faloutsos [14], es representada la interacción entre los subscriptores mediante un grafo donde los nodos representan subscriptores y las aristas indican que entre los subscriptores se ha compartido un SMS. Cada subscriptor cuenta con 12 características extraídas de él, de sus vecinos y de todas las interacciones entre ellos. Entre las características extraídas están: grado de aristas entrantes, grado de aristas salientes, peso entrante, peso saliente, cantidad de vecinos, cantidad de triángulos, promedio de peso entrante, promedio de peso saliente, máximo peso entrante, máximo peso saliente, entre otras.

Teniendo que los datos guardados son de un período de tiempo de 183 días se forma una matriz de tiempo por característica para cada subscriptor. Para cada matriz de características

se realiza el mismo proceso que se describe a continuación con el fin de detectar un cambio de comportamiento en algún subscriptor en un instante de tiempo $m$. Después del analista haber definido la ventana de tiempo $W$ en 7 días, se calculan los coeficientes de correlación de Pearson entre los vectores de características de la matriz $M_k$. Con los coeficientes se generan las matrices de correlación para cada instante de tiempo.

A partir de la matriz de correlación $K_m$ se obtiene el vector de comportamiento $v_m$ y el vector típico $v_{m-1}$, el cual se calcula promediando los vectores de comportamiento detectados en los $W$ días anteriores. Luego se calcula el cambio métrico entre los dos vectores $o(v_m, v_{m-1})$ y se reportan los cinco subscriptores con el valor más elevado.

### *3.3.2. Basado en comportamiento por intervalos de tiempo*

Una propuesta de detección de anomalías basado en comportamiento por intervalos de tiempo es la del motor de detección de anomalías de VoIP (VoIPADE por sus siglas en inglés) [15], la cual detecta anomalías en las llamadas de salida. En dicha propuesta las llamadas reportadas como anómalas son aquellas que tienen alta frecuencia de llamadas a destinatarios con altas tarifas de pago, o poca cantidad de llamadas, pero de larga duración a dichos destinatarios. En la figura 6 se muestra el esquema de esta técnica.

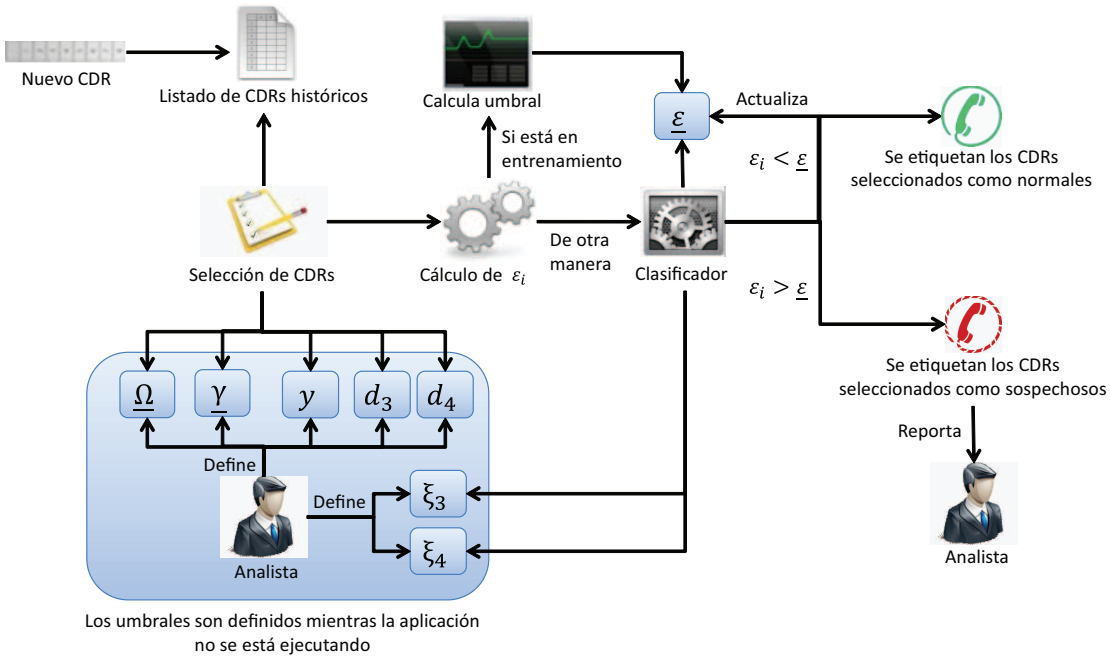


**Fig. 6.** Esquema de detección de anomalías basado en comportamiento por intervalos de tiempo.

Al generarse un nuevo CDR es almacenado en un listado de CDRs histórico, de donde se extraen los que serán procesados. La aplicación comienza por una etapa de entrenamiento, para la cual el analista define un intervalo de tiempo $d_4$. El analista además define umbrales para la frecuencia $\underline{\Omega}$, duración $\underline{\gamma}$ y selecciona el tipo de llamada $y$ que desea analizar. La aplicación busca los CDRs que los valores de sus características sobrepasan los umbrales establecidos por el analista. Si el tipo de llamada del CDR no se incluye en los tipos de llamada seleccionados por el analista el CDR no es seleccionado. En el caso de la frecuencia y la duración, si una de las dos sobrepasa el valor predefinido se selecciona el CDR, de otra manera no es seleccionado. Los tipos

de llamada pueden ser: emergencia, internacionales, nacionales, móviles, servicios y de PRS. Las características utilizadas de un tipo de llamada son las frecuencia y la duración.

Para detectar comportamientos anómalos la aplicación usa el algoritmo de distancia Hellinger [16], como se explica a continuación. En el período de entrenamiento se calcula la probabilidad $\dot{P}(h_k)$ de ocurrencias de una característica $h_k$ que cumpla con el valor establecido por el analista. El valor de $\dot{P}(h_k)$ es el resultado de dividir el número de CDRs que tienen una característica que cumple con el criterio establecido por el analista entre el total de CDRs en el período de entrenamiento.

Para la etapa de detección, el analista define un intervalo de tiempo $d_3$ en minutos que indica el lapso de tiempo después del cual la aplicación busca los CDRs generados en dicho intervalo. En cada intervalo se calcula la probabilidad $\ddot{P}(h_k)$ de ocurrencias de una característica $h_k$ de la misma forma que se calcula $\dot{P}(h_k)$, solo que esta vez para el intervalo de tiempo $d_3$. Después de tener las probabilidades el valor de la distancia de Hellinger $\epsilon$ se calcula como se muestra en la ecuación 5.

$$\epsilon = \left(\sqrt{\ddot{P}(h_k)} - \sqrt{\dot{P}(h_k)}\right)^2. \tag{5}$$

Luego de haber obtenido el valor de $\epsilon$ se procede a calcular un umbral dinámico. El valor del umbral $\underline{\epsilon}$ se deriva a partir del valor de $\epsilon$ y es calculado utilizando el algoritmo de Jacobson [17] (ver ecuaciones 6, 7, 8).

$$\epsilon'_i = \epsilon'_{i-1} + \xi_1 \cdot (\epsilon_i - \epsilon'_{i-1}), \tag{6}$$

$$\tilde{\epsilon}_i = \tilde{\epsilon}_{i-1} + \xi_2 \cdot ((\epsilon_i - \epsilon'_{i-1}) - \tilde{\epsilon}_{i-1}), \tag{7}$$

$$\underline{\epsilon} = \xi_3 \cdot \epsilon'_i + \xi_4 \cdot \tilde{\epsilon}_i. \tag{8}$$

La variable $\epsilon'_i$ es el valor esperado de $\epsilon$ en el $i$-ésimo intervalo, $\epsilon_i$ representa el valor de $\epsilon$ en el $i$-ésimo intervalo y $\tilde{\epsilon}_i$ es la desviación media en el $i$-ésimo intervalo. Las variables $\xi_1$ y $\xi_2$ son coeficientes para hacer que $\epsilon'_i$ con la ayuda de $\tilde{\epsilon}_i$ converja lo más cerca posible a $\epsilon$. El valor de la variable $\xi_3$ define en el algoritmo que tan sensible debe ser $\underline{\epsilon}$ a los cambios. La variable $\xi_4$ define en el algoritmo que tan adaptable debe ser $\underline{\epsilon}$ a los cambios de comportamientos en las llamadas.

El analista define el valor para las variables $\xi_3$ y $\xi_4$. Sin embargo los valores de $\xi_1$ y $\xi_2$ son por defecto $\frac{1}{2^3}$ y $\frac{1}{2^4}$ respectivamente, aunque pueden ser modificados por el analista. Durante la etapa de detección el umbral solo se calcula para los intervalos de tiempo donde $\epsilon_i < \underline{\epsilon}$. El valor de $\underline{\epsilon}$ a medida que se va actualizando se salva en una base de datos. La salva posibilita que si la aplicación se interrumpe o reinicia, puede ser cargado el umbral y se evita realizar nuevamente un entrenamiento.

Cuando se cumple que $\epsilon_i > \underline{\epsilon}$ se está en presencia de una anomalía en el $i$-ésimo intervalo de tiempo. Todos los CDRs seleccionados en ese intervalo para calcular $\epsilon_i$ son enviados a una base de datos de alerta, para ser analizados posteriormente por el analista.

#### *3.3.3. Basado en distancia*

La propuesta que utiliza este enfoque procesa un listado histórico de CDRs, extrayendo características como son la duración y frecuencia de las llamadas de los subscriptores. Dichas características se usan para representar la interacción entre dos subscriptores en una ventana de tiempo

$W$ definida por el analista, mediante un vector teniendo en cuenta el tipo de llamada. La figura 7 representa el esquema del enfoque en análisis.

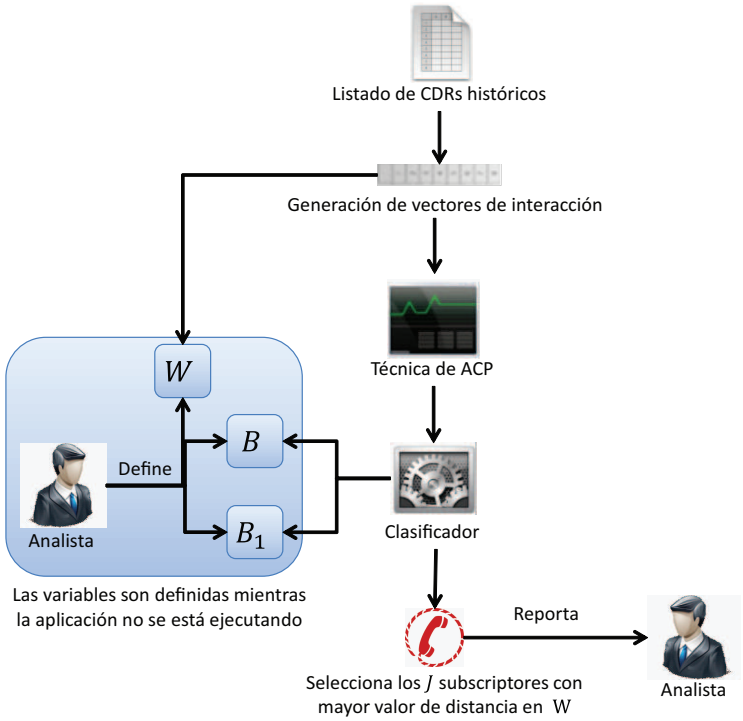


**Fig. 7.** Esquema de detección de anomalías basado en distancia.

En Nithi y Dey [18] se propone una solución en la cual se combinan los tipos de llamadas con sus características para formar un vector (ver figura 8). El tipo de llamada se clasifica como llamada de entrada o de salida, donde $\gamma$ representa el tipo de duración. El tipo de duración es definido como larga o corta, a partir de un umbral establecido por el analista.

