ELSEVIER

Computer Networks 37 (2001) 221–236

COMPUTER NETWORKS

www.elsevier.com/locate/comnet

# Business models and transactions in mobile electronic commerce: requirements and properties

Aphrodite Tsalgatidou [a,*], Evaggelia Pitoura [b]

[a] *Department of Informatics and Telecommunications, University of Athens, GR 157 71 Athens, Greece*
[b] *Department of Computer Science, University of Ioannina, GR 45110 Ioannina, Greece*

**Abstract**

Advances in wireless network technology and the continuously increasing number of users of hand-held terminals make the latter a possible channel for offering personalized services to mobile users and give pace to the rapid development of mobile electronic commerce (MEC). MEC operates partially in a different environment than Internet e-commerce due to the special characteristics and constraints of mobile terminals and wireless networks and the context, situations and circumstances in which people use their hand-held terminals. In this paper, we discuss the business models in MEC and transaction modeling issues pertinent for the business models and the environment. © 2001 Elsevier Science B.V. All rights reserved.

*Keywords:* Mobile computing; Electronic commerce; Business models; E-commerce transactions

## 1. Introduction

Advances in wireless network technology and the continuously increasing number of users of hand-held terminals make the latter a new channel for offering personalized services to mobile users and give pace to the rapid development of e-commerce conducted with portable devices. The basic requirement for mobile e-commerce and personalized services is that the mobile wireless devices are either directly "Web-enabled" or at least "WAP-enabled" [11] or "I-mode-enabled", i.e., "Internet-enabled". Current estimates of the number of such wireless Internet-enabled devices range from 134 to 330 millions around the year 2003 [3].

Mobile electronic commerce (MEC) or mobile e-business refers to e-commerce (e-business) activities relying solely or partially on mobile e-commerce transactions. As a mobile e-commerce transaction we define any type of transaction of an economic value that is conducted through a mobile terminal that uses a wireless telecommunications network for communication with the e-commerce infrastructure. MEC operates partially in a different environment than e-commerce conducted in fixed Internet, due to the special characteristics and constraints of mobile terminals and wireless networks and the context, situations and circumstances in which people use their hand-held terminals. MEC has a number of business, technical and legal implications that are different from e-commerce in the fixed Internet setting. Most notably, location-based products and services is a completely new business, technical, and legal area that is typical only of MEC. MEC

* Corresponding author.
*E-mail addresses:* afrodite@di.uoa.gr, afrodite@tifu.jyu.fi (A. Tsalgatidou).

1389-1286/01/$ - see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S 1 3 8 9 - 1 2 8 6 ( 0 1 ) 0 0 2 1 6 - X

becomes interesting with the huge proliferation of the WWW-based business-to-consumer (b-to-c) e-commerce in Internet since 1995 and the simultaneous and huge proliferation of digital wireless telecom networks throughout the world.

In this paper, we first present the special characteristics that differentiate MEC transactions from traditional e-commerce transactions, in particular the constraints of mobile terminals and wireless networks and the context, situations and circumstances in which people use their hand-held terminals. Then, we focus on how these characteristics affect MEC business models and the important role that mobile network operators (MNOs) can play in this context. We also present a formal model for MEC transactions and their properties.

The remainder of this paper is structured as follows. In Section 2, we survey some of the issues that differentiate MEC from traditional electronic commerce. In Section 3, we present the business models and the main players in MEC. In Section 4, we focus on MEC transactions, their properties and how they relate to business models. Section 5 concludes the paper.

## 2. Peculiarities of the wireless environment

### 2.1. Implications of the mobile terminals

Mobile devices that are of interest to MEC can be divided into four categories based on their processor, memory and battery capacity, application capabilities (SMS, WAP, Web, I-mode), as well as physical size and weight. These categories are (from weakest to strongest): usual voice handsets with SMS capability, WAP phones, communicators/PDA with wireless communication capability (e.g. [5]), and finally laptops with wireless communication facilities. To be easily carried around, mobile devices must be physically light and small. The smaller and lighter the devices are, the *more portable* they are. In addition, a mobile device should be a multipurpose device (voice phone, data transmitter, PDA, etc.) so that the user does not need to carry too many gadgets. Portability considerations, in conjunction with a given cost and level of technology, will keep mobile elements having less resources than static elements. In particular:

- The devices have small screens and small, multifunction keypads; the former fact necessitates the development of appropriate visual user interfaces, different from the PC or laptop.
- They have less resources than static elements, including memory, disk capacity (usually absent from the three lower classes) and computational power than traditional computing devices.
- Portable devices rely for their operation on the finite energy provided by batteries. Even with advances in battery technology, this energy concern will not cease to exist. This is because the conserved energy depends primarily on the weight volume of the battery. Different technologies have in this respect different coefficients, but the law is the same.
- There are higher risks to data stored and transactions performed in mobile devices, since it is easier for mobile devices to be accidentally damaged, stolen, or lost than fixed devices.

### 2.2. Implications of the wireless networks

The necessary networking infrastructure for wireless mobile computing in general combines various wireless networks including cellular, wireless LAN, private and public radio, satellite services, and paging [10]. As compared with wireline networks, wireless communications add new challenges:

- *C-autonomy*. The handsets in the wireless radio networks are normally not always communicating with the network infrastructure, i.e., they are unreachable. There are numerous reasons for this behavior that can be described under C(ommunication)-autonomy [8]. First, disconnections may be voluntary, e.g., when the user deliberately avoids network access during nighttime, or while in a meeting, or in other places where the user does not want to be disturbed. In cases that the handset does not have voice capabilities, and thus disturbing is not a big issue, it is still often reasonable to cut the wireless communications with the network to reduce cost, power consumption, or bandwidth use. The break in on-going communication or incapability to set up any communication can also happen against the will of the user, e.g., when a user enters a physical area where there is not any or not enough field strength for a successful communication (a typical example is a train entering a tunnel, which often leads to an abrupt decrease in the field from the device point of view), battery becomes suddenly empty, or hand-over between base stations does not succeed and the connection is therefore lost.

When analysing the different situations, one must differentiate between non-reachability of the device from the network because the user wants to exhibit her C-autonomy and non-reachability of the device against the will of the user. The latter can be called *disconnection*in the strong sense, if there was an ongoing connection between the terminal and the network when the device became unreachable for the network. However, if the user just shuts down the radio transmitter in the middle of a connection, then this is a disconnection only from the network point of view and it is a voluntary disconnection from the user's point of view.

