Mobile Commerce Regulatory Frameworks under Digital Convergence

Jari Veijalainen

University of Potsdam, Hasso-Plattner-Institute, Prof.-Dr.-Helmert-Str. 2-3
D-14482 Germany
j.veijalainen@hpi.uni-potsdam.de

Abstract
Mobile Commerce can be seen as a special case of E-commerce, conducted using portable wireless terminals. The terminal base ranges from laptops to smallest mobile phones and further wireless gadgets. The wireless terminal base uses a certain wireless spectrum portion that is regionally regulated by authorities and the ways of usage standardized by industry consortia (e.g. 3G standards) or government. The end-to-end protocols are regulated by de facto standardization bodies (TCP/IP by IETF). The higher layer protocols are standardized by industry consortia but not by authorities. The digital contents and services that are the actual object of trading in M-commerce are again regulated by regionally valid legislation, including E-commerce legislation, privacy protection, and regulations concerning harmful and criminal contents. All these are factors that make the M-commerce market fragmented. A further reason is that mobile commerce happens from machine to a human user. That is, the information at the user interface of the device and the contents must be understandable in a natural language. This imposes a further fragmentation upon the M-commerce market and makes offering Location-Based Services (LBS) problematic for visitors. Digital convergence is just about to take a new step when mobile-TV is being introduced into mobile terminals. The TV-contents has again its own regulation that is different from the above regulations. This causes tensions within and between the regulatory bodies. The new portable terminals can also produce huge amounts digital contents including pictures and videos. Should these be regulated in TV or M-commerce framework, if they become object of commercial activity? The central conclusion is that authority regulation needs to be re-organized. But how?

Keywords: M-commerce regulation, Regulatory Frameworks, Mobile-TV

1. Introduction

The term Mobile commerce (M-commerce) was coined towards end of 1990’s soon after the term Mobile Internet was introduced, meaning that “Internet” and its contents and services would become accessible from wireless telecom terminals. The term M-commerce was introduced without a clear meaning and it is still lacking a single widely accepted definition. For most people the term M-commerce refers to electronic commerce (E-commerce) activities performed by people “while on the move”. Thus, M-commerce involves E-commerce transactions (OECD, 2002; Veijalainen, 2003), where a mobile terminal and a wireless network are used to conduct them. M-commerce takes thus advantage from the E-commerce infrastructure developed for Internet E-commerce. M-commerce transactions might indeed be an alternative to a regular E-commerce transaction (such as buying a book) performed using a workstation and wired network, but the limitations of the smart phones/PDAs, for instances user interface limitations, are such that it is not always attractive to perform typical Internet E-commerce transactions on them. Another limitation is the
transfer capacity and cost of the many existing wireless technologies although the former is not anymore significant for the High-Speed Downlink Packet Access (HSDPA) (ETSI 2004), Wi-Fi or WiMAX (IEEE, 2004) links; the practical transfer capacity for these is roughly 10 Mbps.

The more and more powerful mobile telecom terminals and PDAs are almost always carried by their owners and they can be positioned dynamically. Using the actual position of a person makes it possible to offer him or her services and contents that are related with the place he or she is in. Location-Based Services (LBS) are a new type of E-commerce transactions that are not possible or reasonable for wired terminal. Using the actual position of a person makes it possible to offer him or her services and contents that are related with the place he or she is in. In a more general setting, not only place, but a larger context that might be related with the current place of the user (train, car, boat, home, church...) or otherwise further specified by sensors and/or the user him-or herself, can be used to tailor the offered services and content. In addition, because the terminals are carried by the people, they are just like wallets or watches. They can indeed also store electronic cash, credit card information, tickets, certificates of the Public Key Infrastructure (PKI), etc. Thus, they can assume the role of an e-wallet, as well as function as authentication and authorization devices in various business and other situations. This kind of “close-proximity” M-commerce is another potentially new form made possible with the highly portable terminals. NTT DoCoMo has been strongly developing ways to apply their terminals in these new ways. Lawson and NTT DoCoMo will introduce during spring 2006 mobile terminal-based payments in Lawson convenient stores. The DoCoMo’s iD™ credit card payment and Osaifu-Keitai™ phones with wallet technologies will be applied in these contexts (NTT DoCoMo 2006). ToruCa™ information capture service will function within a year in all 8300 convenient stores of Lawson. Earlier, NTT DoCoMo has already earlier made contracts that make it possible to use the phones as train tickets (NTT DoCoMo 2005).

