Clashing over the NFC Secure Element for Platform Leadership in the Mobile Payment Ecosystem

Jan Ondrus
ESSEC Business School
ondrus@essec.edu

Abstract

This paper explores how the control over the Near Field Communication (NFC) secure element could provide platform leadership in the mobile payment ecosystem. During a decade of trials, firms from different industries struggled to launch proximity mobile payment platforms to pay at physical point-of-sales. Mobile payments platforms are complex socio-technical systems that need to combine an array of resources and capabilities from different firms. The implementation of NFC platforms requires to store financial credentials of consumers in a secure element (SE). The industry designed several competing platform architectures for the secure element: SIM-centric, device-centric, SD-centric, and Host Card Emulation (HCE). Each of these alternatives gave more or less control to the different firms involved in the mobile payment ecosystem. For years, firms promoted the architecture which would give them more control. Over the time, this struggle just hindered the emergence of successful mobile payment platforms. The recent emergence of new actors (e.g., Google, Apple, Samsung) raised optimistic prospects for future developments. This paper provides a strategic analysis of the mobile payment ecosystem by studying the relationship between SE architectures and the establishment of platform leadership by incumbents and emerging actors.

1 Introduction

Since the late 1990s, mobile network operators (MNOs) and mobile handset manufacturers have worked to transform mobile phones into wallets. They realized that mobile phones have the potential to become a life-centric device for daily activities. With the emergence of electronic transactions in e-/m-commerce, mobile phones happened to be suitable and capable to drive remote/distant and proximity payments. For the latter, the plan was to digitalize the cash-based transactions. Limiting the use of cash could offer significant benefits for consumers and merchants. In addition, mobile payment platform providers could charge a commission on transactions which were previously free.

The implementation of mobile proximity payment platforms is so complex that there are only few schemes which entered the mass market. One of the main challenge faced by the industry was to find the appropriate technology. Proximity payments with a mobile device had to be easy to use. Unfortunately, the technology embedded in mobile phones did not offer obvious options. Some early platforms used SMS, USSD, QR code and infrared technologies. The most advanced solutions implemented an RFID-based technology for contactless payments. While SMS and USSD were standards used by the telecom industry, the alternative technologies had to be adapted to each implementation. As a result, the upfront investments were important for designing a mobile payment platform.

Some companies such as SK Telecom in Korea invested a massive amount of money into the infrastructure of Moneta [11]. They had to equip consumers with compatible handsets while merchants required new payment terminals (i.e., RFID dongles). Other companies were more fortunate due to an existing local standard. In Japan, NTT DoCoMo was able to roll out a mobile payment platform (Osaifu Keitai) based on Sony’s Felica technology. Felica was a de facto standard used by the electronic smartcard systems for public transportation and payments. Therefore, merchants and other service providers were already in possession of the adequate equipment. NTT DoCoMo had to provide mobile handsets with the technology and the services that users needed. The complex alignment between the technology required on the merchants and consumers side could be achieved without hurdles.
In other developed countries, numerous firms tried to emulate NTT DoCoMo’s Osaifu Keitai implementation success. Unfortunately, the majority of ecosystems did not have the appropriate market structure, technological infrastructure, and more importantly the institutional environment [18]. Having unfavorable settings, mobile payment platform providers had to improvised with either available technologies or to design a new one. Fortunately, in 2004, Nokia, Philips and Sony established the Near Field Communication (NFC) Forum. In 2006, the first Nokia 6131 embedding NFC technology was offered to the market. The NFC Forum aimed at providing an international standard for mobile services, especially mobile proximity payments. Today, all major handset manufacturers are supporting the NFC standards. Moreover, the financial industry issues more actively contactless EMV payment cards, which are compatible with NFC. Consequently, merchants are progressively getting equipped with contactless NFC payment terminals. With the current technology trends in the mobile and financial industries, the whole mobile payment ecosystem will inevitably be NFC-compatible in the next few years. The ongoing deployment of NFC follows a technology-push, as opposed to a market-pull strategy. Consumers and merchants will get the NFC technology in their hand even if they did not specifically asked for it.

The emergence of NFC brought encouraging prospects for the mobile payment providers. Yet, an essential issue had to be solved: the implementation of the secure element (SE). The NFC standard has several designs to secure transactions [6]. This paper focuses on the three most popular secure element architectures: SIM-centric, device-centric, and Host Card Emulation (HCE). In other words, the consumers’ credential can be stored in different places, under the control of different stakeholders. The firm which controls the secure element would own the consumer relationship and more likely to become the key actor in a given ecosystem.

