Métodos estatísticos para detecção de bots maliciosos

Caio Henrique Assad Racy RA: 141020091
Orientador: Profº.Drº. Kelton Augusto Pontara da Costa

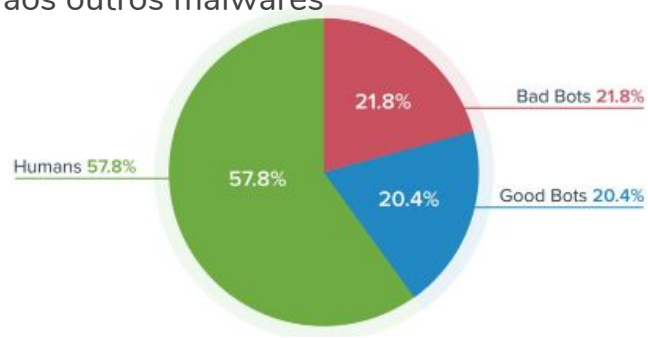
Unesp - Universidade Estadual Paulista
”Júlio de Mesquita Filho”
Faculdade de Ciências
Departamento de Computação
2018

Conceituação

- **Bots:** Ataques cibernéticos ocorrem constantemente e tem o intuito de capturar informações ou até mesmo utilizar-se de outros sistemas para ganho pessoal, e é nesse contexto que surge o conceito de bot, palavra inglesa derivada de “robot” (robô) que significa software de automatização de procedimentos.
- **Fake Followers:** Seguidores falsos são aquelas contas do Twitter criadas especificamente para aumentar o número de seguidores de um alvo ou contratante. Seguidores falsos são perigosos para a plataforma social, já que eles podem alterar conceitos como popularidade e influência na mesma portanto, impactando a economia, a política e a sociedade

Problema

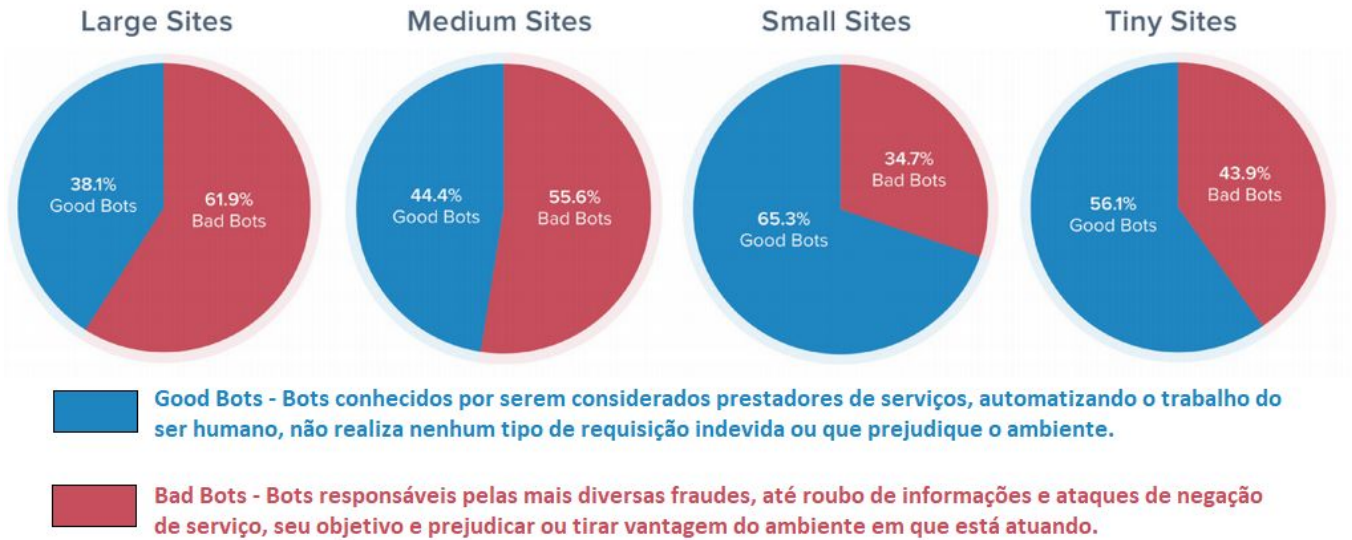
- Bots sendo utilizados como facilitadores de trabalhos “manuais”
- Quantidade de bots cada vez maior
- Mais da metade dos bots existentes na rede, são maliciosos
  - Segundo a Distil Networks cerca de 40% de todo tráfego na rede é de bots e desses 40%, mais de 50% são de bots maliciosos
  - Sites de porte consideravelmente maiores são os maiores alvos de bots maliciosos. Incluindo redes sociais (sites gigantes)
- Bots são mais difíceis de se detectar se comparados aos outros malwares
  - Devido sua inatividade seletiva.
  - Não agem de forma constante.



