Applications of minimum spanning trees

Short list[1]

- Building a connected network. There are scenarios where we have a limited set of possible routes, and we want to select a subset that will make our network (e.g., electrical grid, computer network) fully connected at the lowest cost.
- Clustering. If you want to cluster a bunch of points into k clusters, then one approach is to compute a minimum spanning tree and then drop the k-1 most expensive edges of the MST. This separates the MST into a forest with k connected components; each component is a cluster. (I confess I'm not very clear on whether anyone uses this clustering method in practice, and if so, what domains it is useful in.)
- Traveling salesman problem. There's a straightforward way to use the MST to get a 2-approximation to the optimal solution to the traveling salesman problem, and the Christofides' heuristic uses MSTs to get a 1.5-approximation. (One could reasonably question how real-world this is, though, as there are other approximation algorithms for the traveling salesman problem that will typically do even better in practice.)

1 Network design

Minimum spanning trees have direct applications in the design of networks, including computer networks, telecommunications networks, transportation networks, water supply networks, and electrical grids. [2]

One example would be a telecommunications company laying cable to a new neighborhood. If it is constrained to bury the cable only along certain paths (e.g. along roads), then there would be a graph representing which points are connected by those paths. Some of those paths might be more expensive, because they are longer, or require the cable to be buried deeper; these paths would be represented by edges with larger weights. A *spanning tree* for that graph would be a subset of those paths that has no cycles but still connects to every house; there might be several spanning trees possible. A *minimum spanning tree* would be one with the lowest total cost, thus would represent the least expensive path for laying the cable.

2 Muddy city problem

Once upon a time there was a city that had no roads. Getting around the city was particularly difficult after rainstorms because the ground became very muddycars got stuck in the mud and people got their boots dirty. The mayor of the city decided that some of the streets must be

[1]http://cs.stackexchange.com/questions/29305/applications-of-min-spanning-trees

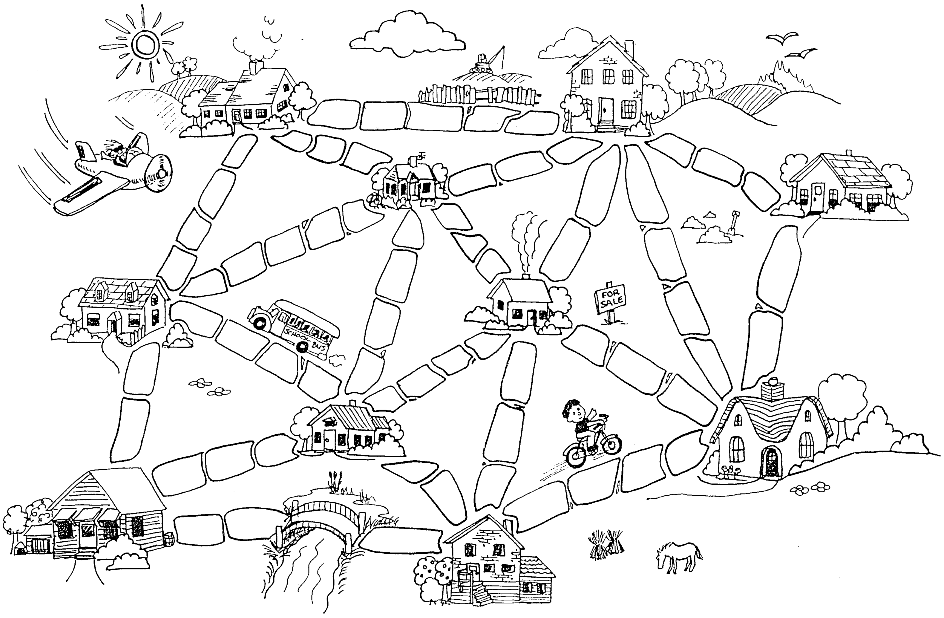
[2]https://www.quora.com/What-is-a-real-time-practical-application-of-the-Minimum-Spanning-Tree-MST

paved, but didnt want to spend more money than necessary because the city also wanted to build a swimming pool. The mayor therefore specified two conditions: [3]

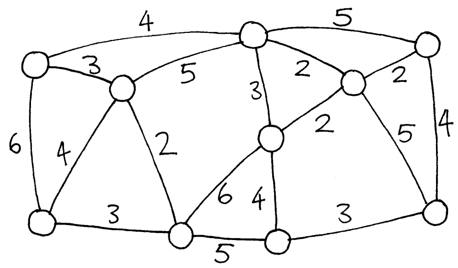
1. Enough streets must be paved so that it is possible for everyone to travel from their house to anyone elses house only along paved roads, and

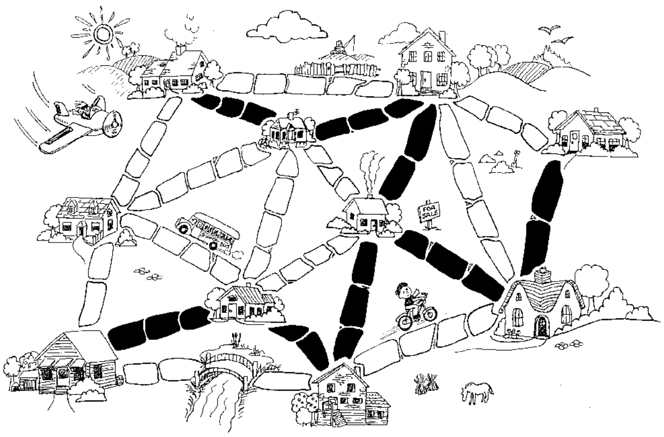
2. The paving should cost as little as possible.

Here is the layout of the city. The number of paving stones between each house represents the cost of paving that route. Find the best route that connects all the houses, but uses as few counters (paving stones) as possible.



Solution: the graph (for another muddy city) and the paving.





3 Other practical applications

Other practical applications based on minimal spanning trees include: [4]

- Taxonomy.
- Cluster analysis: clustering points in the plane, single-linkage clustering, graph-theoretic clustering, and clustering gene expression data.

[3]http://computing2school.com/category/computer-science-unplugged-2/part-ii-algorithms/lesson-9-minimal-spanning-trees

[4]https://www.quora.com/What-is-a-real-time-practical-application-of-the-Minimum-Spanning-Tree-MST

- Constructing trees for broadcasting in computer networks. On Ethernet networks this is accomplished by means of the Spanning tree protocol.
- Image registration and segmentation.
- Curvilinear feature extraction in computer vision.
- Handwriting recognition of mathematical expressions.
- Circuit design: implementing efficient multiple constant multiplications, as used in finite impulse response filters.
- Regionalization of socio-geographic areas, the grouping of areas into homogeneous, contiguous regions.
- Comparing ecotoxicology data.
- Topological observability in power systems.
- Measuring homogeneity of two-dimensional materials.
- Minimax process control.