2016 IEEE 24th International Conference on Network Protocols (ICNP)
Poster Paper

ASCEND: A Search Engine for Online Industrial Control Devices

Xuan Feng*, Qiang Li†, Haining Wang ‡, Limin Sun*
* Institute of Information Engineering, Chinese Academy of Sciences, China
† School of Computer and Information Technology, Beijing Jiaotong University, China
‡ Department of Electrical and Computer Engineering, University of Delaware, USA

***Abstract*—Industrial control system (ICS) devices with IP addresses are accessible on the Internet and play a crucial role for critical infrastructures like power grid. However, there is a lack of deep understanding on these devices' characteristics in the cyber space. In this paper, we propose *ASCEND*, a search engine for online industrial control devices. *ASCEND* analyse 17 industrial protocols and use it to discover almost all online ICS devices in the IPv4 while reducing the noise of industrial honeypots. It provides a big picture of online ICS devices: who are using ICS devices; where they are located and what functions these ICS device have. In order to demonstrate how *ASCEND* works, we have implemented the prototype system and verified it in the real-world experiments.**

I. Introduction

The number of industrial control system (ICS) devices with computing and communication capabilities is rising rapidly, and they are crucial for infrastructure-critical systems, such as power, oil, and gas pipelines, water distribution, and wastewater collection systems. Supervisory control and data acquisition (SCADA) is often used to control remote ICS devices with coded signals, and these ICS devices are typically computer-based systems with access to the Internet. The deep understanding on characteristics of these online ICS devices in the cyber space is highly desirable for the interests of infrastructure-critical systems, which is unfortunately still missing.

The first characteristic of online ICS devices is where they are located. The distribution features of ICS devices would give a big picture about network measurement. We could keep track of important devices that can be directly accessible from the Internet. The second characteristic is who uses these ICS devices. It can help us understand why they are deployed on the Internet. Owner can save time and money to pay attention to these online devices. The third one is what functions these online ICS devices have. There are power plants, solar panel and power controller that can use these industrial protocols. Various types of industrial devices can be found on the Internet for remote access and and data transmission. Digital footprints of ICS devices can help secure network security.

In this paper, we propose *ASCEND*, a search engine for online industrial control devices. It use industrial protocol banners to discover online ICS devices in a real-time and non-intrusive manner. Table I shows 17 typical industrial protocols that are widely adopted in industrial control systems. These ICS devices typically run a variety of industrial protocols in the application layer over standardized TCP/IP protocols. There is abundant information encapsulated in the packet header of these protocols, such as device vendor, type and function. We could use such information to identify industrial control devices by analyzing the contents. *ASCEND* sends common ICS protocol requests to remote networks and determine whether it runs the ICS protocol or not by matching the receiving packet with pre-defined words. The approach design and prototype system is detailed in our previous paper [1].

TABLE I: 17 industrial control protocols for online ICS devices

Category	ICS Protocols
TCP-Based	OMRON FINS, HART-IP, Siemens S7, Mddbus, IEC 104, DNP3, EtherNet/IP, Tridium Niagara Fox, PCWorx, ProConOS, CodeSys, Red Lion Crimson V3, General Electric SRTP, CSPV4, Automatic Tank Gauge
UDP-Based	BACnet, OMRON FINS, MELSEC-Q, HART-IP

We deployed it on cloud computing server to collect online exposed ICS devices. After the ICS data collection, all the search result is transmitted and register to our local data center, then index using the Elasticsearch framework for quickly search. *ASCEND* exposes data to researchers through a public search engine and REST API. The user also can use the web interface to perform full-text searches about the online collection result. To observe the characteristics of online ICS devices in a relatively long run, *ASCEND* provides the following three service:

- Tracking online ICS devices. *ASCEND* stores all the history data of the online ICS devices into the MongoDB and build a fast index for the latest search result with E-lasticSearch. Users would simply use *ASCEND* to explore these ICS devices and keeping their history track.
- Showing the Big Picture. *ASCEND* use open database MaxMind's GEOIP [2] to map between ICS device address and location. It also provides online rDNS to map each device address to its domain name. Users would use *ASCEND* to know where ICS devices are located and who is using them.
- Exploring ICS devices. We extract these information

† Qiang Li is corresponding author.

