Detecção de Fraudes: Uma Revisão Sistemática

Jean Avila Rangel[1], **Maria Claudia Figueiredo Pereira Emer**[1], **Adolfo Gustavo Serra Seca Neto**[1]

[1]Universidade Tecnológica Federal do Paraná (UTFPR)
Avenida Sete de Setembro 3165 – 80230-901 – Curitiba – PR – Brasil

jean_rangel94@hotmail.com, mciemer@gmail.com, adolfo@dainf.ct.utfpr.edu.br

Abstract. *Fraudulent activity is an action of a person or a group of people aiming to gain resources in an illicit manner. This article presents a systematic review of the literature in the area of fraud detection to identify the techniques used. These results suggest that most authors divide the problem into areas where fraud occurs most often. The most cited areas were frauds on credit cards, insurance and telecommunication systems. Within each area, specific computational techniques are used for each situation. This work identified, categorized and presented the most commom technical areas guided by previous studies. The main techniques were identified using supervised and unsupervised algorithms. As a contribution, this article presents a discussion on the subject, where the possibility of future research was found as benchmarks to compare tests of algorithms and tools used to locate fraud in the public sector applied, not well covered in the studies found.*

Resumo. *Uma atividade fraudulenta é uma ação realizada por uma pessoa ou um grupo de pessoas visando obter vantagem individual sobre determinado serviço ou recurso. Este artigo apresenta uma revisão sistemática na literatura na área de detecção de fraudes e identificar os tipos e técnicas empregadas. Os resultados obtidos sugerem que a maior parte dos autores divide o problema em áreas nas quais as fraudes ocorrem com maior frequência. As áreas mais citadas foram fraudes em cartões de crédito, em seguradoras e em sistemas de telecomunicação. Dentro de cada área, técnicas específicas da computação são utilizadas para cada situação. Este trabalho identificou, categorizou e apresentou as técnicas e áreas mais abordadas pelos estudos anteriores. As principais técnicas constatadas foram as que utilizam algoritmos supervisionados e não-supervisionados. Como contribuição, este artigo apresenta uma discussão sobre o assunto, no qual foi constatada a possibilidade de futuras investigações, como propor benchmarks para padronizar os testes aplicados nos algoritmos e utilizar ferramentas para localizar fraudes em órgãos públicos, pouco abordados nos estudos encontrados.*

1. Introdução

A detecção de fraudes é utilizada para resolver problemas variados como, por exemplo, reduzir falhas de segurança em sistemas nos quais há gasto de recursos com usuários mal intencionados. A resolução do problema de detecção de fraudes é obtida, em grande parte, com análise computacional de dados. Mais especificamente, a área de mineração

de dados é aplicada, utilizando conhecimentos da estatística, matemática e aprendizado de máquina [Wongchinsri and Kuratach 2016].

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. Desta maneira, este estudo visa conhecer abordagens gerais e adaptáveis em detecção de fraudes para gerenciar os dados armazenados pelas organizações.

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, foram encontrados algoritmos para detectar fraudes em sistemas financeiros ou de cartão de crédito [Chan et al. 1999], [Chandola et al. 2009] e [Abdallah et al. 2016]

O objetivo geral deste trabalho é identificar e categorizar técnicas e ferramentas computacionais para detecção de fraudes dentro de áreas distintas.

Como objetivos específicos, este trabalho busca:

- Descobrir a importância das áreas de mineração de dados e aprendizado de máquina e suas utilizações em detecção de fraudes;
- Examinar se os estudos realizados em detecção de fraudes apresentam oportunidades para áreas pouco ou não abordadas;
- Elencar como os autores validaram suas pesquisas nas áreas de detecção de fraudes;
- Agrupar quais tecnologias foram as mais utilizadas para cada contexto.

A realização da revisão sistemática proposta por este trabalho obtém motivação devido à crescente expansão do estado da arte no assunto [Pejic-Bach 2010]. Em pesquisas realizadas em bases de dados por trabalhos realizados na área de detecção de fraude, notou-se grande produção de artigos anteriores ao ano de 2010, porém com um significativo aumento na porcentagem de publicações posteriores. Este fator indica a ampliação do interesse da comunidade científica no tema.

Embora cartões de crédito e sistemas de telecomunicações tenham sido criados há décadas, ainda estão em crescimento em utilização, resultando no aumento de possíveis fraudes que podem gerar prejuízos [Abdallah et al. 2016]. No início dos anos 2000, as palavras chave *fraud detection* já contavam com mais de 80 patentes registradas [Bolton et al. 2002].

A seguir, na Seção 2, será descrito o referencial teórico que situará o assunto abordado. Na sequência, a metodologia utilizada para realizar esta revisão sistemática será apresentada e discutida na Seção 3. Apresentando o estado da arte, os trabalhos encontrados serão elencados, analisados e discutidos na Seção 4. No final do texto haverá

a finalização desta revisão sistemática, em que ocorrerá uma discussão, na Seção 5, e a conclusão, na Seção 6.

2. Referencial teórico

Esta Seção fornece um panorama geral da detecção de fraudes, situando o problema e relacionando pesquisas.

A área de detecção de fraudes possui sua base estabelecida no campo da mineração de dados, na qual a computação é utilizada para gerar informação a partir de dados coletados, geralmente, de maneira automática. Os dados iniciais não apresentam sentido para seu possuinte, porém, no âmbito empresarial, a quantidade de informação tem se mostrado relevante para o sucesso de corporações.

A Figura 1 apresenta a compreensão de [Fayyad et al. 1996] sobre a descoberta de conhecimento em bases de dados.

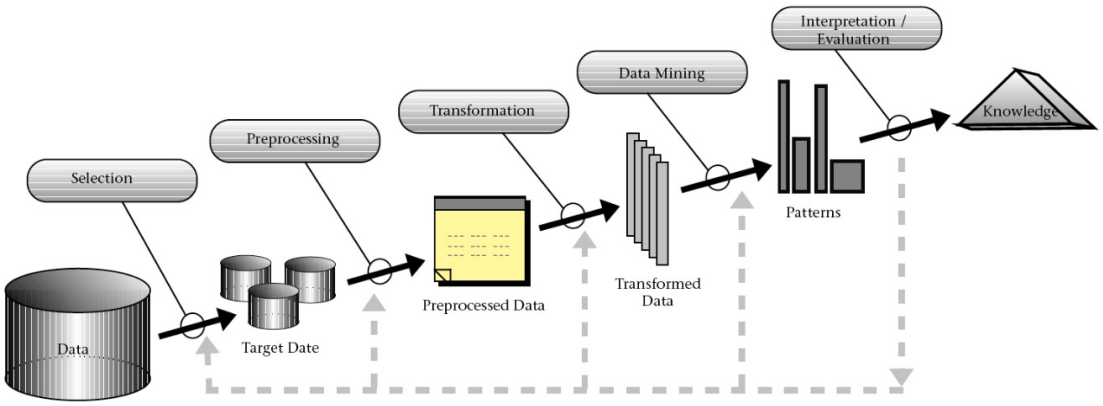


Figura 1. Mineração de dados para ganho de informações. Fonte: [Fayyad et al. 1996]

Na Figura 1, temos um processo evolutivo para a mineração de dados pode ser observado. Inicialmente, os dados de um conjunto são selecionados e separados em conjuntos menores para melhor manuseio e visibilidade. Em seguida, os dados são pré-processados e transformados para serem legíveis por algoritmos computacionais na etapa de mineração de dados. Após os resultados serem exibidos, um humano com conhecimento sobre a base de dados inicial avalia as informações fornecidas pelos algoritmos para adquirir algum conhecimento que não se possuía até então.

O maior processamento computacional do processo, em comparação entre as seis etapas, ocorre na etapa de mineração de dados, na qual algoritmos complexos são executados para gerar informações. Todavia, o maior gasto temporal se encontra nas outras etapas, pois torna-se necessário o trabalho manual para trabalhar com os dados.

A mineração de dados foi utilizada pela primeira vez de maneira empírica em setores do comércio. As empresas possuíam dados de compras de seus clientes e, com o cruzamento das informações, conseguiam predizer o perfil do usuário.

Em uma rede de supermercados, o caixa eletrônico registra todos os produtos que os clientes compram na sexta-feira e, como padrão, muitos adquirem itens em comum. Como estratégia de *marketing*, o supermercado pode gerar mais lucro após possuir o conhecimento nesse novo perfil identificado, pois pode usar a informação a seu favor.

A área de detecção de fraudes obteve início em pesquisas nas áreas da estatística e da matemática. No âmbito computacional, [Fawcett and Provost 1997] realizaram uma das primeiras e principais abordagens encontradas na literatura. Em seus estudos, analisaram e criaram um sistema para detecção de fraudes em sistemas de telecomunicação. Em seus estudos, constaram que indivíduos fraudadores estão constantemente mudando suas táticas para burlar sistemas, que devem ser adaptáveis às novas situações.

De acordo com a Associação dos Examinadores Certificados de Fraudes (Association of Certified Fraud Examiners), a definição de fraude é: o uso de forma incorreta de algum setor ou recurso para o aumento dos benefícios individuais [Abdallah et al. 2016] e [Allan and Zhan 2010].

Idealmente, o custo para detectar uma fraude deve ser menor do que o gasto gerado por ela. Para ilustrar uma situação hipotética, uma empresa fictícia que contém um milhão de dados em um arquivo pode possuir 1% de falsos alarmes. Esta quantia, embora aparentemente pequena, representará para a equipe responsável em avaliar fraudes cerca de 10.000 avisos. Alocar funcionários para verificar cada uma destas informações terá um custo considerável para a empresa e deverá ser comparado com o prejuízo de deixar as fraudes acontecerem.

Posteriormente, utilizando dados reais fornecidos por companhias bancárias, [Chan et al. 1999] realizaram um trabalho em detecção de fraude em cartões de crédito e obtiveram um aumento nos seus resultados em comparação com técnicas de detecções de fraudes utilizadas anteriormente. Nas constatações, indicaram que os métodos mais potentes para detectar fraudes são utilizados por corporações bancárias. Todavia, as organizações não costumam revelar suas tecnologias de forma aberta, pois possíveis fraudadores podem se aproveitar do conhecimento e utilizar os sistemas com má fé.

Segundo [Fawcett and Provost 1997], há muitas técnicas para detecção de fraudes. As mais difundidas e utilizadas são as que produzem detecções por meio de regras pré estabelecidas e as que elaboram a comparação entre valores de dados. Estas classificações geraram duas ramificações iniciais em pesquisas da área.

Uma das ramificações, a 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 predizer como a fraude ocorre. A desvantagem está na possibilidade de descobrir somente fraudes já conhecidas, pois os dados serão comparados levando em conta as informações obtidas anteriormente.

O outro ramo na detecção de fraudes ocorre por meio de comparação de valores sem muita informação sobre os dados atuais ou anteriores, sem saber se os valores similares do passado apresentaram fraudes ou não. Como principal exemplo, um desvio no padrão comportamental de alguma variável pode ser identificado comparando seus dados com outras variáveis similares. Sua principal vantagem é a disponibilidade de um amplo número de algoritmos destinados para a tarefa [Fawcett and Provost 1997]. O problema apresentado pela técnica é a possibilidade do indivíduo fraudador conhecer e inserir dados

fraudulentos de maneira que o sistema não consiga o identificar como um usuário malicioso, não sendo tão eficiente quanto a primeira técnica, que utiliza informações prévias para prever novas situações.

Provost, ao comentar sobre o trabalho de [Bolton et al. 2002], indica que é importante não visualizar a detecção de fraudes somente como uma área, mas desmembrá-la em sub-áreas para promover a interdisciplinariedade. Com isto, constata-se que a detecção de fraudes não possui uma técnica universal, pois ela contém diversas abordagens devido às mutações de contexto.

