Lesson for lab 4: Scripts, utility tools, and trace analysis

Rahul Hiran

TDTS11:Computer Networks and Internet Protocols

Objective

- Lab introduction
- Collecting data
- Introduction to tools
  - grep, sed, sort, awk
- Format of data to be used with tools
- Example task

About Lab 4

- Problem-based learning
  - Number of tasks to solve
  - Apply your knowledge, techniques to get solutions
- Data collection
  - Collect your own trace
  - Borrow a trace from another student in TDTS11

Collecting your own trace

- Steps
  - Close all applications, empty local cache in browser
  - Start wireshark and browser
  - Get the list of 25 most popular websites from alexa.com
  - Every new minute, go to the i’th most popular site
  - Stop capturing with wireshark

Demo: Collected trace file

- Demo of file
- File in binary format → export to text format from wiresha
- Show the file
- Issues working with file

grep- demo

- Pattern matching/ search text
- Show demo

sed demo

- To transform text
- Typical format:
  - Sed ’s/regexp/replacement/g’
- Demo

sort- demo

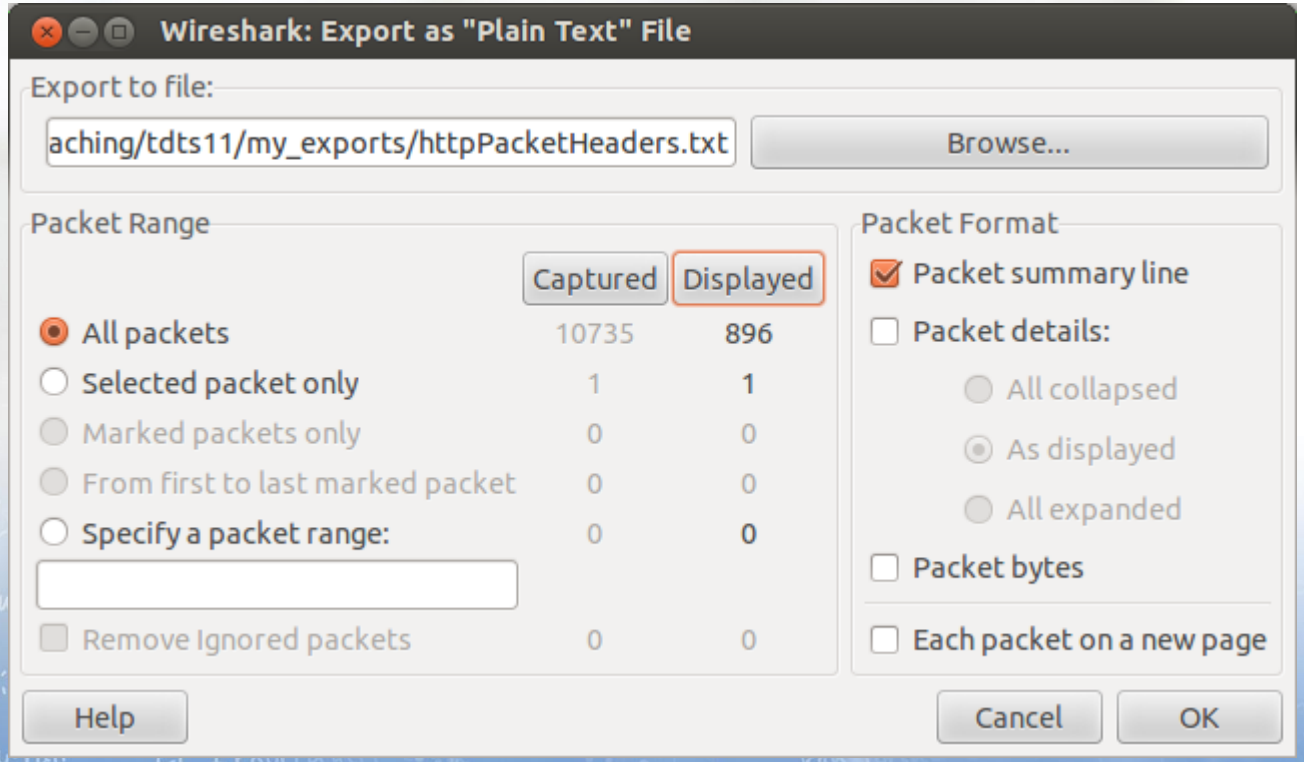
- Sort the output
- Demo
  - Necessary when you want to find the count of unique entries

awk- demo

- Programming language
- Text processing tool
- Typically used for data extraction
- Demo