En este caso el analista define como llamadas largas ($\gamma = 1$) aquellas que su tiempo de duración $\overrightarrow{\gamma}$ es superior al 95 % de las otras llamadas, y llamadas cortas ($\gamma = 0$) aquellas que $\overrightarrow{\gamma}$ es inferior al 75 % de las otras llamadas. Por lo cual se define el tipo de duración como:

$\gamma = 0$ si $\overrightarrow{\gamma} < 75\,\%$ de las otras llamadas
$\gamma = 1$ si $\overrightarrow{\gamma} > 95\,\%$ de las otras llamadas

Para cada tipo de llamada teniendo en cuenta a $\gamma$, se busca su frecuencia. La frecuencia para cada tipo de llamada viene dada por $\check{\Omega}_\gamma$ que representa la frecuencia de las llamadas de entrada de tipo de duración $\gamma$, y $\hat{\Omega}_\gamma$ que representa la frecuencia de las llamadas de salida de tipo de duración $\gamma$ (ver figura 8). La frecuencia indica la cantidad de llamadas de un mismo tipo entre dos subscriptores. Dependiendo de una frecuencia $\Omega$ de un tipo de llamada, se puede definir el tipo de interacción entre dos subscriptores como alto $\overline{f}(\Omega)$ o bajo $\underline{f}(\Omega)$. Teniendo en cuenta que $\ddot{\Omega}$ es el por ciento que representa $\Omega$ de las otras llamadas, los tipos de interacción se definen mediante las siguientes funciones difusas:

$\underline{f}(\Omega) = 1, \overline{f}(\Omega) = 0$ si $\Omega < 95\,\%$ de la frecuencia de llamadas,

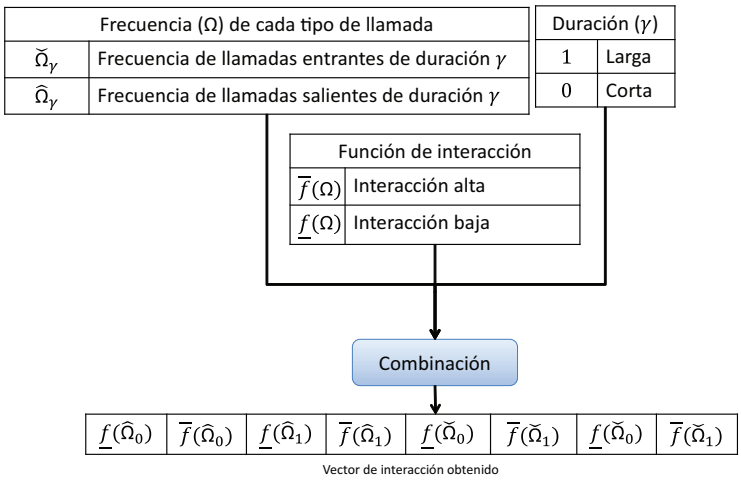


**Fig. 8.** Ejemplo de creación de un vector.

$\underline{f}(\Omega) = 0, \overline{f}(\Omega) = 1$ si $\Omega > 99\,\%$ de la frecuencia de llamadas,

de otra manera véase ecuación 9.

$$\underline{f}(\Omega) = \frac{0,99 - \ddot{\Omega}}{0,99 - 0,95}, \overline{f}(\Omega) = \frac{\ddot{\Omega} - 0,95}{0,99 - 0,95}. \tag{9}$$

Combinando las características con el tipo de llamada queda representada por un vector la interacción entre dos subscriptores en la ventana de tiempo $W$ (ver figura 8). A los vectores de interacción generados les aplican una técnica de análisis de componentes principales (ACP) para identificar los componentes con mayor variación. Los tres componentes con mayor variación se escogen para representar el conjunto de datos transformado.

Al conjunto de datos transformado se le aplican un algoritmo basado en la distancia de Ramaswamy [19] para detectar las anomalías. Dicho algoritmo parte de las variables $B$ y $B_1$ que representan números enteros no negativos definidos por el analista. Luego para cada subscriptor busca la distancia de los $B$ vecinos más cercanos y va actualizando una lista que contiene los $B_1$ subscriptores con los mayores valores de distancia. Al terminar de procesar todos los subscriptores se retorna la lista, la cual es chequeada por el analista.

### *3.3.4. Basado en contactos*

El enfoque basado en contactos se centra en el seguimiento y análisis de los contactos que cada subscriptor tiene. El método comienza cuando se genera un nuevo CDR por un subscriptor $x$, el cual se comprueba si es nuevo o no a partir de un listado de CDRs históricos. En caso de $x$ ser nuevo se pasa a la fase de entrenamiento donde se guardan en una lista los contactos con los cuales interactuó (envió o recibió SMS) en ese tiempo. Después de terminada la fase de entrenamiento se pasa a la fase de prueba, donde los contactos con que interactúa cada subscriptor son controlados en una lista. Luego la cantidad de nuevos contactos es calculada a partir de las listas obtenidas. Teniendo en cuenta la cantidad de nuevos contactos y un umbral predefinido, el algoritmo marca el comportamiento como anómalo o normal (figura 9).

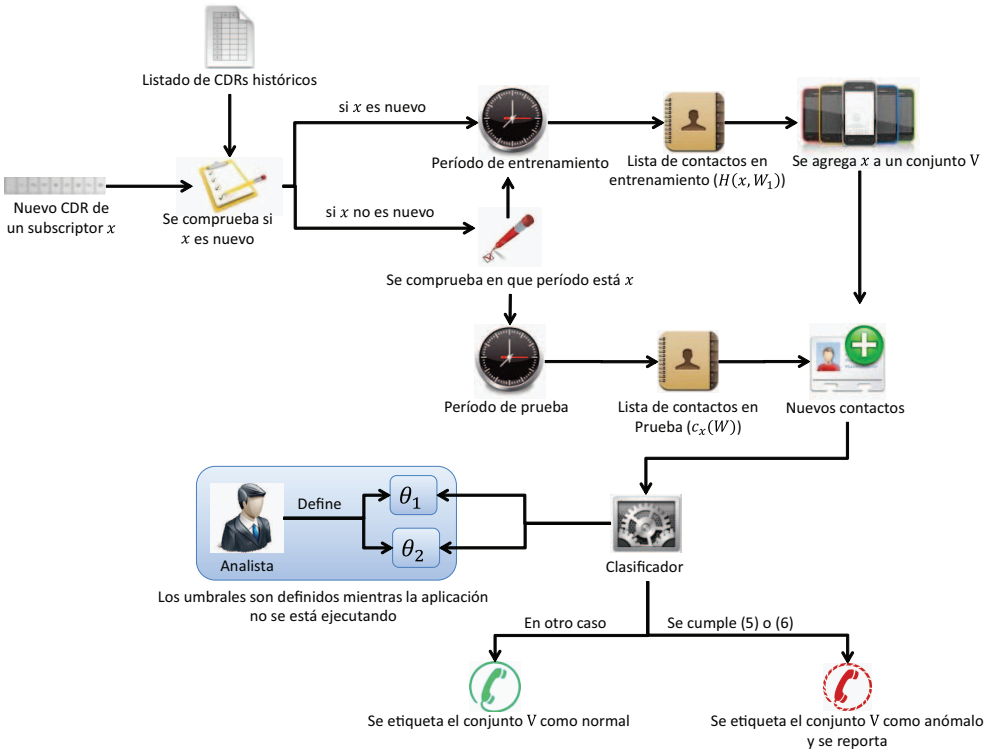


**Fig. 9.** Esquema de detección de anomalías basado en contactos.

En la propuesta de Murynets y Jover [20] cuando un subscriptor nuevo $x$ en la red envía o recibe el primer SMS, comienza el período de entrenamiento. El intervalo de tiempo $d_1$ que marca la duración del entrenamiento puede variar de un día a una semana en dependencia de la carga de datos que genere $x$. Durante este período se guarda una lista $H(x, d_1)$ con los subscriptores con los cuales interactuó $x$. Los subscriptores son agrupados teniendo en cuenta los contactos de cada cual. El listado de contactos agregados al conjunto $V$ durante el período de entrenamiento es definido como $H(V, d_1)$. Después del entrenamiento, $x$ es asignado a uno de los grupos existentes basándose en su historial de comunicación $H(x, d_1)$. Si los contactos de $H(x, d_1)$ no están contenidos en ninguno de los listados $H(V, d_1)$ de los grupos existentes se crea un grupo nuevo.

En la etapa de prueba cuando llega un nuevo CDR de un subscriptor $x$ de un conjunto $V$ los contactos actuales $c(x, d_2)$ son controlados durante el intervalo de tiempo $d_2$ que representa la duración de la etapa de prueba. Un nuevo contacto es un SMS enviado o recibido de un subscriptor en la etapa de prueba que no fue observado durante el entrenamiento. El umbral $\theta_1$ representa el mínimo de nuevos contactos de los subscriptores de un conjunto y $\theta_2$ es el umbral que establece el mínimo de nuevos contactos en un conjunto. Ambos umbrales son predefinidos por el analista. Teniendo en cuenta esto, si se cumple una de las siguientes condiciones:

$$\sum_{x \in V} |c(x, d_2) \setminus H(x, d_1)| > \theta_1, \tag{10}$$

$$\sum_{x \in V} |c(x, d_2) \setminus H(V, d_2)| > \theta_2, \tag{11}$$

se está en presencia de una anomalía en el conjunto $V$ y se reporta.

### *3.3.5. Mapas auto-organizados*

La técnica analizada basada en mapas auto-organizados utiliza como repositorio un listado de CDRs históricos tanto para la etapa de entrenamiento como para la de prueba. Los mapas auto-organizados son redes neuronales no supervisadas, donde a cada neurona se le asocia un vector de peso. Los CDRs son usados para generar vectores de características como se explica a continuación. En la etapa de entrenamiento los vectores son procesados por un algoritmo que actualiza los vectores de peso de las neuronas de tal forma que se acercan al vector de entrada. Una vez entrenada, la red es usada para clasificar patrones de entrada similares. Un grupo de patrones similares tiende a controlar una neurona específica, la cual resultará la más activada frente a los patrones más parecidos a su vector de pesos. Por lo que la clasificación consiste en introducir un vector de entrada y en dependencia de la unidad más activada, la red determina si es fraudulento o no (figura 10).

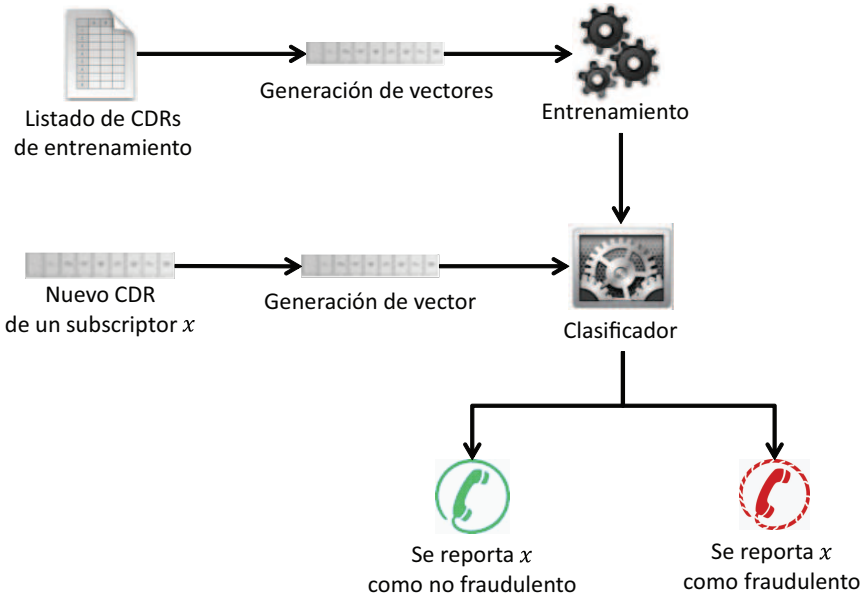


**Fig. 10.** Esquema de las técnicas basadas en mapas auto-organizados.

En la propuesta de Abidogun [21] para cada cliente se genera un vector de características de cinco elementos. Las características que contiene el vector son: el destino de llamada, el ID del celular utilizado, el código del área de ubicación, fecha y hora de la llamada y duración de la llamada. Las características destino de llamada, ID del celular utilizado y código del área de ubicación tenían datos simbólicos y fueron llevados a numéricos. El campo de fecha y hora fue transformado, se definió como horario pico desde las 7 AM hasta las 8 PM y horario bajo desde las 8 PM hasta las 7 AM.

El entrenamiento se realiza aplicando el algoritmo de Mapa de Kohonen y su proceso se describe a continuación. Primero se define una retícula rectangular donde cada vector de peso que inicialmente contiene valores aleatorios puede tener ocho vecinos. Para cambiar los valores de los vectores de peso de la vecindad se utiliza una función de tipo gaussiana, la cual hará que el cambio de valores disminuya con la distancia. Los vectores del conjunto de entrenamiento son procesados por el algoritmo. Luego el vector de peso más cercano de entre los del mapa es calculado, y se le denomina ganador. Después el valor de ese vector de peso y de otros vectores

de peso en su vecindad cambia de forma que se acerquen más al vector de entrada. El tamaño de la vecindad disminuirá a lo largo del entrenamiento.

Después del entrenamiento queda preparada la red neuronal que es de tipo mapa auto-organizado (SOM por sus siglas en inglés) conformada por dos capas, una de entrada y una de salida. La red permite agrupar los vectores de entrada similares. La distancia entre los grupos implica diferencia, pero la cercanía no implica semejanza. La distancia es determinada calculando la diferencia cuadrática entre los vectores de peso que son vecinos en el mapa entrenado. El valor obtenido indica que tan semejantes o diferentes son los grupos. Generalmente los grupos con grandes valores de distancias representan comportamiento anómalo, por lo cual los vectores que son clasificados como de ese grupo son reportados como anómalos. El resultado posteriormente es comprobado por un analista de fraudes.