3. Metodologia

Ao primeiro momento de uma revisão sistemática, o planejamento da estratégia de pesquisa norteia a procura de publicações na área abordada. Esta Seção é destinada a detalhar o processo metodológico utilizado para orientar a revisão sistemática proposta.

A revisão da literatura foi baseada no trabalho apresentado por [Kitchenham and Charters 2007], no qual os elementos de pesquisa são definidos para desenvolver e relatar a revisão.

A revisão sistemática é, de acordo com [Kitchenham 2004], uma forma de identificar, avaliar e interpretar todas as pesquisas relevantes para uma determinada situação.

A autora indica que todos os estudos individuais em determinada área são categorizados como estudos primários, e as revisões sistemáticas, são chamadas de estudos secundários.

As razões para desenvolver revisões sistemáticas, segundo a autora, são: (1) sumarizar as tecnologias e técnicas existentes; (2) identificar as lacunas presentes nas pesquisas para sugerir novos estudos; (3) prover um arcabouço para os autores obterem embasamento científico; e (4) examinar se as evidências empíricas dão suporte ou contradizem as hipóteses teóricas.

As questões de pesquisa (QP) foram desenvolvidas seguindo o modelo de revisões sistemáticas proposto por [Kitchenham and Charters 2007]. As questões de pesquisa para o desenvolvimento deste trabalho são:

- **QP1:** *Quais são as áreas de detecção de fraudes mais estudadas na literatura?*
- **QP2:** *Como a literatura categoriza as técnicas de detecção de fraudes?*
- **QP3:** *Quais foram os problemas mais relatados pelos autores?*
- **QP4:** *Como os autores testaram e validaram suas pesquisas?*
- **QP5:** *Em qual área há pouca abordagem no estudo de detecção de fraudes?*
- **QP6:** *Há espaço para futuras pesquisas na área?*

Foram selecionadas as bibliotecas digitais a seguir:

- ACM Digital Library[1]

[1] `http://dl.acm.org/`

- IEEE Xplore Digital Library[2]
- ScienceDirect[3]
- SpringerLink[4]

A tarefa selecionar os artigos pode ser visualizada no diagrama da Figura 2, na qual as etapas foram divididas em início, desenvolvimento e conclusão. Na parte inicial, o tema é definido e publicações naquele tema são pesquisadas. Posteriormente, se após a busca por artigos a base de referências não estiver sólida para um bom estudo, a etapa de busca é refeita até obter uma boa base do estado da arte na área. Por fim, os artigos selecionados são estudados.

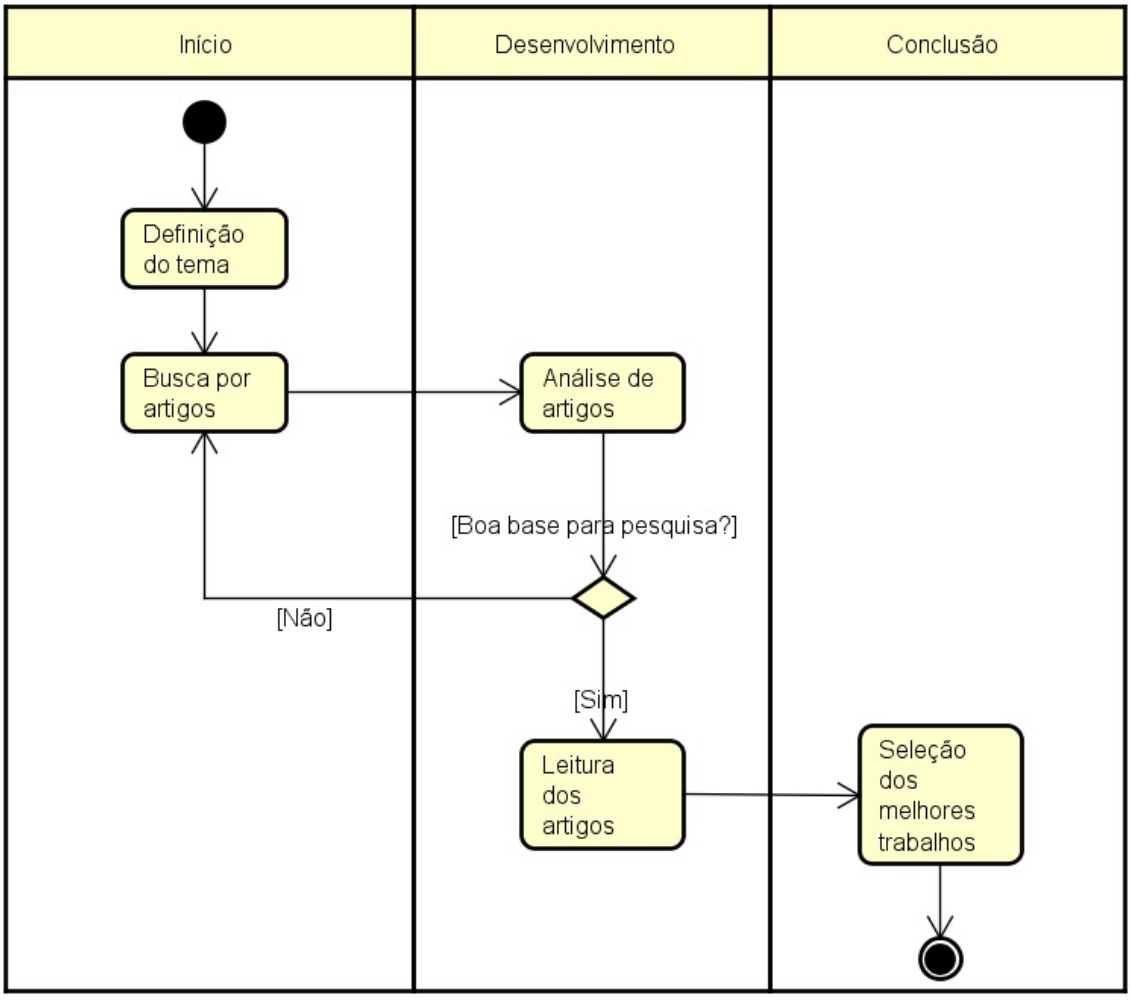


Figura 2. Diagrama de sequência indicando as etapas de seleção de artigos para a pesquisa.

Como critérios de seleção e ordenação para a pesquisa, a revisão escolheu os artigos com maior número de citações e relevância para a área. As medições foram feitas através dos filtros de pesquisa das bases escolhidas.

[2]http://ieeexplore.ieee.org/
[3]http://www.sciencedirect.com/
[4]http://link.springer.com/

O principal critério de seleção foi a proximidade com o tema destacado na introdução deste trabalho, em que a detecção de fraudes deveria ser a principal abordagem do artigo. A proximidade com o tema foi identificada a partir do título, resumo e introdução dos artigos. Além disso, os artigos encontrados como referência na maioria dos trabalhos atuais que tratam do tema de detecção de fraudes foram incluídos no estudo.

Os artigos que fugiam do tema proposto, não possuíam citações ou relevância nas bases de dados ou eram anteriores ao ano de 2006 foram excluídos. Também foram excluídos da seleção, os artigos contendo assuntos relacionados a detecção de fraudes em:

- Imagens, vídeos ou impressões físicas;
- Contextos biológicos, como análises de DNA e moléculas;
- Componentes eletrônicos;
- Dispositivos de segurança eletrônica relacionados a vírus, *firewalls* e *cyber* ataques.

Para esta revisão sistemática, obteve-se foco em detecção de fraudes em dados computacionais como números e informações em tabelas, arquivos de texto, planilhas eletrônicas, etc. Isso justifica a exclusão de artigos tratando de áreas como processamento de imagens, biologia, eletrônica e segurança de redes de computadores.

[Flegel et al. 2010] demonstraram em seus estudos que técnicas para detectar invasores em redes de computadores, embora teoricamente similares com técnicas para identificar usuários maliciosos no contexto de detecção de fraudes, não são indicadas.

A mesma situação ocorre nos contextos apresentados anteriormente, como detecção de anomalias em imagens, componentes eletrônicos (incluindo hardwares como roteadores) e sistemas biológicos.

Outro fator que levou a pesquisa de publicações a excluir os fatores anteriores foi o grande número de publicações focadas em detecção de fraudes em dados financeiros, que se aproximam do objetivo geral deste trabalho.

Para a busca de publicações, foram considerados somente trabalhos que realizaram uma revisão no estado da arte em detecção de fraude, na língua inglesa e publicados no período entre 2006 e agosto de 2016. As palavras chave *fraud detection survey* foram utilizadas para a busca, que obteve artigos selecionados com a leitura na respectiva ordem: título, resumo e artigo completo.

Foram considerados somente trabalhos em que nenhum novo método ou algoritmo era apresentado ou testado, para aumentar a confiabilidade e imparcialidade dos autores.

Na primeira fase da pesquisa, foi considerada a utilização da combinação das palavras chave a seguir, que posteriormente foram eliminadas da metodologia, pois não retornaram resultados satisfatórios:

- *fraud detection AND state of art OR systematic review OR meta analysis*;
- *anomaly detection AND survey OR state of art OR systematic review OR meta analysis*;
- *outlier detection AND survey OR state of art OR systematic review OR meta analysis*;

- *deception detection AND survey OR state of art OR systematic review OR meta analysis*.

Os resultados contendo os temas *anomaly, outlier* e *deception detection* juntamente com *survey* retornaram trabalhos duplicados em comparação com a pesquisa utilizando as palavras chave *fraud detection survey* ou fora do escopo. A grande parte foi eliminada da seleção utilizando os métodos de exclusão citados anteriormente. Por esta razão, essas palavras chave não foram consideradas para a pesquisa de artigos.

Outras palavras que podem retornar revisões sistemáticas na língua inglesa são: *state of art, systematic review* ou *meta analysis*. Contudo, a utilização dessas palavras chave também se mostrou redundante em comparação à palavra *survey*, pois trouxe muitos resultados duplicados ou em áreas indesejadas para esta revisão sistemática, como processamento de imagens, biologia, eletrônica e segurança de redes de computadores.

Portanto, o conjunto de palavras chave que obteve resultados satisfatórios para a revisão sistemática foi:

- *fraud detection survey*.

Conforme apresenta a Figura 3, no início da revisão sistemática, foram encontrados 101 artigos com potencial para estruturar este trabalho. Destes, dois não estavam disponíveis para consulta, pois requisitavam nível de assinatura na base de dados superior à acadêmica. Após a leitura do título dos 99 artigos restantes, 49 foram removidos, seguindo os critérios de exclusão. Posteriormente foi realizada a leitura do resumo dos artigos. Destes, 22 trabalhos foram selecionados para prosseguir a revisão sistemática proposta por este trabalho. Os trabalhos estão divididos entre conferências e revistas, conforme visto na Tabela 1.

Tabela 1. Artigos publicados em conferências e revistas

Categoria	Trabalhos	Porcentagem
Conferência	12	55%
Revista	10	45%

4. Estado da arte

Esta Seção irá apresentar os estudos de autores na área de detecção de fraudes. Conforme descrito na Seção de metodologia, foram considerados artigos que realizaram uma revisão sistemática na literatura sobre o tema.

4.1. Revisões sistemáticas gerais em detecção de fraudes

De acordo com o *Basel Committee on Bank Supervision*, existem fraudes de nível interno e externo. Fraudes em nível interno se referem às ocorridas quando empregados cometem fraudes contra a própria organização. Fraudes de nível externo correspondem a clientes ou indivíduos distantes das organizações que obtém proveito de falhas variadas [Abdallah et al. 2016].

