Evolution of Charging and Billing Models for GSM and Future Mobile Internet Services

John Cushnie[1], David Hutchison[1], Huw Oliver
Internet Mobile Solutions Laboratory
HP Laboratories Bristol
HPL-IRI-2000-4
14th July, 2000*

mobile service providers, 2nd and 3rd generation mobile, network charging

[1] Distributed Multimedia Research Group, Lancaster University, Lancaster, U.K.
* Internal Accession Date Only Approved for External Publication

1. Introduction

The Global System for Mobile (GSM) was first introduced in 1992 with approximately 23 million subscribers, rising to over 200 million in 1999 on over 300 GSM networks [1]. The aim was to provide a global mobile telephone network that could be implemented using standard building blocks not tied to specific hardware vendors. The uptake of GSM by subscribers is far higher than any industry predictions and typifies the 1990's and the increasing need for personal mobility.

The 1st generation GSM mobile telephone networks provide subscribers with high quality voice communications and low bandwidth (9.6Kb/sec) data connections for FAX, Short Message Service (SMS) and full dial-in connection to the Internet for email and web browsing, usually requiring a mobile computer or intelligent handset. The addition of overlay communication protocols, such as Wireless Application Protocol (WAP) [2], allow mobile handsets on 1st generation GSM networks to be used for secure connection applications such as mobile banking and other transaction based services.

International roaming agreements between the numerous mobile telephone network providers allow subscribers to be reachable almost anywhere in the world where there is GSM coverage using the same telephone number and handset. Satellite based services such as GlobalStar [3] and ICO [4] allow GSM subscribers to further expand their network coverage and availability using the same mobile communications infrastructure.

The increasing use of mobile telephones and devices for data communication drives the need from the market for a fast, reliable and available infrastructure. GSM proposes to provide the required infrastructure using 2nd and 3rd generation GSM which introduce new technology that allows increased data bandwidths and new data services [1]. 2nd generation GSM introduces the General Packet Radio Service (GPRS) and 3rd generation GSM introduces the Universal Mobile Telecommunication System (UMTS).

The introduction of 2nd and 3rd generation (2G and 3G) GSM technology brings convergence of GSM mobile telephone networks with the Internet. Packet Switching [5] is being introduced as the switching mechanism for data calls and internet sessions, in contrast with the current circuit switching implementations currently used in 1st generation GSM and fixed line telephony networks. Due to the quality of service limitation of packet switching protocols and Voice over IP (VoIP), circuit switching may still be used for voice communications on 2nd and 3rd generation GSM networks.

The 2G and 3G technologies bring the Internet to the mobile subscribers. The same services available from the Internet today, including email, secure transactions and Web browsing become available on mobile telephone devices, using the standard infrastructure of the Internet.

In order for the mobile telephone networks to be able to offer these additional services to the customers there is a requirement for the recovery of the infrastructure investment cost. This is a prime justification and motivation for charging and billing for telephone network usage together with the need for generating commercial profits for telephone network shareholders and companies. Charging may also be used to provide congestion control in under-provisioned and over-subscribed networks. This may be achieved by placing price premiums on network bandwidth, Quality of Service (QoS) or utilisation, which may then make the network self-regulating at a first level, or by reducing usage with additional charges that fewer subscribers may be prepared to pay for.

2nd and 3rd generation GSM networks present the network operators with many charging and billing challenges. GPRS may be charged for by packet usage counting, but the associated cost of measuring the packet usage may be greater than their actual worth. The experience gained with charging and billing with GPRS will prove valuable when UMTS is being rolled out in GSM networks.

There are various proposed economic and technical models for charging and billing for Internet usage. Most of these are equally suitable for charging and billing of mobile telephone network traffic, especially with 2nd and 3rd generation GSM systems.

This paper first describes the convergence of GSM networks with the Internet and then examines the methods for collecting the billing information and then the processing of this billing information. Future research that needs to address the suitability of the currently proposed charging models is discussed. This includes investigation and development of new or combined charging models that may be used together with benchmarking and testing of the charging models with real or simulated data to illustrate their possible application and suitability.

2. GSM Mobile Networks and the Future Internet

1st generation GSM networks [1] provide high quality digital telephony with low bandwidth (9.6Kb/sec) data communications for FAX and SMS. Implementations of GSM are typically multi-vendor and consist of a layered architecture including the mobile telephones, the telephone network and the subscriber invoices and bills, as shown in figure 1 below:

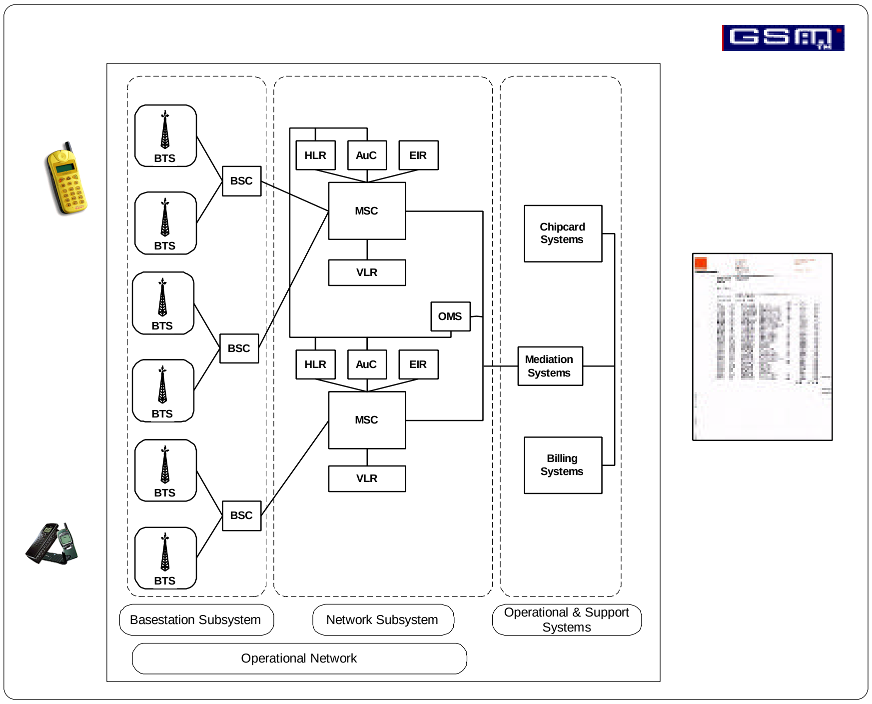

Figure 1 GSM Network Architecture

The Base Station and the Network Subsystems are often referred to as the Operational Network (ON), and is usually physically distributed around the area of coverage of the GSM network. The ON elements are often sited remotely with wide area networking (WAN) connectivity to the rest of the network to allow centralised remote administration of the network. The Base Transmitter Stations (BTS) and the Base Station Controllers (BSC) provide the air interface for GSM, which is then circuit switched [5] using the telecommunications industry standard SS7 by the Mobile Switching Centers (MSCs) in the ON. Additional Gateway MSCs allow switching to other mobile and fixed line telephone networks, allowing interconnection and roaming. Billing tickets for all calls made in the network are produced on the MSCs, usually based on subscriber Ids in the network.

Operational Support Systems (OSS) are usually centralised in a data center or across several data centers for disaster and fault tolerance. The OSS provides the interface to the customer invoices and bills, and normally includes systems for billing, subscriber administration, GSM Subscriber Identification Module (SIM) chipcard production, fraud detection, voicemail and off-line data-mining and analysis systems. Most mobile telephone networks de-couple the ON from the OSS using Mediation Systems or Mediation Devices (MD). These systems are used to collect billing data from the ON and also to manage the subscriber databases in the ON elements. In some networks the MD forms part of the billing system.

The collection of the billing data is normally via high-speed communication links using reliable data protocols such as File Transfer and Management (FTAM) and X.25. Once billing data is collected centrally it can be processed into subscriber invoices and bills using dedicated billing systems and the mobile network's charging tariffs. The billing data can also be further processed by additional data-mining systems to detect subscriber's usage patterns, possible fraud detection and subscriber profile surveying.

With 2nd generation GSM the General Packet Radio Service (GPRS) [1,6,7] is introduced providing an overlay service for Internet access that shares the same air interface as 1st generation GSM. The design goal behind GPRS is to provide high-speed Internet data communications for mobile devices and subscribers using the existing 1st generation GSM air interface, thereby minimising the cost impact on the existing installed network infrastructure.

GPRS provides a direct interface to the Internet services for GSM mobile telephone devices, and is implemented in an existing GSM network with the addition of two new ON elements the Signalling GPRS Service Node (SGSN) and the Gateway GPRS Service Node (GGSN) as shown in figure 2 below:

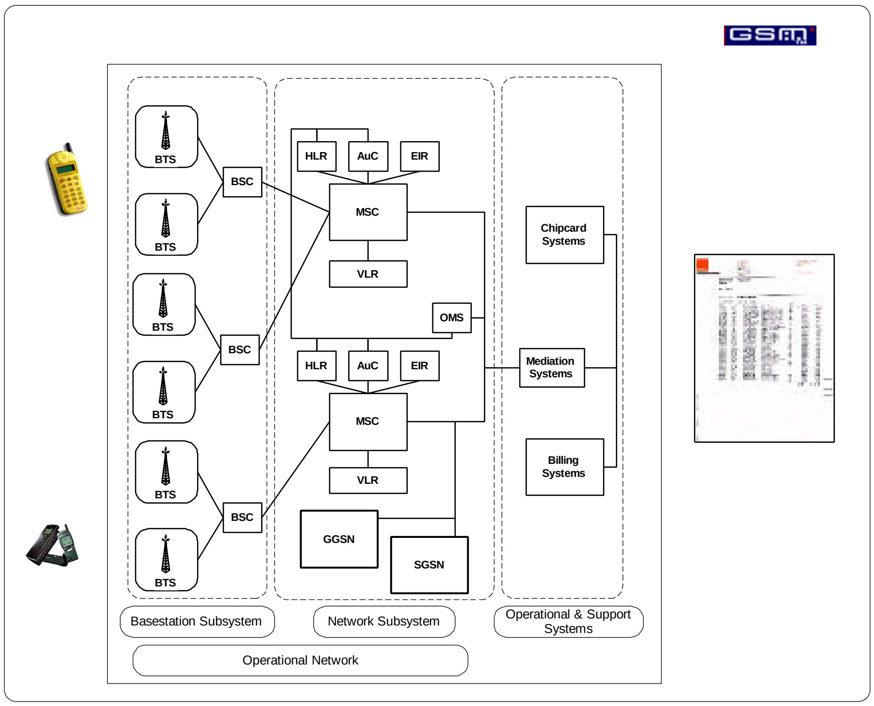

Figure 2 GSM Network Architecture with GPRS

Additional modifications to the existing BTS and BSC to include Packet Control Units are also required so that the network is GPRS aware. The two new ON elements provide the interface between the GSM air interface and the TCP/IP network used for the GPRS specific traffic, (i.e. Internet sessions used for email, http, ftp etc). GPRS has the advantages of digital telephony of GSM combined with the increased bandwidth of 115Kb/sec over the air interface for data traffic. The actual bandwidth of the channels or gateways to the Internet is dependent on the network provisioning of the mobile network operator. The GGSN and SGSN in the ON provide the switching for the mobile data sessions and use packet switching [5]. GPRS data sessions are routed by the MSCs as for 1st generation GSM with the SGSN and GGSN routing the Internet sessions to the TCP/IP network, using packet switching [5]. Circuit switching such as SS7 used in mobile and fixed line telephony relies on network resource allocation when the call is set up to guarantee the Quality of Service for the call. This approach carries a high overhead and does not allow for adaptive routing or switching when the overall network load is variable. Packet Switching makes full use of the available bandwidth of the underlying network, but often has a reduced Quality of Service, and is suited to 'bursty' network traffic including Internet protocols such as http, ftp, email etc, where guaranteed qualities of service are not a top priority. In addition to introducing TCP/IP packet switching GPRS equipped mobile networks may roll in IPv6 [8] as the preferred IP protocol. This will allow the large number of addressable network nodes that will be required when there is a high saturation of mobile devices requiring Internet connectivity.

The GGSN and SGSN produce billing tickets and statistical data relating to Internet traffic usage generated by GPRS calls and sessions. GPRS also introduces the new concept to the mobile subscriber of having endless connections to an IP network that may stay open at low cost, giving the advantages of keeping connection times and cost to minimum. Voice calls on GPRS handsets are routed by the MSCs using the same circuit switched channels and protocols as 1st generation GSM calls.

GPRS brings with it a new set of parameters to the challenge of billing and charging subscribers for using the GSM mobile networks. Mobile telephone network subscribers are normally charged on a time and usage basis for the high quality telephony. With GPRS there are new possibilities to charge the subscribers for how much data or bandwidth they use in the network, in addition to the amount of talk-time consumed. This shares commonality with the possibilities currently being proposed for general Internet charging and billing. As in the Internet the cost of packet usage counting may be more expensive than the value of the packets being counted. These new challenges need to be met and addressed by the network operators.

3rd generation (3G) mobile telephone networks arrive with the introduction of Universal Mobile Telecommunication System (UMTS). This will be based on the standard ON and OSS GSM architecture with the addition of UMTS specific Network Elements. It will build on the infrastructure installed for GPRS with a marked increase in maximum bandwidth to 2Mbits/sec [1].

Typical GPRS mobile handsets may include display screens for multimedia content together with some kind of keyboard or touch-screen entry and the microphone and earphone required for GSM telephony. Supported applications for 2nd and 3rd generation GSM may involve Internet intensive activities such as web browsing and email communication, as well as traditional mobile telephony. Table 1 below shows a comparison of the bandwidth and communication rates achievable with the different generations of GSM networks [1].

GSM Generation	Date Introduced	Technology	Data Bandwidth
1st	1992	Voice	N/A
1st	1995	SMS & Mobile Data inc. FAX	9.6 Kb/sec Internet via Modem
2nd	2001	GPRS	115 Kb/sec Direct Internet connection
3rd	2002/3	UMTS	2 Mb/sec Direct Internet connection

Table 1 GSM Architecture Generations

The introduction of GPRS is considered a stepping stone to the promises and functionality of UMTS and high-speed access to Internet services and Networking. Many mobile telephone network operators view the experience to be gained with GPRS and the associated billing issues essential for the implementation of systems for UMTS and it may be too late to learn when UMTS is implemented and available for the mass market. The systems and methods developed for GPRS charging and billing need to be compatible with the requirements of UMTS to ensure preservation of investment and knowledge.

