Adam Smith Goes Mobile: Managing Services Beyond 3G with the Digital Marketplace

James Irvine

University of Strathclyde - Department of Electronic and Electrical Engineering – 204 George St. – Glasgow G1 1XW – Scotland

Ph.: + 44 141 548 4072, Fax: + 44 141 552 4968, e-mail: j.m.irvine@strath.ac.uk

ABSTRACT

The next generation of mobile communications systems is expected to offer new business opportunities to existing and new market players. A market-based middleware framework has been recently proposed whereby service providers, independent of network operators, are able to tender online service contracts to network operators in a dynamic and competitive manner. This facilitates a seamless service provision over disparate networks in a consumer-centric manner. Service providers select network bearers according to the network operators' ability to meet the QoS target, which in turn is influenced, among other things, by user's price and quality requirements. The benefits of this proposal are the complementarity of numerous network resources, the decoupling of services and networks in a self-organising distributed environment, and increased competition to consumers' advantage.

1. INTRODUCTION

The provision of services as systems evolve beyond the 3rd generation is very challenging. Users will want access to more services, and the services themselves will be more diverse and have a greater range of requirements. Higher capacities will be needed, forcing a move to smaller cell sizes and air interfaces optimised to the specific environment. The trend towards hybrid networks with multiple air interfaces is likely to increase. Existing management approaches are not flexible enough to adapt to these changes.

An economy is an example of a highly complex distributed system. The Scottish economist, Adam Smith, argued in his seminal work *The Wealth of Nations* that a market was the optimal way of allocating a society's resources and satisfying both producers and consumers [3]. He also defined a number of principles required for such a market process to operate fairly which still hold true today.

In this paper, we describe the recent digital marketplace proposals which allow use a market to control services within a mobile communication system. We explain the features of the marketplace and the reasoning behind the concept. We also place the digital marketplace in context with some other similar proposals for the distributed control of systems.

2. EXISTING TRENDS IN SERVICE PROVISION

It is almost certain that the environment for the provision of services beyond 3G will have changed considerably from the current view as 3G is being deployed. Two significant trends are towards a separation of service and transport provision, and more flexible pricing strategies.

2.1 Separation of Service and Transport

Many countries are moving to introduce competition in utility services. A feature of such services, which interestingly does not apply to mobile radio, is that they are often provided over a fixed infrastructure which is difficult to duplicate. Examples include electricity, gas or water supply. In these cases, the infrastructure is shared and separate service providers compete to supply customers over this infrastructure.

This separation between service providers and those providing transport – network operators – was envisaged in the UMTS business model. Mobile radio does not require a single infrastructure, but the registration and paging arrangements do require them to be connected to a single network. An example of the separate service provider concept already in existence in 2G systems is Mobile Virtual Network Operators (MVNOs) [9], who sell there own services but use a separate network operator to connect to their users.

2.2 Pricing

Consumers tend to dislike uncertainty. When offered a choice between fixed monthly charges and a price per call, there is a tendency to select the fixed price option even where there may be savings by working on a per call basis. However, demand based charging has the advantage of forcing more efficient use of scarce resources.

Fitkov-Norris and Khanifar [2] consider the effect on usage of a mobile communication system with dynamic pricing, and argue that dynamic pricing can be used in order to allow a network operator to maximise revenue. However, although they consider the effects of user mobility so that users may wait until they move out of a high price area before making a call, their system does not consider the possibility of users switching between operators. This paper contains useful references to other work on dynamic pricing in the communications and utility fields.

The separation of service providers and network operators gives a further opportunity for dynamic pricing, since the network operator can charge a dynamic price to the service provider without the customer being aware of the fact. This is commonly done in the case of electricity supply, for example. Different service providers can offer different service packages and pricing structures to users. In general such systems work well; the recent problems in California being the result of regulation on the end price

