

# Uma Introdução ao Combate Automático às *Fake News* em Redes Sociais Virtuais

Paulo Márcio Souza Freire  
Ronaldo Ribeiro Goldschmidt

Instituto Militar de Engenharia  
Simpósio Brasileiro de Banco de Dados 2019 (34ª edição)

25 de setembro de 2019

- 1 Considerações Iniciais
- 2 Fundamentos
  - Definição de Fake News
  - Comportamento disseminativo das Fake News
- 3 Trabalhos Relacionados
  - Proposta de Modelo Comparativo
  - Revisão dos Trabalhos Relacionados
  - Datasets
- 4 Problemas em Aberto
- 5 Estudo de Caso em Detecção Automática de Fake News
  - Fake News Detection in Social Networks via Crowd Signals
- 6 Considerações Finais
- 7 Referências

Considerações Iniciais

Fundamentos

Trabalhos Relacionados

Problemas em Aberto

Estudo de Caso em Detecção Automática de Fake News

Considerações Finais

Referências

# Considerações Iniciais

## • Redes Sociais

- Crescimento e importância [Vosoughi et al., 2017]
- Facilidade de divulgação [Wang et al., 2018]
- A ameaça das Notícias Falsas [Conroy et al., 2015]
- O desafio das *Fake News* [Shu et al., 2017]

## ● Poder de influência das *Fake News* na sociedade

[Farajtabar et al., 2017]

- Notícias falsas relacionadas às eleições presidenciais americanas - 2016
  - As notícias falsas publicadas nos últimos 3 meses, no *Facebook*, tiveram 37 milhões de compartilhamentos
  - As 20 principais notícias, criadas por sites fraudulentos, tiveram 1,5 milhões de engajamentos
- Terror em uma pizzeria na capital *Washington* - 2016
- Notícias falsas teriam levado a linchamentos na Índia (*WhatsApp*) - 2018

- **Necessidade de Combate Automático às *Fake News* nas Redes Sociais** [Flintham et al., 2018] [Wang et al., 2018] [Zhou et al., 2019] [Campan et al., 2017] [Kshetri and Voas, 2017]

- Volume
- Velocidade

Considerações Iniciais

**Fundamentos**

Trabalhos Relacionados

Problemas em Aberto

Estudo de Caso em Detecção Automática de Fake News

Considerações Finais

Referências

Definição de Fake News

Comportamento disseminativo das Fake News

# Fundamentos

## • Definição de *Fake News*

- Publicações intencionalmente e verificadamente falsas (Disinformation)  
[Reis et al., 2019] [Zhou et al., 2019]
- Publicações falsas, independente da sua natureza intencional  
[Sharma et al., 2019] [Castelo et al., 2019]



**Áreas Correlatas à *Fake News*** [Vosoughi et al., 2017]  
[Li et al., 2015] [Shu et al., 2017] [Braz and Goldschmidt, 2017]  
[Cazalens et al., 2018] [Hendrikx et al., 2015]

- Classificação de Rumores
- Descoberta da Verdade
- Detecção de *Clickbait*
- Detecção de *Bots*
- Checagem de fatos
- Sistemas de Reputação

- Definição de *Fake News* adotada

- Publicações intencionalmente e verificadamente falsas (Disinformation)

- **Comportamento disseminativo das *Fake News***

- **Divulgação (Câmara de Eco)**

- **Publicação**

- **Propagação**

- Razões para o poder divulgação
  - Fatores inerentes ao homem [Shu et al., 2017]
    - Confirmação da opinião
    - Ganhos x perdas
    - Aceitação dos outros
  - Carência legislação punitiva
  - Ganho financeiro [Kshetri and Voas, 2017]
  - Facilidade de criação de contas [Conroy et al., 2015]
    - Bot
    - Humano (Trolls)
    - Cyborg

