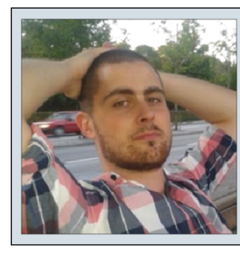
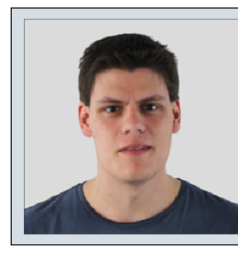
MADALGO CENTER FOR MASSIVE DATA ALGORITHMICS


Edvin Berglin
Aarhus University

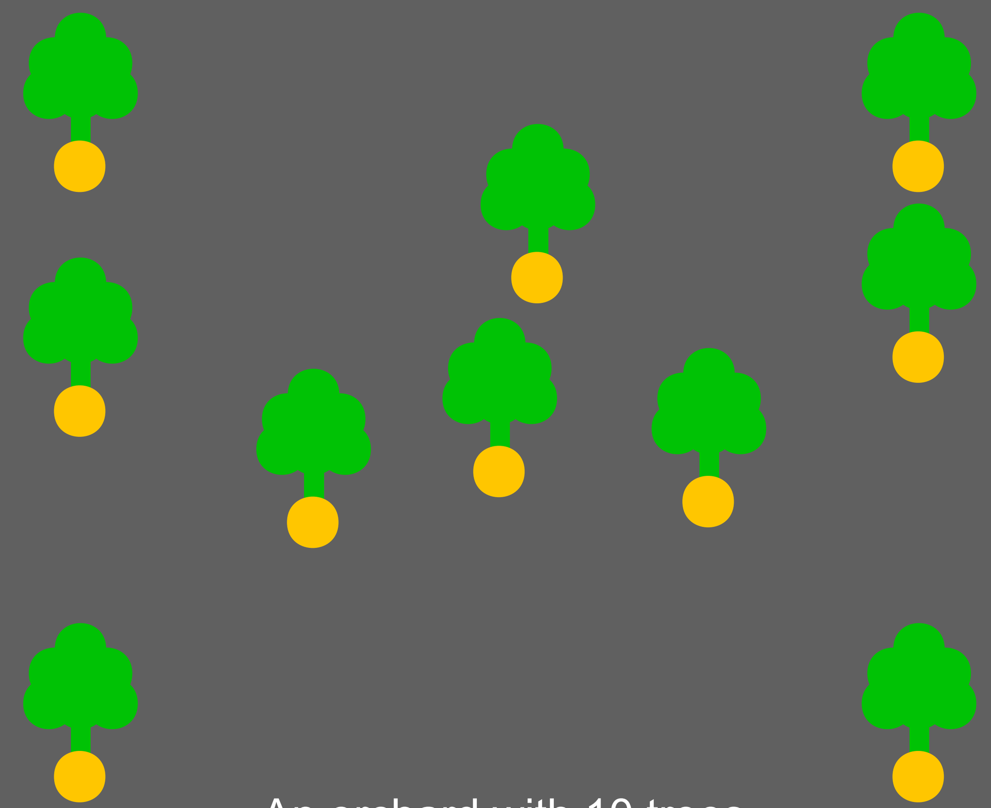

Ingo van Duijn
Aarhus University

Danmarks Grundforskningsfond
Danish National Research Foundation

Covering Points with Lines

Felling Trees and Covering Lines

You are given an orchard with trees and a machine that can efficiently fell trees when moving in a straight line. What is the minimal number of directions you need to move your machine in such that you can clear the whole orchard?



An orchard with 10 trees.

In the above orchard there are many possibilities of positioning the machine such that it can fell three trees. However, it is impossible to clear four or more trees in a single direction. We can therefore argue that in this case we cannot clear the orchard with only three directions.

In the problem *Line Cover*, the above puzzle is stated as follows:

Given a set of n points in the plane, is it possible to place k lines such that all points are covered by at least one line?

