Research on Experiential Solidification Learning Technology for Intelligent Mining of Power System Vulnerabilities

Di Zhao¹, Yan Li¹, Yang Dong¹, Jiaming Ni²

¹State Grid Tianjin Information & Telecommunication Company, Tianjin, 30010, China
²State Grid Tianjin Electric Power Company, Tianjin, 30010, China

Abstract. Aiming at the problems that the experience of network security personnel in power system is difficult to solidify, and the basic and repetitive vulnerability mining work is intelligently replaced by the lack of experience guidance, this paper studies the experience solidification learning technology of power system vulnerability intelligent mining. Format specification and parsing language of vulnerability intelligent mining experience tree are proposed, which can transform flexible and complex vulnerability mining experience into a language that can be recognized and understood by machine. At the same time, it puts forward the solidification and intelligent learning technology of vulnerability mining experience, through the automatic pervasion test for the target, combined with reinforcement learning algorithm, intelligent learning generates vulnerability mining experience tree, and finally forms the high-efficient transformation of the red team personnel experience and the intelligent substitution of part of the work.

1. Introduction

With the integration of Internet, mobile Internet and Internet of Things, the informationization of China's power industry is developing rapidly. The massive electric power web system has realized the connection of more than 400 million smart meters, and has satisfied the friendly interactive use of the company's e-commerce and various kinds of marketing apps of more than 180 million registered users. All kinds of interconnection and interoperability of the company, and a large number of business systems constitute a powerful power grid social public service platform. Under the new situation that the new attack means of network security and the 0Day vulnerability derive constantly, the work pressure of the security hidden trouble investigation and the security protection supervision and inspection of the massive electric power web system is huge¹. At present, there are more than 160 network security vulnerability pervasion verifiers in China’s power grid enterprises. Facing such a large number of business system environment, the traditional manual vulnerability mining and supervision and inspection based on professionals can no longer meet the requirements of power security protection. It is urgent to use artificial intelligence technology to solidify the professional vulnerability pervasion experience, and to realize the intelligent substitution of basic repetitive work.

Aiming at the above problems, this paper studies the experience solidification and learning technology of power system vulnerability intelligent mining. Firstly, the overall framework of experience solidification and intelligent learning of vulnerability mining is designed. On this basis, this paper further design vulnerability mining experience solidification script markup language, and standardize its parsing process. Finally, the key steps of the experience solidification and learning of vulnerability mining are designed.
2. The general idea of experience solidification learning in intelligent vulnerability mining of power system

For the research of experience solidification and expert system construction technology of vulnerability intelligent mining, firstly, we need to study the format specification and parsing language of vulnerability intelligent mining experience tree, and transform flexible and complex vulnerability mining experience into a language that can be recognized and understood by the machine; At the same time, the solidification technology of vulnerability mining experience is studied. Before facing the actual Web target system, through the training of pervasion test on the target system is established, and it is associated with the vulnerability mining scripts. Finally the accurate "Web system-vulnerability mining method" relationship is constructed, and the targeted vulnerability mining of the Web system is realized under the guidance of experience. The red team personnel experience into specific steps and rules stored in the vulnerability mining and utilization of the experience tree, and finally formed for the red team personnel experience of the efficient transformation. The overall research route is shown in Fig.1:

![Figure 1. System Frame Diagram](image)

3. Vulnerability mining script markup language DPL and parsing process design

Markup language is a system of annotating documents in a semantically text-sensitive manner. The idea and terminology evolved from the "markings" of a paper manuscript, editors' instructions for revision, which were traditionally written in blue pencil on the author's manuscript. Marking originates from the traditional practice of marking manuscripts. When a manuscript needs to be published, it is first handed over to the Marker, who marks the text on the manuscript to indicate what font, style, and size should be used in what part of the manuscript, and then hands the manuscript over to the person for manual typesetting upon completion of the marking. In addition to taggers, tags are often used by editors, proofreaders, publishers, and graphic designers, as well as by document authors. The main markup languages are GenCode, troff, nroff, TeX, Scribe, GML, SGML, HTML, XML and so on.

The DPL language designed in this paper refers to the characteristics of extensible markup language. Aiming at the special data structure of pervasion experts' experience, a markup language is developed to standardize the input parameter specification, implementation method and result specification of vulnerability mining verification script. The invention has the advantages of fast reading, simple structure, easy transplantation and high retrieval efficiency. The main function of DPL interpreter is to
analyze the expert experience of DPL format and initialize the experience tree and decision automata according to the experience content. The location of the DPL language interpreter in the overall intelligent process exploited by vulnerability mining is shown in Fig.2.

![Figure 2. Schematic diagram of DPL language interpreter](image)

4. Research on experience solidification and learning technology of power system vulnerability mining

The core idea of the experience solidification and learning technique for power system vulnerability mining is to use reinforcement learning\(^4\) to learn how to use the best payload to pervasive specific port service vulnerability and store the learning results in the form of experience tree. This paper collects the multi-dimensional expression of the protocol, port and other related information of the target system, constructs the correlation analysis model of the expression and vulnerability mining script based on the experience tree, and realizes the pruning and optimization of the vulnerability mining experience tree based on the machine learning method\(^5\). The technology uses Metasploit as the underlying tool to automatically perform port scanning and service detection. The main features are as follows:

- Effective vulnerability using: Attempts to exploit vulnerability at least once at a potential vulnerability site;
- Deep pervasion: If successfully pervasive a target, it will further pervade the internal server;
- Self-regulated learning: Reinforcement learning is used to learn how to exploit vulnerabilities without the need to prepare training data\(^6\);
- Strong intelligence gathering capability: Automatically access the target intelligence, including operating system, port service information, mainly using the following methods: Port Scan, Naive Bayes for HTTP analysis, content exploration\(^7\).

