TechXplore

Fake-Waking Voice Assistant: Research provides new insights into the 'fake wake phenomenon'

September 27 2021

Credit: Unsplash/CC0 Public Domain

The international researcher team has discovered new insights and valuable finding that manufacturers of voice assistants can utilize to enhance the security of their voice assistants as well as the privacy of users. The team investigated the eight most popular English and Chinese voice assistants regarding the "fake wake phenomenon."

Voice assistants typically actively listen to the environment for their native "wake words" such as "Alexa," "OK Google" or their brand names, to get activated. The fake wake phenomenon appears when the voice assistant detects false wake-up words, so-called "fuzzy words" from, e.g., through conversations or TV shows running in background. These incorrectly recognized words can be used by attackers to activate voice assistants without the user's awareness.

For the first time, the team led by Prof. Wenyuan Xu, Dr. Yanjiao Chen and Prof. Ahmad-Reza Sadeghi has succeeded in automatically and systematically generating their own false wake words instead of screening audio material. The generation of these fuzzy words started with a known initial word such as "Alexa." The researchers neither had access to the model that recognizes the wake-up words nor to the vocabulary on which the voice assistant is based. In this context, they also investigated the root causes for the acceptance of incorrect wake-up words.

First, they identified the features that most frequently contributed to the acceptance of fuzzy words. The determining factors focused only on a small phonetic section of the word. However, the voice investigated assistants were also able to activate false words that differed significantly from the real wake-up words, whereby, surrounding noises, the volume of the words and the gender of the speaker hardly played a role.

Using genetic algorithms and machine learning, more than 960 custom fuzzy words in English and Chinese could be generated, which activated the "wake-up word detector" of the voice assistants. On the one hand, this shows the severity of the fake wake phenomenon, and on the other hand, it provides a deeper understanding of it.

The impact of fake wake phenomenon can be significantly reduced by retraining the wake word detector of the voice assistant with the generated fuzzy words. This allows the voice assistant to more accurately distinguish between fake and real wake-up words. Voice Assistant manufacturers can highly benefit from the results of this research and re-train the existing models to make them more accurate and less vulnerable to fake wake attacks, and thus increase the security of own products and improve user's privacy.

More information: Yanjiao Chen et al, FakeWake: Understanding and Mitigating Fake Wake-up Words of Voice Assistants. arXiv:2109.09958v1 [cs.LG], arxiv.org/abs/2109.09958

Provided by Technische Universitat Darmstadt

Citation: Fake-Waking Voice Assistant: Research provides new insights into the 'fake wake phenomenon' (2021, September 27) retrieved 9 April 2024 from https://techxplore.com/news/2021-09-fake-waking-voice-insights-fake-phenomenon.html