Sample problem Proposal (Junior), 2022

Singapore International Modelling Challenge

Sample Problem Scope

There are two parts to the challenge. Part 1 contains individual questions to be answered by hand, while Part 2 is to be completed using computer program in Python, together with a report. This sample problem gives an illustration of the type of questions to be expected in the main challenge, while the subject matter, breath and depth may be different. For part 2 of the challenge, you will be attending trainings to learn some techniques and methods in Machine Learning.

Sample Challenge Problem

1.1 Efficient Networks

This sample question is adapted from the SIMC 2016 Part A[1]. In this question we imagine a large network of computers, linked by cables. Initially, we have a network of n^2 computers, arranged in a square grid. Neighbouring computers are joined by a cable, as shown in Figure 1.1 (left). Formally defined, the computers are the nodes (circular dots) in the network, while the cables are the links (lines) connecting nearby nodes.

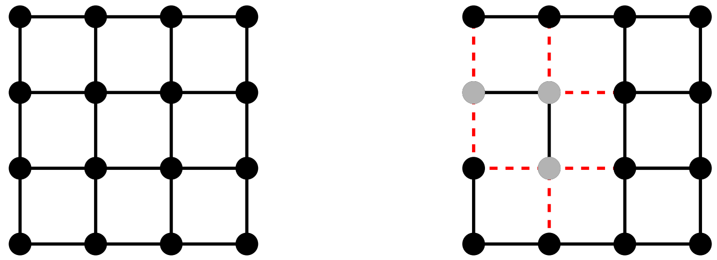

Figure 1.1: Example network with $n = 4$. Left: The original network. Right: The breaking of 7 connections (red dashed links) has disconnected a set of 3 computers (grey nodes) from the rest of the network.

If say two cables break from a corner computer, then clearly that computer is cut off from the rest of the network. We are concerned with some cables breaking so that some k of the computers all become unable to communicate with the remaining $n^2 - k$. For example, if 7 cables break as shown in Figure 1.1 (right) then $k = 3$ computers become disconnected from the remaining 13. In this case, the network disintegrates into two separate connected clusters: one cluster consisting of the 3 grey nodes, and the other consisting of the 13 black nodes. We say the sizes of the clusters are 3 and 13 respectively.

Question 1a. If $n = 10$, what is the smallest number of cables whose failure would disconnect

[1] https://www.nushigh.edu.sg/qql/slot/u90/2020/Admissions/Outreach/SIMC/Challenge%20Resources/2016/SIMC2016_ChallengeQuestion.pdf.

some set of $k = 4$ computers from the rest?

Question 1b. If $n = 10$, what is the smallest number of cables whose failure would disconnect some set of $k = 36$ computers from the rest?

Question 1c. If $n = 10$, what is the smallest number of cables whose failure would disconnect the network, such that the size of the largest cluster is $k = 36$?

Suppose now that our concern is to minimise transmission times. For example, sending a message from A to B in Figure 1.2 would need 4 cables to be traversed (in other words, the 'distance apart' of A and B is 4), which we assume takes 4 units of time. We have the option of adding some 'superhighways': a superhighway is a connection between two computers, using a very fast cable, so fast that the time needed for a signal to traverse it is negligible. For example, if we add a superhighway between C and D in Figure 1.2 then the time needed for a message to go from A to B is down to 3 – so the 'distance' between them is now down to 3. We also call the maximum distance between any pair of nodes as the *diameter* of the network.

