

Midterm 1 - Q3

Alara Dirik

April 27, 2020

Problem

The "Community Structure in Social and Biological Networks" paper by Girvan [1] discusses an edge-weight algorithm. Suppose you are given undirected graph $G=(V, E)$ and a weight function $w : E \rightarrow \mathbb{R}$ which maps edge $(i, j) \mapsto w_{ij}$.

a) Consider the bottom up approach of constructing dendrogram as described in the paper. What kind of data structure would you use for it? Why?

A listed-link would be ideal for constructing the dendrogram described in the paper. A linked list is a linear data structure where each node contains a data field and a pointer (link) to the next node in the list. Thus, as the dendrogram is constructed with a bottom up approach, we can store the edges and then add pointers to their parent components as the construction progresses. This approach would also allow us to identify sibling communities by checking if they share a parent community efficiently.

b) What is the number of partitions of a set with N elements? No proof is required. Investigate and write your findings.)

The total number of partitions of a set with N elements is the Bell number B_N . Bell numbers satisfy the recursion $B_{n+1} = \sum_{k=0}^n \binom{n}{k} B_k$. The Bell numbers may also be computed using the Bell triangle in which the first value in each row is copied from the end of the previous row, and subsequent values are computed by adding two numbers, the number to the left and the number to the above left of the position. The Bell numbers are repeated along both sides of this triangle. The numbers within the triangle count partitions in which a given element is the largest singleton [2].

References

- [1] Michelle Girvan and M. E. J. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99 12:7821–6, 2002.
- [2] Wikipedia. Partition of a set — Wikipedia, the free encyclopedia, 2004. [Online; accessed 22-July-2004].