We view regulation to be an enabler and limiting factor for an environment. The inherent structure is such that there is an organizational body (government, standardization body, industry grouping) that controls in a form and/or usage of certain technological artifacts. The limits of the controlled domain can be given by the technology or by a geographic region or both. An example of a “global” standard is the IP protocol the design of which is controlled by IETF (IETF, 1981; Johnson, 2004). An example of a national wireless standard is e.g. the (essentially Chinese) 3G standard TD-SCDMA (Kowalke, 2006; TD-SCDMA Forum, 2005).

A more extensive analysis of the permanent structures in M-commerce environment is described in (Veijalainen 2003). In this paper we investigate the currently happening changes due to convergence and more detailed inherent structures in the current M-commerce environment. We first describe in section 2 convergence phenomena. Starting from the perspective of a terminal we subsequently investigate in section 3 which M-commerce services it reaches at various layers. These are essentially determined by the regulatory, business, and technical borders. In section 4 various regulatory settings and policies and their effects on M-commerce are discussed. Section 5 concludes.

2. Digital convergence and the borders of regulatory spheres

Digital convergence is a techno-social process where analog information processing, transfer and storage are replaced by their digital counterparts. Digitally encoded information has many advantages over analogically encoded information: a common format as sequences of bits, error correction while transferred and stored, encryption, and compression. Processing instructions, i.e. programs, are coded as bit strings, like any other data, and add flexibility and intelligence. It also enables the use of the same components in very different applications and
in greater production runs. 2G and satellite TV networks started using digital techniques in the 1990's, and currently the process of converting terrestrial radio and TV networks into their digital equivalents is under way (DAB, DVB-T, DVB-H). The third generation cellular networks (3G, cf. IMT-2000, UMTS) are being implemented and deployed (Kaaranen 2001; ITU 1996) and the use of 4G is under active investigation. First of the 4G types have been adopted in Korea in the form of Portable Internet. Networked computers, utilizing the Internet Protocol, have created a global delivery channel for music and other digital content, bypassing the conventional system of delivering physical copies. The most prominent examples are the I-Pod and the I-tunes services. TV and radio broadcast through the Internet is also commonplace and expected to grow and possibly replace traditional TV channels with Internet based TV. Cognitive Radio (Cognitive Radio, 1999) can also send and receive digital information over a large spectrum of radio waves.

One visible development is the convergence of terminals. This means two-fold development. First, new functionalities are being added to voice telecom terminals. These include digital cameras, mobile TV, FM radio, media player, various sensors and readers, and numerous applications that can be packaged into a "multimedia computer". Second, the terminals run many protocol stacks on the lower OSI layers (2G, 3G, Bluetooth, IrDA, DVB-H, etc.). They also integrate many network, transport and application protocols (e.g. TCP/IP+HTTP) facilitating thus the “Always Best Connected” (ABC) principle and a wide range of application support.

Terminal convergence is the first step in the network convergence. In the next step the existing wire-line infrastructures might merge to a backbone network based on the Internet Protocol, resulting in a true network convergence. This is much more difficult, because of the high invested value of the current infrastructure as compared to the terminal convergence. The ultimate end of this development might be that all wire-line and wireless access networks use the same packet based backbone to transmit any possible digitized information, be it data, emails, or streaming data like video or audio, images, etc. This would evidently have a devastating effect on the mobile and other industry structures and require completely new thinking in regulation related to different information industries associated with the change (media, banking, voice, etc.)

The current environment is represented from terminal point of view in Fig.1. It shows how the same terminal can be a citizen in many wireless networks. It means that for each arrow coming or going to the terminal there are possibly a piece of special hardware and a protocol stack (software) supporting the communication. The two-way arrows mean two-way communication, one-way arrows one-way communication. The protocol stack and the overall system architecture are controlled by the actors mentioned in the ovals. This view covers primarily OSI-layers 1-3 (physical, link, and network layer).