Traditionally, the payment ecosystem is dominated by the banks and financial institutions. With NFC, their established position in the market could be questioned. When MNOs and handset manufacturers master the technology, they have no incentive to promote an NFC architecture that would give way control to incumbents actors in the payment industry. Therefore, the GSMA promoted the SIM-centric model1, which naturally gave more power to MNOs. On their side, the handset manufacturers endorsed the device-centric (also called SE Embedded) architecture. The secure element is stored on the motherboard of the phone itself. Therefore, manufacturers can leverage more control over the NFC implementation. Google used such architecture for its Wallet and the Nexus S. By having the SE inside the Nexus phone, Google could rule the ecosystem. More recently, Apple used a hybrid model which is based on the SE embedded architecture and cloud-based model. While a secure element is present inside the iOS device, the payment credentials are uploaded ”over the air” onto the phone using a tokenization mechanism. This architecture creates an ecosystem where MNOs have no control over the mobile payment platform. Based on these few example, one can start to understand the importance of the SE implementation to shape mobile payment ecosystems.

This paper explores how the control of the NFC secure element can provide platform leadership in a mobile payment ecosystem. By systematically analyzing the different SE implementations and their consequences on the participating stakeholders, this study provides an explanation of how mobile payment ecosystems can/will develop in different contexts. The use of framework based multi-sided platforms theory aims at structuring the analyses. Moreover, several real-world cases are used to illustrate the different architectures and the resulting ecosystem configurations. This topic is timely as new entrants, such as Google and Apple recently started to offer mobile payment services and strive for platform leadership.

In the next section, we discuss the concept of mobile payment and related research. We also identify some issues that we want to address with our study. In Section 3, we present our framework based on multi-sided platform theory. Section 4 provides a description and analysis of each SE implementation. Finally, Section 5 concludes the paper with some discussions and proposes further research.

2 Background

Although the term “mobile payments” is often used for different types of scenarios, this paper focuses exclusively on mobile proximity payments. That is paying for goods and services using a mobile phone at a physical point of sale (POS) terminal. These types of payments can happen offline as the mobile device embeds the payment means, such as a debit/credit card or a stored value account. Nowadays, many refer to these platforms as mobile wallets.

Academic research on mobile payments started

soon after Coca Cola experimented vending machines which could accept SMS payments in Finland in 1997. After almost a decade of research, a literature review was published to provide guidance to researchers [8]. The main motivation of this review was that mobile payments issues were not fully understood by the academic community. Interestingly, the review revealed that a considerable number of the articles focused mainly on two issues: technology (e.g., security, protocols, systems architectures) and consumer adoption. Despite their respective importance, these focuses could only provide a partial view on mobile payments. Already in 2005, several researchers insisted that multi-perspective approaches should be used to properly study mobile payments [20, 26]. Following this suggestion, a number of multi-perspective frameworks were proposed to study more than one issue at a time [3, 7, 23, 25]. More recently, Hedman and Henningsson studied how technological payment innovations influence payment ecosystems. They also explained that digitalization of payments has caused ecosystem instability by impacting the competitive and collaborative dimensions of ecosystems [13]. In his paper, Lim describes the different efforts of firms involved in mobile payment to form consortia [16]. In addition, the research of Kazan and Damsgaard constructed frameworks to study digital payment systems and to analyze the strategies of current market actors [14, 15]. To further understand the role of the actors, Gaur and Ondrus examined what strategic assets make the presence of banks and financial institutions mandatory in mobile payment ecosystems [12]. Moreover, with the arrival of new market actors to mobile payment markets, most notably Google, Paypal, Apple and AliBaba, Ondrus et al. have started to investigate their impact on ecosystems [22]. Taking into account national specificities, Magnier-Watanabe proposed that the development of mobile payment systems depends on country-level institutional constraints (i.e., compliance with industry- and resource-based/dependent views) [18].

A number of recent studies identified the reasons for the failures of mobile payment platforms. The lack of collaboration between multiple stakeholders, the difficulties in finding win-win business models, and the lack of standardization [2, 9, 11, 17, 24] seems to have impeded the developments of mobile payment platforms over the years.