Considerações Iniciais

Fundamentos

**Trabalhos Relacionados**

Problemas em Aberto

Estudo de Caso em Detecção Automática de Fake News

Considerações Finais

Referências

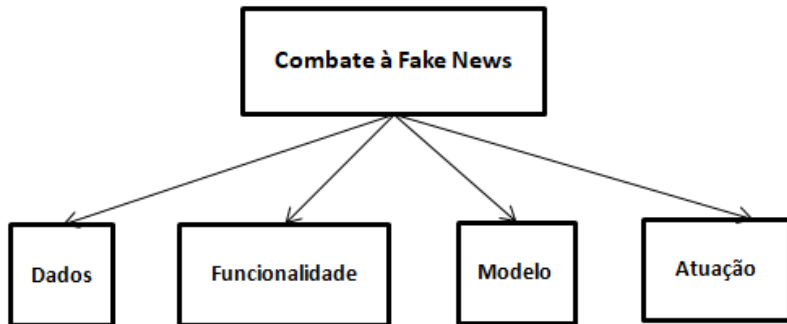
Proposta de Modelo Comparativo

Revisão dos Trabalhos Relacionados

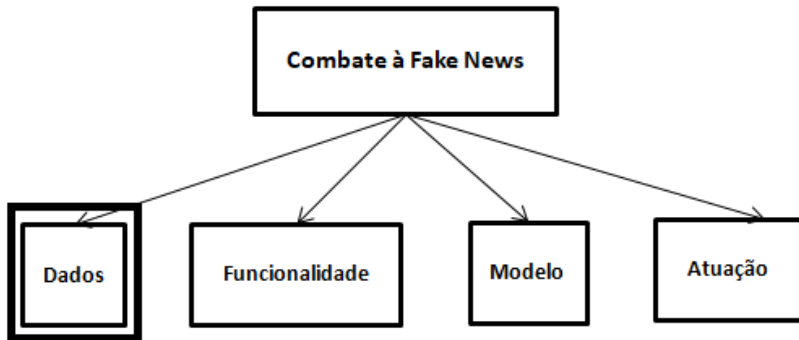
Datasets

# Trabalhos Relacionados

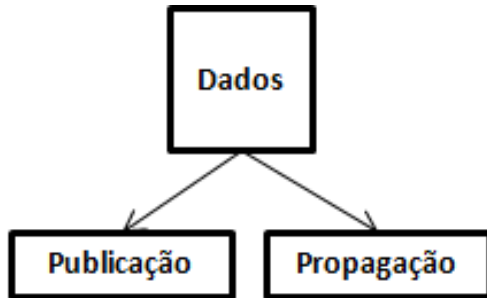
## Proposta de Modelo Comparativo de Abordagens Computacionais para Combate Automático às *Fake News*



## Proposta de Modelo Comparativo de Abordagens Computacionais para Combate Automático às *Fake News*



## Proposta de Modelo Comparativo de Abordagens Computacionais para Combate Automático às *Fake News*





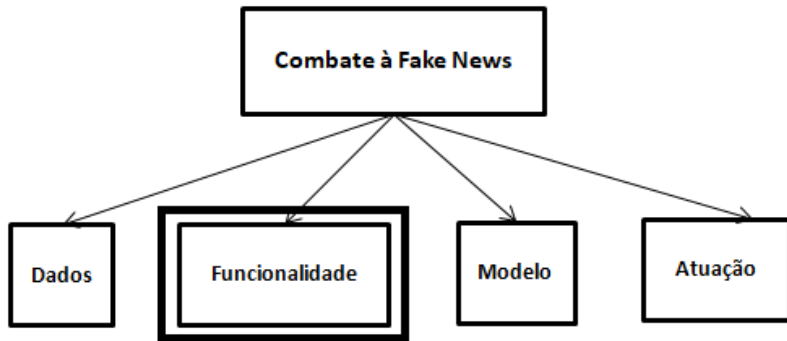
## Dados

- Publicação
  - Notícia
    - Mídia (Texto/Áudio/Imagem)
    - Conteúdo (Léxica/Sintática/Semântica/Legibilidade)
  - Usuário
    - Tipo (Humano/Bot/Cyborg)
    - Perfil
    - Reputação
  - Assunto
    - Especificidades (Relacionados/Controversos/Tópicos)
    - Relevância
  - Temporalidade

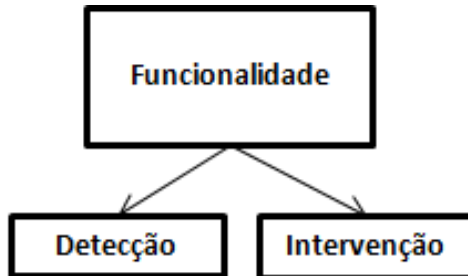
## Dados

- Propagação
  - Contribuição
    - Mídia (Texto/Áudio/Imagem)
    - Conteúdo (Léxica/Sintática/Semântica/Legibilidade)
  - Usuário
    - Tipo (Humano/Bot/Cyborg)
    - Perfil
    - Reputação
  - Assunto
    - Especificidades (Relacionados/Controversos/Tópicos)
    - Relevância
  - Temporalidade
  - Rede

