EMERGING MODELS FOR MOBILE COMMERCE

A Consult Hyperion white paper

www.chyp.com

May 2002

Version 1

Contents

List of Figures

Please note that all figures and charts not created by the author specifically for this chapter were obtained from publically–available sources and every effort has been made to correctly label that source.

ABSTRACT

Getting from the general view that m–commerce will be huge to the specifics of which business models will come to dominate the sector is difficult. Firstly, no–one knows anything about mobile data and so we have little to go on to make predictions. Secondly, the speed and unpredictability of technological evolution make it hard to stabilise the platform for services. Thirdly, there are legal, political and social issues yet to be resolved. This paper attempts an overview of some of these factors and combines it with experience gained advising leaders in the m–commerce field to try and create a useful framework for setting business strategy for m–commerce.

1 Mobile Commerce

> “Consumers and businesses are beginning to shift their perception of the mobile handset from that of a voice telephony device to that of a personal e-commerce device.”

In the medium term, m–commerce is likely to account for a fraction of the volume of the mobile services (as shown in Figure 1) but operators hope that it will account for a significantly higher fraction of revenues as other services (eg, messaging) are commoditised.

What the m–commerce services will be, and how revenues will be earned from them, is hard to predict. In the absence of useful analytical techniques, one way forward is to develop a vision for these services to fit in to. Looking forward in this way, it is first necessary to distinguish between early experiences with wireless e–commerce and future m–commerce.

Wireless e–commerce, which has dominated the scene in North America, uses the mobile network and handset solely as a “pipe”: a cut–down version of the web accessed via a mobile handset. It does not make any use of the characteristics of the mobile network—location information, subscriber authentication or even calling line ID—to enable or enhance services.

M–commerce, by contrast, draws on the services provided by the mobile network and makes them an integral part of the commercial transactions conducted over the network. These transactions will be, primarily, those that are timely, simple and location–based [2].

1.1 M–COMMERCE TECHNOLOGY

This paper certainly isn’t the place for a detailed discussion of mobile communications standards, but it is important to establish the basic characteristics of the technology, as we enter the transition period between “2nd Generation” digital mobile maturity (2G) and the “3rd Generation” Universal Mobile Telephone Service (UMTS). In conceptual terms, the most important general trend is that the US market will converge with the European and Far Eastern markets (although the US market will remain a multiple standards environment, as shown in Figure 2, and China may have its own standard).

While UMTS is a year away, interim solutions (generally referred to as enhanced 2nd Generation, or 2.5G, services) are being deployed.The most significant of the 2.5G European services is the General Packet Radio Service (GPRS). This is the “always on” data service, deployed by most operators by the end of 2001, albeit in limited fashion and with a less than enthusiastic early response. GPRS may then be followed by Enhanced Data for the GSM Environment (EDGE), providing services at up to 384Kb/s, comparable with 3G, before operators begin to deploy UMTS using either W–CDMA or CDMA2000 standards. This evolution is shown in Figure 3 below.

1.2 THE 3G/4G COMMERCE PLATFORM

Having discussed the technology, it is important to note that the major transition from 2nd to 3rd generation mobile isn’t really technological—the general public don’t know or care about Subscriber Identification Modules (SIMs) or Code Division Multiple Access (CDMA)—but is actually mental. Consumers and businesses are beginning to shift their perception of the mobile handset from that of a voice telephony device to that of a personal e–commerce device: a personal, portable, convenient and effective transaction device.

There will be a wide variety of 3rd and then 4th generation services. Some future services are obvious—repeated surveys, including Jupiter (8/00), say that 49% of consumers want the mobile Net to deliver e–mail , and the evidence from i–Mode in Japan confirms the central role of messaging [5]. The only existing 2G messaging service (SMS, or “g–mail”) is a runaway success now exceeding 15 billion messages per month (note that worldwide e–mails exceed 100 billion per day) and has steadily come to dominate SMS traffic. Other future services are, of course, wholly unknown.

The story of SMS demonstrates clearly how all forecasts of 2.5G and 3G services have so far been overestimates, whereas forecasts of 2G services have been underestimates [6]. SMS traffic is now forecast to hit 80 billion messages per month worldwide in 2005 and there is already a migration path in view: SMS, enhanced SMS (eg, Ericsson EMS, Magic4 and Nokia smart messaging) and then some kind of Multimedia Messaging Service (MMS). Note also that SMS is no longer a GSM–only service: it was

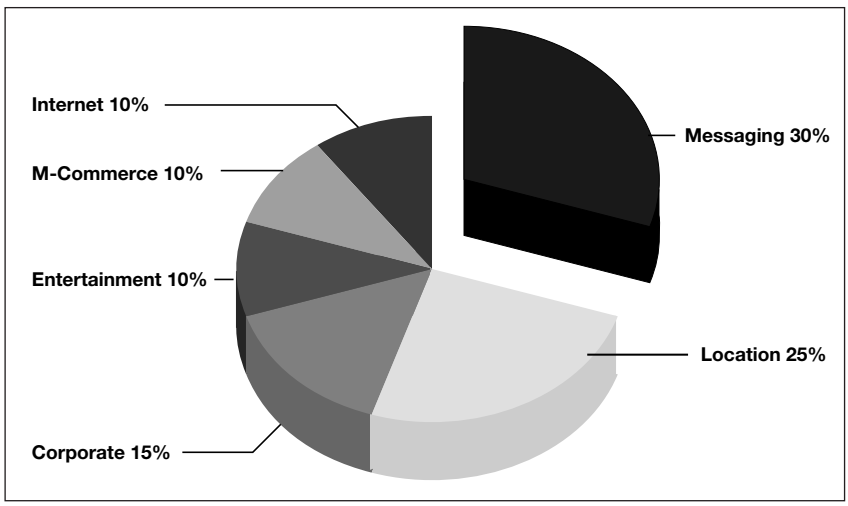

Figure 1. Mobile Services Forecast

(Source: Mobile Lifestreams, 4/01).

