CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

Tesina TIR

Analisi di traffico reale su una rete domestica

ver. 1.0

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

SOMMARIO

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

INTRODUZIONE

Lo scopo del presente elaborato è di raccogliere ed in seguito analizzare i dati di una rete domestica per rispondere al seguente quesito della tesina:

Analisi di traffico reale in una rete domestica

Ai candidati viene proposto di raccogliere ed in seguito analizzare i dati di una rete domestica, proponendo come tesina finale una analisi dei dati ottenuti sotto diversi punti di vista, riferiti ai protocolli di trasporto (es. percentuale dati UDP e TCP, durata connessioni TCP, valori RTT e congestion window stimati, ecc.)

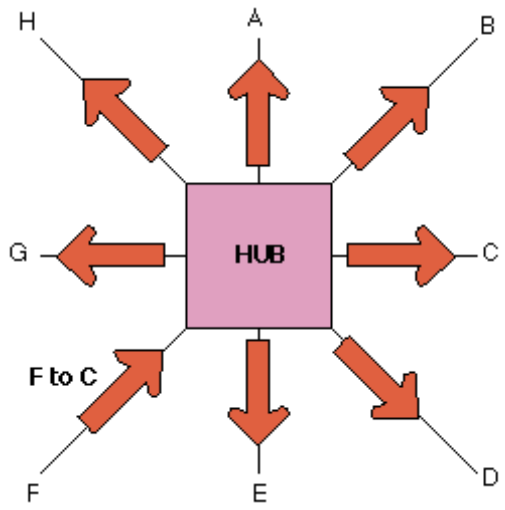
Workflow proposto:

1. Raccolta dati in formato PCAP di una rete domestica secondo diversi orario di utilizzo
2. Valutazione di software di analisi esistenti, o preparazione software di analisi che fornisca i dati richiesti sul dataset
3. Presentazione dei risultati tramite tabelle e analisi degli stessi in un elaborato

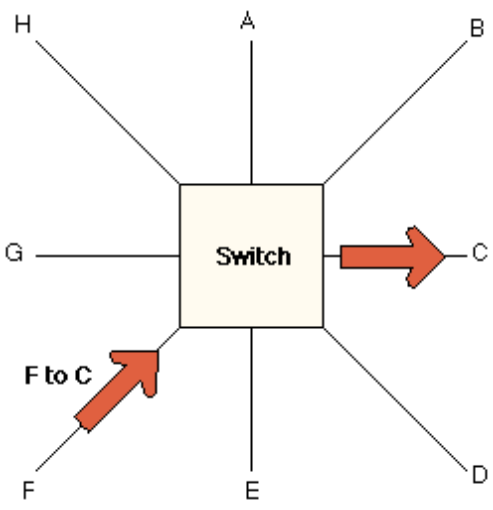
L'importanza dell'access point per l'esecuzione dell'esperimento

La comunicazione fra due dispositivi appartenenti alla stessa rete locale, quale ad es. una rete domestica, avviene mediante un dispositivo di rete che funge da nodo di smistamento; tale dispositivo è posizionato nel centro stella ed è dotato di più porte: i dati in arrivo da una qualsiasi porta vengono inoltrati su una o su tutte le altre porte a seconda che il dispositivo sia uno switch o un hub.

La differenza sostanziale fra hub e switch è che il primo, lavorando a livello 1 (fisico) del modello ISO/OSI, replica incondizionatamente il traffico proveniente da una qualsiasi porta su tutte le altre (fig. 1a), con ovvi problemi di congestione della banda disponibile ed il conseguente aumento delle collisioni, mentre il secondo, lavorando a livello 2 (data link) del modello ISO/OSI, inoltra i dati mediante una corrispondenza MAC address destinazione – porta (fig. 1b), ottimizzando l'uso della banda.



(fig. 1a - hub)



(fig. 1b - switch)

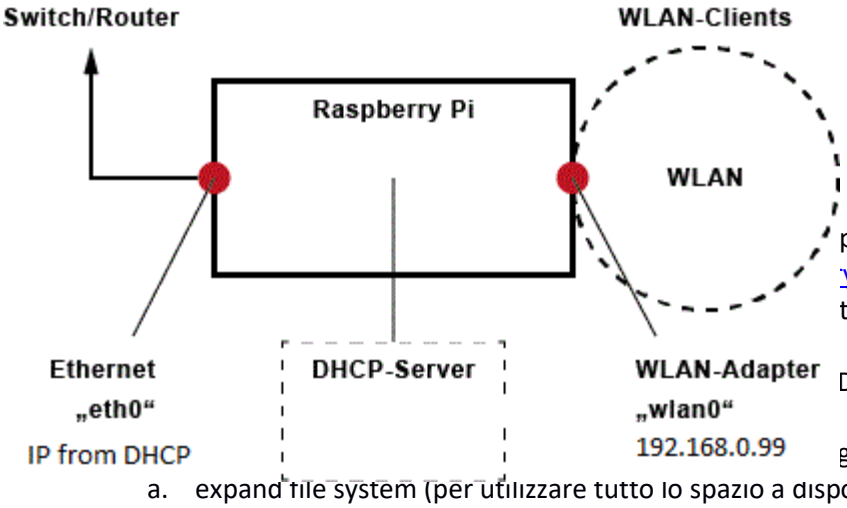
L'utilizzo degli switch è subentrato anche nelle reti domestiche e se da un lato migliora le prestazioni generali dall'altro aggiunge alcune complicazioni qualora si voglia raccogliere il traffico dell'intera rete LAN in quanto il dispositivo attaccante, in virtù del principio di funzionamento dello switch, potrà apprezzare esclusivamente il proprio traffico.

