DENIS MARCEL FERNANDES

# Detecção de fraudes em transações bancárias online via dispositivos móveis

São Paulo
2014

**DENIS MARCEL FERNANDES**

# Detecção de fraudes em transações bancárias online via dispositivos móveis

Monografia apresentada ao PECE – Programa de Educação Continuada da Escola Politécnica da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Especialista em Tecnologia da Informação.

Orientador: Prof. Dr. Stephan Kovach.

Versão corrigida de acordo com as Diretrizes 2013

São Paulo

2014

Dedico esta monografia aos meus pais, ao meu irmão, à minha família e aos meus amigos, pelo apoio e compreensão durante a realização desta monografia

# RESUMO

As transações bancárias online estão se tornando o principal meio de se realizar uma transação financeira. Este aumento é notável em razão ao crescimento de utilização dos *smartphones* (dispositivos móveis) e do acesso a redes móveis.

O aumento da utilização de transações *online* proporciona também que indivíduos desonestos (fraudadores) comecem a direcionar suas atenções nessa nova área em expansão.

A detecção de fraudes em transações bancária *online* tenta verificar se uma transação é legítima ou fraudulenta. Para essa detecção é preciso aplicar um método estatístico que irá comparar a transação atual com a base de dados daquele usuário. Se a transação atual for muito distinta dos dados armazenados é possível que esta seja uma fraude.

Este trabalho pretende apresentar um modelo de arquitetura para detectar fraudes em transações bancárias *online* efetuadas a partir de um dispositivo móvel. Esse modelo utiliza como um dos atributos de detecção, a localização do dispositivo móvel onde é efetuada a transação. Acredita-se que com uso deste atributo será possível aumentar a taxa de acertos com relação aos métodos convencionais de detecção de fraudes.



**Palavras-Chave:** Fraude. Detecção de Fraude. Transação Bancária *Online*. Localização. Dispositivo Móvel.

# ABSTRACT

The on-line banking are becoming the main means of conducting financial transactions. This increase is remarkable due to the growing use of smart phones (mobile devices) and access to mobile networks.

Increasing use of on-line transactions also provides that dishonest people (fraudsters) begin to direct your attention in this new growing area.

Fraud detection in on-line banking transactions attempts to verify that a transaction is legitimate or fraudulent. For this detection is necessary to apply a statistical method that will compare the current transaction with that user database. If the current transaction is very distinct from the stored data is possible that this is a fraud.

This work intends to present an architectural model to detect fraud in online banking transactions from a mobile device. This model uses as an attribute of the detection, the location of the mobile device, where the transaction is performed. It is believed that with the use of this attribute is possible to increase the hit rate over conventional fraud detection methods.



**Keywords:** Fraud. Fraud Detection. On-line Banking Transaction. Location. Mobile Device.

# LISTA DE ILUSTRAÇÕES

# LISTA DE TABELAS

# LISTA DE ABREVIATURAS E SIGLAS

| | |
|---|---|
| **GPS** | Global Positioning System |
| **IP** | Internet Protocol |
| **PA** | Perfil Atual |
| **PH** | Perfil Histórico |

# SUMÁRIO

# 1. INTRODUÇÃO

## 1.1. Considerações iniciais

A segurança em todos os meios de trabalho é uma das mais importantes áreas a serem priorizadas e frequentemente aprimoradas. Na tecnologia da informação isso não é diferente. Ao se tratar de instituições financeiras, a segurança se torna um dos alicerceis desta instituição.

Uma instituição para adquirir, manter e não perder clientes deve primeiramente, definir uma infraestrutura de segurança confiável ao seu consumidor. Para isso a instituição deve comprovar que essa infraestrutura é segura.

A prevenção da fraude consiste em aplicar algumas medidas que impedem o usuário de realizar uma determinada fraude (BOLTON, HAND, 2002). Em uma transação online, um bom exemplo de prevenção de fraude seria a senha que cada usuário possui. Esta senha é para ser única e por isso é uma prevenção para possíveis fraudadores.

Os mecanismos de segurança evoluíram consideravelmente. Mesmo com esse avanço é quase impossível afirmar que uma empresa que se utiliza da Tecnologia da Informação esteja segura somente com prevenção de fraudes.

A engenharia social são técnicas aplicadas em grandes empresas para buscar falhas em sistema de segurança da informação. A aplicação benéfica da engenharia social se aplica em aperfeiçoar o sistema de segurança da informação. As práticas benéficas utilizadas com frequência são investimentos em treinamentos e conscientização dos funcionários com relação a engenharia social. A aplicação maliciosa tem como objetivo explorar falhas para a equisão de dados sigilosos (ALVES, 2010).

De acordo com Kovach (2011), o problema maior está na engenharia social aplicada maliciosamente. Uma prática de engenharia social maliciosa são de interações entre pessoas, em que uma mal intencionada consegue obter informações sigilosas de uma outra sem que esta saiba. Atualmente, o método mais comum da engenharia social é o *phishing*. O *phishing* é mais utilizado em e-mails através de *spam*. O *spam*

é um termo utilizado para enviar e-mails para um grande número de pessoas sem que estas tenham sidas solicitadas.