## Proposta de Modelo Comparativo de Abordagens Computacionais para Combate Automático às *Fake News*



## Proposta de Modelo Comparativo de Abordagens Computacionais para Combate Automático às *Fake News*



## Funcionalidade

- Detecção
  - Classificação
  - Probabilidade
  - Peso
  - Pertinência

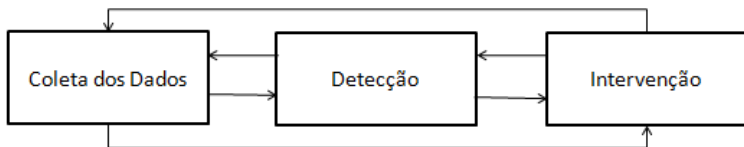
## Funcionalidade

- Detecção
  - Autenticidade
  - Intencionalidade

## Funcionalidade

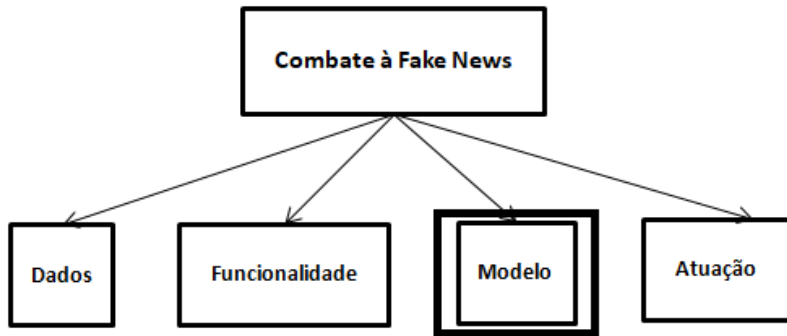
- Intervenção
  - Bloqueio (Reativa)
  - Mitigação (Reativa / Proativa)

## Funcionalidade

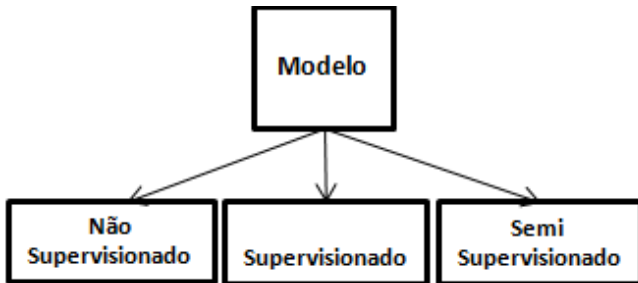




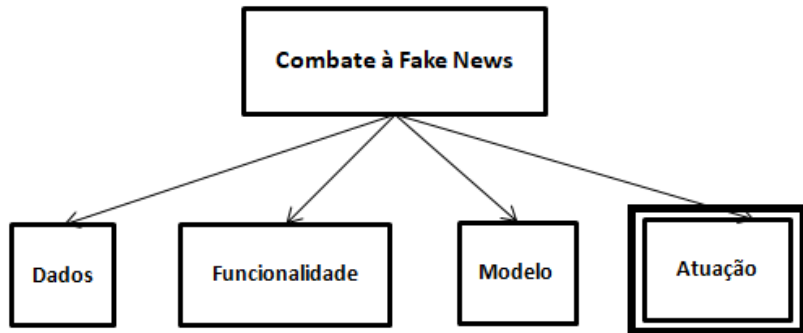
## Proposta de Modelo Comparativo de Abordagens Computacionais para Combate Automático às *Fake News*



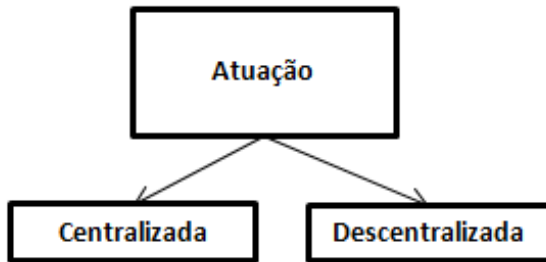
## Proposta de Modelo Comparativo de Abordagens Computacionais para Combate Automático às *Fake News*



## Proposta de Modelo Comparativo de Abordagens Computacionais para Combate Automático às *Fake News*



## Proposta de Modelo Comparativo de Abordagens Computacionais para Combate Automático às *Fake News*



Id	Dados Publicação Usuário							
	Notícia		Usuário			Assunto		Temporalidade
	Mídia (Texto, Áudio e Imagem)	Conteúdo (Léxica, Sintática, Semântica e Legibilidade)	Tipo (Humano, Bot e Cyborg)	Perfil	Reputação	Especificidades	Relevância	
T1	Texto	Léxica e Semântica						
T2	Texto	Léxica, Sintática, Semântica e Legibilidade						
T3	Texto e Imagem	Léxica						
T4	Texto	Léxica e Semântica		X				X
T5	Texto	Léxica e Semântica		X	X			
T6	Texto	Léxica		X				
T7	Texto	Semântica						
T8	Texto	Léxica e Semântica		X	X			X
T9					X	Assuntos controversos		
T10	Texto e Imagem	Léxica						
T11				X				X
T12	Texto	Léxica e Semântica						
T13								
T14					X			X
T15								
T16	Texto	Léxica e Semântica		X				X
T17	Texto	Léxica e Semântica		X		Relaciona Assuntos		
T18	Texto	Semântica						
T19	Texto	Semântica						
T20			Bot	X				X
T21	Texto	Léxica e Semântica						
T22	Texto	Léxica e Sintática						
T23	Texto	Semântica						
T24					X			
T25	Texto	Léxica e Semântica		X		Análise dos Tópicos		X
T26	Texto	Léxica e Semântica		X		Análise dos Tópicos		