Fonte: Distil Networks (2018)

Problema

Distribuição do tráfico de bots e humanos por tipo de site



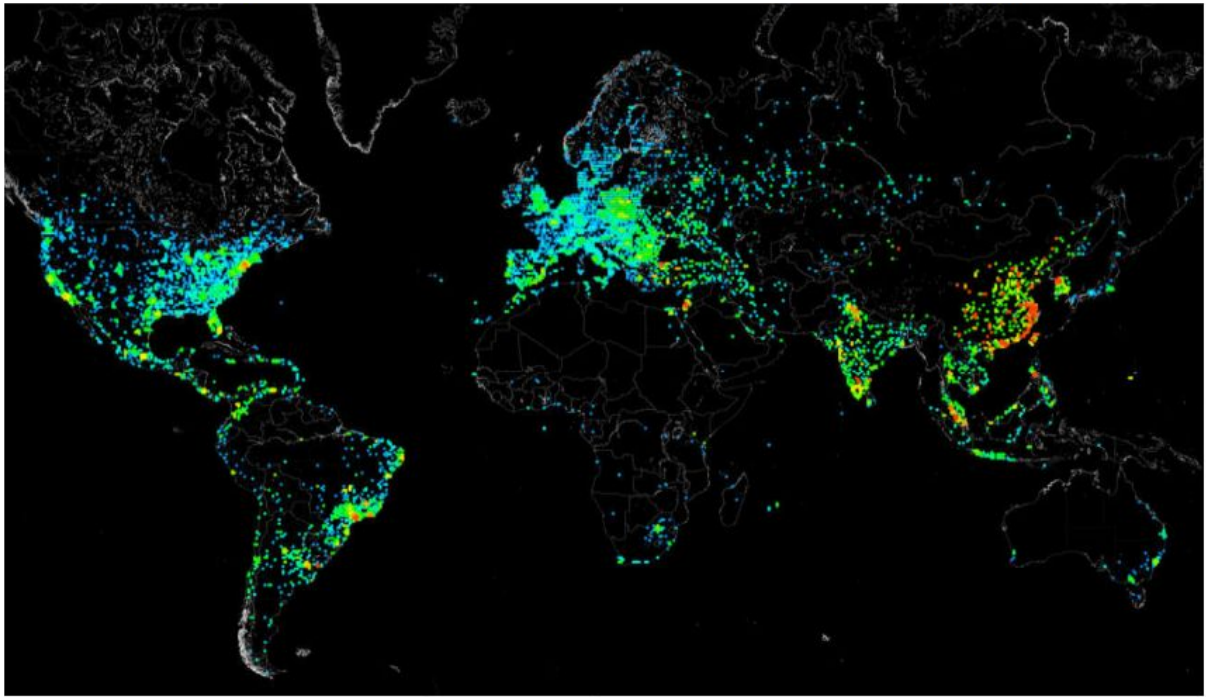
Fonte: Distil Networks (2018)

Justificativa

- Dados obtidos geralmente levam muito tempo de pesquisa
- Análise demorada pode atrapalhar na identificação de bots maliciosos
  - Métodos estatísticos como facilitadores
  - Amostragem x população
- Carna Bot demonstrou como é fácil sua propagação pela rede
  - 420 mil computadores infectados em 2012
  - Quantidade enorme de dados para se trabalhar
  - Apelidado de Census 2012

Justificativa

Quantidade de IPs afetados mundialmente que sofreram ping pelo bot



Fonte: Census2012 (2012)

Objetivo

- O objetivo é analisar pelos métodos estatísticos de média, variância e desvio padrão, além de Regressão Linear Simples (RLS) a possibilidade de se detectar bots maliciosos do Twitter, com objeto de estudo, os bots do tipo “Fake Followers”, ou seja, falsos seguidores.

