

# Detecção de Fraudes – Uma Revisão Sistemática

**Aluno:** Jean Avila Rangel

**Orientadores:** Adolfo Neto e Maria Claudia Emer

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
(UTFPR)**

*Metodologia de Pesquisa*

1

Detecção de Fraudes – Uma Revisão Sistemática

*Curitiba, 2016*

# Agenda

- Tema;
- Motivação e justificativa;
- Objetivos;
  - Geral;
  - Específicos.
- Metodologia;
- Referencial teórico e estado da arte;
- Indicação da Revista;
- Cronograma;
- Conclusão.

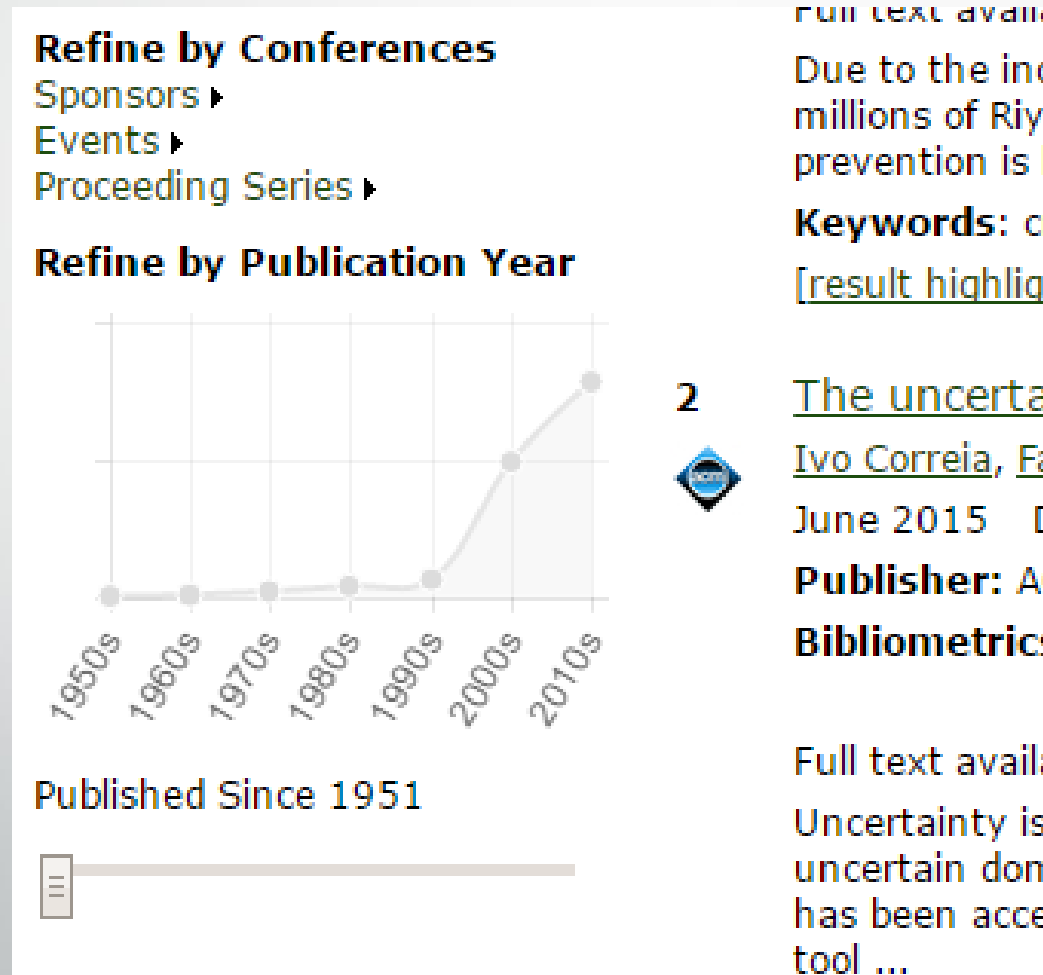
# Tema

- Uma atividade fraudulenta se caracteriza pela ação de um ou mais pessoas para ganhar proveito individual de determinada situação [1];
- A detecção de fraudes é uma área que visa encontrar falhas ou comportamentos inesperados em dados [4];
- O estudo da detecção de fraude pode englobar muitas áreas, como cartões de crédito e sistemas médicos.