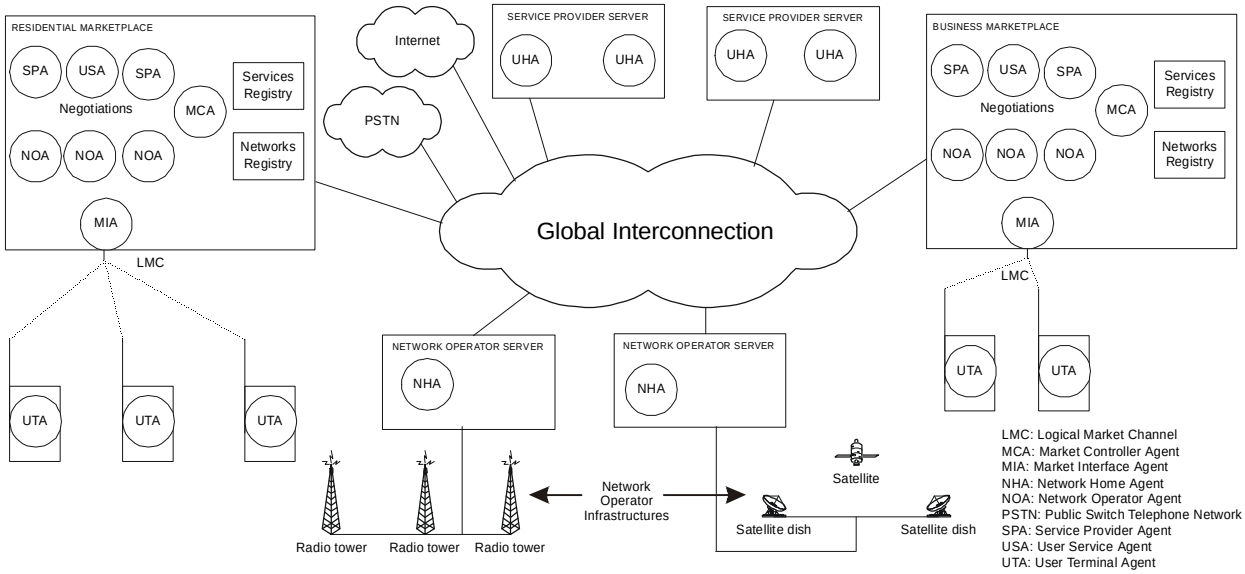

Figure 1 Digital Marketplaces

to users without corresponding regulation on the price from producers. Returning to Smith, in a true market-based system, requirements to regulate prices are replaced by a requirement to regulate the market to ensure free competition.

3. REQUIREMENTS OF A MARKET BASED SYSTEM

In *The Wealth of Nations* [3], Smith argues that for a market to establish a fair price in return for a commodity, producers and customers must have free access to the market. Smith also argues that such a system will produce a satisfactory outcome for producers and consumers, and also allocate the resources of the society optimally.

The market itself can be relatively easily provided. Agents have been proposed as a means for controlling telecommunication systems [4]. A marketplace can be provided as a venue for these agents to negotiate.

The requirement that the market operate 'freely' is, however, much more difficult to meet. The requirements can be broken down as follows

1. Producers must be able to advertise their wares and be able to sell to any consumer.
2. Consumers must be able to buy from any producer they wish to.
3. 'Products' – i.e., the service, must be well described.
4. Producers and consumers must be able to trust each other.
5. Given that the radio interface is unreliable, and so 'product quality' cannot be absolutely guaranteed, there must be a fair method of specifying how much confidence a consumer can place on the 'product'.

An implied consequence of the first two requirements is that there is more than one provider and that the services from different providers are interchangeable. This is termed *substitutability* [10]. It does not imply that the services have to be identical; just that the consumer is able to choose one service over another, otherwise an effect monopoly operates. However, the mobile market is already sufficiently mature that this is the case.

Having specified the requirements, we will now describe the digital marketplace environment which fulfill them and so allow a market to operate.

4. THE DIGITAL MARKETPLACE

The Digital Marketplace is a central point for trading transport and services. Many different marketplaces would be distributed throughout the network, each one dealing with trading for a relatively homogenous part of the system (Figure 1). Marketplaces are agent based, with different parties to the negotiation represented by agents. Users wishing to make a call, or their service provider, will send a Service Agent to the marketplace in order to seek bids from network providers willing and able to provide the required service to the mobile. Each network providing coverage in the area of the marketplace will have a Network Agent. These Network Agents will make bids to the Service Agents, which will choose the most appropriate (cheapest, highest quality, etc) for its needs. The business relationships between these entities is shown in Figure 2. The business model exhibits some similarities with the already proposed TINA-C business model [12]. However, it has been elaborated introducing the ability that both service and transport could be negotiated.