In spite of a number of studies on mobile payments, the "mobile payment" phenomenon still raises numerous questions. The efforts of investigating issues from multiple perspectives improved the general understanding of mobile payment ecosystems. However, more research has to be done to fully grasp the current issues that need to be solved in order for mobile payments to enter the main market and reach mass adoption. We need to better understand how technology choices impacts strategy and vice versa. As explained previously, the secure element implementation greatly influences the ecosystem architecture. However, the problems go beyond the simple inclusion or exclusion of specific stakeholders. The growth rate of a mobile payment platform can be accelerated, decelerated, or even stopped based on a technology or strategic choice.

3 Multi-sided platform theory to study mobile payments

The concept of multi-sided platforms is essential when studying mobile payments. Similarly to numerous current high technology products and services, mobile payments are examples of technology platforms [14]. A platform is defined as "a set of stable components that supports variety and evolvability in a system by constraining the linkages among the other components" [4]. The mobile payment platform and the infrastructure represent the set of stable components. The other components are the merchants and consumers who are connected by a buyer-seller linkage. Mobile payment platforms are multi-sided as they bring together more than one group of users (i.e., consumers, merchants) to the platform.

A challenge for multi-sided platforms is the "chicken-and-egg" problem that has to be solved in order to grow [5]. More precisely, consumers would only join a mobile payment platform if a decent number of merchants already joined, and vice versa. There are different strategies to solve this problem. For example, the providers can equip the merchants for free while consumers start adopting the scheme with some discounts. Depending on the growth rate of each user groups, the providers have to provide more or less incentives. A harmonious growth of a multi-sided platform requires a careful orchestration of these incentives (i.e., discounts, subsidies).

Due to co-dependence issues, mobile payment platforms require a complex coordination between multiple interdependent firms from different industries. Each stakeholder control or possess resources that are required to design the platform. Depending on their architecture and SE implementation, mobile payment platforms can involve (or not) actors such as mobile network operators (MNOs), banks, financial institutions, payment networks, payment service providers, technology providers, mobile handset manufacturers,
payment terminal manufacturers and other third parties (e.g., Trusted Service Manager).

Eisenmann et. al. [10] proposed a classification of the different actors involved in a platform. They start with platform sponsors which represent the group who exercise property rights, determine platform participants and develop technology. Then, there are the platform providers who serve as the primary point of contact of the platform users with the platform. Finally, there are the demand-side and supply-side platform users. Based on the classification, Ondrus et al. proposed three levels: i) the provider level, ii) the technology level, and the user level [21]. Figure 3 illustrates the different levels for a generic mobile payment ecosystem. The provider level comprises the strategy issues and the roles of platform providers such as MNOs, financial institutions, banks, handset manufacturers. The technology level deals with issues such as standardization and interoperability. The user level is mainly concerned with the dynamics of the two groups of users that are the consumers and merchants.

4 A strategic analysis of NFC secure element architectures

This section provides a strategic analysis of the main SE architecture: SIM-centric, device-centric and Host Card Emulation (HCE). We did not include the SD-centric and other models that are not actively used by the industry. The framework previously introduced is used to evaluate the impact of each implementation on the providers, the technology, and the users. Some real-world examples illustrate the different architectures in order to provide a more realistic picture of the current situation. These short case studies were built using secondary data from academic research papers, websites, industry reports and white papers.

4.1 SIM-centric model

The SIM (Subscriber Identity Module) card has been a secure element of choice for MNOs for many years. The SIM card stores the information used to authenticate and identify a specific user on a mobile network. In other words, a SIM card is a secure passport that allow users to access mobile networks. The SIM card also defines the relationship between the issuer (generally an MNO or MVNO) and the consumers.

The SIM card is comparable to the EMV chip found on payment smartcards. Therefore, MNOs proposed to securely store financial credential on the SIM card for NFC mobile payment platforms. Obviously, this option would give MNOs an important role to play in the ecosystem, as exclusive owners of the SIM card.

In terms of the provider level, the SIM-centric architecture gives a clear advantage to MNOs. They can control any information that is uploaded to the SIM. Therefore, other actors are obliged to work with MNOs, and vice versa. The collaboration can be organized in two ways: a joint-venture or a rental-deal. In a joint-venture, each involved firm is supposed to provide access to their respective resources and actively participate to the design, implementation and promotion of the platform. In a rental-deal, the mobile payment platform providers have to pay a rental fee to MNOs to securely store financial credentials and any applications on the SIM card.