# Motivação e justificativa

- Estudar uma área que está em constante acréscimo de importância no Estado da Arte;
- Há um grande número de revisões sistemáticas na área;
- Em trabalhos futuros, pretende-se aplicar ferramentas para detectar fraudes em dados de auditoria e controle de órgãos públicos.

# Gráfico de publicações na ACM



# **Resultados em bases com as palavras-chave “Fraud Detection Survey” no período entre 2006 e 2016:**

- **ACM Digital Library – 8 resultados;**
- **IEEEXplore – 51 resultados;**
- **ScienceDirect – 14 resultados;**
- **Springer Link – 170 resultados;**
- **Google Scholar – 17.600 resultados.**

# Objetivos

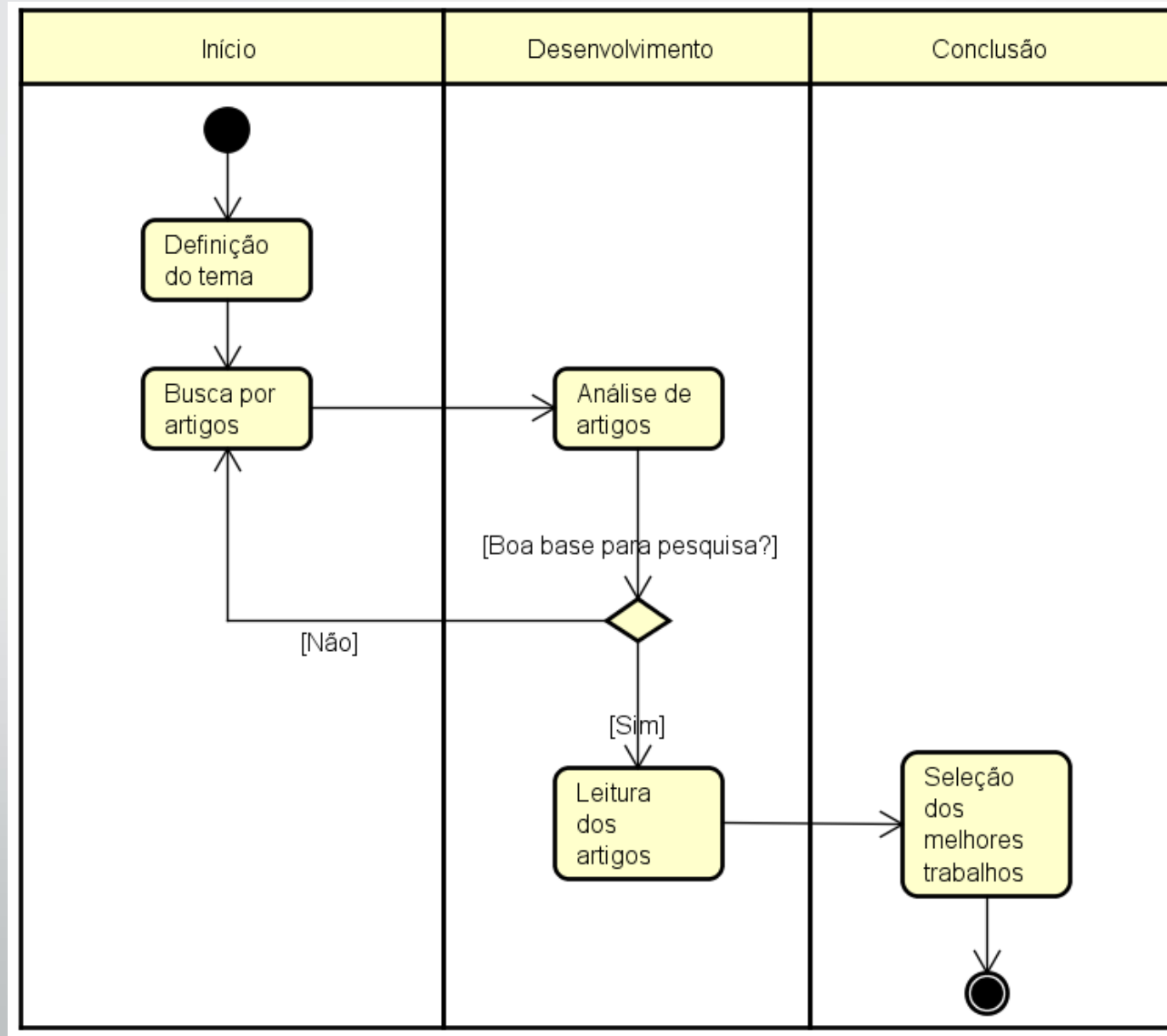
- Geral;
  - O objetivo geral deste trabalho é identificar, categorizar e agrupar técnicas e ferramentas para detecção de fraudes.
- Específicos.
  - Examinar se os estudos realizados em detecção de fraudes possuem abertura para áreas pouco ou não abordadas. Caso haja necessidade, o trabalho poderá categorizar possíveis trabalhos futuros em lacunas no assunto;
  - Elencar como os autores validaram suas pesquisas nas áreas de detecção de fraudes;
  - Categorizar quais tecnologias foram as mais utilizadas para cada contexto.