O sistema não tem ciência de quando a prevenção pode falhar, por isso a detecção de fraude deve ser aplicada continuamente (BOLTON, HAND, 2002). Os mecanismos de detecção utilizam algoritmos de *data mining.*

Como exemplo de uma transação financeira pode-se citar o evento de um pagamento de boleto bancário ou de depósito em outra conta. Hoje em dia, uma transação financeira online em tempo real pode ocorrer tanto via Internet através de computadores pessoais quanto pelos *smartphones (mobile banking).* As transações em tempo real estão crescendo a cada dia e se consolidando como o principal meio de transações realizadas. Segundo a Folha de São Paulo de Dezembro de 2013 uma pesquisa feita com as 5 maiores instituições financeiras que atuam no Brasil, cerca de 51% das operações bancárias registradas de Janeiro a Junho de 2013 foram realizadas pela Internet incluindo o *mobile banking.*

Com a evolução da tecnologia, os dispositivos móveis (*smartphones, tablets)* estão, a cada ano, recebendo novas funcionalidades. Uma das características que a maioria dos dispositivos móveis apresentam hoje é a da possibilidade de se utilizar a sua localização como referência para inúmeras aplicações. Esta característica pode ser utilizada para reforçar a capacidade de detecção de fraudes em transações bancárias realizadas através de *mobile banking.*

## 1.2. Objetivo

O objetivo desse trabalho é apresentar uma arquitetura de detecção de fraude em transações bancárias *online* (em tempo real) via dispositivos móveis (*mobile banking*), utilizando a sua localização como um dos atributos para aumentar o desempenho de detecção.

## 1.3. Justificativa

Com o crescimento dos serviços *online* (*e-commerce, mobile banking* entre outros) surgiu a necessidade de se processar uma grande quantidade de informações requisitadas pelos usuários e ainda verificar se este não é um fraudador. Uma das

razões desse crescimento é devido ao fácil acesso as informações de qualquer lugar pelos dispositivos móveis.

Em transações financeiras via internet, tem se o desafio da detecção em tempo real, ou seja, é necessário detectar se uma transação é fraudulenta ou não antes que esta termina, pois os custos são elevados ao se tratar de fraudes em instituições financeiras (EDGE, SAMPAIO, 2009).

Segundo Fawcett e Provost (1997) existem dois aspectos importantes em detecção de fraudes. O primeiro é que uma ação para um determinado usuário que via ser fraudulenta e para outro usuário pode não ser. O segundo problema é que um usuário legítimo pode ser apontando erroneamente como um fraudador.

Deste modo, é necessário um sistema que detecte se um usuário é legitimo ou não com uma grande taxa de acerto.

Este trabalho é baseado principalmente em dois artigos. O primeiro é o de Kovach (2011), neste artigo o autor utiliza a combinação de dois tipos de análise para obter um melhor desempenho na detecção da fraudes. No segundo é um estudo de cenário real realizados por Fawcett e Provost (1997). Neste os autores utilizam a localização das chamadas como um dos atributos para evidenciar a clonagem de celulares.

Existem vários trabalhos publicados sobre detecção de fraudes em cartões de créditos, telefones celulares e redes de computadores, mas poucos na área de transação financeira.

## 1.4. Estrutura do trabalho

A monografia está organizada como segue:

Este capítulo contém a introdução do tema abordado, o objetivo específico da monografia e a motivação para escrever esta monografia.

O capítulo dois explica os aspectos teóricos em que esta monografia está embasada. Neste capítulo contém métodos estatísticos para aprendizagem de padrões e trabalhos correlatos sobre detecções de fraudes em dispositivos móveis.

O capítulo três apresenta uma arquitetura para a detecção de fraudes em transações bancárias *online* que utiliza a informação de localização derivado dos dispositivos móveis como um atributo para melhorar o desempenho do sistema.

O trabalho é finalizado com as considerações finais no capítulo quatro.

# 2. FUNDAMENTAÇÃO TEÓRICA

## 2.1. Prevenção e detecção de fraudes

Fraude é definida no escopo deste trabalho como sendo qualquer acesso não autorizado ou uma transação não autorizada efetuada em uma conta corrente através da Internet.

Prevenção de fraudes consiste de medidas de segurança para evitar que elementos não autorizados efetuem transações em contas não autorizadas (KOVACH, 2011).

A prevenção é feita normalmente durante a fase de *autenticação do usuário* por meio de senhas, frases secretas, dispositivos de geração de códigos secretos *(tokens)*, entre outros.

Entretanto, os mecanismos para prevenção de fraudes disponíveis para aplicações bancárias podem falhar, pois eles não protegem contra "as falhas de segurança dos humanos".

A engenharia social é um dos meios mais utilizados para a obtenção de informações sigilosas e importantes, especialmente dos usuários domésticos.

Entre as abordagens de engenharia social usadas pelos fraudadores, *phishing* através de *e-mails* é uma das formas mais comuns.

Detecção de fraudes consiste em identificar atividades não autorizadas após a falha na prevenção de fraudes. Na prática, detecção de fraudes é aplicada constantemente, pois o sistema não tem ciência de quando a prevenção falha (BOLTON, HAND, 2002).

## 2.2. Métodos de Detecção de Fraudes

Vários métodos são propostos para detecção de fraudes, sendo uns mais adequados que os outros dependendo do domínio de cada aplicação, como de cartões de crédito, de intrusão de computadores, de telefonia móvel, assim como de transações financeiras *online*.

Como exemplos de métodos de detecção, existem os métodos supervisionados baseados em regras de decisão e redes neurais e os métodos não supervisionados baseados em modelos estatísticos que detectam evidências de fraude através de desvios do comportamento normal de usuário.

### 2.2.1. Métodos de detecção supervisionada e não supervisionada

Os métodos de detecção de fraudes podem ser classificados como supervisionados e não supervisionados.

Métodos supervisionados são aqueles em que amostras de comportamentos normais e fraudulentos são usadas para construir modelos que permitem o sistema classificar as novas observações em uma destas duas classes.

Exemplos de métodos supervisionados são aqueles baseados em regras de decisão e em redes neurais artificiais

Uma característica do método supervisionado é que ele é capaz de identificar apenas atividades fraudulentas conhecidas (HILAS, SAHALOS, 2005) e (CORTES, PREGIBON, 2001).

Os métodos não supervisionados apenas procuram observações que são diferentes do comportamento usual. São baseados em métodos estatísticos.

