Design and Evaluation of Enumeration Attacks on Package Tracking Systems

Hanbin Jang, Woojoong Ji, Simon S. Woo, and Hyoungshick Kim(✉)

Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon, Republic of Korea
{hanbin,woojoong,swoo,hyoung}@skku.edu

Abstract. Most shipping companies provide a package tracking system where customers can easily track their package delivery status when the package is being shipped. However, we present a security problem called *enumeration attacks* against package tracking systems in which attackers can collect customers' personal data illegally through the systems. We specifically examine the security of the package tracking websites of the top five popular shipping companies (Korea Post, CJ Logistics, Lotte Logistics, Logen, and Hanjin Shipping) in South Korea and found that enumeration attacks can be easily implemented with package tracking numbers or phone numbers. To show potential risks of enumeration attacks on the package tracking system, we automatically collected package tracking records from those websites through our attack tool. We gathered 1,398,112, 2,614,839, 797,676, 1,590,933, and 163,452 package delivery records from the websites of Korea Post, CJ Logistics, Lotte Logistics, Logen and Hanjin Shipping, respectively, during 6 months. Using those records, we uncover 4,420,214 names, 2,527,205 phone numbers, and 4,467,329 addresses. To prevent such enumeration attacks, we also suggest four practical defense approaches.

Keywords: Package tracking systems · Enumeration attack · Privacy

1 Introduction

Most shipping companies provide a web service to allow people to track their packages and monitor the status of their package information online. A *package tracking number* (PTN) or phone number is popularly used to check and track the real-time package delivery status. That is, if a user enters a valid PTN or his/her phone number, the package tracking website displays the corresponding package status information along with some types of personal information, such as the full or partial name of the sender or the receiver, time-stamps, transit locations, the expected delivery time, etc. Such package tracking systems are widely used in the shipping industry because they are highly usable and convenient for customers to monitor and track their packages without directly logging in to the shipping company's website. As an example, Fig. 1 shows the package tracking website

J. K. Liu and H. Cui (Eds.): ACISP 2020, LNCS 12248, pp. 543–559, 2020.
https://doi.org/10.1007/978-3-030-55304-3_28

provided by UPS (https://www.ups.com). If a user enters a valid PTN into the input text field called '`Track`', the website displays not only the details of the package delivery status but also additional personal information (e.g., name, phone number and address) of the sender or the recipient.

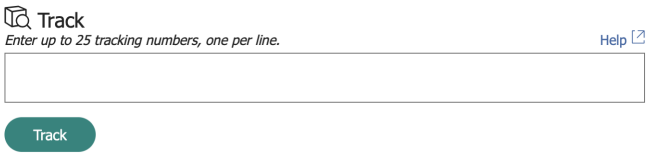

Fig. 1. UPS website for the package tracking service.

However, in most services, we found that explicit user authentication is not required in package tracking websites. We surmise that a package tracking system is generally designed for even non-members of the system to use their services with ease because the sender and/or the recipient can be a non-member of the system who cannot login to the website. At first glance, this package tracking service seems to be a useful feature, because the package tracking status information is only provided to the recipient and/or the sender who know the corresponding PTN or user's phone number. As long as these PTNs or phone numbers are kept confidential among legitimate parties, displaying information can be adequate. However, if those PTNs and phone numbers are guessable, then any 3rd party can also see the displayed information. We are wondering whether this feature can potentially be abused to harvest customers' personal data such as customers' names, phone numbers, and addresses at large scale; those stolen data would be abused or sold for conducting ads or additional cyber criminal activities such as sending spam/phishing messages [13] or creating Sybil accounts [11]. Recently, Woo et al. [15] showed the possibility of *enumeration attacks* with the top three package service providers (FedEx [4], DHL [3], and UPS [5]) in which the enumeration attack is a type of dictionary attack in which an attacker tries each of a list of possible candidate values (in a valid format) to determine the correct secret values (e.g., email addresses, phone numbers, and PTNs) through an online verification tool. However, their work was focused on those three service providers only and did not explain how the existence of these attack vectors is systemically detected and tested. Our work is motivated by extending their research to additional services for generalization and developing a systematic method to analyze the attack vectors related to web enumeration attacks on different websites.

To achieve these goals, we first analyze the main causes of enumeration attacks and then develop a framework to identify attack holes that can be exploited to perform enumeration attacks. As case studies of our framework, we chose the top five most popular shipping companies (Korea Post, CJ Logistics, Lotte Logistics, Logen, and Hanjin Shipping) in South Korea and then analyzed attack vectors of their package tracking websites.