The convergence of terminals opens up new possibilities for M-commerce. So far, M-Commerce has been based on specifically tailored services and contents available in Internet, or directly on contents and services designed for PCs (laptops). Now also TV-contents can be received wirelessly either through special broadcasts (DVB-H, DMB), or over the mobile networks. Because the content is digital, one can set up back-channel through a URI that can refer to an M-commerce server. A similar idea has already been implemented in Japan in radio context (Funk, 2004).

A possible future environment is represented in a layered way, as shown in Figure 2. The idea is that there are less wireless network types than now, perhaps only three or four; satellite navigation system, some kind of successor of a broadband Wi-Fi, and a short range radio network, successor of Bluetooth. Perhaps Infrared systems will also survive. Basically, different actors will control the structure of each layer. We assume in the figure that the protocol layering is still close to that of current Internet. The bottom layer consists of the
physical layer, in the wireless case air interfaces, and the link layer. For wire-line access the authorities do not specify an allowed protocol stack.

On the contrary, in wireless case the spectrum allocation is currently controlled by international body (WRC) that determines which frequencies are allowed for which use globally. These air interface standards are specified by multiple industry consortia or individual actors (such as ETSI, 3GPP, IEEE, Bluetooth consortium, Chinese government etc). The governments control the usage of licensed and unlicensed spectrum in a certain geographic area and the deployment of the networks. Changing this scheme, e.g. allowing more unlicensed spectrum and Cognitive Radio-based terminals that dynamically choose any frequency between 100 MHz and 2.5 GHz, would probably radically change the conditions of various wireless actors and competition between them. The Software radio technology is being developed and soon deployed by US DoD (Wilson, 2006).

Various government entities and NGO authorities currently control the higher layers in the protocol stack. IETF and XML family of formats by W3C controls TCP/IP, UDP and
HTTP. Multimedia formats are controlled partially by industry (e.g. Microsoft video format, Real-audio), industry consortia (DRM, OMA; see (OMA, 2006)), and regional standardization bodies (DVB-A). We believe that these protocols or their successors will survive into the future.

Intellectual property rights and privacy are established by legislation (i.e. regulated by a nation state as a regulator) and technically supported by varying technical mechanisms, such as DRM and encryption. User authentication and authorization at the highest layer can be based on PKI scheme and Certification Authorities (2G and 3G networks have a standardized authentication scheme of their own for terminals at network layer) that are national.

3. Accessibility Spheres of an M-commerce Terminal

Regulation and markets set various limits for interoperability between wireless terminals and servers in the environment. Historically 1G and 2G wireless networks developed into different directions in Japan, USA and Europe. GSM developed in Europe has become a truly global standard in the sense that networks compatible with GSM the standard(s) are deployed all over the world. 3G networks are not really unified either, because IMT-2000 standard allows several incompatible standards (W-CDMA, CDMA-2000 and SD-SCDMA) to be used at the air interface (ITU, 2000).

The air interface is just a necessary condition for a terminal to access M-commerce servers. There are further levels that must be interoperable. We discuss them below.

3.1 **Globally accessible M-commerce servers**

Let us assume that we have a particular wireless terminal T with certain protocol stacks, software to process certain content formats, and authentication mechanism. Let us assume that it in a coverage area of a wireless network X. Let us further assume that M-commerce transactions are only facilitated by server components accessible through an IP number. What is now:

1. The set of all possible servers that can communicate with T over TCP/IP (TCPIP, X)
2. The set of servers that have interoperable protocol stack up to the application layer including contents formats (e.g. HTTP(S)+ HTML/XML, certain multimedia contents) (SHTTP+cont,X)
3. The set of servers that can perform an M-commerce transaction with T, i.e. supporting payments and possibly authentication and authorization, and allow contents to be delivered to T or to user (cf. taxi) accessing the server from X (SMC-t,X).
4. The set of servers that deliver such contents that can be understood by the user of T (Suser-understandable,X).

If we analyze the situation, it can be that X is such that T cannot access it. This can be for technical reasons (no common bandwidth, missing wireless data transmission support at the network or at the terminal, incompatible wireless data protocol stacks at X and terminal T, etc.) or for business reasons (no roaming agreement between the operator of the home network and the operator of X the terminal is stolen, etc.). In this case all the sets above are empty.

