MEASURE – a Summary

Simon Crosby

Multiservice networks and multimedia operating systems have a common problem: resource allocation. Present resource management schemes in communications networks and operating systems are based on a static model of how the users of the resources behave. Multimedia communication and processing, however, are not static. They are rapidly superseded by better versions, and both user control and the content of the media being processed can drastically change the behaviour of an application during its lifetime. Moreover, such applications require minimal Quality of Service (QoS) guarantees of resource availability over a range of timescales, to function properly.

We propose to investigate the technique of dynamic, on-line measurement and estimation of resource demand, rather than static modelling, to provide information to resource management systems. We will apply this technique to three distinct but related domains: public ATM networks, local area ATM networks and the operating system in end-user computers. The goal is to develop measurement techniques and statistical estimators, and to demonstrate their utility in call admission and network dimensioning in ATM networks, and application resource negotiation in a multimedia operating system. Furthermore, we aim to provide a theoretical unification of resource management across the three domains. If our techniques can be used both in end systems and networks, then there is potential to tackle a serious problem facing distributed multimedia systems, namely the mapping of QoS attributes across the end system/network boundary.

The project will approach the problem of measurement and resource allocation in the three domains, using equipment developed at Telia Research in a public ATM network scenario, equipment developed at Cambridge in a local area ATM network, and a multimedia operating system developed by the Pegasus ESPRIT project.

The theoretical foundation of this work is large deviation theory, which enables us to derive, for a given system, a *rate function* which characterises the probability of rare events occurring. The conventional approach to derive the rate function requires the traffic process to be statically characterised, by fitting model parameters to observed data. Our approach, however, estimates the rate functions directly from on-line measurements, thus bypassing the use of models, by exploiting a close analogy between large deviation rate functions and thermodynamic entropy.

Preliminary work applying this method to predict cell loss and delay on the Fairisle ATM LAN has been performed by the Dublin Institute of Advanced Studies and Cambridge, with good results. The theoretical foundations of the approach are now well established. Experience gained from the preliminary experiments will be used to develop algorithms for resource allocation using entropy as a traffic descriptor, and to test them on both Fairisle and the Swedish Public ATM network. The technique of on-line measurement will also be extended to the operating system in the end-user computing platform, to allocate resources for multimedia applications.

Multimedia communication and applications have enormous economic potential. Understanding how to deliver multimedia services in a cost effective way is key to the economic exploitation of these technologies. Cost effective does not simply mean efficient, it means balancing the tradeoff between complexity of algorithm and efficiency of resource usage. This is true whether the resource being consumed is in the network or the workstation. This work is thus relevant to network operators, network equipment providers, operating system providers and application providers.