Optimal network modification to limit outbreak size

a novel application of treewidth

Jessica Enright[1] and Kitty Meeks[2]
[1] University of Stirling
[2] University of Glasgow

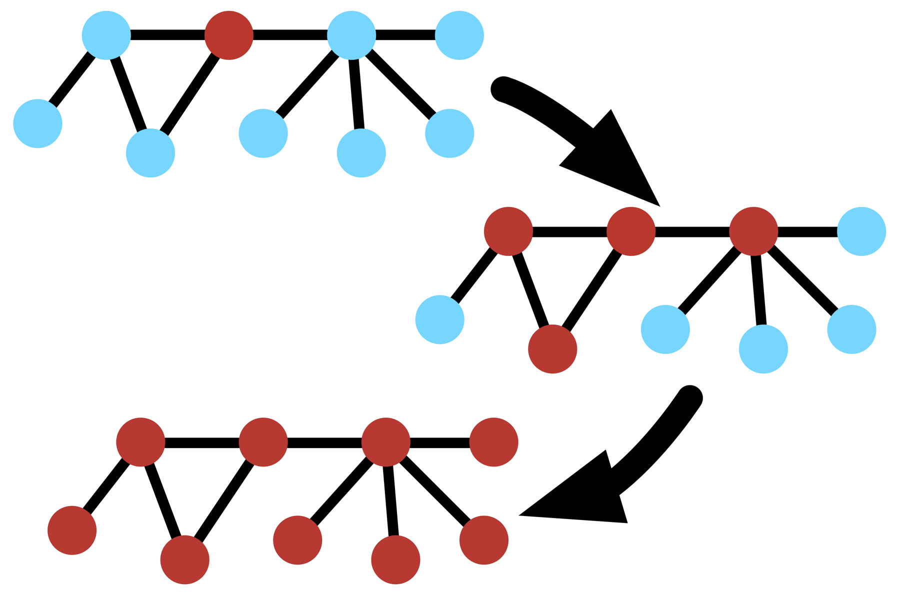
Contagion on graphs

A **graph** (or network) captures pairwise contacts between entities.

We can represent our infectable agents as the **vertices** of the graph, and the potentially infectious contacts between them as the **edges** of the graph. Then contagion introduced to a vertex can spread from vertex to vertex over the edges.

Graph models are most appropriate when we know explicit contacts, or have a good individual-level contact model.

The **outbreak** spreads from node to node over the edges.



Optimal network modification

Making the **minimum cost changes** to a network to give the **largest difference in contagion** spread is of interest to both disease controllers and mathematicians. Related problems are studied as network reliability measures and combinatorial optimisation problems by computer scientists, mathematicians, and engineers.

We want to explore methods that are mathematically provably optimal, computationally feasible, and appropriate for data-derived networks

Our goal is to incorporate some of the well-developed machinery of graph theory into epidemiologically useful tools, benefitting both epidemiologists and mathematicians.

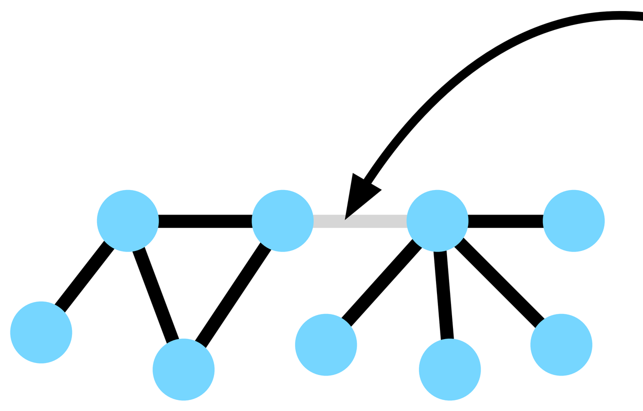
Limiting component size

We have started with optimal edge deletion to limit maximum component size.