Essendo uno dei nostri scopi quello di acquisire l'intero traffico di una rete LAN domestica possiamo mettere in campo diverse metodologie, alcune lecite ed altre illecite, quali ad esempio lo sniffing del traffico mediante un attacco "man in the middle", la sostituzione fraudolenta dell'access point, oppure l'istituzione di un nuovo access point e chiedere la partecipazione allo studio.

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

PREDISPOSIZIONE APPARATO DI TEST

Concordando sull'utilizzo di una metodologia lecita e sul coinvolgimento volontario e consapevole degli utenti all'esperimento, si è scelto di utilizzare un dispositivo che consentisse di attivare una nuova rete wifi e che contemporaneamente alla funzionalità di bridge fra rete wireless e rete cablata consentisse l'acquisizione del traffico. Inoltre, essendo un'attività da svolgere in luoghi diversi si è considerato l'aspetto di utilizzare un dispositivo ed una configurazione che fosse facilmente portabile e facilmente installabile in diversi ambienti e che principalmente non richiedesse personalizzazioni a seconda del luogo.
Per queste ragioni abbiamo scelto di utilizzare un dispositivo Raspberry pi model B, che date le dimensioni ridotte, la presenza di una porta ethernet e diverse porte usb e l'utilizzo mediante sistema operativo linux, si presta bene allo scopo del presente elaborato.
L'idea è di configurare tale dispositivo per connettersi alla rete lan per mezzo di un comune cavo rj45, ottenendo un ip dinamico dall'eventuale DHCP presente nella LAN, di creare una rete wireless per mezzo di un adattatore wifi e di inoltrare tutto il traffico wifi verso la porta ethernet, potendo così intercettare e memorizzare il traffico sull'interfaccia wireless (fig. 2).



passi:
ypi.org/downloads/raspbian);
tare su ON l'opzione "Format Size
DiskImager;
ggi, tra cui:

a. expand file system (per utilizzare tutto lo spazio a disposizione della scheda SD);
b. change user password (utente pi) -> "TesinaTIR2016";
c. boot options -> selezionare "Console" (per avviare il dispositivo senza GUI);
d. Internationalization Options -> Change Locale -> It.UTF8;
e. Internationalization Options -> TimeZone -> Europe -> Rome;
f. Internationalization Options -> Keyboard -> Generic 105 -> Other -> Italian -> Italian WinKeys -> Right Alt (AltGr) -> Right Alt (AltGr) ;
g. Internationalization Options -> WiFi Country -> IT;
h. Advanced -> HostName -> piTIR;
i. Advanced -> SSH -> Yes (per attivare SSH);

6) Riavviare ed accedere con l'utente "pi"
7) Modificare /etc/apt/sources.list, decommentando deb-src per accedere ai repository dei codici sorgenti;
8) Sudo passwd per modificare la password dell'utente root -> "TesinaTIR2016";
9) Sudo apt-get update e poi Sudo apt-get upgrade per aggiornare il sistema

Setup Raspberry – AP

Seguire le seguenti istruzioni per configurare il raspberry come access point:

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

1) Verificare che la scheda di rete wireless si stata riconosciuta dal sistema ed impostare un ip statico per l'interfaccia wlan0:

```
ifconfig
nano /etc/network/interfaces

auto wlan0
iface wlan0 inet static
address 192.168.0.99
netmask 255.255.255.0
```

2) Verificare se la scheda preveda la modalità AP (deve esserci AP in "Supported Interface Mode")

```
iw list
```

3) Installare e configurare modulo per AP. Potrebbe essere necessario utilizzare una versione di hostapd compilata appositamente per la scheda di rete wireless in uso.

```
apt-get install hostapd
nano /etc/hostapd/hostapd.conf

interface=wlan0
driver=rtl871xdrv
ssid=pi_wifi
hw_mode=g
channel=8
macaddr_acl=0
auth_algs=1
wmm_enabled=0
ignore_broadcast_ssid=0
wpa=2
wpa_passphrase=prova2016
wpa_key_mgmt=WPA-PSK
rsn_pairwise=CCMP

nano /etc/dafault/hostapd
specificare "/etc/hostapd/hostapd.conf" in DAEMON_CONF
```

Setup Raspberry – DHCP

Seguire le seguenti istruzioni per attivare il DHCP sull'interfaccia wlan0:

1) Installare e configurare DHCP:

```
apt-get install udhcpd
nano /etc/udhcpd.conf

impostare start-end range 192.168.0.20 -> 50
```

```
interface wlan0
remaining yes
opt dns 8.8.8.8 4.4.4.4
opt router 192.168.0.99

nano /etc/default/udhcpd
remmare #DHCPD_ENABLED="no"
```

Setup Raspberry – IP forward

Seguire le seguenti istruzioni per configurare il forward wlan0/eth0:

1) Attivazione IP forward e creazione regole iptables

```
nano /proc/sys/net/ipv4/ip_forward
indicare 1

/etc/sysctl.conf
Aggiungere net.ipv4.ip_forward=1

sudo iptables -t nat -A POSTROUTING -o eth0 -j MASQUERADE
sudo iptables -A FORWARD -i eth0 -o wlan0 -m state --state
RELATED,ESTABLISHED -j ACCEPT
sudo iptables -A FORWARD -i wlan0 -o eth0 -j ACCEPT

per salvare
sudo sh -c "iptables-save > /etc/iptables.ipv4.nat"

per avvio automatico
/etc/network/interfaces
up iptables-restore < /etc/iptables.ipv4.nat
```

Setup Raspberry – Wireshark

Seguire le seguenti istruzioni per installare Wireshark, un network protocol analyzer:

1) Installazione Wireshark

```
Sudo apt-get install wireshark
nano /usr/share/wireshark/init.lua
disable_lua = true
```

2) Per avviare la cattura del traffico in transito dall'interfaccia wlan0

```
su
sudo tskark -i wlan0 -w /test.pcap
Chmod pi /test.pcap
```

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

VALUTAZIONE SOFTWARE DI ANALISI

Al fine di individuare un software adeguato per l'elaborazione della presente tesina sono stati analizzati e valutati alcuni software soffermandosi sulle caratteristiche elencate di seguito:

- open source/closed source, free/pagamento
- ambiente windows/unix-linux
- linea di comando/ide
- programmazione mediante linguaggio di scripting
- estendibile mediante package e disponibilità di package statistici

La nostra attenzione si è focalizzata su tre prodotti: Matlab, Octave, R.