### 2.2.2. Métodos estatísticos

Métodos estatísticos utilizam métricas e modelos estatísticos para determinar as variações de comportamento dos usuários.

Segundo (KOVACH, 2011), existem duas abordagens baseadas em métodos estatísticos:

- Análise absoluta; e
- Análise diferencial

Na abordagem baseada em análise absoluta, a detecção é feita por meio de algum critério de comparação de um ou mais campos de uma transação com valores fixos preestabelecidos, denominados limiares (ou *thresholds*).

A análise absoluta é útil para detectar atividades fraudulentas extremas, como número elevado de erros de senha.

Na abordagem baseada em análise diferencial, o padrão comportamental dos acessos as contas bancárias são monitorados, comparando suas atividades mais recentes com o histórico de sua utilização. Alarmes são gerados quando o padrão de utilização muda de forma significativamente em um curto período de tempo.

Quando uma nova atividade for observada, o sistema gera um escore que determina o grau da anormalidade do comportamento desta observação. Este escore é gerado como resultado da comparação do perfil da atividade observada com o perfil de comportamento histórico.

Caso não haja anormalidade, o perfil da atividade observada é normalmente fundido com o perfil anterior para se adaptar às variações do comportamento de um usuário com o tempo.

A vantagem de utilizar métodos estatísticos é que eles são baseados em teorias bem conhecidas.

Como algumas das desvantagens podemos citar as seguintes:

- Os detectores podem ser treinados gradualmente até que eles passam a considerar um comportamento anormal como normal;
- A determinação do limiar para que um comportamento seja considerado anormal é difícil de ser estabelecida.

## 2.3. Combinador de Dempster-Shafer

Dentre as possíveis opções, este trabalho optou pela utilização da teoria de Dempster-Shafer (DS) para combinar evidências de fraude estimadas por dois

métodos de detecção, a mesma utilizada na arquitetura proposta por (KOVACH, 2011).

A teoria de Dempster-Shafer é uma teoria matemática de evidências que fornece um arcabouço para combinar fontes de evidências (KOVACH, RUGGIERO, 2011).

Resumidamente, a principal diferença entre a teoria de Dempster-Shafer e a teoria de probabilidade é que a primeira permite representar a incerteza explicitamente. Na teoria de probabilidades, a probabilidade deve ser distribuída igualmente mesmo havendo incerteza. Por exemplo, se não existir nenhum conhecimento a priori, deve-se assumir probabilidade P igual a 1/N para cada uma das N possibilidade pelo *princípio da indiferença*.

A combinação de Dempster-Shafer é feita por meio de uma função que calcula um escore final a partir de duas evidências. Dadas duas probabilidades de evidências $m_1(h)$ e $m_2(h)$, elas podem ser combinadas em uma terceira probabilidade, $m_3(h)$. A expressão é apresentada na Figura 1:

$$m_3(h) = m_1(h) \oplus m_2(h) = \frac{\sum_{x \cap y = h} m_1(x) \cdot m_2(y)}{1 - \sum_{x \cap y = \varnothing} m_1(x) \cdot m_2(y)}$$

Figura 1 – Expressão para o cálculo da combinação de Dempster.

Onde, o símbolo ⊕ representa soma ortogonal.

A expressão da figura 1 é apresentada apenas a título de ilustração. Os detalhes desta teoria poderão ser encontrados em Kovach (2011).

## 2.4. Métrica de Desempenho

As métricas, normalmente utilizadas para avaliar o desempenho de detector de fraudes são as seguintes (FAWCETT, 2006):

- Taxa do Verdadeiro Positivo (*Tvp*) – É a fração de transações fraudulentas que o detector conseguiu classificar como fraudulentas.

- Taxa do Falso Positivo (*Tfp*) – É a fração de transações normais que o detector classificou erroneamente como fraudulentas.

- Taxa de Verdadeiro Negativo (*Tvn*) – É a fração de transações normais que o detector classificou corretamente como normais.

- Taxa do Falso Negativo (*Tfn*) – É a fração de transações fraudulentas que o detector classificou erroneamente como normais.

- Precisão (*Pr*) – É a fração de transações fraudulentas que foram classificadas corretamente.

- Exatidão (*Ex*) – É a fração do número total de transações (normais e fraudulentas) classificadas corretamente.

### 2.4.1. Matriz de Confusão

As métricas de desempenho podem ser derivadas a partir de uma tabela conhecida como *matriz de confusão ou tabela de contingência* (FAWCETT, 2006).

Considerando que um detector de fraudes é um classificador de duas classes, P (Positiva ou Fraude) e N (Negativa ou Legítima), existem quatro possíveis resultados ao classificar um conjunto de transações. Estes quatro resultados podem ser representados em uma matriz conhecida como matriz de confusão, como mostra a figura 2.

**Classificação Correta**

| **Resultado da Detecção de Fraude** | | | |
|---|---|---|---|
| | Verdadeiro Positivo (VP) | Falso Positivo (FP) | P |
| | Falso Negativo (FN) | Verdadeiro Negativo (VN) | N |
| | P | N | |

Figura 2 – Matriz de Confusão.

Onde,

- *VP* = O número de positivos (fraudes) classificados corretamente;

- *FP* = O número de negativos (legítimos) classificados incorretamente;

- *VN* = O número de negativos (legítimos) classificados corretamente;
- *FN* = O número de positivos (fraudes) classificados incorretamente;

As classificações dispostas na diagonal principal (verdadeiro positivo e verdadeiro negativo) da matriz de confusão são as classificações corretas. Os demais campos significam classificações erradas.

Com base nestes campos, as métricas de desempenho são derivadas como segue:

- *Taxa de verdadeiro positivo: Tvp = VP / P*
- *Taxa de falso positivo: Tfp = FP / N*
- *Taxa de verdadeiro negativo: Tvn = VN / N*
- *Taxa de falso negativo: Tfn = FN / P*
- *Exatidão: Ex = VP+VN (P+N)*

Onde,

*P = VP + FN* é igual ao número total de positivos (fraudes);

*N = VN + FP* é igual ao número total de negativos (legítimos).

