Networks

Simple is the best for dynamic routing of telecommunications

from Alistair Mees

TELECOMMUNICATIONS networks have to choose routes dynamically to make the best use of lines available at each moment. If the most direct link for an attempted conversation or data packet is busy, an exchange tries to find a more roundabout way to get the message through. Complex and ingenious schemes have been proposed for doing this automatically. New work shows that a surprisingly simple scheme is not only very good but comes near to making the best use of transmission capacity that could be achieved by any possible scheme. Frank Kelly and Richard Gibbens, in collaboration with British Telecom researchers, studied the absolute bounds on capacity determined by the nature of the telecommunications network, and it was while thinking about these problems that they arrived at the new method.

To get the bounds they showed that although network traffic is stochastic — calls do not arrive in an exactly predictable way — the capacity problem takes little heed of this. When a major trunk link is nearly saturated it might be carrying 1,000 calls, and the standard deviation of the call distribution will then be about 30, which is relatively small. By solving a deterministic problem of multicommodity flow, Kelly and Gibbens could obtain upper bounds on capacity. This particular problem has special structure, and is not unmanageably large, so it can be solved by linear programming: the difficulties with general multicommodity flow problems (see my recent News and Views article in *Nature* **321**, 19; 1986) do not arise here.

The real question is how to get near to these upper bounds. Imagine a trunk network. There will be 50 or so trunk exchanges in the new UK digital network, dealing with long-distance communications traffic in much the same way the national railway services handle long-distance passenger traffic. The trunk network will be fully connected: every exchange is (effectively) connected by its own lines to every other. This means that a call from, say, Manchester to Edinburgh will normally use the direct Manchester–Edinburgh lines. If they are all busy, the Manchester exchange might try to route via Glasgow (or via Southampton, for that matter). If there is no two-link route of this kind, a route with more links might be tried, although there are good reasons why this game quickly becomes self-defeating. An n-link call is a single call using as much capacity as n calls would,

and perhaps causing $n-1$ calls to fail later because there is no route available. As the network became very busy, more and more calls would be routed in this wasteful manner, using a lot of capacity for only a small increase (or even a decrease) in throughput. In other words, n-link routing is unstable.

One-link (direct) routing often gives well below the theoretical network capacity: this is obvious, and is the reason dynamic routing is considered in the first place. Kelly and Gibbens decided to try
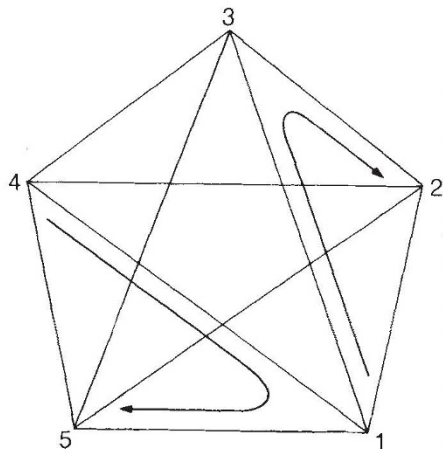
A model of dynamic routing. Calls from 1 to 2 overflow via 3; those from 4 to 5 overflow via 1.

two-link routing in an attempt to get a significant capacity increase while avoiding instability. The problem was to choose which exchange to act as the middleman in a two-link route when calls are blocked because a direct route is busy. One scheme would be to give every exchange a fixed partner that would be asked to try to forward blocked calls; this solution improves capacity but requires careful traffic analysis and in any case needs frequent reprogramming because traffic patterns change over time. It is also inflexible during, for example, breakdowns and unexpectedly high traffic at the partner exchange.

A very simple way around this problem is the following. An exchange puts calls through directly as long as there are lines available. If all direct lines are busy, it chooses a partner exchange at random and asks it to forward the call. If the chosen partner succeeds, it is chosen the next time a call needs to be forwarded, and keeps getting chosen until it fails. ("All lines from Manchester are engaged. Please try later".) Then a partner is chosen at random again, and so on. It can be seen at

once that this scheme will explore the network until it finds a friendly partner, without any fancy pre-analysis of traffic. A few (very few) calls will be lost, but they would have been lost anyway without the dynamic routing. In the long run, the Manchester exchange will spend most of its time with the friendliest partner. It will adapt itself completely automatically in the face of breakdowns or variations in traffic.

This 'sticky random' selection is a neat trick, but there is another difficulty with this scheme (or indeed with any scheme that allows indirect routing): unselfishness towards one's friends is good up to a point, but sometimes it has to be every exchange for itself! If the Glasgow exchange is as kind about forwarding calls from Manchester as it is about putting through calls from its own subscribers, it will be risking instability even in a system where two-link calls are the worst that are allowed. Somehow, we need to penalize two-link calls, at least when the lines are very busy. One way is to have a trunk-reservation parameter k. If there are at least k free lines on a given link, an exchange forwards calls out along that link without question. If there are fewer than k free lines, it refuses to forward calls at all along that link.

Trunk reservation and sticky random-partner selection turn out to work together in a synergistic way. With sensible values of k, Kelly and Gibbens claim they can get so close to the bound on what is achievable that there is no reason even to bother with other schemes unless they have advantages quite other than improved capacity. Perhaps surprisingly, performance is very insensitive to the exact value of k. As long as k is not equal to zero (total unselfishness, causing instability) or comparable with the channel capacity (complete selfishness, causing inefficiency) the system performs well. For links of a few score up to many thousands of lines, taking k anywhere between 5 and 10 is just fine.

British Telecom, Kelly and Gibbens have applied for a patent for the scheme, under the name Dynamic Alternative Routing. It is possible that the new electronic 'System X' exchanges will be given the capacity to use it, and if it proves as effective in practice as it has been in theory and in simulation, it is likely to become the preferred routing method, for its robust simplicity as much as because of its near optimality. □

Alistair Mees is Professor of Modern Applied Mathematics at the University of Western Australia, Nedlands, Western Australia 6009. The work is awaiting patent and is not yet published. Interested readers are asked to contact Dr Frank Kelly or Mr Richard Gibbens at the Department of Pure Mathematics and Mathematical Statistics, 16 Mill Lane, Cambridge CB2 1SB, UK.