Objetivos específicos

- Definir quais métodos estatísticos que melhor consegue analisar o comportamento dos bots do tipo “Fake Followers” do Twitter.
- Identificar características relevantes sobre o comportamento dos bots
- Implementação de um script capaz de carregar e realizar a análise quantidade desses bots, devolvendo resultados expressivos.
- Avaliar os resultados obtidos pelas amostras, com base na população de bots e no intervalo de confiança obtido (95% como padrão do R).

Metodologia

- Utilização da Linguagem R de programação.
- Utilização da IDE RStudio.
- Utilização de base de dados de 2015.
- Aplicar os conhecimentos básicos de estatística em dados quantitativos na base de seguidores falsos:
    - Média, Desvio e Variância
- Aplicar método de regressão Linear Simples em determinadas características da base para identificar relações.
    - Grau de correlação entre as características.
- A partir do resultado, aplicar um teste de Hipótese para comparar as bases de usuários reais e bots.

Metodologia

- Informações relevantes:

Base de usuários reais - "Genuine Users"	
ID	**Quantidade**
Nº de Registros	3474
Nº de Características	42

Tabela 1 – Informações quantitativas dos dados de usuários reais.

Base de seguidores falsos - "Fake Followers"	
ID	**Quantidade**
Nº de Registros	3351
Nº de Características	38

Tabela 2 – Informações quantitativas dos dados de seguidores falsos.

Fonte: Própria do autor

Metodologia

- Características relevantes em comum dos dados:

Características em comum entre os dados	
Nome	**Descrição**
id	chave de identificação
name	Nome do usuário
screen_name	nome da conta do usuário
statuses_count	total de tweets
followers_count	quantidade de seguidores
friends_count	quantidade de amigos
favourites_count	quantidade de favoritos
listed_count	quantidade de listas públicas de um usuário
created_at	data de criação da conta
url	link de referência vinculado à conta
lang	linguagem cadastrada
time_zone	Fuso horário do usuário
location	Localização dos usuários

Tabela 3 – Campos relevantes levados em consideração para o estudo.

Fonte: Própria do autor

Metodologia

Regressão Linear Simples (RLS):

"A análise de regressão diz respeito ao estudo de dependência de uma variável, a variável dependente em relação a uma ou mais variáveis, as variáveis explanatórias, visando estimar e/ou prever o valor médio (da população) da primeira em termos dos valores conhecidos ou fixados (em amostragens repetidas) da segunda."
(GUJARATI; PORTER, 2011)

Metodologia

Regressão Linear Simples (RLS):

- Com o R é possível de se trabalhar com cálculos de regressão Linear e descobrir padrões sobre certas características, com este método é possível dizer qual o grau de correlação entre os dados e como eles estão distribuídos na base (força da regressão).

O grau de correlação é medido a partir do coeficiente de regressão linear (cor):

$0 < cor < 0,7$:	regressão positiva fraca entre as variáveis X e Y
$cor \geq 0,7$:	regressão positiva forte entre as variáveis X e Y
$cor = 0$:	Inexistência de relação linear
$-0,7 < cor < 0$:	regressão negativa entre as variáveis X e Y
$cor \leq -0,7$:	regressão negativa forte entre as variáveis X e Y

Metodologia

Teste de Hipótese

- Na linguagem da estatística, a hipótese estabelecida é denominada hipótese nula ou H_0.
- Em geral essa hipótese é testada com um hipótese alternativa que é totalmente diferente do resultado dela, ou seja uma hipótese H_1.

$$H_0: \quad \text{Bases iguais } (H_0 = 0)$$
$$H_1: \quad \text{Bases diferentes } (H_0 \neq H_1)$$

A teoria do teste de hipótese simplesmente busca por regras e procedimentos para dizer se a hipótese nula deve ser rejeitada ou não. Para chegar em um resultado, avalia-se o valor **p** *ou "p-value", baseado no intervalo de confiança de 95%, ou seja, se o valor for menor que 0,05 é descartada a hipótese* H_0.