### 3.4. Análisis de redes sociales

El análisis de redes sociales es usado por muchas técnicas de detección de fraude, incluyendo muchas de las presentadas en este trabajo. Representando una red social como un grafo, un nodo puede representar un número telefónico y las aristas atributos que definan la relación de comunicación entre dos nodos. Por ejemplo en la figura 11, se muestra un fragmento de una red de telecomunicaciones representado como un grafo ponderado dirigido. En el ejemplo cada nodo representa un número telefónico y las aristas el sentido en que se realizaron las llamadas, además su peso es el total de llamadas realizadas. En dependencia de lo que representan las aristas en el grafo, este pudiera ser no dirigido. Por ejemplo si representaran el total de llamadas realizadas entre dos nodos, no tendría sentido la dirección de las aristas, por tanto sería no dirigido.

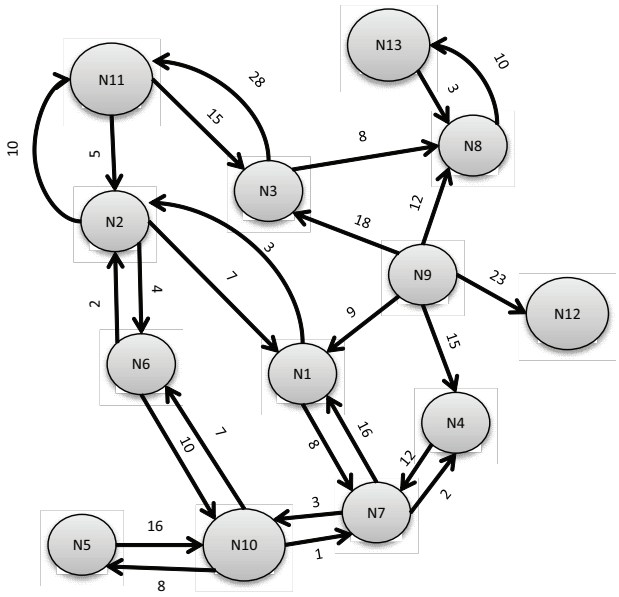


**Fig. 11.** Representación en grafo de un fragmento de red de telecomunicaciones.

### *3.4.1. Grafo de llamadas de voz*

Un grafo de llamadas de voz es un grafo bipartito definido como $G = (\hat{I} \cup \check{I}, E)$, donde $\hat{I}$ representa el conjunto de los subscriptores de origen, $\check{I}$ el conjunto de los subscriptores terminales y $E$ el conjunto de aristas que los relaciona, en una ventana de tiempo $W$ definida por el analista. Una arista $e_{ij}$ conecta a $\hat{x}_i \in B$ y $\check{x}_j \in \check{I}$ si al menos una llamada de voz se realiza de $\hat{x}_i$ a $\check{x}_j$ en $W$. Dentro del grafo de llamadas de voz los componentes desconectados son definidos como subgrafos y son extraídos para ser analizados. Los subgrafos más extensos son descompuestos para luego ser clasificados y determinar si existen en ellos posibles números fraudulentos. En la figura 12 se muestra el esquema del enfoque analizado.

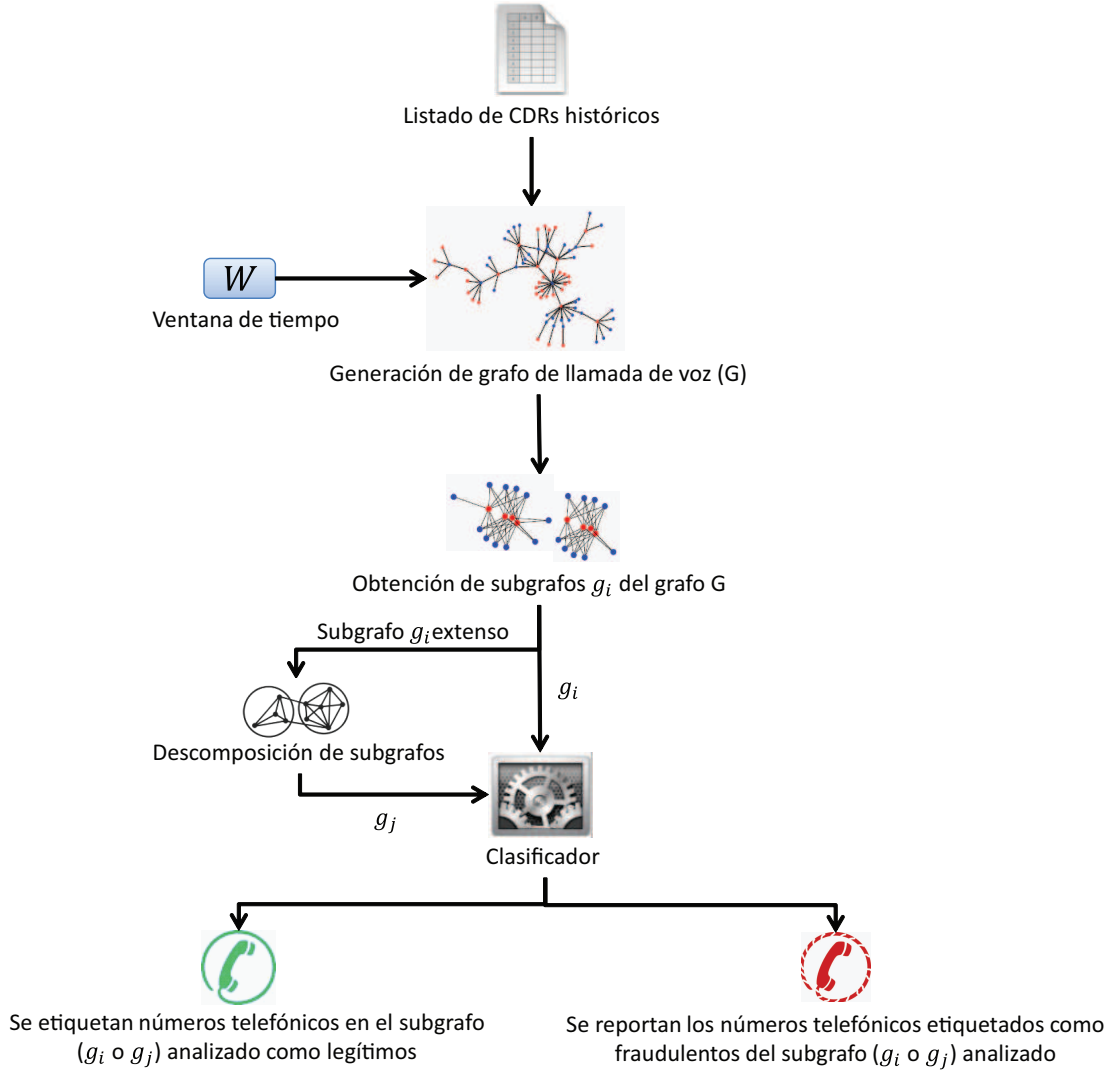


**Fig. 12.** Esquema del enfoque basado en grafo de llamada de voz.

En Jiang et al. [3] se parte de un listado de CDRs de llamadas realizadas por subscriptores de celulares nacionales a números internacionales, en una ventana de tiempo de un mes. Para generar el grafo de llamada de voz las llamadas las analizan en un solo sentido, de los números nacionales (origen) hacia los internacionales (terminales). Se extraen del grafo de llamadas de voz

los subgrafos que al menos tienen un clique. Un clique representa un subgrafo bipartito completo de dos por dos, donde cuatro aristas conectan a dos números nacionales con dos internacionales. Los subgrafos con más de 2000 aristas se descomponen usando Markov Clustering (MCL por sus siglas en inglés). Para detectar posibles números fraudulentos se basan en dos propiedades. La primera consiste en detectar un número telefónico internacional tenga relacionados más de 10 números de nacionales. La segunda se considera que un subgrafo contiene actividades de fraude cuando tiene al menos 5 cliques y los números internacionales que conformen los cliques son considerados posibles números fraudulentos.

Su algoritmo en más del 90 % de los fallos los hace sobre números IRSF de servicio de valor agregado (sección 2.4) que pudieran ser utilizados para cometer fraude, pero que aún no han sido reportados por los subscriptores como fraudulentos. Excluyendo los números PRS y dejando solo los que han sido reportados por los subscriptores, el índice de detección del algorítmo sobre los números de fraude IRSF pudiera exceder el 50 %. Habría que ver que cantidad de números fraudulentos no detectados representa ese 50 % del total. El algoritmo no detecta las actividades de fraude que atraigan menos de diez víctimas en un mes, aunque manifiestan que solo ocurre en menos del 10 % de los casos de fraude, habría que hacerse la misma pregunta de que cantidad representa ese 10 % del total de números fraudulentos.

En esta propuesta se analizan solamente las llamadas realizadas de números nacionales a internacionales omitiendo el sentido contrario, que muy bien serviría de ayuda para identificar fraudes que son generados desde números internacionales. Por ejemplo en un caso de fraude de escaneo aleatorio (sección 2.4) solo se detectaría el fraude si los números a los que les hizo llamadas perdidas, devolvieran la llamada (en esta propuesta específicamente más de 10 números tendrían que devolver la llamada), si la cantidad de llamadas devueltas no excediera el umbral definido simplemente se pasaría por alto el número telefónico fraudulento, que pudo haber hecho llamadas perdidas a una elevada cantidad de números telefónicos. De haber sido analizadas las llamadas entrantes del exterior se pudiera determinar que se estaba en presencia de un fraude de este tipo, disminuyendo así el número de víctimas considerablemente.

Este método también tiene sus ventajas destacándose la detección temprana de los números de fraude IRSF, en más del 80 % de los casos precede los reportes de los subscriptores, llegando en más del 60 % de los casos a detectarlos un mes antes de ser reportados. También detecta características únicas en fraudes típicos de redes celulares.

### *3.4.2. Comunidad de interés*

El enfoque basado en comunidad de interés (COI por sus siglas en inglés) se inicia generalmente a partir de un subscriptor $x$ (ver figura 13). En la base de datos están guardados los CDRs, los cuales contienen información sobre las llamadas realizadas. Entre las informaciones más utilizadas para generar la COI están la duración, la cantidad de llamadas, el subscriptor de origen y el subscriptor de destino. Los CDRs guardados que incluyen a $x$ como subscriptor de origen o subscriptor de destino, son usados para generar un grafo de llamadas donde el vértice central es $x$. El grafo obtenido es conocido como COI de $x$.

En la COI los vértices representan a subscriptores y las aristas la interacción entre ellos. El peso de las aristas está dado por el cálculo de los valores de un atributo de los CDRs, como pueden ser el total de llamadas o el total del tiempo de duración. La distancia entre dos vértices está dada por el número mínimo de aristas que se deben recorrer para unirlos. La profundidad $i$-ésima de la COI es la distancia máxima del vértice central a cada vértice que la conforma. El

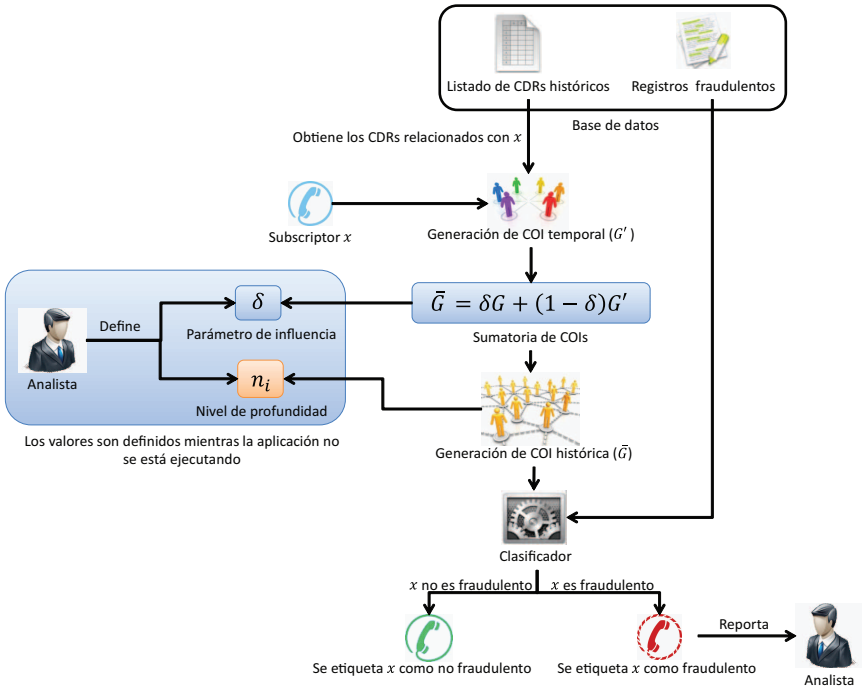


**Fig. 13.** Esquema de métodos basados en COI.

nivel $n_i$ es el conjunto de todos los nodos que están a la misma profundidad y es definido por el analista. En la literatura analizada no se sobrepasa el $n_2$, y se define un umbral para $n_1 = 9$.