Vale lembrar que, um dos objetivos mais importantes de um sistema de detecção de fraudes é identificar transações fraudulentas com menor número de alarmes falsos.

Uma transação legítima que é sinalizada como uma fraude caracteriza um alarme falso (ou falso positivo).

No caso de transações financeiras, por exemplo, o custo de não detectar uma fraude (falso negativo) pode ser bem alta. Por outro lado, disparar alarmes mediante qualquer suspeita pode gerar uma taxa elevada de falsos alarmes (falsos positivos), o que pode gerar insatisfação aos clientes legítimos.

## 2.5. Detecção de Fraudes em Clonagem de Celulares

A fraude em celulares resulta em grande quantidade de despesas para as empresas de telecomunicações. Um tipo de fraude em celulares é a clonagem, e isso ocorre quando um usuário não legítimo consegue fazer ligações utilizando a conta de um usuário legítimo.

As clonagens em celulares ocorrem quando o fraudador consegue o número de identificação e o número de série do celular de um usuário legítimo.

Em (FAWCETT, PROVOST, 1997), o modelo de detecção de fraudes utilizado consiste de um método supervisionado baseado em regras de decisão para gerar perfis de utilização dos usuários e em redes neurais para combinação dos perfis.

Neste modelo, os atributos usados para indicar uma evidência de fraude em clonagem de celulares são: Dia da Semana, Hora inicial da ligação, Duração da chamada, Origem da chamada e Destino da chamada.

Com base nos históricos de cada chamada é gerado um conjunto de regras que caracterizam os perfis dos usuários. Cada regra é implementada em um monitor. Portanto, cada usuário é caracterizado por um conjunto específico de monitores. Ao chegar uma nova chamada, cada um dos monitores emite um escore que é combinado por meio de redes neurais para gerar uma evidência de fraude.

# 3. PROPOSTA DE UMA ARQUITETURA PARA DETECÇÃO DE FRAUDES EM TRANSAÇÕES BANCÁRIAS UTILIZANDO DISPOSITIVOS MÓVEIS

## 3.1. Considerações Iniciais

Neste capítulo será apresentando os atributos mais significativos de transações bancárias *online* utilizando dispositivos móveis. Com estes atributos será proposta uma arquitetura de detecção de fraudes em tempo real baseando-se no modelo proposto em (KOVACH, 2011), porém modificado, com a inclusão de um atributo de localização obtido dos dispositivos móveis. O uso deste atributo seguirá o modelo aplicado por (FAWCETT, PROVOST, 1997) para obter um desempenho maior na detecção de fraudes comparado com os métodos atuais. Com o atributo de localização obtido nos dispositivos móveis e alguns atributos pertencentes a transações bancárias *online*, como por exemplo, a data e a hora, será possível calcular a distância e o tempo de locomoção de um usuário. Ao se calcular a distância e o tempo de locomoção a partir da localização da última transação efetuada, com esse cálculo é possível de verificar se as informações são coerentes.

## 3.2. Atributos Associados a Transações Bancárias Online

Qualquer que seja o método utilizado para a detecção de fraudes deve-se antes de tudo definir os atributos mais significativos para esta ação. Para a escolha correta dos atributos é necessário a definição do domínio da sua aplicação. O domínio de aplicação é de transações bancárias realizadas por usuários em uma única conta corrente. Portanto, os atributos escolhidos são centrados no usuário que acessa uma única conta. O método de detecção de fraude adotado é baseado em análise diferencial.

Na abordagem baseada em análise diferencial, os padrões são obtidos através da monitoração das transações na conta corrente do usuário. Esta monitoração é comparada com um histórico da utilização da conta corrente que representa o comportamento normal do usuário. Caso haja um desvio significativo em relação ao comportamento normal é acionado um alerta de possível fraude.

Tratando-se de transações bancárias online, existem vários atributos relacionados. A figura 3 ilustra alguns atributos mais significativos associados a uma transação bancária *online*.

| Agência | Conta | Data | Hora | Identidade do Dispositivo | Endereço IP | Tipo de Transação | Valor |
|---|---|---|---|---|---|---|---|

Figura 3 – Exemplo de alguns atributos nas mensagens de transação bancária *online* (KOVACH, 2011).

### 3.3. Atributos Baseados em Características de Dispositivos Móveis

Os dispositivos móveis atualmente são embarcados com várias funcionalidades interessantes. Em Fawcett e Provost (1997) um dos atributos utilizados na detecção de clonagens de celulares, é o de "Local de Origem" da chamada. Este atributo está contido nas informações geradas em chamadas telefônicas. Em transações bancárias *online* feitas por meio de dispositivos móveis, está informação pode ser obtida através do *Global Positioning System* (GPS), ou de uma Rede sem fio (*Wi-Fi*). A Rede Sem Fio não fornece uma posição muito precisa quanto a de um GPS, mas a utilização desta localização em detecção de fraudes poderá ser satisfatória.

Segundo notícia publicada no Site Terra em Abril de 2011 (Referência ao jornal *The Guardian*, Reino Unido), a empresa Google ao mapear as ruas da cidade para seu aplicativo *Street View* coletou informações das redes sem fio, por essa razão é possível adquirir uma localização sem muita precisão através de redes sem fio. Além disso, é possível descobrir uma localização aproximada com o endereço do IP que são gerados ao acesso dos usuários aos provedores, segundo o Site GeoIPTool o grau de precisões das informações são de: País – 99,8% de precisão, Região/Estado – 90% de precisão, Cidade 70% de precisão e Latitude e Longitude 52% de precisão.