3. Infrastructure for Charging and Billing

In order to charge for mobile telephony services the network operator has to first capture the network usage of all of the network's users including subscribers and roaming subscribers. This usage data then needs to be processed and then set against the billing and charging models and tariffs in use. The Mobile Switching Centers (MSC) in the Operational Network (ON) produce the billing tickets for all the calls made in the mobile network. Figure 3 below describes the call-flow for mobile to mobile calls and billing ticket generation in a typical 1st Generation GSM mobile telephone network.

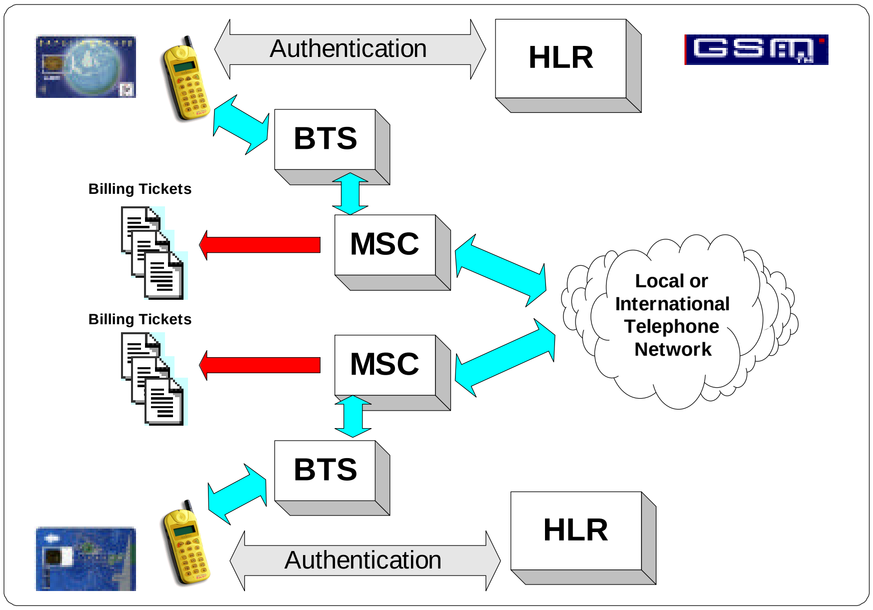

Figure 3 GSM 1st Generation Call Flow

The MSCs produced billing tickets need to be collected and then processed centrally so that the subscriber bills can be produced. The Home Location Registers (HLR) are responsible for authentication of subscribers within the mobile network. The collection of billing tickets is often done by a mediation system. These systems may also carry out vendor specific translations on the billing ticket formats so that multi-vendor ONs can be implemented, or to allow the native billing tickets to be used on commercial billing system, or on other centralised OSS systems used for data-mining. The heterogeneous nature of most mobile networks may be very problematic with many different file formats and standards being involved. Once all the billing tickets have been collected and pre-processed into a standard format that the billing and other OSS systems can understand they may them be used to produce the invoices and bills for the subscribers.

With the introduction of GPRS comes the need to be able to charge for the new GPRS and Internet services and therefore earn a return on the investment in the new technology. The challenge of charging and billing for the GPRS services is simplified by the GPRS solutions from the major vendors all supporting charging functions. The ETSI [1] standards recommend a Charging Gateway Function (CGF) to handle the billing record generation. Two billing records are generated, one by the GGSN for the Internet part and one by the SGSN for the mobility (radio) part.

Current GSM billing systems have difficulties implementing charging for 1st generation non-voice services and it is unlikely that existing billing systems will be able to process the large number of new variables introduced with GPRS.

An added complication for GPRS charging is the overlap and convergence to the Internet and the multitude of diverse systems connected to it. In addition to the inter-charging between the mobile and fixed telephone networks inter-charging between the mobile networks and Internet providers will be required and this will add to the operational costs of running the GPRS services in parallel to the normal GSM network. With the already proposed Internet charging models the inter-charging between the mobile and Internet network providers has the potential to become very complicated and may include requirements for additional billing and charging systems for the required accounting. There may also be requirements to include additional network elements to allow network traffic monitoring, for example packet counting, at the edges of the network. These will allow the inter-charging data to be captured by the mobile network provider and used to verify the cross-charges from the Internet service providers that the mobile network purchases the Internet access from.