Para determinar si $x$ es un posible subscriptor fraudulento dada su COI, es necesario contar con registros de fraudes. Los registros fraudulentos contienen información de fraudes detectados anteriormente, como pueden ser subscriptores fraudulentos detectados o las COIs asociadas a ellos. La clasificación de $x$ se basa en el análisis de su COI teniendo en cuenta los registros fraudulentos, lo cual determina si es un posible subscriptor fraudulento.

En la tabla 1 se comparan dos enfoques distintos de COI, los cuales presentan algunas características similares. Definen la COI histórica de $x$ con un máximo de nueve vértices, los cuales son los de mayor peso histórico en las aristas. El peso de las aristas es la cantidad de llamadas efectuadas. La suma entre dos COI está representada por el operador $\oplus$ y se define como la unión de las aristas y los vértices. Los pesos de las aristas comunes resultantes de la unión se suman y si alguna no está en una COI su valor es cero. La nueva COI histórica $\overline{G}$ de $x$ se obtiene mediante la fórmula $\overline{G} = \delta G \oplus (1-\delta)G'$, donde $G$ es la COI histórica actual de $x$, $G'$ es la COI temporal y $\delta$ es un parámetro definido por el analista para quitarle o darle influencia histórica a $\overline{G}$. Luego de la suma, $\overline{G}$ es ordenada por el peso de las aristas descendentemente y se escogen los nueve primeros subscriptores para conformarla.

### 3.5. Redes bayesianas

Existe una propuesta analizada que se basa en redes bayesianas. Las redes bayesianas se pueden representar como grafos. Los nodos de los grafos representan las variables del dominio del problema. Las flechas entre los nodos representan la relación de dependencia entre ellos. Por ejemplo en

**Tabla 1.** Comparación de dos métodos basados en COI.

| **Artículo** | Cortes et al. [22] | Weigert et al. [23] |
|---|---|---|
| **Generación de la COI temporal** $g$ | Esta se genera diariamente dado que se define un intervalo de tiempo de un día. Al final de cada intervalo de tiempo, se hace la suma entre las dos COIs. | Cuando la cantidad de aristas de $G'$ supera un umbral predefinido se suman las COIs, se iguala a cero el número de aristas de $G'$ y se continúan procesando los CDR. |
| **Parámetro de influencia histórica** $\delta$ | $0 \leq \delta \leq 1$. | $\delta = 0{,}85$ es definido por el analista para darle mayor peso a los datos históricos. |
| **Nivel de profundidad** $n_i$ **de la COI histórica de** $x$ | Es $n_2$ y en $n_1 = 9$. | $n_1 = 9$. |
| **Generación de COI histórica** $\overline{G}$ | Para cada nuevo subscriptor mantiene un registro de la COI histórica en una ventana de tiempo de una semana. El peso de las aristas de los subscriptores no escogidos para conformar $\overline{G}$ es adicionado al peso de la arista que va a un nodo llamado *otros* que es tratado en la COI como un vértice normal. | Los subscriptores no escogidos para conformar $\overline{G}$ son descartados. Para cada subscriptor que generó un CDR mantiene un registro de su COI histórica. De esta forma almacenan todas las COI histórica y las indexan a partir de su vértice central mediante una función hash. Cada vez que se actualiza la BD solo se realizan cambios en la COI histórica de los números que generaron CDRs. Cuando se introduce $x$ al método para ser analizado, se obtiene mediante una función hash su índice y a partir de este se busca y se devuelve su COI histórica. |
| **Registros fraudulentos** | Subscriptores que ya han sido identificados como fraudulentos. | COIs de subscriptores fraudulentos detectados. |
| **Clasificador** | Busca los registros fraudulentos en la COI obtenida. Mientras mayor sea la cantidad de subscriptores fraudulentos que tenga $x$ en su COI, mayor es la probabilidad de que sea fraudulento. Los autores no definen cual es el umbral para saber si es fraudulento o no. | Compara la COI de cada subscriptor fraudulento con la de $x$. Si al menos la COI de un subscriptor fraudulento se intersecta con más del 90 % de la COI de $x$, este representa un posible subscriptor fraudulento. |

la figura 14 se muestra que las variables días laborables, fines de semana y escenario de fraude influyen sobre las variables de frecuencias de llamadas y promedios de duración. Estas dependencias se cuantifican mediante distribuciones condicionales para cada nodo dado sus padres.

Uno de los enfoques de Taniguchi et al. [24] es basado en redes bayesianas, donde un analista en el dominio del problema crea el grafo asumiendo los impactos causales entre las variables. La correspondiente distribución condicional es introducida por el analista quien proporciona criterios sobre las relaciones causales. Teniendo en cuenta la figura 14, cuando se está en presencia de un escenario de fraude suele producirse un aumento de las llamadas internacionales o un aumento en el promedio de duración y frecuencia de las llamadas. De esta forma se puede decir que se está en presencia de una relación de causalidad entre estas variables. Una vez lista la red bayesiana se pueden inferir probabilidades para las variables desconocidas. Las probabilidades se infieren mediante la inserción de evidencias en la red y propagando la evidencias a través de la red mediante el uso de reglas de propagación [25].

En el método analizado se construyen dos redes bayesianas para describir el comportamiento de los subscriptores de telefonía móvil. La primera es construida para modelar el comportamiento bajo la suposición $F$ de que el subscriptor es fraudulento y la otra bajo la suposición $L$ de que el subscriptor es legítimo. La *red de fraude* es establecida a partir del conocimiento del analista. La

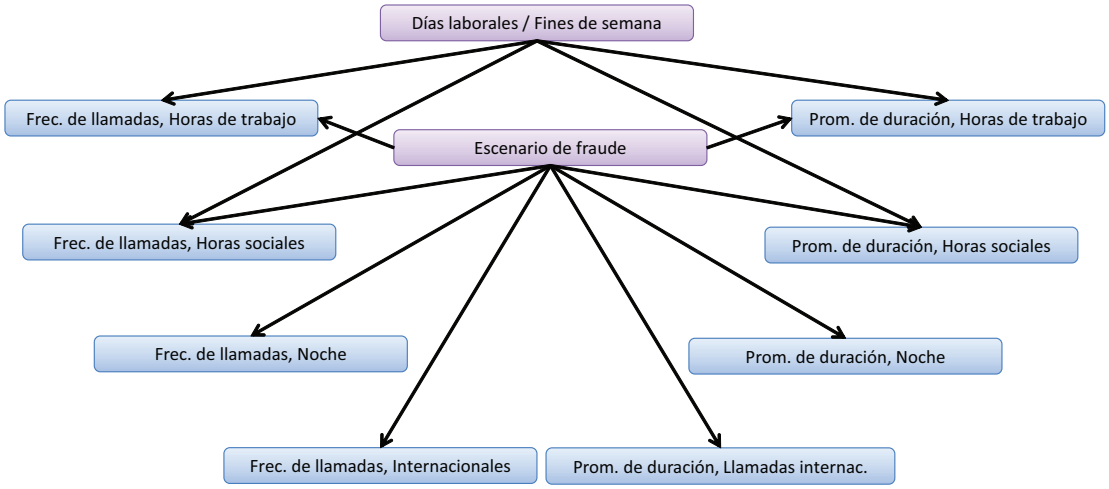


**Fig. 14.** Red bayesiana.

*red de usuarios* es establecida a partir de datos de subscriptores no fraudulentos. En la figura 15 se muestra el esquema del enfoque analizado.

El repositorio de datos consiste en un listado de CDRs históricos. A partir de los CDRs se crea un vector de características que resume el comportamiento de las llamadas de un subscriptor. Las características del vector se pueden ver representadas en las variables de la figura 14. Introduciendo evidencias como el vector de comportamiento $v$, y propagando los datos a través de la red, se obtiene la probabilidad de la medida $\overline{h_k}$ bajo las dos hipótesis. El valor de $\overline{h_k}$ representa la medida de la $k$-ésima característica de $v$. Las probabilidades obtenidas representan en que grado el comportamiento observado de un subscriptor se inclina a un comportamiento fraudulento $P(h_k|F)$ o no fraudulento $P(h_k|L)$. Considerando la probabilidad de fraude $P(F)$ y $P(L) = 1 - P(F)$ y aplicando la regla de Bayes se obtiene la probabilidad de $F$ dada la medida $\overline{h_k}$ (ver ecuación 12).

$$P(F|\overline{h_k}) = \frac{(P(F)P(\overline{h_k}|F))}{(P(\overline{h_k}))}. \tag{12}$$

El denominador $P(\overline{h_k})$ puede ser calculado como se muestra en la ecuación 13.

$$P(\overline{h_k}) = P(F)P(\overline{h_k}|F) + P(L)P(\overline{h_k}|L). \tag{13}$$

El valor de $P(F|\overline{h_k})$ es usado como nivel de alarma. Si $P(F|\overline{h_k}) > z_k$ donde $z_k$ es un umbral definido por un analista para la $k$-ésima característica de $v$, representa que el comportamiento del subscriptor analizado es fraudulento, de lo contrario es etiquetado el subscriptor como legítimo.

### 3.6. Redes Neuronales

Una red neuronal artificial se representa mediante la interconexión de varias neuronas. Cada neurona tiene asociada una función de transferencia. La función genera la señal de salida de la neurona a partir de las señales de entrada [26]. La entrada de la función es la suma de todas las señales de entrada por el peso asociado a la conexión de entrada de la señal. Los parámetros fundamentales de este tipo de red son el número de capas, el número de neuronas por capa y el

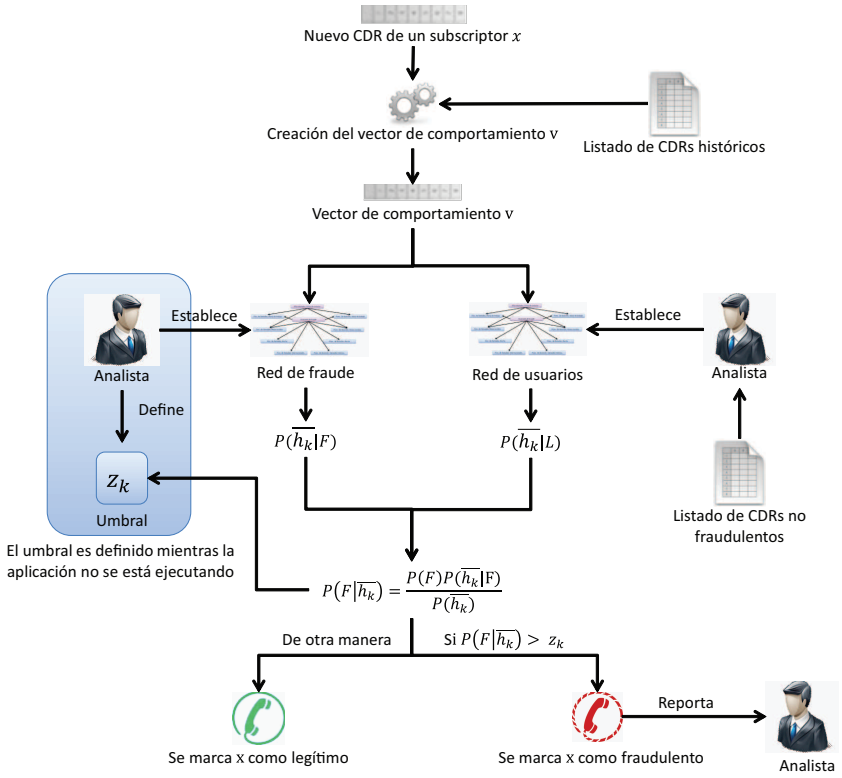


**Fig. 15.** Esquema del enfoque basado en redes bayesianas.

tipo y número de conexiones entre neuronas. En la figura 16 se muestra el esquema de las técnicas basadas en redes neuronales con aprendizaje supervisado.

Los enfoques basados en redes neuronales procesan vectores de características. Las características generalmente usadas son la media y la desviación estándar de la duración y la cantidad de llamadas. Los vectores resumen el comportamiento de los subscriptores en determinadas ventanas de tiempo a partir de un listado de CDRs históricos. Para el entrenamiento se emplean listados de CDRs que incluyen números fraudulentos y legítimos. Los pesos de la red se ajustan basándose en el error de la predicción. La función de ajuste de pesos se define en la ecuación 14.

$$w_{ij}^{*} = w_{ij} + \varphi \cdot \varepsilon_j \cdot \phi, \tag{14}$$

donde $w_{ij}^{*}$ es el nuevo peso, $w_{ij}$ es el peso actual entre la neurona oculta $i$-ésima y la neurona $j$-ésima de la siguiente capa, $\varepsilon_j$ es el error en la neurona $j$-ésima. La variable $\varphi$ define el ritmo de aprendizaje y toma valores entre 0 y 1. Mientras menor sea el valor de $\varphi$ el aprendizaje se hace más lento. Lo cual se debe a que los cambios en los pesos son más pequeños después de cada iteración. A la función de ajuste de pesos en los métodos analizados se le agrega un término $\phi$.