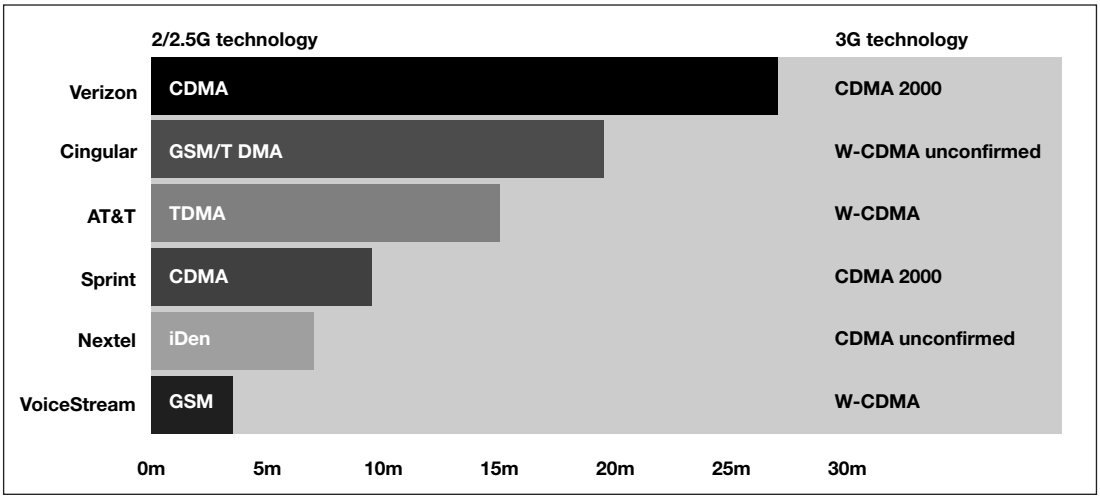

Figure 2. US 3G Technologies

(Source: Merrill Lynch, Operators 3/01).

implemented on TDMA networks last year and is now implemented on CDMA networks as well.

1.2.1 WAP Lessons

The transition away from the mobile as a voice device to the mobile as a transaction device has been reinforced by early experiences with the Wireless Application Protocol (WAP). WAP delivers web–like (or, perhaps, "web lite") functionality to mobile handsets. If enough interesting services are pushed through WAP channels then they may take off (although not, we suspect, until GPRS is widespread) but even then WAP will still lack many of the services necessary for commerce applications: effective navigation, no "cookies" or other session management mechanism, no way of identifying the handset, no end–to–end security, no push facilities and no location–dependent services.

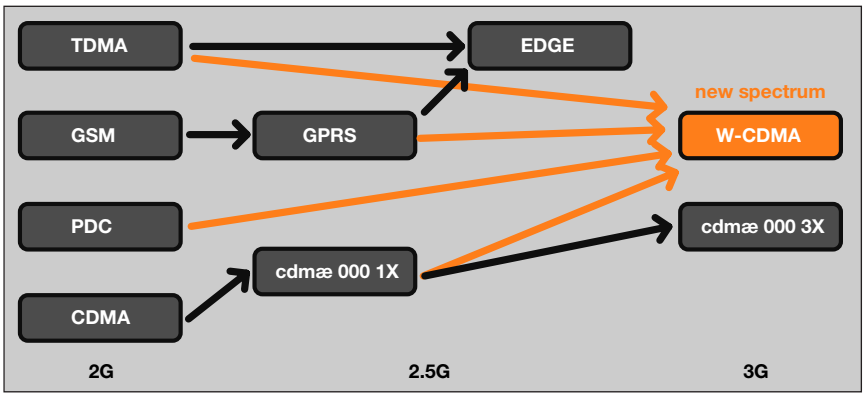

Figure 3. 2G to 3G Roadmap

GPRS (with volume–based charging) and WAP 1.2 handsets (which address some of the transaction issues) are on the horizon, but it takes time for organisations to gain experience and evolve strategies [7]. At the time of writing, the take up of WAP devices in Europe seems to be slowing and usage is low. There are a few interesting bright spots, such as mobile gaming: a third of new AT&T PocketNet subscribers are playing games on their devices, half of all Japanese mobile Internet users subscribe to at least one game service and the UK's Genie says that fantasy soccer league players log on via WAP at least four times per week [8].

While some observers think that the best strategy for organisations is to skip the current generation of WAP, as developing services for it will divert attention and resources from more promising next–generation services [9], we see it as a good way to learn. By using the limited functionality of WAP to begin experimenting with mobile interaction, organisations can develop more robust longer–term strategies to take advantage of the evolution of the sector.

1.2.2 Case Note: i–Mode

Japan's i–Mode, with around 21m subscribers, colour screens and an always–on network is a useful living case study of how consumers react to data services. In the year from its launch it attracted 4.5 million data customers (a sixth of all mobile customers in Japan). Its customers spend a flat $3 per month to get the service and add an average of $9–10 per month on (packet–based, as GPRS will charge) usage. NTT DoCoMo has amassed more than six thousand content providers and caters specifically to the uses that it finds are most popular among customers.

In terms of the actual traffic, entertainment again dominates, and of the sites that charge for their services (the charges are simply added to the subscribers bill, and passed on to the service provider minus a healthy commission) entertainment sites are likewise No.1.

In a clear window on the future, NTT DoCoMo have already signed a memorandum of understanding with Sony on the joint development of an i–Mode Playstation and Vodafone in the UK have followed suit [10].

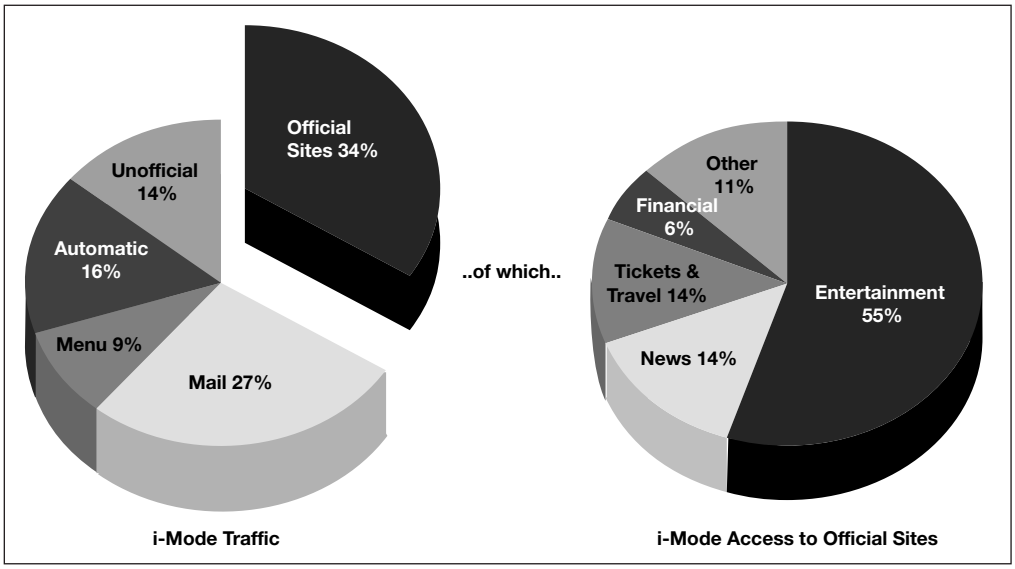

Figure 4. i-Mode Traffic Breakdown

(Source: Author's Analysis).

> "nothing can be called a service until you can bill for it."

Extrapolating the Japanese experience may be risky, however, and many of the things that the Japanese find useful on their mobile handsets may not make sense to Americans (with fast, cheap PC access to the Net) or Europeans (with inexpensive, simple TV access to the Net) [11].