Id	Dados								
	Contribuição		Propagação			Assunto		Temporalidade	Rede
	Mídia (Texto, Áudio e Imagem)	Conteúdo (Léxica, Sintática, Semântica e Legibilidade)	Tipo (Humano, Bot e Cyborg)	Usuário		Especificidades	Relevância		
				Perfil	Reputação				
T1									
T2									
T3	Texto	Léxica						X	
T4	Texto	Léxica e Semântica		X			X	X	
T5				X	X			X	
T6				X				X	
T7									
T8	Texto	Léxica e Semântica		X	X		X	X	
T9									
T10									
T11				X			X	X	
T12									
T13								X	
T14					X		X	X	
T15								X	
T16				X			X	X	
T17									
T18	Texto	Semântica							
T19									
T20			Bot	X			X	X	
T21	Texto	Léxica e Semântica							
T22									
T23									
T24					X			X	
T25	Texto	Léxica e Semântica							
T26									

Id	Modelo			Funcionalidade				Atuação (Centralizada ou Descentralizada)
				Detecção		Intervenção		
	Não Supervisionado	Semi Supervisionado	Supervisionado	Autenticidade	Intencionalidade	Bloqueio (Reativa)	Mitigação (Proativa e Reativa)	
T1			X	X				Centralizada
T2			X	X				Centralizada
T3			X	X				Centralizada
T4			X	X	Análise das características dos usuários			Centralizada
T5		X		X	Pontuação de credibilidade para os usuários			Centralizada
T6			X	X				Centralizada
T7			X	X				Centralizada
T8			X	X	Score para os usuários			Centralizada
T9		X		X	Atribui pesos de confiança aos websites			Centralizada
T10			X	X				Centralizada
T11			X	X				Centralizada
T12			X	X				Centralizada
T13							Reativa	Centralizada
T14			X	X		X		Centralizada
T15							Reativa	Centralizada
T16		X		X				Centralizada
T17			X	X	Associação com o usuário			Centralizada
T18			X	X				Centralizada
T19			X	X				Centralizada
T20		X		X	identificação de bots			Centralizada
T21			X	X	Análise de Sentimentos			Centralizada
T22			X	X				Centralizada
T23		X		X				Centralizada
T24			X	X	Relação entre os usuários			Centralizada (pode ser paralelizada)
T25			X	X	Análise de Sentimentos			Centralizada
T26			X	X				Centralizada

Id	DataSet
T1	FakeNewsAMT (D11), Celebrity (D5) e PoliticalNews (D19)
T2	FakeNewsAMT (D11) e Celebrity (D5)
T3	BuzzFeedNews (2016-10-facebookfact-check modificado) (D3)
T4	Twitter para Automatically Identifying Fake News (D28)
T5	FakeNewsNet1 (D13)
T6	FakeNewsNet1 (D13)
T7	DataSet Emergent (D7)
T8	Twitter e Sina Weibo para CSI (D24)
T9	DistrustRank Datasets (D8)
T10	Twitter e Sina Weibo para EANN (D25)
T11	Twitter e Sina Weibo para Early Detection Through Propagation Path (D26)
T12	Signal Media para Evaluating Machine Learning Algorithms for Fake News Detection (D22)
T13	Soc-LiveJournal (D23) e Twitter Trec (D32)
T14	Facebook para Detective (D9)
T15	Twitter para Mitigation via Point Process (D30)
T16	FakeNewsNet1 (D13)
T17	LIAR (D18)
T18	Twitter e Sina Weibo para TCNN-URG (D27)
T19	DataSet Emergent (D7)
T20	Twitter para Content Polluters (D29)
T21	PHEME (dataset para Rumor)
T22	FakeNewsData1 (D12)
T23	RST-SVM Dataset (D21)
T24	Twitter para TraceMiner (D31)
T25	Large-scale Training Dataset e Small-scale Evaluation Dataset (D17)
T26	PolitiFact para XFake (D20)



