Alexandre Pimentel: Identificação de fraudes feitas por click farms

A fraude de engajamento digital pode ser obtida pelas *click farms* (fazendas de cliques), que utilizam seres humanos para clicar em determinados conteúdos, muitas vezes através de trabalho escravo ou a um custo remuneratório irrisório, exatamente para projetar um rankeamento positivo, com o detalhe de não serem detectadas pelos sistemas de barreira cibernética que identificam quando os cliques são feitos por a tarefa.

Consistem em agrupamentos de pessoas que são contratadas com a finalidade específica de compartilhar conteúdos na Internet, escapando, dessa forma, ao controle cibernético de reconhecimento de robôs. Muitas delas situam-se fora do território nacional, em especial na Ásia, onde a mão de obra pelos serviços de cliques é extremamente barata. Segundo Nishant Kadian, a prova da fraude de cliques é tarefa difícil, pois pressupõe saber quem está utilizando um computador ou outro terminal de acesso à Internet e predizer quais são suas intenções. A tentativa de descobrimento de autoria através da identificação do IP não é eficaz, porém a análise de dados digitais revela o comportamento dos usuários na rede e pode fornecer informações úteis e confiáveis.

No entanto, as métricas cibernéticas, isto é, o método que mensura resultados pela análise dos dados comportamentais dos visitantes de determinado site ou rede social indicam um padrão de comportamento do usuário identificado pelo IP, e quando esse padrão é anormal isso pode sugerir ou indicar a presença de tentativas de fraudes [1]. Será, portanto, pela associação de algumas técnicas de cibersegurança que será possível combater a fraude de cliques.

Estatísticas demonstram que de cada quatro cliques realizados em anúncios um deles é fraudulento, pois aproximadamente 36% do tráfego da internet é gerado por *bots* e *scripts* automatizados. Além dos robôs, as *click farms*, isto é, organizações que podem produzir milhares de cliques em minutos a custos irrisórios de $ 1 para cada mil cliques, também contribuem para esse tipo de fake clicks.

Reportagem publicada no *The Guardian*, que repercutiu uma pesquisa sobre o poder da influência das *click farms* nas decisões dos consumidores, comprovou que 31% dos internautas tomam suas decisões sobre a compra dos produtos oferecidos nos anúncios digitais levando em consideração a avaliação dos demais usuários que "curtiram" ou "não curtiram" o anúncio, esse quantitativo de consumidores *"[…] verificará as classificações e avaliações, incluindo curtidas e seguidores no Twitter, antes de decidirem comprar algo, sugere a pesquisa. Isso significa que as fazendas de cliques podem desempenhar um papel significativo para enganar os consumidores"* [2].

As *click farms* são organizações ilícitas que contratam pessoas com a finalidade de clicar em determinados anúncios online, conforme o interesse do fraudador contratante, que é chamado de "fazendeiro" [3]. É um meio que tenta escapar da vigilância de softwares de proteção contra *click frauds* feitas por robôs, pois, como nas *click farms* os *fake likes* são feitos por seres humanos contratados, a identificação da fraude e sua prevenção são dificultadas [4].

A *click fraud* pode objetivar promover determinado produto, serviço, candidato etc., através de um falseamento de cliques excessivos que se prestam para dar a falsa impressão de adesão do público, bem como para fazer exatamente o contrário, ou seja, para desqualificar determinado adversário. Os especialistas consideram que a detecção de fraude de cliques em campanhas publicitárias pode ser feita através da análise conjunta dos seguintes aspectos:

a) Endereço IP: considerando que os bots executam inúmeros scripts semelhantes no mesmo servidor, ou seja, obedecem a uma série de instruções pré-determinadas para a execução das suas tarefas, isso indicará uma alta densidade de cliques derivados do mesmo IP ou de IPs semelhantes, assim a verificação dos IPs e o seu histórico é um meio de prevenção contra fraudes de cliques, mas esta técnica deve ser associada a outros métodos, pois isoladamente considerada não se constitui num meio plenamente seguro de se chegar à autoria das fraudes, nada obstante não deixa de ser um mecanismo hábil para, no mínimo, a produção de indício probatório;

