

RELATÓRIO DE ANÁLISE

sobre comportamento automatizado de perfis pedintes no Twitter

03 de agosto de 2021

Resumo

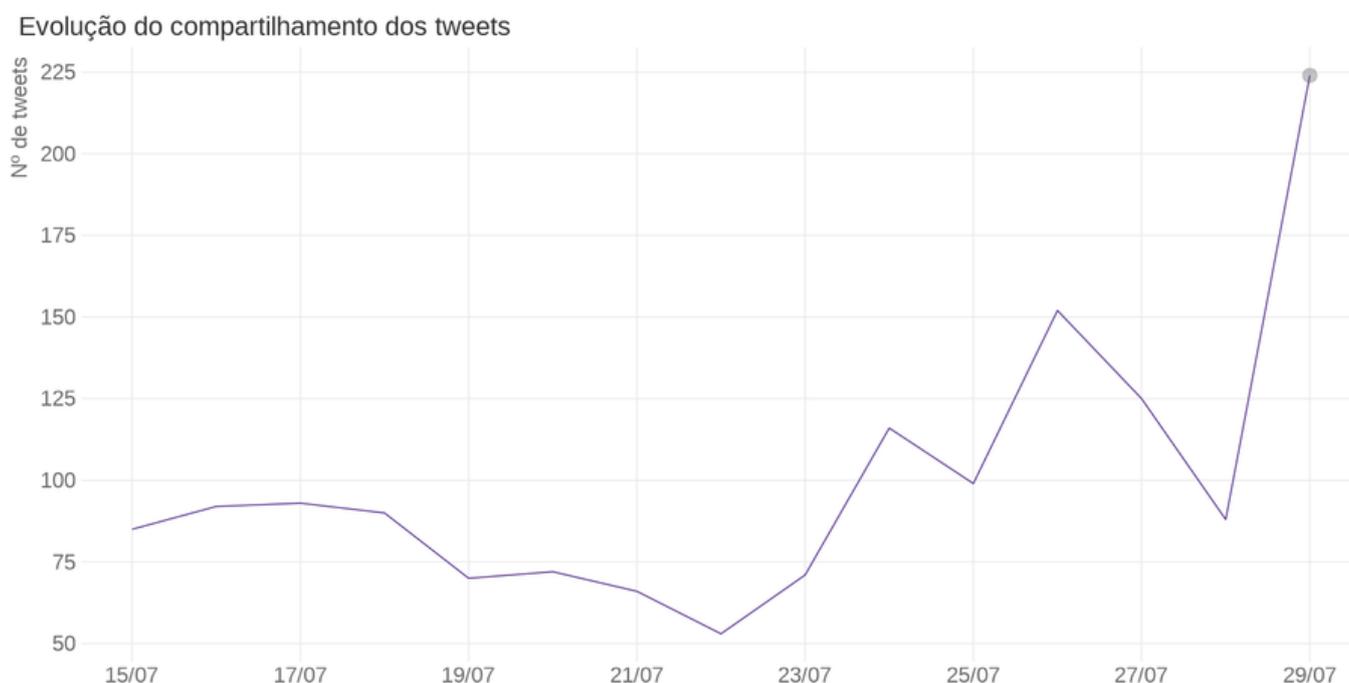
Com a popularização do Pix, meio de pagamento eletrônico que traz maior facilidade para a realização de transferências financeiras, tornou-se cada vez mais comum observarmos perfis no Twitter solicitando algum tipo de doação ou apoio financeiro para compra de comida, remédios e auxílio para tratamentos. O conteúdo dos tweets aparenta ser sempre muito similar e ocorre muitas vezes no formato de resposta a tweets de outros usuários. Isso acende um sinal de alerta sobre o possível uso de técnicas de automação para a publicação desse tipo de tweet e por isso merece atenção. Neste relatório levantamos alguns perfis que solicitam ajuda informando chaves Pix e analisamos com o Pegabot sobre o comportamento automatizado dessas contas.

Os principais pontos que você precisa saber

- 1 Coletamos tweets publicados em um período de 2 semanas de usuários que solicitam ajuda financeira por meio do compartilhamento de chaves Pix.
- 2 Identificamos que 4% dos usuários que participaram do compartilhamento de tweets com pedidos de doação ou auxílio financeiro por meio de Pix tem alta probabilidade de comportamento automatizado. Juntos eles publicaram 10% dos tweets analisados.
- 3 Identificamos usuários que chegaram a publicar mais de 70 vezes durante o período analisado, com conteúdo muito similar e variando as menções à outros usuários.
- 4 Verificamos, a partir da rede de interações que os usuários que mais recebem pedidos de ajuda são perfis verificados ou com grande número de seguidores.