To show the feasibility of enumeration attacks identified by our systematic method, we implemented a tool to automatically collect users' personal data in package tracking services by enumerating a specific range of PTNs or phone numbers. Although we focused on analyzing package tracking systems in South Korea, our attack techniques were not designed to solely work on specific companies or countries. Our framework is generic enough and can be extended to any package tracking systems in the world, which provide web-based status checking information, as shown in Fig. 1. Our contributions are summarized as follows:

- We present a framework to systematically analyze enumeration attacks against package tracking systems. Our research is the first to examine how web enumeration attacks can be tested, exploited, and detected systematically.
- We implemented the automatic enumeration attack generation tool against package tracking systems. Using this tool, we collected more than 4 million package delivery records and identified more than 4 million unique names, 2 million unique phones, and 4 million unique addresses. We clearly show that existing package tracking systems are at a real serious risk of revealing their customers' data.
- We propose four practical defense approaches for package tracking systems such as limiting the number of PTN verification failed attempts, using CAPTCHAs, generating unpredictable PTNs, and minimizing information leakage from those systems to reduce the chance of enumeration attacks. Our proposed defense approaches would be integrated into the existing systems without incurring significant costs.

2 Design of Enumeration Attacks

In this section, we explain how enumeration attacks can be launched automatically to harvest users' personal data from a target website. In the target website, enumeration attacks can be implemented by sending a sequence of request messages for a specific service in the target website and monitoring the corresponding responses in an automated manner. To generate valid request messages, attackers should follow the data formats and protocols used in the service. Figure 2 shows the overview of the automatic enumeration attack testing framework consisting of four steps. In the following sections, we present the process of each step in detail.

2.1 UI Analysis

Given a web page as an input, the goal of this step is to analyze the web page components and discover all web forms (e.g., `<input type="text">`) that can be potentially exploited to perform enumeration attacks. The identified web forms are passed to the step of "data-flow analysis." We note that hidden fields can be often used to implement enumeration attacks. Therefore, we also need to consider hidden fields as candidate web forms for enumeration attacks.

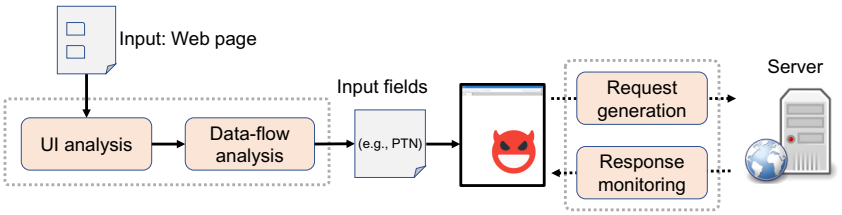

Fig. 2. Overview of the proposed framework for automated enumeration attack testing.

2.2 Data-Flow Analysis

With the web forms delivered from "UI analysis," we narrow down the list of further possible candidate input fields (i.e., web forms) whose values can be enumerated. In practice, it is hard to identify such input fields without any information about input field formats. Therefore, we first collect or generate some initial request message samples and then analyze the format of each input field with those request messages. We can use a heuristic technique to determine whether the values for each input field can be enumerated or countable by checking whether input field values consist of (decimal or hexadecimal) digits only; if an input field value contains characters other than digits, we remove the field from the list of candidate input fields for enumeration attacks because it would be difficult to define a rule to enumerate such input field values in an automated manner.

2.3 Request Generation

Once candidate input fields are determined, the framework computes input field values according to some pre-defined rules to enumerate input field values. Then, a web testing tool generates service request messages containing an input field value and sends it to the target web server.

2.4 Response Monitoring

The final step is to monitor and verify the response from the web server. We can determine whether a service request message (containing enumerated input field values) is correct or not, according to the query response result. The request is successful if the query response is successfully returned; otherwise, it is failed. Next, if successful, the proposed framework extracts the user data from the query result. After finalizing this step, we go back to the step of "request generation." All steps can be repeatedly carried out to harvest a sufficiently large number of user data.

3 Analysis of Services in Package Tracking Systems

To show the feasibility of our framework for performing enumeration attacks in an automated manner, we analyze the services of package tracking systems.

We first aim to investigate the attack surface of package tracking systems, where enumeration attacks can be performed. To achieve this objective, we analyzed the top five most popular Korean package tracking websites (Korea Post, CJ Logistics, Lotte Logistics, Logen and Hanjin Shipping), which can offer several services related package delivery. Table 1 presents the input parameters, which are needed to access each service of package tracking websites.

Table 1. Input parameters needed for each service of package tracking websites.

Company	Checking the status	Changing the drop-off location	Requesting the receipt	Returning the package
Korea Post	PTN	PTN	Authentication code	PTN, Recipient's name
CJ Logistics	PTN	–	–	PTN, Recipient's phone number
Lotte Logistics	PTN	–	–	PTN
Logen	PTN	–	PTN	PTN, Recipient's phone number
Hanjin Shipping	PTN	–	–	–

From Table 1, we can observe that most services can be accessed with PTN and recipient's name or phone number. We aim to exploit those services by enumerating PTNs or phone numbers because they can be enumerated based on our initial PTN structure analysis. Unlike other systems, the receipt requesting service at Korea Post requires an internally generated authentication code (e.g., `Hw17WzULnQ9BgnPZmd`), which we will explain more in the next section.