b) *Click Timestamp:* o instante único evidencia um ponto específico na linha do tempo representado por uma codificação numérica, assim o timestamp mantém a hora em que o clique é feito, possibilitando a constatação do aumento na frequência de cliques durante determinado período [5]. Segundo Kadian, uma grande variedade de cliques com realizados em data e hora semelhantes é um indicativo de fraude de cliques e, mais ainda, se for constatada uma baixa duração associada a uma alta frequência de cliques isso significa uma alta probabilidade de fraude [6];

c) Proteção automatizada por softwares especializados contra *click fraud* — Há softwares de proteção automática contra fraude de cliques que se propõem a oferecer segurança completa para campanhas publicitárias.

As *click farms* podem ser identificadas através de sistemas de cibersegurança baseados em *deep learning* , que são capazes de perceber uma base comum de envio de cliques e, assim, desmascarar falseamento de geolocalização de IPs, bem como de identifica as máquinas que clicam mais de uma vez numa determinada postagem. As fraudes feitas por bots podem ser percebidas por meio de sistemas de IA de contraespionagem. Cuida-se de uma verdadeira guerra cibernética, na qual temos, de um lado, uma organização delitiva cujo objetivo é fraudar o ranqueamento digital de pessoas, produtos, serviços etc., e, de outro, as aplicações de Internet e outras organizações civis e governamentais que se dispõem a enfrentar esse tipo de delinquência.

A práxis comprova que as aplicações de Internet têm sim meios eficazes de controle de conteúdos ilícitos, bem como que as cláusulas dos T&C às quais aderem os usuários e que impõem restrições ao uso de conteúdos falsos, ofensivos, discriminatórios, *cyberbullying* e de controle de plágio são cláusulas válidas e eficazes e, ademais, que podem ser executadas diretamente pelos próprios provedores de aplicações, pois a efetivação dos T&C consiste em espécie de exercício regular de direito.

O inegável domínio que as *big techs* possuem sobre a tecnologia autoriza concluir que elas são as personagens mais bem qualificadas para combater a fraude e a desinformação na internet. Nesse sentido, o TJ-PR negou provimento a recurso de apelação de um indivíduo que não se conformava com a sentença que julgou improcedente pedido de declaração de inconstitucionalidade dos T&C, por supostamente ferirem a liberdade de expressão, visto que permitem o controle de conteúdo pelo próprio provedor, vejamos:

Alegada violação ao direito de liberdade de expressão e controle de conteúdo. Não ocorrência. Rede social que permite o uso de seus produtos e serviços mediante assinatura de termos e condições de uso. Vedação ao compartilhamento de conteúdo ofensivo, degradante, incitação ao ódio, discriminação, bullying. Exercício regular de direito. Autor que violou direito obrigacional e sofreu as sanções advertidas nos padrões da comunidade. Recurso conhecido e não provido [7].

A exclusão de conteúdos ilícitos pelas aplicações de internet é mais do que um direito dessas empresas respaldado pelos seus termos e condições, é um dever constitucional e legal de proteção da dignidade da pessoa humana, do direito à imagem e à honra, assim como pela necessidade de preservação da vida privada. Para mais além da relação jurídica de direito privado mantida entre as aplicações de redes sociais e os seus usuários paira um verdadeiro dever de natureza constitucional, o qual incide independentemente de qualquer regulamentação legal e que enluva tais aplicações e lhes impinge o dever de colaboração na proteção do ordenamento jurídico constitucional.