Disconnections can be categorized in various ways from the point of view of the user, hand-held terminal, or the network infrastructure. Disconnections are either predictable or sudden from some point of view. For example, voluntary disconnections are predictable from the user point of view. From the device point of view they can be sudden. Clearly predictable disconnections from the device point of view include those that can be detected by changes in the signal strength, by predicting the battery lifetime, or by utilizing knowledge of the bandwidth distribution. In general, if the disconnection can be predicted by the device, the latter can usually inform the network infrastructure and the user of the immediate disconnection and then perform it properly. If it is sudden from the device point of view, there is no time or possibility to do anything before the connection breaks. Afterwards, the device can inform the user about loss of connection. These are the most difficult situations from the application point of view. From the communication infrastructure point of view, there is not much difference whether the connection just breaks or whether it knows about it just before it happens; sometime after the disconnection, the resources allocated for the connection will be released in any case.

- *Bandwidth restrictions and network topology.* In the case of many wireless networks, such as in cellular or satellite networks, communication channels have much less transfer capacity than wireline networks. This is caused by the fact that the used modulation and channel allocation schemes designed for voice traffic have rather modest upper bounds. Further, wireless communications are much more error prone than wireline communications and require much redundancy in the channel coding of the payload.
- *Asymmetric communications.* Some wireless networks offer asymmetric transfer capacity for up- and downlink. The asymmetric transfer capacity on uplink and downlink can be applied in a reasonable way if the network offers broadcast facility. This is unfortunately not a strong side of the telecom networks, because they were designed for connection-oriented point-to-point communications. Wireless LANs are better in this respect, because they apply packet broadcast protocols anyhow. GSM networks have broadcast facility on the control channels, but the amount of application data that can be transferred on them is small. The currently very popular short messages (max 160 characters) are an example of such data that is transferred over control channels. If used, e.g., to broadcast multimedia contents over the network, the network would collapse, because controlling the traffic would not be possible any more. Still, the asymmetric transfer capacity is an important asset in cases where the wireless client usually sends a short request and gets a large data set as a response. One should also note that, in general, it costs less to a client in terms of power consumption to receive than to send.
- *Variant bandwidth and bursty traffic.* Currently, multinetwork terminals are emerging that can use several networks to communicate. Typical forerunners are the dual-band devices that are able to use 900 MHz and 1.8 GHz GSM networks. Soon, there will be products that are able to also use WLANs and possibly Bluetooth [2], together with GSM, GPRS, and soon also UMTS network

infrastructure. Wireless technologies (e.g., BT, WLANs, cellular telephony) vary on the degree of bandwidth and reliability they provide. In this respect one can speak of variable bandwidth. Another phenomenon also observable in the wireless world is bursty traffic which is the case with Internet-type networks and this holds in different time scales.

- *Variant tariffs*. For some networks (e.g., in cellular telephones), network access is charged per connection-time, while for others (e.g., in packet radio), it is charged per message (packet). In the WAP environment there is a larger variety of tariffs, e.g., session-based, transaction-based, connection time-based, while in mobile e-commerce the range of tariffs is even wider.
- *Mobility*. GSM infrastructure allows roaming all over the world, i.e., the user can get access to voice and data services basically in any other GSM network (in practice of course the operators must have a roaming contract). Mobility causes diverse phenomena. The available bandwidth might vary, for instance, a mobile terminal may rely on low-bandwidth networks outdoor, while inside a building it may be offered reliable high-bandwidth connectivity or even operate connected via wireline connections. Moreover, there may be areas with no adequate coverage resulting in disconnections while on the move. There may be also variability in the provision of specific services, such as in the type of available printers or local weather reports. Furthermore, the services offered by the telecom network used might differ from those at home. This might have drastic consequences for MEC, if the e-commerce infrastructure used needs them. Finally, the resources available to a mobile element vary, for example, a docked computer or PDA has more memory or is equipped with a larger screen. Mobility also raises very important security and authentication issues.

### 2.3. Usability implications

MEC applications take advantage of mobile communications to offer to consumers and businesses additional benefits as opposed to traditional e-commerce applications.

*Location-awareness*. In mobile computing, knowledge of the physical location of a user at any particular moment is central to offering relevant services. The location of a mobile device is available to the MNO but it can also be found using sensor devices or technologies such as the global positioning system (GPS). GPS uses a number of satellite stations to calculate with great accuracy the location of devices equipped with GPS receivers. There are many examples of location-based electronic commerce applications including: geographically targeted advertising (everyone near a fast-food restaurant gets free (electronic) coupons for the new burger), fleet management, vehicle tracking for security, traffic control, telemetry, emergency services, etc.

*Conditions of usage*. The mobile user may be engaged into another activity, like traveling, meeting people, etc., rather than sitting in front of his/her desk top terminal.

*Adaptivity*. Mobile e-commerce applications should be adapted to the environment of their clients. Adaptability is possible along various dimensions including the type of the device in use, the currently available communication bandwidth as well as location and time.

*Ubiquity*. Mobile communications enhance electronic commerce by making electronic commerce services and applications available *anywhere and at anytime*. Through hand-held devices such as mobile phones, users can be reached at anytime, independent of their location. Mobile computing makes possible that users are immediately notified about particular events. It also enables the delivery of time-sensitive information whose value depends on its timely use.

*Personalization*. The information, services and applications available in the Internet today are enormous. It is thus important that the user receives information that is of relevance. Furthermore, customization is a key issue in using mobile devices because of the limitations of the user interface in terms of size, resolution and surfability. Studies, e.g., [4], show that every additional click reduces the transaction probability by 50%. Thus, MEC applications must be personalized enough to represent information in compact and attractive forms and to optimize the interaction path, enabling the user to reach the desired services with as few clicks as possible.

*Broadcasting*. Some wireless infrastructures, such as cellular architectures and satellite networks, support broadcasting (i.e., simultaneous delivery) of data to all mobile users inside a specific geographical region. Broadcasting offers an efficient means to disseminate information to a large consumer population. This mode of operation can be used to deliver information of common interest to many users such as stock prices, weather information or for advertising.