The addition of GPRS to the mobile network modifies the call flows for Internet packet data as in figure 4 below and includes the required gateway to Internet services and external networks:

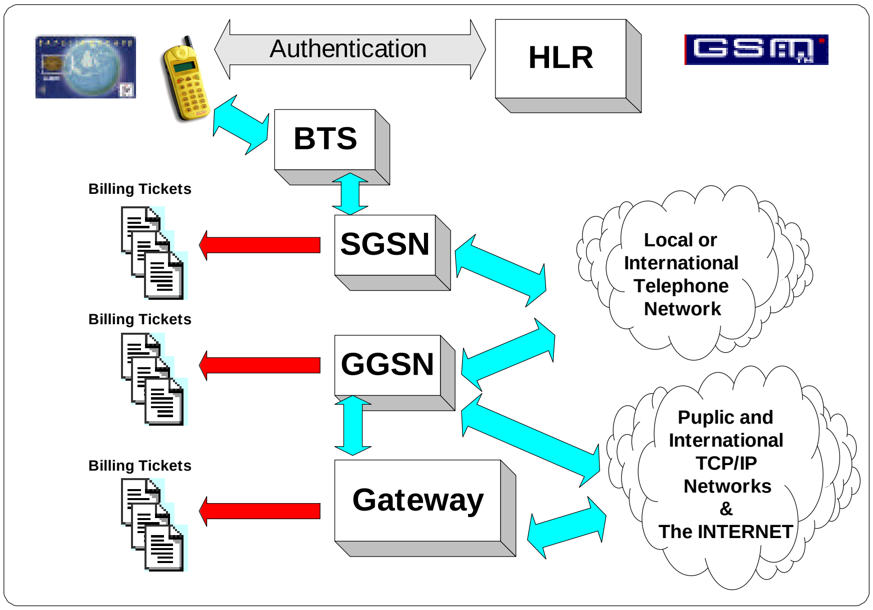

Figure 4 GSM 2nd Generation Call Flow with GPRS

This has the effect of producing more billing tickets and data for processing in the centralised OSS systems from the Gateway GPRS Service Nodes (GGSN) and the Servicing GPRS Service Nodes (SGSN). Network operators may also have packet counting systems in the network that will produce additional billing and charging information that may require processing by the billing systems. It may well be that the cost of measuring the packets is greater than their actual value, both from an infrastructure investment and network traffic cost viewpoint.

The GPRS related billing tickets may be of a different format to the ones produced by the MSCs and may include data on the amounts of packets exchanged during GPRS sessions. The MSCs will produce billing tickets for all voice calls in the network as is normal in 1st generation GSM networks. Extensions to the mediation systems may be implemented for the collection and pre-processing of the GPRS related billing tickets that may then be fed to the billings systems. The above ON systems deal with the capture of the billing data in 1st and 2nd generation GSM systems. For 3rd generation there will be the addition of UMTS Mobile Switching Centers (UMSC) for UMTS specific traffic.

Once the billing ticket information has been collected from the network the mobile network requires a billing and charging system to make sense of all the data and produce the invoices and bills for the subscribers, and also to produce the cross-charge data for partner network providers. The actual cost of providing and maintaining such a billing system can be anything up to 50% of the total infrastructure investment and annual turnover of the mobile network. The billing system therefore needs to be able to provide a good deal of added value to make the investment worthwhile. This provides a valid justification for simplifying the billing function and investigating the use of charging models based on fixed price subscriptions and bulk purchase of talk time and data bandwidth in the network.

Most mobile network operators currently offer contract subscriptions, which include a line rental element plus a contract rate for telephony airtime, usually based on call duration. In addition to the contract subscriptions the network operators offer ‘pre-paid’ contracts where the subscriber pre-pays for the airtime used in the network. From a commercial viewpoint pre-paid makes sense for the network operators, since they will receive payment from subscribers prior to the consumption of resources. This simplifies revenue collection, but with the downside of increased complexity in the ON to prevent subscribers from over-spending on their ‘pre-paid’ subscription.

With the addition of Internet access via GPRS and UMTS existing mobile network subscriptions need to be extended to include charging for the Internet services used by the subscribers. Just how to charge for the Internet services offered to and used by the subscribers is the major challenge to the mobile network providers and will be influenced by many factors.

In most commercial environments some kind of fraud is normally present. The mobile telephone networks are no exception. The vast array of billing ticket information produced by the ON in mobile telephone networks can be processed offline and used effectively for fraud detection. Again the infrastructure investments for such systems are high and their added value has to be proved. Fraud detection fits quite nicely into the billing and charging models and they often go hand in hand. An example of fraud in GSM networks is the running up of large bills on stolen mobile phones. This can be detected using the billing data and the mobile phone being used can be blocked in the network, but incurs a high cost in the real-time monitoring of the network traffic data and the associated systems and personal. With the addition of GPRS and Internet services the opportunity for fraud increases and the network operators need to be aware of the different kinds of fraud that are possible and may occur.

4. Charging Models

There are many charging models that have been proposed [9] for the current and future Internet as well as those traditionally employed by the mobile and fixed line telephone networks. Most, if not all, of the Internet charging models are equally applicable for use in the mobile telephone networks, especially with the introduction of 2^{nd} and 3^{rd} generation GSM systems. Below is a discussion of some of the proposed charging models, and how they can be adapted to the mobile network markets.

