

Unit 11 - Week 9

Course outline

How to access the portal

Week 0 Assignment 0

Week 1

Week 2

Week 3

Week 4

Week 5

Week 6

Week 7

Week 8

Week 9

- Lecture 44 : Decision support Concepts
- Lecture 45 : Understanding the web log
- Lecture 46 : Understanding the web log-II
- Lecture 47 : Using the Web log: Web usage mining
- Lecture 48 : User behaviour modelling from web log
- Week 9: Lecture Material**
- Quiz : Assignment 9
- Feedback For Week 9

Week 10

Week 11

Week 12

Download Videos

Assignment Solution

Assignment 9

The due date for submitting this assignment has passed. As per our records you have not submitted this assignment.

Due on 2019-10-02, 23:59 IST.

1) In the log file the field "Agent" denotes - 2 points

a. The person who is accessing the website.
 b. The page to which the data is send
 c. The URL of the Server
 d. Name and version of the browser

a
 b
 c
 d

No, the answer is incorrect.
Score: 0
Accepted Answers: d

2) Type of the browser cannot be used for user identification because - 2 points

a. Many people can have same browsers
 b. Browsers are unique for each user
 c. cookies are not working on all browsers
 d. None of these

a
 b
 c
 d

No, the answer is incorrect.
Score: 0
Accepted Answers: a

3) In the access log file, the field "referrer" denotes - 2 points

a. The person who is accessing the website.
 b. The page to which the data is send
 c. The URL from which the request has been generated
 d. Version of the browser

a
 b
 c
 d

No, the answer is incorrect.
Score: 0
Accepted Answers: c

4) 2 points

Question 4 to 6 is connected. The description of the problem is given below -

Suppose you are a website administrator of ABC multinational company. In the website, data related to employees are displayed in "employee.php" page, and confidential information related to salary is stored in "salary_sheet.php" page. "salary_sheet.php" page is accessible for authorized users only. On 21/2/2018 morning, you came to know that employee salary sheet data has been hacked and modified. The file modification time is between 4:26 AM to 4:32 AM. You have to check the website log and analyze. Before analysis you have to keep in mind that, IP allocation is dynamic. So, IP may change on every new session.

Sl no	Field 1	Field 2	Field 3	Field 4	Field 5	Field 6
1	203.110.246.11	2/20/2018 4:25:37	GET	HTTP/1.1	http://www.abccompany.org/employee_data.php	Mozilla/5.0 (Window 10)
2	205.123.200.1	2/20/2018 4:26:10	GET	HTTP/1.1	http://www.abccompany.org/login.php	Mozilla/5.0 (Window 10)
3	203.110.246.11	2/20/2018 4:26:15	POST	HTTP/1.1	http://www.abccompany.org/login.php	Mozilla/5.0 (Window 7)
4	205.123.234.1	2/20/2018 4:26:40	POST	HTTP/1.1	http://www.abccompany.org/login.php	Mozilla/5.0 (Window 10)
5	203.110.210.12	2/20/2018 4:27:05	POST	HTTP/1.1	http://www.abccompany.org/login.php	Mozilla/5.0 (Window 7)
6	205.123.200.1	2/20/2018 4:27:30	GET	HTTP/1.1	http://www.abccompany.org/employee_data.php	Mozilla/5.0 (Window 10)
7	203.110.246.11	2/20/2018 4:27:37	POST	HTTP/1.1	http://www.abccompany.org/salary_sheet.php	Mozilla/5.0 (Window 7)
8	205.123.200.1	2/20/2018 4:29:05	GET	HTTP/1.1	http://www.abccompany.org/salary_sheet.php	Mozilla/5.0 (Window 10)
9	205.123.200.1	2/20/2018 4:33:20	GET	HTTP/1.1	http://www.abccompany.org/salary_sheet.php	Mozilla/5.0 (Window 10)

Which entry in the log entry will you suspect as a potential hacker? 2 points

a. Sl no #5
 b. Sl no #6
 c. Sl no#7
 d. Sl no #8

a
 b
 c
 d

No, the answer is incorrect.
Score: 0
Accepted Answers: c

5) Which pair of entries in the log represents probably the same user considering time oriented heuristics? 2 points

a. {Sl no #1 and Sl no#3}
 b. {Sl no #4 and Sl no#6}
 c. {Sl no#8 and Sl no#9}
 d. All the above

a
 b
 c
 d

No, the answer is incorrect.
Score: 0
Accepted Answers: d

6) In the log file the Field 5 resembles which entry in a standard log file? 2 points

a. Method
 b. Agent
 c. Referrer
 d. IP

a
 b
 c
 d

No, the answer is incorrect.
Score: 0
Accepted Answers: c

7) The IP address visible in a website access log is that of the 2 points

a. Server
 b. Last ISP
 c. client
 d. DNS

a
 b
 c
 d

No, the answer is incorrect.
Score: 0
Accepted Answers: b

8) What is path completion in the context of access log mining? 2 points

a. Finding edges of disconnected graph
 b. Inferring missing user references due to caching
 c. Network analysis techniques
 d. Job sequencing problems

a
 b
 c
 d

No, the answer is incorrect.
Score: 0
Accepted Answers: b

9) To build the dynamic part of user behavior graph the matrices required are - 2 points

a. Transpose and Transform
 b. Transport and Transition Time
 c. Transport and Average Transition Matrix
 d. Transition and Average Transition time

a
 b
 c
 d

No, the answer is incorrect.
Score: 0
Accepted Answers: d

10) What are the sessionization strategies in access log mining? 2 points

a. Delay and Response oriented.
 b. Time and Navigation oriented
 c. View and browse oriented
 d. GET and POST oriented

a
 b
 c
 d

No, the answer is incorrect.
Score: 0
Accepted Answers: b