In each package tracking website, the following four services are commonly offered: 1) checking the package delivery status, 2) changing the drop-off location, 3) requesting the package delivery receipt, and 4) returning the package to the sender. Figure 3 presents some examples of each service. However, we note that the above four services can be abused to harvest users' information if we fail to protect PTNs or phone numbers from guessing. The detailed description of each service and the types of displayed personal information are provided as follows:

1) Checking the Package Delivery Status. As shown in Figure. 3a, a user can check the expected delivery time and tracking details. We can obtain the following information through the package delivery status checking service: 1) the package delivery status, time, and item; 2) the sender's *masked* name and address; 3) the recipient's *masked* name and address; and 4) the courier's name

and phone number. For example, the sender's and recipient's names (e.g., ***soo Kim) can be masked to hide their full names, while the courier's full name is revealed. Table 2 summarizes the detailed package, sender, recipient and courier information provided by each company. Not surprisingly, the degree of information provided through this service slightly varies across different companies.

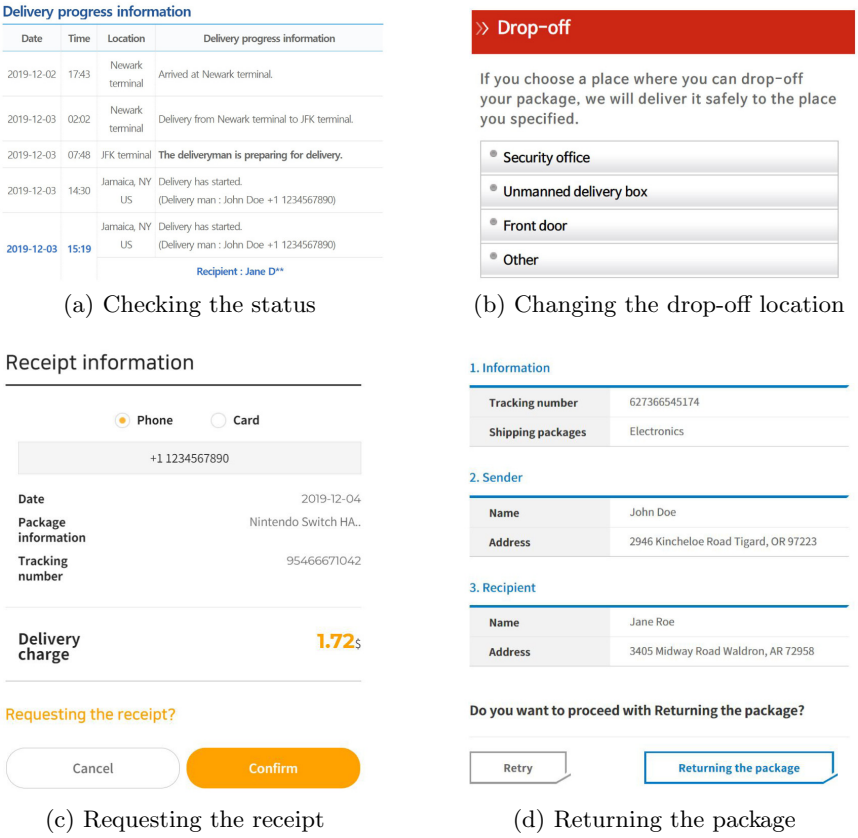

Fig. 3. Four common services in package tracking websites.

Table 2. Information types obtained from the package delivery status checking service.

Company	Package	Sender	Recipient	Courier
Korea Post	Status, Time	Masked name	Masked name	Name, Phone number
CJ Logistics	Status, Time , Item	Masked name	Masked name	Name, Phone number
Lotte Logistics	Status, Time	City	City	Name, Phone number
Logen	Status, Time	Masked name, City	Masked name, City	Name, Phone number
Hanjin Shipping	Status, Time, Item	Masked name	Masked name, City	Name, Phone number

2) Changing the Drop-Off Location. In all services, the package is directly delivered to the recipient's home address by default. However, recipients can often change their final drop-off location. As shown in Fig. 3b, the recipient can choose a drop-off location to security office, unmanned delivery box, front door, or other places.

Although three package service providers (Korea Post, CJ Logistics, and Logen) offer an option to change the drop-off location, CJ Logistics, and Logen do not provide any information through this service. From only Korea Post, we can obtain the following additional information through the drop-off location changing service: 1) the package item (e.g., electronics, books, etc.); and 2) the recipient's name and address.

3) Requesting the Package Delivery Receipt. Senders and recipients can further request the receipt of payment for a proof of the package delivery. Two package service providers (Korea Post and Logen) offer an option to display the receipt of the payment for the package delivery, as shown in Fig. 3c. In particular, we can obtain the following auxiliary information through the receipt requesting service: 1) the package item, 2) the sender's name, phone number and address, and 3) the recipient's name, phone number, and address. Table 3 summarizes the auxiliary information types provided by each company.

Table 3. Information types obtained from the receipt requesting service.