At the technology level, MNOs have to upgrade their SIM cards to be compliant with NFC standards. These cards are commonly called NFC SIM cards. Moreover, in the SIM-centric model, the handsets have to be certified by the mobile payment platform providers. The hardware and software implementations of NFC could differ and generate incompatibility with the overall infrastructure in place.

The user level is concerned with the availability of compatible devices and an easy registration/enrollment process. Unlike merchants with payment terminals, consumers are more sensitive about the handset they use. Therefore, if the mobile device of choice is not NFC compatible, consumers are unlikely to compromise for another model, just because it would be certified for mobile payments. Therefore,
MNOs and other mobile payment platform providers are dependent on the handset manufacturers. Concerning the registration process, the SIM-centric approach can be confusing as eligible consumers have to first possess the right mobile device, obtain a new NFC SIM card, register their payment card with their MNO and bank. The payment experience is as easy as a contactless card payments. Consumers just have to tap their phone on the terminal.

The SIM-centric approach has been chosen for the Cityzi project in France [1]. MNOs, financial institutions and other firms (public transportation, parking, museum, and retailers) created a consortium to offer NFC services throughout whole French cities. Despite 6.5 million consumers using with one of the 75 NFC mobile phones and about 20% of merchants equipped with contactless payment terminals\(^2\), the Cityzi model still struggles to impose itself in the French market. Worse, banks and MNOs are now launching their own wallets. A group of major banks have formed an alliance to offer the Paylib mobile wallet\(^3\). Paylib is now exclusively used for online purchases, while extension to physical payments with NFC is not excluded. The participating banks have their own NFC mobile payment wallet (which currently uses a SIM-centric model). French MNO Orange also launched their own wallet called Orange Cash\(^4\). As for now, the SIM-centric architecture has been chosen. While these wallets offer a rather easy registration process for Android users, the consumers still need to possess a compatible/certified mobile device.

Due to the complicated process and hardware requirements, the SIM-centric approach still experienced a limited diffusion. Moreover, this NFC architecture requires a collaboration between the telecom and financial industry. Previous research [2, 9, 11, 17, 24] showed that this collaboration is difficult, if not impossible, to orchestrate. Despite excellent technological features, the SIM-centric architecture struggles to impose itself as a dominant design. Other actors in the ecosystem tend to favor other architectures that gives less control to the MNOs.

4.2 Device-centric model

The device-centric model implies that the secure element is directly embedded into the device. In other words, there is a secure chip on the motherboard of the mobile device. The security features of the chip are rather similar to the ones on a SIM card. However, the chip cannot be removed or transferred to another device.

In terms of the provider level, the device-centric model provides more control to mobile handset manufacturers in the mobile payment ecosystem. An embedded secure element in the device allows actors to free themselves from the dominance of the MNOs, which they gain with their SIM card. The difficulty of implementing this model is that MNOs can refuse to distribute the handsets to their users. Worse, MNOs can even ban them from their network. Therefore, there is still room for a collaboration with MNOs, at least for the distribution.

The technology level is mainly concerned the availability and compatibility of the handsets for consumers. The implementation of the secure element should follow the EMV secure specifications in order to be accepted by the financial institutions. Moreover, there is not real dependance on the SIM card as the secure element is located in the device.

For the user level, consumers have a limited choice of exiting mobile handsets that embed NFC secure elements. Due to this restriction, only manufacturers with a large market share (e.g., Apple and Samsung) can impose the device-centric model.

In 2010, the first version of Google Wallet required a Nexus phone which embedded a secure element. Therefore, only a handful of consumers could use the service, as the platform growth was directly linked to the sales of the Nexus line. Apple Pay use a variant model based on an embedded secure element within the iPhone 6. Apple could justify this strategy as iPhones dominate the sales ranking, without the same device fragmentation that other manufacturers experience by selling multiple phone models.

The device-centric model is difficult for one firm to implement as it requires a large market share in the mobile device industry. Google tried to impose this model with their Nexus phone. However, the sales did not allow to grow the platform. Later, Google decided to change their strategy by providing in newer versions of Android a new software layer that can support the Host Card Emulation (HCE) architecture.

