

Detecção de Fraudes: Uma Revisão Sistemática

Jean Avila Rangel · Maria Claudia
Figueiredo Pereira Emer · Adolfo
Gustavo Serra Seca Neto

Received: date / Accepted: date

Resumo Insert your abstract here. Include keywords, PACS and mathematical subject classification numbers as needed.

Keywords Fraud detection · Anomaly detection · Deception detection · Standard deviation detection · Detecção de fraude · Detecção de anomalia · Detecção de engano · Detecção de desvio padrão

1 Introdução

Nesta seção são apresentados alguns conceitos básicos sobre os temas abordados nesta pesquisa, junto com as motivações, justificativas e objetivos encontrados.

1.1 Contexto

A detecção de fraudes é utilizada para resolver problemas variados, sendo geralmente utilizada para reduzir falhas de segurança em sistemas onde há gasto de recurso com usuários mal intencionados.

Segundo Fawcett and Provost [1997], há muitas técnicas para detecção de fraudes. As mais difundidas e utilizadas são as que produzem detecções por

Jean Avila Rangel
Federal University of Technology - Paraná, 3165 Sete de Setembro Avenue, Curitiba, PR
80230-901, BRA

Maria Claudia Figueiredo Pereira Emer · Adolfo Gustavo Serra Seca Neto
Academic Department of Informatics, Federal University of Technology - Paraná, 3165 Sete
de Setembro Avenue, Curitiba, PR 80230-901, BRA
E-mail: mciemmer@gmail.com

Adolfo Gustavo Serra Seca Neto
E-mail: adolfo@dainf.ct.utfpr.edu.br

meio de regras pré estabelecidas e as que realizam a comparação entre valores de dados. Estas classificações, ao longo do tempo, geraram ramificações em pesquisas.

Detecção de fraude por meio de regras determina que uma fraude será detectada devido ao conhecimento que a equipe adquiriu observando fraudes recorrentes em outrora. A vantagem nesta técnica está em predir como a fraude ocorre, porém tem como desvantagem a possibilidade de descobrir somente fraudes já conhecidas.

O outro grande ramo na detecção de fraudes ocorre por meio de cálculos e comparação de valores. Como principal exemplo, um desvio no padrão comportamental de alguma variável pode ser identificado comparando seus dados com dados anteriores ou de outras variáveis similares. Sua principal vantagem é a utilização de algoritmos já conhecidos na literatura [Fawcett and Provost, 1997]. O problema apresentado pela técnica é a possibilidade do indivíduo fraudador inserir dados fraudulentos de maneira em que o sistema não consiga o identificar como um usuário malicioso.

1.2 Motivação

O controle organizacional de uma empresa ou setor pode garantir a sua boa estabilidade. Soluções para detectar, prevenir e evitar fraudes podem se tornar uma ferramenta para auxiliar a economia de recursos [Chan et al, 1999].

A detecção de fraudes vem sendo utilizada há muito tempo para controle organizacional e econômico [Seyedhossein and Hashemi, 2010]. Em estudos realizados no estado da arte, houve uma grande apresentação sobre algoritmos para detectar fraudes em sistemas financeiros ou de cartão de crédito [Chandola et al, 2009] e [Abdallah et al, 2016].

Em pesquisas realizadas em bases de dados procurando trabalhos realizados na área de detecção de fraude, notou-se a produção de muito material anteriormente ao ano de 2010, porém com um significativo aumento na porcentagem de publicações posteriores a 2010. Este fator indica o crescimento do interesse em pesquisar determinado assunto e é uma das motivações que o presente trabalho apresenta.

1.3 Justificativa

Em todo contexto organizacional, cada ambiente possui um tipo de tratamento e armazenamento de dados. Portanto, ambientes podem possuir dados organizados de maneira estável por induzirem normativas e tecnologias padronizadas ou fazer exatamente o contrário, possuindo dados desorganizados com poucas conexões lógicas.

Para resolver o problema, técnicas e ferramentas para detectar fraudes em sistemas governamentais estão amplamente difundidas na literatura. Este trabalho busca encontrar relações entre as soluções apresentadas e verificar se existem lacunas onde a área ainda pode ser explorada.