Metered Charging

This pricing model is already in use with many Internet service providers (ISPs) and European mobile and fixed line telephone companies. The model involves charging the subscriber for the connection to the service provider on a monthly basis and then charging for metered usage of the service. The usage is usually measured in units of time and there is often a ‘free’ period of usage included with the monthly fee. Variations on this model include having scaled subscription charges that increase with the metered usage.

The use of this model in 2^{nd} and 3^{rd} generation GSM networks may become commercially problematic since subscribers may leave GPRS sessions open endlessly without the handset being powered on. Metered charging based on time for such usage may prove prohibitive. However, if the usage is based on other session parameters, for example number of packets transmitted/received, then the commercial impact becomes less and the model may be usable in mobile telephone networks for data.

Fixed Price Charging

This pricing model is similar to that used by some US fixed line telephone networks for local call charging. The network service provider sets a fixed rental charge for the telephone connection and all local calls are then free of charge with metered charging used for long-distance calls.

The advantage of this charging model is that call data for local calls does not need to be collected and processed, providing a commercial saving for the network operator in the billing systems and mediation systems infrastructure.

Disadvantages of this model include no added revenue for the service providers in times of above average usage on the network, and congestion may also become an issue if the network is under provisioned for the number of possible subscribers at peak times. This provides a strong argument for using charging and billing to improve congestion control by dissuading subscribers from using the network through higher cost for the provided services.

Packet Charging

Packet Charging is specific to Packet Switching [5] networks and involves the capturing and counting the number of packets exchanged in a session. This is a proposed method of metering Internet traffic and being able to cross-charge between networks as well as ISP and mobile subscribers. This model requires the implementation of packet counting systems in the network and complex billing systems that can process the packet data on a subscriber and customer basis.

The advantage of this method of charging is that the absolute usage of the network and services can be metered, calculated and billed for very accurately, as long as the packet information can be captured efficiently.

The major disadvantage of Packet Charging is that the cost of measuring the packets may be greater than their actual value, both from an infrastructure investment and additional network traffic viewpoint. This may lead to packet charging being used as a policing tool to ensure that network bandwidth is used efficiently and not over consumed by the network subscribers, rather than as a direct charging model.

Expected Capacity Charging

This charging model [9] allows the service provider to identify the amount of network capacity that any subscriber receives under congested conditions, agreed on a usage profile basis, and charge the subscriber an agreed price for that level of service. The subscribers are charged for their expected capacity and not the peak capacity rate of the network. Charging involves using a filter at the user network interface to tag excess traffic; this traffic is then preferentially rejected in the network in the case of network congestion but is not charged for; charges are determined by the filter parameters.

This model has the advantage that the price to the subscriber is fixed and predictable which in turn permits the network provider to budget correctly for network usage. The expected capacity model also gives the network provider a more stable model of the long-term capacity planning for the network. This model fits is well with mobile telephone networks and the administration of the agreed expected capacity would be done as part of the normal subscriber administration tasks.

One disadvantage is that the network operator has to police the actual capacity of the network used by subscribers and act accordingly by limiting the subscribers service to what has been purchased, or by invoicing the subscriber for the extra capacity used, on a metered tariff for example.

Edge Pricing

Proposed in [10] this model charges for the usage at the ‘edge’ of the network scope for the subscriber, rather than along the expected path of the source and destination of the calling session. The networks in turn then cross-charge each other for the usage at the network ‘edges’. Edge pricing refers to the capture of the local charging information. Once captured the information can be used for any kind of charging including metered, fixed or expected capacity, for example. Past research [13] has shown that much of the observed congestion on the Internet is at the edges of the individual networks that make up the Internet. The use of edge pricing may be effective as a policing method to monitor and alert the network operators to such congestion.

This approach has the advantage that all session data can be captured locally and does not involve exchanging billing data with other networks and partners for subscriber billing, as for current roaming arrangements between mobile telephone networks.

A disadvantage with this model is the lack of visibility of the routing via external networks and the costs of that traffic to both networks. The cost of collection of the data may again also be an influencing factor in the selection of this method, as for Packet Charging above. The cost of collecting the edge usage information may be in excess of the value of the collected information.

Paris-Metro Charging

This charging model, proposed in [11], introduces the concept of travel class, as used on public transport systems, to network traffic and relies on providing differentiated levels of service based on customer usage pricing only. The scheme assumes that subscribers will assign a preferred travel class with an associated cost for their different network traffic. The class assigned may be simplified to first and second class, as used on the Paris Metro system that inspired this charging model. The choice of class may be made dynamic and the subscriber may also use the throughput of the network to determine which class to use for their required traffic. The network may become self-regulating at periods of high usage. When the network becomes congested and all the capacity in first class is filled subscribers may downgrade to second class to improve their own network performance.

