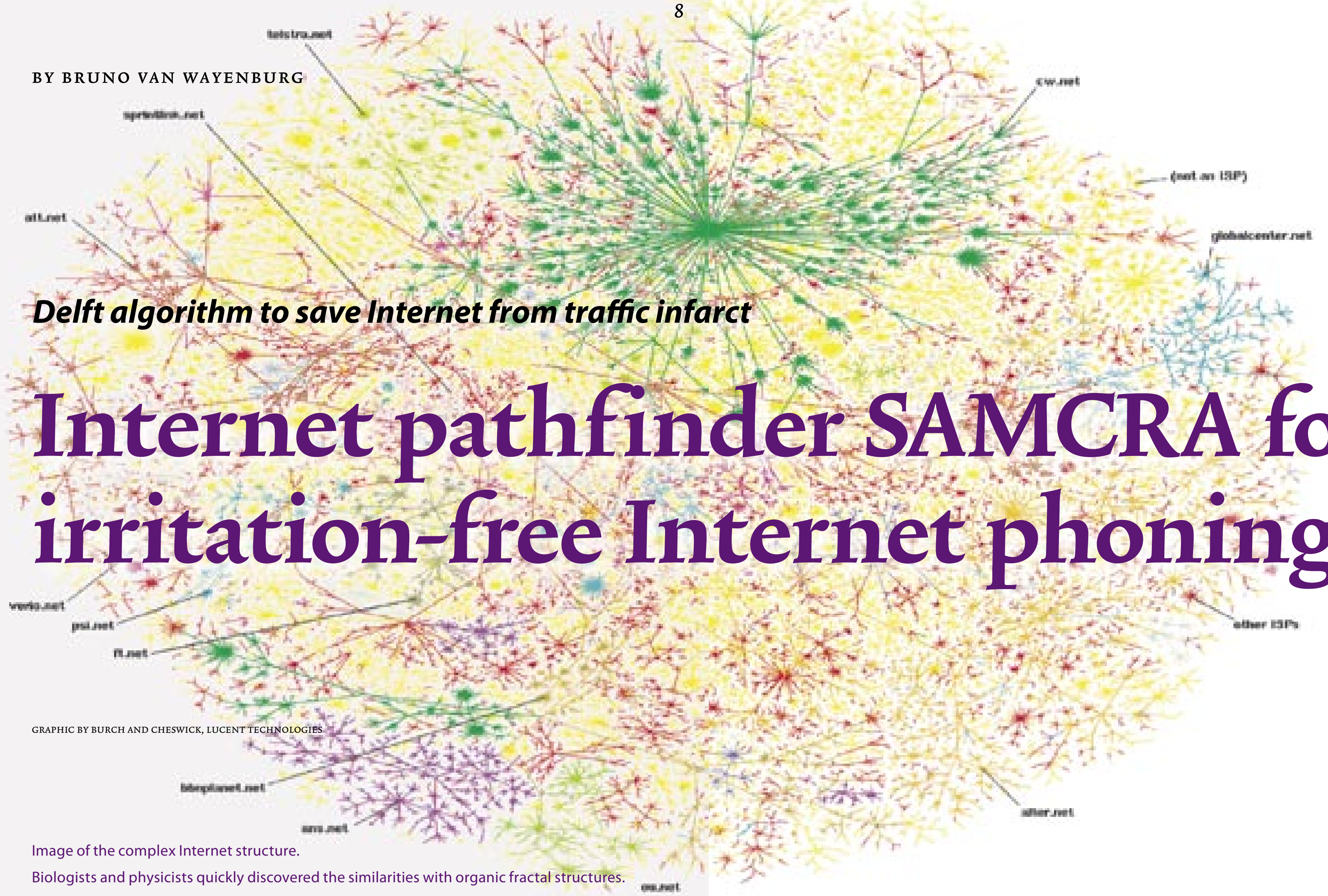
BY BRUNO VAN WAYENBURG

Delft algorithm to save Internet from traffic infarct

Internet pathfinder SAMCRA for irritation-free Internet phoning



GRAPHIC BY BURCH AND CHESWICK, LUCENT TECHNOLOGIES

Image of the complex Internet structure.
Biologists and physicists quickly discovered the similarities with organic fractal structures.

