# Hackers Can Tell What Netflix Choices You Make And Spy On Your Mind

*Netflix*

Netflix made a splash when it debuted *Black Mirror: Bandersnatch* in December, a "choose your own adventure"-style movie that put viewers in charge of their cinematic destiny. It has since invested in even more interactive programming, including a live-action show featuring survivalist Bear Grylls. But researchers at the Indian Institute of Technology, Madras have found that whether you're warping a dystopian future, adventuring in the Minecraft universe, or staring down a boa constrictor, your data may not be as secure as you think.

Netflix took the big, difficult, commendable step of encrypting all of its video streams in 2016 to better protect user privacy. That layer of security makes it much more difficult for a "man in the middle" between Netflix's servers and a user's browser to track what customers watch. In practice, though, the researchers say that they can analyze Netflix's encrypted interactive video traffic to find clues about what users are watching, and which choices they've made in their movie journeys.

"I work on analysis of encrypted network traffic, and when we stumbled upon this Netflix movie *Bandersnatch* it was something very new," says Gargi Mitra, a PhD student at IIT Madras. "But when I was looking at the choice-making interactions it turned out that they are similar to other kinds of interactions in web applications and web sites I study. So I tried out some of my techniques and we were able to determine which options the viewer chooses."

> "Encryption may hide content, but it does not hide traffic patterns, and traffic analysis can reveal important secrets."

Vitaly Shmatikov, Cornell Tech

Though Netflix added HTTPS in 2016, numerous studies have found that a man in the middle with access to encrypted video data can use attributes of the stream, particularly the size of the data files Netflix sends users, to figure out what people are watching. Past research has shown that even if an attacker doesn't have access to network traffic, they can still use malicious web scripts to get the information, and "fingerprint" the videos Netflix users are watching.

The IIT researchers can go even deeper to monitor interactive choices. They found that at each decision branch Netflix considers one of the options the "default," or more likely choice, and the other to be a backup. The service stages the default video so it's ready to play and then sends a packet that contains a notation file, known as a JSON file, that contains information about the viewer's choice. The researchers found that they could use characteristics of the pre-queued stream, the size of the JSON file, and a file header used by the encryption protocol—known as the SSL record length—to determine which choice the user made at each step.

In an analysis of choice data from 100 viewers, the researchers determined the decisions correctly 96 percent of the time.

Netflix says that such an attack would be difficult to carry out in practice, because it requires access to network traffic for analysis. The company also points out that it's much more difficult to identify and contextualize video stream data out in the open internet versus in a controlled research scenario.

The IIT Madras researchers point out, though, that it is feasible for attackers to trick users into connecting to rogue routers or access points. And even if you're not worried about that, your internet provider and VPNs inherently sit in this position. They may be interested in tracking what web users are watching and choosing on streaming services, because it can help them sell ads or monetize their own offerings.

Whether you voted to make Bear Grylls eat a bug is not especially sensitive personal data. But the findings fit into larger concerns about minimizing and safeguarding data generated by interactive media and streaming services more broadly. In February, University College London researcher Michael Veale discovered through a GDPR request that Netflix retained records on user Bandersnatch choices. IIT's Mitra points out that while this type of information may seem minor, it may include some decisions that people could want to keep private, such as whether characters will consume illegal drugs.

"This research confirms an important lesson that has been demonstrated time and again," says Vitaly Shmatikov, a data privacy and network security researcher at Cornell Tech, who worked on a 2017 study about identifying Netflix users' browsing traffic. "Encryption may hide content, but it does not hide traffic patterns, and traffic analysis can reveal important secrets *without* breaking encryption. As video systems become more adaptive and interactive, traffic analysis will reveal more information about users’ private choices." Shmatikov was not involved in the IIT research.

The IIT Madras researchers say that they submitted their findings to Netflix's bug bounty program, and the company acknowledged their validity. The submission was rejected for being "out of scope," though, because the data leak partly stems from the encryption protocol, which Netflix does not control. The researchers say that Netflix could mitigate the data exposure by changing how it compresses the JSON files, making them harder to distinguish and analyze in the encrypted stream. But the researchers also note that other types of analysis attacks may exist that would defeat even that added protection.

Web encryption is vital to security and privacy online, but the group's findings are a reminder that the web community needs to develop better ways of masking encrypted streams so they aren't inadvertently leaking some information—and revealing your late-night Netflix habits in the process.