If the terminal T is able to access X and run TCP/IP towards some servers, then STCP/IP,X is not empty. In general it holds that level n+1 set of servers is a subset of the level n set of servers. Thus, Suser-understandable,X is the smallest subset among the servers. This is of course relative to the capabilities of the user, because different people can understand only one or few languages (at least to such extent that the contents or service is sufficiently
understood by the user to make sense to purchase it). English is currently the most spoken foreign language, so if services or contents are offered in a language different from the language of the domicile of the server, then the best idea is probably to use English.

Are the sets of servers above dependent on the terminal only? No. If the same terminal and user accesses IP network through another wireless network Y (Y being, say, Wi-Fi instead of W-CDMA), then the set STCPIP,Y might change from STCPIP,X. This is because the W-CDMA operators might have such servers in its IP network that are only accessible from its wireless network (in Fig. 1 we have therefore a separating line between operator-internal IP-network and the common Internet). Second, some operators might limit the access to certain Internet sites (IP addresses) for various reasons. Third, the servers might allow access from certain wireless networks, but not from all. Thus, changing the access network from X to Y changes the set of accessible servers in general. The change can happen at the...
spot, i.e. from the same geographic location where there is coverage of several wireless networks or X and Y can cover different points on earth.

This brings up the next question. Because we are interested in M-commerce the physical movement of the users on the earth is inherent in the concept. What are the geographic locations where the user can access a certain set of servers, say a subset of Suser-understandable,X, Sfavorite,X, while moving around? Let Locations(T,User, S...,X) = {(x,y)} S...,X is accessible from point (x,y) on earth} be a set of locations from where the set of servers S...,X is accessible; each pair (x,y) is a WGS-84 coordinate for an x-y grid of 10x10 meters on earth. The larger the set Locations(T,User, Sfavorite,X) is, the larger the possibility of the user to access favorite services and contents, while on the move. In general, if X is a wireless telecom network and a place with coordinates (x,y) belongs to Locations(T,User, Sfavorite,X), then the whole coverage area of X belongs to it. Further, foreign networks with data roaming might belong to Locations(T,User, Sfavorite,X). If X is a Wi-Fi network, its coverage can be 100 meters in diameter, or it can have thousands of base stations in various cities, each of which covers a small spot on earth.

Looking at terminals, the more frequency bands, protocol stacks, and content formats a terminal supports, the more widely in the world it can be used to access the M-commerce servers. Thus, if T’ has all capabilities that T has, but in addition more of relevant ones, then Locations(T,User, Sfavorite,X) is a subset of Locations(T’,User, Sfavorite,X).

3.2 Locally Accessible Spheres While Roaming

Whereas the above setting answers the question “from where can I access my known favorite M-commerce providers with my own terminal”, another question is also interesting for moving customers. Namely, “how can I access local services with my terminal from where I currently am?” Typical example might be to order a taxi in a foreign place. Coming back to the modeling above, the set Locations(T,User, Suser-understandable,Y) should contain the current coordinate of the user (x,y) and a suitable server S that provides the service relevant for user in location (x,y), when accessed through wireless network Y that has coverage in (x,y)). There are challenges on each level 1-4 above in this case. At level 1-2 these should be solvable, because both TCP/IP and the contents formats can be assumed to be globally specified and processing support implemented at the terminal should be there. But at level 3 M-commerce transactions should be performed. This usually already precludes usage of natural language, in addition to the interoperability of the M-commerce transaction support between the local environment and at terminal. Should this function, the level 4 remains open. It might be that Locations(T,User, Suser-understandable,Y) is even empty, or that the servers in the set Suser-understandable are not relevant for the task at hand, i.e. they do not provide services for the local environment the user is in.

There is a further problem that should be resolved by standardization. Namely, how to find the server or servers S that might potentially provide a local service for the user. The problem can be formulated by requiring that one must find a server in Suser-understandable,Y such that it provides a suitable service for User in location (x,y). This can be modeled as a requirement that from coordinates (x,y) and the request the user has req(goal,lang) the network infrastructure must be able to deduce the address of the server S that would provide to the user the service expressed with the goal in language lang. This is a service discovery task guided by semantics and position.

Another issue is to use the terminal as a means to perform local payments (cf. above the Lawson example in Japan). Whereas plastic cards function all over the world because of standardization of their physical dimensions and encoding of the data on the physical carrier (magnetic stripe, IC), the same is not as obvious for wireless terminals. There are however,
no inherent technical reasons for not having standard technology both at the terminal side and at the infrastructure side that would make global usability of terminals as credit cards impossible. The business and legal reasons are here more important obstacles. Unless there is strong will of global actors to push this kind of payment technology to global use, it will not happen. What kind of alliances could do this is for further study.