This charging model may work well in GPRS and UMTS networks and allow subscribers to prioritise network traffic, for example business emails may be considered more important that personal email so the cost penalty for first class may be considered appropriate for business email.

An advantage of this charging model is the flexibility given to the network subscribers and also the control they have over the cost of their network traffic.

This model has the disadvantage of introducing mathematical complexity to the network's behaviour and a tariff class decision overhead to the network subscriber. For this scheme to work the class decision may need to be made automatic, and may involve extensions to network communication protocols and the application software handling the network traffic. This charging model also requires the network bandwidth to be segmented and therefore does not allow multiplexing.

Market Based Reservation Charging

This charging model discussed by [12] and usually attributed to Mackie-Mason, introduces the concept of a public auction of bandwidth or network resources. The network subscribers place monetary bids that will influence the quality of service they receive from their network-based applications. This model may be used in the mobile telephone networks by having subscribers to the network maintaining a preferences profile that details the subscriber's bids for the various services used, for example email, voice, http, ftp and SMS. The network provider may then use the subscriber's preference profile when routing the network traffic. In the case of GPRS networks the subscriber preference profiles may be administered via a WWW page and browser or possibly by SMS.

Subscribers have the advantage that they can influence their quality of service from the mobile network by the value they attach to the services they require.

As a disadvantage this charging model introduces some uncertainty to the subscribers with regard to the quality of service in the network. It may also allow some of the subscribers to gain unfair advantage when they have bid for certain services at the expense of other subscribers and network users. This charging model is widely agreed to be unimplementable for Internet networks, and maybe also for the mobile telephone networks.

Summary

The table 2 below compares some of the metrics of the charging models discussed above:

Charging Model	Cost of Implementation	Overhead on the Network	Overhead on Subscribers	Provision for QoS Improvement
Metered Charging	High/Medium	Low	Low	No
Fixed Price Charging	Medium/Low	Low	Low	No
Packet Charging	High	High	Low	No
Expected Capacity Charging	Medium/High	Low/Medium	Medium	Yes
Edge Pricing	Medium/Low	Low/Medium	Low	No
Paris-Metro Charging	Medium	High	Medium/High	Yes
Market Based Reservation Charging	Medium/High	High	Medium/High	Yes

Table 2 Charging Model Comparison

In table 2 above the cost of implementation covers the infrastructure capital investment in new equipment and software to enable the use of the charging model. The overhead on the network of the charging models includes, but is not limited to, the additional network traffic required to implement the model, and the addition of any new systems for data collection and processing over and above the standard GSM building blocks. Overhead to the subscribers include added complexity of tariffs and the maintenance of the subscriber's account to use the charging models efficiently and to avoid excessive charging by the mobile network provider.

From the above discussion of the various charging models it is clear there is much scope for the methods which can be employed to charge subscribers for the services provided in future GSM mobile telephone networks. The technical challenges to capture the GSM based billing ticket data in order to charge subscribers are well known and documented. In addition to these it will be necessary to capture and process the network-related usage, using methods such as packet counting within and on the edge of the networks.

Support from the communication protocols in use on the mobile networks may be required to allow some of the charging models above to be implemented. This is needed to allow Quality of Service provision [14] within the networks and also provide identity tagging or other information based on the charging model requirements. Reservation protocols such as Resource ReSerVation Protocol (RSVP) may be used in the network to provide the support for QoS, and move away from the best-effort service as currently used on the Internet.

5. Conclusions and further work

Not all GSM mobile telephone networks may choose to charge their subscribers in the same way or in the same proportion for the new Internet services available with GPRS and UMTS technology. This is also true with 1st generation GSM mobile networks due to a variety of technical, commercial, geographical and political issues and concerns. In the mobile telephone network market it may make technical and commercial sense to adapt and combine some, or perhaps all, of the above charging models and additional ones into 'unified' flexible models which will cover the more diversified requirements of mobile charging. Figure 5 below shows an example of how the discussed charging models may be combined for charging for GSM voice and Internet services.