Company	Package	Sender	Recipient
Korea Post	–	Name	Name, Address
CJ Logistics	–	–	–
Lotte Logistics	–	–	–
Logen	Item	Name, Phone number, Address	Name, Phone number, Address
Hanjin Shipping	–	–	–

4) Returning the Package to the Sender. Recipients often want to return the received items to the senders. Therefore all package service providers except Hanjin Shipping offer an option to allow users to return the received item through their website. As shown in Fig. 3d, this service typically displays the sender's and recipient's details such as their names and addresses. We can obtain the following user and package information through the package returning service: (1) the package item, (2) the sender's name, phone number, and address, and (3) the recipient's name, phone number, and address. In Logen, the recipient's name and phone number are masked, while the sender's name and phone number are fully visible in plain text. Table 4 summarizes the auxiliary information types provided by each company.

Table 4. Information types obtained from the package returning service.

Company	Package	Sender	Recipient
Korea Post	Item	Name, Phone number, Address	Name, Phone number, Address
CJ Logistics	Item	Name, Phone number, Address	Name, Phone number, Address
Lotte Logistics	Item	Name, Phone number, Address	-
Logen	Item	Name, Phone number, Address	Masked name, Masked phone number, Masked address
Hanjin Shipping	–	–	–

4 Experimental Results

To show the feasibility of the proposed enumeration attacks presented in Sect. 2, we implemented a tool to perform enumeration attacks in Python 3.7. We also used an open source automated web testing tool, `Selenium` (https://selenium.dev), to modify query cookies and HTTP headers. For testing, we executed this tool on the Ubuntu 16.04 64-bit running on an Intel(R) Core(TM) i5-6500 CPU (with 16 GB RAM), equipped with a 100 MB LAN connection.

If the package tracking record is successfully displayed on a web page while performing enumeration attacks, we can extract specific customer's data from the web page. Interestingly, in some services (e.g., the receipt requesting service at Logen, the drop-off location changing service at Korea Post, and the package returning service at CJ logistics), customers' data is not directly visible in the web page because values are presented in hidden fields in the HTML source file. Therefore, we use Chrome Devtools (https://developers.google.com/web/tools/chrome-devtools) to extract hidden field values from the source file.

4.1 Enumeration Attacks with PTNs

We first manually collected several initial PTNs used in the four package tracking systems (Korea Post, Lotte Logistics, Logen and Hanjin Shipping) to analyze the underlying structure of valid PTNs for each system. On the other hand, for CJ Logistics, we did not use PTNs because we found that enumeration attacks can be more effectively implemented with phone numbers on package return service – PTNs can additionally be obtained with phone numbers.

We used Naver (https://www.naver.com/), which is the most popular search engine in South Korea, to collect valid PTNs. For the initial PTNs collection, we searched web pages containing specific keywords such as "package tracking number" and then extracted strings in the format of package tracking number from the search results. As a result, we obtained the following number of initial PTNs: 1,518 for Korea Post; 770 for Lotte Logistics; 1,693 for Logen; and 1,366 for Hanjin Shipping. With those initial seed PTNs, we can analyze the valid PTN formats used for each tracking system, and they are summarized in Table 5, where all service providers' PTNs consist of digits only.

Table 5. PTN formats and maximum possible PTN spaces.

Company	PTN format	Example	Max. space
Korea Post	13-digits	1102914267781	10^{13}
CJ Logistics	10- or 12-digits	101835579911	10^{12}
Lotte Logistics	12-digits	101821471776	10^{12}
Logen	11-digits	12796430323	10^{11}
Hanjin Shipping	12-digits	304139498250	10^{12}

Even though it is not feasible to correctly guess a specific PTN having the range of 10–13 digits, enumeration attacks can be practically performed because the goal of enumeration attacks is just to identify any valid PTNs rather than to find a specific PTN. Furthermore, we found that PTNs are not randomly generated. Therefore, we can efficiently find new valid PTNs from existing PTNs. That is, given an initial PTN, we generate a candidate PTN by increasing a certain number and try to search for tracking information with the candidate PTN on the tracking service website. If the package tracking information is successfully returned from the website, the information is crawled and stored in a database; otherwise, we sequentially repeated the searching and crawling step with the next candidate PTN. In Sect. 5, from the collected data, we will show the difference between two consecutive PTNs is very small in practice.

4.2 Enumeration Attacks with Phone Numbers

For CJ Logistics, we specifically implemented a new enumeration attack, which uses phone numbers by analyzing its Android application. Specifically, we focused on designing enumeration attacks exploiting the package returning

```
POST /express.xml/delivery.do?cmd=SAF_LIST_RCV_C HTTP/1.1
Host: mobile.cjlogistics.com
...

_
      "P_PARAM": _
          "AUTH_TEL1": "010", PG_MONTH:0,
          "AUTH_TEL2": "1234",
          "AUTH_TEL3": "5678", PG_NUM:1"
"
```

Fig. 4. Input parameters used to access the customer's service usage history information at CJ Logistics, where a user's phone number (e.g., 010-1234-5678) is divided into `AUTH_TEL1`, `AUTH_TEL2`, and `AUTH_TEL3`.

service, because CJ Logistics' package delivery status checking service only displays customers' masked name instead of their full name (see Table 2).