Generalmente para determinar el número óptimo de neuronas ocultas para resolver un problema específico se realiza por prueba y error. Partiendo de una arquitectura entrenada, se realizan cambios en el número de neuronas ocultas y el número de capas aumentándolo y disminuyéndolo, hasta encontrar la arquitectura que se ajuste a la solución del problema.

La cantidad de neuronas en la capa de entrada esta asociada a la cantidad de características escogidas para determinar el fraude. Los métodos analizados cuentan con cinco neuronas en las

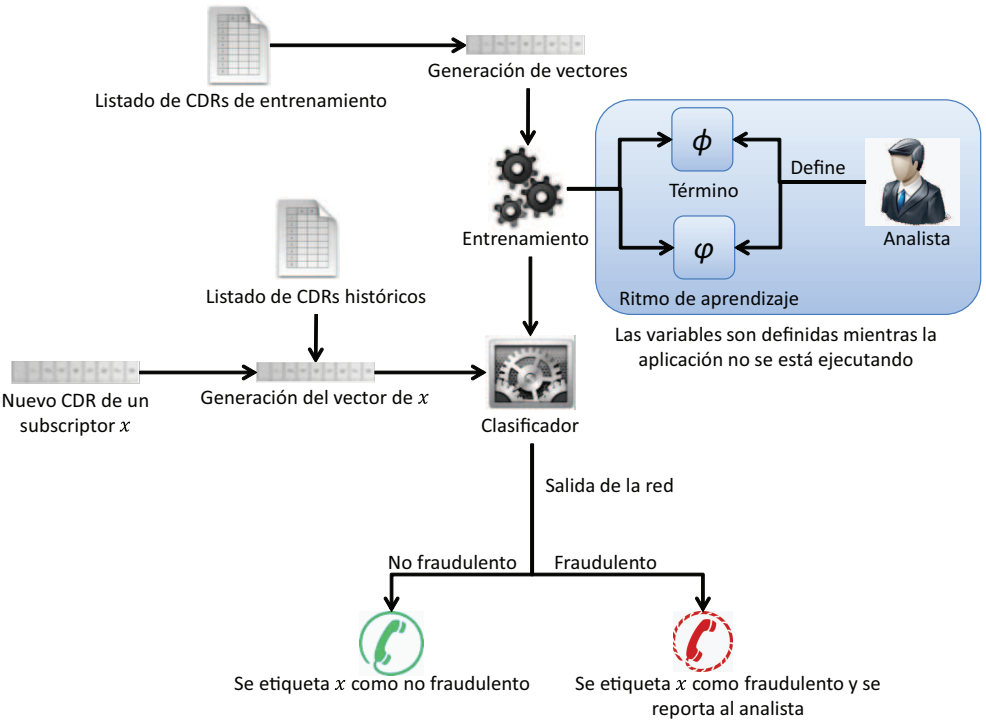


**Fig. 16.** Esquema de las técnicas basadas en redes neuronales con aprendizaje supervisado.

capas ocultas. Se interpreta la salida de la red neuronal, como la probabilidad de que sea fraude dado los datos de entrada. En la tabla 2 se comparan tres enfoques distintos de redes neuronales con aprendizaje supervisado.

### 3.7. Híbridos

El método analizado integra varias técnicas para detectar fraudes. En la propuesta de Panigrahi et al. [30] se integra un modelo basado en reglas, la teoría de Dempster-Shafer [31] y el aprendizaje bayesiano. En la figura 17 se muestra el esquema de este método.

El modelo basado en reglas parte del análisis de un listado de CDRs históricos. Para procesar la información del listado son empleadas dos reglas. La regla $R_1$ se basa en el comportamiento de los subscriptores para detectar fraude. Se define una ventana de tiempo $W$ la cual contiene un conjunto $Z$ de CDRs de un subscriptor. Del conjunto $Z$ se extrae un subconjunto $X$ para formar el perfil del subscriptor y el conjunto $Y = Z \setminus X$ es usado de prueba para detectar un aumento de la duración de las llamadas. A partir de los dos grupos de llamadas se busca si la media de la duración de las llamadas de uno es estadísticamente diferente a la del otro. Para determinar si las medias son estadísticamente diferentes una de la otra se aplica la ecuación 15.

$$q(X,Y) = \frac{\varrho(X) - \varrho(Y)}{\sqrt{\frac{\sigma^2(X)}{|X|} + \frac{\sigma^2(Y)}{|Y|}}}, \tag{15}$$

donde $\varrho(X)$ representa la media de la duración de las llamadas del conjunto $X$, $\sigma^2(X)$ la varianza de las llamadas del conjunto $X$ y $|X|$ la cantidad de llamadas en $X$. En dependencia de si $q(X,Y)$

**Tabla 2.** Comparación de tres métodos basados en Redes Neuronales con aprendizaje supervisado.

| **Artículo** | Taniguchi et al. [24] | Qayyum et al. [27] | Elmi et al. [28] |
|---|---|---|---|
| **Generación de vector** | El vector creado resume las características diariamente durante un período de tiempo no definido. | Se crea un vector histórico que resume el comportamiento de las llamadas de los subscriptores en cada uno de los últimos 4 meses. Se genera un vector actual del comportamiento de las llamadas de las últimas 2 a 3 semanas. | El vector generado contienen un resumen de como se comportan diariamente las características escogidas. |
| **Entrenamiento** | Es entrenada usando la optimización Cuasi-Newton. Se utiliza una técnica de regularización conocida como *decaimiento de peso*. La técnica incluye un término adicional en la función de error $\lambda\Sigma_i w_i^2$. Este término penaliza las redes extensas y por tanto los mapeos complejos [29] reduciendo la varianza. La magnitud de la penalización es determinada por $\lambda$. Para $\lambda = 1$ la red obtuvo el mejor resultado. | A partir de los vectores histórico y actual, se calcula la evolución del comportamiento del subscriptor. A cada caso fraudulento se le asocia una calificación de fraude. Los casos con evoluciones anómalas se les asoció más alta calificación de fraude que a los normales. En la función de ajuste de pesos $\phi = \frac{dp_i(\psi)}{d\psi} * v_i$, donde $v_i$ denota el vector de entrada para la iteración $i$-ésima y la función de transferencia $p(\psi)$ usada es una función sigmoide. | Para la función de ajuste de pesos $\phi = O_i$, donde $O_i$ es la salida de la neurona $i$-ésima oculta. |
| **Clasificador** | Es una red neuronal unidireccional de tres capas. La red cuenta con una neurona de salida binaria. | Es una red neuronal de tipo perceptrón multicapa de tres capas. La capa de entrada esta compuesta por 14 neuronas, y la de salida por una. Si la calificación de la predicción supera un umbral predefinido por el analista, se marca como posible fraude. | Es una red neuronal de tipo perceptrón multicapa con dos capas ocultas. |

es positivo o negativo, definen el grado de variación $\nu(X, Y)$ (ver ecuación 16) de las llamadas de prueba con respecto a las llamadas del perfil. La siguiente fórmula se define para $q(X, Y) > 1$:

$$\nu(X, Y) = 1 - \frac{|\varrho(X) - \varrho(Y)|}{\overline{\gamma}}, \tag{16}$$

donde $\overline{\gamma}$ es la duración máxima de las llamadas en $W$. De no cumplirse $q(X, Y) > 1$, tenemos que $\nu(X, Y) = 0$.

La segunda regla $R_2$ se basa para detectar fraudes en la frecuencia y el tipo de llamada que puede ser internacional, nacional, entre otros. Por ejemplo la regla puede identificar como fraude cuando ocurren muchas llamadas internacionales en el día. Para detectar el incremento brusco de las frecuencias de un tipo de llamadas crean un perfil de frecuencia de llamadas de un subscriptor junto con cada tipo de llamada. Comparan la actividad más reciente con el perfil histórico y calculan la desviación de la frecuencia $\sigma_F$ (ver ecuación 17) de un subscriptor. Para $\dot{\mu}(y) > \ddot{\mu}(y)$ se define la siguiente función:

$$\sigma_F = 1 - \frac{|\dot{\mu}(y) - \ddot{\mu}(y)|}{\bar{\Omega}_y}, \tag{17}$$

donde $\dot{\mu}(y)$ es el promedio de llamadas de tipo $y$ por día, $\ddot{\mu}(y)$ es el promedio de llamadas de tipo $y$ por mes y $\bar{\Omega}_y$ es la máxima frecuencia de llamadas de tipo $y$ por día en un mes determinado. De no cumplirse $\dot{\mu}(y) > \ddot{\mu}(y)$ entonces $\sigma_F = 0$.

Luego es aplicada La teoría de Dempster-Shafer con el propósito de combinar los resultados de las reglas $R_1$ y $R_2$ y obtener una probabilidad $P(F)$ que permita determinar que tan legítima puede ser una llamada. La hipótesis $F$ implica que la llamada es fraudulenta. El valor de $P(F)$ (ver ecuación 18) es calculado combinando $u_1(F)$ y $u_2(F)$, que son asignaciones básicas de probabilidad (ABP) para $R_1$ y $R_2$ respectivamente.

$$P(F) = u_1(F) + u_2(F) = \frac{\sum_{Q_1 \cap Q_2 = \{F\}} u_1(Q_1) \cdot u_2(Q_2)}{1 - \sum_{Q_1 \cap Q_2 = \{F\}} u_1(Q_1) \cdot u_2(Q_2)}, \tag{18}$$

donde $Q_1$ y $Q_2$ son conjuntos de hipótesis. Para obtener los valores de las ABP se tienen en cuenta tres hipótesis: hipótesis $F$ implica que la llamada es fraudulenta, hipótesis $L$ la llamada no es fraudulenta y la hipótesis $S$ implica que la llamada es sospechosa. Las ABP para $R_1$ y $R_2$ son dadas por:

ABP para $R_1$:

$$u_1(F) = \nu(X, Y), \tag{19}$$

$$u_1(L) = 0, \tag{20}$$

$$u_1(S) = 1 - \nu(X, Y). \tag{21}$$

ABP para $R_2$:

$$u_2(F) = \sigma_F, \tag{22}$$

$$u_2(L) = 0, \tag{23}$$

$$u_2(S) = 1 - \sigma_F. \tag{24}$$

El cero en la ABP de $L$ significa que la regla $R_1$ da como cero el grado de probabilidad con respecto a la legitimidad de una llamada. Para obtener el valor de $P(L)$ se tiene que $P(L) = 1 - P(F)$.

El analista define un umbral de probabilidad alto $z_1$ y uno bajo $z_2$ para clasificar el CDR generado por una llamada como se muestra a continuación. Un CDR puede ser etiquetado como legítimo si $P(F) < z_2$, fraudulento si $P(F) > z_1$ o sospechoso si $z_2 < P(F) < z_1$. El CDR $a$ etiquetado como sospechoso es enviado a una base de datos donde se almacenan los CDR sospechosos (BDS) y el subscriptor que realizó la llamada es etiquetado como sospechoso también. Cuando el subscriptor sospechoso realiza otra llamada generando el CDR $a_1$ se le aplica el mismo proceso y en caso de ser etiquetado el CDR como sospechoso se analiza insertándolo a la BDS.

En la base de datos de los CDRs sospechosos, se construye un perfil histórico $N(s_j)$ para clientes individuales y un historial de fraude genérico $N'$ para comparar los nuevos comportamientos de llamadas con los perfiles de fraude genérico. El perfil histórico es construido a partir del comportamiento de las llamadas pasadas de clientes individuales. El historial de fraude genérico es construido a partir de datos de fraudes anteriores.

En este método se definen doce eventos $D_{d,y}$ que representan la cantidad de llamadas de un mismo subscriptor teniendo en cuenta que $y$ es el tipo de llamada y $d$ un intervalo de tiempo. Por ejemplo $y = 0$ representa llamadas internacionales y $d = 1$ representa el intervalo de 0 a 8

horas después de la última llamada, por tanto el evento $D_{1,0}$ representa la ocurrencia de llamadas internacionales en el intervalo definido por $d$.

