World Academy of Science, Engineering and Technology
International Journal of Cognitive and Language Sciences
Vol:17, No:11, 2023

Detection Method of Federated Learning Backdoor Based on Weighted K-Medoids

Authors : Xun Li, Haojie Wang

Abstract : Federated learning is a kind of distributed training and centralized training mode, which is of great value in the protection of user privacy. In order to solve the problem that the model is vulnerable to backdoor attacks in federated learning, a backdoor attack detection method based on a weighted k-medoids algorithm is proposed. First of all, this paper collates the update parameters of the client to construct a vector group, then uses the principal components analysis (PCA) algorithm to extract the corresponding feature information from the vector group, and finally uses the improved k-medoids clustering algorithm to identify the normal and backdoor update parameters. In this paper, the backdoor is implanted in the federation learning model through the model replacement attack method in the simulation experiment, and the update parameters from the attacker are effectively detected and removed by the defense method proposed in this paper.

Keywords : federated learning, backdoor attack, PCA, k-medoids, backdoor defense
Conference Title : ICSLP 2023 : International Conference on Speech and Language Processing
Conference Location : San Francisco, United States
Conference Dates : November 06-07, 2023