A figura 4 apresenta os atributos associados a transação bancária online e o acréscimo do atributo de localização considerados no modelo de arquitetura proposto.

| Agência | Conta | Data | Hora | Identidade do Dispositivo | Endereço IP | Tipo de Transação | Valor | Local de Origem |
|---|---|---|---|---|---|---|---|---|

Figura 4 – Exemplo de atributos nas transações bancárias *online* utilizando dispositivos móveis.

Abaixo estão os significados de cada atributo:

- **Agência -** Número da agência acessada.
- **Conta -** Número da conta acessada.
- **Data -** O dia, o mês e o ano em que esta transação está sendo acessada.
- **Hora -** A hora em que foi acessada esta transação.
- **Identidade do Dispositivo -** Identidade do dispositivo, se não possuir este campo é deixado em branco.
- **Endereço IP -** Neste campo é o endereço de IP que está sendo utilizado.
- **Tipo de Transação -** Indica o tipo de transação que está sendo executada. (Pagamento ou Transferência).
- **Valor da Transação -** Este atributo apresenta o valor da transação.
- **Local de Origem -** Constam neste atributo, a cidade, o estado e o país em que foi realizada a transação, derivados das informações obtidas a partir de endereço IP no caso de Wi-Fi ou de coordenadas no caso de GPS.

Com a utilização dos atributos Data, Hora e Local de Origem é possível verificar se uma transação é uma fraude comparando-a com a transação anterior e verificando o tempo e o local da transação. Na tabela 1 são apresentados alguns dados que servem como exemplo para análises.

| Data | Hora | País | Estado | Cidade |
|---|---|---|---|---|
| 21/09/2014 | 21:30 | Brasil | São Paulo | São Paulo |
| 21/09/2014 | 23:30 | EUA | Califórnia | *Los Angeles* |

Tabela 1 – Exemplo de informações em uma transação bancária *online* via dispositivos móveis fraudulenta.

A distância entre São Paulo Brasil e *Los Angeles* Estados Unidos é de aproximadamente 9.917 km. A diferença de tempo entre as duas transações foi de 2 horas. Utilizando a fórmula para cálculo de velocidade média, apresentado na Figura 5 abaixo, pode-se calcular a velocidade média necessária para ir de São Paulo a *Los Angeles.*

$$v_m = \frac{\Delta s}{\Delta t}$$

Figura 5 – Fórmula para cálculo da velocidade média.

A velocidade média é calculada a partir da distância percorrida (Δs) sobre o tempo que levou para percorrer esta distância (Δt). Por exemplo, a distância percorrida sendo de 9.917 km e o tempo de 2 horas, a velocidade média é de 4958,5 km/h. Esta velocidade para os padrões atuais é impossível de ser alcançada para a locomoção de pessoas.

## 3.4. Modelo Proposto da Arquitetura de Detecção de Fraudes Utilizando Dispositivo Móvel

O modelo de arquitetura proposto é composto de duas vertentes. A primeira se refere à transação bancária e a segunda, as características do dispositivo móvel. As duas vertentes utilizam-se das informações armazenadas (Histórico) para a verificação da transação atual. Esta verificação resulta em um indicio ou não de uma fraude.

Na vertente de transações bancárias, definida como Monitor de Transação Bancária, utiliza um método estatístico baseado na diferença do perfil observado de uma transação (definido aqui como perfil de comportamento atual de um usuário) com a

média histórica das transações (definida aqui como perfil de comportamento normal ou histórico de um usuário). A segunda vertente, definida como Monitor de Dispositivo Móvel, também utiliza o mesmo método estatístico da primeira vertente.

A Figura 6 apresenta a Arquitetura para Detecção de Fraudes Utilizando Dispositivo Móvel.