Teniendo en cuenta el evento $D_{d,y}$ que a ocurrido se calculan $P(D_{d,y}|F)$ y $P(D_{d,y}|L)$. $P(D_{d,y}|F)$ es la probabilidad de ocurrencia de $D_{d,y}$ dado que la llamada es originada por un cliente fraudulento y $P(D_{d,y}|L)$ es la probabilidad de ocurrencia de $D_{d,y}$ dado que la llamada es originada por un cliente legítimo. Para determinar $P(D_{d,y}|F)$ (ver ecuación 25) y $P(D_{d,y}|L)$ (ver ecuación 26) crean dos tablas de consultas. La tabla de frecuencia de llamada de fraude (TFF) se crea a partir del $N'$. La tabla de frecuencia de llamada legítima (TFL) se forma a partir del $N(s_j)$. Para su cálculo usan las siguientes funciones:

$$P(D_{d,y}|F) = \frac{\varsigma(D_{d,y})}{\chi(N')}, \tag{25}$$

$$P(D_{d,y}|L) = \frac{\varsigma(x, D_{d,y})}{\chi(x, N(s_j))}, \tag{26}$$

donde $\varsigma(D_{d,y})$ es la cantidad de ocurrencias de $D_{d,y}$, $\chi(N')$ la cantidad de llamadas en el $N'$, $\varsigma(x, D_{d,y})$ la cantidad de ocurrencias de $D_{d,y}$ del subscriptor $x$ y $\chi(x, N(s_j))$ la cantidad de llamadas en el $N(s_j)$ del subscriptor $x$. A partir de las ecuaciones 25 y 26 se calcula $P(D_{d,y})$ (ver ecuación 27).

$$P(D_{d,y}) = P(D_{d,y}|F)P(F) + P(D_{d,y}|L)P(L). \tag{27}$$

A partir de los valores $P(D_{d,y}|F)$ y $P(D_{d,y}|L)$ obtenidos de las tablas TFF y TFL respectivamente se aplica el aprendizaje bayesiano para obtener $P(F|D_{d,y})$ (ver ecuación 28) y $P(L|D_{d,y})$ (ver ecuación 29).

$$P(F|D_{d,y}) = \frac{P(D_{d,y}|F)P(F)}{P(D_{d,y})}, \tag{28}$$

$$P(L|D_{d,y}) = \frac{P(D_{d,y}|L)P(L)}{P(D_{d,y})}. \tag{29}$$

Si $P(F|D_{d,y}) > P(L|D_{d,y})$ se vuelve a aplicar la teoría de Dempster-Shafer para combinar el valor $P(F|D_{d,y})$ del CDR $a$ con el $P(F)$ del $(a_1)$ generado por el subscriptor en análisis. De no cumplirse la condición anterior solamente se devuelve el $P(F)$ de $(a_1)$. A partir del resultado obtenido se etiqueta el CDR $(a_1)$ con una de las tres categorías establecidas. Este proceso continua hasta que $P(F)$ caiga por debajo del umbral $z_2$ o exceda el umbral $z_1$.

### 3.8. Conclusiones parciales

En lo anteriormente expuesto se evidencia la variedad de técnicas de detección de fraude que han sido propuestas e implementadas. Cada una con sus características particulares que las hacen más ventajosas con respecto a otras.

Por ejemplo los métodos basados en evaluación de reglas son muy eficientes sobre los fraudes que se tiene conocimiento de su existencia. Por otra parte, las técnicas como las basadas en detección de anomalías permiten identificar nuevos tipos de fraudes sin un conocimiento previo de su existencia. Los métodos basados en generación automática de reglas cuentan con características que los hacen muy útiles para la evaluación de reglas. Cuando se procesa automáticamente un

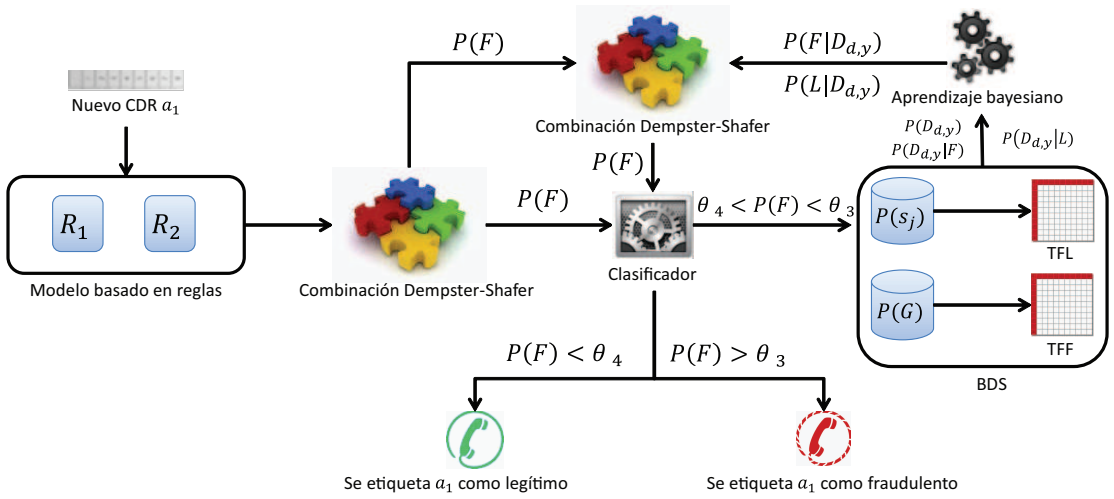


**Fig. 17.** Esquema del método híbrido.

conjunto de datos etiquetados, se pueden obtener reglas que son prácticamente imposibles de generar por un analista. El resultado es un conjunto de reglas que de ser adicionado a la base de un motor de reglas, le permitiría detectar fraudes que anteriormente no podían haber sido definidos mediante reglas por un analista. Lo cual traería considerables mejoras en la eficiencia del sistema de detección de fraudes.

En esta sección se abordaron las técnicas empleadas para la detección de fraudes en servicios de telecomunicaciones. Fueron analizados varios trabajos basados en distintos enfoques y se propusieron recomendaciones para mejorar la eficacia y eficiencia en algunos casos. También fueron mostrados los esquemas para cada uno de los enfoques tratados, los cuales permiten comprender con mayor claridad como funciona cada método. Se presentó la taxonomía para las técnicas de detección de fraude. Existen técnicas para detectar fraudes que se aplican en otros ámbitos, las cuales serán abordadas en la siguiente sección.

## 4. Técnicas de detección de fraude en otras esferas

Como parte de la investigación desarrollada en este trabajo, se analizaron también propuestas de técnicas de detección de fraude orientadas a otras áreas fuera de los servicios de telecomunicaciones (ver sección 4). El análisis efectuado, contribuye a comprender el funcionamiento de dichas técnicas e identificar características que permitan establecer mejoras en la detección de fraudes en los servicios de telecomunicaciones.

Algunas técnicas de detección de fraude utilizadas en otras esferas como en tarjetas de crédito [32], subastas online [33], mercado de valores [34], entre otras; pudieran extenderse para su posterior aplicación en servicios de telecomunicaciones. A continuación es presentado un análisis sobre algunas de las técnicas en cuestión.

En el trabajo de van Leeuwen y Siebes [35] proponen un algoritmo para detectar cambios en una secuencia de datos. Basándose en el análisis del principio de la longitud mínima de la descripción (MDL por sus siglas en inglés), se particiona una secuencia en varias subsecuencias. Luego para cada subsecuencia hacen uso del algoritmo heurístico KRIMP [36] para caracterizar

su distribución con una tabla de códigos. En cada una de estas tablas que se genere por cada subsecuencia indica un cambio en la distribución subyacente. La tabla de códigos es usada como base por el algoritmo propuesto para detectar los cambios en la secuencia. En el experimento realizado sobre el conjunto de datos de ajedrez del repositorio UCI, el algoritmo no detecta las clases más pequeñas. Lo cual representa un problema si se tiene en cuenta que en un conjunto de datos de servicios de telecomunicaciones la clase de fraudes es muy inferior a la de no fraudes.

La propuesta de Zhao et al. [37] se enfoca en la clasificación de secuencias de transacciones usando patrones negativos y positivos. Basándose en ambos tipos de patrones detectados en el entrenamiento, pasan al conjunto de patrones secuenciales clasificables aquellos que tienen un mayor significado para la clasificación. Lo cual se define teniendo en cuenta su frecuencia y el coeficiente de correlación de clase (CCR por sus siglas en inglés) [38]. Posteriormente le asignan una puntuación a los patrones para la clasificación. Luego dada una secuencia, la puntuación de todos los patrones secuenciales clasificables que se detectan es sumada. Después es clasificada la secuencia como fraudulenta o legítima atendiendo a la mayor puntuación obtenida por los patrones negativos o positivos. Entre los problemas se destaca la ventana de tiempo la cual fijaron para cuatro meses, por lo que si una actividad de fraude viene desarrollándose con anterioridad de manera estable se pasaría por alto.

La propuesta de McGlohon et al. [39] se enfoca en un algoritmo basado en vínculos analíticos para el etiquetado de grafos y detección de riesgo. La premisa básica del algoritmo es usar las clases de los nodos vecinos para clasificar un nodo dado. Tienen en cuenta el conocimiento del dominio que se basa en información de patrones de comportamiento conocidos con anterioridad de defraudadores potenciales. Basado en estos patrones y la cantidad que cubran un nodo, los expertos del dominio asignan una puntuación de riesgo inicial a los nodos antes de evaluar las asociaciones de vecindad entre ellos. Para etiquetar un nodo como perteneciente a una clase, la información se infiere de los nodos circundantes. Pasando mensajes iterativos entre los nodos vecinos, donde cada mensaje evalúa la clase a la que pertenece el nodo vecino se infiere la información. Al final del procedimiento se determina una probabilidad estimada que comparada con las puntuaciones de riesgo de los nodos conectados a él o un umbral puede definir a que clase pertenece el nodo. El problema esta en que si un nodo fraudulento tiene como vecinos a nodos con una puntuación de riesgo baja, la probabilidad de que este sea fraudulento es baja.

Li et al. [40] propone la búsqueda de dos patrones. El patrón de agujero negro (en inglés blackhole) es definido como un grupo de nodos que solo tienen enlaces de entrada del resto de los nodos en el grafo. El patrón volcán (en inglés volcano) es lo opuesto del agujero negro, es un grupo de nodos que solo tiene enlaces de salida con el resto de los nodos del grafo. Su enfoque más bien se centra en la detección de patrones de agujero negro. Con la detección de estos patrones en una red de comunicaciones, se puede proporcionar una visión de algunas propiedades estructurales, que ayuden a comprender mejor las interacciones entre algunos nodos de la red. Sin embargo en el algoritmo presentado no tienen en cuenta el peso en los enlaces o aristas para determinar la existencia de este patrón. No tener en cuenta el peso hace que se pierda información que podría resultar ser muy valiosa en la detección de fraudes en servicios de telecomunicaciones. Agregándole peso a las aristas, que represente atributos claves de una llamada y tenerlo en cuenta para reconocer el patrón, llevaría a una mejor aplicación de esta técnica en la detección de fraudes en servicios de telecomunicaciones.

Un grafo de revisión es la propuesta de Wang et al. [41] compuesto por tres tipos distintos de nodos los cuales representan usuarios, sus opiniones y tiendas en línea. El objetivo es capturar relaciones entre estos, con el propósito de determinar cuales usuarios y opiniones son fraudulentos.

Para ello cuando una opinión se desvía o va en contra de la mayoría de las opiniones en una tienda, es marcada como posible falsa, lo cual se tiene en cuenta para definir al usuario que la escribió como un usuario legítimo o fraudulento. Atendiendo al comportamiento de las opiniones de un usuario en varias tiendas, donde sean identificados sus criterios como falsos, se puede identificar y alertar los demás usuarios y tiendas, sobre el usuario fraudulento y sus opiniones.

El grafo de revisión se podría enfocar en alguna medida a la detección de fraudes en servicios de telecomunicaciones. Si se tiene en cuenta que muchos defraudadores usan varios números telefónicos para cometer fraude, ya que de esta manera no exponen un número a alta actividad fraudulenta, sino que se distribuye entre varios haciendo que la actividad para un número disminuya. La baja frecuencia de actividad fraudulenta en un número podría pasarse por alto y no detectar el fraude. Con el uso de este enfoque se podría detectar el fraude en este caso. Para ello se necesitaría determinar si un cliente está subscrito con varios números telefónicos. A partir de lo cual se procede a comprobar teniendo en cuenta todas las llamadas del cliente si su comportamiento se sale de lo normal o lo común.

El trabajo de Boding et al. [42] se basa en un sistema de detección de fraudes para transacciones financieras en linea. Dicho sistema cuenta con una base de datos de reglas de detección de fraude y otra de perfiles de comerciantes. Un perfil de comerciante contiene un nombre, una descripción y un conjunto de reglas. Un comerciante puede tener varios perfiles. Las reglas son asignadas a los perfiles, permitiendo tener una misma regla asociada a más de un perfil.

El proceso de detección de fraude comienza cuando un cliente tiene la intención de adquirir algún bien o servicio en linea. Luego de establecerse la conexión entre el cliente y el servidor del comerciante, el cliente procede a llenar su orden. Una vez terminada la orden se genera una transacción que llega al servidor del comerciante. Los detalles de la transacción como el nombre del cliente, su dirección, su número de teléfono, número de cuenta, artículos comprados, precio de los artículos, entre otros, son enviados al sistema de detección de fraude. El sistema de detección de fraude se encarga de validar si la transacción es legítima o fraudulenta y envía una respuesta al servidor del comerciante. De ser legítima la transacción se procesa la orden y en caso de ser fraudulenta es rechazada.

