Relazione Reti Geografiche

Packet Capture on General Purpose Systems

Egidio Mario 0522500969

Fasolino Lorenzo 05121xxxxx

Sommario

[Definizione Esperimento 3](#_Toc54548285)

[Strategia utilizzata 4](#_Toc54548286)

[Packet Capture (selenium/wireshark/python) 5](#_Toc54548287)

[Analisi pacchetti(python/json) 6](#_Toc54548288)

[Risultati 7](#_Toc54548289)

# **Definizione Esperimento**

Setting

1. Installare wireshark (https://www.wireshark.org/) 2. installare Selenium (https://www.selenium.dev/)

Metodolgia

1. Definire il workload (set dei primi 50 siti Web piu` popolari, <https://www.alexa.com/topsites>)
2. Attivare wireshark e richiedere le pagine attraverso l’utilizzo di Selenium
3. Scrivere uno script che acceda al log di wireshark e calcoli:
   1. Il numero di pacchetti https
   2. il numero di siti terzi contattati

L’esperimento verrà ripetuto 5 volte aumentando di volta in volta il tempo di capture.

1. Exp1. 5 minuti di packet capture
2. Exp2. 10 minuti di packet capture
3. Exp3. 15 minuti di packet capture
4. Exp4. 20 minuti di packet capture
5. Exp5. 25 minuti di packet capture

Analisi dei risultati

1. Quanto spazio su disco serve per immagazzinare i dati dei 5 esperimenti?
2. Quanto tempo impiega lo script per ciascuno dei 5 esperimenti?

# **Strategia utilizzata**

L’esperimento consiste nel catturare ed analizzare i pacchetti http che viaggiano dopo aver invocato i 50 siti più visitati al mondo.

Quindi come prima cosa è stato definito il workload (set dei primi 50 siti Web più popolari, (<https://www.alexa.com/topsites>).

Il passo successivo è stato quello di installare Wireshark (https://www.wireshark.org/) e Selenium (<https://www.selenium.dev/>). Dopodiché è stato realizzato uno script in Python per il lancio automatizzato dei siti web. Lo script è stato ripetuto 5 volte:

1. Exp1. 5 minuti di packet capture

2. Exp2. 10 minuti di packet capture

3. Exp3. 15 minuti di packet capture

4. Exp4. 20 minuti di packet capture

5. Exp5. 25 minuti di packet capture

Ad ogni esperimento sono stati catturati i pacchetti con l’utilizzo di Wireshark.

Dopo averli catturati sono stati esportati come file in formato JSON.

Poi è stato realizzato uno script (sempre in Python) che prende in input i file JSON esportati da Wireshark e calcola:

1. Il numero di pacchetti https
2. Il numero di siti terzi contattati
3. Quanto tempo impiega lo script per ciascuno dei 5 esperimenti

Infine, è stato calcolato quanto spazio su disco serve per immagazzinare i dati dei 5 esperimenti.

# **Packet Capture (selenium/wireshark/python)**

(Qui parliamo dei problemi con i siti che abbiamo sostituito e dello script con selenium)

# **Analisi pacchetti(python/json)**

(Qui parliamo dello script che analizza i JSON)

# **Risultati**

(Qui mettiamo i risultati con dei grafici)