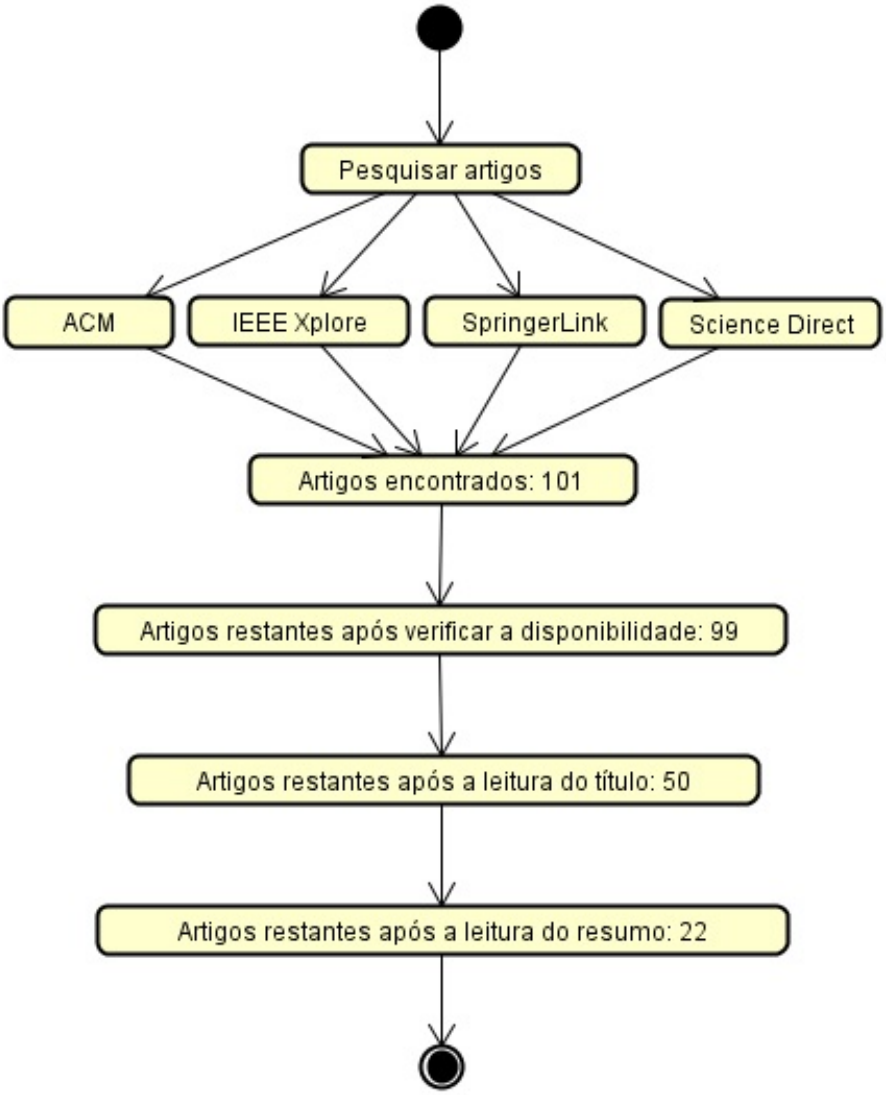


Figura 3. Etapas de inclusão e exclusão dos artigos, seguindo critérios.

Ainda conforme [Abdallah et al. 2016], o índice de fraudes continua a crescer. De 2011 para 2016, por exemplo, obteve-se um acréscimo de 15% nas fraudes cometidas em sistemas de telecomunicações.

[Abdallah et al. 2016] elencaram as áreas mais presentes em trabalhos anteriores e as tecnologias mais utilizadas em cada contexto. Segundo os autores, fraudes ocorridas em bancos compõem 53% dos estudos, seguidas por fraudes em seguradoras (31%), em comércio eletrônico (10%) e em sistemas de telecomunicações (6%). O gráfico com as informações está demonstrado na Figura 4.

A justificativa do maior índice de fraudes ocorrer em operações bancárias é o crescente aumento da utilização de cartões de crédito, que expandiram rapidamente nas últimas duas décadas [Wongchinsri and Kuratach 2016].

[Flegel et al. 2010] indicam que na indústria, a abordagem de detecção de frau-

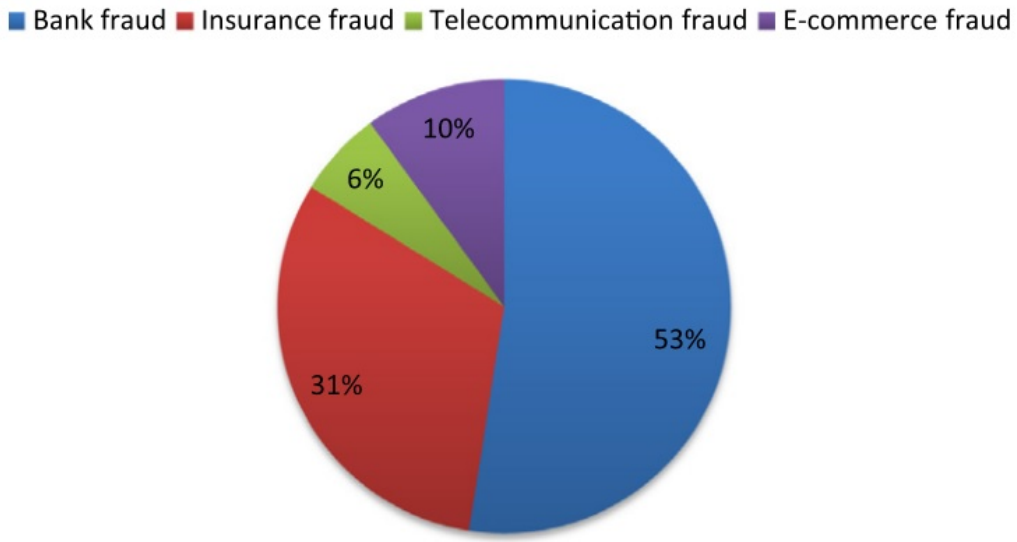


Figura 4. Áreas com maior pesquisa em detecção de fraudes. Fonte: [Abdallah et al. 2016]

des não reflete a grande importância que a literatura dá para o tema. Fora do contexto acadêmico, a maior parte do processo de descobrir uma fraude é realizado por um humano que utiliza ferramentas triviais para controle de dados, como planilhas eletrônicas *Microsoft Excel* ou *OpenOffice Calc*.

Os autores ressaltam que tecnologias para detectar e coibir tentativas de invasão em sistemas de redes de computadores não podem ser comparadas - ou suas técnicas utilizadas - com detecção de fraudes em dados financeiros.

[Flegel et al. 2010] diferenciam as informações financeiras como sendo dados organizados de uma forma diferente com as características que usuários tentando atacar *hardwares* de computadores apresentam.

Contradizendo os argumentos de [Flegel et al. 2010], [Pejic-Bach 2010] realiza uma revisão sistemática considerando artigos do período de 1956 à 2009 e chega a conclusões diferentes. A autora [Pejic-Bach 2010] relata que a academia e a indústria dão grande atenção para detecção de fraudes, e categoriza as subáreas de maneira similar a [Abdallah et al. 2016].

[Bansal 2016] classifica as principais abordagens da detecção de fraudes como aplicações para: cartões de crédito, seguradoras, sistema interno de empresas, sistemas médicos e de saúde e equipamentos industriais. O artigo apresenta as tecnologias para resolver os problemas apresentados, com enfoque na detecção de *outliers*.

Categorizando os *outliers*, os autores denominam: *Point Outliers, Contextual Outliers* e *Collective Outliers*, que são, respectivamente, *outliers* individuais (valores com informações muito divergentes da maioria), *outliers* por contexto (um valor diferente dos outros, porém aceitável naquele contexto, ex: um usuário que gastou além do normal no dia 24 de dezembro, véspera de natal) ou *outliers* que realizam uma formação de cartel fora do padrão esperado (similar ao individual, no qual um grupo de indivíduos possui uma atividade em comum que é divergente da maioria).

[Bansal 2016] indica que para detectá-los, as formas mais difundidas na literatura são: *Distance based outlier detection, Clustering based outlier detection, Density based outlier detection* e *Depth based outlier detection*.

4.2. Técnicas de detecção de fraudes

4.2.1. Mineração de dados por aprendizado supervisionado (algoritmos de classificação) e não-supervisionado (algoritmos de agrupamento)

Grande parte dos algoritmos de mineração de dados trabalham com dados similares a tabelas em bancos de dados, nos quais cada instância está representada em uma linha e seus atributos são indicados em colunas. Nos dados supervisionados, uma coluna (geralmente a última) possui uma informação classificadora, que pode ser uma variável binária ou nominal.

[Zhao 2012] utiliza o conjunto de dados Iris[5] para exemplificar técnicas de mineração de dados utilizando a linguagem de programação R. O conjunto apresenta informações a respeito das dimensões de pétalas e sépalas de três espécies de flores. Pode-se observar na Tabela 2 os cinco primeiros registros do Iris *dataset*. Os nomes dos atributos estão presentes na primeira linha horizontal, e dão sentido para cada instância, em suas respectivas linhas. No conjunto utilizado, a última coluna de informação é referente a classe que a instância faz parte. A coluna está indicada com o título classe sublinhado.

Tabela 2. Cinco primeiras instâncias do *dataset* Iris. Fonte: [Zhao 2012]

	Atributos					Classe
Instâncias	***Id***	***Sepal.Length***	***Sepal.Width***	***Petal.Length***	***Petal.Width***	***Species***
	1	5.1	3.5	1.4	0.2	setosa
	2	4.9	3.0	1.4	0.2	setosa
	3	4.7	3.2	1.3	0.2	setosa
	4	4.6	3.1	1.5	0.2	setosa
	5	5.0	3.6	1.4	0.2	setosa

Com o foco em detecções de anomalias, [Ahmed et al. 2015] limitam-se a trabalhar com dados não-supervisionados, que são dados sem classificações a respeito de seus conteúdos, pois pontuam que a maior parte dos dados reais não apresentam algum atributo classificador.

Os autores indicam que a aplicação de técnicas de clusterização/agrupamento (como o algoritmo mais popular para a área, o *K-Means*) é comum para detectar anomalias nesses tipos de dados. Além das classificações apresentadas por [Ahmed et al. 2015], o estudo indica a fraude no mercado de ações como um nicho entre os criminosos, no qual pessoas que possuem informações internas antes do público geral as utilizam para realizar mudanças nos lucros.

Em contraste com os dados não-supervisionados, os dados supervisionados são informações em que os pesquisadores possuem controle de quais instâncias representam ou não uma fraude. Um exemplo de dados supervisionados são tabelas com informações

[5] `https://archive.ics.uci.edu/ml/datasets/Iris`

de gastos em um cartão de crédito, em que o banco especificou previamente quais dados são considerados fraudulentos [Akoglu et al. 2015] e [Branco 2016]. Para os dados supervisionados (nos quais alguma parte das informações contém algum registro classificador como fraude), os algoritmos especializados são treinados utilizando aprendizado de máquina para tentar prever novos valores sem classificação.

Portanto, enquanto os métodos de clusterização (não-supervisionados) tentam identificar padrões por não possuírem alguma informação da classificação de determinada instância, os algoritmos de classificação supervisionados realizam o treinamento sabendo quais instâncias são fraudulentas ou verossímeis.