The growth of the Internet is going to grind to a halt unless fundamental changes are made in its organisation, predicts Piet Van Mieghem, Professor of Telecommunication Networks at TU Delft. Routing, the information highway's traffic control system, is already a cause for concern. And, as in the case of road congestion, simply creating additional capacity will not provide relief.
Van Mieghem's solution: simply offer better Internet connections for sale in order to guarantee the quality of the connection for Internet phoning or video conferencing and ensure the service proceeds without interruption. Internet service providers could then really start making money from the Internet.
But this calls first of all for smart routing software which is guaranteed to find the best route through the network, preferably without having to compute for days. With the routing algorithm SAMCRA, Van Mieghem and his Network Architectures & Services (NAS) research group at Delft University of Technology have found a prospective solution that, theoretically, seems to have the best qualifications. A comprehensive simulation and test programme recently carried out together with American researchers demonstrated that the computation method also satisfied the high requirements under practical conditions.

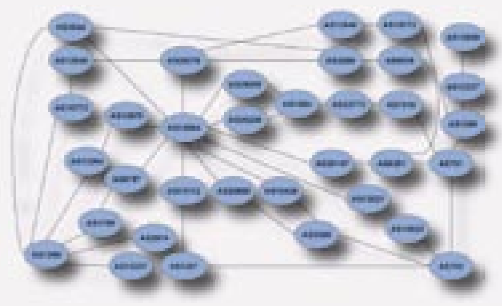
Professor Piet Van Mieghem is convinced of one thing: Internet has to do much better! True, you can already listen to the radio through the World Wide Web, you can telephone and you can surf anywhere you want. And e-mails almost always reach their destination.
'However, no provider will guarantee that you can telephone flawlessly across the Internet or that you can hold an undisrupted video conference', according to Van Mieghem. 'There are simply no guarantees of quality.'
Segments of sounds or images may disappear, you may have to wait too long for an answer or the line just fails no matter how broadband your connection is. Yet the 'old-fashioned' alternative, the telephone, has been operating for decades with virtually none of those problems.
According to Van Mieghem, the design of the Internet needs a major overhaul. Providers will have to start offering quality guarantees for Internet connections and clients will have to start paying for them. Moreover, the control of the World Wide Web will become significantly more complex.

Secondary role The root of the problem is that the Internet was never built to transmit data within a particular time span.
'Basically, it was purely a data network in which time was initially a secondary concern', explains Van Mieghem.
Any file that is transmitted over the Internet, whether it is an e-mail, a webpage or an audio signal, is split up into packets, each of which is sent individually into the network. That network is an enormous, relatively unorganised system of interconnected computers. Each packet follows its own route before they are all neatly aligned again at the destination: the e-mail, the web page or the sound segment has arrived. How long it takes depends on the travelling time of the last packet, a fairly randomly determined parameter.
Moreover, data can also be lost, sometimes because of bad connections but mainly because the routers, busy junctions in the network, become temporarily overloaded and throw away incoming packets. These losses can be rectified by a system of confirmation of receipt and resending but that causes additional delays. Experts refer to this approach as "connectionless". Not because there is no connection between sender and recipient but because no fixed connection is reserved for a single data flow. Every packet finds its own way, almost haphazardly. The advantage is that the network is not particularly hierarchical and can easily cope with local malfunctions (the data simply take a detour).
In addition, the network can easily be extended without central control.
The counterpart of a connectionless network is the "connection-oriented" approach, exemplified for instance by the worldwide telephone network, which is composed of a strictly hierarchical pattern of local, national and international exchanges and connections. For a single telephone call to the United States a connection route is reserved for the duration of the conversation. Everything is designed to transport the data flow as flawlessly as possible, without losses and preferably as quickly as possible, because even a delay of more than a tenth of a second is a nuisance in a conversation.
According to Van Mieghem and many Internet developers like him, what is actually required is a sort of hybrid between the two extremes. This would involve a monitoring system that gives certain packets a higher priority, like registered mail and reserves bandwidth for it at a surcharge for the user.