We found that the CJ Logistics' Android application provides a login option for users' phone numbers. For example, when the login process has been successfully completed with a phone number, the user's service usage history can be accessed for 90 days. Figure 4 shows the example request message to obtain the customer's service usage history information. Therefore, if we modify the `B_PARAM` field with another valid phone number in Fig. 4, we can easily obtain the corresponding customer's entire 90 of days service usage history information.

In the service usage history information, each transaction record is composed of PTN, masked recipient's name, city, and item information. Therefore, if we have a customer's phone number, we can obtain all those information. In fact, the South Korea's phone number format has 11-digits (e.g., 010-1234-5678) as shown in Fig. 4. At first glance, the theoretically possible space of 11-digits seems sufficiently large to resist against guessing because an attacker would try 10^{11} number of guesses at the worst case. However, phone numbers are not random in practice; the first three digits (i.e., "010") of phone numbers are always the same. Furthermore, in the second part, there are some specific 4-digits that appear more frequently. For example, the 4-digits between 0000 and 1999 are reserved for the Korean government. Therefore the actual phone number space is much smaller than our expectation, making enumeration attacks feasible.

4.3 Summary of Enumeration Attack Results

In Table 6, three possible attack results are presented for each service. Specifically, "Attacked" means when the service can be executed with artificially generated input parameters (e.g., PTN, phone number, and/or name) in a short time (e.g., within a minute), and "Not Attacked" represents when we failed to find a method for enumeration attacks. "Not Applicable" indicates when the service is not provided or there is no personal information provided by the service.

Table 6. Summary of our attack analysis results for the top 5 package delivery providers in South Korea.

Company	Checking the status	Changing the drop-off location	Requesting the receipt	Returning the package
Korea Post	✓	✓	✗	✓
CJ Logistics	✓	—	—	✓
Lotte Logistics	✓	—	—	✓
Logen	✓	—	✓	✓
Hanjin Shipping	✓	—	—	—

✓ Attacked ✗ Not Attacked — Not Applicable

As explained in Sect. 4.1 and 4.2, we can successfully perform enumeration attacks on all the services, requiring PTN alone as input parameter (see Table 1).

At first glance, it does not seem straightforward to implement enumeration attacks on the other four services (the receipt requesting service at Korea Post, the package returning service at Korea Post, the package returning service at CJ Logistics, and the package returning service at Logen), because those services require some other parameters in addition to PTN.

However, for the package returning service at Korea Post, CJ Logistics, and Logen, we can still perform enumeration attacks efficiently. For the package returning service at Korea Post, two input parameters (PTN and the recipient's name) are needed (see Table 1). In this case, we first obtain the recipient's name with a PTN through the drop-off location changing service and perform enumeration attacks on the package returning service with a PTN and the recipient's name. Figure 5 illustrates this process in detail.

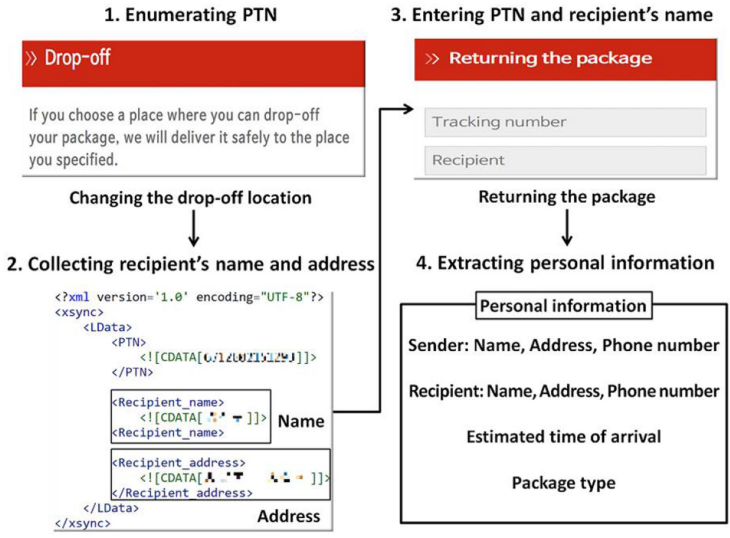

Fig. 5. Process of enumeration attacks on Korea Post: 1) we first perform an enumeration attack with PTNs on the drop-off location changing service; 2) we extract a recipient's name from the search result if a valid PTN is entered; 3) we then execute the package returning service with the obtained PTN and name; and 4) we extract the target personal information from the search result of the package returning service.

Similarly, for the package returning service at Logen, we first obtain the recipient's phone number with a PTN through the receipt requesting service (see Table 3) and perform enumeration attacks on the package returning service with the PTN and the recipient's phone number. For the package returning service at CJ Logistics, we can use phone numbers instead of PTNs for enumeration attacks. As explained in Sect. 4.2, we can obtain PTNs with phone numbers by modifying the parameters to access the user's service usage history at its Android application. In summary, we only failed to perform enumeration attacks

on the receipt requesting service at Korea Post because this service requires an internally generated authentication code to access (see Table 1).