O problema apresentado pela utilização de dados supervisionados é a dificuldade de encontrar novas anomalias idênticas as já reconhecidas [Ahmed et al. 2015]. Ao final da revisão sistemática, os autores indicam que uma técnica universal para detectar fraudes ainda está para ser encontrada devido à grande variação no contexto da anormalidade.

A tarefa de selecionar os atributos que levarão para o algoritmo algum acréscimo de informação para um melhor resultado deve executada por um humano. Nenhum algoritmo de pré-processamento é inteligente o suficiente para compreender os dados de uma base e selecionar automaticamente os dados que achar relevante de uma forma melhor que uma pessoa experiente na área.

Um dos fatores que contribui com o aumento do processamento computacional é a maldição da dimensionalidade (*curse of dimensionality*), em que muitos atributos desprezíveis no conjunto de dados levam o algoritmo a ter menor eficiência [Hodge and Austin 2004].

Na etapa da seleção de atributos, a dificuldade está em nivelar o tamanho do *dataset* para ficar pequeno o suficiente para que tenha sentido ao humano que irá manuseá-lo e ao algoritmo que irá computá-lo em tempo hábil e grande o suficiente para abstrair a maior parte das possibilidades.

Conforme visto na Figura 5, a porcentagem de dados fraudulentos em um *dataset* é bem menor que a porcentagem de dados verossímeis. Por esta razão, se ocorrer na etapa de seleção de atributos o descarte de uma parte significante de dados fraudulentos, as instâncias não terão grande visibilidade, tanto em algoritmos supervisionados quanto não-supervisionados.

Os dados representados na Figura 5 são de um contexto real, originários de compras virtuais realizadas por clientes em um site da internet. Os dados representam 100.000 transações de 73.729 clientes no período de 98 dias. O *dataset* possui 97.346 transações normais, restando 2654 exemplos fraudulentos.

Portanto, na etapa de pré-processamento das informações, uma pessoa que possui conhecimento sobre os dados da base pode realizar a seleção de atributos.

Realizando uma generalização, a Figura 6, proposta por [Allan and Zhan 2010] demonstra que os dados classificados como fraudulentos geralmente aparecem em menor número em comparação com dados legais. O conjunto de dados utilizado para o aprendizado de um algoritmo supervisionado é geralmente menor do que os dados que serão utilizados para teste ou uma situação real.

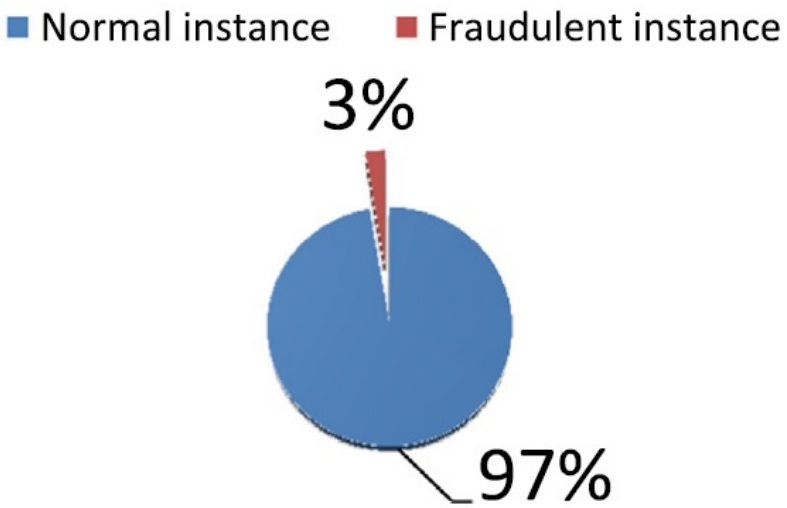


Figura 5. Porcentagem de instâncias normais e fraudulentas num *dataset* real. Fonte: [Abdallah et al. 2016]

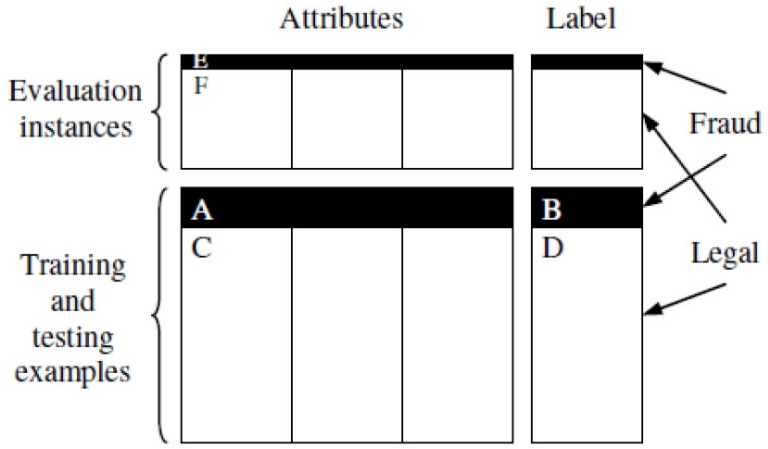


Figura 6. Distribuição padrão na característica dos dados. Fonte: [Allan and Zhan 2010]

Em casos de dados não-supervisionados, não há o campo *label* do atributo, no qual geralmente é dada a informação de que aquele valor é ou não uma fraude.

[Rebahi et al. 2011] discutem a abordagem baseada em aprendizado, similar ao método supervisionado, porém computacionalmente mais simples. Os autores indicam que a característica principal desta técnica está na necessidade de definir regras iniciais. Portanto, dados e conhecimentos de atividades fraudulentas são necessários para construir alarmes, que serão disparados quando dados naquelas características forem encontrados no sistema.

[Pejic-Bach 2010], [Wang 2010] e [Raj and Portia 2011] analisaram trabalhos que utilizam algoritmos destinados a detectar fraudes. Com as pesquisas, constatam uma grande frequência na utilização de determinadas abordagens, como as técnicas de redes neurais, lógica *fuzzy*, algoritmos genéticos, computação evolucionária (ou evolutiva), programação genética e otimização por nuvem de partículas. [Raj and Portia 2011] cons-

tataram, além das técnicas apresentadas anteriormente, os métodos de aprendizado por Inferência Bayesiana, Teoria de Evidência de Dempster-Shafer, *BLAST-SSAHA Hybridization* e Modelo oculto de Markov.

A Tabela 3 apresenta um resumo das abordagens e técnicas de mineração de dados mais utilizadas em classificação e clusterização (agrupamento), conforme apresentado por [Abdallah et al. 2016]. Todavia, é importante ressaltar que os algoritmos indicados como técnica para determinada categoria podem ser utilizados por outro método ou mesmo em conjunto.

Tabela 3. Técnicas e categorias mais utilizadas para mineração de dados. Fonte: [Abdallah et al. 2016]

Método	Categoria	Técnica
Supervisionado	Classificação	Algoritmo genético
		Árvores de decisões
		Baseado em regras
		Inferência Bayesiana
		K-Nearest neighbours (K-NN)
		Naive Bayes
		Modelo oculto de Markov
		Redes neurais artificiais
		Support vector machine (SVM)
Não-supervisionado	Clusterização	Baseado em distância
		K-Means
		Lógica Fuzzy
		Mistura gaussiana
		Principal component analysis (PCA)

Além das técnicas para realizar mineração de dados apresentadas na tabela anterior, técnicas destinadas à visualização e estatística dos dados também foram comentadas pelos autores. O objetivo dessas técnicas é somente apresentar para o usuário uma forma de organização para melhor compreensão dos dados. Em poucos casos, foram utilizadas análises de regressão.

[Allan and Zhan 2010] também indicam que é possível o cruzamento de técnicas, no qual pode-se executar métodos semi-supervisionados em instâncias em que parte possuem classificação, e outra parte não. Métodos semi-supervisionados também podem ser aplicados em sequência e métodos supervisionados e não-supervisionados sendo utilizados em conjunto de maneira híbrida.

Para testar as técnicas e novos algoritmos, muitos autores não utilizam dados reais, obtidos por eles, e recorrem a algum repositório de *datasets*. O mais conhecido na área de mineração de dados é o UCI Machine Learning Repository[6], que contém diversos conjuntos de dados para métodos supervisionados e não-supervisionados.

Para a utilização da maioria das técnicas apresentadas nesta Seção,

[6]https://archive.ics.uci.edu/ml/datasets.html

[Ahmed et al. 2015] observam a grande utilização da ferramenta Weka[7] e da linguagem R[8].

O *dataset* Iris é o conjunto de dados mais conhecido na área de mineração de dados [Tan et al. 2005]. No conjunto Iris, informações a respeito do tamanho de sépalas e pétalas de flores classificam-nas em três grupos distintos.

Utilizando a linguagem R, [Zhao 2012] apresenta exemplos para utilização dos algoritmos, conforme visualizado na Figura 7, que apresenta o resultado do diagrama de caixa no *dataset* Iris. O diagrama de caixa calcula a mediana dos valores do atributo escolhido e seus quartis. No exemplo, pode-se visualizar um círculo representando um *outlier* na classe das flores do tipo virgínica.

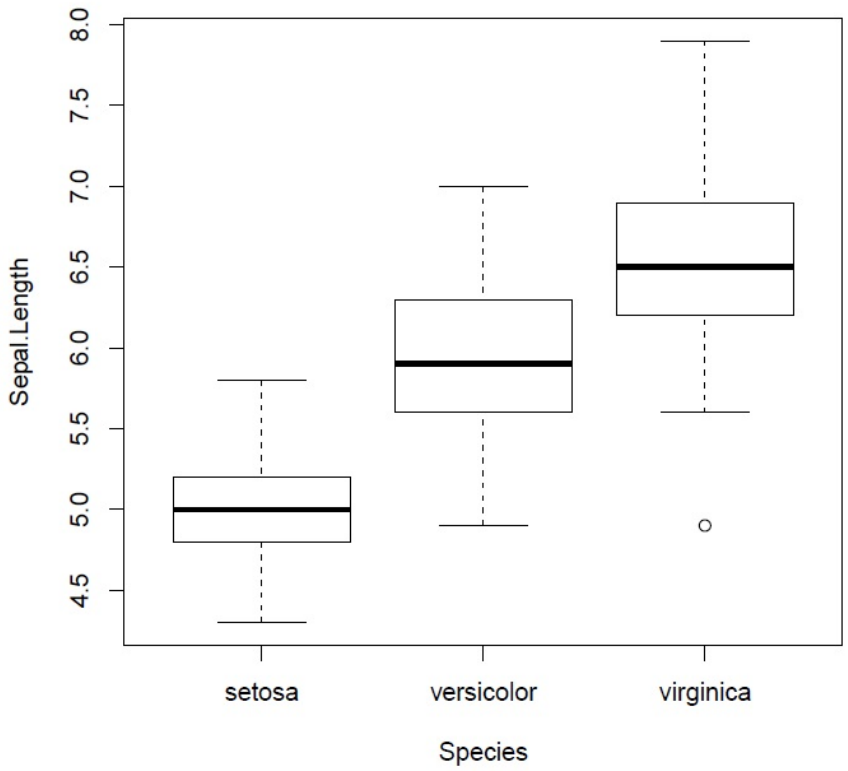


Figura 7. Diagrama de caixa. Fonte: [Zhao 2012]

Como função de visualização de dados, [Zhao 2012] apresenta um gráfico tridimensional gerado sobre o Iris *dataset* utilizando três atributos (dois das sépalas e um da pétala), conforme representado na Figura 8.

Exemplificando um algoritmo de clusterização com a Figura 9, na utilização do K-means nos dados da sépala constata-se que no mínimo dois grupos distintos poderiam ser observados, caso não houvesse a classificação presente no conjunto de dados.

