

# Homework 6

## COM402 - Information Security and Privacy 2018

- The homework is due **Sunday, June 3, 2018 at 23h55 on Moodle**. Submission instructions are on Moodle. Submissions sent after the deadline **WILL NOT** be graded.
- In the event that you find vulnerabilities, you are welcome to disclose them to us (can even have a bonus !)
- For Exercises 2 and 4, do not forget to submit your source files to Moodle along with your tokens.

### Exercise 1: NOP sled galore!

It's your good friend NOP sled from the buffer overflow exercise! You don't need to do anything for the first part of this homework. Isn't that awesome? Go ahead and skip to the exercise 2!

### Exercise 2: Blockchain

Everybody has a blockchain - so why not you too? We set up a simple blockchain at `com402.epfl.ch` that waits for new blocks. You don't have to include any broadcast protocol like in Bitcoin, but we ask you to perform some mining and send the block to the com402 server.

#### Installation

To interact with our cothority where the blockchain is handled, you need to install some packages using `pip`. **Please note that only versions  $\geq 3.5$  of Python work!** If you get

```
async def getStatus(url):  
    ^  
SyntaxError: invalid syntax
```

during the installation, this means you're running a version of Python that is too old. In this case, install the packages using:

```
pip3 install cothority~=0.1.0 websockets protobuf
```

and you're ready to go.

## Communicating with the blockchain

To communicate with the blockchain, you have two methods:

```
# returns a list of all blocks that are stored in the skipchain
# identified by the genesis string.
# Input:
# - url where to contact the skipchain
# - genesis-block (in binary format)
# Output:
# - all blocks since the genesis-block
# Error: websockets.exceptions.ConnectionClosed indicates that
# something went wrong. The error message has more information.
cothority.Cothority.getBlocks(url, genesis)
```

```
# Asks the skipchain to add a block
# Input:
# - url where to contact the skipchain
# - block to store on the skipchain
# Output:
# - previous block of the blockchain
# - latest block of the blockchain which should be yours
# Error: websockets.exceptions.ConnectionClosed indicates that
# something went wrong. The error message has more information.
cothority.Cothority.storeBlock(url, block)
```

For creating a block, you can use the convenience method:

```
# returns a new block that appends to the last block
# Input:
# - last valid block from the blockchain
# - data that will be stored in the new block.
cothority.Cothority.createNextBlock(last, data)
```

The URL for the blockchain is

```
ws://com402.epfl.ch:7003
```

And the genesis-id (which corresponds to the skipchain-id):

```
0d8d8f2012b4905b90dd179c060bf56b36fffed4b285067898a642347a9c7621
```

## Mining

To mine a block, you need to create binary data that conforms to some SHA256 value. The data format is:

```
nonce = data[0:32]
```

```
hash_of_last_block = data[32:64]
full_email = data[64:]
```

You find the hash of the last block in the return value of `getBlocks`. Only the last block is important. To create a new block, you can use the `createNextBlock` method and give it your data.

**For the data to be accepted by the server, the first three bytes of the SHA256 hash of the data has to be 0!** This is what the nonce is for. You either use it as a counter or take random data (both should give the same result) and iterate, creating new SHA256 hashes, **until you find one hash that gives 0x000000 in the beginning**. With our simple sample-script it takes about 4 minutes, so don't worry if you don't immediately find a valid hash! For testing reasons you can start with 0x0000 to see if your method works. Even though it will be rejected by the blockchain...

If you find a better way than to try out different combinations, be sure to write a paper about it ;)

## Adding the block

Once you found a correct data for your block, create a block with `createNextBlock` and submit it to the blockchain with `storeBlock`. Be sure to check the return message. In case the block is not accepted, you will get an error message.

You can also check on <http://com402.epfl.ch/static/conode.html> for error messages from the nodes.

## Finished

You need to add at least 3 blocks to the blockchain to get your token from the website.