Matlab

Matlab è senza dubbio il prodotto leader per questa tipologia di applicazioni, infatti le caratteristiche del prodotto e dei package presenti di default e acquistabili da terze parti coprono qualsiasi esigenza, superando di gran lunga le altre piattaforme, ma purtroppo è a pagamento.

Octave

Octave è concettualmente paragonabile a Matlab, open source e free, per tutte le piattaforme, supporta quasi tutte le istruzioni base di matlab. Purtroppo i package sviluppati per MatLab spesso non sono compatibili.

R

R è una delle migliori piattaforme per l'analisi di dati e possiede circa 2000 package per l'analisi statistica, che è sostanzialmente il campo d'interesse per l'elaborato. E' disponibile per tutte le piattaforme ed è free.

La tabella seguente (tab. 1) riepiloga le caratteristiche di ogni software, il cui confronto ci ha fatto propendere per la scelta di R.

software	Open source	Unix/linux	Ide	Scripting	packages	free
matlab	No	Si	Si	Si	Si	No
octave	Si	Si	No	Si	Si	Si
r	Si	Si	Si	Si	Si	Si

(tab. 1)

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

INSTALLAZIONE DI R SU UBUNTU

Per installare R mediante apt è necessario aggiungere il repository del CRAN che contiene i sorgenti per il sistema operativo ubuntu ed aggiungere la chiave pubblica per verificare l'integrità e l'autenticità dei pacchetti scaricati.

```
sudo sh -c 'echo "deb http://cran.rstudio.com/bin/linux/ubuntu trusty/" >> etc/apt/sources.list'
gpg --keyserver keyserver.ubuntu.com --recv-key E084DAB9
gpg -a --export E084DAB9 | sudo apt-key add -
```

Lanciare i seguenti comandi che consentono dapprima per aggiornare l'elenco dei package e quindi di installare R.

```
sudo apt-get update
sudo apt-get -y install r-base
```

Per testare l'installazione digitare il comando "R" che dovrebbe produrre la visualizzazione dei dati relativi alla versione corrente, poi q()+invio per uscire:

```
R

R version 3.2.1 (2015-06-18) -- "World-Famous Astronaut"
Copyright (C) 2015 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

…
```

Per lo svolgimento del progetto è stata selezionata un'ide compatibile con R chiamata Rstudio; per la sua installazione è necessario scaricare il package da https://www.rstudio.com/ ed installarlo da shell con il seguente comando:

```
sudo dpkg -i rstudio-<version>.deb
rstudio
```

Per lanciare gli script sviluppati per il presente elaborato è necessario installare dal prompt di R alcuni package aggiuntivi:

```
Install.packages("stringr")
Install.packages("argparser")
install.packages("roxygen2")
install.packages("devtools")
```

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

METRICHE

Al fine di valutare il traffico acquisito sotto diversi punti di vista abbiamo pensato di definire alcune metriche che ci avrebbero guideranno sia nello sviluppo degli script che nell'interpretazione dei dati:

TCP vs UDP (script 1):

- totale flussi tcp/udp
- totale segmenti/dati tcp
- totale segmenti /dati udp
- campionamento numero flussi tcp e udp (ogni secondo)
- campionamento segmenti tcp/udp (ogni secondo)
- campionamento dati tcp/udp (ogni secondo)
- andamento ack duplicati

Andamento RTT nel tempo (script 2):

- min, max, media

Distribuzione durata dei flussi (script 3):

- TCP vs UDP

Distribuzione trasmissioni flussi (script 4):

- TCP vs UDP

Analisi instaurazione e finalizzazione flussi TCP (script 5):

- numero di flussi tcp con three way handshake corretto
- numero di flussi tcp terminati da FIN
- numero di flussi tcp terminati da RST
- grafico andamento dei reset nel tempo TCP vs UDP

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

SOFTWARE SVILUPPATO

Questa sezione illustra i passi necessari per il download, la configurazione e l'uso dei vari script sviluppati per il presente elaborato da applicare ad un file in formato PCAP che consentano di dare una risposta alle metriche stabilite nel precedente paragrafo.

Git

L'intero progetto e la documentazione è scaricabile da Git:

```
git clone https://github.com/DanieleCampagnoli/TesinaTIR
```

Struttura del software e contenuto dei file

Il software è contenuto in una cartella chiamata TesinaTIR la quale contiene due sottocartelle:

- scripts: contiene i vari script che partendo da un file PCAP estraggono metriche sopra indicate producendo i relativi grafici;
  Per lanciare un qualsiasi script presente in tale cartella è sufficiente sostituire la x con il numero di script desiderato digitando il seguente comando:

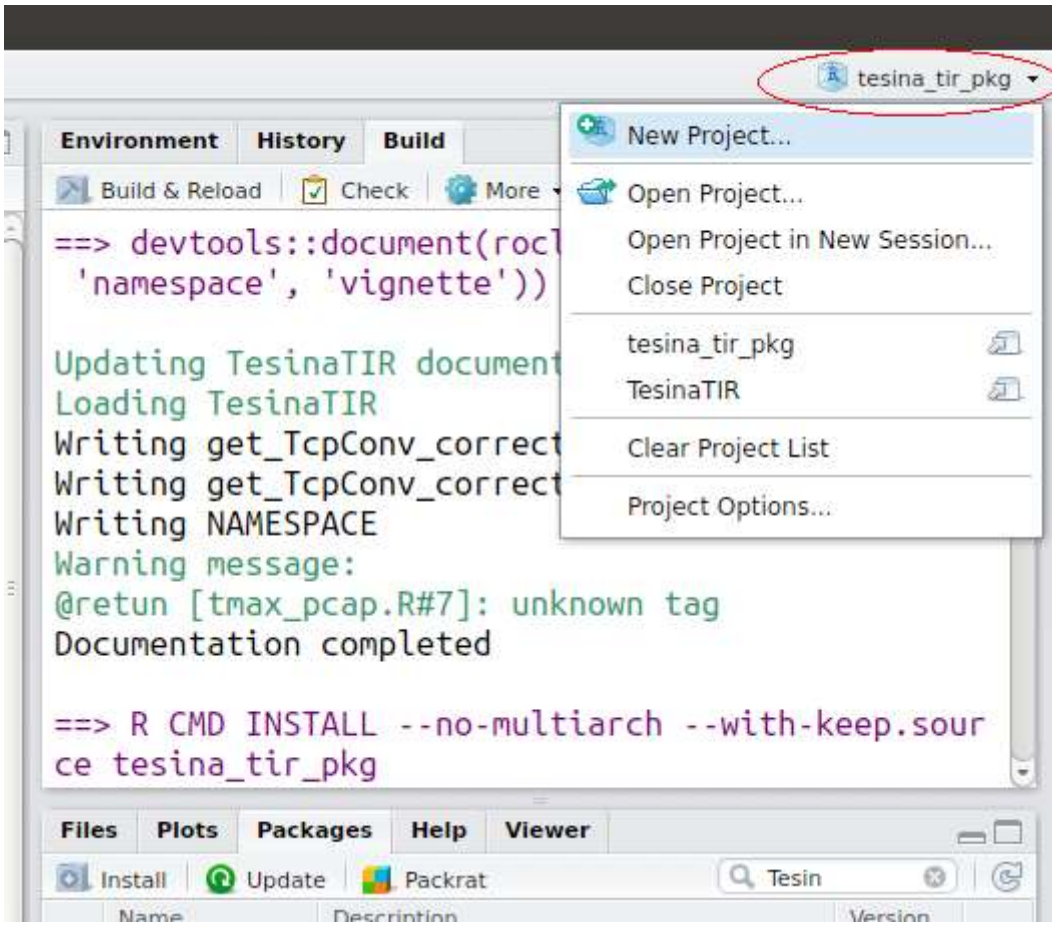
  ```
  Rscript script_x.R -h
  ```

  Al suo interno è inoltre presente una sottocartella pcap per contenere i flussi acquisiti e da elaborare;
- tesina_tir_pkg: package di R che contiene le funzioni utilizzate da dagli script inclusi nella cartella precedente ed è strutturata come segue:
  - R: contiene il codice sorgente delle funzioni
  - man: contiene il manuale in formato .Rd
  - inst/extdata: contiene un pcap di esempio associato al package
  - altri file necessari alla gestione del progetto

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

Installazione TesinaTIR

Di seguito verranno elencate le operazioni per utilizzare il software sviluppato. I vari sottoparagrafi devono essere eseguiti in sequenza e la procedura di installazione è stata testata solo su ubuntu.
Importare il package in rstudio cliccando su “Open Project”



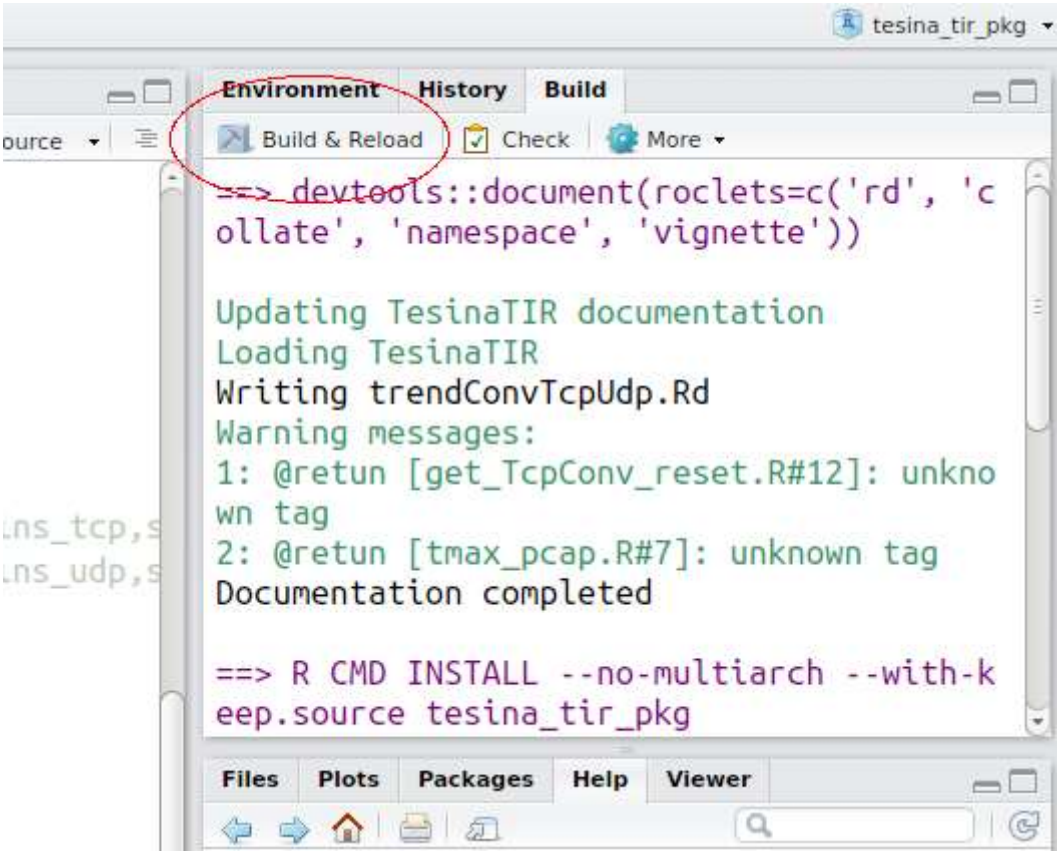
CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