## 3. Business models

In order for companies to be successful in m-commerce, they need to evaluate innovative new strategies that capitalize both on the power of the Internet as well as on the changes in market demands. This requires that new business models be created to offer a new way to deliver value to customers. By *business model* we mean a logical architecture for product, service and information flows, including a description of the involved business actors and their roles, as well as sources of revenue.

A company in order to adopt a successful business model for MEC has to take into account:

- Core competencies.
- The special characteristics and constraints of mobile terminals and wireless network.
- The different context, situation and circumstances that people use their mobile terminals.
- Internet e-commerce models.
- Market needs.
- Other actors and players in the field.
- Previous success stories.

The second and third factors above have been analyzed in the previous section. In the following we examine some of the most representative Internet e-commerce models that can be used as a basis for m-commerce models. We then examine the main players in m-commerce value chain and the m-commerce business models that are based on Internet e-commerce business models. Furthermore, since the role of the MNO is of strategic importance, we examine some scenarios of the various roles it can play.

### *3.1. Business models in Internet e-commerce*

E-commerce over the Internet has a great impact to the business world. Many organizations have already embarked upon reengineering efforts in order to keep or create a competitive business advantage in a changing business environment and a lot of new e-commerce business models have emerged. The emerged e-commerce models vary in their degree of innovation. Some of them are not 'new' but rather an automation of models existing in traditional commerce, see for example the 'e-shop' which is actually an electronic version of traditional ways of selling. Other models, like for example the 'e-auctions', 'e-procurement' or 'trust services' automate services and functions that have been provided since years and at the same time they add new functionality. A third category of models, like for example the 'value chain integrators' are models which do not exist in traditional commerce as they provide services which are totally dependent upon information technology and cannot be done at all in a traditional form. E-commerce models also vary depending on the degree of the supporting functionality, from models providing only one marketing function over the Internet (e.g., the e-shop) to models providing a fully integrated functionality (e.g., value chain integration) [6].

Some of the most successful business models in Internet e-commerce are listed below:

- *E-shop*. It is a business model according to which individual shops sell various goods, like for example cigars (www.stogies.com), tickets (www.travelocity.com) or food (www.pizzahut.gr). A collection of e-shops under a common umbrella giving entry to individual e-shop gives pace to another business model that is called *E-mall*. The common umbrella is usually a well-known brand that can also provide additional support like for example a commonly guaranteed payment method.
- *E-auctions*. It automate the traditional bidding process over the Internet. They can also additionally support contracting, payments and delivery processes.
- *Third party marketplace*. It is a model used by companies that wish to outsource their Web

marketing activities. These market places like for example Citius (http://www.citius.fr) or Trade-Zone (http://tradezone.onyx.net/) offer a user interface to the supplier's product catalogues and may additionally support payment, logistics, ordering, secure transactions etc.

• *Value-chain integration model.* Value chain integration uses Internet technology to improve communication and collaboration between all parties within a supply chain. Value-chain integration is necessary if vendors are to coordinate between "upstream" suppliers, internal operations (e.g., manufacturing processes), and "downstream" shippers and customers effectively. With this model processes once perceived as internal to the company, now span the entire value chain. Service providers integrate their operations directly into the processes of their customers. With this model every company in the chain performs a set or sequence of activities to produce its products. The links between those activities provide a prime opportunity for competitive advantage, whether due to exceptional efficiency or some form of product differentiation. This chain of partners that work in sequence to create, market and move goods and services grows ever more complex. Based on the sorts of intimate trading relationships central to the integrated value chain model, modern business partnerships are eradicating duplication, irrelevant hand-offs and rework, ensuring that processes run smoothly and effectively.

• *The hub or the portal model.* This business model focuses on the integration of multiple steps of the value chain with the potential to exploit the information flow between those steps. Thus, the hub or the portal not only supports services already offered by its customers but also offers added value services. Customers of the hub/portal have an easy and cheap access to services provided by the participating companies (e.g., e-banking or phone banking, loan applications and so on) and to the added value services offered by the portal through a uniform interface. The benefits of the participating companies is the minimized implementation effort for modernizing their business, the potential to form alliances with the other participating parties and to offer their services via the value chain integrator platform as an integrated services. In the latter case, the customer has the impression that all the provided services are offered by the same institution, a kind of virtual merge of the involved parties.

• *Value chain service providers.* They usually support a specific function in the value chain, such as electronic payment, logistics or distribution (e.g., FedEx (http://www.fedex.com) or UPS (http://www.ups.com)). The service provided (e.g., e-payment) has usually an interesting degree of innovation as it offers over the Internet a traditional function (e.g., accounts management) enhanced with new functionality (e.g., Internet smart card support). The direct benefit for businesses using the provided services is low cost of investment for modernizing their operation and providing to their customers advanced services like e-payment.

• *Information brokerage.* The vast amount of information available in the Internet has resulted in the emergence of a whole range of information services which add value by providing information search, customer profiling, business opportunities brokerage, investment advice and so on, see for example http://www.yahoo.com, www.altavista.com. The revenues from this business are coming either through subscription fees or through advertising.

• *Trust services.* Companies adopting this business model (e.g., Belsign (http://www.belsign.be)) offer traditional trust services over the internet with the addition of new functionality, such as encryption and public and private key management, which cannot be available without IT.

• *Process outsourcing model.* Process outsourcing is the delegation of one or more business processes to an external provider who owns, manages and administers the selected processes. For example, take Ford Motor Company who decided that the manufacture of cars will be a declining part of its business and instead they will concentrate in future on design, branding, marketing sales and service operations. Like all modern carmakers, Ford has outsourced the supply of entire sub-systems – from engines and suspension assemblies to car interiors. In such situations, suppliers application systems are automatically kept abreast of requirements (via EDI).

• *Virtual organization.* A virtual organization may be a temporary or permanent collection of geographically dispersed individuals, groups, organizational units – which do not necessarily belong to the same organization – or entire organizations that depend on electronic linking to complete the production process. Onsale (www.onsale.com) is an example of such a model which has successfully created a market-space for not only selling online, but also exploited organizational opportunities afforded by emerging technologies, by building a truly virtual organization where inventory is minimal, geographical proximity is not a major issue, and information and partnering are of paramount importance. Onsale is an on-line auction house that sells computers, chips, peripherals, and other computer related add-ons through the World Wide Web. With knowledge and information around the on-line auctioning process being its primary resources, the relatively small company gains significantly in size and functionality by partnering with other organizations and outsourcing non-essential activities.