O problema computacional que ainda ocorre é a não detecção de todos os falsos negativos (onde ações fraudulentas não são detectadas), bem como a classificação equivocada de falsos positivos (ações legítimas caracterizadas, erroneamente, como fraudulentas). Identificar uma forma para reduzir a incidência de erros é uma importante contribuição para a área.

1.4 Objeto de Estudo

O objeto de estudo do trabalho consiste em encontrar técnicas de detecção de fraude aplicáveis a contextos gerais e que apresentem bons resultados e desempenho. A eficácia e utilidade dos algoritmos computacionais serão analisadas e consideradas para realizar estudos futuros onde serão aplicadas comitadamente.

1.5 Objetivos

Neste seção são apresentados os objetivos que são almejados na finalização do trabalho.

1.5.1 Objetivo Geral

Este trabalho possui como objetivo geral realizar uma revisão sistemática em trabalhos e técnicas voltados para detecção de fraudes. Desta forma, algoritmos desenvolvidos para inspecionar desvio de padrões, comportamentos indevidos ou fraudes serão identificados e estudados através dos trabalhos publicados no Estado da Arte.

1.5.2 Objetivos Específicos

Para encaixar o objeto que está sendo estudado com o problema apresentado, pretende-se - após realizada a pesquisa entre trabalhos na área - encontrar uma ou um conjunto de técnicas apropriadas para resolver o problema computacional apresentado.

Paralelamente com o objetivo geral, espera-se encontrar as carências que os trabalhos estudados demonstram. As falhas cometidas por autores passados sobre o assunto podem gerar novas pesquisas e um melhor aprofundamento sobre o tema.

Também se espera realizar uma revisão sistemática considerando os trabalhos que foram publicados após as últimas revisões sistemáticas serem escritas.

Todas as técnicas para detecção de fraudes que forem encontradas e consideradas relevantes serão apresentadas e descritas.

2 Materiais e Métodos de Pesquisa

Nesta seção serão apresentados os materiais e métodos que guiaram a pesquisa realizada.

A revisão da literatura foi baseada no trabalho apresentado por Kitchenham and Charters [2007], onde os elementos de pesquisa são definidos para, em seguida, a revisão ser realizada e seus resultados registrados como relatos.

2.1 Estratégia de pesquisa

Ao primeiro momento do trabalho de revisão sistemática, o planejamento da estratégia de pesquisa norteia a pesquisa em publicações na área abordada.

Para a busca de publicações, foram considerados somente trabalhos na língua inglesa e publicados no período temporal entre 2006 e 2016. As palavras chave *fraud detection*, *deception detection* e *anomaly detection* foram utilizadas para a busca, que obteve artigos selecionados com a leitura na respectiva ordem: título, resumo e artigo completo.

Foram selecionadas as bibliotecas digitais abaixo para realizar a pesquisa. O motivo da escolha foi a grande quantidade de material publicado com relevância científica.

- ACM Digital Library¹
- IEEE Xplore Digital Library²
- ScienceDirect³
- SpringerLink⁴

2.2 Processo de seleção dos periódicos

A tarefa para selecionar os periódicos pode ser visualizada no diagrama da Figura 1, onde as etapas foram divididas em início, desenvolvimento e conclusão. Na parte inicial, o tema é definido e publicações naquele tema são pesquisadas.

Após os artigos serem analisados, caso se encontre pouca base para a pesquisa, a tarefa de busca por artigos é refeita e somente após a estabilização de uma boa base para pesquisa, os trabalhos são estudados e os melhores são elencados.

3 Cronograma

Nesta seção é apresentado o cronograma que gerenciou a revisão sistemática até o momento.