# Exercise 3: Network Security

## Part 1: Analyzing network traces

### Background

You will [find](#) four packet traces (pcap files) that can be read by the Wireshark tool (among other tools). You will need to investigate these traces to answer the questions below. To get started you will want to understand how to use Wireshark's filtering capabilities. Your solution will be a file `solutions.txt` with answers to the questions below.

### Trace 1: HTTP

1. Give three websites (domain name and IP addresses) visited from source IP address `192.168.0.100` (**this does not include the URLs related to online ads, showing images, internet telemetry, API calls, etc.**).
2. Give three search queries and the domain of the site for each query made from source IP address `192.168.0.100`.

### Trace 2: FTP

FTP is the file transport protocol. There is a lot of information about it on the internet.

1. What is the user name and password used to connect to the FTP server?
2. List all files that were downloaded from the FTP server.
3. List the full path for two files (in different directories) on the FTP server that were **NOT** downloaded.

### Trace 3: Traceroute

traceroute is a tool used to determine the route between two IP addresses. You can find information about it on the internet. Wikipedia has a nice introduction:

<https://en.wikipedia.org/wiki/Traceroute>.

1. Briefly describe how the traceroute tool works including which network protocols are in use.
2. Give the source IP address that issued the traceroute command and the destination IP address.
3. List the IP addresses on the route **between source and destination (meaning: do not include the IP addresses of the source and destination)**.

### Trace 4: POP

The post-office protocol (POP) is used for email.

1. What is the POP username and password?
2. How many emails are in the user's mailbox?
3. Give the contents of from, to, subject, and date for **one** email message.
4. What email client (application) and operating system is this person using to send and receive email?

## Deliverables

Along with the pcap files, we provide you with a skeleton of solutions.txt, which contains the questions above. You will put your answers in this file, and submit it to Moodle. Make sure that you include your full name and SCIPER at the beginning of solutions.txt by replacing [Student name] and [SCIPER].

**IMPORTANT:** You will notice that solutions.txt contains many lines of the form [some text here]. Please replace these pieces of text with your answer. **DO NOT REMOVE ANY TEXT THAT IS NOT ENCLOSED WITH [ and ]!** Notice that there are different formats for different answers. For instance, your answers for the 2nd and 3rd questions of *Trace 2: FTP* should be a single line that is comma separated.

## Part 2: Anomaly detection

Now, you will programmatically analyze a pcap file to detect suspicious behavior. Specifically, you will be attempting to identify port scanning. You can download the pcap file [here](#).

Port scanning is a technique used to find network hosts that have services listening on one or more target ports. It can be used offensively to locate vulnerable systems in preparation for an attack, or defensively for research or network administration. In one kind of port scan technique, known as a SYN scan, the scanner sends TCP SYN packets (the first packet in the TCP handshake) and watches for hosts that respond with SYN+ACK packets (the second handshake step). Since most hosts are not prepared to receive connections on any given port, typically, during a port scan, a much smaller number of hosts will respond with SYN+ACK packets than originally received SYN packets. By observing this effect in a packet trace, you can identify source addresses that may be attempting a port scan. Your task is to develop a Python program that analyzes a pcap file in order to detect possible SYN scans.

To do this, you will use `dpkt`, a library for packet manipulation and dissection. It is available in most package repositories. You can find more information about `dpkt` [here](#), and view documentation by running `pydoc3 dpkt`, `pydoc3 dpkt.ip`, etc.(or `pydoc`). There's also a helpful tutorial [here](#). Or [here](#).

**Your program will take the absolute path of the pcap file (it has to be named `com402_hw6_ex3_p2.pcap`) to be analyzed as a command-line parameter. So if the pcap file is located at `/home/ceyhun/last_hw`:**

```
python3 detector.py /home/ceyhun/last_hw/com402_hw6_ex3_p2.pcap
```

The output should be the set of IP addresses (one per line) that (1) sent more than 3 times as many SYN packets as the number of SYN+ACK packets they received and (2) sent out at least one SYN packet but has not received any SYN+ACK packets. Your program should silently ignore packets that are malformed or that are not using Ethernet, IP, and TCP.