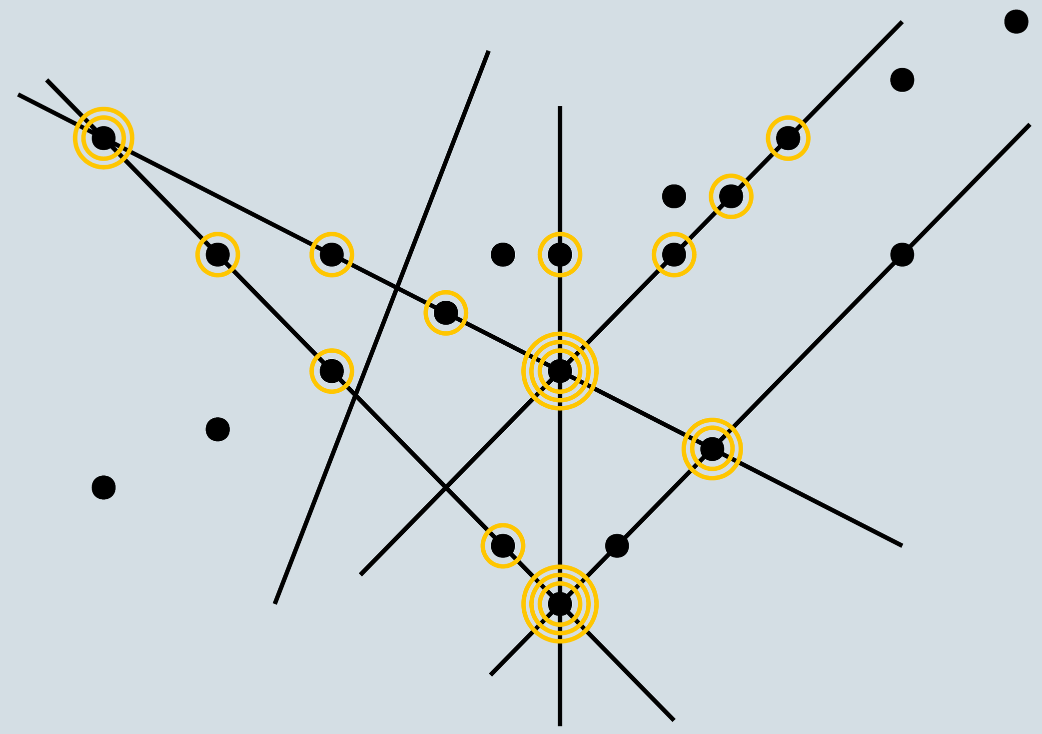
This problem is NP-hard and it is therefore likely that the best way of solving it is not much different than trying all solutions in a clever way. To understand what might be clever we observe he following:

- If the point set is random we do not expect any three points lying on one line, which means k must be at least $\frac{n}{2}$ in order to cover the points.
- If k is small compared to n and it is possible to cover the points, many of the points must be collinear.
- $k+1$ collinear points can only be covered by k lines if a single line covers all the points.

The above indicates that a good chance of finding the solution quickly is to try lines covering many points first. In fact, the last observation implies that we can **greedily** pick a line covering more than $k+1$ points. Since remaining lines can cover at most k points, the remaining instance (called the **kernel**) is only solvable if no more than k^2 points remain. This is exactly the reason why the above puzzle is unsolvable for $k=3$.

Incidences

An **incidence** is a point-line pair such that the point lies on the line. If a point set allows many ways of placing a line such that it cover many points, intuition tells that this point set contains many incidences.



An arrangement of 21 points and 6 lines with 19 incidences

The famous Szemerédi-Trotter theorem [1] upper bounds the number of incidences in a set of n points and m lines:

$$\#incidences = O((nm)^{\frac{2}{3}} + n + m)$$

A corollary of this theorem states that there are *few* ways of placing a line covering *many* points. These **rich** lines are the starting point of our algorithm; by trying rich lines first we hope to reduce the size of the problem as quickly as possible. The key observation for our algorithm is that if the solution has few (or no) rich lines, then we **implicitly** reduce the input size by essentially the same argument as the k^2 kernel.

Algorithm

The algorithm is in essence very simple:

- Given the points, determine all possible lines covering at least 2 points.
- Sort all those lines by how many points they cover.
- Starting at the richest line and continue in descending order, recurse in two branches:
  - add the line to a partial solution
  - discard the line from being in the final solution
- When the partial solution has size k, verify if all points are covered.

