Reports on Twitter hack confirm existence of admin tool used for social media censorship

Kevin Reed
24 July 2020

News reports published in the week since hackers breached Twitter security and hijacked dozens of high-profile accounts have confirmed that administrators working for the social media company use a dashboard to blacklist and censor content down to the level of specific users and their individual tweets.

A report in the *New York Times* on July 17 confirmed that a screenshot, originally shared by *Motherboard* on the day of the hack and reported previously here on the WSWS, is authentic and shows that Twitter's backend administrative tool includes several buttons including "Search Blacklist," "Trends Blacklist" and "Notifications Spike," among others.

Along with these features—showing clearly that Twitter employees have the ability to "throttle" or "shadowban" tweets or users on their platform—the tool also can be used to modify email addresses associated with accounts. This latter function enables the transfer of a Twitter account from one person to another and is one of the features of the dashboard that the hackers used to gain control of the high-profile accounts.

On the afternoon and evening of July 15, anonymous hackers accessed Twitter's internal employee dashboard and hijacked the accounts of major public figures—including Barack Obama, Kanye West, Bill Gates and Elon Musk—each of which have tens of millions of followers.

In what Twitter described as a "coordinated social engineering attack," the hackers repeatedly published tweets from the breached accounts as part of a scheme to dupe the public out of money in the form of Bitcoin cryptocurrency.

Messages posted by the hackers said, in the case of Barack Obama's account for example, "I am giving back to my community due to Covid-19! All Bitcoin sent to my address below will be sent back doubled. If you send $1,000, I will send back $2,000!"

On that day, Twitter management responded in its own tweets that it was aware of the attack and was working to "fix it." However, after hours of struggling to regain control of their own platform, Twitter was forced to block all verified accounts from posting any tweets.

In subsequent reports from Twitter, the company admitted that the hackers gained control of 130 accounts and had downloaded data from at least eight of these using the "Your Twitter Data" user tool. In a corporate blog post on Saturday, Twitter reported that the hackers "manipulated a small number of employees and used their credentials to access Twitter's internal systems … the attackers were able to initiate password reset, login to the account and send Tweets."

In an update to the blog post on Wednesday, the company reported, "We believe that for up to 36 of the 130 targeted accounts, the attackers accessed the DM inbox, including 1 elected official in the Netherlands. To date, we have no indication that any other former or current elected official had their DMs accessed." DM stands for direct messaging which is private communications between individual Twitter users.

Finally, Twitter said, "We're embarrassed, we're disappointed, and more than anything, we're sorry. We know that we must work to regain your trust, and we will support all efforts to bring the perpetrators to justice." The day after the event, Twitter's shares fell 4 percent, wiping out $1.3 billion in value.

The *Times* story on Friday included interviews with four individuals who participated in the hacking operation. The report said, "The interviews indicate that the attack was not the work of a single country like Russia or a sophisticated group of hackers. Instead, it was done by a group of young people—one of whom says he lives at home with his mother—who got to know one another because of their obsession with owning early or unusual screen names, particularly one letter or number, like @y or @6."

Although the *Times* published a screenshot of the Twitter admin dashboard, the journalists and editors focused entirely on the fact that someone named "Kirk" transmitted it to the youthful hackers as proof that he had breached Twitter security and gained access to the backend tool and did not raise a single question about the functionality that is clearly displayed in the image.

By not raising any questions about the admin panel, the *Times* is essentially functioning as a PR operation for Twitter. Far from providing an explanation of the blacklisting and account manipulation features of their backend tool, according to The Verge, Twitter has been busy "removing images of the screenshot from its platform and in some cases suspending users who continue to share it."

Other tech news sites have also raised questions about the admin panel such as Reclaim the Net, which published an article on July 15 entitled, "Leaked screenshots appear to show internal Twitter tool that can blacklist users from search and trends."

Referring to such practices as "shadowbanning," Reclaim the Net reporter Tom Parker wrote, "At the start of the year, Twitter officially made shadowbanning, a controversial practice that involves limiting the distribution or visibility of user posts in a way that's difficult to detect, part of its terms of service. Now new leaked screenshots from Motherboard appear to show an internal Twitter user administration tool that can be used by Twitter staff to blacklist user accounts from search and trends."

Parker continued, "In June, the pop culture satirical news account Price of Reason documented how Twitter had shadowbanned one of its viral tweets that made fun of HBO Max's controversial decision to stop Bugs Bunny's hunter adversary Elmer Fudd using a gun in its Looney Tunes remake.

"The account owner noticed a drastic slowdown in engagement after his tweet had started to go viral and discovered that he was being blacklisted from Twitter search, causing both his account and the tweet to be scrubbed from search results. 'It's as if neither it or I ever existed,' Price of Reason told Reclaim the Net."

Another story on Telecoms.com by Scott Bicheno on July 15 said, "We don't know exactly what being blacklisted entails, but the name of the tools strongly implies accounts can be prevented from appearing in searches and trending lists, even while they're still otherwise active. We're also not aware of any precedent for accounts being notified when they are placed on one of these blacklists, which adds weight to claims that Twitter seeks to manipulate conversation on its platform through the means of 'shadow banning.'"

These revelation about the details of Twitter's methods for manipulating the account activity of users—either by shadowbanning or by deleting accounts entirely—shows that censorship is a top level activity and priority of the social media company.

On Tuesday, Twitter announced that it had removed 7,000 accounts allegedly active in spreading far-right QAnon conspiracy theories. The essential views of the QAnon "movement" are the belief that Donald Trump is waging a secret war against an elite cabal of devil worshiping pedophiles who rule the world.

QAnon also claims that the Democratic Party is behind international crime rings and that "deep state" figures are waging a war against Donald Trump. The group's content has spread widely on Facebook, TikTok, Twitter and YouTube.

In a tweet from Twitter Safety, the company said, "We've been clear that we will take strong enforcement action on behavior that has the potential to lead to offline harm. In line with this approach, this week we are taking further action on so-called 'QAnon' activity across the service."

According to a report in Reuters, Twitter will roll out the account suspensions this week and are expected to impact about 150,000 accounts globally. It said the initial 7,000 accounts were removed for "violating the company's rules against spam, platform manipulation and ban evasion."

The mass shutdown of Twitter accounts is a blatant act of censorship by the social media platform. It is not the responsibility of the giant tech monopolies—Google, Facebook, Twitter, etc.—to determine what is authentic or truthful information and what can be seen or read by the public.

The use of advanced software tools to throttle tweets or ban groups such as the far-right QAnon proponents is part of a much larger effort by the corporate and financial elite and capitalist state to gain control of online and social media information and communications. Under conditions of growing working class struggles, these measures are ultimately aimed at blocking and shutting down the dissemination of revolutionary socialist ideas among masses of people all over the world.