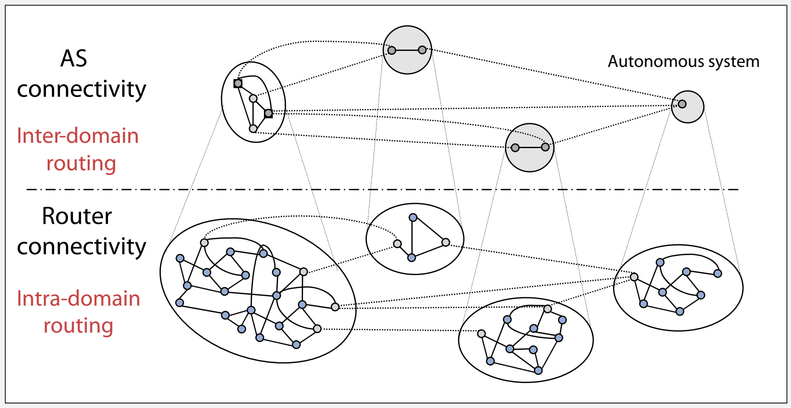
"Quality of Service" (QoS) is the somewhat meaningless technical term that network specialists use for this approach of price and quality differentiation. As a matter of fact there was once a QoS-based data communication protocol, the "connection-oriented" ATM protocol for networks, which was developed by the International Telecommunication Union (ITU) standardization sector and ATM Forum, the foundation for broadcasting networks, in the late eighties and early nineties.
'It never took off, partly due to the Internet hype', says Van Mieghem with regret.

Problem child However, what is needed now is to improve the Internet machinery, says Van Mieghem. With the current state-of-the-art the international computer network will not be able to grow much further.
'The approach to Internet extensions has always been that if you just increase capacity everything will be all right', the professor says. 'That is basically how the Internet was built. But that won't work any longer.'
Routing, the switching of data flows over the Internet through busy exchanges or routers, is gradually starting to become a problem child. The Internet comprises some fifteen thousand subnetworks of major Internet providers, or 'Autonomous Systems' (AS). To find the best routes within their own subnetworks the providers use the OSPF (Open Shortest Path First) protocol, which is based on a computation method developed in 1959 by the Dutch mathematician Edsger Dijkstra (1930-2002). The major advantage of this 'Dijkstra algorithm' is its lightning speed, even in large networks. As the number of routers in the network increases the computation time to find the optimum path only increases a little more than proportionally.
'That is really awesome', according to Van Mieghem.
At a higher level, however, in powerful routers that make up the network backbone, the 'Border Gateway Protocol' is active. It does not determine the route using 'Dijkstra' but on the basis of arrangements made between Internet providers. Consequently, it does not necessarily result in the shortest route.
'Border Gateway Protocol is a really complex system with all kinds of updates, attenuators and other routing tricks to guide traffic properly', Van Mieghem says. 'Routing specialists are already discovering that BGP cannot be extended much further.'
The complexity and inefficiency are already causing major problems, for instance in the event of a BGP router failure. An alternative solution is a much more extensive type of monitoring, with which it is also possible at the highest level to obtain an overview of the entire network and to plan the best routes.
A QoS system with price and quality differentiation offers that possibility, according to Van Mieghem, because the system already needs an overview of the network for the division of the capacity among data packets with higher and lower privileges. In 1998, after much debate this led the Internet standards organisation IETF (Internet Engineering Task Force) to take over the 'Quality of Service' idea, which was then already ten years old, and this boosted research and interest in QoS systems. Experimental versions of such systems, for instance with twelve different priority classes, are being tested by dozens of research groups.