Sobre os tweets coletados

Coletamos e selecionamos um conjunto de tweets que mencionaram a palavra *pix* juntamente com alguma das palavras *ajuda*, *ajudar*, *ajude*, *ajudem* e *Deus* desde o dia 15 de julho. Removemos desse conjunto os tweets de usuários que só publicaram uma única vez e tweets que, depois de um processo de limpeza para remoção de menções, hashtags e caracteres especiais, não apresentaram conteúdo repetido na base de dados. Em seguida, uma inspeção nos dados nos permitiu ainda remover alguns tweets fora de contexto: tweets que mencionam *pix* e alguma variação da palavra *ajuda* e *Deus*, mas que não solicitavam auxílio ou doações. Nosso objetivo com isso foi manter, dentro do possível, apenas tweets cujo conteúdo contém pedidos de ajuda financeira.



Como resultado, a base de dados contém 1496 tweets, não considerando RTs, referentes ao período de 15 dias. No total, 181 usuários foram responsáveis pelas publicações. No gráfico acima mostramos a evolução dessas publicações, contabilizando a quantidade de tweets publicados por dia.

Esses perfis têm comportamento automatizado?

Utilizamos o [Pegabot](#) para verificar sobre a existência de comportamento automatizado por parte dos usuários que publicaram os tweets coletados. O Pegabot analisa, individualmente para cada perfil, os dados públicos das contas disponíveis por meio de consultas à API do Twitter. Esses dados incluem, por exemplo, nome do perfil, descrição, quantidade de perfis seguidos e seguidores, número de postagens, além de uma amostra dos tweets da linha do tempo para identificar hashtags e menções. Com base nessas informações, o Pegabot estabelece um conjunto de 4 critérios, temporal, usuário, rede e sentimento, que juntos indicam a probabilidade de comportamento automatizado de uma conta, dando uma pontuação de 0 a 100. Quanto mais alto o valor, maior a chance do perfil ser automatizado.

Do total de

181

usuários únicos verificados

7 (4%)

apresentaram alta probabilidade de existência de comportamento automatizado

Esse resultado é referente a usuários que pontuaram com valor de, no mínimo, 70% de probabilidade na análise total do Pegabot. Esses perfis foram responsáveis por um volume de 10% (157) de publicações, volume relevante considerando a quantidade de usuários e o fato de estarmos analisando apenas tweets com conteúdo muito específico.

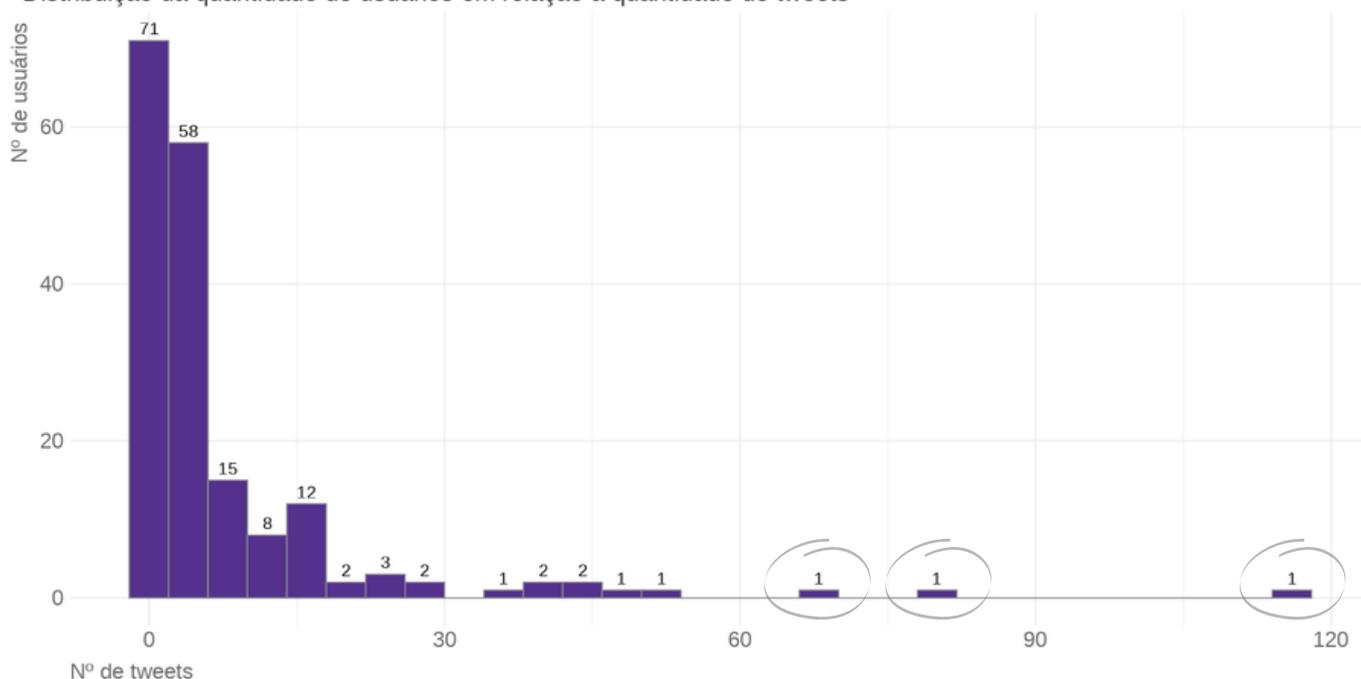
Verificamos também que no dia de maior pico de publicações, 29/07, 2 usuários com alta probabilidade de comportamento automatizado publicaram 28,6% dos 224 tweets compartilhados.