5 Analysis of Collected Data

During 6 months (from May 2019 to November 2019), we collected package delivery records as follows: 1,398,112 for Korea Post; 2,614,839 for CJ Logistics; 797,676 for Lotte Logistics; 1,590,933 for Logen; and 163,452 for Hanjin Shipping. We collected at least 700,000 package delivery records from all providers except Hanjin Shipping. Hanjin Shipping blocked the IP addresses used for our experiments. We surmise that Hanjin shipping only used a proper security solution to block the IP addresses used for generating a large volume of suspicious queries within a short time interval.

From the collected package delivery records, we count each type of customers' personal data categorized by sender's and recipient's name, phone number, and address, after removing duplicated customer data. The results are presented in Table 7.

Table 7. Numbers of customers' personal data categorized by name, phone number and address.

Company	Sender			Recipient		
	Name	Phone number	Address	Name	Phone number	Address
Korea Post	19,712	1,207	18,012	1,220,350	822,159	1,212,029
CJ Logistics	128,426	52,329	56,062	1,811,325	724,824	1,734,623
Lotte Logistics	7,268	1,844	5,004	–	–	–
Logen	204,405	92,755	140,847	1,388,539	980,222	1,366,470
Hanjin Shipping	–	–	–	–	–	154,207
Total	359,811	148,135	219,925	4,420,214	2,527,205	4,467,329

Personal Information. As shown in Table 7, the number of senders' data is less than the number of recipients' data because a vast majority of senders are professional sellers or companies, while most recipients are normal customers. Therefore, we note that recipients' personal information appears more attractive to attackers than senders' information. For recipients' data, we collected 4,420,214 names, 2,527,205 phone numbers, 4,467,329 addresses, respectively, in total. Perhaps, such people's personal information could be abused to conduct additional cyber criminal activities such as sophisticated spam/phishing attacks [8,13] and Sybil accounts creation [11], and invade user privacy. For example, we found that some military officers used their military rank as a part of their name (e.g., Captain John Doe). In this situation, their private home address or the location of a military base can be exposed to the public including potential attackers. Furthermore, a celebrity's phone number and/or home address can be potentially revealed by linking his/her publicly known other information (e.g., real name,

location of home address). Previous studies demonstrated that the inclusion of more detailed contextual information would increase the success probability of phishing attacks [9].

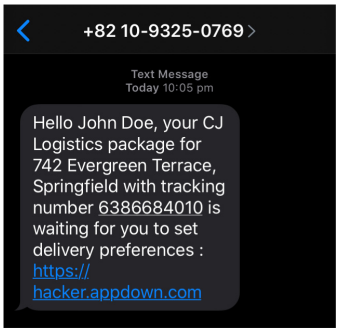

Fig. 6. Example of targeted SMiShing attacks.

Figure 6 shows a targeted SMiShing attack, which is a type of phishing communication that is sent to a victim's mobile phone through an SMS message. In this example, the personal information (name, PTN, and address) about a victim (John Doe) is added to deceive the victim into believing that this SMS is sent from the original shipping company (CJ Logistics) in order to entice the victim to click the link to the attacker's website.

Predictability of PTN. Furthermore, we examine patterns in a sequence of PTNs to predict PTNs. We specifically measure the difference between two consecutive PTNs ($\Delta PTN(i) = PTN(i+1) - PTN(i)$) where $PTN(i)$ is the ith PTN in the sequence of PTNs. We calculate the cumulative distribution functions (CDF) of $\Delta PTN(i)$ (from 1 to 20) with all the collected PTNs from Korea Post, CJ Logistics, Lotte Logistics, Logen, and Hanjin Shipping tracking systems, respectively.

Figure 7 shows the calculated CDFs for Korea Post, CJ Logistics, Lotte Logistics, Logen, and Hanjin Shipping, respectively, where the X-axis represents the difference between two successive PTNs[1], and the Y-axis represents the cumulative percentage of the number of PTNs with less than or equal to ΔPTN. We can see that in most cases, the gaps between two successive PTNs are smaller than 20, indicating that PTNs can be efficiently enumerated in a sequential manner.

6 Possible Defense Mechanisms

In this section, we suggest three possible defense mechanisms to mitigate the security threats.

[1] We denote $\Delta PTN(i)$ for all i in the collected PTNs as ΔPTN.

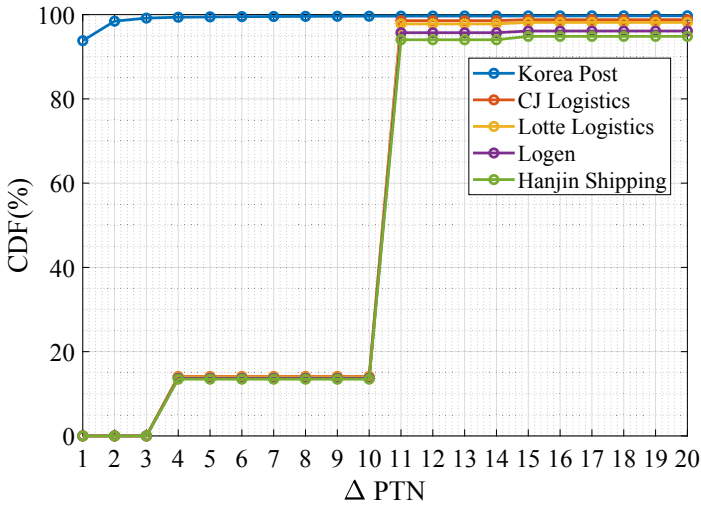

Fig. 7. Cumulative distribution function (CDF) of ΔPTN for Korea Post, CJ Logistics, Lotte Logistics, Logen, and Hanjin Shipping.