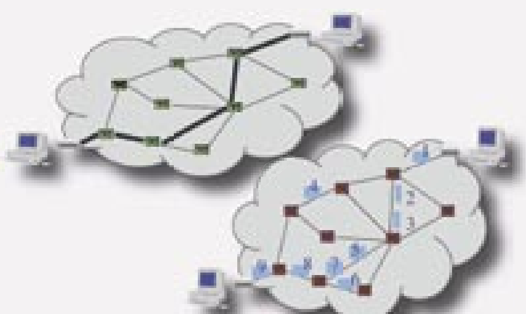
Infamous class Yet making the switch over is not so easy. From a computational point of view QoS routing is much more complicated than regular routing. The Dijkstra algorithm has just one criterion per subpath between two routers: the delay experienced by the data on the way. For efficient QoS routing, by contrast, several criteria have to be considered, including bandwidth, delay and the percentage of packets lost.
By contrast with 'Dijkstra', searching for a route that fully satisfies several criteria is an 'NP-complete problem', according to Dr Fernando Kuipers, who was until recently a doctoral student in Van Mieghem's group. In other words, QoS routing belongs to an infamous class of computation problems for which in some cases the computation time increases exponentially with the scale of the problem, in this case the size of the network. Quite simply, for some network structures every possible route must be computed, and with every node the number of possible paths may increase drastically by the number of nodes in the network. A network with several dozens of nodes may easily have billions times billions of possible routes and it will take even the most powerful computers days or weeks to compute them all. Such waiting periods are no option for the Internet.



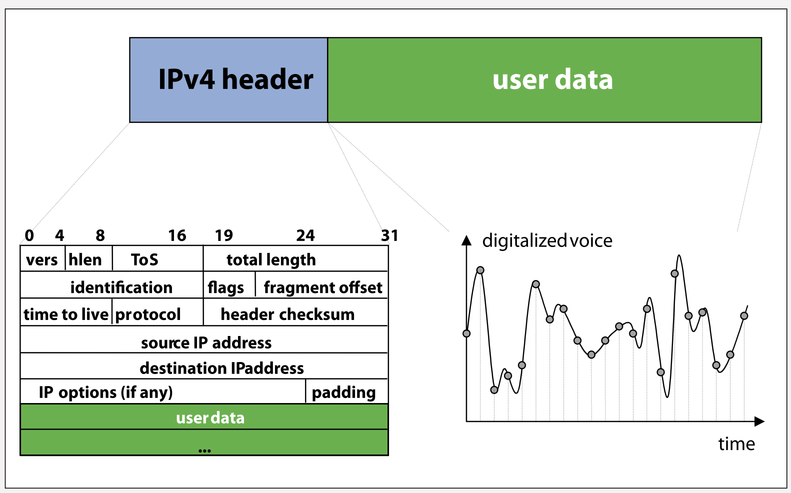
Schematic overview of a small part of the global Internet. Each of these AS (Autonomous System) ovals represents a network of an Internet service provider (ISP) as seen from the Amsterdam Internet Exchange. The Internet comprises some 15,000 autonomous systems all over the world.



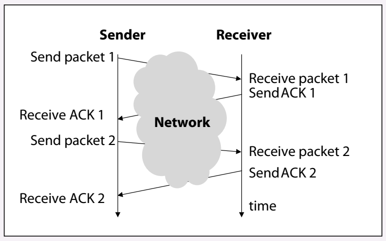
The Internet actually consists of two hierarchical layers. At the top is the AS layer used for traffic between Internet providers (known as inter-domain routing). Below it is the IP layer, where communication between users of the same Internet provider takes place (known as intra-domain routing).



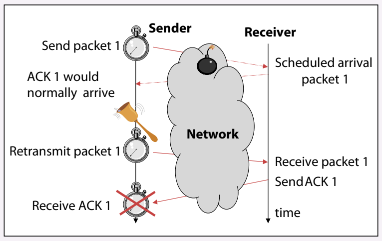
There are two ways to get a packet from A to B. Telephone is an example of the 'connection-oriented approach', which is jargon for the situation where all packets of a conversation follow the same path from A to B. 'Connectionless' transmission basically means that every packet of a data flow may follow a different path from A to B. The Internet is a classic example of 'connectionless' transmission



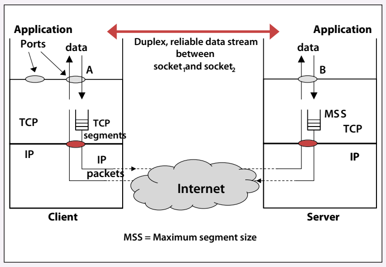
All information transmitted via the Internet is digitised. The data is sent as a data packet. A control section is placed in front containing among other things the IP address of sender and recipient, the total length of the packet and the protocol used to generate the data (the information).