978-1-5090-3281-5/16/$31.00 ©2016 IEEE

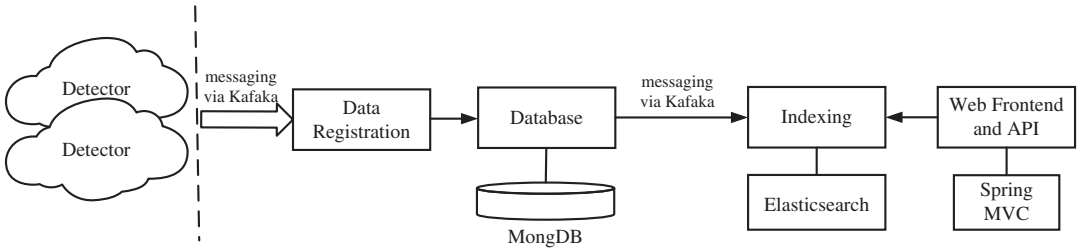

Fig. 1: *ASCEND* Architecture: system implementation component diagram.

(device type) based on ICS protocol packet format payload. *ASCEND* infer their functions based their detail information and provide restful API for users to further explore these online ICS devices.

II. Implementation

In this section, we describe our software prototype and system implementation. In order to demonstrate how *ASCEND* works, we have implemented a prototype system as a self contained piece of software based on open source libraries.

Figure 7 shows our system implementation component diagram. Detector modular are deployed in the cloud server, collecting online ICS devices. It first send a single TCP packet that has "SYN" filed in header and "NULL" data payload to determine whether the host is alive or not. If the host responds with "SYN-ACK" in the TCP header or response the pre-generated packet, we put this host into the set of candidates, otherwise throw it away. Then detector modular would use industrial protocol (Table I) to identify whether the candidate is a ICS device.

The collection data are encapsulated as an unified message format and transmitted via Apache Kafka [3]. Kafka provides the functionality of a messaging system with a fast and scalable manner. The data format is the tuple as $(IP, Protocol, Type, Time, Data)$. The data registration modular keeps the tracks of online ICS devices. If a new ICS device tuple appear, we add it in the register list. If an existing tuple comes, we update it and replace its timestamp. The major reason is that dynamic host configuration protocol (DHCP) would lead to devices dynamically join and leave the network.

After the data registration stage, the data is then stored into the NoSQL database MongoDB. MongoDB provides the support for fast reading and writing JSON-like documents. Then we adopt Elasticsearch [4] to build the index about device data. Elasticsearch is a open source search and multitenant-capable full-text search engine. When we want to see a certain period of ICS devices, the database transfer the corresponding dataset via Kafka to our index modular.

ASCEND provides a programmatic API following the semantics of a REST API. For example, we can get the distribution of online ICS devices during a certain time, by sending a GET request to our system. The system would feed the data as JSON file as response data. A graphical user interface is also needed for general users to esaily access *ASCEND*,

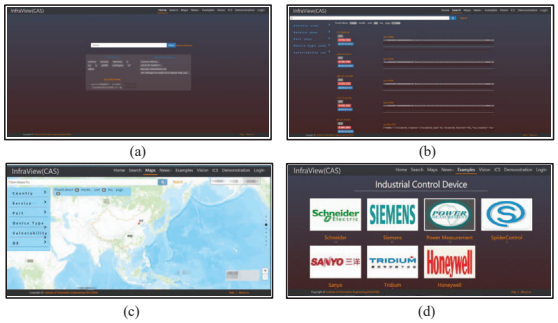

Fig. 2: : The functions of *ASCEND* search engine provided.

this frontend web interface is implemented using Spring Web model-view-controller (MVC) framework. Figure 2 shows the four webpages about the function of our system. Figure 2(a) is the front page of our search engine; the part (b) is to show a list of ICS device search results; Fig. 2(c) shows the geographical distribution of online ICS devices and Fig. 2(d) provides the exploration for ICS devices.

III. Conclusions

ICS devices with IP addresses are accessible on the Internet and play a crucial role for critical infrastructures. In this paper, we propose *ASCEND*, a search engine for online industrial control devices. It could discover ICS devices in a real-time and nonintrusive manner and characterized these devices on the Internet. We have implemented the prototype system and provide three functions for characterizing online ICS devices: (1) tracking online ICS devices; (2) showing the big picture and (3) exploring ICS devices.

IV. Acknowledgments

This work was supported in part by the Fundamental Research Funds for the Central Universities (K15RC00060) and the Research on the universality of mobile sensing-based indoor fingerprint positioning (No. 61401300).

References

[1] X. Feng, Q. Li, H. Wang, and L. Sun, "Characterizing industrial control system devices on the internet," *24th IEEE International Conference on Network Protocols (ICNP 2016)*, 2016.
[2] Maxmind geoip2. [Online]. Available: https://www.maxmind.com/en/geoip2-services-and-databases
[3] Apache kafka, messaging. [Online]. Available: http://kafka.apache.org/
[4] Elasticsearch, search ands analyze data in real time. [Online]. Available: https://www.elastic.co/products/elasticsearch