Alguns estudos [1,2] apontam sobre o fato de que a regularidade de postagens é um importante aspecto a ser considerado no processo de identificação de perfis automatizados. Nesse sentido, um olhar focado no critério temporal, analisado pelo Pegabot, indica que a quantidade de perfis com comportamento automatizado chega a 16,6% do volume de usuários. Juntos eles reúnem 22,6% do volume de publicações.

Sobre a frequência de postagens

O número de publicações que um usuário realiza no Twitter pode variar por uma série de motivos, como a finalidade da conta e número de seguidores, por exemplo. Para entender melhor sobre o comportamento dos usuários no que diz respeito ao volume de postagens, olhamos aqui para a frequência de postagens por usuários. Consideramos aqui todos os 181 usuários, e não somente os usuários identificados com comportamento automatizado pelo Pegabot. Numa primeira análise, notamos um desbalanceamento na distribuição da quantidade de usuários em relação ao número de tweets que eles publicam: 57,4% dos usuários analisados fizeram até 3 publicações. No gráfico a seguir apresentamos a distribuição, onde contabilizamos o total de usuários responsável por publicar um determinado número de tweets. Verificamos, a partir dessa análise, usuários que chegaram a fazer mais de 70 publicações durante as duas semanas:

Distribuição da quantidade de usuários em relação à quantidade de tweets



Sobre o conteúdo publicado

É comum observar em publicações de usuários que procuram por doações e ajuda financeira uma semelhança no conteúdo publicado por eles. Além disso, esses tweets são acompanhados de menções à outros usuários, indicando que eles foram diretamente marcados ou que ocorreram em resposta a outros tweets publicados. Pensando nisso, em um dos procedimentos de limpeza que aplicamos nos tweets coletados, e que também serviu de apoio para filtrar tweets relevantes, removemos menções, hashtags, pontuações, emojis, urls e caracteres especiais. Dessa forma, foi possível quantificar quais conteúdos mais aparecem na base coletada. Listamos abaixo os três principais e a quantidade de vezes que aparecem. Apesar dos tweets listarem dados pessoais como CPF e endereço, aqui optamos por anonimizar essas informações para não expô-las excessivamente.

“*Você sabia que com apenas R\$ 1,00 você pode ajudar a financiar a nossa luta? Faça sua contribuição e nos ajude chegar a Brasília para lutar contra o PL 490 e o Marco Temporal. Os Direitos dos Povos Indígenas estão em risco. Contribua! PIX: XXX*”

