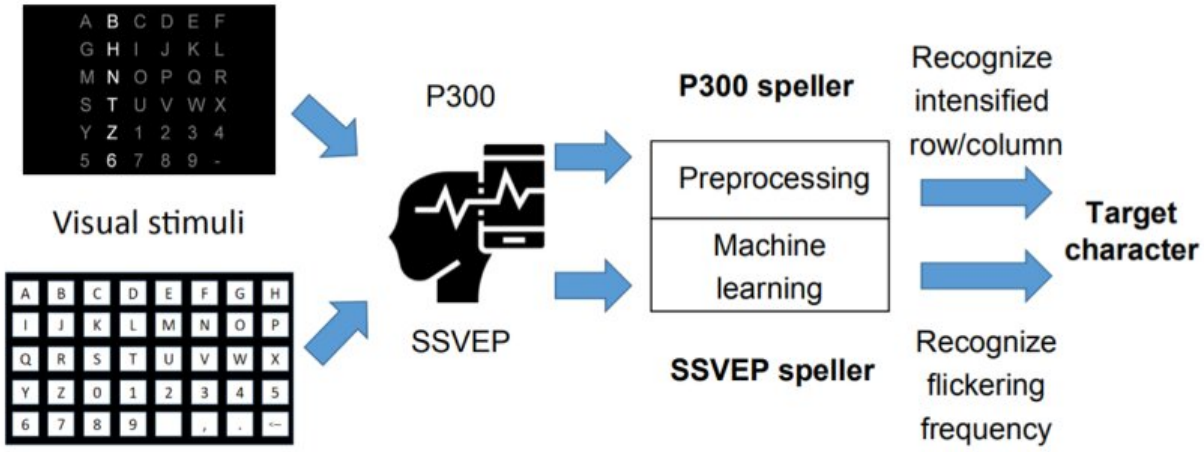
Are brain-computer interface spellers secure?

October 5 2020



For each speller, the user watches the stimulation interface, focusing on the character he/she wants to input, and EEG signals are recorded and analyzed by the speller. The P300 speller first identifies the row and the column that elicit the largest P300, and then outputs the character at their intersection. The SSVEP speller identifies the output character directly by matching the user's EEG oscillation frequency with the flickering frequency of each candidate character. Here is a demo of an SSVEP based BCI speller, developed by a co-author of this paper (Tzyy-Ping Jung) and his collaborators: Credit: @Science China Press

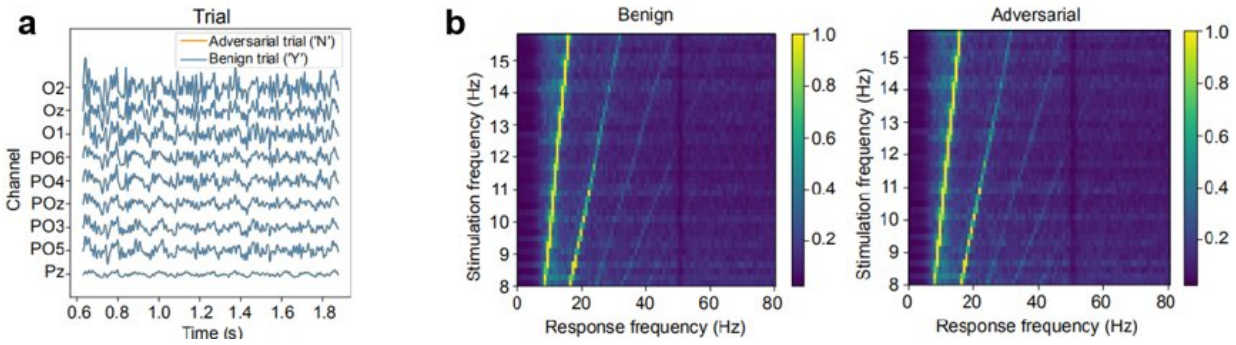
Brain-computer interfaces (BCIs) record and decode brain signals to construct a communication pathway, which allows people to interact with a computer by thought directly. BCIs have been used in a broad range of applications, including neuroscience, psychology, clinical

rehabilitation, and entertainment. As one of the most commonly used BCI systems, the electroencephalogram (EEG)-based BCI speller, which allows the user to input text to a computer using EEG signals, is particularly useful to severely disabled individuals, e.g., amyotrophic lateral sclerosis patients, who have no other effective means of communication with another person or a computer. However, one very important question remains: are these BCI spellers really secure?

In response to this question, a BCI research team, led by Professor Dongrui Wu from Huazhong University of Science and Technology (HUST), has recently published an article in the Beijing-based *National Science Review*, which shows that the output of BCI spellers may be easily manipulated by tiny adversarial noise, exposing a critical security concern in EEG-based BCIs.

"It shows for the first time that one can generate tiny adversarial EEG perturbation templates for target attacks for both P300 and SSVEP spellers, i.e., mislead the classification to any character the attacker wants, regardless of what the user intended character is. The consequence could range from merely user frustration to severe misdiagnosis in clinical applications," they stated in their article entitled "Tiny noise, big mistakes: adversarial perturbations induce errors in Brain-Computer Interface spellers".

"More seriously, these perturbation templates are so tiny that one can barely distinguish the adversarial EEG trial from the original EEG trial. When drawn together, the signals look almost completely overlapping," they added. "The adversarial perturbation templates can even stay imperceptible to some widely-used approaches for evaluating the quality of EEG signals."



a) SSVEP signals before and after adversarial perturbation (almost completely overlapping). The output text was changed from "Y" to "N"; b) Spectral analysis for benign and adversarial SSVEP signals. Credit: @Science China Press

"These spellers' ability to defend adversarial perturbation templates is totally different from their robustness to random noise," they further emphasized. "Even the BCI spellers which show excellent performance against random noise can be manipulated by these deliberately-designed perturbation templates with a high successful rate."

They also stated that this security concern is not only specific to the victim models used in these EEG-based spellers, but also other popular BCI systems. "It should be noted that the goal of this study is not to damage EEG-based BCIs. Instead, we aim to demonstrate that serious adversarial attacks to EEG-based BCIs are possible, and hence expose a critical security concern, which has received little attention before," said Professor Wu. "Our further research will focus on addressing this security issue and making BCI systems safer."

More information: Xiao Zhang et al, Tiny noise, big mistakes: adversarial perturbations induce errors in Brain-Computer Interface spellers, *National Science Review* (2020). DOI: 10.1093/nsr/nwaa233

Provided by Science China Press