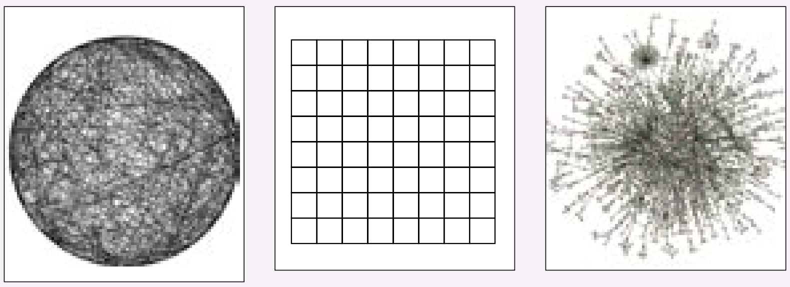
When data packets are transmitted over the 'connectionless' Internet some packets may not reach their destination. To ensure that IP packets that have not arrived are still sent TCP (Transmission Control Protocol) uses a technique in which the receipt of the packet is confirmed with an 'ACK' message.



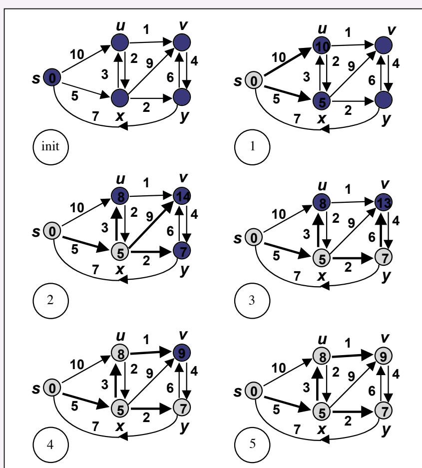
When source and recipient decide to communicate through the TCP protocol the source expects a confirmation (ACK message) for every data packet sent. However, the source will not wait forever for such an ACK. If the packet is lost in the network the source will wait for a certain period. If it has not received any message within the anticipated period the lost packet is resent. This process is repeated until a confirmation has been received.



Communication between two e-mail programs on different computers over the Internet follows a layered structure. The top layer is known as the application layer where the programs generate information. On the layer below it TCP divides the information into segments. The size of the segments depends on the available capacity of the connection A-B. The TCP segments are sent as IP packets over the Internet, whereupon TCP regroups them into a single correct information flow at the recipient's end.



There are various types of network topologies. Two extreme types are the random topology and the lattice topology. The lattice topologies are highly regular, while their counterpart, the random topologies, are irregular. The topology of the Internet probably lies somewhere between these two extremes. The third topology, which besides for instance the protein structure probably also includes the Internet, is referred to by physicists as 'scale-free'.



Dijkstra(G,w,s)
G: graph
$w(i,j)$: weight link from node i to j
s : source
$d[i]$: sum of link weights from s to i

Initialize-Single-Source(G,s): $d[i] = 0$
$S = \{\}$
$Q = \{V\}$
while Q is not empty
  do i = Extract-Min[Q]
    $S = S \cup \{i\}$
    for each neighbour j of i
      do if $d[j] > d[i] + w(i,j)$
        then
          $d[j] = d[i] + w(i,j)$
          predecessor[j] = i

The Dijkstra algorithm finds the shortest paths from the starting point (s) to all other nodes in the network. The distance from s to all other nodes d[i] is initially infinite, except for s itself: d[s]= 0. Each time Dijkstra chooses the node i for which d[i] is smallest and checks whether it can improve the path to the neighbours of i by going through i. Once a node has been chosen it is never chosen a second time. So when all nodes have been chosen 'Dijkstra' stops and the shortest paths have been found.