The lesson that should be learned is that i–Mode's integrated billing architecture (Figure 5), which made it easy and convenient for customers to pay service providers, was crucial to the proposition. I–Mode recently announced their first moves to open up the billing system beyond network–deliverable content and subscribers can now buy from vending machines using their handsets.

1.3 CONTENT ISN'T KING

As the cost and capacity equations change with the growth of 3G, the mobile world needs innovation to come in and exploit the possibilities. The operators may not be best placed to make these innovations: the Economist Intelligence Unit forecast only 15% of future network revenues will come from network provisioning with the balance coming from services. The success of these services may depend on how personalised they are (to complement the mobile handset's nature as a personal device). The first thoughts of operators (and mobile portals) have been to develop "content" for mobile services, and this was originally forecast (Figure 6) to generate half of their revenues.

The realisation of such value–added revenues is not straightforward because the way in which value–added services are developed does, in the mobile world, have a certain tension associated with it. The operator "walled garden" (portal) in particular does not seem to provide an environment for the energy and innovation that we have come to associate with the internet world, while both operators and equipment suppliers are competing to define the infrastructure for future services.

Content models for operator portals are divided into three: portals with direct billing by the operators (the operator mall), portals where customers pay content partners directly but the portal gets a slice (the walled garden) and advertising or sponsored portals [13]. Since there has been no way for mobile users to simply and inexpensively pay service providers directly (outside of the i–Mode environment), that model hasn't been tried.

If there is a content–oriented strategy for m–commerce it may be in the form of aggregation. Web aggregators have been active for some time. Yodlee, to pick just one, can snatch data from more than 1,400 web sites (including major banks, stockbrokers, news sites and so on). They are not popular with (eg) banks—who refer to aggregators as 'screen scrapers'—because customers go to the aggregator and so never look at the advertisements on the banks site. But in the mobile case, you can't look at advertisements anyway. A single WAP page that gives your account balance, Visa balance, Air Miles balance and so on might conceivably be useful, but it seems to me that unless the aggregation portal provides easy entry to transactional services then it will eventually be bypassed.

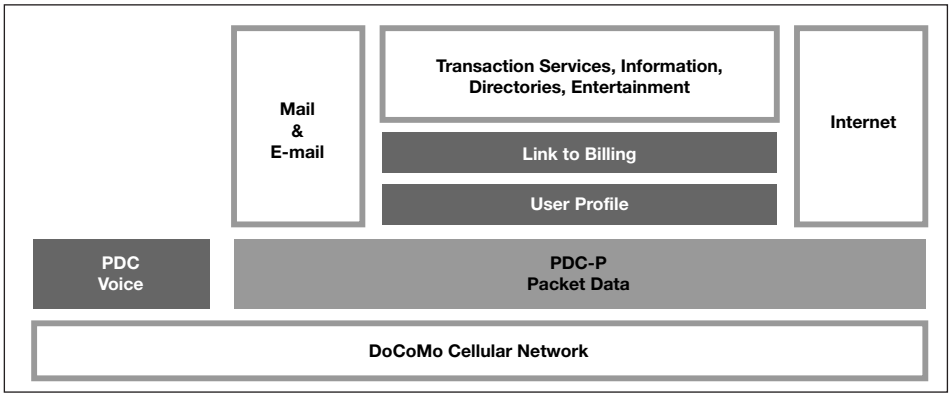

Figure 5. i-Mode Architecture

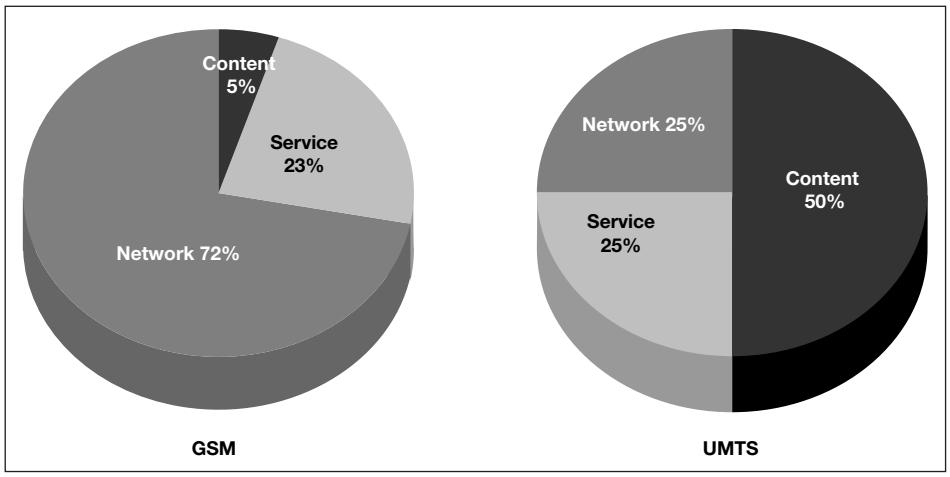

Figure 6. 2G/3G Value Shares
(Source: ADL/Alcatel, 2000).

1.4 MOVING TO TRANSACTIONS

1.4.1 New Business Model

If content isn't going to provide the core business model, then what will? One way to proceed with analysis is to note the old telecommunications maxim that "nothing can be called a service until you can bill for it". For mobile services to continue to evolve and to keep their dynamism, they must evolve standard platforms not just for the bearer and switching business but for third party services and it is already clear that personalisation and transaction management are where the highest added value opportunities are. Efforts to apply another level of infrastructure above mobile networks have so far been ad hoc and it is clear that the ability to, for example, collect micropayments for third–party mobile services has been a significant block on the development of such services. 3G services will never take off unless this problem is addressed [14].

Putting aside the payment problems associated with transactions, personalisation should be the first step although identifying the subscriber (for the purposes of providing persistent services) and delivering a consistent interface to them across multiple devices (eg, their mobile phone and their PDA) is, for the time being at least, very difficult to achieve. The proliferation of smart devices (see Figure 7) brings its own problems because of the sheer complexity of formatting the data: just trying to display the news can be a hassle, especially since most content isn't XML at the moment so the 'black boxes' that magically take web content to wireless don't solve the whole problem. Nevertheless, this kind of personalised information, especially when combined with location services, looks like being popular.

It is clear that service providers with more transactional services are finding a positive response from consumers. The standard European case study is Scandinavian bank Nordea, which has around 1.6m online customers and was one of the first banks to offer mobile service. Just to use one transactional service to illustrate the point: Nordea allowed customers to instruct SWIFT transfers via WAP. After only four months, some 20% of SWIFT payments come from this interface (helped, of course, by the fact that customers save by using WAP instead of going in to a branch) [15]. From early experiences such as these, looking ahead it is already clear that identity, authentication and payments will be central to the 3G and future 4G mobile network propositions [16]. These topics are covered in more detail in sections 2 and 3 of this paper.