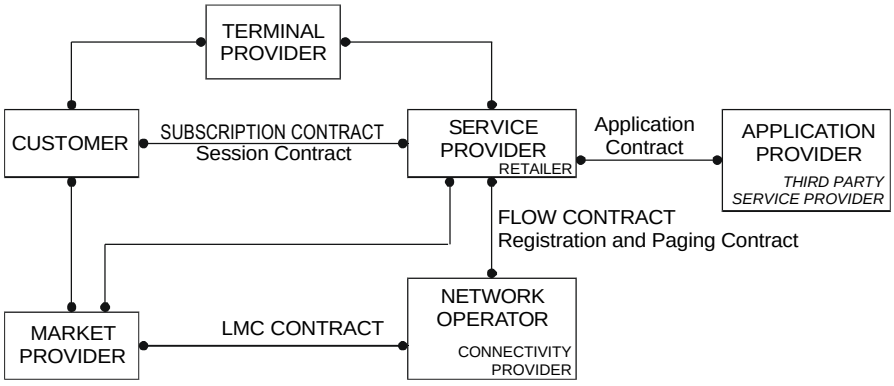

Figure 2 Business relationships in the digital marketplace

5. MANAGING TRADING IN THE MARKET

In order for the digital marketplace for provide a free market, the five requirements identified above must be satisfied. The first two requirements relate to accessibility, the third to transparency (networks may be using different technologies, so a common description of the service to be provided must be defined), and the final two to accountability.

5.1 The Requirements of Accessibility

Criteria (1) is relatively simple to achieve, as it simple requires that all producers – network operators – register in the marketplace. If the marketplace consists of a number of separate entities distributed throughout an area, this has a useful bonus. Network providers which only provide service over small areas can still register in their local market and attract customers. This large base of small suppliers helps to introduce competition in to the market. Small providers, using the IEEE 802.11 WLAN standard, are already offering services in airport terminals and restaurants, but lack a way of advertising their service to passing user terminals.

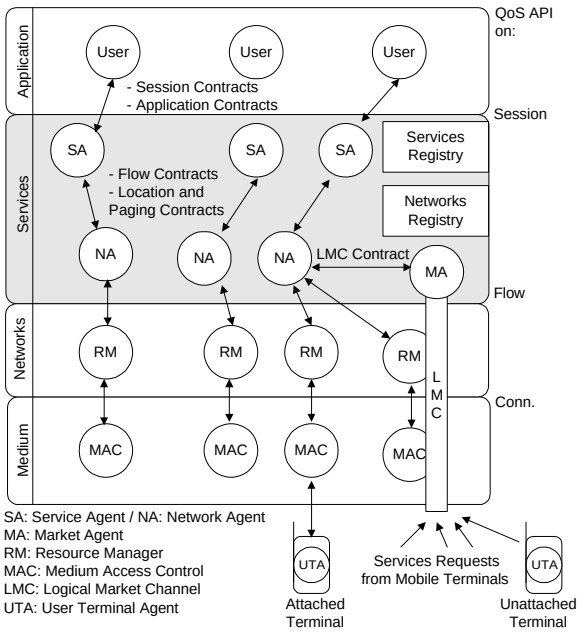

Figure 3: Digital Marketplace within a layered structure

In order to achieve requirement (2), that any consumer can purchase from any producer (technology permitting), each radio access network supplies a Logical Market Channel (LMC) (Figure 3), a common access channel which mobiles can use to initiate transactions. To ensure independence, these control channels will be contracted with network operators by the marketplace. The marketplace may be run by a regulatory body, or more likely by a consortium of network operators or single network operator on behalf of the regulatory body. Due to the nature of the marketplace, it is quite likely that a user may contract for a service with a different network operator from the one providing the LMC. System efficiency can be improved by sharing these channels so that only one or two are provided per access technology. Each Service Agent, created in the market in response to the call initiation, registers with the Market Agent and can therefore verify that its request has reached the marketplace intact.