To solve the above problems requires certainly global (de facto?) standards, because a terminal can roam from any place on earth to another and thus a local solution will work neither for payments nor for locally accessible services.

4. Emerging Challenges for Various Regulation Bodies

4.1 The Impact of Regulation at Physical/Link, Network/Transport, Application/Contents and Business Model Levels

What has regulation to do above? If we look at an ordinary user, he or she cannot do much, but to use a terminal and hope that it works. All the hardware and software and provided to the terminal by manufactures and the functionality is pre-scribed by various standards and operators that follow the local rules imposed by local authorities ruling in a certain geographic area. These rules also prescribe what kind of business practices the operators are allowed to follow, e.g. is the lock-in of customers allowed by terminal properties or access to Internet prohibited or limited (“walled garden”). These restrict the set STCP/IP,X above.

The set SHTTP(S)+cont,X is determined by the interoperability of the application protocols and content formats. It is an interesting question, who is responsible in the mobile ecosystem that terminals and the contents are interoperable? Because the servers can be in different countries than the terminals this cannot be solved by local authorities and even if they could solve these kinds of problems, they would not want to. Evidently, the solution relies at terminal manufacturers and businesses offering the contents that should coordinate their actions. Or the terminals must be able to download new software on demand to be able to present various contents.

The set SMC-t,X are then again influenced by the local authorities that say how E-commerce is to be conducted and by the businesses that follow a code of conduct. The user cannot but comply with the rules. The more different ways to conduct M-commerce transactions there are, the more functionality the terminal should have. This might concern application level protocols, authorization mechanisms, encryption, and payment modalities that all should find technical support in the entire protocol stack. Also DRM support might be needed, before the M-commerce site would give out the contents. As above, the interesting question is, which actor is responsible that the terminal is interoperable at this level.

The language issue at business model level is mainly a business issue, because in a free market situation authorities are not going to compel businesses to offer contents in a certain language or prohibit it to be offered in another. This actually shows also that language communities that are too weak to support an E- and M-commerce market might be in trouble.

Another thing is that for people that visit the geographic area of another language community, the services should be understandable. The community might be so small that it might not have resources to produce these, unless the revenues generated by the visitor population are large enough. What is a minimal size of population that can support a viable M-commerce market?

4.2 M-commerce Regulation from National Point of View
As was discussed above, fragmentation of the M-commerce market can result of incompatible solutions at various technical levels and in this respect national regulators can choose different policies. An issue that not fully understood currently is the interdependencies of the different technical levels. To which extent e.g. does certain radio technology (GSM, UMTS, Wi-Fi, WiMAX, but especially Cognitive/Software Radio technology) used within the wireless access networks have influence on the higher technical layers (network/transport layer, application layer/compatible data formats) and further, on the M-commerce business models? Should governments choose regulation policies that work against fragmentation at all levels and foster global M-commerce? Or allow or even foster fragmentation so that only domestic population can engage in the M-commerce activities? Or be neutral? The current approach in various parts of the world has been to allow certain radio technologies for the licensed bands and - quite naturally - be ignorant on the usage of the unlicensed band (Wi-Fi, WiMAX). The rest of the technical layers are left for the business actors to consider and shape. These are tied, though, by the consumer protection legislation and further legislation for E-commerce and competition to certain practices and business models. These must be reflected also to the technical realization of the systems (e.g. privacy concerns).

One important question from regulation point of is, what are the accepted ways to compete among operators and other actors on the wireless market? In USA and Japan several radio access technologies are allowed for 2G and 3G networks and by choosing an operator the customer is at the same time locked to a certain radio technology and changing operator has meant changing also terminal. In Europe 2G telecom networks have been based on GSM standards and thus the same terminal could have been taken to the other network, though often the change from one operator to another leads to change of the terminal, as well. A few years ago regulation was implemented in Europe that allows the user to keep the old number while changing operator. This has fostered competition. From M-commerce point of view this should be also beneficial, because the newer terminals have more memory and processor capacity and can often run applications and handle many video and picture formats. Thus the palette of contents that can be traded to the terminals becomes larger.

