Jian-hua LI, 2018. Cyber security meets artificial intelligence: a survey. *Frontiers of Information Technology & Electronic Engineering*, 19(12):1462-1474. https://doi.org/10.1631/FITEE.1800573

Cyber security meets artificial intelligence: a survey

Key words: Cyber security; Artificial intelligence (AI); Attack detection; Defensive techniques

Corresponding author: Jian-hua LI
E-mail: lijh888@sjtu.edu.cn
ORCID: Jian-hua LI, http://orcid.org/0000-0002-6831-3973

Introduction

When artificial intelligence (AI) meets cyber security, the cross-disciplinary studies focus on two aspects:

1. AI technologies, such as deep learning, can be introduced into cyber security to construct novel smart models to implement malware classification and intrusion detection, and threat intelligence sensing.

2. AI models will face various cyber threats, which will disturb their sample, learning, decision. Thus, AI models need specific cyber security defense and protection technologies to resolve the problems of adversarial machine learning, privacy-preserving machine learning, secure federated learning, etc.

Artificial intelligence against cyberspace attacks

1. Traditional machine learning schemes
   (1) *k*-nearest-neighbor;
   (2) Support vector machine;
   (3) Decision tree;
   (4) Neural network.

2. Deep learning solutions:
   (1) Deep belief network;
   (2) Recurrent neural network;
   (3) Convolutional neural network;
   (4) Automatic encoder based solutions.

API 11	API 17	API 8	API 7	API 2	API 12	Result
158	190	210	231	55	87	Normal(10.0)
125	201	166	105	8	112	Malware(73.0)
97	130	290	303	72	21	Malware(14.0)
130	78	194	316	21	4	Malware(7.0)
21	96	203	255	43	53	Malware(3.0)
58	166	189	178	19	22	Normal(20.0)
85	167	158	214	6	20	Malware(3.0)

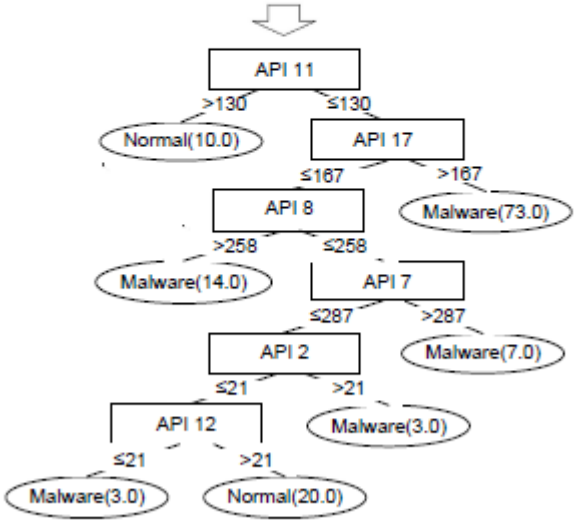

Fig. 2 A decision tree construction for malware detection

Security threats and defensive techniques of AI

1. Adversarial attacks on AI:
   (1) Inject poisoning data;
   (2) Add a small amount of modified Images;
   (3) Imposing only a little adversarial perturbation;
   ….
2. Defense methods against adversarial attacks:
   (1) Modify training process and input data;
   (2) Modify network;
   (3) Use an additional network.