1.5 LOCATION, LOCATION

It could well be that the most important factor in established mobile operators' positions in the m–commerce value chain will come from their control over the location information that merchants and others will need to provide effective mobile services [18]. Strategy Analytics predicts annual revenue from location–based services (LBS) in Western Europe will reach $9 billion in

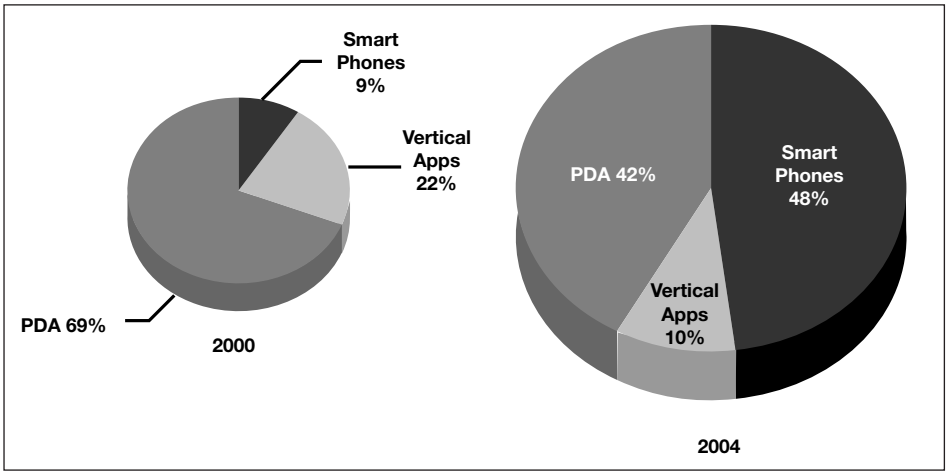

Figure 7. Smart Handheld Device Shipments
(Source: Initiative Europe, 11/00).

2005 (and another $7 billion in North America) but the m–commerce revenues dependent on these services may be significantly higher [19]. Similar figures come from Ovum (1/01), who predict that mobile location services will be generating around $5 billion per annum in 2004 (and by 2006 will have more than half a billion users). They also predict that it will be m–commerce that drives the mobile location–based applications. In the UK, for example, Vodafone has already launched a relatively simple service, based on cell ID, through its Vizzavi portal. In addition to the technological and business pressures for LBS, there are good reasons why law enforcement and government agencies want to see it introduced. In the U.S., federal law will require mobile operators to be able to identify the position of anyone making a call to the emergency services by the end of 2002. This "E911" initiative will create an instant mass–market for LBS in the US.

How the location of the device is established varies: it might be by GPS where there is a GPS unit built in to (for example) a laptop or it might be through the mobile networks. In the case of mobile network LBS, there are two ways to establish location: the handset can work out where it is by triangulating from base stations (which requires hardware in the handset) or the network can work out where the handset is.

What's not covered by current protocols or initiatives is the really big issue: who is allowed to access LBS and under what circumstances. Here, there is a vast difference between using GPS and using mobile networks because the GPS satellites don't know where you are but the mobile networks do. In other words, my device can obtain its location from GPS and then communicate this under my control. In the mobile case, however, the network can figure out where I am and pass it on to other people without having me in the loop at any point. If I call in sick, should my boss be able to call up a web page and see where I am? In the short term, it may be that the major beneficiaries will not be mobile operators or mobile consumers but lawyers!

"..smart card based payment schemes have the advantage that they can be used in both the real and virtual worlds."

2 Mobile Money

2.1 MOBILE PAYMENTS

For mobile operators and service providers, the issue of how mobile services will be paid for is crucial. It is also very interesting, since at the time of writing there are no clear winning strategies visible in the marketplace. The source of the problem lies in the evolution of the mobile business, where the operators have had to concentrate on transport services (for voice and messages) without having to worry about either the purpose or value of the transactions they are supporting [20]. This has a very negative impact. The real innovation in mobile content will come from small start–ups, but they will invest only if they can see a payback. At present, operator revenue sharing is the only way of getting paid and if they won't, the providers don't. A typical case was that of Citikey (a British WAP startup) which filed for bankruptcy because it couldn't capitalise on the services it was offering while the operators raked up the traffic revenues from users accessing Citikey's content [21].

It isn't just network services or value–added services (eg, messaging) that are the issue here. The vision of the handset as mobile transaction device is widely shared. Consumer transactions over mobile phones are estimated at $13 billion by 2003 and there are those who think that, in the long run, the SIM will be the only smart card that consumers need and the global telecommunications billing infrastructure will be the dominant form of electronic cash (e–cash) in the future.

2.2 PAYMENT TECHNOLOGY DRIVERS

While there are various digital money mechanisms that could be used for Net payments, smart card–based payment schemes—whether the smart card comes from the bank or from the operators inside the phone—have the advantage that they can be used in both the real and virtual worlds. People are already using both bank and operator smart cards for payment throughout Europe and there are already schemes in place using electronic purses (smart cards storing digital money) over networks, although these have so far failed to gain mass market acceptance.

There is tremendous interest in this field because (smart or otherwise), bank payment products are not necessarily optimal for the new mobile environment. The lessons from the Internet in this respect are interesting, particularly with respect to interpersonal payments

> “The story of SMS demonstrates clearly how all forecasts of 2.5G and 3G services have so far been overestimates, whereas forecasts of 2G services have been underestimates.”

where one might have expected electronic cheques and such like to be popular. New methods of making account–based transfers between individuals range from simple e–mail services to more sophisticated Net–based services that begin to overlap with digital cash and blur the boundaries between flexible account payments and cash replacement (on the Net, although not in physical commerce). Paypal was the first to take off and remains an excellent case study in the symbiotic relationship between new ways of doing business (eg, eBay) and new payment mechanisms.

2.3 PAYMENTS INFRASTRUCTURE

2.3.1 EMV

In a relatively short time almost all European consumers, and many others around the world, will have a bank–issued smart electronic payment card. Even in the supposedly smart card unfriendly US, the first few million are already in the hands of consumers. The driver is the replacement of dumb (magnetic stripe) credit, debit and charge cards with their smart equivalents. These cards will be based on EMV: the Europay–MasterCard–Visa standard. Most schemes for moving existing ‘dumb’ credit, debit and charge cards over to smartcards have declared EMV compliance to be one of their goals. Even France, where smart payment cards have been in use for many years, has decided to switch from proprietary standards to EMV.

The current plan is for all European payments cards to be replaced by smart cards by 2005. This mass issuing of EMV cards is relevant to mobile commerce. While the original reason for issuing EMV cards was fraud reduction (card fraud in France is substantially less than in e.g. the UK), the advent of the superhighway has changed the dynamic. A smart credit card can be used with a two–slot mobile phone to pay for a theatre ticket, for example, quickly and conveniently. Since the banks are issuing the cards anyway, many of them (perhaps even most of them) will choose to issue cards that have both EMV and electronic purse applications on them. It’s therefore likely that in only 3 or 4 years every European with a mobile phone with have an EMV card and an electronic purse as well, so that the full range of consumer payments can be handled in the mobile environment.