The wireless telecom operators seem to be currently in the centre of the M-commerce development. NTT DoCoMo is the leading example in this respect. It created the technology for all the technical layers and set up the business model framework thus creating a new market. The central M-commerce regulation questions are indeed related with the operators. Is M-commerce at all possible without a strong operator like NTT DoCoMo? What should be done in Europe in this respect? This is a question that is primarily addressed to the EU commission.

Another emerging challenge is the horizontal convergence that causes different networks and their terminals to converge. In Germany there is an interesting situation around mobile-TV. According to (Meier 2006) German states are in the process of granting licenses and thus wireless spectrum to DMB standard –based mobile-TV broadcasting. Other actors are in favor of DVB-H standard and are waiting for the licenses for it. Thus, it is possible that this market will be fragmented at the wireless network level. Further, there is an additional discrepancy concerning the rights, in this case the WM soccer contents. The rights owner has sold the rights to the main TV channels (ARD and ZDF) and claims that they are not allowed to send the broadcasts to the mobile TV-enabled terminals. The broadcasters claim that they have the right to use any technical platform to distribute the contents.

4.3. Portable Terminals as Sources of Information
Hundreds of millions of modern mobile terminals have already a camera integrated and producing video clips are also possible. Thus, they can be used to produce contents. This can also be commercialized by making the contents accessible against payment. Although there is legislation in place that regulates the rights of the content owner, there aspects that might need further attention. For instance, if private people begin to make photographs of celebrities and sell them. Or in a soccer match if thousands of users are videoing the match, the coverage of certain details can be better than in the official broadcast. These can be shared by the users or also sold to the broadcasters.

If a TV-station broadcasts a video clip made with a private mobile terminal, should the purchase of the contents be considered as mobile commerce or as normal TV business? Does the broadcast channel (terrestrial or mobile –TV broadcast) have influence on the assessment of the nature of this kind of activity? This kind of regulation is regional and it addresses the usage of the terminal for certain purposes. Different countries can easily decide differently in these matters.

5. Conclusions

Regulation must be understood as activity that is performed by authorities on one hand and by industry and standardization organizations on the other hand. The regulation performed by authorities is valid in a special geographical area (country, state) and can concern allowed technologies and market conditions. The standardization bodies specify technologies that are potentially globally deployable, but will be deployed by in certain geographic areas based on market conditions and authority regulation in the geographic area.

We argue in this paper that regulatory framework for M-commerce consists of three main layers that are controlled by different actors. We assume that the wireless terminals will have and IP address and they can thus directly access M-commerce servers in IP network. The lowest layer are comprised of those OSI layers that are below IP layer (physical and link layer). Above that there is a homogeneous TCP/IP layer that isolates the heterogeneous wireless protocol layers and facilitates global access to TCP/IP enabled servers.

Above TCP/IP layer there are application protocols, especially HTTP(S) and various content formats, like video and audio formats. Further, authentication and authorization mechanisms are at this level. These are imposed by regulations of local authorities and implemented within the infrastructure, including the terminal. Harmful and criminal contents regulation is dealt with at this level, too. Thus, there are two layers where local authorities regulate and a layer on between that is globally regulated by IETF.

The convergence is changing the entire wireless environment. The harbinger is the mobile-TV that brings together the previously separate wireless networks at the terminal. From M-commerce point of view this opens a new channel to distribute contents and also a back channel that can be used to launch M-commerce transactions based on the TV-contents. Further, from regulatory point of view, Mobile-TV-contents and mobile-TV usage might require new legislation. It is also an interesting example concerning the impact of regulation on deployment of new technologies. If e.g. one decides that each mobile-TV owner must pay for a full-priced TV license, the technology will most probably not lead to a profitable business.

Physical mobility of people and terminals is an essential ingredient of M-commerce and access to local services while roaming is an important difference between it and other forms of E-commerce. The analysis shows that more regulation is needed to solve this. The problem lies in the fact that local authority regulation has effect on the highest layer necessary in M-commerce transactions and these are different form region to region. Thus, the problem is difficult.
In general, the digital convergence will most probably lead to convergence of the networks and also to a new regulatory and technical environment for M-commerce. This development can take different paths in different markets due to different regulation by the authorities. To understand this requires more research on regulation and its relationship with market development and new technologies.

References