Resultados

- Os resultados obtidos da primeira análise pode ser comparado nas tabela a seguir:

Comparação de Médias			
Característica	**Usuários Reais**	**Usuários Bots**	**p-value**
followers_count	1393,21	17,94	$< 2,58 * 10^{-6}$
friends_count	633,24	370,05	$< 2,2 * 10^{-16}$
statuses_count	16958,22	71.89	$< 2,2 * 10^{-16}$
favourites_count	4669,62	4,29	$< 2,2 * 10^{-16}$

Tabela 4 – Comparação de médias entre usuários reais e Bots do tipo *"fake followers"*

Comparação de Desvios		
Característica	**Usuários Reais**	**Usuários Bots**
followers_count	17216,66	54,22
friends_count	1600,96	212,55
statuses_count	30696,29	634,97
favourites_count	11527.57	59,55

Tabela 5 – Comparação de desvios entre usuários reais e Bots do tipo *"fake followers"*

Resultados

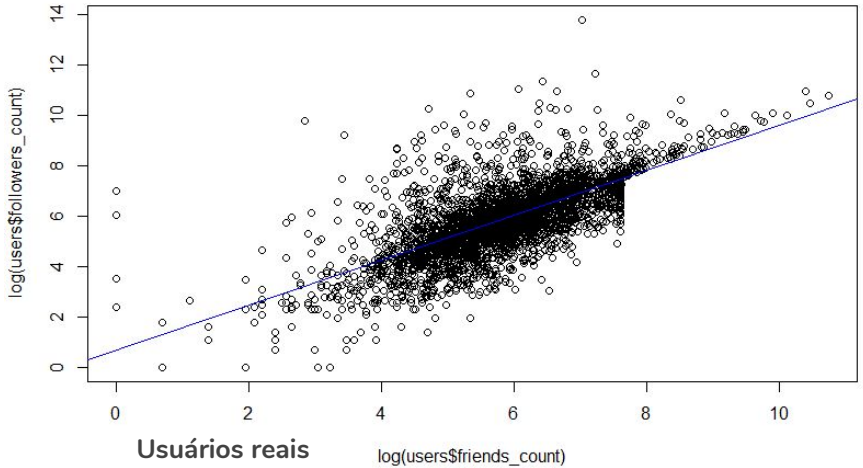
Comparação de Variâncias		
Característica	**Usuários Reais**	**Usuários Bots**
followers_count	296413537	2939.372
friends_count	2563082	45177.68
statuses_count	942261981	403199.2
favourites_count	132884793	3546.381

Tabela 6 – Comparação de variâncias entre usuários reais e Bots do tipo *"fake followers"*

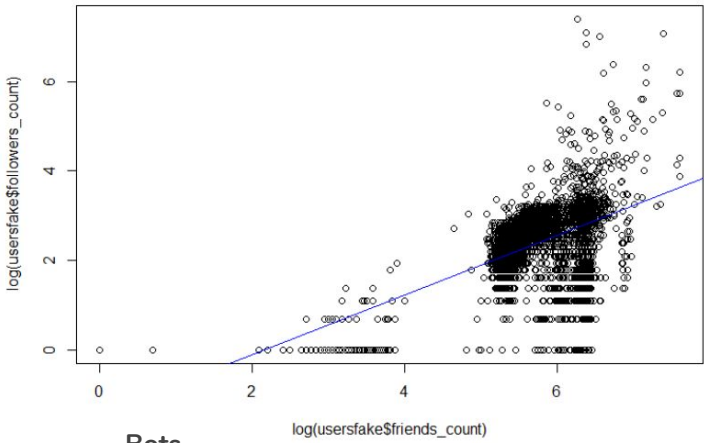
- Possível notar grande diferença nos valores, principalmente na média e nos desvios que são medidas mais utilizadas.
- O teste **t** aplicado faz referência apenas a utilização das médias e por isso foi calculado apenas uma vez.
- Todas as características quantitativas deram o valor de **p** menor que 0,05, ou seja, Hipótese diferente confirmada para todas essas características.