Users or their Service Providers will put in special contracts to the marketplace for paging and location updates to allow them to receive incoming calls. This signalling will normally imply a cost, and some users, particularly those of data services, might not wish to make use of a facility to receive connections.

5.2 The Requirement of Transparency

In order to allow a fair comparison between different network providers, with possibly heterogeneous networks, a generic description of the service is required. This description forms part of a contract between the user or their service provider and the network operator, and it is these contracts that the Network Agents place bids to provide.

Three basic parameters covering what has to be transported are:

- Bit rate (T): the minimum bit rate to be offered.
- BER: a maximum BER acceptable by the service.
- Delay: the maximum delay tolerated by the service.

However, these parameters are not sufficient in and of themselves to describe communications in a variable environment such as a wireless system. Therefore, the concept of a degradation allowance is also defined. The degradation allowance is the proportion of the service which may be allowed to fall below the contracted parameters. A degradation allowance of 0 would imply that the contracted parameters would always have to be maintained for the contract to be met. Such a situation is difficult to guarantee, and so such calls would be expensive. Increasing the degradation allowance would make the contract easier to meet, and therefore is likely to attract cheaper bids from Network Agents. To support the degradation allowance, three additional parameters are defined as part of the contract:

- Degradation Allowance: the proportion of measures which are allowed to be non-compliant with the three performance parameters over a sliding window.
- Sampling rate: the rate at which measures are performed.
- Monitoring Period: the length of the sliding window.

An extension of the fixed contract is a multimode contract where different services are provided at different points in time. Multimode contracts allow more of the quality of service management to be devolved to individual network elements, simplifying network management at the expense of more variability in the service provided to the user (see [6]).

5.3 The Requirements of Conformance

Requirement (4), that network operators and service providers trust each other, is fairly easy to meet using authentication mechanisms, since these parties will have long term relationships within the marketplace. Users which have contracts with service providers will be trusted through their service provider.

A user may choose not to have a relationship with a service provider by generating an agent and access the market directly. If they do this, they are unlikely to be trusted. An 'up front' payment by means for electronic cash could be used in this eventuality.

Requirement (5) is difficult to meet given that the trading system is operating on an individual call level, as it is difficult for a user to make meaningful measurements with regard to overall performance. A user either receives the contracted quality or not. For this reason, the market maintains a reputation for each network operator, which represents their ability to meet past contracts based on information passed to them by Network and Service Agents. Service Agents can take a network's reputation into account when considering bids as well as simply considering the offered cost [1].

6. OPERATION OF THE DIGITAL MARKETPLACE

As an example of how the digital marketplace system works, consider a mobile user which is not attached to any network infrastructure but is wishing to originate an outgoing communications session. The user has a valid subscription contract with a service provider. The user's terminal first scans radio channels seeking an LMC broadcast signal associated with the reachable marketplace. Once the LMC located, the terminal transmits a session contract stating quality and price requirements through the LMC along with its identification and the identification of its associated service provider. The market agent analyses the connection request and forwards it to the service provider. The service provider analyses the request and, if compliant with the subscription contract, calls for bids in the marketplace where the user is located. Registered network operators propose bids on the contract tender(s) and the service provider selects the one(s) which best suits the contract requirements. The sequence of interactions involved in the process is given by the following functional steps (see Figure 4):

1. The user generates a connection request (Conn.Req) through the LMC to the Market Agent (MA). A session contract is embedded in the request.
2. The MA forwards the connection request to the User Home Agent (UHA). The UHA is active on the service provider domain. The UHA location is embedded in the registration request generated by the User Terminal Agent (UTA).
3. The UHA migrates the Service Provider Agent (SPA) to the marketplace where the user is located.
4. The SPA split the session contract into several flow contracts and tenders each flow contract to Network Agents (NA).
5. NAs propose back bids for the flow contracts.
6. The SPA selects the NAs that are the most suitable to support the session.
7. NOAs that have been selected by the SPA to support the flows confirm to the UTA that the flows are established.
8. Once the session communication ends then the UTA releases each flow by sending connection release signals to the selected NAs.
9. NAs reports on commitments to the MA. The MA updates the associated decommitment penalties.
10. The NAs inform the SPA that the session has ended.
11. The SPA updates the UHA state with the billing information related to the session.
12. The SPA is removed from the marketplace.