5. Execution process

The main process of this technology is divided into the following four steps: Intelligence gathering, pervasion, post-pervasion, report generation.

5.1. Intelligence gathering

In the intelligence gathering stage, nmap is mainly used to scan the ports, and obtain the necessary information of the target’s port information, service information and operating system information in the pervasion stage. In addition to using nmap, HTTP pages crawled on the web port are predicted using Navie Bayes and analyzed for content. But whatever the method, the ultimate goal is intelligence gathering. Once the information is gathered, two data structures are constructed: Exploit_tree and target_tree (built only at the first run, and then run to read both data structures directly from the local
file). Exploit_tree is an abstraction of the exploit module in Metasploit, organized as a multi-tree structure. Target_tree is the abstraction of target intelligence and is also organized as a multi-tree structure.

5.2. Intelligent pervasion training

The training and testing of reinforcement learning model mainly take place in the pervasion stage.

A) Problem abstraction

The state of the problem is abstracted as a 5-dimensional vector (ostype, servicename, serviceversion, exploitmodule, target). The action is defined as selecting a payload available under the current state. If the payload is used to successfully pervade the target, a positive Reward is returned and the episode ends. Otherwise, a negative reward is returned and the state is converted.

B) Training process

First, start the thread in the main method and train the model:

```python
if __name__ == '__main__':
    ...
    # Execute learning.
    for worker in threads:
        job = lambda: worker.run(exploit_tree, target_tree, saver,
                                  env.save_file)
        t = threading.Thread(target=job)
        t.start()
```

In the Worker_thread.run method, the Environment.run method is executed iteratively, each time the execution of the method corresponds to an episode:

```python
class Worker_thread:
    def run(self, exploit_tree, target_tree, saver=None, train_path=None):
        while True:
            if self.thread_type == 'learning':
                # Execute Learning thread.
                self.environment.run(exploit_tree, target_tree)
                # Stop learning thread.
                if isFinish:
                    ...
            else:
                ...
```

The Environment class has two important class member variables: Agent and Metasploit classes, in which the Metasploit class is the real interactive environment of Agent, and executes actions through the execute_deploy method of the Metasploit class, and dynamically converts the environment. When episode ends, the global parameter network is updated.

A FIFO queue named memory and a LocalBrain class member are maintained in the Agent class to hold n-step information, and each call to the advantage_push_local_brains() method evaluates the non-bootstrap rewards before n-step:

\[
R_{t-n+1:t} = R_{t-n+1} + \gamma R_{t-n+2} + \gamma^2 R_{t-n+3} + \cdots + \gamma^{n-1} R_t
\]

And the state \(s\), action \(a\), and \(R_{t-n+1:t}\) before n-step, the state \(s_{\_}\) after n-step are sent to the training sample queue of self.brain.

5.3. Post-pervasion

Experienced pervasion person can pervasive further into the internal server during the pervasion phase of a session that was successfully opened. Because the infiltrator/system cannot connect directly to the internal server, further pervasion is performed through the first pervaded server. The technology repeats these processes on the pervasive machine, thus achieving deep pervasion.
5.4. Experience generation
When pervasion is complete, the vulnerability information for successful pervasion is aggregated to generate a pervasion result report.

6. Summary
Aiming at the problem that the experience of vulnerability mining is difficult to solidify and the machine is lack of experience guidance, this paper innovates the solidification method of manual mining experience and instructs the machine to construct the experience of vulnerability mining through intelligent training. It puts forward the format specification and parsing language of the experience tree of vulnerability intelligent mining, at the same time, puts forward the solidification technology of the experience of vulnerability mining, transforms the experience of network security professionals into the concrete steps and rules experience tree, and finally forms the substitution ability for the experience of network security professionals and the work of intelligent vulnerability mining.

ACKNOWLEDGMENTS
This work is Supported by the science and technology project of State Grid Tianjin Electric Power Company: "Research on power information system vulnerability mining and monitoring analysis technology based on machine learning"

References
[1] Jia Yan, Qi Yulu, Shang Huaijun, Jiang Rong, Li Aiping. A Practical Method of Constructing Network Security Knowledge Map [J]. Engineering, 2018, 4 (01): 117-133.
[2] Liao Xi. Overview of Network Security Vulnerability Pervasion Testing Framework [J]. Computer Products and Circulation, 2018 (10): 36.
[3] Lin Jing. Research on Heterogeneous Data Exchange Model Based on XML Technology [J]. Journal of Jiujiang University (Natural Science Edition), 2017, 32 (03): 89-91.
[4] Lu Zhike, Kang Xiaofeng, Sui Zhenyi, Wang Hao, Liu Jiqing, Sun Chongchong. Web Application Vulnerability Scan Detection System [J]. Software Guide, 2019, 18 (08): 186-190 +195.
[5] Mei Guojun. Research and Implementation of Vulnerability Mining Based on Genetic Algorithm and Model Constraints [D]. Beijing University of Posts and Telecommunications, 2019.
[6] Chen Yu, Zhao Jingya, Zhu Qinghua, Liu Yong. Adaptive Adjustment of WLAN Nodes Based on Reinforcement Learning [J]. Computer Engineering and Design, 2019, 40 (09): 2422-2427.
[7] Bie Yidong. Research and Implementation of Specific WEB Vulnerability Analysis System [D]. Nanjing University of Posts and Telecommunications, 2017.