### Deliverables and getting the token

You can get your token by submitting the file with the IP addresses. Each line should only have a single IP address in the following format `xxx.xxx.xxx.xxx`. You are also going to submit your Python program (on Moodle when you turn in the homework) that accomplishes the task specified above, as a file named `detector.py`. It would be ideal if you work with `dpkt 1.9.1`, as this is the version we have in our servers.

## Exercise 4: PIR with a twist

In this exercise you will implement a very simplified version of the Riffle protocol (<https://people.csail.mit.edu/devadas/pubs/riffle.pdf>). Riffle is a communication system with

strong anonymity, for all the detail please read the paper. The emphasis in the exercise will be on the PIR (Private Information Retrieval) phase of the protocol. The original protocol consists of 2 phases: the setup phase where clients and servers exchange cryptographic keys and the communication phase where clients upload/download messages from the servers. In our simplified case, we will have the following steps in the protocol:

- **Setup phase:**
  - Servers generate their public/private key pairs and start listening for UDP packets on a given ip and port
  - Client, let's call it *Com402Client*, first generates a shared key (AES key) for each of the servers
  - For each server, client encrypts the shared key with server's public key and send it to the server
  - So, at the end of the phase we will have:
    - Let's say  $m$  servers running, each of them having a shared key which is shared with *Com402Client*
    - *Com402Client* which has all the shared keys  $sk1, sk2, \dots, skm$
- **Communication phase - Upload:**
  - *Com402Client* client generates  $n$  chat messages and onion-encrypts them with the shared keys. Onion encryption of a message  $msg$  means the following:  

$$cipher = AES\_sk1(AES\_sk2(\dots(AES\_skm(msg))\dots))$$
  - *Com402Client* then sends  $n$  onion-encrypted messages to the primary server
  - Servers then onion-decrypt all the messages:
    - Primary server removes one layer of encryption - decrypts the ciphertext with his shared key, and send the result to the next server in chain
    - Next server does the same, until eventually the last server in the chain obtains the plaintext message.
    - The last server then broadcasts the plaintext msg and all the servers store it.
  - Order of onion-encryption/decryption is determined by the servers ids. Each server has an assigned integer  $id$ , and the chain is defined by the increasing order of ids (server0 -> server1 ->...).
  - So, at the end of this phase all the servers have  $n$  plaintext chat messages. All the messages have an index, so the servers have them in the same order as well.
- **Communication phase - Download:**
  - In this phase a client, let's call *StudentClient*, want to download a message with a specific index, but doesn't want servers to learn which index is that. So, *StudentClient* performs a simple PIR.
  - *StudentClient* generates  $m$  bitmasks(one for each server) of length  $n$  (total number of messages), in such way that when XORed, the result will be the target index (index of a message which *StudentClient* wants to retrieve).

- *StudentClient* sends bitmasks to the servers and asks each server to select messages according to the bitmask, XOR the selected messages and send the result back
- *StudentClient* can recover the desired message by combining all the responses from the servers.
- If this is not clear to you, please read the PIR phase in the paper!

We provide you with two skeleton scripts, one for the servers and one for the clients. You can download them on the com402 website. In the scripts you will find detailed instructions on what and how you should implement. We will test and grade two things:

1. Your server code: once you finish your implementation upload your script to com402
2. Your client code: we will grade only the client which performs the PIR

**Update: When submitting your solutions, please submit your server and client separately (i.e. submit your server and get a token for it, then submit your client and get a second token for it)**

You are provided with enough code to make the whole system work locally on your machine, so please test thoroughly before submitting the solution.

Note: you will have to install the pycrypto package by:  
pip3 install pycrypto

Good luck!