Midterm exam

Fall 2022

Analysis of Network Data Tuesday 10/25

Please do not turn the page until instructed to do so.

This exam has 6 questions.

You can pick 4 of the 6 questions to submit for a total of 100 points.

Every question you pick will be equally weighted.

If you pick more than 4 questions, 4 of them will be selected at random for grading.

There are 6 pages in the exam.

The exam is **a take-home exam**.

It is due by class time (2:00 pm Central time) on Tuesday 11/01.

You are **not** allowed to collaborate or communicate with others about the questions in this exam.

You are **perfectly allowed** to discuss and collaborate with me!

Answer <u>every question</u> to the best of your knowledge.

Make sure to show your work for **partial credit**.

Some questions (like, for example, Question 2) will ask you to use Python, Gurobi, networkx or the software of your choice. You do **not** have to solve this one if you prefer to avoid coding during the exam.

Good luck!

Question 1: Smaller proofs and algorithms

(a) A matching is a set of edges such that no two edges include the same node. An independent set is a set of nodes such that no two nodes are adjacent (connected by an edge). If M is a matching and $\mathcal{I}$ is an independent set, show that $|M| + |\mathcal{I}| \leq |V|$.

(b) A cycle and a cut-set (set of edges that separate the network in two sets S and $\overline{S}$) intersect each other in an even number of edges. True or False? Prove or provide a counter example to the statement.

(c) Consider a weighted, directed graph $G(V, E, w)$, where all weights w on the edges are positive (i.e., $w_{ij} > 0$). Then, build two new weighted graphs $\hat{G}(V, E, \hat{w})$ and $\tilde{G}(V, E, \tilde{w})$ such that the weights are the negatives of the ones in G (i.e., $\hat{w}_{ij} = -w_{ij}$) and such that the weights are the reciprocals of the ones in G (i.e., $\tilde{w}_{ij} = \frac{1}{w_{ij}}$), respectively. Prove or disprove the following:

1. Every shortest path in $\hat{G}$ is a longest path in G.
2. Every shortest path in $\tilde{G}$ is a longest path in G.

(d) Given an edge-weighted directed graph with non-negative weights, design an algorithm for finding the shortest path from a given source s to a terminal t when you can change the cost of *any edge* to 0. What is the computational complexity of the algorithm?

Question 2: Emergency crew routing

Last winter, a number of cities in the north states of the US faced threatening ice, snow, and cold weather conditions that caused provision of electricity to become intermittent. As a measure to prevent that and improve societal welfare, state and city authorities decided to locate repair crews and equipment in certain strategic areas in order to quickly assess, respond to, and repair such damages.

We represent this problem below on a graph. Figure 1 represents all candidate locations where crews can be stationed (in red), all arcs are given travel times, and all power system components (in blue) that can fail. If a power component is located in the middle of a street, then only half the traversal time needs to be "paid" before the crew reaches the power station. Assume that no streets are ever closed down, and hence the repair crews can always travel to address the issues arising in any power system component.

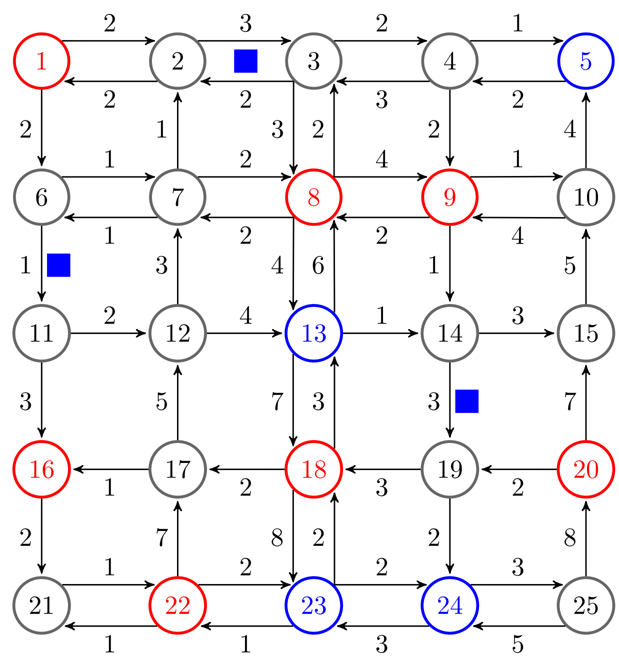


Figure 1: The network for Question 2. Note that the nodes in red are location candidates where your emergency crew can be stationed, while the nodes in blue are ones where damaged power station components can be found. If a blue node exists in the middle of a street, then you only need to "pay" half the arc time to get there from a starting node.

(a) Assume that your 2 crews are located in nodes 1 and 20. Formulate an optimization problem to evaluate the quality of the site selections, by deciding on which power systems get repaired by each crew, and the route that a crew would select to get to the damage site the fastest. Define your parameters, your variables, your constraints, and your objective function.

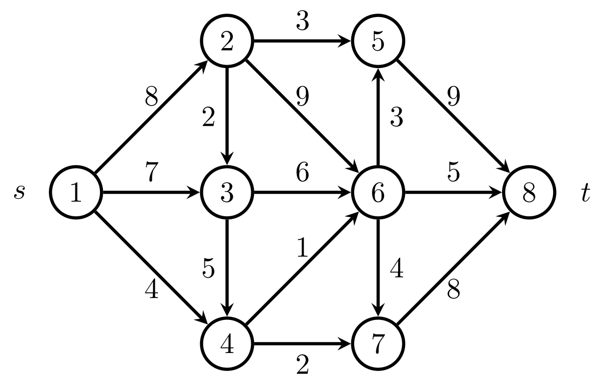
(b) Optimizing the same decisions as before, consider the case where the location crews are not known in advance, but instead are another decision variable. Formulate the problem under the assumption that the **2** crews are to be located on any node. Once more carefully define all your parameters, variables, constraints, and your objective function.