The synergy between the bank–issued smart payment card—and, for that matter, perhaps a bank–issued smart identity card in the future—and the mobile handset looks very powerful. The bank card provides security and payment services, the handset provides the most flexible and convenient platform for using those services.

Therefore, perhaps the most flexible approach comes through the addition of an external smart card interface.

2.3.2 Case Note: Iti Achat

France Télécom launched a service whereby consumers can pay for goods (which they have ordered using a voice service) by inserting their bank card into the external slot in their mobile handset. The pilot was called “Iti Achat”, a name that may still be in use, although the operational service is now called CB Payments on Mobile.

At the beginning of 2001 there were about 140,000 dual slot handsets in use in France and the take up was actually accelerating. This has led some observers to predict that, by the end of this year, there will be more mobile phones capable of accepting French banks’ smart cards than POS terminals in the country (ie, 500,000). The phones currently cost less than 500FF (for the WAP, tri–band dual slot phone) and this price is set to fall as more networks offer the service and more manufacturers provide the handsets (currently provided by Motorola and Sagem).

It is quite clear from the early results in France that by far the most attractive application of the combination of the dual slot phone and the smart bank card, as judged by consumer usage, is the loading of pre–paid mobile phones.

Overall, the success of this approach may depend on the ability of service providers (eg, banks) to fund the handsets with the additional slots as it is not clear that operators (especially given their expenditure on 3G licences) will have any business case for doing so themselves.

2.3.3 The Server Wallet

The server wallet seems to have become the main thrust in m-commerce payments. The idea of a wallet, somewhere on the network, that can store both payment information and other data (insurance data, health data, corporate data and so on) is appealing. While it may seem clumsy in the short term to have to enter user names and passwords via the mobile keypad, in the long term one can easily imagine voice and other biometric authentication being used to provide highly secure, highly convenient and flexible access to personal data. While some consumers are happy to carry phones, digital assistants, pagers and other devices, most are not. Therefore if the phone can be used to provide a single interface to server side data one might expect it to become popular with consumers.

> If the operator SIM can both identify and authenticate access to third-party systems, then the handset becomes a passport to a variety of other services (of which payment is only one).

But where will the consumers' wallet come from? Is it the mobile operator, the customer's bank or some other third party. There is a reasonable amount at stake in this discussion, because the wallet provider will naturally become a favourite online destination of consumers. Given the previous analysis, factoring in national technological and non–technological factors, it seems a fair prediction that in some national markets the operator might well become the provider of choice, while in other markets banks may be able to capitalise on their natural position as the consumers' trusted companion.

2.4 BILLING AS PAYMENT

Even in their current limited form, consumers do seem to be happy with paying for stuff via their phone bill. In iPIN's AT&T trial the 12,000 users spent an average of $6 per month, with the average transaction between $0.25 and $1 [23]. Consumers seem warm to a micropayment solution (actually a microbilling solution) that accumulates small charges to their bill at the end of the month. Operators, though, are still left with the problem of collecting the transaction data, billing, debt recovery and so on and so on. Passing the buck on to the phone company might seem attractive to a service provider, but it just isn't going to work.
Billing systems are expensive and complicated—figures are hard to come by, but the whole process of metering, billing and collection probably accounts 20–50% of the cost of a phone call—and to be effective they have to do a lot more than send out bills.

Even though early evidence suggests that consumers would use operator–provided microbilling solutions, these may still not be optimal because of the problems associated with extending billing systems. An example from the utility sector is cautionary. Independent Energy, a UK gas and energy supplier, went into receivership in September 2000 despite having pre–tax profits of £22m (up from under £2m the previous year) in March. The company had signed up a large number of small industrial and commercial customers but found it impossible to issue bills against a range of complex tariffs: by mid–year it was unable to collect from some two–thirds of its customer base. Cashflow was placed under serious pressure and the company ran out of working capital [24].

2.4.1 The New Purse

The use of the mobile handset as a payment device for impulse purchases at unattended POS may become very significant. The now–infamous Sonera Coke machine demonstrates how such a system might work and there are already other suppliers working to develop infrastructure. It isn't only operators, but third–party service providers who are pushing forward in this area. Coca–Cola and its' local bottling partners, to give an important example, are to invest $100m in bringing 500,000 vending machines online. The technology will allow customers to make cashless purchases and give bottlers greater flexibility in managing inventory. About 60,000 machines in the US, Australia and New Zealand are already on line [25].

The use of the handset as a payment device is spreading. Scandanavians pay for their car parking using their phones in several cities. Hundreds of thousands of French pre-paid users top up their accounts by slotting their bank card into their phones and Japanese consumers buy from vending machines by dialling them (lifting profits per machine by more than two-thirds in the process).

If operators do, in fact, decide that they will offer micropayments via their billing relationship (whether open account, pre–paid or whatever) while passing larger payments on to another account manager (a bank, credit card company or perhaps a specialist provider of lines of credit) then there may be no need for a second slot, chip or anything else. If the operator's SIM (WIM, SWIM etc) can both identity and authenticate access to third–party systems, then the handset becomes a passport to a variety of other services (of which payment is only one). This may be achieved through the MVNO route. Assuming that some kind of standard could be assembled between the operators so that consumers could roam and still make payments, then there has to be suspicion that operators could dominate the remote and unattended small retail e–payments sector.

2.4.2 Case Note: Movilpago

Movilpago is a joint venture between Banco Bilbao Vizcaya Argentaria (BBVA) and the mobile operator Telefonica. The system has been in pilot in Spain and was due to go live at the time of writing. The key requirements for the service were [28]:

- Use existing networks, handsets and SIMs.
- Create a new payment system, not a layer on legacy systems, capable for handling micropayments as well as large payments.
- Permit P2P as well as remote, unattended and physical POS.

The system can work in several ways. In a traditional merchant environment, the customer either tells the sales assistant their mobile phone number or (in larger retailers) allows the sales assistant to

> “Industry observers estimate that perhaps half of the $200 billion that banks currently earn from payments worldwide could be captured by new entrants but this income could be outweighed by new income in the online world.”

scan their phone using a special barcode reader. The POS terminal sends the phone number, a description of the goods and the payment amount to Movilpago.

Movilpago makes an Unstructured Supplementary Services Data (USSD: essentially a transactional version of SMS) call to the customers handset and sends the “invoice” and amount. The customer authorises the transaction by punching in their PIN code. All of this takes a few seconds.

The service costs the customers nothing and the charge to merchants (apart from the special POS interface if they choose to have one) is “comparable” to credit cards. The system currently offers two payment options to customers:

- A pre–paid network wallet (separate from the operators’ pre–paid wallet) that can be loaded manually or automatically.
- A post–pay (against a bank account) option.