'This NP completeness tends to deter researchers', Kuipers says. 'Their reasoning is that if you cannot solve it exactly easily, you just have to make a stab at it. That leads to 'heuristics', computation methods based on reasonable sounding rules of thumb. For example, preferably look for short routes through the network even though a detour may cause less delay. This means that heuristic methods do not guarantee that the optimum path is found, but at least the computation time remains limited (even though that cannot always be guaranteed either).'
Van Mieghem: 'There is currently a boom in publications in which people compare their heuristic with a different heuristic which happens to be a little better or a little worse.'
Van Mieghem himself, who was until 1998 employed by telecom company Alcatel in Antwerp, initially also followed that strategy and developed the heuristic algorithm TAMCRA, which stands for 'Tunable Accuracy Multiple Constraints Routing Algorithm'. Unfortunately, Alcatel decided not to develop the idea further at the time.'

SAMCRA Van Mieghem then started in Delft where he and his doctoral student Kuipers decided to have another shot at finding a non-heuristic, exact solution. This resulted in a new variant: SAMCRA (Self Adapting Multiple Constraints Routing Algorithm). SAMCRA is an algorithm that does guarantee it will find the path that best meets with all criteria (or reports with certainty that there is no such path). Van Mieghem acknowledges that in theory this can take a very long time.
'But NP-complete means explosive computation time in the worst case.'
In practice, there were never any problems of extremely long computation times with realistic network structures, the Delft researchers noticed. To substantiate that observation, Kuipers carried out a comprehensive test programme in which he tested SAMCRA for a large variety of network structure types. Eventually he found four conditions that all have to be met for NP-complete behaviour to emerge. Most of these conditions were not very realistic and are unlikely to occur in a regular computer network. One such condition was a clearly negative correlation between, for instance, delay and the rate of packet loss. However, connections between two routers in fact tend to lose more packets when delays increase, Kuipers explains. The researchers considered it highly unrealistic that the four conditions would coincide.
'The beauty of it is that you can conclude that realistic networks can still be computed properly with an exact algorithm', Van Mieghem states. 'We have developed more algorithms, but SAMCRA is our thoroughbred.'
Together with researchers at the University of Arizona and the University of Texas, Austin, they compared SAMCRA with a large number of heuristic algorithms by simulating a series of realistic networks on a computer.
'In every case SAMCRA outperformed the others', Kuipers reports.
The two cannot however prove that SAMCRA will never behave as NP-complete in realistic networks, Van Mieghem admits. 'If you could do that, you would not be far from winning a major prize in mathematics.'
Kuipers recently obtained a doctorate *cum laude* for his work on the algorithm, which the network researchers are trying to disseminate among Internet and network experts.
'The code is on our site', says Van Mieghem invitingly. 'Another nice aspect', he suggests, 'is that SAMCRA can also be applied to other routing problems. Route planning for vehicles, for instance: while Dijkstra's algorithm can simply find the shortest route, SAMCRA could also take into account delays caused by congestion or traffic lights or, if you want, the number of restaurants along the road.'
Kuipers was once approached by a robotics researcher who wanted to use SAMCRA for an independent robot. The robot worked on the basis of information from various cameras. When integrating the various video images problems similar to those of routing cropped up. Other applications, for instance for logistic problems or when searching large databases, are also conceivable.
Van Mieghem: 'I welcome any suggestions'.

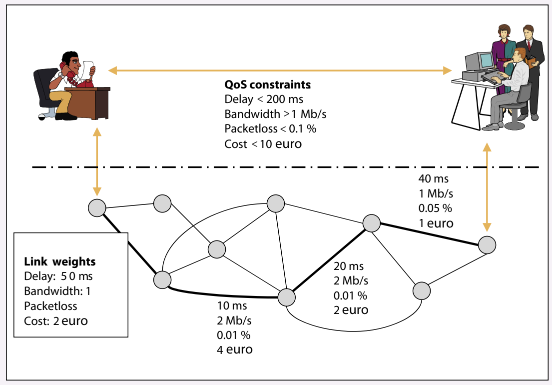
Business model 'Broad application of SAMCRA and the quality differentiation of QoS would immediately help Internet providers to improve their 'business model', the professor says. 'Today most Internet providers charge a flat fee for which you can surf as much as you want (up to a certain limit).