(c) Formulate and solve problem (b) in Gurobi.

Question 3: Minimum cuts

(a) Find a minimum cut of the following network shown in Figure 2.

Figure 2: The network for Question 3 part (a). All arcs are given a capacity.



(b) Propose an *efficient* algorithm (i.e., an algorithm that runs in polynomial time) that can decide whether an edge $(i, j) \in E$ belongs to **all minimum cuts** between a source s and a terminal t. Derive the computational complexity of the algorithm.

(c) Propose an *efficient* algorithm (i.e., an algorithm that runs in polynomial time) that can decide whether an edge $(i, j) \in E$ belongs to **at least one minimum cut** between a source s and a terminal t. Derive the computational complexity of the algorithm.

(d) We define a node i as being *downstream* if for all minimum cuts node i is on the side of the terminal node. On the other hand, node i is called *upstream* if for all minimum cuts node i is on the side of the source node. Finally, node i is called *central* if it is neither upstream nor downstream. Recommend an algorithm that finds all central nodes in a network. Apply that algorithm to the network of part (a) and find all central nodes in the network.

(e) *Arc connectivity* refers to the minimum number of arcs that need to be removed from a connected network in order for it to become disconnected. Recommend an approach to calculate the arc connectivity of a network in polynomial time (as a hint, consider a way to solve the problem by solving $n - 1$ maximum flow problems).

Question 4: Eulerian cycles and extensions

In class, we saw the Eulerian path and cycle problem. As a reminder, a graph is Eulerian if it has an Eulerian cycle. In an Eulerian cycle, we may start a path from a node and end in the same node after traversing every arc exactly once. Answer the following questions.

(a) Consider a simple and undirected graph $G(V, E)$. Assume you "double" every edge in the network. With "doubling" we mean that for every edge (i, j) you add a second copy of it and you now have two of each. Call this new (multi)graph $\hat{G}(V, \hat{E})$. Is $\hat{G}$ Eulerian or not? Prove it or provide a counterexample.

(b) Let us consider the following problem. A person is visiting an area to inspect all the streets (that is, they need to traverse all edges). They will park at a node and then start inspecting street by street by traversing them one at a time. They can traverse a street multiple times if they need to, but they would prefer not to. Based on your answer in part (a), devise a plan for them to find a cycle that uses at most $2|E|$ streets from and to their parking spot after having visited every edge in the network at least once.

(c) For the same problem in part (b), can you devise a polynomial time algorithm that produces a cycle from and to the parking spot that traverses every edge at least once and that is of cost at most $|E| + |V| - 1$?

Question 5: Vehicle fleet planning

A company wants to plan its fleet size for the next 5 months. At the moment, they do not own any trucks, however they will need to lease multiple ones for each of the following months. Their demands are as follows:

Month	1	2	3	4	5
Trucks needed	20	25	40	20	25

The leasing company they work with offers only two plans:

- leasing a truck for 2 months at a cost of \$1,500;
- leasing a truck for 3 months at a cost of \$2,000.

Answer the following questions:

(a) Formulate the problem of minimizing the cost of the lease while satisfying every period demand.

(b) Using the formulation of part (a), create a *minimum cost flow problem* and draw its corresponding network.

(c) Right now the company plans to do the following:

1. Lease 15 trucks for 3 months at month 1.
2. Lease 15 trucks for 2 months at month 1.
3. Lease 15 trucks for 3 months at month 3.
4. Lease 10 trucks for 2 months at month 3.
5. Lease 10 trucks for 2 months at month 5.

Verify that this is indeed a feasible solution to your original planning problem formulation of part (a) and to the minimum cost flow problem of part (b).

(d) The leasing company imposes that no lease can be made for more than 15 vehicles at a time. Moreover, the company board has indicated that they will not allow any solution where more than 20 trucks are leased but not used from month to month. With this new information (capacity constraints!), and starting from the previous solution, perform one iteration of bounded network Simplex on the network you have identified in part (b).

Question 6: An evacuation planning problem

During an evacuation, people living in each of the locations in a network are to be routed towards one or many shelter (safe) locations. Assume you are provided with an undirected transportation network $G(V, E)$ (as an example, you can use the provided network emulating the main street arteries of West Virginia). Further assume that your shelter is randomly located somewhere in the state (i.e., simply select a random node s anywhere in the network to serve as your shelter). Answer the following questions.

(a) Consider the following problem: first, identify a minimum cost spanning tree in the whole network. This is ensured to lead everyone (eventually) to the shelter. Then, calculate the average time it takes every node to get to the shelter. Report this average time using networkx and the West Virginia network provided.

(b) The minimum spanning tree is a little too strict, in the sense that all nodes are equally important. Instead, assume that a set of nodes $S \subset V$ are the most important locations in the network (e.g., they have the biggest population; they are located in especially dangerous locations; they have more vulnerable populations; etc.). Solve a Steiner tree problem where you connect the set of nodes $S \cup s$ (i.e., the important locations and the shelter). Then, identify the shortest path from each of the non-important locations to any of the nodes in the Steiner tree. Again, report the average time it takes for each location to evacuation in the West Virginia network.

(c) Formulate the optimization problem of identifying an evacuation tree of minimum cost. This would be a tree, rooted at the shelter, where every node has **only one path to safety**. Formulate the problem and solve it using Gurobi and networkx. Report the average time it takes every node to get to safety for the West Virginia network provided.