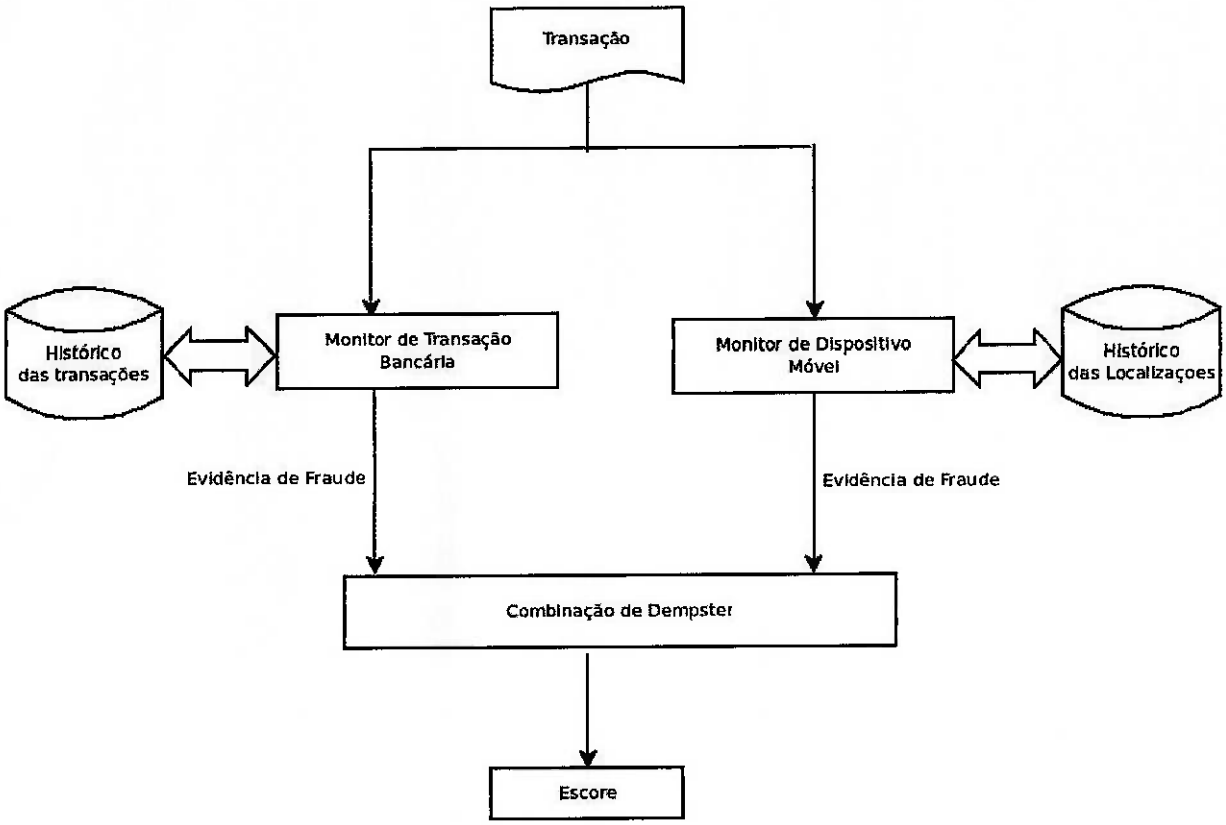


Figura 6 – Arquitetura para detecção de fraudes.

Ao realizar uma transação é acionado em paralelo o Monitor de Transação Bancária e o Monitor de Dispositivo Móvel. Os monitores fazem a comparação do perfil de comportamento atual da transação com o perfil normal do usuário utilizando o método estatístico definido. As evidências (escores) produzidas por cada análise, tanto no Monitor de Transação Bancária como no Monitor de Dispositivo Móvel são combinadas para produzir um valor final. Esse valor é comparado com um determinado limiar pré-definido. Se ele estiver acima desse limiar, um alarme é disparado. A combinação dos dois monitores será realizada pelo método de Dempster.

Em transações bancária *online* é necessário um tempo de resposta imediato. Como o serviço é em tempo real o método de detecção de fraude deve analisar as informações de forma ágil. No modelo de arquitetura proposto foi utilizado os monitores em paralelo para se obter o menor tempo possível de resposta da verificação das informações contidas na transação.

O detalhamento do Monitor de Transação Bancária e o Monitor de Dispositivo Móvel serão apresentados em sequência.

### 3.5. Monitor de Transação Bancária

O Perfil Atual (PA) são os atributos utilizados no Monitor de Transação Bancária obtidos a partir dos atributos da transação atual. A Figura 7 apresenta o Perfil Atual.

**Perfil Atual**

| Agência | Conta | Data | Hora | Identidade do Dispositivo | Endereço IP | Tipo de Transação | Valor |
|---|---|---|---|---|---|---|---|
| pa(0) | pa(1) | pa(2) | pa(3) | pa(4) | pa(5) | pa(6) | pa(7) |

Figura 7 – Atributos do Perfil Atual de Transação Bancária.

O Perfil Histórico consiste de atributos obtidos a partir do valor médio dos atributos das transações do usuário ao longo do tempo. Basicamente, a comparação entre o perfil atual (PA) e o perfil histórico (PH) é realizada por meio de comparação de cada um dos atributos individualmente. Se o Perfil Atual desviar muito do comportamento médio do Perfil Histórico pode ser uma indicação de fraude.

O Perfil Histórico pode ter mais atributos que o Perfil Atual. Por exemplo, se pa(2) for a data da transação atual e ph(2) for a data da última transação, poderá ter um atributo, ph(x), indicando o tempo médio entre as transações realizadas pelo usuário.

A Figura 8 define o Perfil Histórico tendo n atributos, este número de atributos é pré-definido para ser utilizado no método estatístico.

$$PH = \{ ph(0), ph(1), ph(2), \ldots, ph(n-1) \}$$

Figura 8 – Atributos do Perfil Histórico do Monitor de Transação Bancária.

A Figura 9 mostra que, ao chegar uma Transação Atual (Perfil Atual), ela é comparada com a média das transações anteriores (Perfil Histórico) aplicando um método estatístico apropriado para cada atributo. Após essa comparação, a evidência de fraude (escore) é obtida.