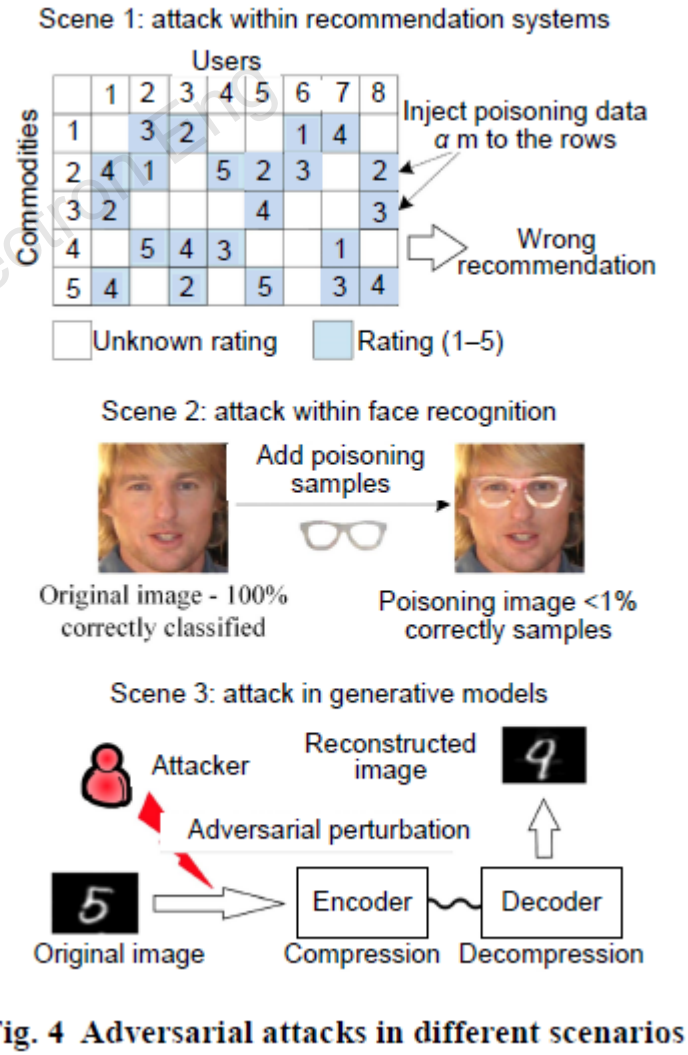

Fig. 4 Adversarial attacks in different scenarios

Security threats and defensive techniques of AI

Construction of safe AI system:

1. Safe distributed ML/DL systems:
   (1) Privacy-preserving deep learning;
   (2) Secure federated deep learning.

2. Machine learning classification over encrypted data:
   (1) machine learning classification protocols over encrypted data:
   - hyperplane decision
   - Naïve Bayes
   - decision trees

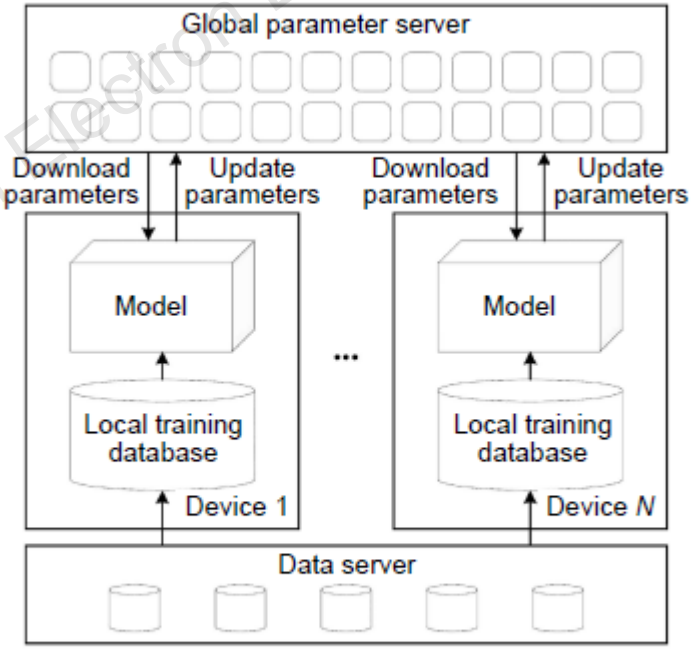

Fig. 5 Safe distributed machine learning/deep learning systems

Conclusions

1. We have summarized the integration of AI and cyberspace security from two aspects:

   Review the use of AI related technologies (ML/DL) to detect and resist various types of attacks in cyberspace.

   Review various attacks from which AI systems may suffer in an adversarial environment, and the defense strategies for different kind of attacks.

2. Discuss how to build a safe AI system in a distributed ML/DL environment.

3. While using AI to protect cyberspace security, building a secure AI in the future also requires sufficient attentions.