There is an IVR alternative for handsets incapable of placing USSD calls, where the customer calls Movilpago on receipt of a payment instruction and confirms the payment by a voice call.

There is no doubt that Movilpago have the ambition to become a payment brand on par with Visa or MasterCard. They are currently planning to be in 30 countries by 2002 and are forecasting 100 million customers (and 5 million merchants) in the scheme by 2005.

2.5 OTHER PLAYERS

Other players are pursuing similar concepts. Paybox, for example, in Germany (where there are now nearly 50 million mobile phone subscribers and only 40 million mobile subscribers).

While banks may see it as imperative to provide retail payment for the new channels, their efforts to date have been poor. Several years after the web burst on to the business scene, credit cards are still the primary means of retail payments, with all of their attendant problems and limitations. Typing your name, delivery address and a 19–digit European debit card number into a WAP phone is hardly the most convenient way to buy a £1 lottery ticket (to pick an often–used example).

Banks don’t make very much money from payments. The top 25 banks in the US derive just 7% of their operating income from payment revenues. It’s also worth noting that economic theory on the operations of commercial banks cannot, by itself, explain why they provide payment services on such a large scale [29]. This limits their enthusiasm for investment in new payment infrastructure: this is bad for everyone, not just banks, because the economy needs the new products.

It is possible to see a near future payment sector, then, where operators offer micropayments via their billing relationship (whether open account, pre–paid or whatever). The operators will also authenticate users for larger payments that are passed on to another account manager (a bank, credit card company or perhaps a specialist provider of lines of credit). The bank brand is no longer seen at point–of–sale (perhaps the operators could get together and create their own version of Visa or Europay and the bank provides “white label” services to competing operators).

2.5.1 Case Note: Bank of Montreal

In the North American market, the Bank of Montreal’s Veev mobile banking service was one of the earliest entrants and remains one of the most advanced services in the world [32]. The service was launched back in May 1999 on a platform provided by 724 Solutions and Bell Mobility and then later in the year extended into the US via Harris Bank (a wholly owned subsidiary of Bank of Montreal) [33]. The service itself has steadily expanded from information access, through trading and on to an aggregation portal for customers’ financial information giving one password access to information from any of a customer’s financial institutions. The aim is to create a mobile financial services and “lifestyle” portal, and it already offers non–financial services including sports scores and news headlines as well as the usual stock listings.

At the time of writing, Veev is available over a variety of networks (including Rogers AT&T Wireless and Clearnet PCS) and is now available in both English and French.

The service provides access to a server wallet via mobile or internet connection and has the usual range of services (balance enquiries, bill payments and so on). Interestingly, in the market trial, 41% of the participants used the services in cars (!) and another 18% in transit. Less than 10% of access was from home and only 21% from work [34].

3 Mobile Identity

The idea of the SIM containing some kind of digital passport, which the mobile operator allows other service providers to use (as shown in Figure 9) is particularly appealing. This passport has immediate application as a means to authenticate server–side wallets both for payment and identity purposes. Note that since real mobile PKI, implementing the WAP 1.2 standard, is not yet here there is a need for some interim (SIM Toolkit, for example) security infrastructure to build on the GSM CLI plus PIN baseline [35].

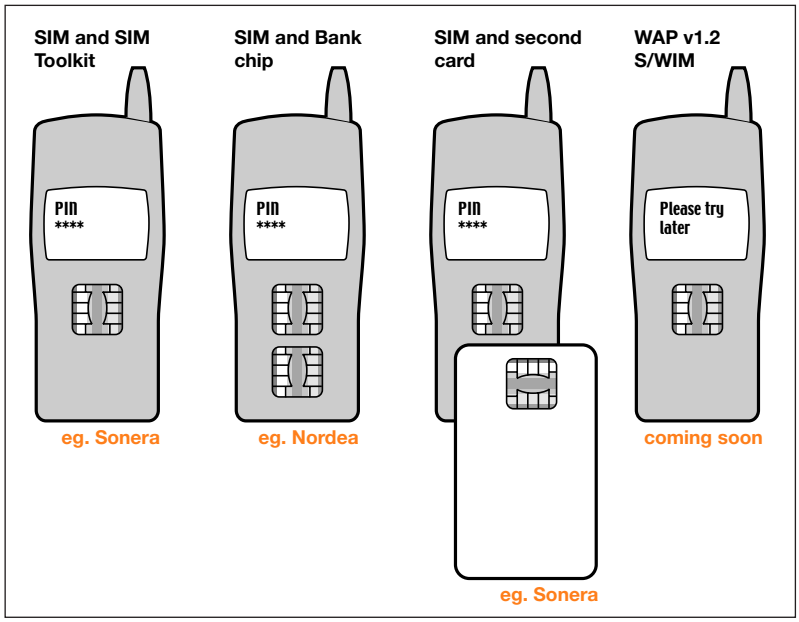

Figure 8. Handsets and Chips

In a few years time it is highly likely that every mobile phone on the planet will have one or even more smart cards inside it. Who controls these smart cards and the applications on them will have a profound effect on the shape and dynamic of the industry that is built on top of them. Drawing on a lesson from one of the early leaders—Nordea—yet again, it is clear that an important element of their successful strategy has been the integrated security platform. Consumers used the same log on procedure and security scheme (based on lists of code numbers) whether accessing via PC, WAP or any other channel. This strategy has now been 'exported'. Sonera, the telecoms operator, allows customers to 'sign' orders using their Nordea security codes: this model is a clear precursor of the more generalised role of the mobile handset as an authentication device.

Figure 9. Mobile Operator acting as Registration Authority

3.1 MOBILE PKI

Many Wireless PKI components from vendors are still in development, and there is also a reluctance for vendors to develop products in new areas where agreed standards are not in place (eg, to put in place extended functionality between the WAP browser and smart cards in a 2nd slot or transferring URLs pushed via SMS to the browser). Future versions of WTLS will offer:

- Authentication of the user. By authenticating the link (initially established by PIN, later by biometrics) between the user and the telco digital ID stored in the handset/SIM, the operator at a stroke removes a whole raft of problems for service providers.
- Encryption of individual messages (eg, for secure email) and the digital signing of messages. In the UK, at least, there are very few contracts that require a written signature on paper hence the potential for widespread and early adoption.
- Push alerts to handset.
- The Wireless Identity Module (WIM).

The WIM will provide a secure container for security data such as cryptographic keys. The WIM is, of course, only one of the options for connecting a smart card holding encryption keys with a handset (as shown in Figure 10). It may be part of the SIM or a separate smart card chip either internal or external to the handset (ie, 2–slot phones). It is highly likely that the WIM functionality will be integrated with storage of other sensitive data, such as payment, or personal data subject to the EU Data Protection Directive, and will be the basis of a personal data container or "passport" for Net use.