¹ <http://dl.acm.org/>

² <http://ieeexplore.ieee.org/>

³ <http://www.sciencedirect.com/>

⁴ <http://link.springer.com/>

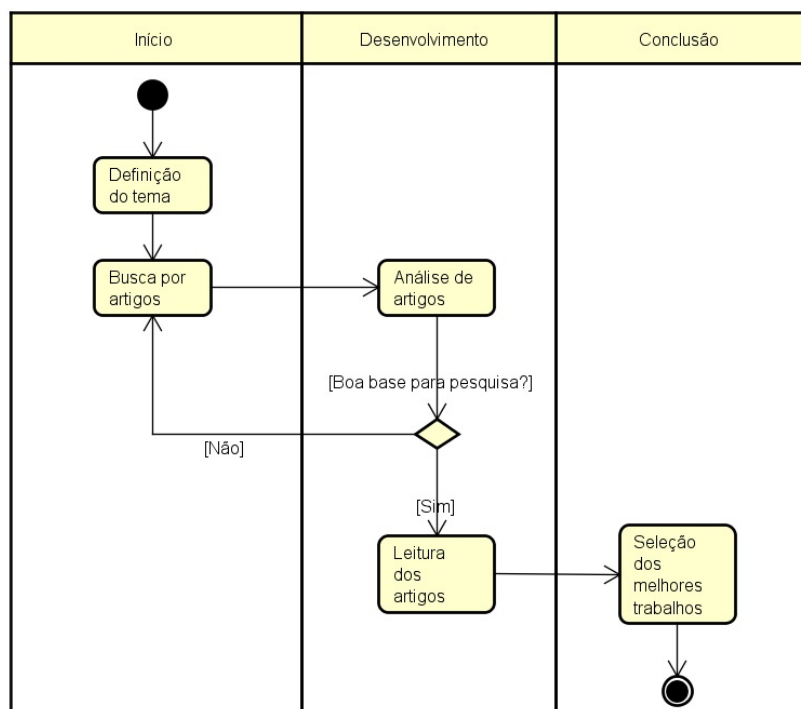


Figura 1 Diagrama de sequência indicando as etapas da pesquisa

Conforme visto na Figura 2, o escopo do trabalho foi delimitado pelas tarefas: Definir tema; Definir objetivos, justificativa e motivação; Escolher o local para publicação; Elaborar o método de pesquisa; Revisar a literatura e selecionar trabalhos; Produzir artigo científico; e Publicar o artigo finalizado.

As tarefas foram divididas entre os meses de junho a setembro e receberam a classificação: Concluída ("OK", em verde); Em andamento ("Andam.", em amarelo); e Não iniciada ("Não Inic.", em vermelho).

Tarefa	Mês			
	Jun	Jul	Ago	Set
Definir tema.	OK			
Definir objetivos, justificativa e motivação.	Andam.			
Escolher o local para publicação.	Ok			
Elaborar o método de pesquisa.	Andam.			
Revisar a literatura e selecionar trabalhos.		Não Inic.		
Produzir artigo científico.		Não Inic.	Não Inic.	Não Inic.
Publicar artigo finalizado.				Não Inic.

Figura 2 Cronograma do planejamento do projeto

4 Projeto de Pesquisa do Mestrado

Na finalização da revisão sistemática, espera-se a publicação dos resultados obtidos na revista *Data Mining and Knowledge Discovery*, a qual possui a classificação de qualidade Qualis B1, elencada pela CAPES em sua última análise.

Referências

- Abdallah A, Maarof MA, Zainal A (2016) Fraud detection system: A survey. *Journal of Network and Computer Applications* 68:90 – 113, DOI <http://dx.doi.org/10.1016/j.jnca.2016.04.007>, URL <http://www.sciencedirect.com/science/article/pii/S1084804516300571>
- Chan PK, Fan W, Prodromidis AL, Stolfo SJ (1999) Distributed data mining in credit card fraud detection. *IEEE Intelligent Systems and their Applications* 14(6):67–74, DOI 10.1109/5254.809570
- Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: A survey. *ACM Comput Surv* 41(3):15:1–15:58, DOI 10.1145/1541880.1541882, URL <http://doi.acm.org/10.1145/1541880.1541882>
- Fawcett T, Provost F (1997) Adaptive fraud detection. *Data Mining and Knowledge Discovery* 1(3):291–316, DOI 10.1023/A:1009700419189, URL <http://dx.doi.org/10.1023/A:1009700419189>
- Kitchenham B, Charters S (2007) Guidelines for performing systematic literature reviews in software engineering
- Seyedhossein L, Hashemi MR (2010) Mining information from credit card time series for timelier fraud detection. In: *Telecommunications (IST), 2010 5th International Symposium on*, pp 619–624, DOI 10.1109/ISTEL.2010.5734099