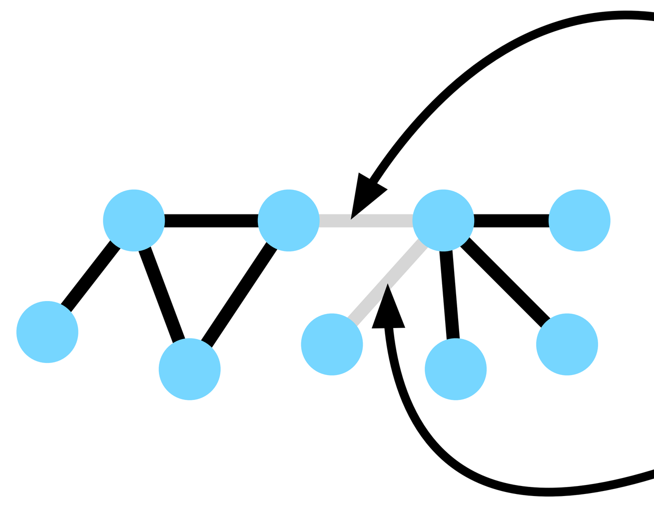
A **component** of a graph is a joined-up set of vertices. The size of the largest component in a graph is an upper bound on the size of the largest possible outbreak on that graph.

We have focussed on **edge removal** because, in the setting of listed trading partnerships between cattle holdings in Scotland, a limitation on the link is far more feasible than a limitation on the farm.

As this problem is intractable in general, we have developed an algorithm applicable to graphs limited by treewidth, a classic graph parameter.



We can delete **1** edge to limit the maximum outbreak size to **5** infected nodes



We can delete **2** edges to limit the maximum outbreak size to **4** infected nodes

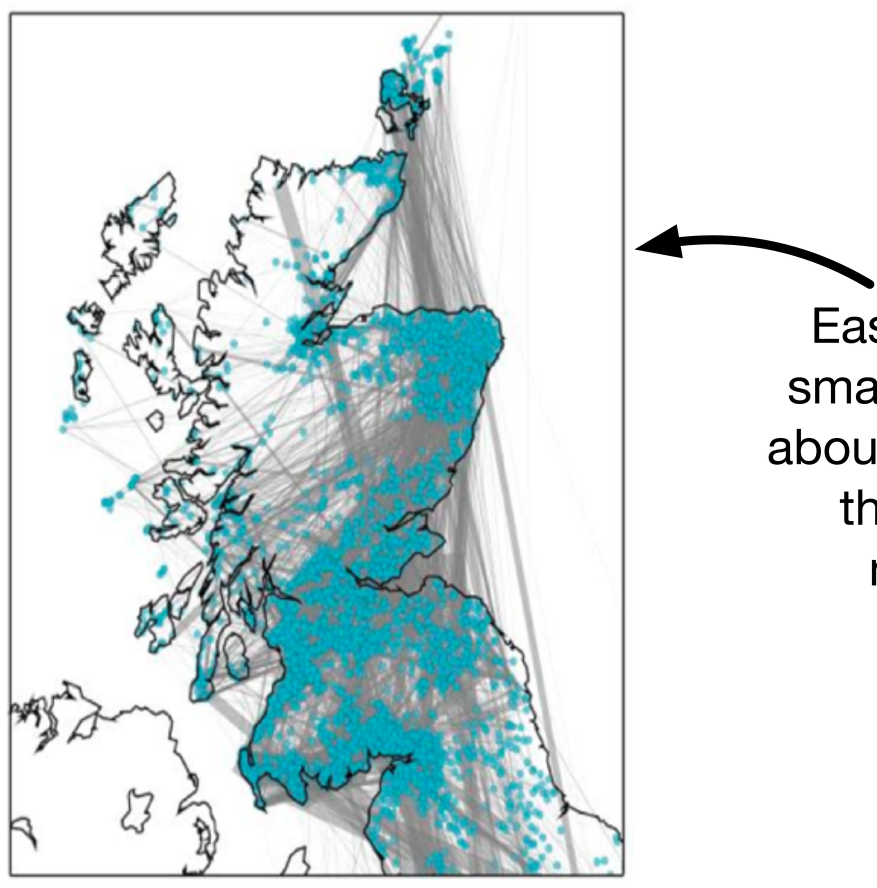
Treewidth

When a problem is intractable on graphs in general, a popular approach is to limit the graph by some parameter. **Treewidth** is a very popular graph parameter.

When a graph has **limited treewidth**, we can produce a mapping of the nodes of the graph to the nodes of a tree with only a few nodes of the graph at each node of the tree. Given this representation, many hard problems can be **solved efficiently**.

Our algorithm uses standard techniques for processing representations of limited treewidth.

We have found that large cattle trading networks (as well as some other data-derived networks) have low treewidth. This means that many efficient graph algorithms will work on these networks to solve problems that are usually intractable.



Easy enough with a small graph, but what about a large graph like the cattle trading network here?