# Metodologia

- Realizar um levantamento no Estado da Arte buscando aproximar do objetivo encontrado [5];
- Desenvolver a procura em bases de dados globais na área da informática científica.



# Método de Pesquisa

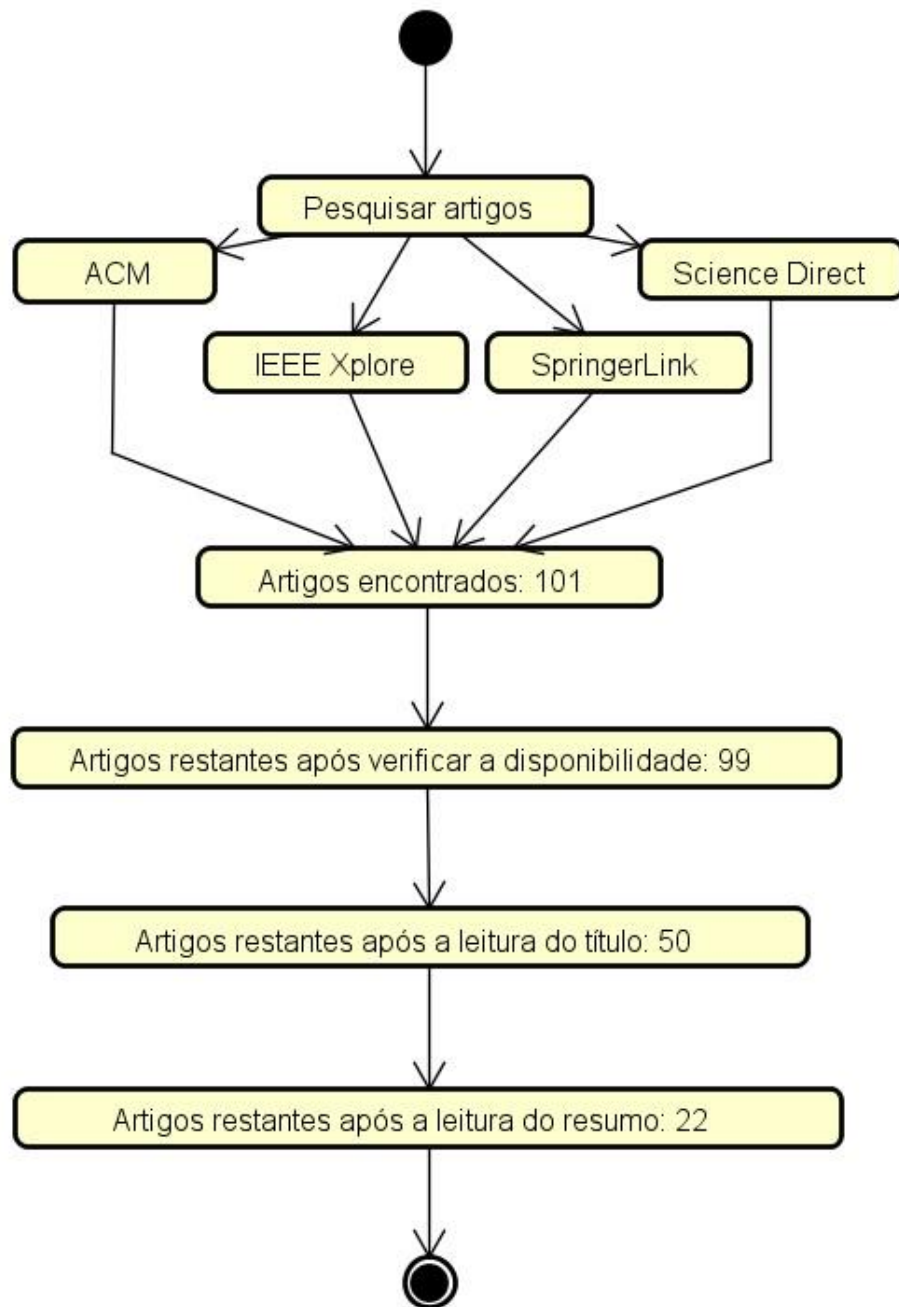


# Critérios para exclusão de artigos

- Foram considerados somente trabalho com relevância científica, na língua inglesa e no período de 2006 a 2016;
- Após a leitura do título, resumo e artigo completo, foram excluídos os artigos que possuíam como tema central:
  - 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.

# Foram desconsideradas a utilização das palavras chave

- 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.



Categoria	Quantia	Porcentagem
Conferência	12	55%
Revista	10	45%

# Questões de pesquisa (QP)

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

# Agenda

- Tema;
- Motivação e justificativa;
- Objetivos;
  - Geral;
  - Específicos.
- Metodologia;
- **Referencial teórico e estado da arte;**
- Cronograma;
- Indicação da Revista.

# Objetivos da utilização de sistemas para detectar fraudes

- Redução de custos com usuários maliciosos [6];
- Alertas de dados errôneos ou detentores de ruídos;
- Visualização de possíveis padrões indesejados;
- Melhor monitoramento das informações.

# Técnicas utilizadas

- Baseada em aprendizado (criação de alarmes);
- Grafos:
  - Aprendizado supervisionado (métodos de classificação. Ex: redes neurais);
  - Aprendizado não-supervisionado (métodos de *clusterização*. Ex: *K-Means*);
  - Aprendizado híbrido.