Id	Dados									URL
	Publicação			Usuário	Propagação			Usuário	Rede	
	Texto	Áudio	Imagem		Texto	Áudio	Imagem			
D1	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	No Dataset	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	<a href="https://github.com/thiagovas/bs-detector-dataset">https://github.com/thiagovas/bs-detector-dataset</a>
D2	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>		No Dataset	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>		No Dataset	<a href="#">Link para notícia</a>	<a href="https://github.com/gsanita/BuzzFace">https://github.com/gsanita/BuzzFace</a>
D3	No Dataset		<a href="#">Link para imagem</a>	No Dataset	No Dataset			<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	
D4	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	No Dataset	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	<a href="https://github.com/BuzzFeedNews/2016-10-facebook-fact-check">https://github.com/BuzzFeedNews/2016-10-facebook-fact-check</a>
D5	No Dataset			No Dataset						<a href="http://lit.eecs.umich.edu/downloads.html#undefined">http://lit.eecs.umich.edu/downloads.html#undefined</a>
D6	No Dataset			No Dataset				No Dataset	No Dataset	<a href="http://composocial.github.io/CREDBANK-data/">http://composocial.github.io/CREDBANK-data/</a>
D7	No Dataset			No Dataset						<a href="https://github.com/FakeNewsChallenge/fnc-1">https://github.com/FakeNewsChallenge/fnc-1</a>
D8	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	No Dataset						
D9				No Dataset				No Dataset		
D10	No Dataset	<a href="#">Link para notícia</a>	<a href="#">Link para notícia</a>	No Dataset						<a href="https://github.com/jgoebck/fakenews">https://github.com/jgoebck/fakenews</a>
D11	No Dataset			No Dataset						<a href="http://lit.eecs.umich.edu/downloads.html#undefined">http://lit.eecs.umich.edu/downloads.html#undefined</a>
D12	No Dataset									<a href="https://github.com/BenjaminDHorne/fakenewsdata1/blob/master/Horne2017_FakeNewsData.zip">https://github.com/BenjaminDHorne/fakenewsdata1/blob/master/Horne2017_FakeNewsData.zip</a>
D13	No Dataset		<a href="#">Link para imagem</a>	No Dataset				No Dataset	No Dataset	<a href="https://github.com/KaiDMML/FakeNewsNet">https://github.com/KaiDMML/FakeNewsNet</a>
D14	No Dataset		<a href="#">Link para notícia</a>	No Dataset				No Dataset	No Dataset	<a href="https://github.com/KaiDMML/FakeNewsNet">https://github.com/KaiDMML/FakeNewsNet</a>
D15	No Dataset		<a href="#">Link para imagem</a>	No Dataset						<a href="https://www.kaggle.com/mrisdal/fake-news/data">https://www.kaggle.com/mrisdal/fake-news/data</a>
D16	No Dataset			No Dataset						
D17	No Dataset			No Dataset						<a href="http://dws.informatik.uni-mannheim.de/en/research/twitter-fake-news-detection">http://dws.informatik.uni-mannheim.de/en/research/twitter-fake-news-detection</a>
D18	No Dataset			No Dataset						<a href="https://github.com/nishitpatel01/FakeNewsDetection/tree/master/liar_dataset">https://github.com/nishitpatel01/FakeNewsDetection/tree/master/liar_dataset</a> ou <a href="https://www.cs.ucsb.edu/~william/software.html">https://www.cs.ucsb.edu/~william/software.html</a>
D19	No Dataset			No Dataset						<a href="https://osf.io/ez5q4/">https://osf.io/ez5q4/</a>
D20	No Dataset			No Dataset						
D21	No Dataset			No Dataset						
D22	No Dataset			No Dataset						
D23									No Dataset	<a href="https://snap.stanford.edu/data/soc-LiveJournal1.html">https://snap.stanford.edu/data/soc-LiveJournal1.html</a>
D24	No Dataset			No Dataset	No Dataset			No Dataset	No Dataset	<a href="https://github.com/majingCUHK/Rumor_RvNN">https://github.com/majingCUHK/Rumor_RvNN</a> ou <a href="http://alt.qcri.org/~wgao/data/rumdet.zip">http://alt.qcri.org/~wgao/data/rumdet.zip</a>
D25	No Dataset		No Dataset	No Dataset						
D26				No Dataset				No Dataset	No Dataset	Twitter 15 e 16 ( <a href="https://www.dropbox.com/s/7ewzdrbelpmrxu/rumdet2017.zip?dl=0">https://www.dropbox.com/s/7ewzdrbelpmrxu/rumdet2017.zip?dl=0</a> ) e Weibo( <a href="http://alt.qcri.org/~wgao/data/rumdet.zip">http://alt.qcri.org/~wgao/data/rumdet.zip</a> )
D27	No Dataset			No Dataset	No Dataset					False ( <a href="https://drive.google.com/open?id=1WRoRVV9j4CstMFKDwP7DVGAFJZX4t5a">https://drive.google.com/open?id=1WRoRVV9j4CstMFKDwP7DVGAFJZX4t5a</a> ) e True( <a href="https://drive.google.com/open?id=1JgbW4suN2yWHt65P4QU8HkrB30MHsuo">https://drive.google.com/open?id=1JgbW4suN2yWHt65P4QU8HkrB30MHsuo</a> )
D28	No Dataset			No Dataset				No Dataset	No Dataset	
D29				No Dataset				No Dataset	No Dataset	
D30									No Dataset	
D31				No Dataset				No Dataset	No Dataset	
D32	No Dataset			No Dataset				No Dataset	No Dataset	<a href="https://trc.nist.gov/data/tweets/">https://trc.nist.gov/data/tweets/</a>

# Problemas em Aberto

## Problemas em Aberto

- Carência de *datasets*
- Trabalhos que levem em consideração aspectos temporais do ciclo de vida da *Fake News*
- Estudos que analisem o aspecto intencional
- Extração de características a partir de imagem e/ou áudio
- Métodos que abordem características baseadas na rede que representa a propagação da notícia