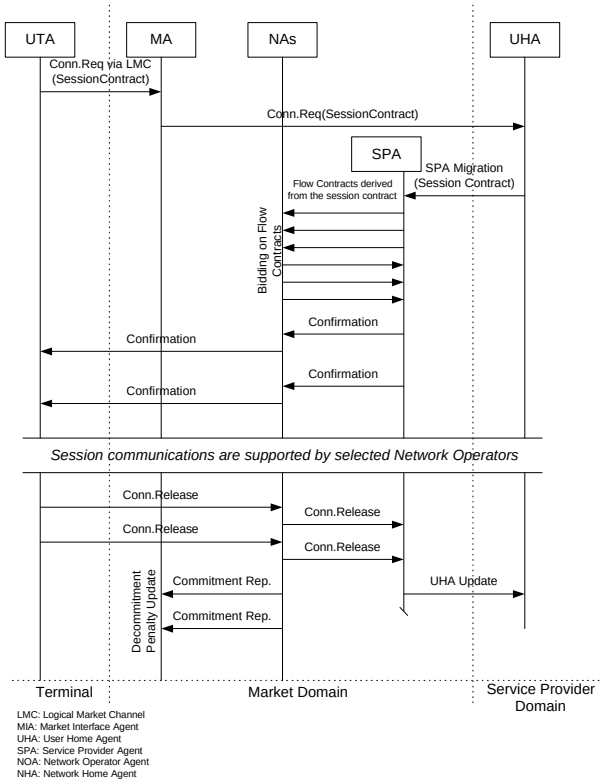

Figure 4 Establishment of an Outgoing Session

The previous scenario was concerned with the trading of a set of flow contracts for the support of a communications session. Another type of contract is introduced in the framework to enable service providers to keep track of the location of their mobile users in order to route incoming sessions. The *location and paging contract* is negotiated in a digital marketplace and states that the network operator will keep track of the user location and page the user when requested by the service provider. The dissociation of communications services, as specified by flow contracts, and location and paging services means that a user might be paged for incoming calls by a network operator but could use the services of another network operator for the transport of the call traffic.

This dissociation of signalling also means that the costs of signalling become explicit, at least to the service provider (which is likely to provide a 'package deal' to the user). This avoids cross-subsidy between users which is another important feature of a fair market.

Pricing strategies within the digital marketplace can be simple or complex depending on the wishes of the operators. The strategies for service providers and network operators will be different.

A simple strategy for a network operator would be to charge a fixed price for each resource, accepting calls only where a commitment over a certain level can be given, and blocking calls when no resource is available. A complex strategy would see prices raised to cover the level of commitment which is being offered, as well as

the prices charged by other operators and the level of remaining resources an operator has.

A service provider, on the other hand, will have an agreement negotiated with a user covering the type of calls which will be served and the cost of each call. The cost to the user is likely to be, at least relatively, fixed. For this reason, if the cost in the market is high, a service provider may block a call even though resources were available to serve it. However, it would only be able to do this if its customer agreement allowed it to do so. A user may pay a premium for special service, requiring the service provider to facilitate its connection whatever the cost.

Example negotiating strategies and price fluctuations are described in [1].

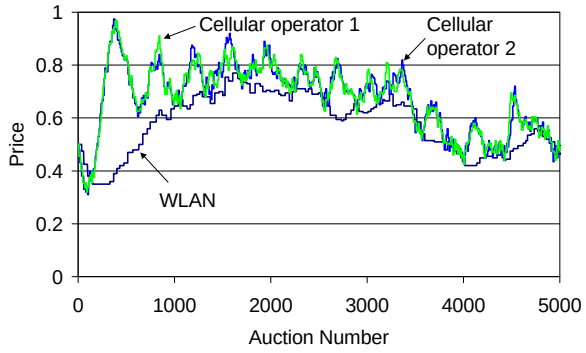

Figure 5 Pricing for cellular vs WLAN.