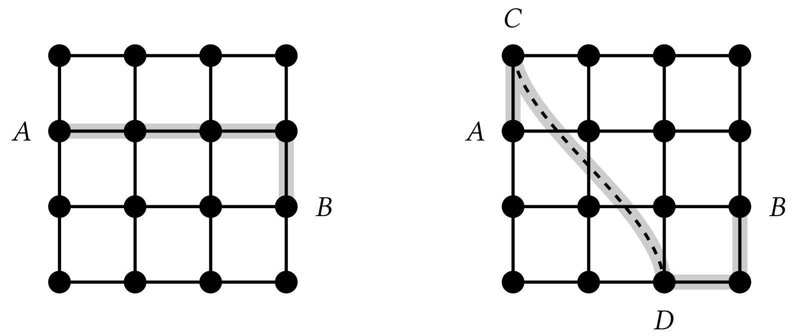

Figure 1.2: Example network with $n = 4$. Left: The distance between A and B in the original network is 4. The highlighted path is one of several length − 4 paths between A and B. Right: After the superhighway between C and D is added, the distance between A and B is reduced to 3. The highlighted path is one of two length − 3 paths between A and B.

Question 1d. For $n = 5$ as shown in the left of Figure 1.3, if we are only allowed to add one superhighway that originates from X, where can the other end of the superhighway Y be, such that the maximum distance (*diameter*) that any signal has to travel in the resulting network is the smallest? Note that we label the position of X as (3,1), since it is in the third row and first column from the bottom left.

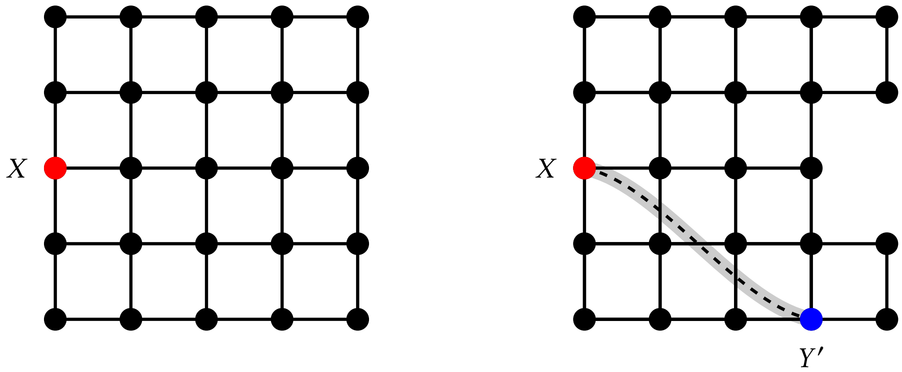

Figure 1.3: Network with $n = 5$, and X is the starting point of the superhighway. Left is the original network. Right is the network with the node at position (3,5) removed, and a superhighway is build between X and Y'.

1.2 Modelling and Machine learning

In this section, you will need to use computer programs to solve the problems. You can choose to use any open source package in Python if needed.

Question 2a. Write a computer program to compute the diameter of the network for any end node Y of the superhighway in Question 1d, and use it to find the solution for Question 1d.

Question 2b. One day the computer at the position (3,5) of the network is removed from the network as shown in the right of Figure 1.3, and no signal can pass through it. Write a computer program to find out the new superhighway that minimize the diameter of the network.

Question 2c. Construct a machine learning model to predict the network diameter value for given superhighway XY in the network in the right of Figure 1.3, where X is the node (3,1) as shown, and Y is any other node. You will need to construct training data first, i.e. input features of every superhighway and output label which is the diameter associated with the superhighway.

Hint: For example, the input feature to the machine learning model can be the location of the end node, and the output is the diameter. To illustrate, for the end node Y' at position (1,4) (blue color in the right of Figure 1.3), its relative distance with X is (-2,3), and the diameter of the network with superhighway connecting X and Y' is 7. You can then write the input feature of end node Y' as (1,4,-2,3), and output label is 7. There are 24 nodes in the network, and one of which is X. Therefore, Y can take any of the remaining 23 nodes, so that there should be 23 samples to train your machine learning model.

Question 2d. If instead of node (3,5) in question 2b, another computer or computers in the network is/are removed, is the machine learning model trained in Question 2c able to find out the new diameters from different superhighways? What if instead of network size $n = 5$, n is a very large value?