

# COMMUNITY DETECTION

How to split the King's Reign



## The algorithm

Girvan-Newman's Community detection is an algorithm that works on a network of elements (called *nodes*). Its aim is to identify communities among the network, by operating on the hidden relationships and connections between the nodes. It can compute a few thousands of nodes. The most common uses are in social media, to detect people with similar interests and make them connect (or keep them tightly connected).

Welcome to *Biska*, the Realm of Flying Islands. The King has reached a certain age and must divide his Kingdom for his four sons. He wants to find four regions and interrupt the bridges that connect them. To do this in the most efficient way he must interrupt the smallest number of bridges. He asked you, his Strategist, to help him. The best tool to divide the Kingdom is the *Community Detection Machine*. Learn how to split the King's inheritance.

## The tools you need

To identify the most strategic bridges on which to place checkpoints, it is crucial to understand the concept of Edge Betweenness: a value associated to each bridge, based on shortest path. Let's consider a small network to make things easier.

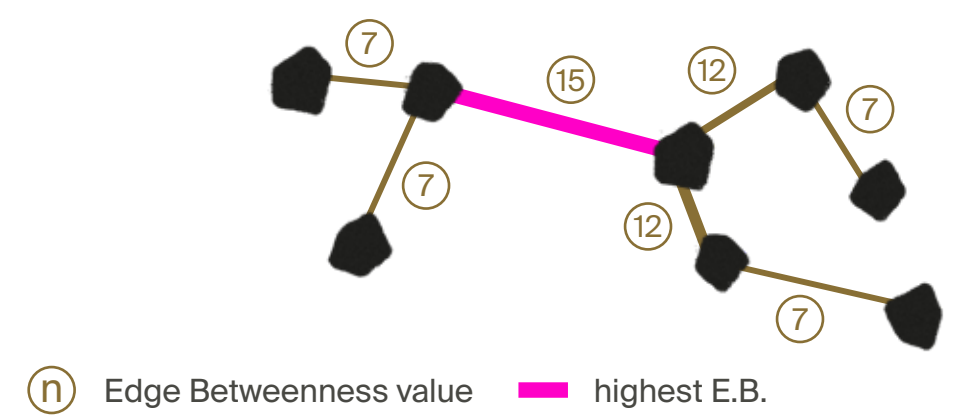
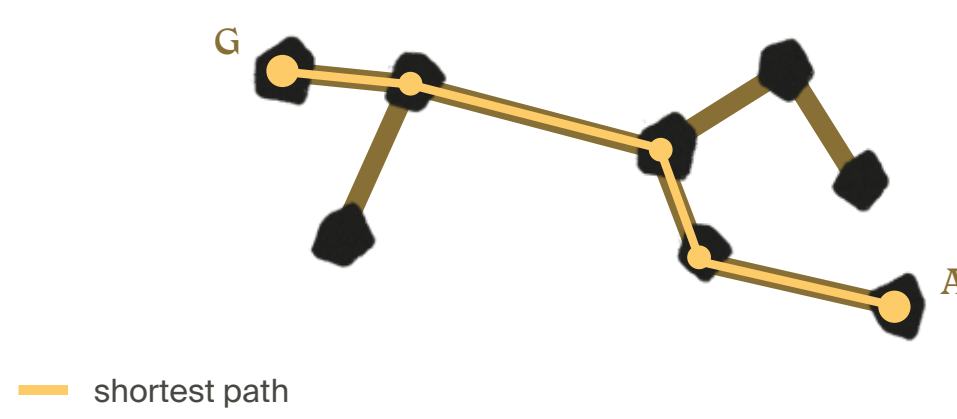
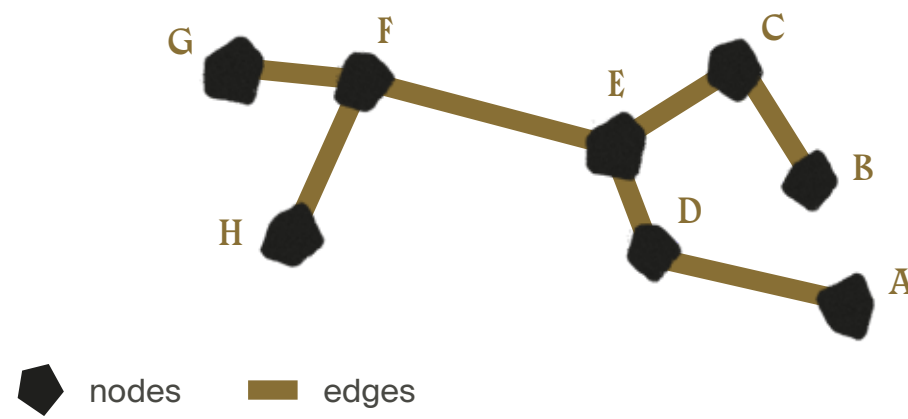
## Shortest path