Limiting the Number of PTN (or Phone Number) Verification Failed Attempts: Limiting the number of PTN (or phone number) verification failed attempts can be the first line of defense to prevent enumeration attacks because a large number of PTN (or phone number) verification failed attempts are necessarily induced while performing enumeration attacks. In practice, this approach can be implemented by simply counting the number of verification failed attempts from a specific client. Hence, we can apply this policy in package tracking systems with a low deployment cost. The idea of limiting the number of attempts from a specific client (e.g., with an IP address) or imposing a minimum time interval between failed attempts is not new [6]. However, none of the package tracking systems that we tested to limit the number of failed attempts, and seem to be considering enumeration attacks[2]. If we deploy the policy of "maximum failed attempts allowed," attackers would try to change their strategy into more complicated enumeration attack scenarios (e.g., [11]) with multiple hosts and diverse query patterns, leading to the increase in attackers' efforts.

Using CAPTCHA Challenges: Another promising approach is to use CAPTCHA [16] challenges to hinder automated attempts which are necessary to perform enumeration attacks. However, the use of CAPTCHA challenges can incur the usability cost of taking the time to solve CAPTCHA challenges for

[2] We believe that Hanjin Shipping would use a DDoS mitigation solution at the network level rather than the policy of "maximum failed attempts allowed" at the web application level because we cannot access the website itself when we queried multiple times within a short time interval.

normal users. Therefore, it seems better to combine this approach with our first recommendation – we can ask users to solve a CAPTCHA problem only when the number of PTN (or phone number) verification failed attempts is greater than the maximum number of failed attempts allowed (e.g., five) or suspicious query patterns are detected.

Generating Unpredictable PTNs: The problem with PTNs is that they are highly predictable and can easily be enumerated[3] (see Fig. 7). Therefore, we need to change the existing structure of PTNs by reserving at least some reasonable number of digits (e.g., 6 digits) in PTNs to represent a random number, which makes PTNs harder to enumerate within a reasonable time. A cryptographically secure pseudo-random number generator (CSPRNG) such as Fortuna [12] can be used to generate random digits in an unpredictable manner.

Minimizing Information Leakage: Current package tracking systems provide unnecessary personal information about sender and recipient in their online website, which can be viewed and harvested by a third party. To address this problem, we suggest that package tracking systems should not provide any personal information (e.g., name, phone number, address, etc.) about the sender (or recipient) with a PTN alone. That is if a user enters a PTN or his/her phone number, package tracking websites can show the only information about package status such as current package location and estimated delivery date, but no user-related personal data. For some situations where senders' or recipients' personal data is needed (e.g., some recipients may want to contact their senders), however, sender's (or recipient's) personal data can be additionally provided only when the recipient (or sender) successfully logs-in into the package tracking website.

7 Ethical Considerations

The main motivation of our experiments is to show the risk of potential enumeration attacks on package delivery service and discuss effective defense mechanisms to mitigate such attacks. Therefore, we only checked service providers' responses for our enumeration attack attempts; however, actual user data were not stored. Furthermore, we queried the websites' tracking services at a very slow rate to minimize adverse impacts on the websites' normal operations. Finally, we reported the discovered design flaws and our recommendations to shipping companies running those services.

8 Related Work

In recent years, the possibility of enumeration attacks has been intensively studied in social network and instant messenger services.

[3] We surmise that PTNs may contain some meaningful information (e.g., location and time) about package delivery records because they have a well-formatted structure.

Balduzzi et al. [7] discussed the possibility of enumeration attacks to automatically harvest active email addresses by using Facebook's friend-finder feature. About 10.4 million e-mail addresses were tested and more than 1.2 million user profiles were found to be associated with these addresses. To fix this problem, Facebook employed several defense mechanisms such as detecting suspicious query patterns and using CAPTCHA challenges. More recently, however, Kim et al. [11] showed that an advanced enumeration attack scenario with a few Sybil accounts can evade those defense mechanisms in the real-world situations. Similar problems related to enumeration attacks were also reported in instant messenger services. Schrittwieser et al. [14] presented an enumeration attack to collect 21,095 live phone numbers from WhatsApp within less than 2.5 h. Kim et al. [10] also collected 50,567 users' phone numbers, names, and profile pictures from KakaoTalk (https://www.kakaocorp.com/service/KakaoTalk?lang=en) through enumeration attacks. Gupta et al. [8] demonstrated that the collected phone numbers could potentially be abused to perform sophisticated targeted phishing attacks or a larger phishing campaign. Recently, Woo et al. [15] demonstrated the possibility of *enumeration attacks* with the top three package service providers (FedEx [4], DHL [3], and UPS [5]). However, their work was focused on those three service providers only and did not explain how enumeration attacks can be systematically detected and tested. We extend their study to additional services for generalization, and further develop a framework to systemically analyze the attack vectors related to enumeration attacks on websites.