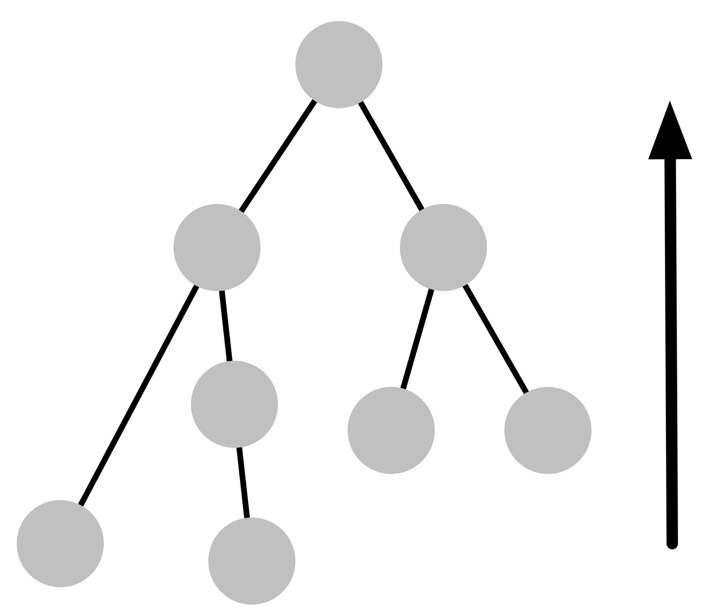
How does the algorithm work?

We map the graph onto a tree decomposition, and then work from the bottom up.

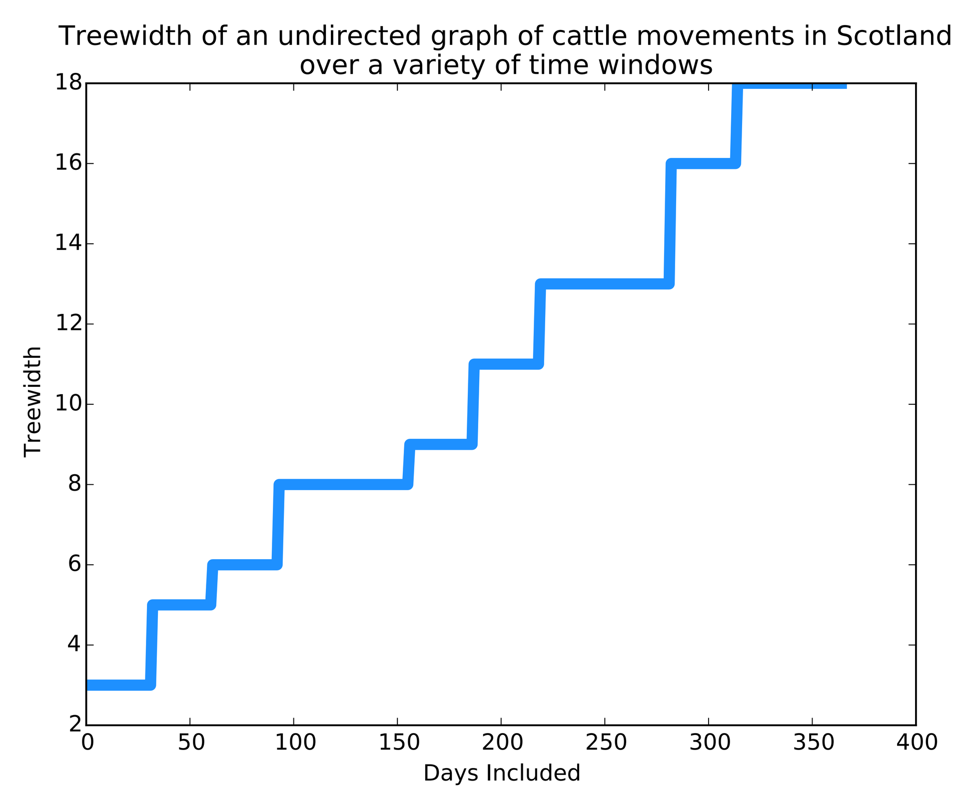
This algorithm takes time bounded by a function of the form:

$$O((wh)^{2^w} n)$$

treewidth (w), max component size (h), number of nodes (n)



Even when we include 200 days of cattle movements in Scotland, the graph that we make from those movements has treewidth less than 10, so we can feasibly use our algorithm!



The authors gratefully acknowledge funding and data provision from the Scottish Government as part of EPIC: Scotland's Centre of Expertise on Animal Disease Outbreaks.

Want to learn more? Email jae@maths.stir.ac.uk

arXiv:1504.05773