## Problemas em Aberto

- Pesquisas que, ao invés de realizarem uma classificação binária, utilizem probabilidades e/ou pertinências na detecção
- Utilização de um comitê de classificadores para determinar se uma notícia é *fake*
- Utilização de modelos não supervisionados ou semi-supervisionados
- Estudo sobre o comportamento distinto da *Fake News* em diferentes comunidades (escolar, trabalho e etc) e/ou redes sociais (*Weibo, WhatsApp e etc*)

## Problemas em Aberto

- Classificar os usuários de *Fake News* com o objetivo de identificar o seu tipo (*humanos, bots e cyborgs*)
- Trabalhos relacionados à intervenção de *Fake News*, tanto para bloqueio quanto para mitigação
- Abordagens que atuem descentralizadas na rede
- Abordagens que utilizem o assunto para a análise da notícia
- Pesquisas que levem em consideração a reputação dos usuários

# Estudo de Caso em Detecção Automática de *Fake News*

## Fake News Detection in Social Networks via Crowd Signals [Tschitschek et al., 2018]

- A reputação do usuário é expressa em função do seu histórico de acertos e erros em suas opiniões sobre as notícias
- Propõe um método chamado *Detective* que classifica uma notícia, como *fake* ou não, a partir de *Crowd Signals*.
- Este *Crowd* é formado pelas opiniões dos usuários, juntamente com as suas respectivas reputações.
- O *Detective* usa Inferência Bayesiana para detectar *Fake News*

## *Detective*

$$\begin{vmatrix} \theta_{u,\bar{f}} & 1 - \theta_{u,f} \\ 1 - \theta_{u,\bar{f}} & \theta_{u,f} \end{vmatrix}$$

onde:

- $\theta_{u,\bar{f}} = P(Y_u(x) = \bar{f} \mid Y^*(x) = \bar{f})$
- $1 - \theta_{u,\bar{f}} = P(Y_u(x) = f \mid Y^*(x) = \bar{f})$
- $\theta_{u,f} = P(Y_u(x) = f \mid Y^*(x) = f)$
- $1 - \theta_{u,f} = P(Y_u(x) = \bar{f} \mid Y^*(x) = f)$



## Detective

$$\begin{vmatrix} \theta_{u,\bar{f}} & 1 - \theta_{u,f} \\ 1 - \theta_{u,\bar{f}} & \theta_{u,f} \end{vmatrix}$$

$$\begin{cases} P(Y^*(a) = f) = \omega \cdot \prod_{u \in \psi^t(a)} \theta_{u,f} \cdot \prod_{u \in \pi^t(a) \setminus \psi^t(a)} (1 - \theta_{u,f}) \\ P(Y^*(a) = \bar{f}) = (1 - \omega) \cdot \prod_{u \in \psi^t(a)} (1 - \theta_{u,\bar{f}}) \cdot \prod_{u \in \pi^t(a) \setminus \psi^t(a)} \theta_{u,\bar{f}} \end{cases}$$

onde:

- $\pi^t(a)$

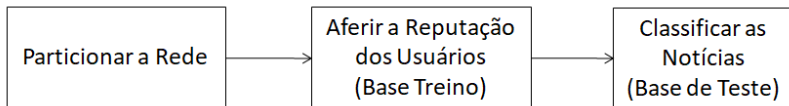
usuários que visualizaram a notícia  $a$  no final do tempo  $t$

- $\psi^t(a)$

usuários que opinaram que a notícia  $a$  é *fake* no final do tempo  $t$

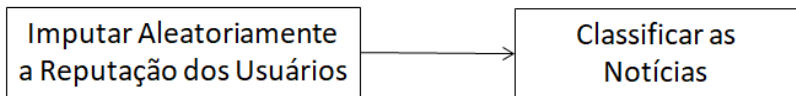
## *Detective*

- Execução do Experimento



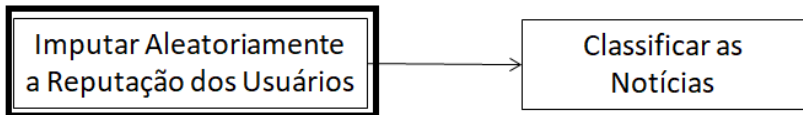
## *Detective*

- Execução do Experimento (Ótima)