• *Application service provider (ASP).* It is another Internet business model where a company that plays the role of ASP gets the licenses for certain application software and rents it to client companies. Thus, the ASP hosts, delivers and supports applications for its client companies. ASP can be application and system integrator and implementor or software vendor and can provide data center infrastructure along with on-going support. A specific example of ASP is the payment service provider, which is a business model by itself. See for example www.bibit.com that provides a payment service that supports many Internet payment methods including the new SET protocol and conventional credit card payments.

There are more business models like virtual communities or collaboration platforms and it is likely that even more models will emerge, while others will be not further promoted if they are not proved to be successful. Creating these new business models is feasible only because of the openness and connectivity of the Internet.

To avoid confusion, we would like to mention that a role of an actor in a specific business model, e.g., the role 'Value Chain Service Provider' in an 'integrated value chain' business model, can be also a business model itself, depending on the way we view things. For example, when we examine the function of UPS (www.ups.com) as such, we can consider that they have adopted the 'Value Chain Provider' business model in their business since they support a specific function in an integrated value chain. But, if we consider the business 'Hub/Portal' model, then, UPS plays the role of the 'Value Chain Provider' in the integrated value chain. Also, a combination of business models is possible. Furthermore, a business model can be seen from many points of view, e.g., the 'Payment Service Provider' can be seen as a business model by itself, or as a specific case of the 'ASP' business model or as a another type of business model, namely, 'Value Chain Service Provider' depending on the way the provided payment service is provided and the specific role of the payment service provider in the value chain. Therefore, since the same term can be used under different meanings, it is important when we refer to a particular business model to define clearly the architecture for the provided products and services as well as of the information flows. A description of the involved business actors and their roles as well as sources of revenue is also required.

### 3.2. MEC main players and business models

The players in the mobile business value chain are: *technology platform vendors* that offer operating systems and microbrowsers (e.g., Phone.com, Symbian, Ericsson, Microsoft), *infrastructure and equipment vendors* that offer the network infrastructure (e.g., Nokia, Ericsson, Siemens, Motorola, Lucent), *application platform vendors* who provide middleware and standards (like Nokia, Phone.com, Ericsson, ETSI, WAP Forum, UMTS Forum), *application developers* who offer mobile platform applications (like Yomimedia, WAPIT, Add2Phone), *content providers* (like Reuters, Yahoo), *content aggregators* (e.g., Digitallook.com), *mobile portal providers* for application aggregation, MNOs, *mobile service providers* who act as an intermediary between the operator and the customer for mobile phone contracts and terminals and handset vendors and retailers.

Regarding the participating entities in an m-commerce transaction, they depend on the underlying business model. In general, the following entities are the main participants in an m-commerce transaction:

- *Customer* who is mainly mobile. The place s/he is when the transaction is initiated can be different from a place s/he is when s/he receives the service, pays and the transaction is committed. These places can be in different countries or in the border of countries and the consumption of the services s/he has acquired can take place in a third country.
- *Content provider* who provides specific contents to a customer through a WAP Gateway which can be hosted at the MNO or through a portal that can be hosted at the operator's WAP server or anywhere else.
- *Mobile portal* that offers personalized and localized services to customers to minimize the required navigation by the user.
- *Mobile network operator*. The role of the operator is vital for the MEC. Depending on where it stands in the whole value chain of m-commerce, its role can vary from a simple mobile network provider to an intermediary, portal or trusted third party.

Depending on the way the above actors are related to each other, we have a number of different m-commerce business models some of which are outlined below:

- *Content providers.* This is a business model (and a role at the same time) which comes natural for companies like Reuters or traffic news providers or stock exchange information providers. They can offer their contents by directly contacting customers or via mobile portals. Apart from companies of this type, this type of business model is adopted by smaller companies or individuals who develop some content for mobile handsets and they offer it (or sell it) to some software companies (e.g., Yomimedia) who subsequently offer it to mobile customers.
- *Mobile portals* that offer personalized and localized services to mobile customers. Thus, a customer moving from Finland to Greece, when in Greece should be able to access a Greek portal that supports not only the local language but also his 'home' language, and provides local-specific information like tourist attractions, local restaurants, etc. Thus, mobile portals differ from Internet portals, as the needs of mobile customers and the characteristics of a mobile terminal differ from those of a customer sitting in front of his/her desktop or portable computer. A mobile portal (m-portal) is characterized by a greater degree of personalization and localization. Localization means that a mobile portal should supply information relevant to the current (geographical) location of the user. Information requirements may include for example, restaurant bookings, hotel reservations, nearest petrol station, yellow pages, movie listings and so on. Personalization applies to any kind of information provided by m-portals including location-specific information. User's profile, cultural interests, past behavior, situation and location should be taken into account for the provision of such personalized and localized services. www.yourwap.com is a representative example of a mobile portal.
- *WAP Gateway providers* (for m-commerce over WAP). The WAP Gateway providers can be considered as a special case of the ASPs Internet business model. The service they provide is the WAP Gateway to service providers who do not want to invest in a WAP Gateway. Revenues of this model depend on the kind of agreement the WAP Gateway provider has formed with its clients.
- *Service providers* provide services to customers either directly or via a mobile portal or via a wap gateway of another company or of the mobile operator. The service they provide may depend on specific contents they have acquired from content providers.

According to the business model supported, revenues are shared among the content providers, service providers, WAP gateway or portal providers and the MNO. A combination of the above business models/roles together with some already existing roles and business models in Internet e-commerce (e.g., payment service provider, financial institution) results in more complex business models for m-commerce. Each player in order to adopt the most profitable business model, has to take into account the various factors mentioned before, i.e., its core competencies along with the special characteristics of m-commerce environment and gain benefit from experiences and success stories from Internet business models.

Since the role of the operator can be quite complex and affect also the billing and payment of m-services, it is separately discussed in the following sections.

### 3.3. The role of the MNO

M-commerce brings challenging opportunities for the MNO, the role of which can vary from very simple and passive to very active and dynamic by being strategically positioned between customers and content/service providers and offer portal or trusted third party services.