Resultados

Quantidade de Seguidores x Quantidade de Amigos



Usuários reais



Bots

Equação da reta: $f(x) = 0,6660 + 0,8938x$

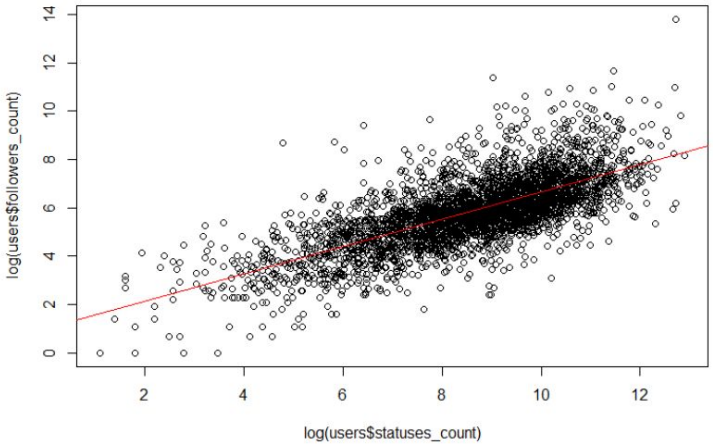
$f(x) = -1.4464 + 0.6686x$

Grau de correlação: 0,6885 correlação positiva fraca (< 0,7)

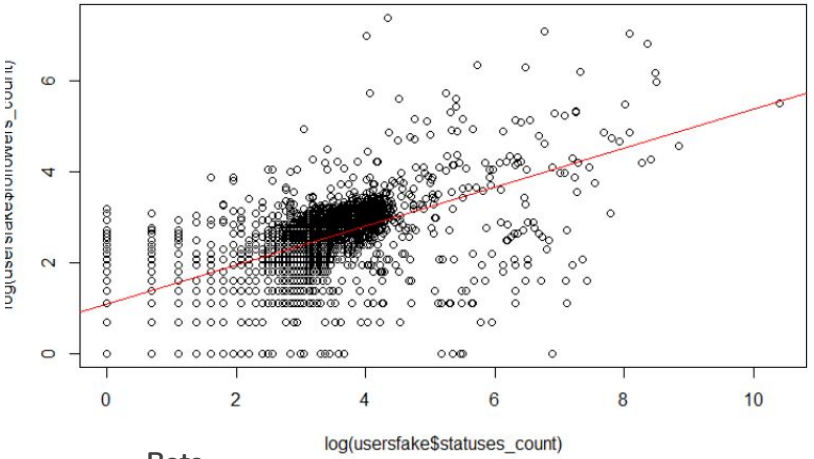
0,5435 – positiva fraca (< 0,7)