A jurisprudência respalda esse entendimento. Em demanda que tramitou pela primeira instância do TJ-MG, o juiz Sérgio Luiz Maia manteve a higidez dos T&C do Google, quanto à vedação de veiculação de conteúdos plagiados:

"[…] o réu provou que o conteúdo informativo utilizado pelo autor não estava de acordo com os termos e condições previstos no contrato, uma vez que a autora estava utilizando de conteúdo copiado/pouco original enquanto a Google já disponibilizava previamente que este tipo de conteúdo não era passível de monetização em sua plataforma, portanto, a Google não agiu de forma injustificada ao realizar o cancelamento do contrato. (TJMG 0043030-76.2017.8.13.0382)".

A liberdade de expressão é uma garantia constitucional não absoluta, sendo passível de limitação quando exorbitante, excitatória ao ódio, violência ou discriminação. Ao tempo em que o artigo 5º da CF garante o direito à liberdade de expressão, *pari passu*, também o limita à medida que, igualmente, protege o direito à honra, à privacidade… Importa acrescentar que o artigo 3º, I, da CF estipula que constituem objetivos fundamentais da República Federativa do Brasil a promoção do bem de todos, sem preconceitos de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação.

Assim, os conteúdos incitatórios podem e devem ser suprimidos pelos provedores de redes sociais, ainda que se trate de conteúdos discriminatórios implícitos, pois os conteúdos subliminares podem denotar a mesma semântica e a mesma ofensa de texto explícito, por isso devem ser suprimidos. Os algoritmos de IA são capazes de captar o real sentido de textos e áudios ou vídeos pela análise crítico-algorítmica do discurso [8].

A propósito, em 2021 o Facebook e o Twitter bloquearam as contas de Donald Trump, após constatarem que a convocação por ele feita para que os seus eleitores marchassem até o Capitólio no dia 6/1/2021, quando o Congresso iria sufragar o resultado das eleições de 2020, resultou numa insólita invasão por vândalos apoiadores do ex-presidente, protesto do qual resultou na morte de quatro pessoas, sendo que uma delas ocorrera dentro do Capitólio [9].

Semelhantemente, no Brasil, o viés identitário-neototalitário de grupos políticos que atuam em redes sociais, identificados como ‘milícias digitais’, também perpetram ataques à democracia e ao Estado de Direito. Em conclusão, o momento é deveras oportuno para se debater o papel das aplicações de redes sociais na defesa do Estado de Direito, sobretudo quando ainda está a tramitar PL que regerá a matéria.

[1] KADIAN, Nishant. *Click fraud prevention – Identify and reduce bot traffic in your paid ads.* Publicado em 07 de junho de 2019. Tradução livre.

[2] ARTHUR, Charles. *Facebook is riddled with click farms where workers sit in dingy rooms, bars on the windows, generating 1,000 likes for $1*. The Guardian, fevereiro de 2013.

[3] *In: How low-paid workers at 'click farms' create appearance of online popularity*. The Guardian, 02/08/2013.

[4] LEE, Munson. *What is a click farm? In:* https://pt.wikipedia.org/wiki/Fazenda_de_cliques#cite_note-1. Acesso em 14 de junho de 2022.

[5] KOTSUBO, Hugo. *O que é o timestamp. In:* https://hkotsubo.github.io/blog/o-que-e-timestamp. Acesso em 07 de janeiro de 2021.

[6] KADIAN, Nishant. *Click fraud prevention – Identify and reduce bot traffic in your paid ads.* Publicado em 07 de junho de 2019. *In:* https://mfaas.com/resources/click-fraud-prevention/. Acesso em 30 de dezembro de 2020.

[7] BRASIL, TJ-PR – APL: 00847723720178160014. Pub: 22/04/2020.

[8] FERNANDES, Paula. *Microsoft lança Security Copilot: O futuro da cibersegurança com inteligência artificial.* 28/3/2023. *In:* https://news.microsoft.com/pt-pt/2023/03/28/microsoft-lanca-security-copilot/.

[9] *In: Facebook e Instagram bloqueiam conta de Trump por tempo indeterminado.* G1. *In:* https://g1.globo.com/economia/tecnologia/noticia/2021/01/07/facebook-bane-conta-de-donald-trump.ghtml.