Fig. 1 depicts the various roles of the MNO in relation with the degree of its involvement in the m-commerce value chain and the degree of the importance of its role. It seems that the simplest and most passive role that the MNO can play is to just provide the mobile network infrastructure and let the customer communicate and negotiate directly with the various content/service providers or other portals (Scenario 1). Profits of the MNO come from the offered wireless connection.

In addition to this, the operator can have three escalating roles:

- Host a WAP Gateway for enabling the exchange of information between a customer holding a WAP terminal and an Internet merchant who does not provide WAP-compliant contents.
- Act as a portal offering advertising services and providing search facilities while enabling connection with the content/service provider.
- Act as an intermediary and trusted third party.

The last two scenarios are the most interesting and they are analyzed in the following.

#### 3.3.1. The MNO offers portal services

Instead of simply facilitating transactions between customers and content providers, the operator can have an escalated role by providing portal services. In other words, the operator can act as a portal, facilitating customers to locate appropriate service providers and at the same time enabling content providers to reach customers via the operator's portal.

MNOs have a number of advantages over other portal players as they have an existing customer relationship and personal data and they can identify the location of the subscriber. Usual Internet portals are neither able to incorporate location-specific information nor do they have the data and knowledge of each customer that the mobile operator has. In many instances, the 'traditional' portal player knows only an email address that can also be virtual without any information about customer's identity. Moreover, the traditional portal does not usually have a billing relationship with the customer (with the exception of portals like Compuserve who are also

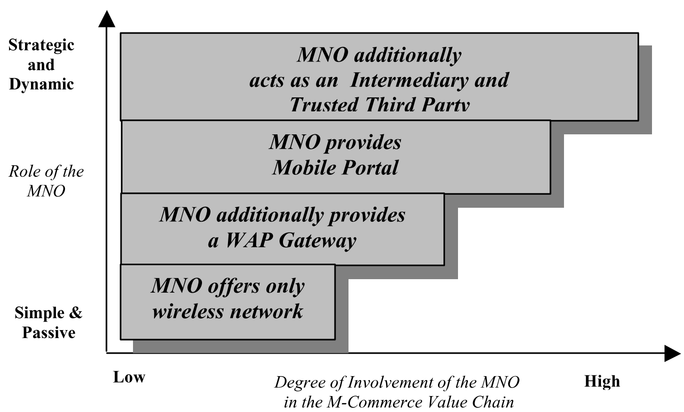



Fig. 1. The MNO in the m-commerce value chain.

Internet service providers). Therefore, it is very natural for a MNO to offer mobile portal services to its customers and additionally offer services similar to the ones offered by Internet portals, for example search services, lists of content and service providers with provided services, products and prices, respective comparisons, etc. In other words, the operator can be the front end for a number of suppliers. The customer in this way can choose the supplier with the best offer from quality of service (QoS) and financial point of view.

The operator can make profit by making special agreements with content/service providers. Direct contact of the customer with a provider, after the first contact through the portal, can be allowed or not depending on the operator's policy.

*3.3.2. Operator acts as an intermediary/trusted-third-party (TTP)*

In this case the role of the operator has been escalated into a more dynamic one in the MEC market (see Fig. 2) and provides whatever a portal provides and in addition:

- *Provides bundled services*, i.e., provides offers with a combination of various purchases from different suppliers with discounts. So, for example, if customer wants to buy product A from supplier X, s/he may find an offer by the telecom operator about two products A and B from different supplier at a lower price.
- *Acts as a front-end to the bank*, i.e., the customer pays to the operator who, in this case, is also responsible for payment refund to the customer if the latter is not satisfied with the products.
- *Provides security and payment services to suppliers*, i.e., acts as an ASP to the suppliers.
- *Acts as trusted-third-party* in cases the customer wants to buy a number of goods from various suppliers that must interoperate. Fulfillment is a critical part of an e-commerce transaction and essential for building trust relationships with customers. If the product arrives in bad condition or too late or does not arrive at all the customer has to be protected and the liability issues must be very clear. For example, say that a customer wants to buy a digital camera and a computer from two different suppliers and that the customer requires these two products interoperate in order to easily download images from the camera to the computer. The operator can guarantee properties such as money atomicity and product atomicity, i.e., the operator guarantees that the customer pays only if both products are delivered in good condition and are interoperable. This alleviates the customer from negotiating with different suppliers and from trying to find liabilities in case that something goes wrong. The operator is responsible in this case for the whole billing process, thus the customer pays just the operator who is then responsible for distributing the money to the camera and computer suppliers. Depending on the underlying agreement, some IPR owners may get paid by the content providers rather than the operators.

The role that the operator eventually plays is dependent on what is allowed by the relevant legislation. For example, current legislation in

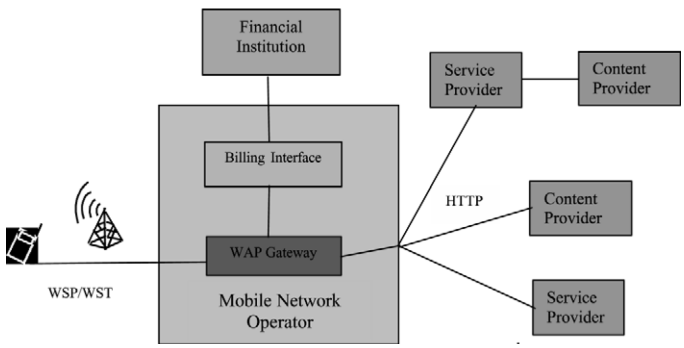



Fig. 2. The role of MNO has escalated to intermediary and TTP.

Finland does not allow telecom operators to charge for contents/services exceeding a certain amount of money (60 mk). This means, that the operator may need to find subsidiary companies in order to act as TTP and/or acquire licenses for a bank. Indeed, there are operators who have started doing this. In addition to this role, the MNO can act as Internet service provider. It is clear that the operator can play such a role not only in m-commerce but also in Internet commerce.

## 4. Transaction models for MEC

Traditionally, transaction models have been used in database systems to reason about and enforce correctness properties of the execution of database programs [1]. In the context of MEC, we use *transaction models* to formally describe the interaction (protocols) among the participants, and *implementation properties* to describe the requirements concerning the execution of such protocols.

### 4.1. MEC transaction models

Transactions in MEC involve a number of players, typically customers, merchants and often banks, MNOs and possibly other authorities, and the physical supply chain. They run their own local programs as part of a MEC transaction instance.