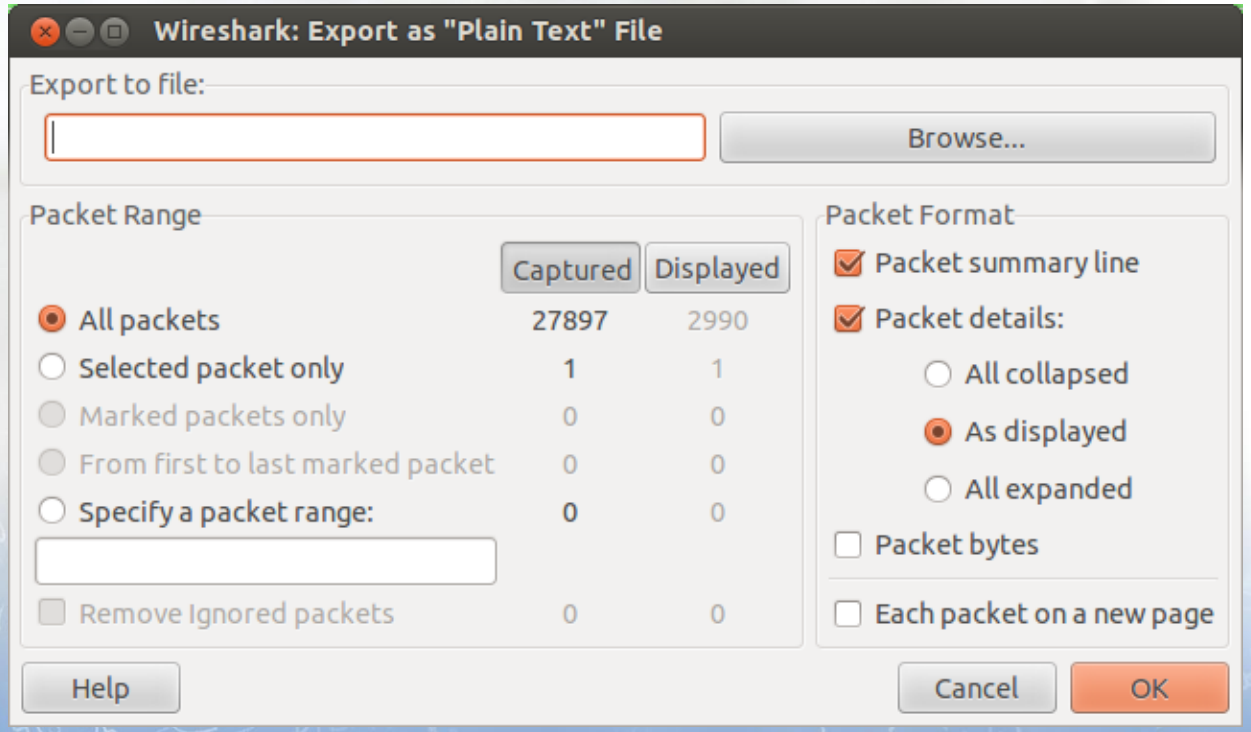
Export data-1

File → export → as “plain text” file



Export data-2

File → export → as “plain text” file



General flow to solve the tasks

- Understand the question
- Figure out which field(s) and protocol(s) to look for
- Export data for that field and protocol in a text file (by ign
- Use the commands and tools to get the solution

Example task

- Task:
  - For entire trace find how many bytes were sent/received?
- Which field, protocol?
  - Show the wireshark screen→ no field in http→ but tcp has ”Len
- Export required data
  - export data with tcp details
- Use commands/tools to get the answer
  - Show the command to get total number of bytes transferred

Example task: commands to use

- Use commands such as cut, awk, and grep to derive the
- Two alternatives
  - With the cut and awk commands
  - Only using the awk command
- There could be more ways to get the answer
  - Choose commands/tools carefully
  - Explore and find the tools that you find most useful

Questions…?

Linköping University

expanding reality

www.liu.se

Examples-1

1. Get unique domains that were contacted:

```
cat httpDetails.txt | grep Host | head -504 | tail -314| grep -v ubuntu | sort | uniq | sed 's/\r\n//g'| awk -F ':' '{print $2}' | sort | uniq | wc
```

2. Get sum of all the data that was transferred and received

```
cat tcpDetails.txt | grep Transmission Control Protocol | awk -F ',' '{print $6 }' | awk -F ':' '{print $2}{sum+=$2}END{ print Final sum is:sum}'
```

OR (with cut command)

```
cat tcpDetails.txt | grep Transmission Control Protocol | cut -d ',' -f6 | awk -F ':' '{print sum+=$2}'
```

Example-2

3. Get unique urls from all the connections:

```
cat httpDetails.txt  | grep '     GET' | tr -s ' ' | cut -
d ' ' -f 11 | more
```

or

```
cat httpDetails.txt  | grep '     GET' | awk '{print $10}'
| more
```

4. About tr command:

```
echo "ttttt ttttt" | tr 't' 'a'
echo "ttttt ttttt" | tr -s 't' 'a'
```