The problem with the above algorithm is that the deeper it recurses, the less rich the considered lines become. The incidence bound gives that there are few lines covering many points, but there might be many lines covering few points.

The solution is to switch to a $O^*(2^n)$ algorithm as soon as few enough points remain to make it an efficient switch.

New Results

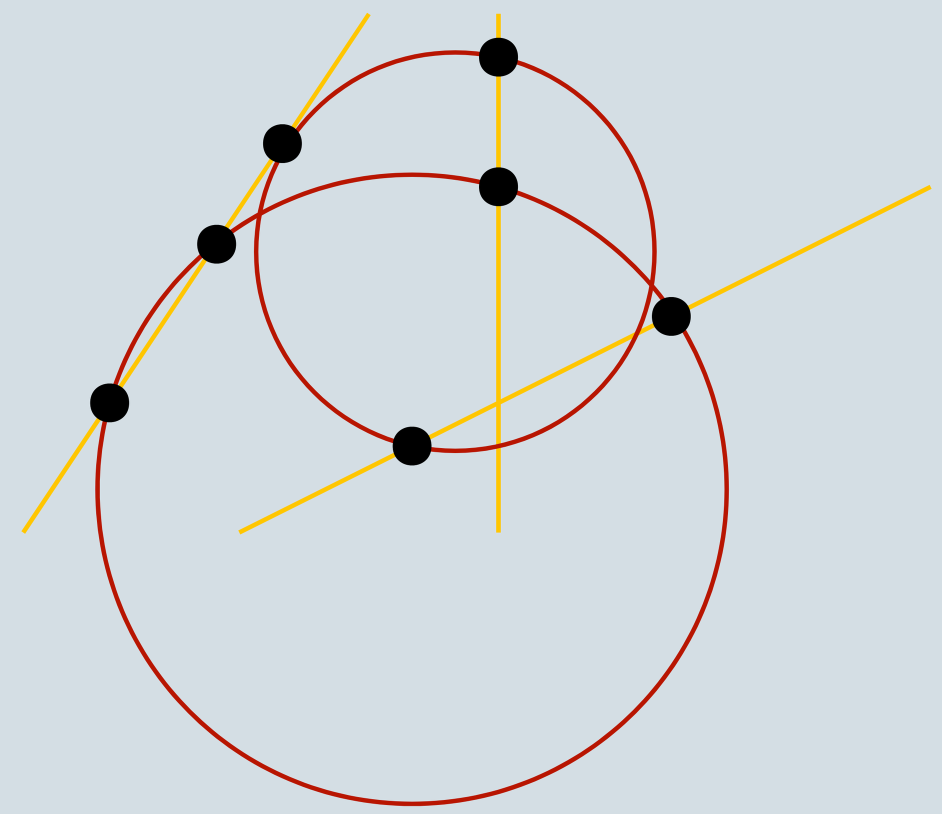
Our main result is a more refined analysis of the simple recursive algorithm. In particular, we use incidence bounds to prove a limit on the initial branching. When the incidence bound becomes too weak, we switch to a base case algorithm that is unaffected to the number of incidences.

- Previous best: $O\left(\left(\frac{k}{1.35}\right)^k\right)$
- Our result: $O\left(\left(\frac{k}{\log k}\right)^k\right)$

Previous work has also focused on a wide variety of covering problems where the covering object is something other than a line. For covering objects for which incidence bounds similar to the Szemerédi-Trotter theorem exist, we have successfully generalised our method.

Our generalisations:

- *Curve Cover:* wide variety of d-dimensional curves
  - Based on incidence bound from [2]
- *Plane Cover:* covering points with planes in 3 dimensions
  - Based on incidence bound from [3]



Points can be covered by different objects other than lines.
This point set requires fewer circles than lines.

References

[1] E. Szemerédi and W. T. Trotter Jr. *Extremal problems in discrete geometry*. Combinatorica, 1983.

[2] J. Pach and M. Sharir. *On the number of incidences between points and curves*. Combinatorics, Probability and Computing. 1998.

[3] G. Elekes and C. D. Tóth. *Incidences of not-too-degenerate hyperplanes*. Proc. 21st Annual Symposium on Computational Geometry. 2005.