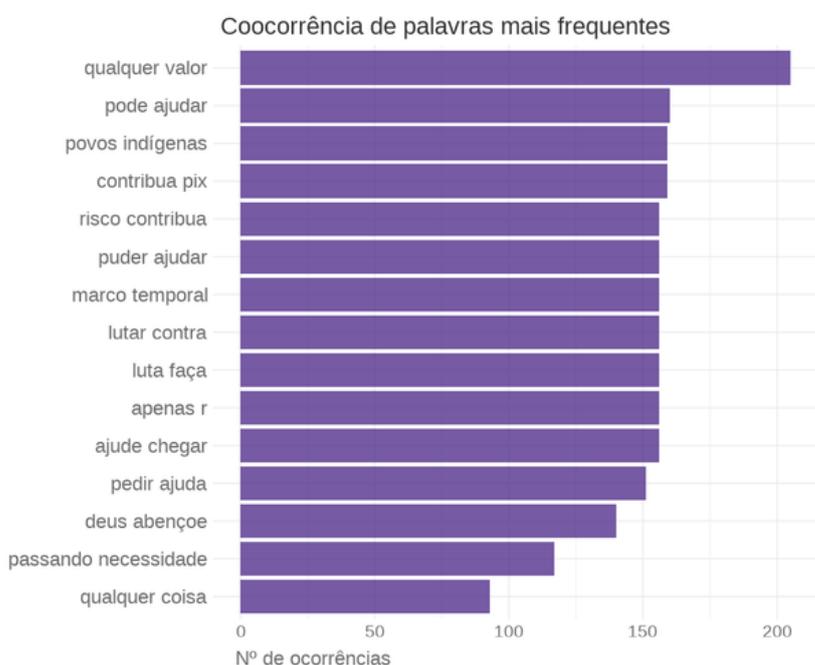
155 vezes, publicado por 2 usuários

“*□ALGUÉM PODE ME DOAR UMA CESTA BASICA? □POR FAVOR ESTAMOS SEM DISPENSA EM CASA. □ ENDEREÇO:XXX □Pix: XXX Meu instagram pra mais informações:@XXX □Deus abençoe a todos.□ #□*”

59 vezes, publicado por 1 usuário

“*Imploro por ajuda pq não tenho mais pra onde ir oro a Deus que ilumina no coração de alguém pra que possa nos abençoar e nós tirar dessa vida de sofrimento ã tenho nada nesse frio pros meus filhos quem tiver misericórdia dé nós qualquer coisa ajuda meu pix e meu cpf XXX*”

46 vezes, publicado por 1 usuário



Ainda sobre o conteúdo dos tweets, apresentamos ao lado o resultado da análise de coocorrência de termos, ou *bigram*, que busca por termos que frequentemente aparecem juntos em um *corpus*. O *corpus*, no nosso caso, é o conjunto de todos os tweets coletados.

A ocorrência dos termos *povos indígenas* e *marco temporal* é um reflexo do conteúdo mais compartilhado que listamos na análise anterior e os termos *apenas r* aparecem devido aos tweets que especificam valores esperados para doação.

Considerações finais

Trouxemos neste relatório um levantamento sobre a existência de comportamento automatizado na veiculação de tweets contendo pedidos de auxílio financeiro por meio do Pix. Focamos os processos de filtragem e análise de dados especialmente a partir de tweets que apresentaram conteúdo repetido, o que consideramos ser, para o contexto deste relatório, um indício relevante de automação. Destacamos que, assim como qualquer ferramenta de detecção de comportamento automatizado, o Pegabot tem limitações especialmente por conta das técnicas utilizadas para automação de perfis no Twitter serem atualizadas com frequência, impedindo que determinadas contas sejam identificadas pelo algoritmo. Da mesma forma, o Pegabot tem passado por um processo constante de avaliação e ajustes, para se manter atualizado frente às novas estratégias adotadas por esses perfis automatizados. Por esse motivo, procuramos não apontar sobre esses perfis de forma direta, mas sim entender sobre o uso de automação de forma mais ampla, identificando comportamentos que fogem do esperado ou observado.

Referências

- [1] Freitas, Carlos, et al. "Reverse engineering socialbot infiltration strategies in twitter." 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE, 2015.
- [2] Chu, Zi, et al. "Detecting automation of twitter accounts: Are you a human, bot, or cyborg?." IEEE Transactions on dependable and secure computing 9.6 (2012): 811-824.

Glossário

Frequência de tweets: indica quantos usuários publicaram um determinado número de tweets.

RT: republicação de um tweet próprio ou de outra pessoa que, para o relatório, contenha algum dos termos consultados na coleta de dados.

Reply: resposta dada à um tweet que, para o relatório, contenha algum dos termos consultados na coleta de dados.