

MSc project ideas - to be supervised by Keith Briggs 2009

[home](#)

- [papers](#)
- [thesis](#)
- [talks](#)
- [meetings](#)
- [records](#)
- [maths notes](#) «
- [software](#)
- [languages](#)
- [music](#)
- [travel](#)
- [cv](#)
- [students](#)
- [maps](#)
- [place-names](#)
- [people](#)
- [photos](#)
- [train data](#)
- [links](#)
- [site map](#)

I work at BT Research in Martlesham. The ideal arrangement is for you, the MSc student, to work here for all or most of your project. We are at a very lively technology park, and you would have at least daily meetings with me and very good computing facilities (i.e. linux). But if this is not feasible, we can negotiate another arrangement.

These are only rough sketches for projects. I find that it is better to start this way, and then fine-tune the detail to suit the student's abilities and interests. I prefer to leave some open-endedness, to provide a research challenge. A good MSc project can result in a publication.

I like mathematical computing. Most of these projects have that theme. The emphasis is on the concrete. This way of working often involves *experimental mathematics*, in which we treat the computer as a laboratory for developing and testing conjectures. This need not mean the work lacks rigour - some algorithms on which we would work would involve only integer arithmetic, and thus the results are exact. Even when floating-point arithmetic is needed, there are more reliable ways than hardware arithmetic available. So, these projects do not generally involve conventional numerical analysis.

I can probably only supervise 2 (maybe 3) of this projects at one time. So if you are interested, contact me asap!

For past projects I have supervised, see [here](#).

0. Quantifying TV whitespace for cognitive radio operation

(To be jointly supervised by Maziar Nekovee.)

Cognitive radio is being intensively studied for opportunistic access to the so-called TV whitespace: large portions of the VHF/UHF TV bands will become available after the digital switchover. Using accurate digital TV (DTV) coverage maps together with a database of DTV transmitters, we would like to develop an accurate methodology for identifying TVWS frequencies at any given location in the United Kingdom.

A preliminary study has already proved the worth of this approach; we wish to take it further and make the results much more precise. The project largely involves computational geometry: such ideas as convex hulls and point-in-polygon tests. These ideas are well worth knowing about for other applications. This is a project with direct applicability to a currently hot topic in telecoms, as well as using fundamental geometric algorithms that need to be well-designed to make the whole computation feasible.

1. The maths of MIMO

MIMO refers to multiple antenna systems currently being investigated for use in wireless data networks. There is a lot of nice maths used in this topic, especially for the decoding - maximum likelihood estimation, singular-value decomposition, etc.

I think a good project would be to study the basics and write an exposition emphasizing the mathematics that the method depends upon. There is plenty of engineering expertise in my research lab which we could draw on. An open-ended research question would be to what extent exact maximum likelihood estimation can be used, or is its computational complexity too high?

2. SAT

The SAT (satisfiability) problem is a very basic problem in logic: to determine whether a given Boolean function has an input which makes the output true. There are many good algorithms now available to solve SAT problems, and many practical problems are equivalent to, and can be translated into SAT.

This project would make a comparison of available solvers, and study their application to some practical problems in optimization, scheduling, and graph coloring. The fun part would be writing programs to translate these into SAT!

3. Maximum-likelihood period estimation

This project would be to study the paper (only 4 pages!):

Maximum-likelihood period estimation from sparse, noisy timing data, by Robby McKilliam and I. Vaughan L. Clarkson

and write a program to implement the method. This is an important problem area with applications in telecoms.

4. Statistics of train arrival delays

A previous project looked at statistical models in the delay to departure time of trains. An interesting result was obtained - that the delays follow a q -exponential distribution. This project would do the same thing for arrival delays. NB: if you are interested in this project, please tell me asap, because we should start collecting data well before the project starts.

5. Chromatic number of geometric random graphs

Geometric random graphs are models of wireless communication networks. The chromatic number tells us how many frequency channels are needed to operate the network without interference.

I am currently doing some work on how the chromatic number of purely random graphs (without the geometric structure) depends on the density (number of links per node). I find some very interesting results - namely, if the graph is big enough, the chromatic number is almost surely determined. This project would be to examine the analogue of this behaviour for geometric random graphs. It will not be exactly true, but might be in some appropriate limiting cases. There is already quite a lot of work in this area which would have to be taken into account.

6. Semidefinite programming and graph theory

Semidefinite programming (SDP) is a kind of generalized linear programming. Many practical optimization problems can be formulated as SDPs. There has been rapid progress in the last 20 years on understanding the theory of SDP, but the development of SDP software that is convenient to use has lagged behind. Good software is now available. This project would use `sdpsol` for small test problems, and then move to `DSDP` for large problems. The aim of the project is to see what performance one can achieve on real graph theory problems of the types that come up in network modelling. Such problems include the Lovasz θ number (a lower bound for the chromatic number), the maximal stable set problem, and the maximum cut problem. An important outcome of the project would be a determination of how the computation time scales with problem size. Geometric optimization problems could also be studied.

Skill in C programming is needed. This would be a good project for someone wanting a career in practical optimization.

7. Fast random selection

Consider the problem: a large number N of objects are presented to us one by one, and we wish to either:

- Select exactly some specified number $n \leq N$ of them
- Select each with some specified probability p

The problem becomes hard to solve efficiently when p is small, because to generate a random number which usually results in a rejection is inefficient. It would be better to skip over a block of items. Ways to do these were proposed by Vitter in *ACM Trans Math Soft* 13, 58 (1987).

This project would investigate these methods and check their efficiency in practice. There are applications to the generation of large random graphs.

8. Random sampling of set partitions

A partition of $[n]=\{0,1,\dots,n-1\}$ is an assignment each element to a subset called a block, without regard to the labelling of the block, or order of elements within a block. The block structure of a partition may be indicated by listing the elements and separating blocks with a symbol such as $|$. Thus for $[5]$, $01|2|34$ and $2|10|43$ are the same partition.

If there are k blocks, we speak of a k -partition. The number of k -partitions of $[n]$ is given by the Stirling number of the second kind $S(n,k)$, which can be computed from the recursions $S(n,k)=S(n-1,k-1) + kS(n-1,k)$, $S(n,1)=S(n,n)=1$. The total number of partitions of $[n]$ is given by the sum over k of $S(n,k)$, and this is called a Bell number $B(n)$.

There are algorithms for generating all partitions in a Gray code order, due to Ehrlich and Ruskey. In such a sequence, only one element moves to a new block on each step. Beyond about $n=15$, there are too many partitions to allow us to construct all of them in reasonable time.

Thus, for large sets, simulations must rely on a uniform random sampling procedure. This project would investigate such methods and produce a well-designed C library for use in other projects.

9. Fast counting

This project is inspired by the paper [HyperLoglog: the analysis of a near-optimal cardinality estimation algorithm](#) of Flajolet et al. How do we estimate (rapidly, and with minimal storage) the number of distinct items in a very large set (or data stream)? In other words, the problem concerns estimating the cardinality of a multiset. Flajolet et al. previously developed their *loglog* algorithm, and I worked on an efficient C implementation of this. They have now improved this with the *hyperloglog* algorithm, and this project would be to implement this and compare its performance in practice with alternatives.

The algorithm should have practical applications in informatics; for example, counting the number of different packet types in a network.

This page was last modified 2008 Nov 18 09:52 by keith.briggs@bt.com