Many customers are already concerned about shipping companies that have maintained customers' personal data insecurely. For example, personal information of thousands of FedEx customers was exposed [1] because of the insecure cloud storage server. Also, USPS exposed 60 Million user information due to the flaws in its APIs [2].

9 Conclusion

In this work, we examined the possibility of enumeration attacks on existing package tracking systems. We developed effective enumeration attack scenarios for the websites of top shipping companies (Korea Post, CJ Logistics, Lotte Logistics, Logen and Hanjin Shipping) in South Korea. Our experimental results demonstrate that those companies do not fully consider a reasonable level of security practices to protect their customer data. We collected a large number of package delivery records from those companies' websites and finally extracted 4,420,214 names, 2,527,205 phone numbers, and 4,467,329 addresses in total through our enumeration attack implementations in an automated manner. To address this security concern, we suggest four practical defense approaches such as limiting the number of PTN verification fail attempts, using CAPTCHA challenges and generating unpredictable PTNs to prevent enumeration attacks.

Although our analysis and observation are package tracking system-specific, they could offer valuable lessons for other websites that provide services with tracking numbers alone. As part of future work, we plan to implement a generic tool for testing the possibility of enumeration attacks on websites.

Acknowledgement. This work was supported in part by the NRF of Korea (NRF-2019R1C1C1007118), the ITRC Support Program (IITP-2019- 2015-0-00403), and the ICT R&D Programs (No. 2017-0-00545).

References

1. FedEx Data Breach (2018). https://www.informationsecuritybuzz.com/expert-comments/fedex-data-breach/. Accessed 14 Oct 2019
2. USPS Site Exposed Data on 60 Million Users (2018). https://krebsonsecurity.com/2018/11/usps-site-exposed-data-on-60-million-users/. Accessed 14 Oct 2019
3. DHL global (2019). http://www.dhl.com/en.html. Accessed 14 Oct 2019
4. Fedex (2019). https://www.fedex.com. Accessed 14 Oct 2019
5. UPS (2019). https://www.ups.com. Accessed 14 Oct 2019
6. Alsaleh, M., Mannan, M., van Oorschot, P.C.: Revisiting defenses against large-scale online password guessing attacks. IEEE Trans. Dependable Secure Comput. **9**, 128–141 (2012)
7. Balduzzi, M., Platzer, C., Holz, T., Kirda, E., Balzarotti, D., Kruegel, C.: Abusing social networks for automated user profiling. In: Jha, S., Sommer, R., Kreibich, C. (eds.) RAID 2010. LNCS, vol. 6307, pp. 422–441. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15512-3_22
8. Gupta, S., Gupta, P., Ahamad, M., Kumaraguru, P.: Exploiting phone numbers and cross-application features in targeted mobile attacks. In: Proceedings of the 6th Workshop on Security and Privacy in Smartphones and Mobile Devices (2016)
9. Hong, J.: The state of phishing attacks. Commun. ACM **55**(1), 74–81 (2012)
10. Kim, E., Park, K., Kim, H., Song, J.: Design and analysis of enumeration attacks on finding friends with phone numbers: a case study with kakaotalk. Comput. Secur. **52**, 267–275 (2015)
11. Kim, J., Kim, K., Cho, J., Kim, H., Schrittwieser, S.: Hello, Facebook! here is the stalkers' paradise!: design and analysis of enumeration attack using phone numbers on Facebook. In: Proceedings of the 13th International Conference on Information Security Practice and Experience (2017)
12. McEvoy, R., Curran, J., Cotter, P., Murphy, C.: Fortuna: cryptographically secure pseudo-random number generation in software and hardware (2006)
13. Palmer, D.: Phishing attack: students' personal information stolen in university data breach (2019). https://www.zdnet.com/article/phishing-attack-students-personal-information-stolen-in-university-data-breach/. Accessed 30 Dec 2019
14. Schrittwieser, S., et al.: Guess who's texting you? Evaluating the security of smartphone messaging applications. In: Proceedings of the 19th Annual Symposium on Network and Distributed System Security (2012)
15. Woo, S., Jang, H., Ji, W., Kim, H.: I've got your packages: harvesting customers' delivery order information using package tracking number enumeration attacks. In: Proceedings of The Web Conference (WWW 2020) (2020)
16. von Ahn, L., Blum, M., Hopper, N.J., Langford, J.: CAPTCHA: using hard AI problems for security. In: Biham, E. (ed.) EUROCRYPT 2003. LNCS, vol. 2656, pp. 294–311. Springer, Heidelberg (2003). https://doi.org/10.1007/3-540-39200-9_18