3.2 IDENTITY SERVICES

One obvious business for a mobile operator is that of a Registration Authority (RA), the body linking key pairs and certificates to real users. Since the operator knows its customers (leaving aside the issue of pre–paid mobiles for the moment) it is relatively easy for them to convert the existing customer relationship into a relationship involving keys and certificates. I might, for example, have special PIN number printed on my mobile bill that enables me to go to a web site and get a certificate issued testifying to the fact that I am a subscriber. The mobile operator would

> “The use of the mobile device as a transporter of keys and certificates is likely to be central to future mobile propositions.”

not necessarily need to run the Certification Authority (CA), the high–security business linking key pairs into certificates behind one or more RA operations, and might prefer to have this outsourced to specialist CA service providers (banks, even) or perhaps come together to pool resources in establishing mobile sector–specific services.

If the operators were to focus on RA services—integrated into existing business structures (eg, call centres)—and focus on shared, sectoral CA/CR/CRL services, then it ought to be possible to envisage a number of lines of the business that could come to market quite quickly and both provide real income and support e–commerce and other services.

Services that implement PKI to support real services in the mobile environment already exist. Finnish broker eQ Securities and Sonera SmartTrust launched the eQ Free mobile broker service earlier this year, using PKI to provide security on both WAP and non–WAP phones. In Hong Kong, where some 40% of handsets will be WAP by the end of this year, Charles Schwab is launching a similar service to trade Hong Kong and US stocks. Confirmations will be in real–time, with a 13 second target to complete a mobile trade [37]. There is little doubt that the deployment of end–to–end mobile security, based on PKI, will engender a range of real business services. Note that these services are not all retail consumer services and the corporate sector is likely to be a big user of Internet–mobile integration, but who is going to make their corporate address book and calendar available through WAP unless there is an appropriate degree of security associated with the new channel?

It would be relatively simple for an operator to issues certificates, whether for fixed or mobile use, that would become indispensable online passports for customers. But if (as discussed earlier) the best structure would be for operators to focus on RA activities and share other elements of the infrastructure, then it’s worth exploring the different RA models that might work. Whether the CA is operated by an operator consortium or by some joint venture between operator and other sectors is also interesting but outside the scope of this paper. There are at least three RA scenarios, as illustrated in Figure 9:

- Anyone could turn up at the operator’s web site (or WAP site) and present a public key for signing. Thus, much as BT’s Trustwise does now in the UK, the operator is issuing anonyms.
- Subscribers could turn up at an operator web site and present some PIN code or password that the operator has mailed them (to prove that they are a subscriber) and be issued with a pseudonym.
- Customers of third parties (eg, banks) could turn up in person to have their certificate issued. If the keys and certificates are being carried around in mobile phones or smart cards then this is the way to deploy a very secure and very flexible infrastructure.

The mobile operator would immediately become a major user of the certificates, since it would want to shift as much customer service activity as possible to the Net immediately and the certificates would provide a convenient (setting aside issues of consumer education…) means to facilitate this.

Self–provisioning of tariff packages and so on would be made more cost–effective. The use of the mobile device as a transporter of keys and certificates is therefore likely to be central to future mobile propositions. DoCoMo has said that it plans to bring forward the launch of its 4G service—in which mobile handsets operating at 20Mb/s will act as authentication devices for all kinds of new applications—to 2006 [38].

Note that there is already a Mobile Electronic Signature Consortium (MESC)—including smart card manufacturers (Gemplus, Schlumberger etc), software companies (Brokat etc), banks (Hypo, Bank of Tokyo etc) and others—working to standardise these kinds of communications between the mobile operator and third–party service providers.

3.2.1 The Privacy Play

Digital IDs provide a particularly effective way for third parties to validate credentials: if I have a certificate from British Airways (BA) giving my Executive Club number, a travel agency web site can easily check the validity of the certificate using BA’s public key. As noted, there may be significant business opportunities for the issuers of credential (rather than identity) certificates [39].

Most current schemes—the banks’ Identrus scheme and government schemes such as, for example, the Finnish national scheme—are predicated on the binding of digital IDs to the real (authenticated and unambiguous) identities of individuals and organisations. But the idea of a single digital ID per person is in many ways unappealing. PKI should allow people to choose which digital ID to use

in different transactions, thereby partitioning their identity. Also the ID should be visualised as supporting not just one single identity per individual but as supporting multiple identities related to all forms of personal, social, professional, commercial and corporate identities. Each of these identities may have multiple attributes, credentials and reputation attached [40].

3.2.2 Case Note: Vodafone Airtouch UK Government Pilot.

In March 2001, Vodafone began a pilot project with the UK government to use mobile digital signature technology for a real service. The pilot, involving an Executive Agency of the Department of Trade & Industry (DTI), used SIM–based digital signature techniques—in this case using the Sonera SmartTrust platform, as shown in Figure 10—to allow users to browse and make selections using a WAP site and then digitally sign information sent from the WAP site via the SMS channel. The programme uses the GlobalSign certificate authority, in which Vodafone has a significant stake.

In this pilot programme, fifty government staff completed and electronically signed their travel or subsistence forms while away from the office, using a Siemens handset fitted with a signature–enabled SIM. When prompted, users entered a PIN number to release a cryptographic signature that is more secure than handwritten signatures as it guarantees the signature is unique and ensures the integrity of the data being signed.

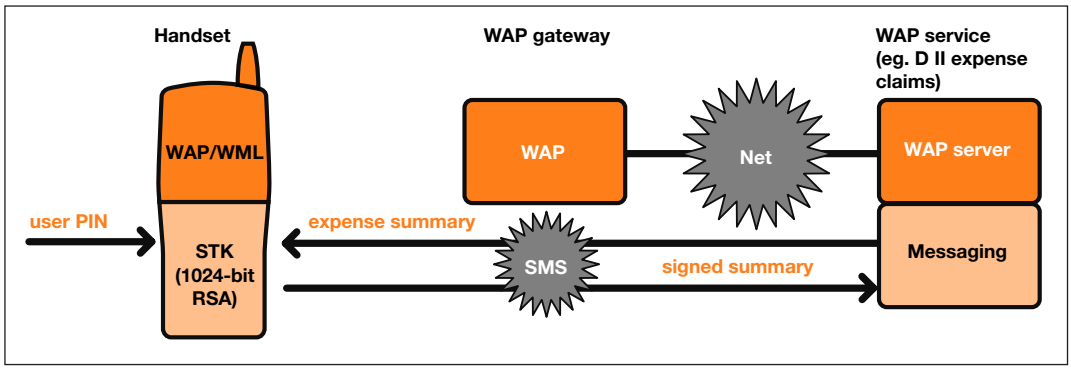

Figure 10. Vodafone/Smartrust Schematic

Conclusion

The infrastructure for future m-commerce services is about digital money and digital identity. Without these key elements in place, mobile services of all kinds will fundamentally limited. A recent global survey of mobile users showed that the gap between the services that consumers access and the services that they would like to access is as wide as ever. In the UK, just 4% of mobile users did something transactional with their handset last month (presumably something involving reverse-charged SMS) while three-quarters said that they would have liked to. Digital money and digital identity implementations can exploit fundamental properties of the 2G, 2.5G and 3G networks and devices to provide an enabling platform for a whole new landscape of mobile, mobile-enabled and mobile-enhanced consumer, business and even government services.