They sell the bandwidth under an obligation to provide best efforts, not with a quality guarantee.' But to keep the revenues coming in once everyone is on the Internet and there is no longer any market growth you are going to need price differentiation, Van Mieghem predicts.

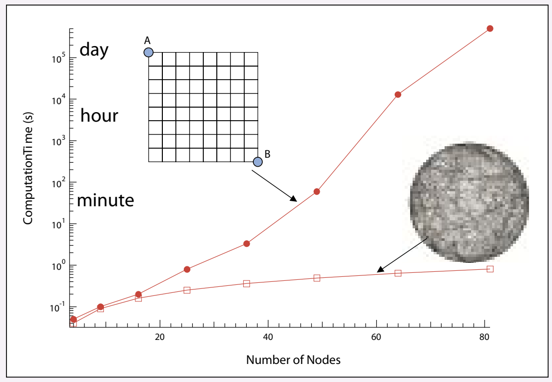
'In America there have been studies into what types of infrastructures are profitable in the long run', Van Mieghem explains. 'They showed that the postal mail service and railroads, for instance, barely pay their way. Only the telephone network, in this case America's AT&T, generated a return after sixty years.' According to Van Mieghem, apart from the monopoly position enjoyed by the telephone companies this was also due to the 'connection-oriented' organisation of the telephone network where you pay more for more extensive facilities (international phone calls are more expensive than local ones) but get guaranteed quality for your money.

SAMCRA could be a first step in that direction for the Internet too, Van Mieghem thinks, on the way to the ultimate goal of the 'Holy Grail'. 'When we have reached that point you will be able to get the quality of connection you need for any application at any time and anywhere and at a reasonable cost.' 'Who needs all those quality gradations and guarantees? Banks that want guarantees that their money transactions are safe', the professor suggests, 'but also companies that want to hold video conferences or to telephone over the Internet. I know plenty of people who say: Voice over IP (internet telephony - editor) works fine for me. But what if 20 million people want to use it rather than 200,000 as at present, and not only in the evening? Then of course there are other applications that you and I cannot even think of yet. E-mailing complete movie libraries, real-time graphic applications in 3D, who can say? History has shown that if you offer more bandwidth people will think of something to fill it.'