As an example, Figure 5 shows the example of a marketplace with three operators, two cellular operators competing with a WLAN operator. The cellular operates cover the entire area of the marketplace, while the WLAN operator has more limited coverage, and so cannot offer such a high commitment. Users in the system are split between 'business' users, who are prepared to pay a high price for service, and 'personal' users, who are price consious and will defer a call if the price is high. The load in this case steadily increases from the start of the simulation where the available traffic is 75% of capacity to about the 2500 auction mark where it is 25% above capacity, then falling back to the 75% level. As load increases, the cellular operators, which can serve all users, are able to significantly increase their prices. As the load increases still further, the WLAN operator, although operating in limited areas, is able to raise its prices to similar levels, but cannot exceed the cellular operators' prices. At this point most personal users do not find the prices worthwhile, so business users predominate. As load decreases, the WLAN has to maintain a slight discount to the cellular operators to maintain loads and is not as flexible in responding to fluctuations in load as it has fewer users in its coverage area.

7. MULTIPLE MARKETPLACES

The forgoing description of the digital marketplace concentrates on its use to negotiate a connection from a radio access network to a fixed network and a service on that network. However, the concept applies equally within these networks so that calls may be routed through the influence of multiple marketplaces between there origin and destination. Intermediate routing may be influenced by the prices obtained in each of these markets.

There are two options for controlling the call in this circumstance: retained responsibility and devolved responsibility [11]. With retained responsibility, the SPA migrates to each marketplace in turn to negotiate a leg of the call. With devolved responsibility, the SPA negotiates with a NA which is willing to transport the call to its destination, and that NA may then use pre-existing agreements with network providers in other marketplaces to complete the call or itself conduct the necessary negotiations if it cannot transport the call itself. The second option is likely to carry a premium, since the network operator is risking its reputation through the actions of the other network providers it employs to carry the traffic.

8. RELATED PROPOSALS

The IST SHUFFLE programme [5] is currently examining somewhat similar ideas for the distributed control of 3G networks [7]. They too proposal an agent based system for the control of resources, although focus is more on the control of resources within a network than to the control for services between, as well as within, networks. They consider dynamic agreements between service providers and network operators on a call by call basis as being too demanding in terms of computation and signalling, and settle instead for what amounts to a futures trading system, where service providers purchase the option to use network operators in the future rather than when the call is set up.

Such a proposal can be considered as a subset of the digital marketplace proposal, as there is nothing to prevent the contract negotiated between the NA and the SPA covering several calls. Indeed, in a similar manner, the MVNO system is another specific example, with a very long term arrangement between a service provider and a single network operator. However, the full benefits of the digital marketplace in terms of open competition and the ability for local suppliers to operate effectively in the market only accrue when there is the possibility (not necessarily realised) of switching to another operator for the call. Fixed long term agreements distort the market and break the supply/demand cost/price relationship, and either the producer or the consumer will lose out [3].

Negotiation procedures in the digital marketplace do not add greatly to signalling since although it is possible to renegotiate a contract with a network operator within an ongoing call should quality degrade, this is unlikely if multi-mode contracts of the type described in [6] are used. Where the additional signalling does become significant is if the service provider tries to route a call through several digital marketplaces and negotiate in each [11].

9. DISCUSSION

One of the framework's key features is to let suppliers and customers of communications services negotiate in a fair and competitive environment. The price charged for services will therefore be competitive and dynamically variable in line with demand and supply. In current mobile systems, the pricing schemes have been borrowed from the ones established for fixed telephony:

time-based pricing (off peak and peak periods) with a subscription fee. The price paid by end users does not therefore reflect the underlying radio resource cost. For instance, it is less resource-consuming for a network operator to accommodate a communications session in the residential area of a city during the office hours. Why, then, during the office hours, a user should be charged at a peak rate when still located in the residential area? Why is a user in the residential area charged as much as another user in the business area where the radio resources are scarcer? This is not an optimal use of resources, which would encourage use where resources are plentiful and discourage their use when resources are scarce. This requires pricing strategies more representative of the use of radio resources; a price for peak price geographical zone and a price for off-peak geographical zone. Applications that would benefit from this scheme are callback services [8] and range from email download to automatic organiser update. The definition of pricing strategies goes beyond the scope of this article but it is expected that the digital marketplace based framework will allow the introduction of fairer pricing schemes and novel applications that efficiently exploit the radio resources.