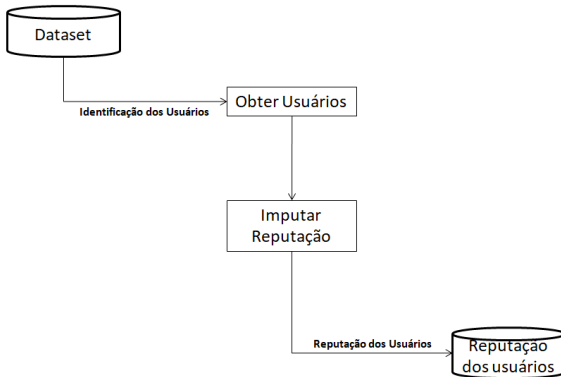


## *Detective*

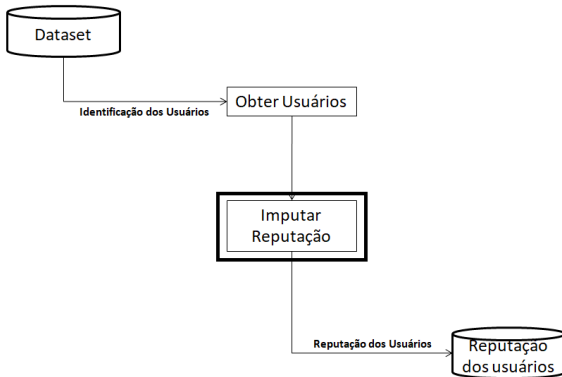
- Execução do Experimento (Ótima)



## Imputar Aleatoriamente a Reputação dos Usuários



## Imputar Aleatoriamente a Reputação dos Usuários



## Imputar Aleatoriamente a Reputação dos Usuários

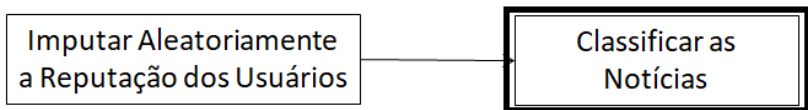
$$\begin{vmatrix} \theta_{u,\bar{f}} & 1 - \theta_{u,f} \\ 1 - \theta_{u,\bar{f}} & \theta_{u,f} \end{vmatrix}$$

onde:

- *Bom* ( $\theta_{u,\bar{f}} = \theta_{u,f} = 0.9$ )
- *Indiferente* ( $\theta_{u,\bar{f}} = \theta_{u,f} = 0.5$ )
- *Spammer* ( $\theta_{u,\bar{f}} = \theta_{u,f} = 0, 1$ )

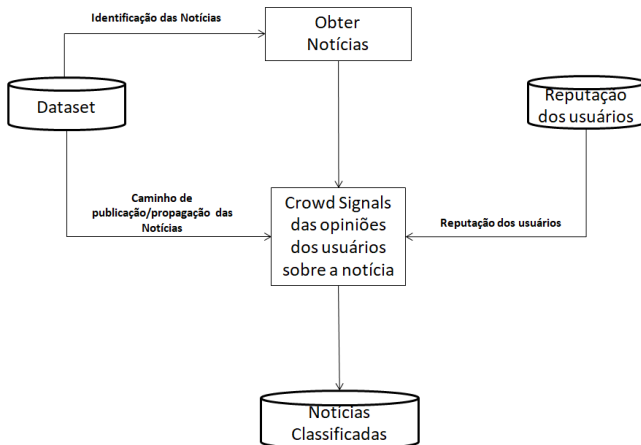
## *Detective*

- Execução do Experimento (Ótima)

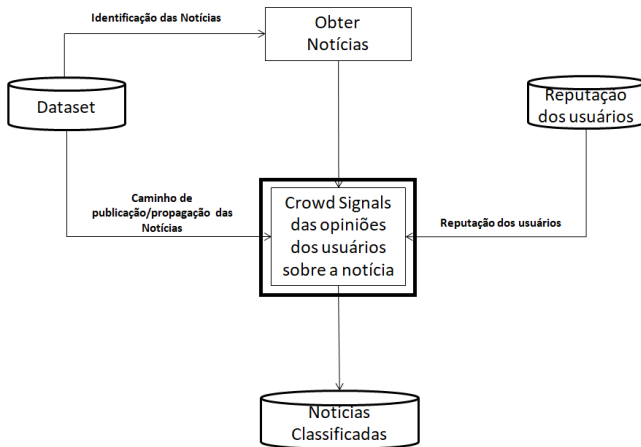




## Classificar as Notícias



## Classificar as Notícias



## Crowd Signals

- Inferência Bayesiana

$$\begin{cases} P(Y^*(a) = f) = \omega \cdot \prod_{u \in \psi^t(a)} \theta_{u,f} \cdot \prod_{u \in \pi^t(a) \setminus \psi^t(a)} (1 - \theta_{u,f}) \\ P(Y^*(a) = \bar{f}) = (1 - \omega) \cdot \prod_{u \in \psi^t(a)} (1 - \theta_{u,\bar{f}}) \cdot \prod_{u \in \pi^t(a) \setminus \psi^t(a)} \theta_{u,\bar{f}} \end{cases}$$