Impostare le seguenti opzioni:

```
Menù Build -> Configure Build Tools -> Build Tools
  -> (baffetto su Generate Documentation with Roxygen)
```

```
Menù Build -> Configure Build Tools -> Build Tools->Configure
  -> (baffetto su tutte le opzioni)
```

cliccare su “Build & Reload” per installare e caricare la libreria in r



Lanciare I seguenti comandi per testare l'installazione del package

```
library(TesinaTIR)
pcapName<-system.file("extdata", "dump.pcap", package = "TesinaTIR")
tmax<-tmaxPcap(pcapName)
```

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

Documentazione

La documentazione del software sviluppato è suddivisa in due parti:

- documentazione degli script
- documentazione del package tesina_tir_pkg

La documentazione degli script è ottenibile lanciando lo script da terminale con il seguente comando, dove x indica il numero dello script.

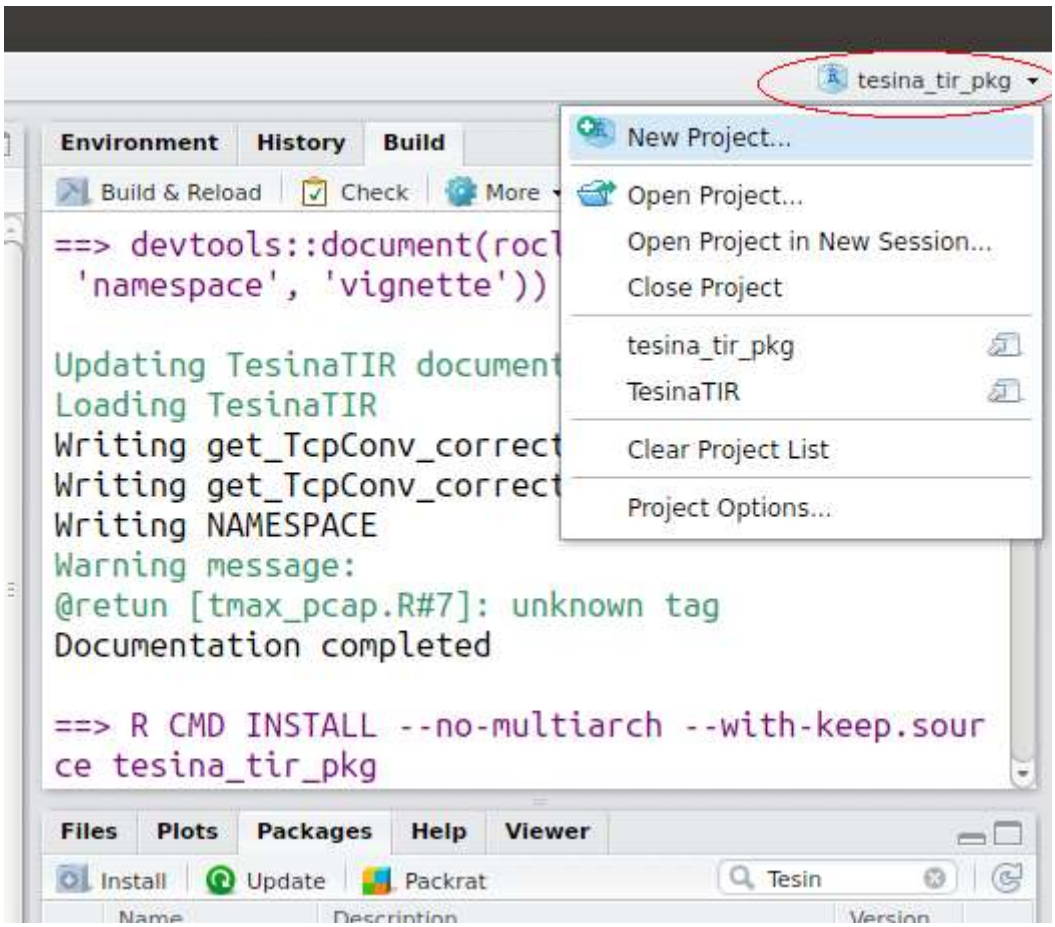
```
Rscript script_x.R -h
```

Di seguito è descritto un esempio di esecuzione di uno script.

```
Rscript script_1.R -i ./pcap/GazzettaDiModena_20161120.pcap -o ./out_x
```