It is accepted that while most economists would agree that a fair market is in practice the optimal way to allocate resources, it is not globally optimal. The market in pizzas is a good example. In the centre of most cities it is possible find several fast food pizza restaurants, all selling basically the same product. It would be more efficient to provide only a single restaurant, to which anyone wishing to purchase a pizza would go. This would equate to the case of a centrally planned economy. However, while a centrally planned economy can in theory be more efficient than a market-based one, in practice there are inevitably errors in estimating demand leading to shortages or surpluses, and consequent inefficiencies.

The traditional mobile radio system is a centrally planned system. Resources are shared between radio access points in a centrally controlled manner. For such a system to operate, we require perfect information on the state of each part of the system. If such information is available, then they will outperform a digital marketplace based system. However, as systems get larger and more diverse, this becomes increasingly impractical, just as it does on an economic scale.

10. CONCLUSIONS

This article has presented the digital marketplace for the management of mobile communications services. The framework has been developed to meet the need of the next generation of mobile systems by offering service providers a market-based middleware over which they can develop applications independently from the underlying network infrastructure. This is made possible with the development of a generic quality-based interface between the network infrastructure and the application entities. Network operators also benefit from the framework in the sense that by registering in a marketplace, they can serve customers, which are not necessarily their direct subscribers.

A free market will benefit the producers and consumers in the market, but there are also benefits to the regulatory authorities. The regulatory is presumably interested in ensuring free competition, but an open market will encourage investment by new players in the industry, such as local operators, which is important given the costs of providing the increased capacity new services require. The presented approach for the management of communications services represents an enabling technology for the development of next generation of communications systems in a multi-vendor, multi-technology and multimedia environment.

ACKNOWLEDGEMENTS

The author is grateful to his colleagues, Gwenäel le Bodic, John Dunlop and Clare McKeown, who have also worked on the ideas reported in this paper.

The work reported in this paper has been part of the Services Work Area of the Core I Research Programme of the Virtual Centre of Excellence in Mobile & Personal Communications, Mobile VCE, www.mobilevce.com, whose funding support is gratefully acknowledged by the author. More detailed information and software tools on this research are available to Industrial Members of Mobile VCE.

REFERENCES

[1] Le Bodic, G. Girma, D., Irvine, J., and Dunlop, J., 'Dynamic 3G Network Selection for Increasing the Competition in the Mobile Communications Market', VTC2000-Fall, Boston, September 2000.

[2] Fitkov-Norris, E. D. and Khanifar, A., 'Dynamic Pricing in Cellular Networks, A Mobility Model with a Provider-Oriented Approach', 3G2001, London, March 2001.

[3] Smith, A, 'An Inquiry into the Nature and Causes of the Wealth of Nations', 1776 (rep. Penguin, 1982)

[4] Hayzelden, A. and Bourne, R., 'Agent Technology for Communication Infrastructures', John Wiley & Sons, Ltd., 2002.

[5] http://www.ist-shuffle.org/

[6] Le Bodic, Irvine, J., G. Girma, D., and Dunlop, J., 2000, 'Co-operative Service and Link Adaptation Techniques for the Support of Multi-media Applications over Wireless Channels', Proc. WPMC, Bangkok, Thailand.

[7] Buckle, P, Dinis, M, and Cuthbert, L, 'Distributed Intelligent Control and Management for 3G Networks', SCI2000, Orlando, FL July 2000.

[8] A.T. Campbell, G. Coulson and M.E. Kounavis: 'Managing Complexity: Middleware Explained', *IEEE IT PRO*, pp. 22-28, September-October 1999.

[9] Office of Telecommunications, Mobile Virtual Network Operators: OFTEL Inquiry into What MVNOs could offer Consumers, June 1999.

[10] Wellman, M. P.: 'A market-oriented programming environment and its application to distributed multicommodity flow problems', *Journal of Intelligence Research I*, 1-23, 1993.

[11] Irvine J, McKeown C & Dunlop J, 'Managing Hybrid Mobile Radio Networks with the Digital Marketplace', VTC2001-Fall, Atlantic City, 2001.

[12] TINA 1.0 Deliverables and Specifications www.tinac.com/specifications/specifications.htm