To calculate the Edge Betweenness, the Machine first identifies the shortest paths between each island. It indicates how to get from G to A crossing the smallest number of bridges.

⚠ In case two paths cross the same number of nodes, the actual length of the bridges connecting them is calculated to find out the shortest one.

## Edge Betweenness

Edge Betweenness (E.B.) is then computed, as a value obtained by summing all the shortest paths passing through a given bridge. The bridge with the highest value is the most strategic one.



## The Machine at work

### I. Find the strategic bridge

After having computed E.B. of the entire network, the first action is to choose the most strategic bridge: the edge with the highest E.B.

### II. Interrupt the strategic bridge

The algorithm now removes the edge with the highest E.B. since it is the most strategic on nodes communications.

### III. Split the Kingdom

After interrupting the strategic bridge, the network has been split in different groups highlighted in violet.

## community detection machine

There are four steps repeating themselves in a circular way. The Machine goes on until there are no more bridges, but you can decide which configuration of the Kingdom suits the King's desires.

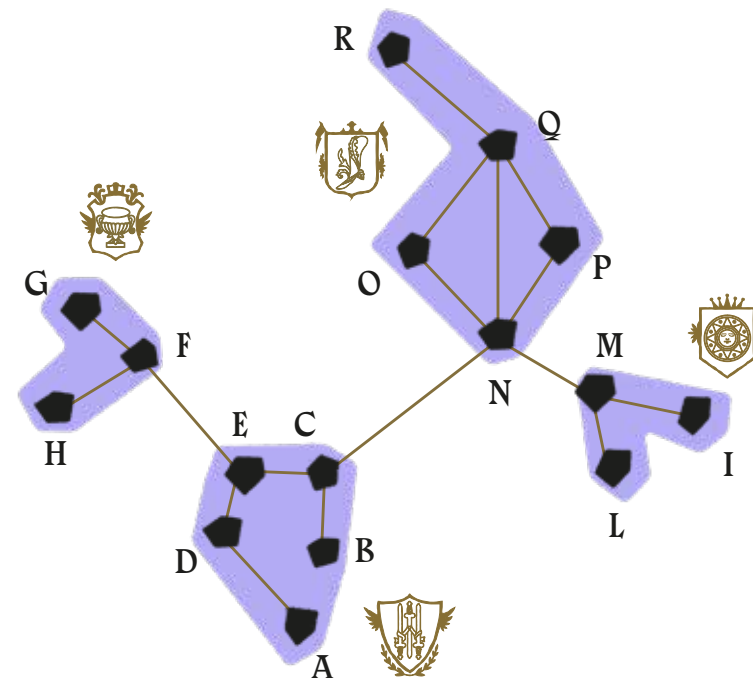
### IV. Spin to recompute

The last action is to recompute E.B. for the new configuration of the network. Some bridges will be missing due to the fact they have been interrupted.

## Results

Aiming to useful results, it is necessary to decide what step of the algorithm is meaningful to consider. The bridges cut in the iterative process will be those connecting larger or smaller communities identified within the starting network.

The most widely used output of the algorithm is the dendrogram (a), a tree graph used to represent the process of grouping elements, not only based on spatial position but also on other kinds of relations. In this case, the King decided to stop at the second step of the process, as it was the one that generated four regions with the least number of interrupted bridges.



a - dendrogram