In order to arrive at a formal model for the MEC transaction specifications, we make the following assumptions: Each participant above hosts one or more *workflow specifications* ($WFS_i$). A workflow specification includes a set of *tasks* to be executed plus a set of *dependencies* that determine a partial order according to which the tasks should be executed. The workflow specifications of the participants are tied together with one or several *protocol specifications* ($PSI_i$). So, the *abstract specification* model of a MEC transaction is a graph, whose nodes are the workflow specifications $WFS_1, \ldots, WFS_n$ and there is an arc $WFS_i \rightarrow WFS_j$, labeled with $PSI_k$, if $PSI_k$ is a protocol that is *compatible* with $WFS_i$ and $WFS_j$ and might send a message from $WFS_i$ to $WFS_j$ in some execution. *Compatibility* is a formal correctness requirement meaning that the protocol $PSI_k$ is defined by the workflow specifications (seen as protocol state machines) of the protocol and the messages being sent. Notice that the protocol topology can combine more than two workflow specifications into the same protocol.

An important question is how many different protocol types are there and what are their topologies. Conceptually, one can assume that one MEC transaction type covers all possible players. Thus, there would be only one protocol whose specifications and executions have the transactional properties and that consists of all the workflow specifications and the message arcs between them. However, this is not probably manageable for many reasons. Therefore, we envision that the formal transaction specification model is a graph that is not necessarily even connected. There might be also arcs that only go in one direction, i.e., the protocol just starts another workflow instance and does not wait for a response or mediates some information for another workflow instance that is in the middle of its execution waiting for it. For example, a message sent by a protocol run at a customer site (e.g., payment) may not wait for the reply from the protocol run at the merchant site (e.g., a receipt of payment notification).

Notice that the graph above contains all possible execution structures that might unfold during concrete execution. Thus, it contains both physical infrastructure transactions where the goods are physically transported between the merchant and the customer, as well as those cases, where the customer gets the digital goods through the communication network. It also contains different forms of authentication methods that are of course used in concrete cases in different ways.

What would then be the execution model? It seems appropriate to apply the model developed for the S-transactions, the execution graphs [9]. In this model, the workflow executions unfold to a tree where actions on the paths of the tree are either local (database) transactions or communication actions (receiving or sending a message). The transactional properties characterizing the executions are expressed through the correctness properties on the paths.

What are the transactional properties in this case? Traditionally, in database systems, transactions are used to encapsulate operations and provide atomicity, consistency, isolation and durability (the ACID properties).

*Atomicity.* A MEC transaction is *atomic* if either all operations of the transaction are executed or none is. Atomicity is an important property; however, many electronic commerce protocols do not provide atomic transactions. Tygar [7] defines three levels of atomicity to properly characterize electronic commerce protocols: money atomicity, goods atomicity and certified delivery. With *money atomic protocols*, funds are transferred from one party to another without the possibility of the creation or destruction of money. Money atomicity is a basic level of atomicity that all electronic commerce protocols must satisfy. *Goods-atomic protocols* are money atomic, and also effect an exact transfer of goods for money. That is, with goods-atomic protocols, a good is received if and only if the money is transferred. *Certified delivery protocols* are money atomic and goods-atomic protocols that also allow both a merchant and a customer to prove exactly which goods were delivered. Such protocols are helpful for scenarios where merchants or customers may be malicious. In terms of our formal model, atomicity can be achieved by requiring that such messages are sent and such actions occur on the paths describing the activity of the merchant, customer, bank/credit card company, and physical delivery infrastructure that cause the above atomicities to be reached.

Techniques that are used to guarantee transaction atomicity in database systems (e.g., two-phase commitment) are not suitable to guarantee transaction atomicity in electronic commerce because of autonomy and dishonest participants. The main reason is that in electronic commerce, a set of dishonest or malicious participants may cause arbitrary failures. For instance, a customer may deny receipt of electronic goods by pretending a system failure or command commits and aborts for participants in the same transaction.

*Consistency.* MEC transactions must preserve consistency at various levels. For instance, if database systems are involved in the transaction, database consistency, that is preservations of the database integrity constraints, must be maintained. For instance, a customer should not be allowed to draw funds from an account if this would result into a negative balance. In general, if more than one player is involved in the EC transaction, distributed constraints may need to be enforced; for instance a positive balance among various credit institutions may need to be preserved.

*Isolation.* Isolation ensures that the various steps of a MEC transaction do not interfere with steps of other MEC transactions. For example, if a customer buys a product through its mobile phone, this purchase should not affect other transactions being made simultaneously at the device or at the merchant server. This can be guaranteed with traditional transactional mechanisms at the sites involved in the workflow execution.

*Durability.* The durability property guarantees that once a MEC transaction completes its execution, its results become permanent even in the presence of failures. For example, after the completion of an electronic purchase, the corresponding funds transfer from a bank is permanent even if the network fails.

In MEC, enforcing the ACID properties is complicated. Mobile devices may be unreachable and unable to participate in a distributed protocol. Furthermore, due to device limitations, the part of the protocol executed on a mobile device should be lightweight. Finally, the fact that a mobile device may be easily stolen or lost must be taken into account and transactional mechanisms combined with security.

Another issue that complicates the enforcement of the ACID properties for MEC transactions is the requirement for anonymity. Some customers do not want their identity to be disclosed. Anonymity is often preserved by adopting a token-based model. Tokens are used as a form of currency: similar to coins, they are used to purchase a good but they do not reveal the identity of the holder. The requirement for anonymity complicates the development of protocols that also require unique instance identity – which is the case with MEC transactions.

### 4.2. Requirements for MEC transaction implementations

We have already discussed C-autonomy of the terminals. It requires from the implementation primarily that the workflow instance states are durably stored into a database and messages sent between the participants are either stored while the user is unable or unwilling to receive or send them or that they are resent when the communications are reestablished.

*Hostility*. In principle, we cannot assume that all participants are honest. The MEC protocol should provide enough mediated and stored information so that dishonest merchants, customers or other players can be disclosed later. This requirement influences the protocols themselves and also the actions included into the workflows.