Nos casos para classificação, [Zhao 2012] apresenta como um dos exemplos uma árvore de decisão para o conjunto de dados Iris. Na árvore da Figura 10, [Zhao 2012]

[7] http://www.cs.waikato.ac.nz/ml/weka/
[8] https://www.r-project.org/

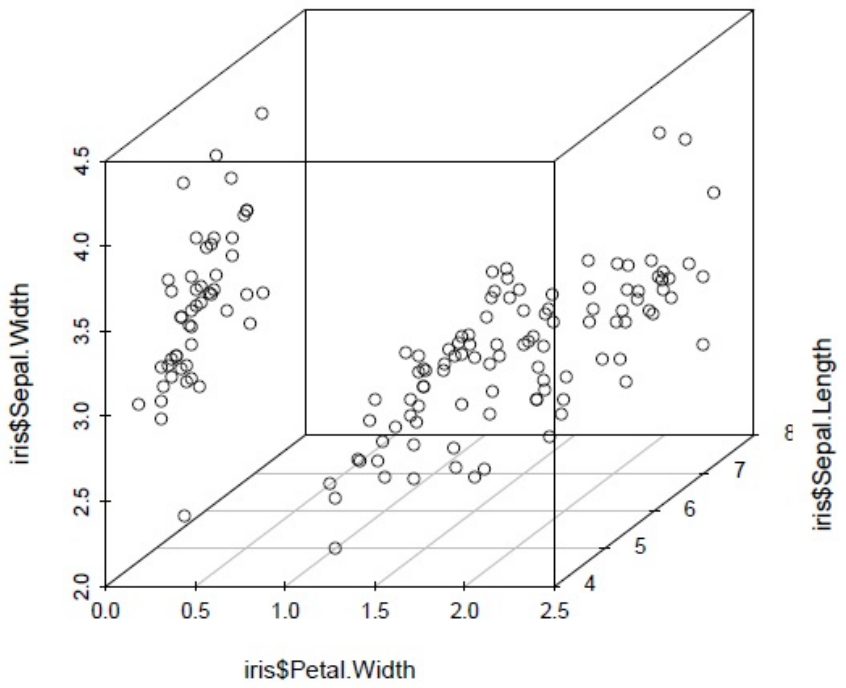


Figura 8. Gráfico tridimensional gerado a partir do cruzamento de três atributos. Fonte: [Zhao 2012]

identifica alguns fatores considerados pelo algoritmos para predizer a classe que alguma flor pode pertencer.

Na matriz de confusão gerada após a criação da árvore de decisão e exibida na Tabela 4, os dados tiveram seus atributos classificadores desconsiderados e então classificados conforme o aprendizado de máquina considerou as características dos demais atributos. No final do processo, as instâncias com classificação são comparadas com as classes preditas pelo algoritmo, gerando a matriz de confusão. No exemplo apresentado, somente duas flores foram consideradas de uma classe errada, havendo uma confusão entre as classes virgínica e versicolor.

Tabela 4. Matriz de confusão. Fonte: [Zhao 2012]

testPred	setosa	versicolor	virginica
setosa	10	0	0
versicolor	0	12	2
virginica	0	0	14

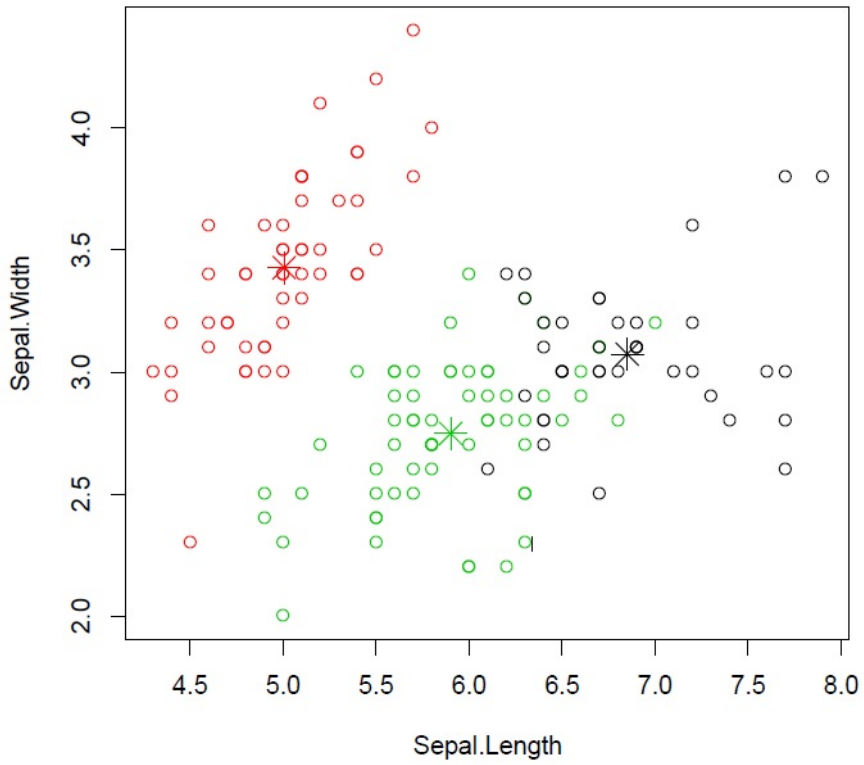


Figura 9. Algoritmo de agrupamento K-means. Fonte: [Zhao 2012]

4.2.2. Tipos comuns de anomalias em dados

Focando nos métodos não-supervisionados, [Ahmed et al. 2015] identificam em sua revisão da literatura em detecção de anomalias os três tipos de anomalias mais comuns.

Os grupos são similares aos grupos identificados por [Bansal 2016]. [Ahmed et al. 2015] indicam: anomalia pontual, onde um dado está fora do padrão de um conjunto de dados; a anomalia contextual, onde um dado está fora do padrão de um conjunto de dados considerando o seu contexto (usando como exemplo, novamente, o acréscimo de gastos com cartões de crédito no dia 24 de dezembro, que é considerado uma exceção por contexto pelas compras natalinas); e anormalidades coletivas, onde espécies de cartéis e conjuntos de dados são identificados fora do padrão dos demais dados.

Nas representações da Figura 11, pode-se constatar as diferenças entre os conjuntos de anomalias. No contexto do mundo real, com dados não-fictícios, a diferença não será tão trivial para ser visualizada.

A Figura 11a apresenta o indivíduo anormal mais recorrente, em que uma única instância possui dados divergentes da maioria das outras instâncias, que formaram um *cluster* (agrupamento). Para o especialista em detecção de fraudes, é uma tarefa relativamente simples tratar esse indivíduo.

Na Figura 11b, há outras instâncias similares com a instância anormal, construindo um novo agrupamento de *outliers*. Caso o grupo seja relativamente grande em

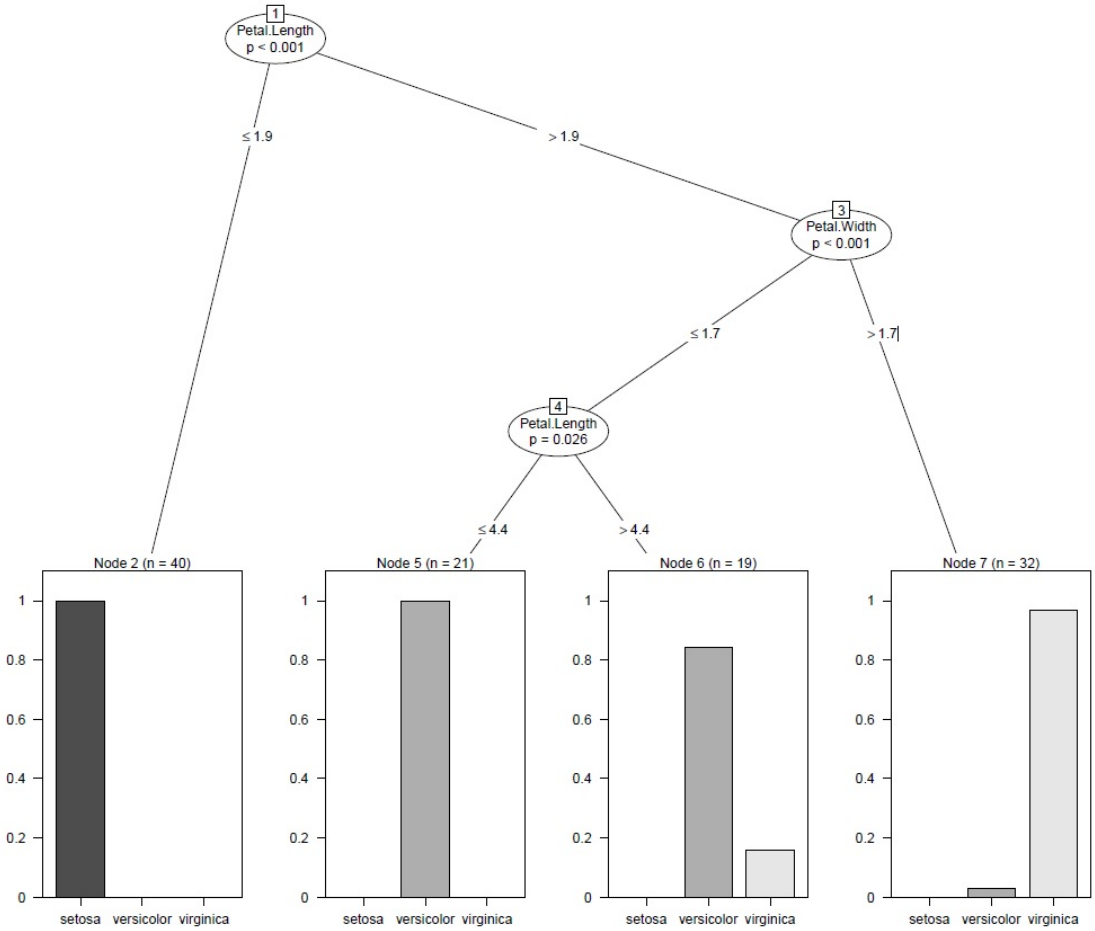


Figura 10. Algoritmo de classificação gerando uma Árvore de Decisão. Fonte: [Zhao 2012]

comparação com o conjunto normal de instâncias, o profissional responsável por detecção de anomalias pode considerar o conjunto erroneamente como um novo agrupamento normal, e então visualizar os dados com dois *clusters*.

Para considerar um grupo diferente do agrupamento geral, a distância entre os elementos nos grupos também pode ser considerada, sendo chamada de densidade. A diferença de densidade pode ser visualizada na Figura 11c, onde um grupo está próximo de outro, porém cada instância do grupo menor está mais próxima do centro.

Com o exemplo anterior, ao remover a indicação de um novo conjunto de anormalidades por densidade, notamos que o agrupamento está próximo ao agrupamento principal. Conforme demonstra a Figura 11d, isso gera possibilidades para o profissional de detecção de fraudes generalizar os grupos e tratá-los da mesma maneira, como um só.

Os autores [Ahmed et al. 2015] indicam a possibilidade de utilizar um conjunto de dados para construir fronteiras para novos dados, porém, é necessário eliminar manualmente os *outliers* do conjunto.

Como pode ser visualizado na Figura 11e, o padrão de identificação é criado a

partir dos dados atuais, e então pode ser aplicado em novos dados. Todas as instâncias que estiverem fora dos limites (circuladas em vermelho) serão consideradas anomalias. Caso um dado esteja presente entre os dois agrupamentos, conforme a representação da figura, ele será considerado normal, pois o grupo menor não foi removido para a criação da fronteira.