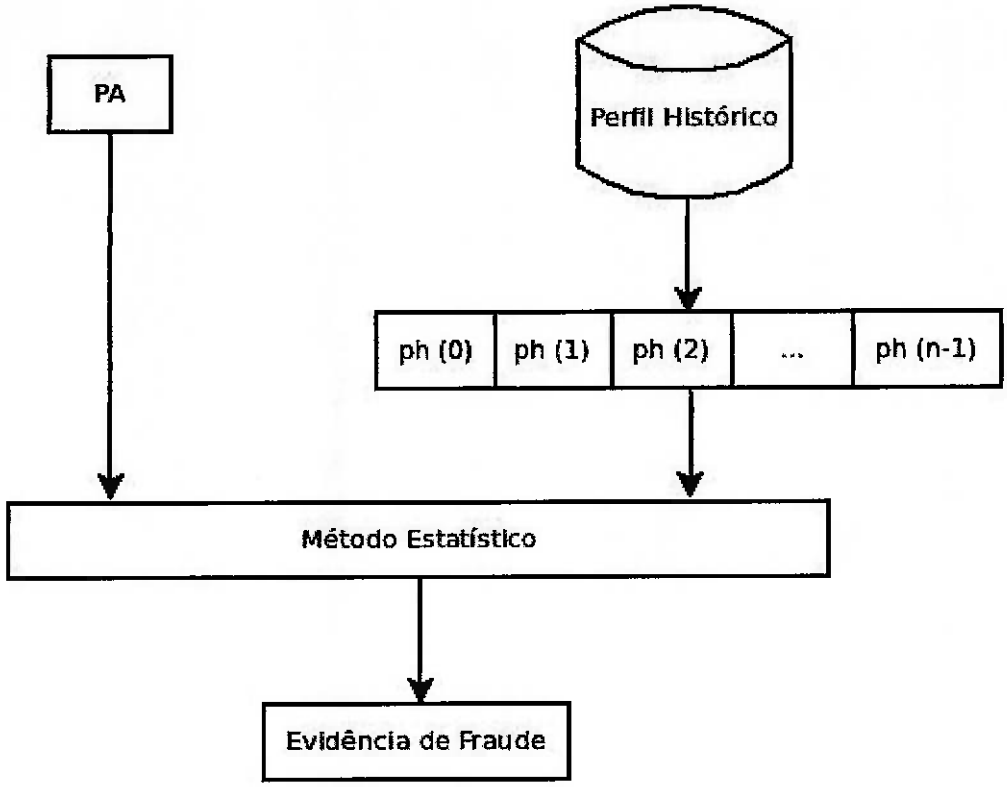


Figura 9 – Verificação da evidência de fraude (escore) no Monitor de Transação Bancária.

### 3.6. Monitor de Dispositivo Móvel

O Monitor de Dispositivo Móvel utiliza o mesmo método estatístico do Monitor de Transação Bancária.

O Perfil Atual se refere aos atributos mais significativos do Dispositivo Móvel. De acordo com Fawcett e Provost (1997), os principais atributos estão apresentados na Figura 10.

| Perfil Atual | |
|---|---|
| Data e Hora da Transação | Local de Origem |
| pa(0) | pa(1) |

Figura 10 – Atributos do Perfil Atual de Dispositivo Móvel.

A Figura 11 apresenta os atributos do Perfil Histórico do Monitor de Dispositivo Móvel.

| Perfil Histórico | | |
|---|---|---|
| Data e Hora da Última Transação | Local de Origem da Última Transação | Limiar da Velocidade |
| ph(0) | ph(1) | ph(2) |

Figura 11 – Atributos do Perfil Histórico do Monitor de Dispositivo Móvel.

Abaixo estão os significados de cada atributo:

- **Data e hora da última transação** – ph(0) armazena a data e a hora da última transação normal.
- **Local de Origem da última transação** – ph(1) é armazenado o local de origem da última transação.
- **Limiar da velocidade** – Em ph(2) é armazenado um limiar pré-definido da velocidade em que um usuário possa atingir. O valor da velocidade média calculado no momento é comparado com esse limiar.

A Figura 12 apresenta a análise do método estatístico referente ao Monitor de Dispositivo Móvel.

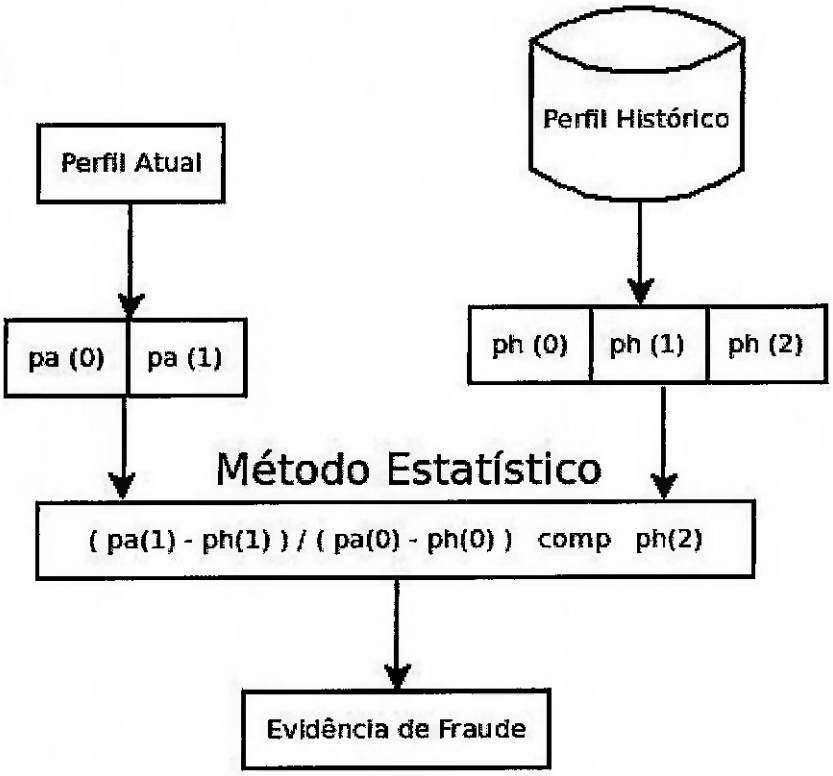


Figura 12 – Verificação da evidência de fraude no Monitor de Dispositivo Móvel.

Na análise, primeiramente é calculada a distância[1] entre pa(1) (Local de Origem) com ph(1) (local de Origem da Última Transação), através da diferença entre as suas coordenadas. Em segundo lugar, sobre esta distância é aplicada a diferença entre pa(0) (Data e Hora da Última Transação) com ph(0) (Data e Hora da Transação Atual), diferença do tempo. O resultado obtido é comparado com ph(2) (Limiar da Velocidade). Ao se aproximar do limiar de ph(2) o grau da fraude vai aumentando até atingir o valor de ph(2) determinado como o valor máximo para evidenciar uma fraude. Posteriormente é gerado uma evidência de fraude (escore) para o Monitor de Dispositivo Móvel.

[1] Embora neste trabalho a distância entre dois pontos seja representada por uma simples diferença de suas coordenadas, na prática, a distância entre duas coordenadas determinadas por GPS é normalmente calculada usando a fórmula de Haversine. Esta fórmula leva em conta a curvatura da Terra. http://pt.wikipedia.org/wiki/Fórrmula_de_Haversine

## 3.7. Combinação dos Resultados Obtidos dos Monitores e Atualização do Perfil Histórico

### 3.7.1. Combinação dos Resultados dos Monitores

Conforme exemplificado no capítulo 2, este trabalho não tem como objetivo o entendimento da teoria de Dempster-Shafer.