*Security*. Security is a key issue in MEC transactions. In general, security and trust in e-commerce transactions can be achieved with the following security functions:

- Authentication and non-repudiation: each party needs to be able to authenticate its counterpart, i.e., to make sure that the counterpart is the one s/he claims to be and that s/he does not deny later on agreements s/he has approved earlier.
- Integrity: each party needs to make sure that received messages are not altered or fabricated by other than their counterpart.
- Confidentiality: each party wants to keep the content of their communication secret.
- Message authentication: each party wants to make sure that the received messages do really come from his counterpart.

Techniques like the asymmetric cryptographic algorithm (also called public key algorithm) are used to achieve these results, with a certification authority (CA) which issues certificates and a public key infrastructure (PKI) for generating, revoking, updating, recovering etc., certificates and keys. Three different security techniques are currently used for hand-held devices: (a) smart trust type solution cards (offered by Sonera, www.sonera.fi), where the PKI private key is on the SIM card and it is used for authentication and non-repudiation, (b) the solution offered by Nokia, Merita (www.merita.fi), Visa where authentication and other security mechanisms are incorporated into the software and hardware of the hand-held device in such a way that the device has a credit card capability and (c) the solution provided by Motorola and MasterCard which offer hand-held devices equipped with a credit card reader.

*Vulnerability*. Mobile terminals are susceptible to being lost, stolen or misused. Data can be lost including the workflow state. Worse, they can be stolen and ongoing transactions can be altered (e.g., paid goods redirected to other addresses) or new transactions started.

*Device limitations*. The terminals do not have much memory or fast processors. Thus, running time or space consuming algorithms can be slow or impossible.

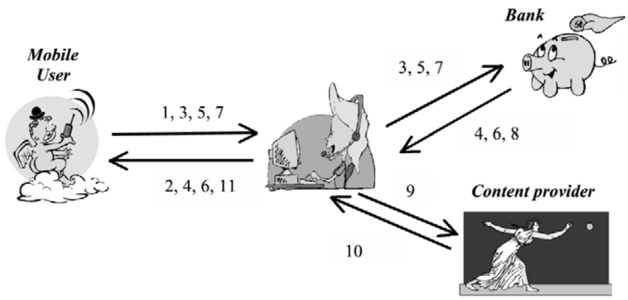
*Expensive communications*. The protocols should be designed in such a way that there is as little as possible traffic over the wireless link.

### 4.3. Case study

In order to better demonstrate the above transactional properties and requirements we will use the following case study. It is past midnight and a traveler holding a mobile phone is in the middle of an unknown place. The traveler wants to find the shortest route to the nearest first class hotel with available rooms. In the following, we examine the process of ordering this information service and paying for the service charge to the mobile portal by a direct bank transfer. The example could be extended with the customer booking for the hotel again by paying via direct bank transfer, however, for reasons of simplicity we consider that the customer pays only for the information s/he has asked for. After the information is received, the transaction is concluded. In this transaction, we consider four players:

- The mobile user.
- The mobile portal.
- The bank.
- The content provider.

Fig. 3 depicts the participating parties and the way they communicate. At each participant there is a workflow instance running. For example, the workflow process that is running at the mobile portal site includes tasks that are responsible for




1. Mobile user connects to mobile portal and asks specific service
2. Portal asks user to pay through bank
3. User connects to bank server (portal is the proxy for the bank) and provides necessary information
4. Bank asks user to confirm payment
5. User confirms payment and provides confirmation code
6. Bank asks user to press end-button
7. User presses end-button and confirms successful payment
8. Bank confirms payment to portal
9. Portal identifies user's location and requests content providers the required information
10. Content provider provides requested information to the portal
11. Portal provides requested service to the customer

Fig. 3. Participating parties in a mobile transaction and coordinating messages of the whole process.

instance for the identification of the location of the user and the calculation of the shortest route for the nearest first class hotel with available rooms. The arrows represent messages that coordinate the advancement of the processes and can be considered as steps of the global process. The messages among two participants follow the *protocol specification* defined between these participants. The actions taken at each participant are instances of the *workflow specification* for the participant.

Processing of the activities between the user, the portal, the content provider and the bank should be seen as one distributed transaction that should be atomic, consistent, isolated and durable.

To achieve *atomicity* either all or none of the operations of the transaction must be executed successfully. That is, if an operation (e.g., confirmation of payment) fails, the other operations of the transaction should also fail. Besides this general atomicity requirement, the execution of the transaction should also satisfy the three basic levels of atomicity: money atomicity, goods-atomicity and certified delivery. Money atomicity requires that no money is created or destroyed. In addition to money atomicity, goods-atomicity requires that the goods, in our case, the information about available hotels and driving directions, is received if and only if, the payment is successful. Finally, to achieve certified delivery, all participating entities must keep sufficient evidence to prove that they have delivered the requested goods, in case of disputes. For instance, in our example, the content provider must be able to provide evidence that it has provided the mobile portal with the required information about hotel availability.

Due to the limitations of mobile devices and wireless networks, outlined in Section 2, ensuring atomicity in a mobile transaction is not an easy task, and requires further consideration. For example, in order to deal with the terminal running out of battery or coverage while paying, the mobile device should have failure resiliency and capacity to recover the process into a consistent state. In practice, this would require that the data obtained from the portal that contains the transaction identifier and the details of the payment should be stored persistently along with the process state indication "transfer in progress". The application at the bank should also accept a new connection attempt after the crash with the same transaction id.

*Consistency* in our example is threefold:

1. Database consistency must be ensured in each individual step of the processes running at the participating entities, e.g., funds cannot be transferred if the credit limit of the traveller's account is exceeded as a result of the transfer.
2. Each one of the running processes should end in an acceptable end state. For example, the

process at the bank should move (or not move) the funds and announce the success (or the failure) of the transfer to both the traveller and the portal. If it stops before the announcements, this means that it did not run until an acceptable end. Similarly, the process at the portal should end either with granting the service or denying it and announcing the result to the traveller.
3. Third, all the processes should end in a globally coherent state that satisfies the distributed atomicity requirement.

*Isolation* refers to the lack of interference among transactions concurrently running at the participants. For instance, say that at the same time, a different user at the same location asks for a similar service but requests the cheapest available room. Even if these two requests are executed concurrently, they should not affect each other.

*Durability* should also be guaranteed. It must be ensured that the process states, including the sent messages, are recorded in a crash-resilient way. In our example, if the bank system crashes after a successful completion of the transaction, it must have recorded the money transfer.