Finalmente, uma das tarefas que geralmente está presente nos algoritmos de agrupamento (métodos não-supervisionados) é indicar previamente o número de grupos que os dados possuem. No algoritmo *K-means*, este número é indicado pela letra k. Portando, a complexidade computacional dessa técnica é baixa, ocorrendo na ordem O(n), onde n representa o número de pontos no conjunto de dados [Ahmed et al. 2015].

Nas Figuras 11f e 11g, visualizamos o mesmo *dataset* com a separação em três e dois grupos, respectivamente. Dependendo do contexto, as duas formas podem estar corretas. Nessa situação, é responsabilidade do indivíduo conhecedor dos dados indicar o número k de *clusters* que as informações possuem. Para tentar solucionar o problema, o algoritmo *X-means* se propõe a identificar o número de grupos de uma maneira automática, porém é menos eficiente do que uma pessoa que conheça o conjunto dos dados [Ahmed et al. 2015].

4.3. Áreas de detecção de fraudes

Realizando uma pesquisa de artigos na área de detecção de fraudes com sistemas inteligentes no período de 1956 à 2009, [Pejic-Bach 2010] encontrou 36 artigos após a aplicação de seus filtros de pesquisa, que foram rigorosos para eliminar os mais de 2000 artigos encontrados e conseguir trabalhos extremamente relevantes.

Como um dos últimos critérios de exclusão adotados, [Pejic-Bach 2010] decidiu considerar apenas os trabalhos nos quais os autores possuíam mais de dez citações e eram autores de pelo menos três artigos similares.

Devido ao seu filtro divergente da maioria das revisões sistemáticas da área, [Pejic-Bach 2010] indicou que a grande parte dos artigos encontrados tratavam sobre a detecção de fraudes em sistemas de telecomunicação (dez artigos), seguido por fraudes em auditorias e seguradoras, posteriormente por sistemas financeiros e uma pequena parte em outras aplicações, como comércio eletrônico.

Outra razão do aparecimento dos sistemas de telecomunicação, auditorias e seguradores em maior número ao de sistemas financeiros pode ter sido o fato do período temporal da procura de artigos. Houve aumento significativo no uso de cartões de crédito somente após a década de 80.

[Abdallah et al. 2016] apresentam, conforme apresenta a Figura 12, as áreas que identificaram como mais recorrentes dentro do assunto de detecção de fraudes. Os autores consideraram detecção de fraude em tempo real. Portanto, nota-se a presença das áreas financeiras, como cartões de crédito, compras pela internet, seguradoras e sistemas de telecomunicação.

Nas seguintes sub-seções, serão categorizadas e generalizadas as áreas mais recorrentes nas revisões sistemáticas executadas que foram escolhidas por este trabalho.

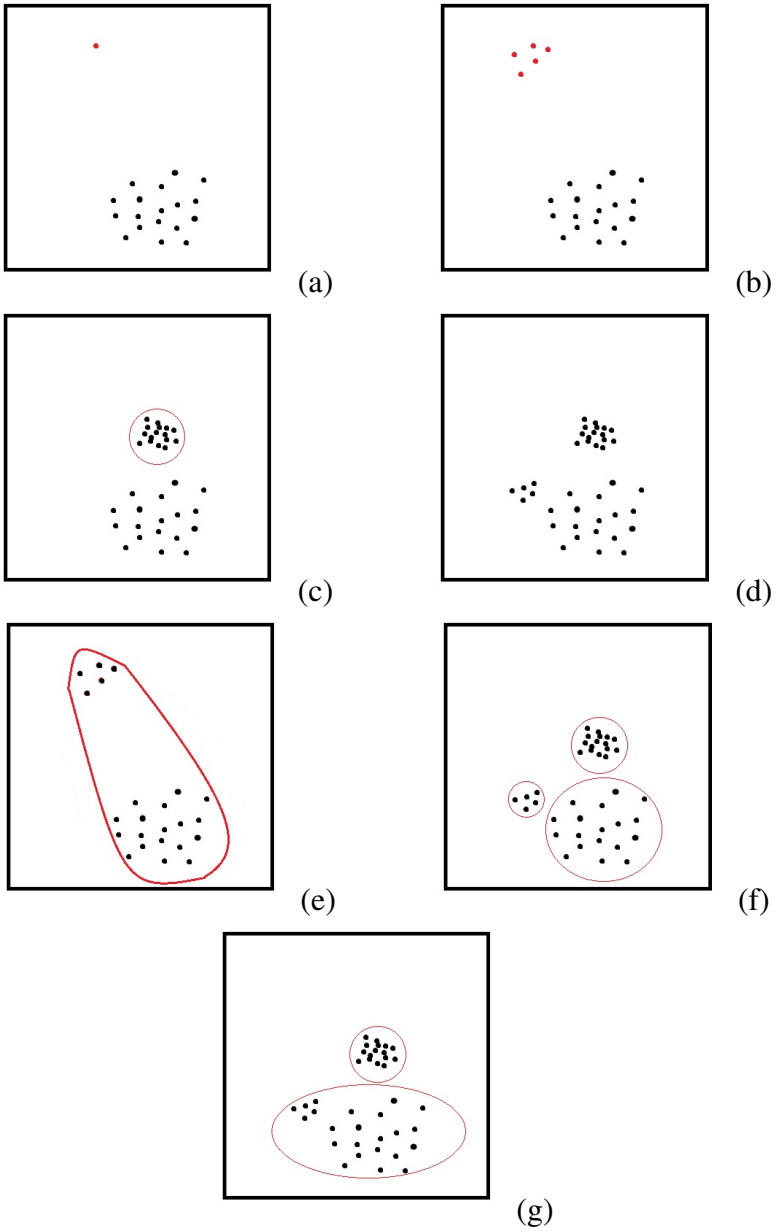


Figura 11. Tipos de anomalias

4.3.1. Sistemas financeiros e cartões de crédito

[Gullkvist and Jokipii 2013] apresentam a importância de indicadores de risco para auxiliar as corporações a detectar fraudes em setores financeiros. Segundo os autores, as fraudes podem ser prevenidas (saber que uma fraude está para ocorrer), detectadas (observar uma fraude que está ocorrendo) e investigadas (possuir o registro das fraudes que já ocorreram). Em qualquer caso, é importante que os sistemas forneçam aos responsáveis por categorizar ações fraudulentas o indício de risco na fase mais inicial possível.

Para setores específicos de detecção de fraudes, a identificação de anomalias em tempo real se demonstrou importante. [Edge and Falcone Sampaio 2009] realizaram um levamento das técnicas financeiras no contexto de detecção de fraudes.

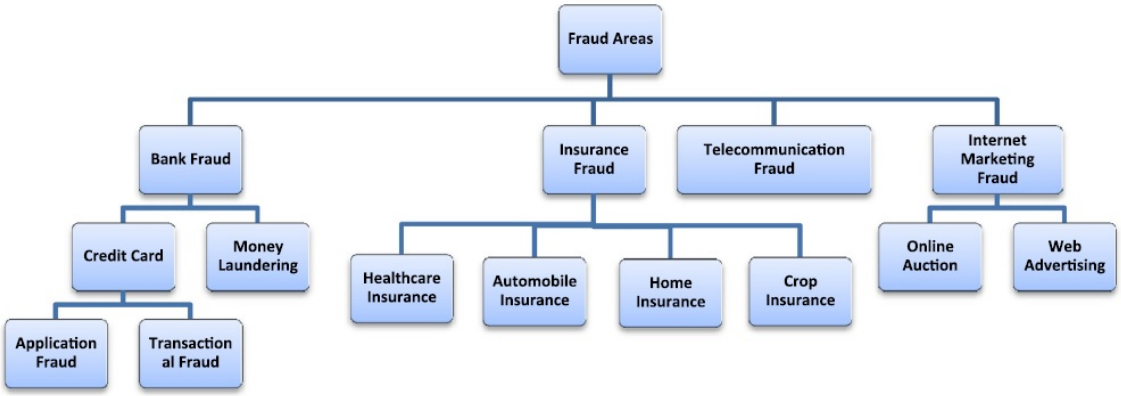


Figura 12. Áreas com grande concentração de fraudes. Fonte: [Abdallah et al. 2016]

Em cartões de crédito, o maior interesse nas pesquisas acontece no ramo de detecção de fraudes em tempo real, pois o prejuízo ocorrido após algum gasto ou transação indevida ser aprovada pode ser irreversível, ou a correção mais custosa do que o valor perdido.

Desenvolvendo um panorama geral na área, [Allan and Zhan 2010] compartilham seus resultados e constatam um problema para a literatura sobre o fato da detecção de fraudes em cartões de crédito possuir técnicas privadas para a indústria. Segundo os autores, as companhias bancárias não possuem interesse em compartilhar seus métodos para detectar fraudadores, pois essas informações poderiam ser utilizadas de forma indevida por criminosos.

[Perlich et al. 2007] e [Kanapickiene et al. 2015] realizaram trabalhos relacionando as áreas financeiras com a introdução de suas técnicas para detectar fraudes. Nas constatações, indicam que os indivíduos fraudadores devem ser considerados como internos ou externos à empresa, similarmente a categorização da *Basel Committee on Bank Supervision*. Por conter dois tipos de características, cada uma das abordagens de fraudes deve ser considerada distinta da outra. Em fraudadores internos, englobam-se os próprios funcionários e gerentes das organizações. Em usuários externos, são considerados os clientes mal intencionados ou criminosos.

4.3.2. Seguradoras

Empresas que prestam serviços como seguros para saúde e automóveis também são alvos de fraudadores.

[Pejic-Bach 2010] indica que a necessidade de resposta em tempo real é pouco relevante neste caso, já que os algoritmos irão trabalhar com dados que foram armazenados num grande período de tempo.

[Bauder et al. 2016] e [Li et al. 2008] realizaram revisões na literatura visando ampliar a abordagem de detecção de fraudes na área da saúde. Ambos os estudos constataram que a área financeira de detecção de fraudes é geralmente adaptada de forma genérica para ser utilizada no contexto da saúde.

Determinadas abordagens são ineficientes em certos casos, indicando uma necessidade de aprofundamento nos estudos. [Bauder et al. 2016] também indicam que são utilizadas técnicas supervisionadas, não-supervisionadas e híbridas no contexto de saúde, no qual geralmente as técnicas não-supervisionadas prevalecem, devido à dificuldade (na maioria das vezes ética) de obter determinados dados médicos.

4.3.3. Sistemas de telecomunicação

[Fawcett and Provost 1997] indicam que com o aumento no número de telefones celulares depois da década de 90, muitos usuários tiveram seus aparelhos clonados. A identificação de uma ligação de um cliente verdadeiro deve ser rapidamente diferenciada de uma fraude. Neste caso, a velocidade do algoritmo é importante, pois como no sistema de transações bancárias, o prejuízo é computado no momento que a fraude está ocorrendo.

4.3.4. Comércio eletrônico e vendas online

Tendo um relacionamento próximo com a utilização de cartão de crédito, o setor de comércio eletrônico está em constante crescimento. Segundo a *Internet Complaint Center*, os crimes de vendas online estão entre os mais perigosos [Abdallah et al. 2016].

Ainda segundo [Abdallah et al. 2016], é importante ressaltar que uma fraude presente em alguma compra online pode ser feita de ambos os lados: tanto de um comprador, quanto de um vendedor. O grande crescimento de sites como o *eBay*, no qual um comprador e vendedor podem ser potenciais criminosos, gera uma preocupação sobre a veracidade das transações online.

Um fator considerável para o ramo de compras online fraudulentas serem tão frequentes é a facilidade de execução, pois os equipamentos necessários para a produção da fraude são de nível técnico básico, como um computador com internet.