## Resultado de uma rodada de execução do experimento

Método	BuzzFeed	PolitiFact
<i>Detective</i>	<b>0.9890</b>	<b>0.9791</b>

# Considerações Finais

# Referências

## Referências

- [Braz and Goldschmidt, 2017] Braz, P. and Goldschmidt, R. (2017).  
Um método para detecção de bots sociais baseado em redes neurais convolucionais aplicadas em mensagens textuais.  
In *SBSeg 2017*, pages 501–508.  
10/11/2017.
- [Campan et al., 2017] Campan, A., Cuzzocrea, A., and Truta, T. M. (2017).  
Fighting fake news spread in online social networks: Actual trends and future research directions.  
In *2017 IEEE International Con on Big Data (Big Data)*, pages 4453–4457.
- [Castelo et al., 2019] Castelo, S., Almeida, T., Elghafari, A., Santos, A., Pham, K., Nakamura, E., and Freire, J. (2019).  
A topic-agnostic approach for identifying fake news pages.  
In *Companion Proceedings of The 2019 World Wide Web Conference, WWW '19*, pages 975–980, New York, NY, USA. ACM.

- [Cazalens et al., 2018] Cazalens, S., Lamarre, P., Leblay, J., Manolescu, I., and Tannier, X. (2018).  
A content management perspective on fact-checking.  
In *Companion Proceedings of the The Web Con 2018, WWW '18*, pages 565–574, Republic and Canton of Geneva, Switzerland. International World Wide Web Con Steering Committee.
- [Conroy et al., 2015] Conroy, N., Rubin, V., and Chen, Y. (2015).  
Automatic deception detection: Methods for finding fake news.  
*Association for Information Science and Technology*, 52:1–4.
- [Farajtabar et al., 2017] Farajtabar, M., Yang, J., Ye, X., Xu, H., Trivedi, R., Khalil, E., Li, S., Song, L., and Zha, H. (2017).  
Fake news mitigation via point process based intervention.  
In *Proceedings of the 34th International Con on Machine Learning - Volume 70, ICML'17*, pages 1097–1106. JMLR.org.



- [Flintham et al., 2018] Flintham, M., Karner, C., Bachour, K., Creswick, H., Gupta, N., and Moran, S. (2018).  
Falling for fake news: Investigating the consumption of news via social media.  
*In Proceedings of the 2018 CHI Con on Human Factors in Computing Systems, CHI '18*, pages 376:1–376:10, New York, NY, USA. ACM.
- [Hendrikx et al., 2015] Hendrikx, F., Bubendorfer, K., and Chard, R. (2015).  
Reputation systems: A survey and taxonomy.  
*Journal of Parallel and Distributed Computing*, pages 184–197.
- [Kshetri and Voas, 2017] Kshetri, N. and Voas, J. (2017).  
The economics of fake news.  
*IT Professional*, 19(06):8–12.
- [Li et al., 2015] Li, Y., Gao, J., Meng, C., Li, Q., Su, L., Zhao, B., Fan, W., and Han, J. (2015).  
A survey on truth discovery.  
*ACM SIGKDD Explorations Newsletter*, 17:1–16.

- [Reis et al., 2019] Reis, J. C. S., Correia, A., Murai, F., Veloso, A., and Benevenuto, F. (2019).  
Explainable machine learning for fake news detection.  
In *Proceedings of the 10th ACM Conference on Web Science, WebSci '19*, pages 17–26, New York, NY, USA. ACM.
- [Sharma et al., 2019] Sharma, K., Qian, F., Jiang, H., Ruchansky, N., Zhang, M., and Liu, Y. (2019).  
Combating fake news: A survey on identification and mitigation techniques.  
*ACM Trans. Intell. Syst. Technol.*, 10(3):21:1–21:42.
- [Shu et al., 2017] Shu, K., Sliva, A., Wang, S., Tang, J., and Liu, H. (2017).  
Fake news detection on social media: A data mining perspective.  
*SIGKDD Explor. Newsl.*, 19(1):22–36.
- [Tschitschek et al., 2018] Tschitschek, S., Singla, A., Gomez Rodriguez, M., Merchant, A., and Krause, A. (2018).  
Fake news detection in social networks via crowd signals.  
In *Companion Proceedings of the The Web Con 2018, WWW '18*, pages 517–524, Republic and Canton of Geneva, Switzerland. International World Wide Web Con Steering Committee.

[Vosoughi et al., 2017] Vosoughi, S., Mohsenvand, M. N., and Roy, D. (2017). Rumor gauge: Predicting the veracity of rumors on twitter. *ACM Trans. Knowl. Discov. Data*, 11(4):50:1–50:36.

[Wang et al., 2018] Wang, P., Angarita, R., and Renna, I. (2018). Is this the era of misinformation yet: Combining social bots and fake news to deceive the masses. In *Companion Proceedings of the The Web Con 2018, WWW '18*, pages 1557–1561, Republic and Canton of Geneva, Switzerland. International World Wide Web Con Steering Committee.

[Zhou et al., 2019] Zhou, X., Zafarani, R., Shu, K., and Liu, H. (2019). Fake news: Fundamental theories, detection strategies and challenges. In *Proceedings of the Twelfth ACM International Con on Web Search and Data Mining, WSDM '19*, pages 836–837, New York, NY, USA. ACM.