References

1. GPRS services to have only 2 million users by year end, says report in Wireless Internet. (20): p. 11 (1st Feb. 2001).
2. Schmidt, C. Get the Low Down on Consumer Trends in proc. of Mobile Content and Entertainment, MforMobile (Cannes: May 2001)
3. Birch, D. 4th Generation Services in proc. of Information Society Technologies, EU (Helsinki: Nov. 1999)
4. Williams, M. News Goals for PHS in Telecom 99 Daily section): p.13 (16th Oct. 1999).
5. Lai–Hung, J. i–Mode in Europe—Help or hindrance in proc. of Mobile Content and Entertainment, MforMobile (Cannes: May 2001)
6. Short message service propagates to new markets in WAP Week. 2(3): p. 12–13 (Apr. 2001).
7. Karliner, M. WAP isn't yet right, but the time is still now in Internet Markets. 1(5): p. 12 (23rd Jun. 2000).
8. Mobile gaming set to score far beyond hard–core community in Mobile Internet. 1(13): p. 1–2 (6th Oct. 2000).
9. Nielson, J. The Backlash Begins in The Industry Standard. p. 116–118 (14th Aug. 2000).
10. Vodafone mobiles to link up with Sony Playstation in The Industry Standard Europe. (12th Feb. 2001).
11. In the laboratory in The Economist. p. 101–102 (2nd. Dec 2000).
12. Nokia touts mobile internet spec in Wireless Briefs. p. 1–2 (22nd Nov. 2000).
13. Tiller, A. Billing on the run in European Communications. p. 89–91 (Spring 2000).
14. Jagoda, A. 3G Business Services and Applications in proc. of Mobile ISP Forum, IIR (London: May 2000 2000)
15. Harald, B. The EMPS Concepts in proc. of Proton World Forum, (Ghent: Sep. 2000)
16. Birch, D. My generation in Global Telecoms Business. (53): p. 38–40 (Nov. 2000).
17. More Phones Will Use SIM Cards in Card Technology. p. 14 (Jul. 2000).
18. Birch, D. It's location, location, location in The Guardian (Online section) (14th Dec. 2000).
19. Rockhold, J. Where It's At in Wireless Review. (30th Nov. 2000).
20. Wingfield, M. What's the service, where's the bill in Billing International. p. 10–13 (May 2000 2000).
21. Johnson, S. Sharing mobile wealth in The Industry Standard Europe. p. 61 (18th Jan. 2001).
22. SET redundant, say Europe's cards executives (Laffety) at <http://www.lafferty.com/> (on 29th March 2000 2000).
23. Davis, J. Spare Change Agents in Business 2.0. p. 204–214 (October 1999).
24. Taylor, A. Terminal failure prompted by pursuit of small customer in Financial Times section): p.16 (9th September 2000).
25. Liu, B. Coke is IT in Marconi vending machine deal in Financial Times section): p.28 (18th May 2000).
26. Helanto, P. The Opportunities and Challenges for Operators of Enabling Payments via Mobile Terminals in proc. of Mobile Commerce European Congress, IIR (London: Oct. 2000)
27. The fall and rise of the mobile virtual network operator in Roam. (1): p. 30–32 (February 2000).
28. Movilpago in proc. of Mobile Banking, European Financial Managers' Association (Paris: Jan. 2001)
29. Radecki, L. Banks' Payment–Driven Revenues in Federal Reserve Bank of New York Economic Policy Review. (forthcoming) (July 1999).
30. Ayres, C. Orange plans to launch virtual bank in The Times section): p.27 (14th July 2000 2000).
31. Jefferey, C. Banks face four possible outcomes in Electronic Payments Intl. (161): p. 10–12 (Nov. 2000).
32. Horton, C. Take it to the bank in Mobile Communications International. (78): p. 101–110 (Feb. 2001).
33. Ryan, V. Budding developments in e-commerce in Telephony. (27th March 2000 2000).
34. Dickelman, M. Bank of Montreal Veev Initiative: Access Anytime Anywhere in proc. of M–Commerce for Banking, IQPC (London: June 2000)
35. Sporn, U. Leveraging Mobile Commerce by Signature Services in proc. of Secure Mobile Commerce, IBC (Paris: Jan. 2001)
36. Springham, J. Keeping the airwaves secure in Mobile Europe. p. 17–20 (Mar. 2000).
37. eQ begins mobile broker service in Virtual Finance International. 00(03): p. 3–4 (9th February 2000).
38. DoCoMo to start 4G four years early in The Guardian section): p.25 (19th Mar. 2001).
39. Birch, D. Digital Passports in Parliamentary IT Briefing—IT and Public Policy. (6): p. 6 (February 2000).
40. Shaw, S. Reflecting Relationships with PKI in ID World. 1(2): p. 21–30 (May 1999).

About the Author

dave@chyp.com

David G.W. Birch is a Director of Consult Hyperion—the IT management consultancy that specialises in e–commerce—which he helped found after several years working as a consultant in Europe, the Far East and North America. A physicist by training, Dave has lectured on the impact of new communications technologies to MBA level. He has been on the editorial board of the Financial Times Virtual Finance Report and Microsoft's Finance on Windows, as well as the editorial advisory board for European Business Review. He chaired the Centre for the Study of Financial Innovation's first working group on the Internet and Retail Banking and is the moderator of First Tuesday's Wireless Wednesday resource for mobile entrepreneurs. He has written for publications ranging from The Guardian to the Computer Law and Security Report and is a media commentator on electronic commerce issues.

Our Business

Consult Hyperion is an independent information technology consultancy. We have over sixteen years experience of helping organisations to derive maximum benefit from technology. Our success lies in our ability to apply a unique blend of business and systems analysis skills to deliver a strong interface between business and technology.

We specialise in transactions and security, providing professional services from the conceptual stage of a project through to completion. We advise on business and technical strategy, performing requirements analysis and producing technical specifications for internal development or external procurement. We can also help to dramatically shorten project life cycles by building proof of concept and pilot services.

Our practical experience with key technologies - including cryptography, smart cards, wireless devices, messaging, internet and digital TV - and more than a decade spent working with international blue–chip clients on innovative projects, make us uniquely qualified to advise on turning great business ideas into working systems.

Consult Hyperion
8 Frederick Sanger Road, Surrey Research Park, Guildford, Surrey GU2 7EB, England
T: +44(0)1483 301793 **F**: +44(0)1483 561657 **E**: info@chyp.com **W**: www.chyp.com