



UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA
DISCIPLINA: INTRODUÇÃO A PROGRAMAÇÃO
PROFESSOR: WESLEY EMMANUEL MARTINS LIMA

SEGUNDA AVALIAÇÃO

1. O Squid é um famoso software livre usado para controlar e registrar os acessos à internet em uma rede de computadores. Ele armazena as páginas utilizadas mais recentemente em disco(CACHE) para responder aos clientes da rede que a solicitarem em um tempo menor. Para cada acesso feito a um site na internet por um usuário, ele gera uma linha em um arquivo de log com um formato padrão descrito no exemplo abaixo:

```
1347542919.400 53 10.2.10.26 TCP_MISS/200 4596 GET
http://l.yimg.com/zz/combo? - DIRECT/200.152.174.74
application/javascript
```

Os campos deste arquivo são separados por espaços. São eles:

1. Data no formato de marca de tempo do Unix(número de segundos transcorridos desde 1/1/1970). Para converter esta data para um formato legível, você pode utilizar a biblioteca datetime como no exemplo a seguir:

```
import datetime
timestamp = 1347542919.400
value = datetime.datetime.fromtimestamp(timestamp)
print(value.strftime('%Y-%m-%d %H:%M:%S'))
```

2. Tempo gasto em milisegundos para atender a requisição.
3. Endereço IP do cliente
4. Código resultante do acesso. Alguns deles são:
 - a. TCP_MISS – Objeto obtido diretamente da internet;
 - b. TCP_HIT – Objeto obtido da cache do squid;
 - c. TCP_DENIED – Acesso negado pelo controle de acesso
5. Quantidade de bytes entregue ao cliente;
6. Método de requisição do objeto pelo cliente: GET ou POST;
7. Endereço acessado(URL)
8. hífen
9. Indica o endereço IP da máquina contactada para a obtenção da url solicitada. Caso a página seja obtida da cache, este campo aparece com a palavra NONE. Caso contrário, aparece a palavra DIRECT seguida do endereço IP da máquina contactada.
10. O tipo de conteúdo retornado pela solicitação da URL;

De posse destas informações, escreva um programa em python que gere uma série de relatórios com as seguintes informações:

1. Número de sites diferentes acessados por cada cliente IP;



UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA
DISCIPLINA: INTRODUÇÃO A PROGRAMAÇÃO
PROFESSOR: WESLEY EMMANUEL MARTINS LIMA

- a. Dê a opção de listar os acessos de todos os clientes ou de um IP específico.
2. Número total em bytes de dados transferidos por cada cliente;
 - a. Dê a opção de listar o total de todos os clientes ou de um IP específico.
3. Percentual de sites encontrados na cache e de sites trazidos diretamente da internet;
4. Lista de sites com acesso negado para cada cliente;
 - a. Dê a opção de listar os acessos negados de todos os clientes ou de um IP específico.
5. Lista de sites acessados por um cliente em um determinado dia;
 - a. Solicite o IP do cliente e a data.
6. Lista de quantitativos, por conteúdo, dos sites acessados por um dado cliente;
 - a. Solicite o IP do cliente.

O programa deverá ser baseado em um menu com as 6(seis) opções descritas acima. Ao escolher uma dada opção, o programa deve solicitar as informações necessárias para exibir o que é solicitado. Deve haver uma opção no menu para carregar os dados de acesso de um arquivo e outra opção para salvar os dados de todos os relatórios acima em um arquivo texto a ser especificado pelo usuário.

Caso o usuário tente visualizar qualquer um dos relatórios sem antes carregar um arquivo de acessos, deve ser gerado um alerta para o usuário informando tal fato.

A seguir, o formato do arquivo texto a ser gerado contendo o relatório:



UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA
DISCIPLINA: INTRODUÇÃO A PROGRAMAÇÃO
PROFESSOR: WESLEY EMMANUEL MARTINS LIMA

Sites diferentes por IP:

10.2.5.23 – 3 sites

10.2.40.21 – 10 sites

Total de bytes transferidos por IP

10.2.3.54 – 2304 bytes

10.2.5.67 – 123434 bytes

Percentual de sites na cache e de acesso direto

Cache: 64%

Direto: 36%

Lista de sites com acesso negado

10.2.3.45

www.playboy.com

www.prive.com

-

10.2.5.75

www.xxx.com

www.hot.com

-

Quantitativos de acessos por conteúdo

text/html – 203

image/gif – 349

application/javascript – 493