Si nos enfocáramos en las telecomunicaciones se pudiera sustituir los perfiles de comerciantes por perfiles de llamadas. De esta forma tendríamos perfiles identificados por el tipo de llamada: internacional, nacional, emergencia, etc. Donde los perfiles pudieran compartir las mismas reglas o tener reglas específicas asignadas.

## 4.1. Conclusiones parciales

El análisis efectuado en esta sección, permitió conocer soluciones que han sido implementadas en otras esferas para detectar fraudes. Las recomendaciones hechas en cada propuesta analizada en esta sección, pueden servir como punto de partida para futuras investigaciones. Las cuales podrían estar enfocadas en estudiar posibles modificaciones en las soluciones con el fin de aplicarlas en los servicios de telecomunicaciones. Para desarrollar nuevas propuestas se hace necesario conocer cuales son los problemas existentes que dificultan el proceso de detección de fraude, lo cual se aborda en la proxima sección.

## 5. Conclusiones generales sobre el estado del arte

Constituye un gran reto elaborar una solución que detecte con alta precisión una gran variedad de fraudes en servicios de telecomunicaciones. En la detección de anomalías establecer un umbral en los datos para diferenciar un comportamiento normal de uno anómalo resulta ser muy difícil y no muy preciso. Contar con métodos para tener presente que no se disparen falsos positivos en los resultados es de gran importancia, ejemplo de esto es la existencia de líneas calientes. Los números pertenecientes a líneas calientes pueden ser de hospitales, emergencias, hoteles, entre otros. De igual manera las fechas de cumpleaños, aniversarios, etc, pueden ser la causa de detección de muchos falsos positivos.

Un factor importante a tener en cuenta antes de desarrollar una técnica de detección de fraude es el avance en el desarrollo de los servicios de telecomunicaciones. Con el desarrollo de las telecomunicaciones el comportamiento de las llamadas telefónicas que pudiera ser definido en la actualidad como normal o legítimo, en un futuro podría no ser suficientemente representativo para determinar si es un comportamiento fraudulento o no. Lo cual se debe al desarrollo de los servicios y la disponibilidad de llamadas, dándole posibilidad a los clientes de efectuar llamadas como pudieran ser las internacionales, con una mayor frecuencia y por largos períodos de tiempo a tarifas mucho más reducidas.

Realizar dichas llamadas con esos tipos de características llevaría a un cambio en el comportamiento de las llamadas en el futuro. Si se tiene en cuenta que la tendencia en la actualidad de un subscriptor común, es la poca frecuencia de llamadas internacionales y por muy breves períodos de tiempo. Por lo cual, los patrones que pudieran definir como un comportamiento sospechoso la alta frecuencia de llamadas internacionales o períodos largos de tiempo, en un futuro podrían representar un comportamiento normal.

Identificar las estructuras comunes de cada uno de los distintos tipos de fraude, sigue siendo uno de los principales problemas para todos los que desarrollan técnicas para su detección. Esto se debe a que por lo general las relaciones establecidas entre números telefónicos, basándose en las llamadas tanto de entrada como de salida, conectan muchas comunidades formando extensos subgrafos donde pueden haber presentes varios tipos de actividades de fraude.

Obtener una base de datos real de telecomunicaciones sobre la cual se puedan ejecutar los experimentos de las distintas propuestas para poder determinar con cual se obtiene una mejor presición en las detecciones de fraude resulta ser prácticamente imposible por razones de privacidad y limitaciones legales. No es posible hacer una comparación entre los resultados de autores distintos ya que las bases de datos sobre las que fueron evaluados no son las mismas.

La detección de fraudes en servicios de telecomunicaciones representa un amplio campo para la investigación. La aplicación de distintos enfoques para la detección, así como la diversidad de técnicas para cometer fraudes fueron analizados en este trabajo. De los enfoques tratados muchos han sido puestos en práctica, sin embargo se continúan haciendo nuevas propuestas para mejoras técnicas y para darle solución a problemas específicos.

Procesar el gran volumen de datos que es almacenado por las compañías de telecomunicaciones, se hace prácticamente imposible para una persona. Por lo cual el uso de técnicas de minería de datos se hace indispensable para la detección de fraudes en servicios de telecomunicaciones. Entre las ventajas que ofrecen estas técnicas se destacan las posibilidades que brindan para detectar patrones y comportamientos que permitan identificar cuando se está en presencia de una actividad fraudulenta.

El análisis de técnicas de detección de fraude empleadas en otras esferas, permite tener en cuenta soluciones propuestas en otras areas que puedan ser modificadas con el fin de aplicarlas en los servicios de telecomunicaciones.

Los problemas detectados sobre la base del estudio realizado representan un punto de partida para futuras investigaciones. Conocer como funcionan las técnicas para cometer fraude, así como los métodos de detección analizados, representa una base sobre la cual consolidar una propuesta que contribuya reducir los daños ocasionados por actividades fraudulentas.

## Referencias bibliográficas


1. Laleh, N., Azgomi, M.A.: A taxonomy of frauds and fraud detection techniques. Communications in Computer and Information Science, Springer **31** (2009) 256–267
2. Páez, P.G.: El negocio del fraude en la industria de las telecomunicaciones. [citado 29 de agosto de 2013]. Disponible en Internet: http://elmayorportaldegerencia.com/index.php/publicaciones-gerenciales
3. Jiang, N., Jin, Y., Skudlark, A., Hsu, W.L., Jacobson, G., Prakasam, S., Zhang, Z.L.: Isolating and analyzing fraud activities in a large cellular network via voice call graph analysis. In: Proceedings of the 10th international conference on Mobile systems, applications, and services. MobiSys '12, New York, NY, USA, ACM (2012) 253–266
4. Almeida, M.P.: Classification for fraud detection with social network analysis. Tesis de Maestría, Instituto Superior Técnico de Lisboa, Portugal (October 2009)
5. Moudani, W., Chakik, F.: Fraud detection in mobile telecommunication. Lecture Notes on Software Engineering **1**(1) (2013)
6. Becker, R.A., Volinsky, C., Wilks, A.R.: Fraud Detection in Telecommunications: History and Lessons Learned. Technometrics **52**(1) (2010) 20–33
7. Ruiz-Agundez, I., Penya, Y.K., Garcia Bringas, P.: Fraud detection for voice over ip services on next-generation networks. In: Proceedings of the 4th IFIP WG 11.2 International Conference on Information Security Theory and Practices: Security and Privacy of Pervasive Systems and Smart Devices. WISTP '10, Berlin, Heidelberg, Springer-Verlag (2010) 199–212
8. Burge, P., Shawe-taylor, J., Cooke, C., Moreau, Y., Preneel, B., Stoermann, C.: Fraud detection and management in mobile telecommunications networks. In: Proceedings of the European Conference on Security and Detection ECOS '97, London, UK, IEEE (1997) 91–96
9. Verrelst, H., Lerouge, E., Moreau, Y., Vandewalle, J., Stormann, C., Burge, P.: A rule based and neural network system for fraud detection in mobile communications. In: Proceedings of European Conference on Circuit Theory and Design (ECCTD'99), Stresa, Italy (2003) 843–846
10. McGibney, J., Hearne, S.: An approach to rules-based fraud management in emerging converged networks. In: Proceedings of 2nd IEE/IEEE Irish Telecommunications Systems Research Symposium, ITSRS 2003, Dublin, Ireland (May, 2003)
11. Rosset, S., Murad, U., Neumann, E., Idan, Y., Pinkas, G.: Discovery of fraud rules for telecommunications-challenges and solutions. In: Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '99, New York, NY, USA, ACM (1999) 409–413
12. Quinlan, J.R.: C4.5: programs for machine learning. San Francisco, California, Morgan Kaufmann (1993) 302 p.
13. Rajani, S., Padmavathamma, M.: A model for rule based fraud detection in telecommunications. International Journal of Engineering Research & Technology **1**(5) (July, 2012) 1–7
14. Akoglu, L., Faloutsos, C.: Event detection in time series of mobile communication graphs. In Army Science Conference, Orlando, Florida, USA (2010)
15. Singh, G.: Voip anomaly detection engine (voipade). [citado 16 de septiembre de 2013]. Disponible en Internet: http://www.terena.org/activities/campus-bp/pdf/gn3-na3-t4-ufs134.pdf
16. Sengar, H., Wang, H., Wijesekera, D., Jajodia, S.: Detecting voip floods using the hellinger distance. In IEEE Transactions on Parallel and Distributed Systems **19**(6) (June 2008) 794–805
17. Jacobson, V.: Congestion avoidance and control. SIGCOMM Computer Communication Review, New York, NY, USA **25**(1) (1995) 157–187
18. Nithi, Dey, L.: Anomaly detection from call data records. In: Proceedings of the 3rd International Conference on Pattern Recognition and Machine Intelligence. PReMI '09, Berlin, Heidelberg, Springer-Verlag (2009) 237–242

19. Ramaswamy, S., Rastogi, R., Shim, K.: Efficient algorithms for mining outliers from large data sets. In: Proceedings of the 2000 ACM SIGMOD international conference on Management of data. SIGMOD '00, New York, NY, USA, ACM (2000) 427–438
20. Murynets, I., Jover, R.P.: Anomaly detection in cellular machine-to-machine communications. In IEEE International Conference On Communications, Budapest, Hungary (2013)
21. Abidogun, O.A.: Data mining, fraud detection and mobile telecommunications. Call Pattern Analysis With Unsupervised Neural Networks. Tesis de Maestría, Universidad de Western Cape, South Africa (2005)
22. Cortes, C., Pregibon, D., Volinsky, C.: Communities of interest. In: Proceedings of the Fourth International Conference on Advances in Intelligent Data Analysis, Cascais, Portugal. (2001) 105–114
23. Weigert, S., Hiltunen, M., Fetzer, C.: Mining large distributed log data in near real time. In: Managing Large-scale Systems via the Analysis of System Logs and the Application of Machine Learning Techniques. SLAML '11, New York, NY, USA, ACM (2011) 51–58
24. Taniguchi, M., Haft, M., Hollmén, J., Tresp, V.: Fraud detection in communications networks using neural and probabilistic methods. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '98), Seattle, WA, USA **2** (1998) 1241–1244
25. Jensen, F.V.: Introduction to Bayesian Networks. 1st edn. Springer-Verlag New York, Inc., Secaucus, NJ, USA (1996)
26. Oluwagbemi, O.O.: A comparative analysis of different neural networks performances in the prediction of superimposed fraud in mobile phone usage. Covenant University College of Science & Technology, Department of Computer & Information Sciences Ota, Ogun State Nigeria. (2010)
27. Qayyum, S., Mansoor, S., Khalid, A., Khushbakht, K., Halim, Z., Baig, R.R. In: International Conference on Information and Emerging Technologies (ICIET), Karachi, Pakistan. (June 2010) 1–5
28. Elmi, A.H., Ibrahim, S., Sallehuddin, R.: Detecting sim box fraud using neural network. Lecture Notes in Electrical Engineering, Springer **215** (2012) 575–582
29. Bishop, C.M.: Neural Networks for Pattern Recognition. Oxford University Press, Inc., New York, NY, USA (1995)
30. Panigrahi, S., Kundu, A., Sural, S., Majumdar, A.: Use of dempster-shafer theory and bayesian inferencing for fraud detection in mobile communication networks. In Proceedings of the Australasian Conference on Information Security and Privacy (ACISP), Townsville, Queensland (2007) 446–460
31. Chen, T.M., Venkataramanan, V.: Dempster-shafer theory for intrusion detection in ad hoc networks. IEEE Internet Computing, Piscataway, NJ, USA **9**(6) (November 2005) 35–41
32. Chaudhary, K., Yadav, J., Mallick, B.: Article: A review of fraud detection techniques: Credit card. International Journal of Computer Applications **45**(1) (May 2012) 39–44
33. Lin, S.J., Jheng, Y.Y., Yu, C.H.: Combining ranking concept and social network analysis to detect collusive groups in online auctions. Expert Systems with Applications, Pergamon Press, Inc., Tarrytown, NY, USA **39**(10) (August 2012) 9079–9086
34. Golmohammadi, K., Zaiane, O.R.: Data mining applications for fraud detection in securities market. In: Proceedings of the 2012 European Intelligence and Security Informatics Conference. EISIC '12 (2012) 107–114
35. Leeuwen, M., Siebes, A.: Streamkrimp: Detecting change in data streams. In: Proceedings of the 2008 European Conference on Machine Learning and Knowledge Discovery in Databases - Part I. ECML PKDD '08, Berlin, Heidelberg, Springer-Verlag (2008) 672–687
36. Siebes, A., Vreeken, J., van Leeuwen, M.: Item sets that compress. In: Proceedings of the SIAM Conference on Data Mining (SDM'06), Bethesda, Maryland, USA (2006) 393–404
37. Zhao, Y., Zhang, H., Wu, S., Pei, J., Cao, L., Zhang, C., Bohlscheid, H.: Debt detection in social security by sequence classification using both positive and negative patterns. In: Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases: Part II. ECML PKDD '09, Berlin, Heidelberg, Springer-Verlag (2009) 648–663
38. Verhein, F., Chawla, S.: Using significant, positively associated and relatively class correlated rules for associative classification of imbalanced datasets. In: Proceedings of the 7th IEEE International Conference on Data Mining (ICDM'07), Omaha, Nebraska, USA (2007) 679–684
39. McGlohon, M., Bay, S., Anderle, M.G., Steier, D.M., Faloutsos, C.: Snare: a link analytic system for graph labeling and risk detection. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '09, New York, NY, USA, ACM (2009) 1265–1274
40. Li, Z., Xiong, H., Liu, Y., Zhou, A.: Detecting blackhole and volcano patterns in directed networks. In: Proceedings of the 10th IEEE International Conference on Data Mining (ICDM '10), Sydney, Australia (2010) 294–303

41. Wang, G., Xie, S., Liu, B., Yu, P.S.: Review graph based online store review spammer detection. In: Proceedings of the 11th IEEE International Conference on Data Mining (ICDM '11), Vancouver, BC, Canada (2011) 1242–1247
42. Boding, B.S., Siddens, C.H.: Fraud detection system automatic rule population engine. Patent. US 2012/0278246 A1 (November 2012)

## Anexo 1

Con la intención de que el trabajo sea autocontenido, este anexo contiene los conceptos básicos que se utilizan en el desarrollo del trabajo. Conocer en que consiste cada concepto es necesario para comprender el funcionamiento de los enfoques de detección de fraude presentados.