4.3.5. *Big data* e redes sociais

[Feily et al. 2009], [Mahmood and Afzal 2013] e [Sharma and Mangat 2015] relacionam o crescimento do termo *big data* com a necessidade de inspecionar os dados para detectar problemas.

A maior dificuldade, segundo os autores, é gerenciar a grande quantidade de informação - geralmente desordenada - que a nova abordagem da computação produz. [Sharma and Mangat 2015] ainda indicam em suas pesquisas que os próprios dados provenientes de *big data* podem ser utilizados para auxiliar mecanismos de detecção de fraudes. Para ilustrar a situação anterior, [Liu and Chawla 2015] e [Yu et al. 2016] discutem que as novas abordagens em detecção de fraudes provavelmente acontecerão com mais frequência em redes sociais. Com uma grande quantidade de dados para analisar, as novas ferramentas de comunicação proporcionam oportunidades para detectar grupos de pessoas e premeditar suas atitudes, podendo ser utilizadas para prevenir ataques terroristas, por exemplo. Os autores ressaltam que as técnicas de anomalias pontuais e grupos de anomalias estão presentes nessa área, onde os algoritmos baseados em grafos são os mais

utilizados. [Pejic-Bach 2010] e [Allan and Zhan 2010] também pontuam que ameaças terroristas podem ser visualizadas com a análise de dados através de redes sociais.

4.4. Resposta às principais perguntas

Após o término da revisão sistemática, as questões de pesquisas levantadas na Seção de metodologia podem ser respondidas. As constatações estão presentes abaixo:

- **QP1:** *Quais são as áreas de detecção de fraudes mais estudadas na literatura?*
  A maioria dos trabalhos obteve as seguintes áreas de fraudes em comum: bancos, dispositivos de telecomunicação, seguradoras, comércio eletrônico, leilões virtuais, mercado de ações e dados jurídicos. Os autores, com seus respectivos países, que citaram determinadas áreas estão relacionados nas Tabelas 5 e 6.

- **QP2:** *Como a literatura categoriza as técnicas de detecção de fraudes?*
  Genericamente, os algoritmos para detectar fraudes realizam cálculos em dados supervisionados e não-supervisionados. Em métodos baseados em dados supervisionados, o pesquisador já possui a informação de qual dado é fraudulento. Em métodos baseados em dados não-supervisionados, ocorre o contrário. Há alguns autores que propõem uma abordagem híbrida ou por meio de regras.

- **QP3:** *Quais foram os problemas mais relatados pelos autores?*
  Algumas áreas específicas não possuem uma abordagem mais aprofundada. Geralmente, os pesquisadores utilizam técnicas genéricas para detectar fraudes nos mais variados contextos. Um exemplo é a grande utilização de técnicas das áreas econômicas e financeiras (exaustivamente estudadas) sendo aplicada para detectar fraudes em sistemas de saúde.

- **QP4:** *Como os autores testaram e validaram suas pesquisas?*
  Os autores utilizaram dados variados para cada contexto, como dados para técnicas supervisionadas e não-supervisionadas. Para comparar e validar com outros algoritmos, os autores executam técnicas conhecidas da literatura em dados de sua escolha e, após, repetem o teste utilizando seus novos métodos para comparação de valores. Geralmente, os algoritmos são comparados utilizando bases de dados fornecidas na internet para estudos, porém alguns autores recorrem a dados reais para obter mais veracidade em seus resultados.

- **QP5:** *Em qual área há pouca abordagem no estudo de detecção de fraudes?*
  Em setores específicos e que estão em recente crescimento de interesse pela comunidade. Para padronizar os resultados de futuros autores, dados de *benchmarks* seriam possibilidades para novos estudos. A criação de um padrão mínimo a ser seguido por autores para realizar experimentos seria um passo importante para a área, pois além de não possuírem dados padronizados, os estudos anteriores não mostram uma uniformidade na execução.

- **QP6:** *Há espaço para futuras pesquisas na área?*
  Sim. Embora os principais tópicos tenham sido muito discutidos, melhorias gerais nas técnicas ainda são possíveis. Reduzir a incidência de falsos negativos e

Tabela 5. Áreas mais abordadas pelos autores no período entre 2006 e 2016.

Áreas	Autores	País
Bancos ou cartões de crédito	[Abdallah et al. 2016]	Malásia
	[Ahmed et al. 2015]	Austrália
	[Akoglu et al. 2015]	EUA
	[Allan and Zhan 2010]	EUA
	[Bansal 2016]	Índia
	[Branco 2016]	Portugal
	[Edge and Falcone Sampaio 2009]	Reino Unido
	[Pejic-Bach 2010]	Croácia
	[Perlich et al. 2007]	EUA e Brasil
	[Raj and Portia 2011]	Índia
Seguradoras	[Abdallah et al. 2016]	Malásia
	[Ahmed et al. 2015]	Austrália
	[Akoglu et al. 2015]	EUA
	[Bansal 2016]	Índia
	[Branco 2016]	Portugal
	[Pejic-Bach 2010]	Croácia
Cuidados da saúde	[Abdallah et al. 2016]	Malásia
	[Bansal 2016]	Índia
	[Bauder et al. 2016]	EUA
	[Li et al. 2008]	EUA
	[Pejic-Bach 2010]	Croácia
Telecomunicação	[Abdallah et al. 2016]	Malásia
	[Ahmed et al. 2015]	Austrália
	[Akoglu et al. 2015]	EUA
	[Allan and Zhan 2010]	EUA
	[Branco 2016]	Portugal
	[Pejic-Bach 2010]	Croácia
Comércio eletrônico	[Abdallah et al. 2016]	Malásia
	[Akoglu et al. 2015]	EUA
	[Pejic-Bach 2010]	Croácia
Leilões virtuais	[Abdallah et al. 2016]	Malásia
	[Pejic-Bach 2010]	Croácia
Mercados de ações	[Ahmed et al. 2015]	Austrália
	[Pejic-Bach 2010]	Croácia

falsos positivos sem gasto excessivo de recursos humanos e computacionais são esperados para o futuro da área.

Junto a esse fator, realizar a detecção de fraude em tempo real, coibindo os falsos negativos e falsos positivos de maneira satisfatória também é almejado para novas publicações. Finalmente, como mencionado por autores da área, a adaptação das

Tabela 6. Áreas mais abordadas pelos autores no período entre 2006 e 2016.

Áreas	Autores	País
Dados financeiros ou empresariais	[Akoglu et al. 2015]	EUA
	[Allan and Zhan 2010]	EUA
	[Branco 2016]	Portugal
	[Edge and Falcone Sampaio 2009]	Reino Unido
	[Flegel et al. 2010]	Alemanha e Austrália
	[Gullkvist and Jokipii 2013]	Finlândia
	[Kanapickiene et al. 2015]	Lituânia
	[Pejic-Bach 2010]	Croácia
	[Wang 2010]	China
Redes sociais	[Feily et al. 2009]	Malásia
	[Liu and Chawla 2015]	EUA e Austrália
	[Rebahi et al. 2011]	Alemanha e França
	[Yu et al. 2016]	EUA
Big data	[Mahmood and Afzal 2013]	Paquistão
	[Sharma and Mangat 2015]	Índia
Redes de computadores	[Akoglu et al. 2015]	EUA
	[Allan and Zhan 2010]	EUA
	[Branco 2016]	Portugal
Equipamentos industriais	[Bauder et al. 2016]	EUA
Terrorismo	[Allan and Zhan 2010]	EUA
	[Pejic-Bach 2010]	Croácia

técnicas de detecção de fraudes amplamente estudadas, como setores financeiros, podem ser aplicadas em setores biológicos ou da medicina.
Finalmente, a etapa de pré-processamento foi pouco discutida entre os trabalhos. Aliado a esse fator, maioria das pesquisas possuíam dificuldades para comparação por não terem sido executadas da mesma maneira, abrindo possibilidade para criação de práticas para sugerir rotinas em trabalhos futuros, como a utilização de *datasets* similares.

5. Discussão

Nas seções anteriores, pode-se observar o estado das pesquisas relacionadas à detecção de fraudes. Conforme as observações realizadas, esta Seção visa discutir os fatores em comum entre os trabalhos e os dados obtidos.

O principal desafio apontado pela maioria dos autores encontrados pela revisão sistemática é a evolução das técnicas utilizadas pelos fraudadores para burlar os sistemas.

Em estudos constatados por corporações bancárias, o comportamento esperado por fraudadores mostrou uma mutação quando os criminosos entenderam as técnicas para detectá-los [Bolton et al. 2002]. Indivíduos que outrora utilizavam cartões de crédito somente para efetuar compras com a intenção de não pagá-las, perceberam que poderiam

pagar compras iniciais para serem classificados como usuários comuns. Dessa maneira, possuiriam maior margem para fraudar no futuro.

Entendendo o problema de mutação nas técnicas criminosas, constatamos um contraste com a engenharia de software tradicional, que possui um processo de produção padrão, no qual ocorrem etapas de análise, projeto, codificação, implementação e testes [Sommerville 2011]. Dessa maneira, a constante evolução das características de usuários fraudulentos pode se tornar um desafio para ser trabalhado em um modelo de desenvolvimento tradicional de software.

É importante ressaltar que a maioria dos estudos indicou que a verificação final se algum dado representa ou não uma fraude é uma tarefa recomendada para um humano, que pode ser auxiliado por uma ferramenta computacional, como um algoritmo.

Ao analisar os trabalhos desenvolvidos na área de detecção de fraudes, não foi possível constatar a presença de um padrão no conjunto de dados que pudesse ser utilizado por diversos algoritmos para computação e comparação de valores. Em seus trabalhos, os autores utilizaram os mais variados tipos de dados, sendo possível identificar duas grandes características em comum: houveram autores que utilizaram dados controlados, possivelmente criado por eles; e dados não fictícios, em que utilizaram suas técnicas e algoritmos em informações recebidas de empresas e problemas reais.

O problema, porém, está no fato que os autores comparam os seus algoritmos com os já existentes na literatura utilizando os dados que lhes são convenientes. Esta revisão sistemática propõe como sugestão para trabalhos futuros desenvolver e apresentar um conjunto de dados que possa servir como um *benchmark*, que diversos pesquisadores possam realizar seus estudos nos mesmos ambientes ou com as mesmas condições, com a finalidade de obter testes padronizados.

O tempo de processamento dos algoritmos pode ser ignorado para alguns casos, como a maioria dos algoritmos de clusterização. Em alguns sistemas de classificação, como os que necessitam do aviso imediato de alguma fraude (transação bancária, telefonema clonado, etc), o tempo computacional pode ser relevante.