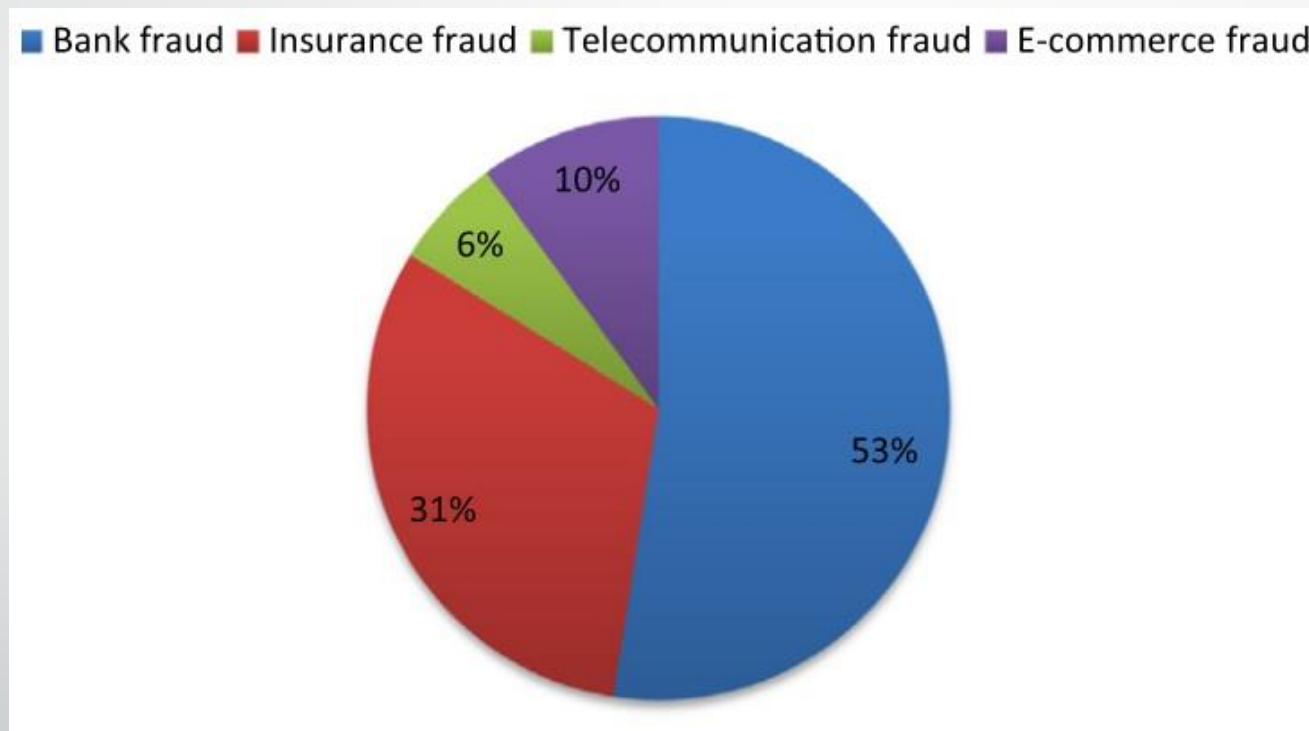
# Tipos de anomalias

- Indivíduo fora do padrão;
- Indivíduo fora do padrão por contexto;
- Grupo fora do padrão.

# Categorias e técnicas utilizadas em métodos supervisionados e não-supervisionados

Método	Categoria	Técnica
Supervisionado	Classificação	Árvores de decisões Redes neurais artificiais Modelo oculto de Markov Inferência Bayesiana
Não-supervisionado	Clusterização	Lógica Fuzzy K-Means

# Áreas mais abordadas em detecção de fraudes



- Fonte: Abdallah et al (2016) [1]

Áreas	Autores
Cartão de crédito	Abdallah et al [2016], Bansal et al [2016], Ahmed et al [2015], Edge and Falcone Sampaio [2009], Allan and Zhan [2010], Pejic-Bach [2010], Raj and Portia [2011], Perlich et al [2007], Kanapickienė et al [2015], Branco [2016], Akoglu et al [2015]
Seguradoras	Abdallah et al [2016], Bansal et al [2016], Ahmed et al [2015], Pejic-Bach [2010], Branco [2016], Akoglu et al [2015]
Cuidados da saúde	Abdallah et al [2016], Bauder et al [2016], Li et al [2008], Bansal et al [2016], Pejic-Bach [2010]
Telecomunicação	Abdallah et al [2016], Ahmed et al [2015], Allan and Zhan [2010], Pejic-Bach [2010], Pejic-Bach [2010], Branco [2016], Akoglu et al [2015]
Comércio eletrônico	Abdallah et al [2016], Pejic-Bach [2010], Branco [2016], Akoglu et al [2015]
Leilões virtuais	Abdallah et al [2016]
Mercados de ações	Ahmed et al [2015]
