Seoul National University

M1522.001400 Introduction to Data Mining

Spring 2016, Kang

Homework 7: Advertising on the Web (Chapter 8)

Due: May 23, 09:30 AM

Reminders

- The points of this homework add up to 100.
- Like all homeworks, this has to be done individually.
- Lead T.A.: Jinhong Jung (montecast9@gmail.com)
- Please type your answers in English. Illegible handwriting may get no points, at the discretion of the graders.
- If you have a question about assignments, please upload your question in eTL.
- If you want to use slipdays or consider late submission with penalties, please note that you are allowed one week to submit your assignment after the due date.

Remember that:

- Whenever you are making an assumption, please state it clearly

Question 1

Define the graph G_n to have the $2n$ node

$$a_0, a_1, \ldots, a_{n-1}, b_0, b_1, \ldots, b_{n-1}$$

and the following edges. Each node a_i, for $i = 0, 1, \ldots, n-1$, is connected to the nodes b_j and b_k, where

$$j = 2i \; mod \; n \text{ and } k = (2i+1) \; mod \; n$$

For instance, the graph G_4 has the following edges: (a_0, b_0), (a_0, b_1), (a_1, b_2), (a_1, b_3), (a_2, b_0), (a_2, b_1), (a_3, b_2), and (a_3, b_3). Answer the following questions. [30 points]

(a) Find one of perfect matchings for G_4. [10 points]

(b) Find one of perfect matchings for G_5. [10 points]

(c) How many perfect matchings do the graphs G_4 and G_5 have? [10 points]

Question 2

Whether or not the greedy algorithm for matching on a bipartite graph gives us a perfect matching for the graph of Figure 1 depends on the order in which we consider the edges. Among the 6! possible orders of the six edges, how many orders do they give us a perfect matching? Give a simple test for distinguishing those orders that do give the perfect matching from those that do not. [35 points]

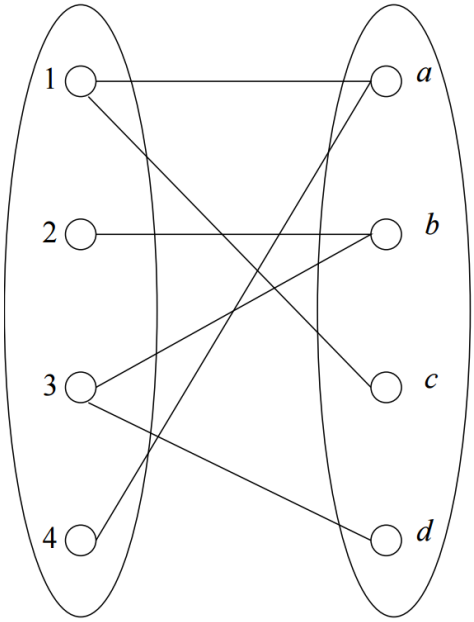

Figure 1. A bipartite graph

Question 3

Suppose that there are three advertisers, A, B, and C. There are three queries, x, y, and z. Each advertiser has a budget of 2. Advertiser A bids only on x; B bids on x and y, while C bids on x, y, and z. Note that on the query sequence xxyyzz, the optimum off-line algorithm would yield a revenue of 6, since all queries can be assigned. Answer the following questions. [35 points]

(a) Show that the greedy algorithm will assign at least 4 of these 6 queries. [20 points]

(b) Find another sequence of queries such that the greedy algorithm can assign as few as half the queries that the optimum off-line algorithm assigns on that sequence. [15 points]