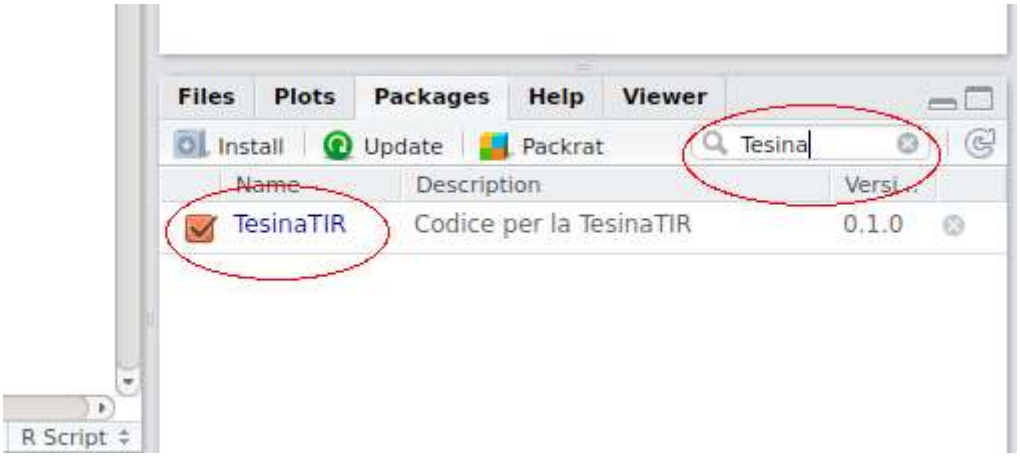
La documentazione del package TesinaTIR è ottenibile nel seguente modo:

- importare il package in rstudio cliccando su “Open Project”



CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

- ricercare il package nel menu packages inserendo nella barra la stringa “TesinaTIR” e cliccare nei risultati su TesinaTIR.



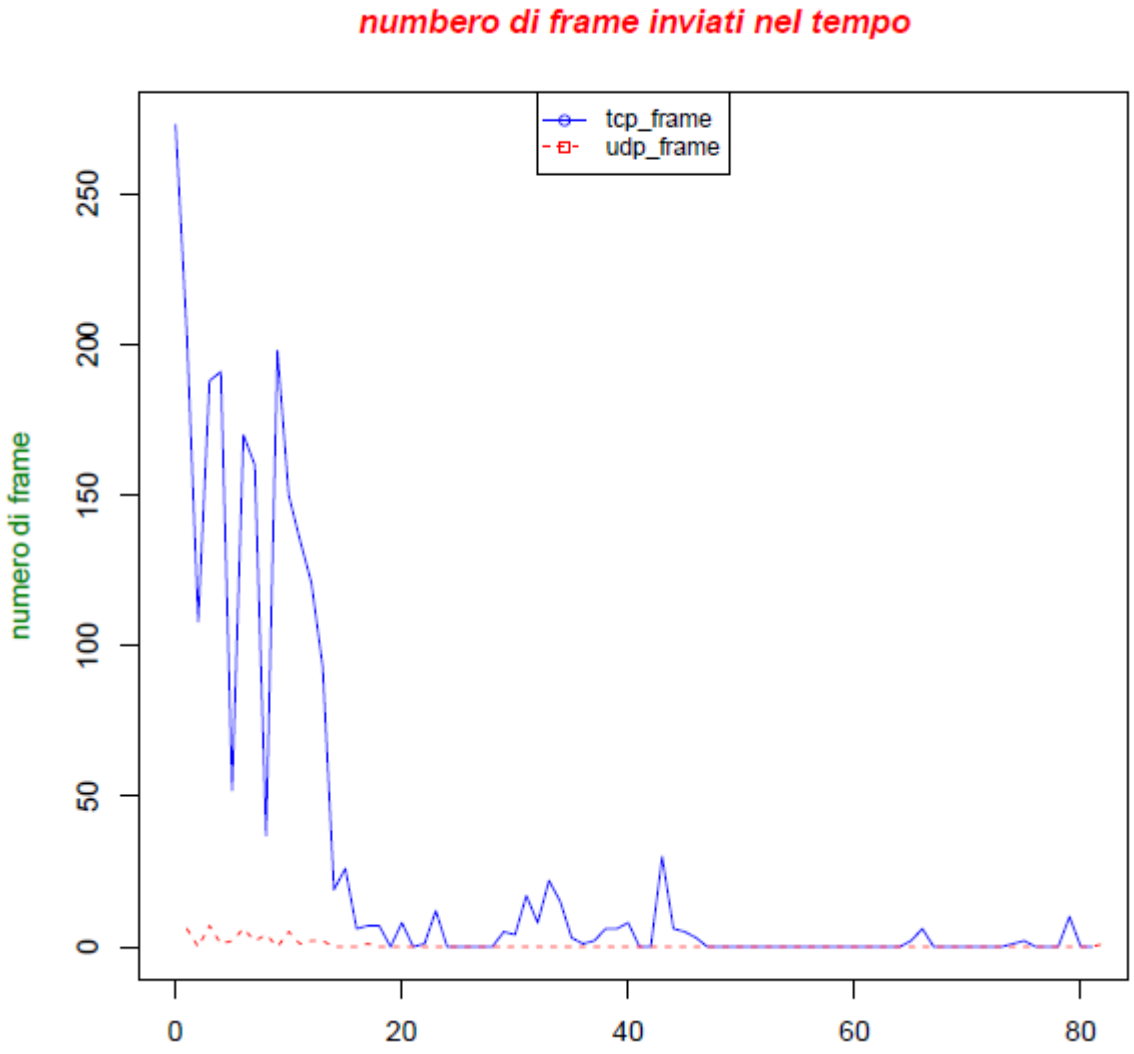
CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508

ELABORAZIONE TRAFFICO E RISULTATI

L'elaborazione delle metriche su un PCAP acquisito ha prodotto i seguenti risultati:

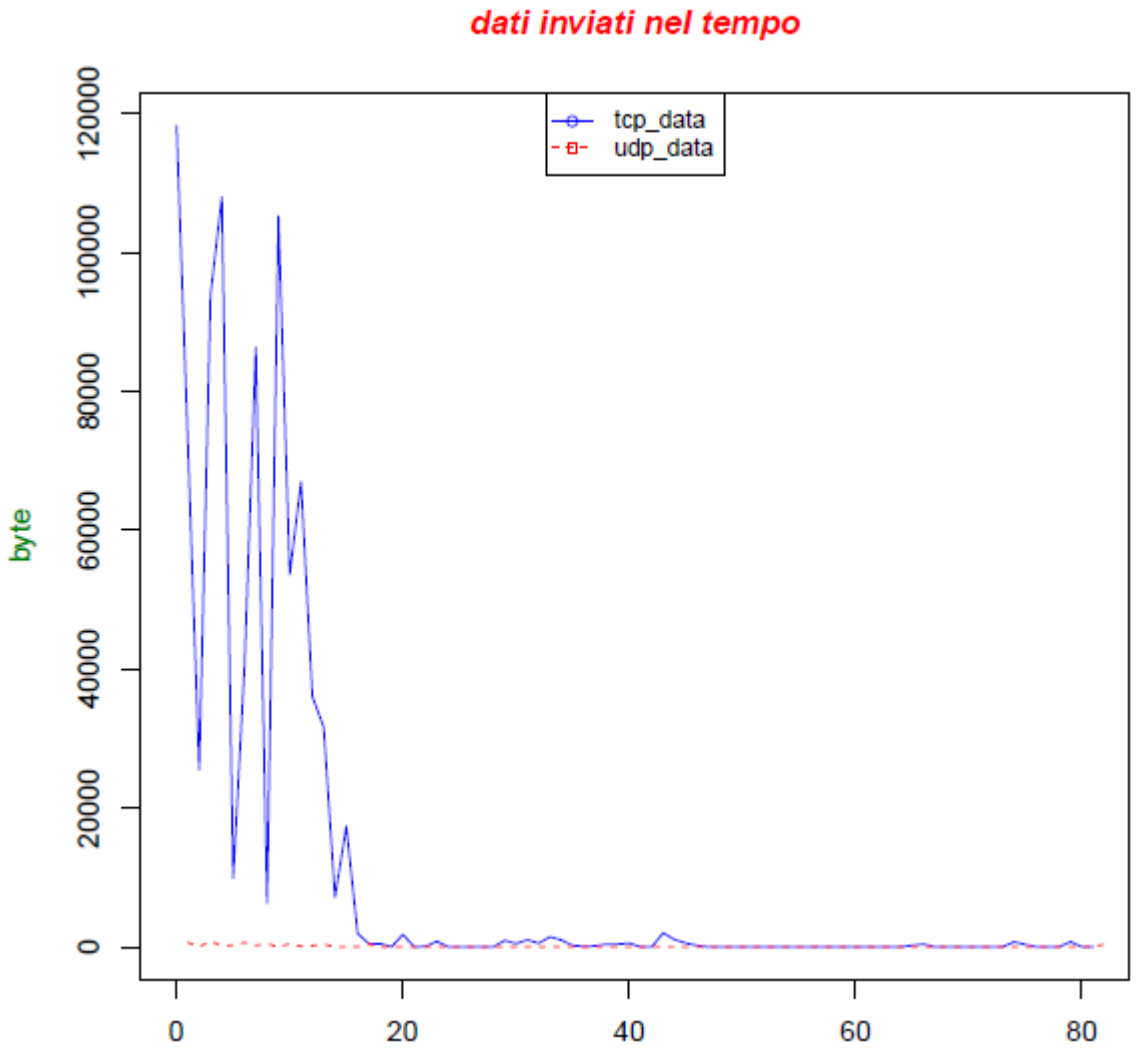
TCP vs UDP (script 1):

- totale flussi tcp/udp
- totale segmenti/dati tcp
- totale segmenti/dati udp
- campionamento numero flussi tcp e udp (ogni secondo)



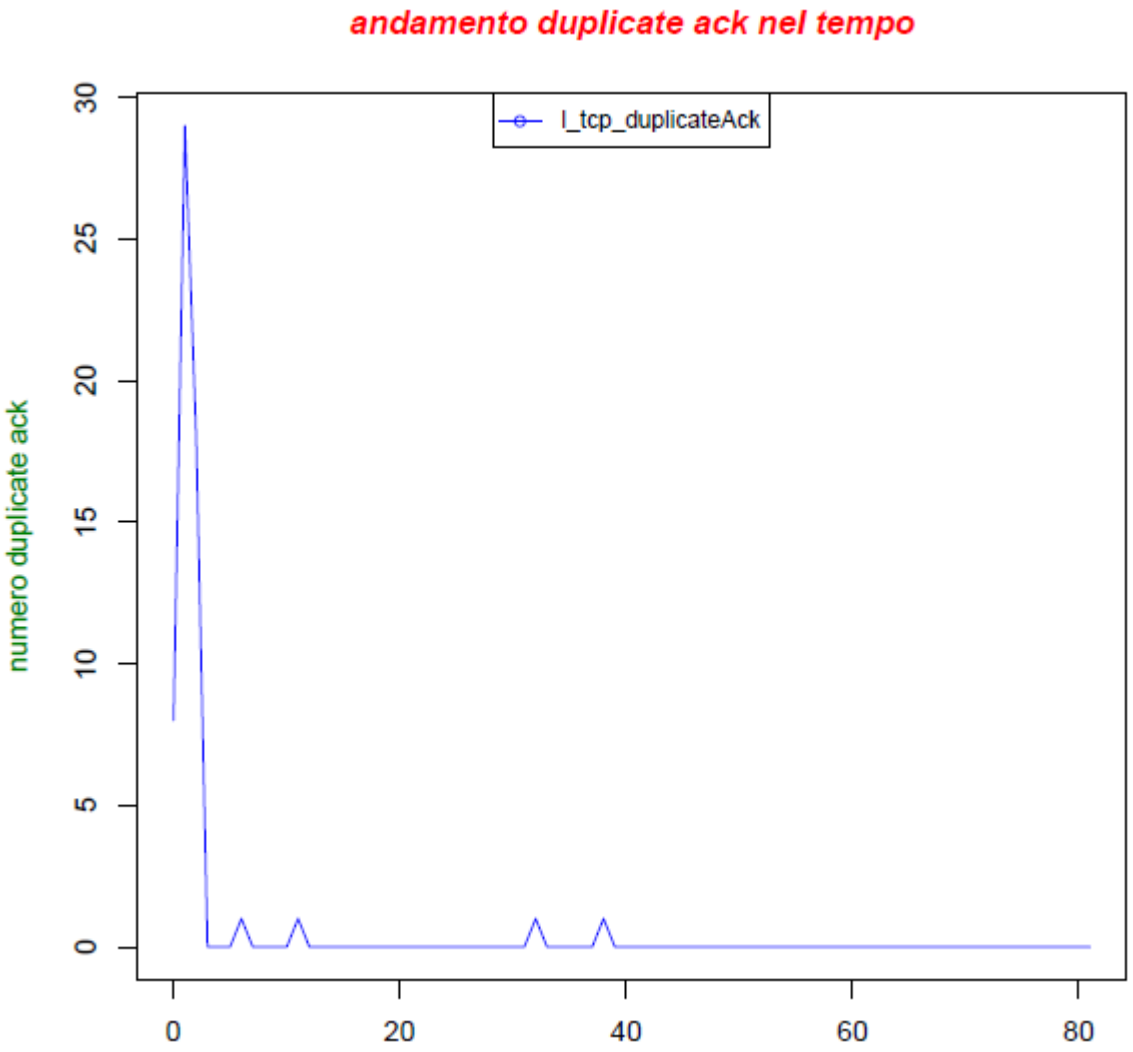
- campionamento segmenti tcp/udp (ogni secondo)
- campionamento dati tcp/udp (ogni secondo)