Dados financeiros ou empresariais	Flegel et al [2010], Edge and Falcone Sampaio [2009], Gullkvist and Jokipii [2013], Allan and Zhan [2010], Pejic-Bach [2010], Wang [2010], Branco [2016], Akoglu et al [2015]
Redes sociais	Yu et al [2016], Liu and Chawla [2015], Feily et al [2009], Rebahi et al [2011]
Big data	Sharma and Mangat [2015], Mahmood and Afzal [2013]
Redes de computadores	Allan and Zhan [2010], Branco [2016], Akoglu et al [2015]
Equipamentos industriais	Bauder et al [2016]
Terrorismo	Allan and Zhan [2010]

# Indicação da revista

- Foco em métodos e técnicas para mineração de dados.

**Data Mining and Knowledge Discovery** (Qualis Capes B1)

<http://link.springer.com/journal/10618>

# Cronograma

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

# Conclusão: respostas para as questões de pesquisa (QP)

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

# Referências

- [1] Abdallah A, Maarof MA, Zainal A (2016) Fraud detection system: A survey. Journal of Network and Computer Applications 68:90 – 113;
- [2] 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;
- [3] Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: A survey. ACM Comput Surv 41(3):15:1–15:58;
- [4] Fawcett T, Provost F (1997) Adaptive fraud detection. Data Mining and Knowledge Discovery 1(3):291–316;
- [5] Kitchenham B (2004) Procedures for performing systematic reviews. Keele, UK, Keele University.
- [6] 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.