Regarding *security*, it can be guaranteed if, for example, the communication with the bank, the portal and the content providers happens in an encrypted form using https. Security requires also the use of authentication mechanisms suitable for the network environment, like the ones mentioned in the previous section. However, all these solutions have weak points [9] and new solutions should be sought.

It has to be noted that, achieving atomicity is further complicated in the case of purchasing tangible goods. This is due to the fact that the process at the portal usually runs days or weeks instead of a few minutes. Also, the customer has to right to return the delivered item in a period of about one month or so, depending on related national legislation. Furthermore, the ordered item may be valuable, e.g., a golden ring. In such cases, it would be useful for the customer to have more check points to the process running at the portal, provided that each check point requires personal identification, like a private key. Such a key should be stored safely in the mobile terminal in order to prevent its misuse in case it is stolen before the transaction has run to an end.

### *4.4. Business models and required transactional properties*

There is a close relation between MEC transaction and business models. The business model provides the abstract definition of a MEC transaction and affects its implementation.

The number of players in each MEC business model determine the different nodes in the graph model of the corresponding MEC transaction, while the interaction among the players in the MEC business model determines the protocol (arcs in the graph) that connect these nodes. For instance, in the e-shop business model, there are two nodes in the transaction model corresponding to the shop (supplier) and the customer. In the third party market place model, there is an additional node, inserted between the shop and customer nodes, that corresponds to the third party. Each node (player in the MEC business model) run its own operation and interacts with the other nodes (players) according to the MEC business model. The more complex the business model, the more involved the transaction model.

Central is the role of the MNO. The case in which the MNO acts as an intermediary trusted party results in a transaction model with a centralized authority (the MNO). In effect, we get a star-like transaction graph, with the MNO as the central node through which all messages are directed.

The business model also affects the implementation of the MEC transactions in various ways. First, the MEC business models determine the abstract transaction properties that each participating entity provides. For instance, in the e-shop business model, the shop must provide consistent transactions (no violation of the constraints inside the shop, e.g., not selling unavailable products). Second, they determine in some degree the type of security provided by the MEC transactions, for example, in the trust business model. Further, in the case of centralized models, as in the case of the MNO being a trusted third party, the central entity can take the responsibility of enforcing the abstract transaction properties.

## 5. Conclusion

MEC is one of the key applications of today's mobile information society, due to the large number of opportunities it opens to all stakeholders: MNOs, service and content providers, finance industry and consumers. In this paper, we discussed business and transaction modeling issues pertinent to MEC. MEC operates partially in a different environment than e-commerce conducted on the fixed Internet, due to the special characteristics and constraints of mobile terminals and wireless networks and the context, situations and circumstances in which people use their hand-held terminals. We focused on how these characteristics affect MEC business models and the important role that MNOs can play in this context. We also presented a formal model for MEC transactions and their properties. Our work is still in progress and it aims to contribute in the development of a standard working framework for mobile e-commerce in order to cope with the current market fragmentation and to ensure interoperability in the global mobile electronic market.

## Acknowledgements

The research was partially supported by the Finnish National Technology Agency (TEKES) and was carried out at the University of Jyvaskyla, Finland, in the framework of the Multimeet mobile (MMM) project, http://www.cs.jyu.fi/~mmm/. The authors would like to thank Prof. Jari Veijalainen for his contribution to earlier drafts of this paper.

## References

[1] P.E. Bernstein, V. Hadzilacos, N. Goodman, Concurrency Control and Recovery in Database Systems, Addison-Wesley, Reading, MA, 1987.
[2] Bluetooth, 2000. www.bluetooth.net (2000).
[3] N. Cravotta, Securing the wireless World Wide Web, 8/2000. Available on-line at http://www.ednmag.com/ednmag/reg/2000/08172000/17tt.htm (2000).
[4] Durlacher Research Ltd., Mobile Commerce Report, February 2000. Available at www.durlacher.com (2000).
[5] Nokia Products 2000. See www.nokia.com (2000).
[6] P. Timmers, Electronic Commerce: Strategies and Models for B2B Trading, Wiley, NewYork, 1999.
[7] J.D. Tygar, Atomicity in electronic commerce, in: Proceedings of the 15th PODC Conference, 1996, pp. 8–26.
[8] J. Veijalainen, Transaction concepts in autonomous database environments, Ph.D. thesis, GMD-Bericht Nr. 183, R. Oldenbourg Verlag, Munich, Germany, April 1990, ISBN 3-486-21596-5.
[9] J. Veijalainen, A. Tsalgatidou, Electronic commerce transactions in a mobile computing environment, in: Proceedings of IS 2000 International Conference on Information Society in the 21st Century: Emerging Technologies and New Challenges, Aizu-Wakamatsu City, Fukushima, Japan, 5–8 November, 2000, Springer, Berlin, 2000.
[10] E.K. Wesel, Wireless Multimedia Communications, Networking Video, Voice and Data, Addison-Wesley, Reading, MA, 1998.
[11] Wireless application protocol. See WAP-forum at www.wapforum.org (2000).

**Aphrodite Tsalgatidou** is a Professor at the Department of Informatics of the University of the Athens, Greece. She holds a M.Sc. and a Ph.D. in Computer Science from the University of Manchester (UMIST), UK. Her scientific interests include requirements engineering, business process modeling and reengineering, workflow systems, virtual enterprises and (mobile) electronic commerce and she has published several articles on these topics. She has chaired two conferences on e-commerce and business process reengineering, she is the managing editor of two electronic journals, she serves on more than 30 scientific committees and on the advisory boards for two international journals.

**Evaggelia Pitoura** received her B.Sc. from the Department of Computer Science and Engineering of the University of Patras, Greece in 1990 and her M.Sc. and Ph.D. in Computer Science from Purdue University in 1993 and 1995, respectively. Since September 1995, she is on the faculty of the Department of Computer Science of the University of Ioannina, Greece. Her main research interests are data management for mobile computing and multidatabases. Her publications include several journal and conference articles and a recently published book on mobile computing. She received the best paper award in the IEEE ICDE 1999 for her work on mobile agents. Evaggelia Pitoura has served on a number of program committees and was program co-chair of the MobiDE01 and MobiDE99 workshops held in conjunction with Sigmod 01 and MobiCom 99, respectively.