A combinação de Dempster é realizada através dos escores de evidências obtidos dos Monitores de Transação Bancária e Dispositivo Móvel. A combinação desses escores resultará em um valor final que será comparado com um limiar pré-definido. Se este valor for superior ao limiar é gerado a indicação de fraude. O valor deste limiar é obtido, normalmente, através de testes.

A combinação das evidências dos monitores distintos possibilita uma redução na taxa de falsos positivos. Segundo Moreau (1999), a combinação de evidências em detecção de fraudes possibilita um desempenho melhor do que somente em abordagens isoladas. Além disso, a análise em paralelo dos monitores possibilita um tempo de resposta eficaz para transações em tempo real.

### 3.7.2. Atualização do Perfil Histórico

A atualização do Perfil Histórico (PH) é realizada somente no final da análise, para as transações caracterizadas como normais. A sua atualização pode ser realizada através do cálculo de média ponderada aplicada para alguns dos atributos da transação (Perfil Atual) e do Perfil Histórico (PH), conforme Kovach (2011).

## 3.8. Análise de Desempenho do Modelo de Arquitetura Proposto

Não existe nenhum experimento prático que possa demonstrar a eficácia do modelo de arquitetura proposto. Entretanto, considerando a ideia proposta em (KOVACH, 2011) onde dois detectores de fraudes diferentes resultaram numa melhoria de desempenho, pode-se supor que o modelo de arquitetura proposto neste trabalho terá um desempenho semelhante com relação a taxa de falsos positivos e verdadeiros positivos. Neste trabalho, o atributo específico utilizado em um dos detectores foi a localização onde a transação é efetuada. Este atributo se baseia no

fato de que duas transações não podem ser efetuadas de lugares cuja distância entre os dois sobre a diferença do tempo das transações seja maior do que um certo valor.

# 4. CONSIDERAÇÕES FINAIS

Neste trabalho foi apresentado um modelo de arquitetura para detecção de fraudes em transações bancárias *online* utilizando dispositivo móvel. O modelo proposto possui duas vertentes na arquitetura: Monitor de Transação Bancária e Monitor de Dispositivo Móvel.

A diferença deste trabalho com os demais é apresentada no Monitor de Dispositivo Móvel, especificamente no atributo de Localização. Com esse atributo pode-se ter mais uma evidência de que uma fraude esteja ocorrendo, desde que duas transações não podem ocorrer em duas localizações muito distantes em um tempo curto.

Com a adição do atributo de localização no método de detector de fraudes é bem provável que diminua a taxa de falsos positivos em transações bancárias *online* que utiliza dispositivos móveis.

Como trabalho futuro, seria interessante explorar ainda mais o atributo de localização, assim como obter informações que possam vir a servir de testes para avaliar o desempenho do modelo de arquitetura proposto.

# 5. REFERÊNCIAS

ALVES, C. B. **SEGURANÇA DA INFORMAÇÃO VS. ENGENHARIA SOCIAL Como se proteger para não ser mais uma vítima**. Centro Universitário do Distrito Federal, Brasília, 2010.

BOLTON, R. J.; HAND, D. J. **Unsupervised profiling methods for fraud detection.** Conference on Credit Scoring and Credit Control 7, Edinburgh, UK, 5-7 September., 2001.

_____., **Statistical Fraud Detection: A Review,** Statistical Science, vol. 17, no. 3,

p. 235-255, 2002.

CORTES, C.; PREGIBON, D. **Signature-based methods for data streams**. Data Mining and Knowledge Discovery, p. 167-83, 2001.

EDGE, M. E.; SAMPAIO, P. R. **Survey of signature based methods for financial fraud detection**. Manchester Business School, University of Manchester, Booth Street East, Manchester M16 6PB, United Kingdom, 2009.

FAWCETT, T. **A Introduction to ROC Analysis.** Pattern Recognition Letters, vol. 27, no. 8, p.861-874, 2006.

FAWCETT, T.; PROVOST, F. **Adaptive Fraud Detection.** Data Mining and Knowledge Discovery, Kluwe, 1, p. 291-316, 1997.

HILAS, C. S.; SAHALOS, J. N. **User profiling for fraud detection in telecommunications networks.** Proceedings of the 5th International Conference Technology and Automation (ICTA'05), Thessaloniki Greece, p. 382-387, 2005

KOVACH, S. **Detecção de Fraudes em Transações Financeiras via Internet em Tempo Real**. Tese de Doutorado, Universidade de São Paulo, São Paulo, 2011.

KOVACH, S.; RUGGIERO, W. V. **Online Banking Fraud Detection Based on Local and Global Behavior,** IPDS, The First International Workshop for Innovative Methods for Intrusion Prevention and Detection Systems, 2011.

MOREAU, Y., VANDEWALLE, J. **Detection of Mobile Phone Fraud using Supervised Neural Networks: A First Prototype**. In: Proceedings of the International Conference on Artificial Neural Networks. 1997

SCIARRETTA, T. **Transações Bancárias On-line Superam Canais de Atendimento tradicionais.** Folha de São Paulo de Dezembro de 2013. Disponível em: <http://www1.folha.uol.com.br/mercado/2013/12/1387272-transacoes-bancarias-on line-superam-canais-de-atendimento-tradicionais.shtml>. Acesso em: 10 jun. 2014.

TERRA NOTÍCIAS. **Google Sabe a Localização do seu Roteador Wi-Fi, diz jornal.** Disponível em: <http://tecnologia.terra.com.br/internet/google-sabe-a-localizacao-do-seu-roteador-wi-fi-diz-jornal,6938ad6ec72ea310VgnCLD200000bbcceb0aRCRD.html>. Acesso em: 15 set. 2014.