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508



- andamento ack duplicati

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508



Andamento RTT nel tempo (script 2):

- min, max, media

Distribuzione durata dei flussi (script 3):

- TCP vs UDP

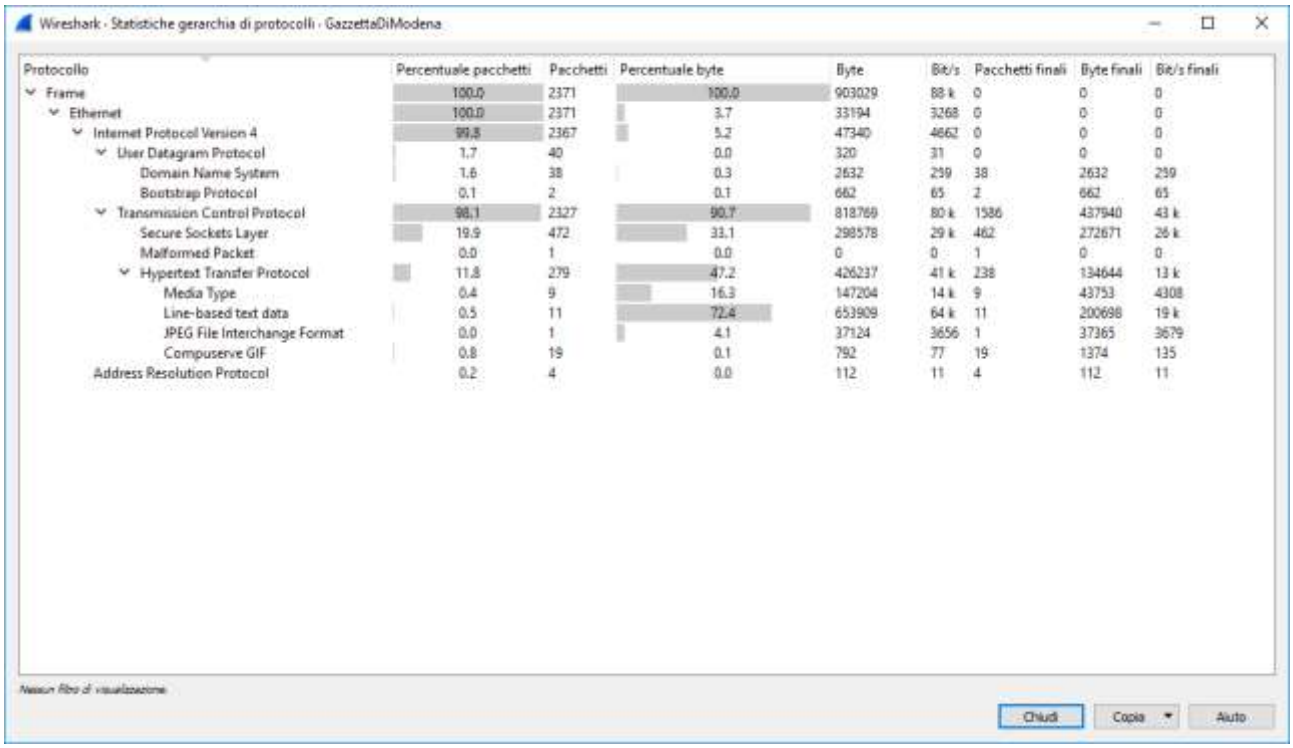
Distribuzione trasmissioni flussi (script 4):

- TCP vs UDP

Analisi instaurazione e finalizzazione flussi TCP (script 5):

- numero di flussi tcp con three way handshake corretto
- numero di flussi tcp terminati da FIN
- numero di flussi tcp terminati da RST
- grafico andamento dei reset nel tempo TCP vs UDP

CAMPAGNOLI DANIELE matr. 94197
MONTAGNOLA GIUSEPPE matr. 95508



Da sistemare:
riferimento RTT con ping
nel primo script replace “frame” con segmenti
nel primo script numBero frame … -> numero
nel primo script togliere INVIATI
script 3, si possono mettere insieme tcp e udp?
script 4, si possono mettere insieme tcp e udp?