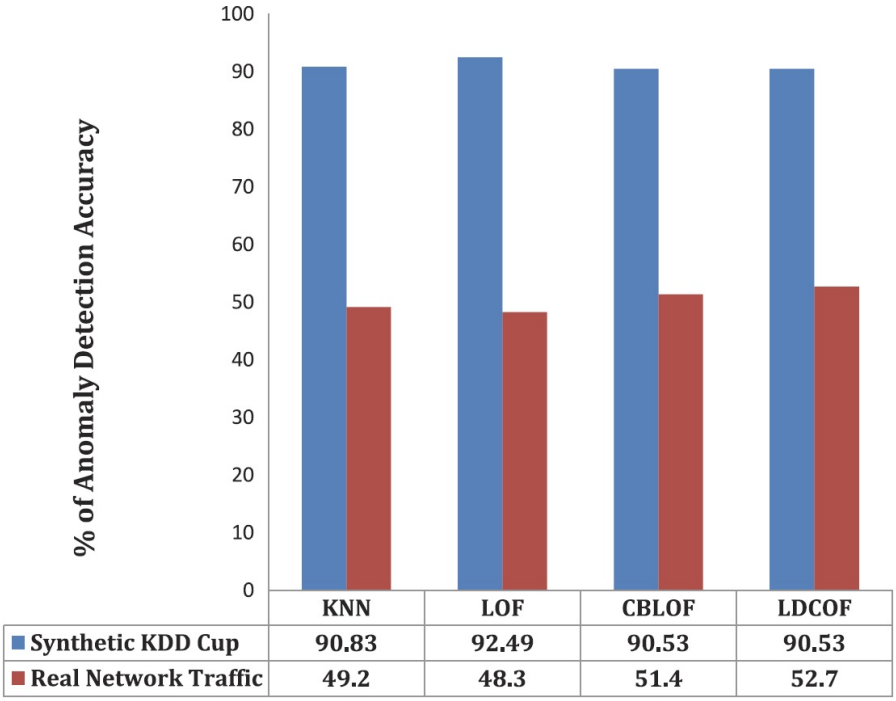
A revisão sistemática deste trabalho não encontrou na literatura estudos específicos para orientar a etapa de pré processamento. Das ferramentas gratuitas e de código aberto mais utilizadas para mineração de dados (Weka[9] e linguagem R[10]), nenhuma apresenta um *framework* dedicado eficiente para pré-processamento. Além das ferramentas gratuitas, [Ahmed et al. 2015] citam eu seu trabalho o software proprietário *SAS*[11], que possui módulos específicos para detecção de fraudes.

O trabalho de revisão sistemática na área de detecção de fraudes que mais se assemelhou ao propósito deste trabalho foi o elaborado por [Ahmed et al. 2015]. Neste trabalho, os autores focaram em detecção de fraudes em sistemas financeiros e compreenderam que a maioria dos casos ocorrem com cargas de dados não-supervisionadas e com muito ruído, por serem do contexto real.

Os autores elencaram que a utilização de dados do mundo real é importante, pois

[9] `http://www.cs.waikato.ac.nz/ml/weka/`
[10] `https://www.r-project.org/`
[11] `http://www.sas.com/en_us/industry/banking/fraud-management.html`

os resultados de algoritmos são voláteis em dados sintéticos. Na Figura 13, pode-se perceber que ao aplicar os algoritmos KNN, LOF, CBLOF e LDCOF em bases de dados sintéticas e em bases de um contexto real (no caso do exemplo, foram utilizadas as informações de tráfego de rede), os resultados se mostram muito mais eficientes em dados gerados de uma maneira mais controlada.



	KNN	LOF	CBLOF	LDCOF
Synthetic KDD Cup	90.83	92.49	90.53	90.53
Real Network Traffic	49.2	48.3	51.4	52.7

Figura 13. Comparação entre algoritmos executando em bases do mundo real e bases geradas de maneiras sintética. Fonte: [Ahmed et al. 2015]

6. Conclusão e trabalhos futuros

Este artigo apresentou os resultados de uma revisão sistemática realizada na área de detecção de fraudes. Pesquisas realizadas para dar o panorama na área foram estudadas e discutidas. Ao obter os principais artigos publicados na área, constatamos um crescimento no índice de publicações na área acadêmica após o ano de 2010. Isto mostra que o assunto está em ascensão de interesse.

Em conclusão, o trabalho indica que a área, embora muito estudada, ainda concentra amplo poder para novas pesquisas. Um setor de pesquisa em ascensão é a adaptação das técnicas difundidas e testadas com experiências positivas para contextos específicos, como dados governamentais.

Como a maioria dos estudos constataram que a influência para algum desenvolvimento científico em alguma área de detecção de fraude vem do interesse financeiro, os sistemas corporativos bancários, de prestação de seguros e de telecomunicações foram os mais citados, pois a geração de capital desses setores é alta.

Posteriormente, o setor de *e-commerce* vem despertando interesse acadêmico e comercial, pois é considerado uma área relativamente nova, se comparada ao sistema de cartão de crédito. Também podemos lembrar que o comércio virtual consegue englobar

muitas categorias citadas como áreas de possíveis fraudadores, com a possibilidade de generalização dos setores.

Junto com a importância do assunto verificada após os estudos, grande parte dos autores destacaram que o problema computacional ocorrido em todos os casos é a não detecção de todos os falsos negativos (em que ações fraudulentas não são detectadas), bem como a classificação equivocada de falsos positivos (ações legítimas caracterizadas, erroneamente, como fraudulentas) com aceitável gasto de recurso. Desta forma, novas abordagens são necessárias para a continuidade do tema.

Como em praticamente todos os casos, a decisão entre dados fraudulentos e não-fraudulentos é atribuída para um humano, os algoritmos e técnicas servem como ferramentas.

Com a revisão sistemática, não foi encontrado um conjunto de ações gerais para os detectores de fraudes utilizarem as ferramentas disponíveis na literatura. Os autores, geralmente, publicam suas descobertas de maneira individual e com uma separação entre o meio acadêmico e profissional.

Comparando todos os trabalhos estudados, houve unanimidade no tópico no que diz respeito ao conhecimento dos dados e a eficiência da mineração e detecção de fraudes. Quanto mais o responsável por conduzir o processo de geração de informação a partir da mineração de dados conhece e sabe interpretar as instâncias e os atributos, melhor será a escolha das técnicas adequadas para o processo e também a sua interpretação do resultado final.

Em trabalhos futuros, pretende-se continuar a pesquisa em detecção de fraudes, aplicando técnicas para auxiliar órgãos públicos a descobrir irregularidades em dados de auditoria e controle e realizando testes no contexto real.

Referências

Abdallah, A., Maarof, M. A., and Zainal, A. (2016). Fraud detection system: A survey. *Journal of Network and Computer Applications*, 68:90–113.

Ahmed, M., Mahmood, A. N., and Islam, M. R. (2015). A survey of anomaly detection techniques in financial domain. *Future Generation Computer Systems*, 55:278–288.

Akoglu, L., Tong, H., and Koutra, D. (2015). *Graph based anomaly detection and description: A survey*, volume 29.

Allan, T. and Zhan, J. (2010). Towards fraud detection methodologies. *2010 5th International Conference on Future Information Technology, FutureTech 2010 - Proceedings*.

Bansal, R. (2016). Outlier Detection : Applications and Techniques in Data Mining.

Bauder, R., Khoshgoftaar, T. M., and Seliya, N. (2016). A survey on the state of healthcare upcoding fraud analysis and detection. *Health Services and Outcomes Research Methodology*, pages 1–25.

Bolton, R. J., Hand, D. J., Provost, F., Breiman, L., Bolton, R. J., and Hand, D. J. (2002). Statistical Fraud Detection: A Review. *Statistical Science*, 17(3):235–255.

Branco, P. (2016). A Survey of Predictive Modeling on Imbalanced Domains. 49(2):1–50.

Chan, P. K., Fan, W., Prodromidis, A. L., and Stolfo, S. J. (1999). Distributed data mining in credit card fraud detection. *IEEE Intelligent Systems and their Applications*, 14(6):67–74.

Chandola, V., Banerjee, A., and Kumar, V. (2009). Anomaly Detection: A Survey. *ACM Comput. Surv.*, 41(3):15:1—-15:58.

Edge, M. E. and Falcone Sampaio, P. R. (2009). A survey of signature based methods for financial fraud detection. *Computers and Security*, 28(6):381–394.

Fawcett, T. and Provost, F. (1997). Adaptive Fraud Detection. *Data Mining and Knowledge Discovery*, 1(3):291–316.

Fayyad, U. M., Piatetsky-Shapiro, G., and Smyth, P. (1996). Advances in knowledge discovery and data mining. chapter From Data Mining to Knowledge Discovery: An Overview, pages 1–34. American Association for Artificial Intelligence, Menlo Park, CA, USA.

Feily, M., Shahrestani, A., and Ramadass, S. (2009). A survey of botnet and botnet detection. *Proceedings - 2009 3rd International Conference on Emerging Security Information, Systems and Technologies, SECURWARE 2009*, pages 268–273.

Flegel, U., Vayssière, J., and Bitz, G. (2010). A State of the Art Survey of Fraud Detection Technology. *Insider Threats in Cyber Security*, 49:73–84.

Gullkvist, B. and Jokipii, A. (2013). Perceived importance of red flags across fraud types. *Critical Perspectives on Accounting*, 24(1):44–61.

Hodge, V. J. and Austin, J. (2004). A survey of outlier detection methodologies. *Artificial Intelligence Review*, 22(2):85–126.

Kanapickiene, R., Grundiene, Ž., Gimzauskiene, E., Duoba, K., Pavie, X., Pinnington, A., Vilkas, M., Kanapickiene, R., Grundiene, Ž., Kanapickiene, R., and Grundiene, Ž. (2015). The Model of Fraud Detection in Financial Statements by Means of Financial Ratios. *Procedia - Social and Behavioral Sciences*, 213:321–327.

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University*, 33(TR/SE-0401):28.

Kitchenham, B. and Charters, S. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering.

Li, J., Huang, K.-Y., Jin, J., and Shi, J. (2008). A survey on statistical methods for health care fraud detection. *Health care management science*, 11(3):275–287.

Liu, Y. and Chawla, S. (2015). Social Media Anomaly Detection : Challenges and Solutions. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2317–2318.

Mahmood, T. and Afzal, U. (2013). Security Analytics: Big Data Analytics for Cybersecurity. *2013 2nd National Conference on Information Assurance (NCIA)*, pages 129–134.

Pejic-Bach, M. (2010). Profiling intelligent systems applications in fraud detection and prevention: Survey of research articles. *ISMS 2010 - UKSim/AMSS 1st International Conference on Intelligent Systems, Modelling and Simulation*, pages 80–85.

Perlich, C., Rosset, S., Lawrence, R. D., and Zadrozny, B. (2007). High-quantile modeling for customer wallet estimation and other applications. *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '07*, page 977.

Raj, S. B. E. and Portia, a. A. (2011). Analysis on credit card fraud detection methods. *2011 International Conference on Computer, Communication and Electrical Technology (ICCCET)*, pages 152–156.

Rebahi, Y., Nassar, M., Magedanz, T., and Festor, O. (2011). A survey on fraud and service misuse in voice over IP (VoIP) networks. *Information Security Technical Report*, 16(1):12–19.

Seyedhossein, L. and Hashemi, M. R. (2010). Mining information from credit card time series for timelier fraud detection. In *Telecommunications (IST), 2010 5th International Symposium on*, pages 619–624.

Sharma, S. and Mangat, V. (2015). Technology and Trends to Handle Big Data: Survey. *2015 Fifth International Conference on Advanced Computing {&} Communication Technologies*, pages 266–271.

Sommerville, I. (2011). *Engenharia de Software*. Pearson Brasil, 9th edition.

Tan, P.-N., Steinbach, M., and Kumar, V. (2005). *Introduction to Data Mining, (First Edition)*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

Wang, S. (2010). A comprehensive survey of data mining-based accounting-fraud detection research. *2010 International Conference on Intelligent Computation Technology and Automation, ICICTA 2010*, 1:50–53.

Wongchinsri, P. and Kuratach, W. (2016). A survey - data mining frameworks in credit card processing. In *2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pages 1–6.

Yu, R., Qiu, H., Wen, Z., Lin, C.-Y., and Liu, Y. (2016). A Survey on Social Media Anomaly Detection. *arXiv preprint*, 18(1):18.

Zhao, Y. (2012). *R and Data Mining: Examples and Case Studies*. Academic Press, Elsevier.