*Análisis de componentes principales*: técnica utilizada para reducir la dimensionalidad de un conjunto de datos. Intuitivamente la técnica sirve para hallar las causas de la variabilidad de un conjunto de datos y ordenarlas por importancia. Para reducir la dimensionalidad retiene aquellas características del conjunto de datos que contribuyen más a su varianza, manteniendo un orden de bajo nivel de los componentes principales e ignorando los de alto nivel. El objetivo es que esos componentes de bajo orden a veces contienen el aspecto más importante de esa información [1].

*Asignación básica de probabilidades*: consiste en una función para asignar a cada hipótesis un valor indicativo de la creencia que dada una evidencia, se deposita en dicha hipótesis. Dicha función es parecida a la función de densidad de probabilidad, pero en la que no se respeta la restricción bayesiana de que la suma de la creencia asignada a las hipótesis originales deba ser uno. Esto quiere decir que confirmar una determinada creencia para una hipótesis no implica confirmar la creencia restante para su negación [2,3,4].

*Correlación*: determina la relación o dependencia que existe entre dos variables estadísticas [5]. Se considera que dos variables cuantitativas están correlacionadas cuando los valores de una de ellas varían sistemáticamente con respecto a los valores homónimos de la otra. Es decir, si los cambios en una de las variables influyen en los cambios de la otra se puede decir que las variables están correlacionadas o que existe correlación entre ellas. La correlación se obtiene dividiendo la covarianza de dos variables entre el producto de sus desviaciones estándar como se muestra en la ecuación 1.

$$\rho(\overline{b_1}, \overline{b_2}) = \frac{\sigma(\overline{b_1}, \overline{b_2})}{\sigma(\overline{b_1})\sigma(\overline{b_2})}, \tag{1}$$

donde $\overline{b_1}$ y $\overline{b_2}$ son variables cuantitativas, $\sigma(\overline{b_1}, \overline{b_2})$ es la covarianza de dichas variables. Los valores de $\sigma(\overline{b_1})$ y $\sigma(\overline{b_2})$ representan la desviación típica de $\overline{b_1}$ y $\overline{b_2}$ respectivamente.

*Covarianza*: es un valor que indica el grado de variación conjunta de dos variables aleatorias. Es el dato básico para determinar si existe una dependencia entre ambas variables. Además es el dato necesario para estimar parámetros básicos, como el coeficiente de correlación lineal. La covarianza entre dos variables aleatorias reales $\psi_1$ y $\psi_2$ se define en la ecuación 2.

$$\sigma(\psi_1, \psi_2) = J[(\psi_1 - J[\psi_1])(\psi_2 - J[\psi_2])], \tag{2}$$

donde $J[\psi_1]$ es el valor esperado de $\psi_1$, conocido también como la media de $\psi_1$.

*Desviación típica*: es el promedio o variación esperada con respecto a la media aritmética. También se puede decir que es una medida de dispersión usada en estadística que nos dice cuánto tienden a alejarse los valores concretos del promedio en una distribución. La desviación típica de un conjunto $T$ de datos cuantitativos se calcula en la ecuación 3.

$$\sigma = \sqrt{\frac{1}{\mid T \mid -1} \sum_{i=1}^{|T|} (\overline{b_i} - \overline{T})^2}, \tag{3}$$

donde $\overline{b_i} \in T$ y $\overline{T}$ es la media aritmética de los valores de $T$.

*Distribución condicional*: Suponiendo que $b_1$ y $b_2$ son variables aleatorias continuas. La distribución condicional de $b_1$ dado $b_2$ define como se distribuyen las probabilidades de los valores de $b_1$ una vez que se conoce el valor que ha tomado $b_2$. Dado un espacio muestral $\aleph$, dos eventos $D_1, D_2 \in \aleph$, donde $P(D_2) > 0$, la probabilidad condicional de $D_1$ dado $D_2$ se define en la ecuación 4.

$$P(D_1 \mid D_2) = \frac{P(D_1 \cap D_2)}{P(D_2)}. \tag{4}$$

*Función gaussiana*: en estadística y teoría de probabilidades, las funciones gaussianas constituyen la función de densidad de la distribución normal, la cual es una distribución de probabilidad límite de sumas complicadas. La función gaussiana se define en la ecuación 5.

$$f(\underline{b}) = \psi e^{-\frac{(\underline{b}-\psi_1)^2}{2\psi_2^2}}, \tag{5}$$

donde $\psi, \psi_1$ y $\psi_2$ corresponden a constantes reales ($\psi > 0$). La gráfica de la función es simétrica y su forma es de campana, por esta razón se denomina generalmente campana de Gauss. La variable $\psi$ representa el alto de la campana que se centra en el punto $\psi_1$, estableciendo $\psi_2$ la anchura de la misma (figura 1).

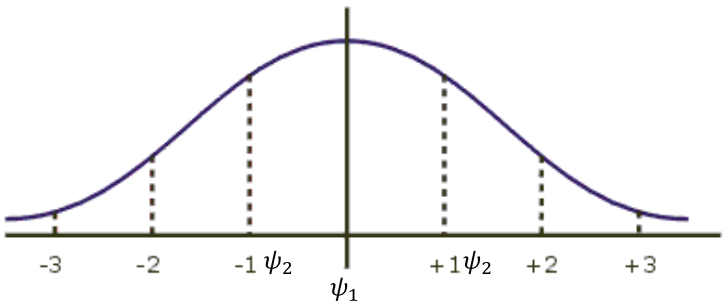


**Fig. 1.** Gráfica de campana de Gauss.

Por esta razón es muy útil para realizar cálculos de estadística. Es correspondiente en el siguiente caso:

$$\psi = \frac{1}{\psi_2\sqrt{2\pi}}, \tag{6}$$

es decir si $\psi$ es igual a la función de densidad de una variable aleatoria con distribución normal de media $\psi_1$ y desviación típica $\psi_2$.

*Función sigmoide*: se trata de una función continua no lineal que posee un rango comprendido entre cero y uno. Cuando es aplicada en las unidades de proceso de una red neuronal artificial significa que, sea cual sea la entrada, la salida estará comprendida entre cero y uno. La función sigmoide viene definida por la ecuación 7.

$$P(\psi) = \frac{1}{1 + e^{-\psi}}. \tag{7}$$

Además tiene una primera derivada no negativa (ver ecuación 8).

$$P(\psi)' = \frac{dP(\psi)}{d\psi}. \tag{8}$$

La gráfica de la función tiene una típica forma de "S" (figura 2).

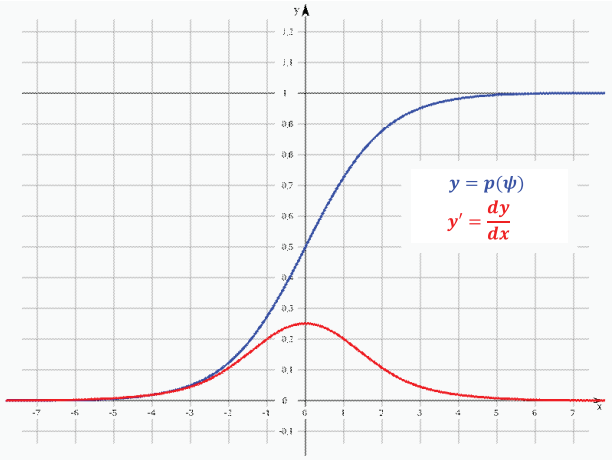


**Fig. 2.** Gráfica de función sigmoide.

*Grafo bipartito*: es un grafo $G = (U, E)$ cuyos vértices se pueden separar en dos conjuntos disjuntos $U_1$ y $U_2$, donde $U_1 \cup U_2 = U$, y $U_1 \cap U_2 = \emptyset$. Las aristas sólo pueden conectar vértices de un conjunto con vértices del otro.

*Matriz de correlación*: es una tabla de doble entrada que muestra una lista multivariable horizontalmente y la misma lista verticalmente y con el correspondiente coeficiente de correlación.

*Máxima correlación*: indica el mayor valor de correlación obtenido por un elemento en un conjunto. Por ejemplo, dado un conjunto de reglas $A$, la máxima correlación de una regla $R$ en el conjunto $A$ se define en la ecuación 9.

$$\overline{r}(R) = \max_{R_i \in A}\{r(R, R_i)\}. \tag{9}$$

*Producto escalar*: el resultado de esta operación es un número o escalar que resulta de sumar las multiplicaciones de las dimensiones de dos vectores. La función para calcular el producto escalar entre los vectores $v_1$ y $v_2$ de dimensión $B$ se define en la ecuación 10.

$$v_1 \cdot v_2 = \sum_{i=1}^{B} v_{1,i} \cdot v_{2,i}, \tag{10}$$

donde el vector $v_1 = (v_{1,1}, v_{1,2}, ...v_{1,B})$ y el vector $v_2 = (v_{2,1}, v_{2,2}, ...v_{2,B})$.

*Red neuronal artificial*: se compone de unidades llamadas neuronas. Cada neurona recibe una serie de entradas a través de interconexiones y emite una salida [6]. La salida viene dada por tres funciones que se explican a continuación [7]. La función de propagación generalmente consiste en la sumatoria de cada entrada multiplicada por el peso de su interconexión. La función de activación que modifica a la función de propagación, y puede no existir, siendo en este caso la salida la misma función de propagación. La función de transferencia se aplica al valor devuelto por la función de activación. Además se utiliza para acotar la salida de la neurona y generalmente viene dada por la interpretación que se le quiera dar a dichas salidas [8].

## Referencias bibliográficas del anexo

1. Jolliffe, I.: Principal Component Analysis. New York, USA, 2nd edition, Springer (2002) 489 p.
2. Llorena, J.M.: Razonamiento bajo incertidumbre. [citado 4 de octubre de 2013]. Disponible en Internet: http://arantxa.ii.uam.es/ jmoreno/razonamiento/tevidencia.htm
3. Augustin, T.: Generalized basic probability assignments. International Journal of General Systems **34**(4) (August 2005) 451–463
4. Tazid, A., Palash, D.: Methods to obtain basic probability assignment in evidence theory. International Journal of Computer Applications **38**(4) (January 2012) 46–51
5. Gustavo, R.: Correlación entre variables. [citado 4 de octubre de 2013]. Disponible en Internet: http://viref.udea.edu.co/contenido/menu_alterno/apuntes/ac36-correlacion-variables.pdf
6. Zhang, G., Patuwo, B.E., Hu, M.Y.: Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting **14**(1) (1998) 35 – 62
7. Hagan, M.T., Demuth, H.B., Beale, M.: Neural network design. Boston, MA, USA, PWS (1996)
8. White, H.: Artificial Neural Networks: Approximation and Learning Theory. Cambridge, MA, USA, Blackwell (1992)

RT_023, febrero 2014
Aprobado por el Consejo Científico CENATAV

**Editor:** Lic. Lucía González Bayona
**Diseño de Portada:** Di. Alejandro Pérez Abraham
RNPS No. 2143
ISSN 2072-6260
**Indicaciones para los Autores:**
Seguir la plantilla que aparece en www.cenatav.co.cu
C E N A T A V
7ma. No. 21812 e/218 y 222, Rpto. Siboney, Playa;
La Habana. Cuba. C.P. 12200
*Impreso en Cuba*