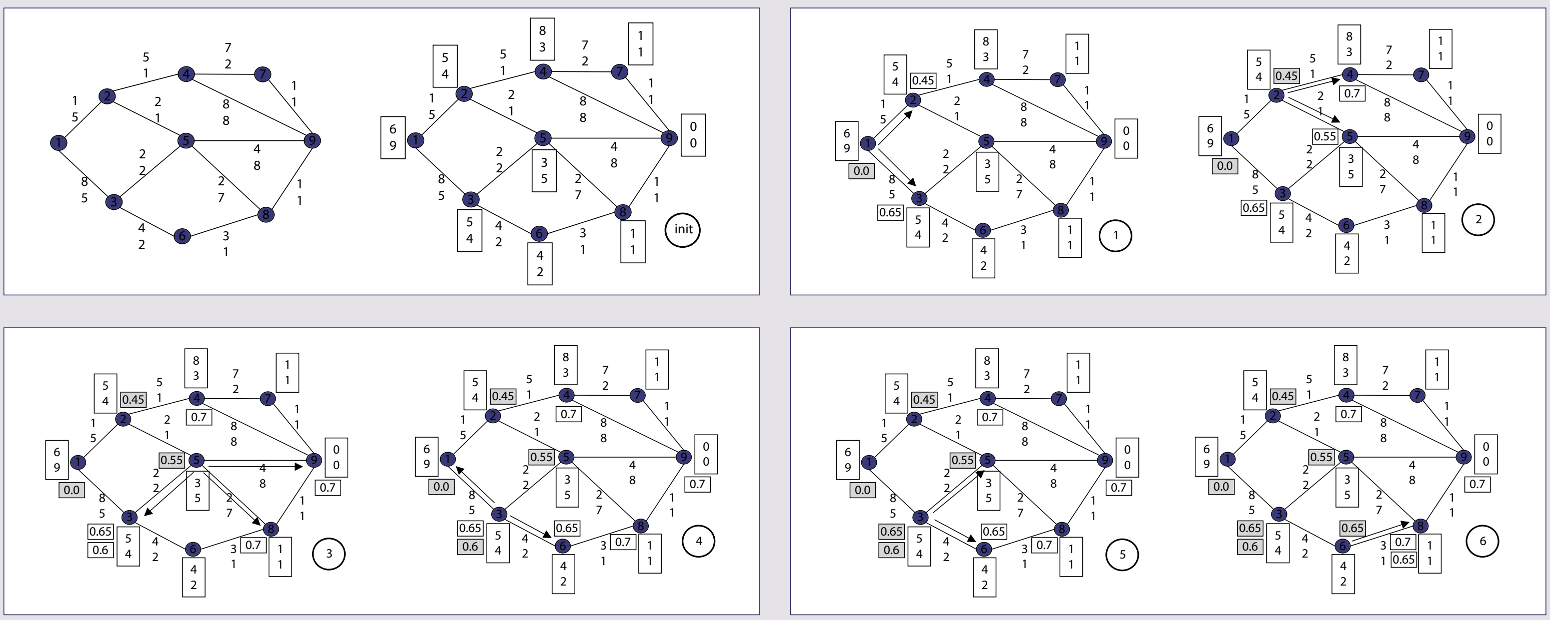
For more information on this subject, please contact Piet Van Mieghem, phone (015) 278 2397, e-mail p.vanmieghem@ewi.tudelft.nl or
Fernando Kuipers, phone (015)278 1347, e-mail f.a.kuipers@ewi.tudelft.nl

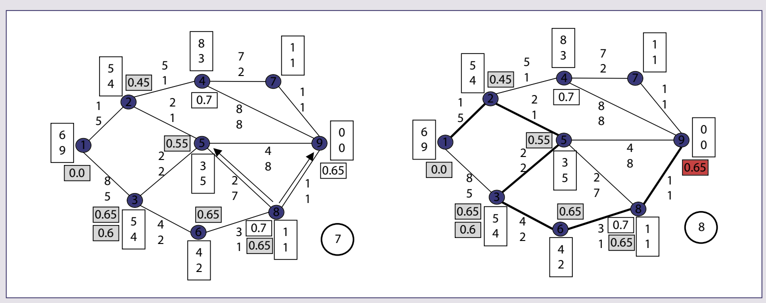


The QoS routing problem involves finding paths that comply with several requirements (e.g. delays, bandwidth, probability of packet loss and costs). For instance, VoIP (internet phoning) can only perform properly if the delay is lower than 200 ms, the bandwidth is not less than 110 Kb/s, the packet loss probability is lower than 1% and costs are as low as possible. The network may be composed of several connections (glass, copper, satellite) each with its own characteristics. The quality properties of such a path are determined by the weights of the various links within that path. The trick with QoS routing is to find the path that remains within the predefined QoS requirements.



NP-complete problems are problems for which in the worst case the number of computation steps and the computation time can very easily get out of hand. The QoS routing problem is such a problem, but fortunately the computation time appears to be very short in practice.





An illustration of SAMCRA. The network is characterised by two parameters: delay and costs. The example shows a search for the path between node 1 and node 9 that complies with the requirements, in this case set at 20,20. That means the delay must be less than 20 units and the costs must not exceed € 20. In the initialisation stage the Dijkstra algorithm is used to find the lower limit for the distance from each node to the destination. These lower limits bounds are shown in the rectangles. SAMCRA (just like Dijkstra) then starts to choose the path with the shortest length and extends this path to the neighbours under certain conditions. SAMCRA's length function is non-linear and it is shown in the small rectangles. Once a path has been chosen it is coloured grey. Step 3 shows that various paths can be established for a node. SAMCRA stops (step 8) when the path with the shortest distance to the destination has been found.