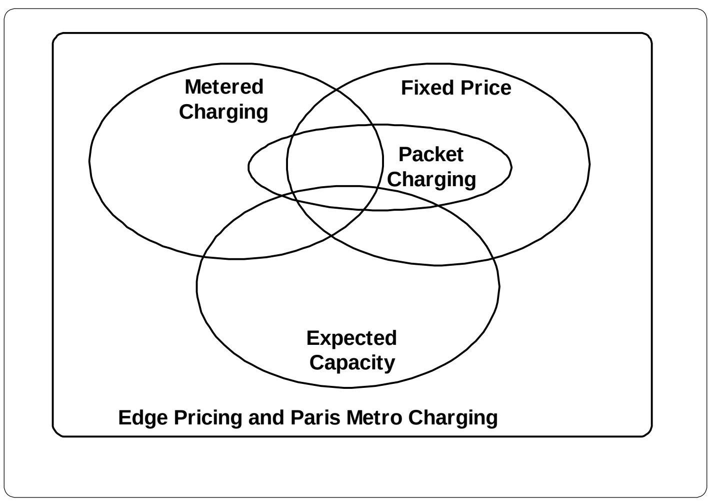

Figure 5 Combining Charging Models

A basic customer subscription may comprise a fixed price tariff that includes free mobile calls up to an agreed limit, plus Metered Charging for extra mobile calls. Packet Charging may be included for Web browsing and email using GPRS for the subscriber's account. Expected Capacity may also be included for GPRS data traffic as well as Edge Pricing and Paris Metro Charging for email and data traffic. Subscriber's requirements vary greatly from students to business users, from light domestic users to heavy personal users. By modifying and combining charging models tariffs can be developed to cover the major demographic groups, which make up the network subscribers.

The above combination of the charging models is a simple example and shows how a mobile network operator may choose to charge and invoice their subscribers. The illustration also shows the overlap between the various pricing models and how the boundaries can be made flexible depending on the subscriber usage profiles. For example, the mobile network operator may use Packet Charging in both Fixed Price and Metered Charging tariffs for some subscribers but only use Packet Charging with a Fixed Price charging model for other subscriber groups or tariffs. Some subscribers may only want to use limited Internet services, for example only text email and no Web browsing. The mobile networks may choose to implement charging models to take care of these limited service requirements

The tariffs for the subscriber may become complicated but may ultimately give the subscribers more control over the way they are charged for using the mobile voice and Internet services and the Quality of Service they receive from the network operator. The network operators will also have the advantage of being able to charge the subscribers for different level of Quality of Service for the different services and network provision.

Once experience of realistic network traffic in the next generation GSM networks has been gained and large quantities of network usage statistics have been collected and analysed, then more comprehensive or simplified charging models may be investigated, developed and prototyped from the ones discussed above. There will always be a trade-off between the complexity of the billing system to be implemented and supported and the advantage the network provider will receive for having the systems in place. Fixed price charging schemes reduce the overhead of the charging and billing systems infrastructure, as they tend to provide the simplest charging scenarios. Usage based charging models provide incremental and harder to predict income for the network providers as well as requiring high investment in the charging and billing infrastructure required. This includes increased cost in network traffic involved in the collection of the billing data required.

When the GPRS networks are up and running and the charging models are in place the economics of the market may take over and under- and/or over-provisioning of network resources may become apparent, especially with the Internet gateway provision. Charging and billing may then be used for congestion control as already proposed for the Internet and not just for recovery of costs for the network operators.

Further work in this area should include the mathematical modelling of the various charging models on simulated mobile network data covering both voice and Internet data services. This should include examining the combining of charging models and the resultant effect on the income for the GSM telephone network provider and also the cost impact on the different types of subscribers using the mobile telephone networks. The IETF and IRTF [15] currently have working groups on Authentication, Authorisation and Accounting with goals that are relevant to this research.

Acknowledgement: We are grateful to HP Labs for the support of John Cushnie's work through an Industrial CASE Award in cooperation with the EPSRC.

References

[1] GSM Association - http://www.gsmworld.com/

[2] Wireless Application Protocol - http://www.wapforum.org/

[3] Globalstar - http://www.gsmworld.com/

[4] ICO - http://www.ico.com/

[5] S. Keshav.: An Engineering Approach to Computer Networking, Addison-Wesley, 1997.

[6] G. Brasche, B. Walke.: Concepts, Services, and Protocols of the New GSM Phase 2+ General Packet Radio Service, IEEE Comms Magazine, August 1997 pp 94-104

[7] J. Cai and D.J. Goodman.: General Packet Radio Service in GSM, IEEE Comms Magazine, October 1997 pp 122-130

[8] IPv6 - http://www.ipv6.com/

[9] D. Clark.: A Model for Cost Allocation and Pricing in the Internet, MIT Workshop on Internet Economics, March 1995

[10] S. Shenker, D. Clark, D. Estrin, S. Herzog.: Pricing in Computer Networks: Reshaping the Research Agenda, ACM Computer Communication Review. 26 1996 pp 19-43

[11] A. Odlyzko.: Paris Metro Pricing: The Minimalist Differentiated Services Solution, AT&T Laboratories Research, April 1999

[12] Z. Fan and E.S. Lee.: Pricing for Differentiated Internet Services, Proc. IDMS'99, Toulouse, France, Oct 1999, published by Springer-Verlag LNCS 1718 (ISBN 3-540-66595-1) pp365-370

[13] A. Odlyzko.: The Internet and other networks: Utilization rates and their implications, Proc. 26th Telecommunications Policy Research Conference, October 1998

[14] A. Campbell, G. Coulson, D Hutchison.: A Quality of Service Architecture, ACM Computer Communications Review, Volume 24, Number 2, 1994, pp6-27
[15] IETF - http://www.ietf.org/ and IRTF – http://www.irtf.org/charters/aaaarch.html