Resultados

Quantidade de Seguidores x Quantidade de Tweets



Usuários reais



Bots

Equação da reta: $f(x) = 0.9969 + 0.5655x$

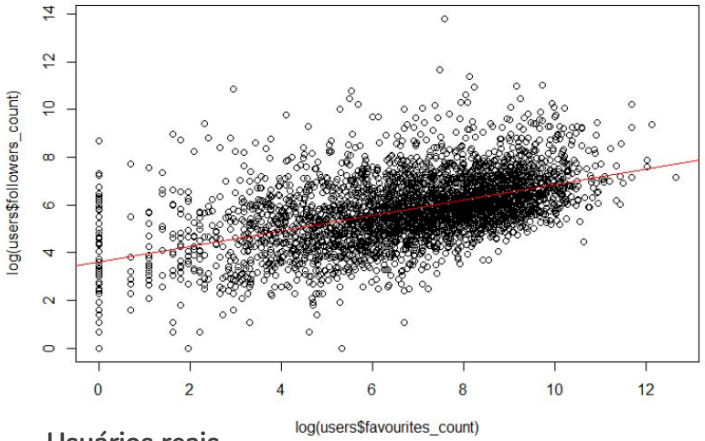
$f(x) = 1.0751 + 0.4312x$

Grau de correlação: 0,7115 – positiva forte

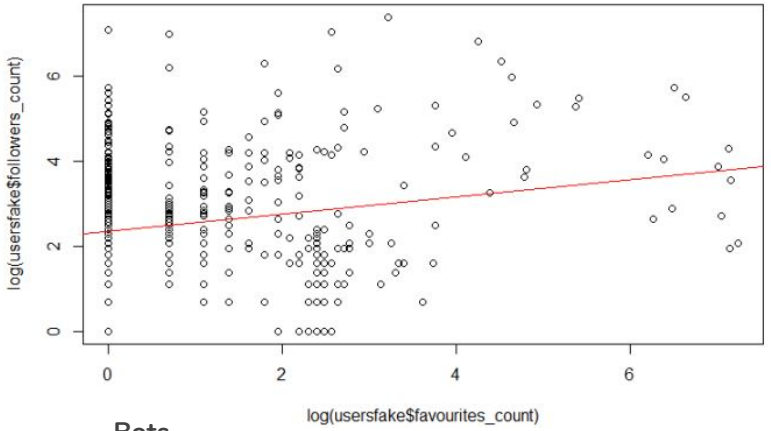
0,6360 – positiva fraca

Resultados

Quantidade de Seguidores x Quantidade de Tweets favoritados



Usuários reais



Bots

Equação da reta: $f(x) = 0,6660 + 0,8938x$

$f(x) = 2.3484 + 0.2033x$

Grau de correlação: 0,4956 – positiva fraca

0,1514 – positiva fraca*

Conclusão

- Concluiu-se deste trabalho que os métodos básicos (Médias, Desvios e Variâncias) conseguem mostrar diferenças consideráveis e por esse motivo, juntamente com o teste T-Student são eficientes para a detecção de bots maliciosos.
- A regressão Linear Simples também se tornou útil visto que houve diferença na correlação dos dados, alguns estão mais relacionados entre si que outros, o que pode influenciar para a detecção.
- O teste de Hipótese é essencial para comparar e analisar os dados das bases e com isso gerar uma diferenciação dos dados, separando-os de forma mais concreta.
- Apesar da Regressão Linear Simples ser muito útil, deve-se tomar cuidado em relação às características analisadas, principalmente com os valores de correlação próximos.

Trabalhos Futuros

- Desenvolver uma aplicação que consiga realizar e captar todos os dados e funções, retornando os valores de forma mais intuitiva do que é retornado no Rstudio.
- Analisar outro tipo de “*Social Bot*”, como por exemplo os “*Spams Bots*”, responsáveis por disparar várias mensagens e informações na redes sociais.
    - Comparar os resultados com os analisados neste trabalho.
- Incrementar com mais metodologias e fazer uma análise mais aprofundada do assunto, criando parâmetros de análise, separando por tipos de “*Social Bot*” analisados.

Referências

- BARBOSA, Kaio et al. **Botnets: Características e Métodos de detecção através do tráfego de rede**. Dissertação apresentada no SBSeg 2014 - XIV Simpósio Brasileiro em Segurança da Informação e sistemas computacionais.
- FREITAS, Carlos et al. **SocialBots: Implicações na segurança e na credibilidade de serviços baseados no Twitter.** Dissertação apresentada à Universidade Federal de Minas Gerais (UFMG). Disponível em: <http://sbrc2014.ufsc.br/anais/files/trilha/ST14-2.pdf>. Acesso em 3 de Junho de 2018
- SILVA, Ana Cristina. **Sistema de detecção de intrusão baseado em métodos estatísticos para análise de comportamento.** Tese de doutorado apresentada à Universidade Federal do Rio Grande do Sul. 2003.
- DataDome. **How to detect malicious bots.** Disponível em: <https://datadome.co/how-to-detect-malicious-bots/>. Acessado em 3 de Junho de 2018.
- DISTIL NETWORKS. **Distil's Bad Bot Report 2018: The year Bad Bots went mainstream.** Disponível em: <https://resources.distilnetworks.com/all-blog-posts/bad-bot-report-now-available>. Acesso em 4 de junho de 2018.