CS648A : Randomized Algorithms
Semester II, 2022-23, CSE, IIT Kanpur

Theoretical Assignment 2

Deadline : 11:55 PM, 15th March 2023.

Most Important guidelines

- It is only through the assignments that one learns the most about the algorithms and data structures. You are advised to refrain from searching for a solution on the net or from a notebook or from other fellow students. Remember - **Before cheating the instructor, you are cheating yourself**. The onus of learning from a course lies first on you. So act wisely while working on this assignment.
- Refrain from collaborating with the students of other groups. If any evidence is found that confirms copying, the penalty will be very harsh. Refer to the website at the link: https://cse.iitk.ac.in/pages/AntiCheatingPolicy.html regarding the departmental policy on cheating.

General guidelines

1. This assignment is to be done in groups of 2 students. You have to form groups on your own. You are strongly advised not to work alone.
2. The assignment consists of 3 problems. Each problem carries 50 marks.
3. **Naming the file**:
   The submission file has to be given a name that reflects the information about the type of the assignment, the number of the assignment, and the roll numbers of the 2 students of the group. If you are submitting the solution of Theoretical Assignment x, you should name the file as **Theor_x_Rollnumber1_Rollnumber2.pdf**.
4. **Each student of a group** has to upload the same submission file separately. Be careful during the submission of an assignment. Once submitted, it can not be re-submitted.
5. Deadline is strict. Make sure you upload the assignment well in time to avoid last minute rush.

How well did you internalize the proof of Chernoff Bound?

(*marks = 50*)
Consider a collection $X_1, \ldots, X_n$ of n independent geometrically distributed random variables with expected value 2. Let $X = \sum_{i=1}^{i=n} X_i$ and $\delta > 0$.

1. Derive a bound on $\mathbf{P}(X \geq (1+\delta)(2n))$ by applying the Chernoff bound to a sequence of $(1+\delta)(2n)$ fair coin tosses.

2. Directly derive a Chernoff like bound on $\mathbf{P}(X \geq (1+\delta)(2n))$ from scratch.

3. Which bound is better?

Estimating all-pairs distances exactly

(*marks = 50*)
Consider an undirected unweighted graph G on n vertices. For simplicity, assume that G is connected. We are also given a partial distance matrix M : For a pair of vertices i, j the entry $M[i, j]$ stores exact distance if i and j are separated by distance $\leq n/100$, otherwise M stores a symbol # indicating that distance between vertex i and vertex j is greater than $n/100$. Unfortunately, there are $\Theta(n^2)$ # entries in M, i.e., for $\Theta(n^2)$ pairs of vertices, the distance is not known. Design a Monte Carlo algorithm to compute exact distance matrix for G in $O(n^2 \log n)$ time. All entries of the distance matrix have to be correct with probability exceeding $1 - 1/n^2$.

Rumor Spreading Problem

(*marks= 50*)
Recall the rumor spreading problem discussed in the class. Show that everyone in the town with population n will know the rumor in $O(\log n)$ days with probability at least $1 - 1/n$.

Reducing diameter of a connected graph by adding $\Theta(n)$ edges

Note: This problem is totally optional and does not carry any marks. This problem will not be asked in any exam or quiz of the course. It is meant for only those students who might go for research in algorithms. The instructor sincerely hopes that at least one student in the entire class will solve it.

Given an undirected connected graph on n vertices. Further, each vertex of the graph adds an edge to every other vertex with probability c/n independently for $c > 0$. Show that, for some large enough constant c, the diameter of this graph is $O(\log n)$ with high probability.