4.3 Host Card Emulation (HCE) model

Host Card Emulation (HCE) is a newer architecture introduced in 2013 to store the critical financial information either in a remote location (i.e., cloud), in the payment application itself, in a secure element or in a secure environment within the mobile device [19].
Table 1: Summary of the analysis

<table>
<thead>
<tr>
<th>SIM-centric</th>
<th>Device-centric</th>
<th>Host Card Emulation (HCE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provider</td>
<td>- Give an advantage to MNOs</td>
<td>- Manufacturers are in control</td>
</tr>
<tr>
<td>- Require inter-industry collaborations</td>
<td>- Could partially bypass MNOs</td>
<td>- More flexibility</td>
</tr>
<tr>
<td>Technology</td>
<td>- Need to distribute specific SIM</td>
<td>- SIM-independent</td>
</tr>
<tr>
<td>- Need to certify handsets</td>
<td></td>
<td>- Require online connection</td>
</tr>
<tr>
<td>User</td>
<td>- Complicated registration/enrollment</td>
<td>- Limited choice of handset</td>
</tr>
<tr>
<td>- Great user experience</td>
<td></td>
<td>- Less intuitive user experience</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Possibility to add extra services</td>
</tr>
</tbody>
</table>

However, the secure element neither have to be located in the SIM, nor in the device. This architecture was adopted by Android in its version 4.4 (KitKat). The mobile wallet/application can virtualize and emulate a payment card when using the NFC interface to communicate with contactless payment terminals. More than for payment purposes, HCE can emulate any loyalty, transportation, access or identification NFC card.

The HCE architecture solves a number of issues at the *provider level*. As previously discussed, mobile payment developments have been hampered by the complexity to orchestrate collaboration in the ecosystem. HCE offers the possibility to organize a mobile payment platform without the mandatory participation of the usual actors, especially MNOs.

In terms of the *technology level*, HCE has several limitations. For example, the card emulation cannot work when the phone is off, as opposed to the SIM- and device-centric models. Moreover, there is usually the need for an online connection to conduct a payment transaction. The experience could be slightly more complicated than the use of a SIM-centric mobile payment platform.

At the *user level*, consumers need a recent Android mobile device equipped with NFC to join a mobile payment platform based on HCE architecture. The enrollment of consumers remains straightforward as the main requirement is to install an application and register the payment credentials. There is no need to visit an MNO or bank branch to signup or acquire additional pieces of equipment.

From mid-April 2014, Google decided to end its support for physical secure element (i.e., device-centric) and migrate the Google Wallet to HCE. The "tap and pay" function would only be possible with HCE. The previous architecture had a slow growth due to the lack of compatible device. The adoption HCE aimed at reviving Google Wallet by offering more possibilities.

In general, HCE allows more flexibility to design a mobile payment platform. The interactions between the actors in the ecosystem are simplified and enable a clearer business model. HCE generated new hopes for the future development of mobile payments. However, there are still technical hurdles to solve in order to propose a simple solution to consumers.

5 Conclusion

The previous Section presented a strategic analysis of the secure element architectures. A summary is presented in Table 1. Using a multi-level framework, different perspectives such as the strategical issues, the technological aspects, and the user-related challenges have been examined. Each model exhibits advantages and disadvantages, and can give more or less power to specific stakeholders. As illustrated with examples, platform leadership is mainly obtained with an exclusive control of the secure element. The classical approach (SIM-centric) offered more opportunities to the MNOs, while device-centric approach gave more control to the mobile handset manufacturers. With the emergence of HCE, new ecosystem configurations are possible. Therefore, it is possible that platform leadership will not necessarily obtained with the control of the secure element.

Apple and Google demonstrated that it was possible to shake the status quo in the mobile payments ecosystem. By using a different technological approach, they have been able to bypass the barriers that were used to defend MNOs’ territories. The same two actors already defeated MNOs’ "walled garden" strategy in the mobile application market. There is a chance that history is about to repeat itself with the mobile payment market. Today, newcomers are benefiting from a favorable regulatory environment. The whole U.S. payment ecosystem has

---

5http://www.nfcworld.com/2014/03/17/328926/google-wallet-ends-support-physical-secure-elements/

6http://www.totalpayments.org/2014/04/10/google-wallet-saved-hce/
to move to EMV, forcing the merchants to update their payment terminals. Moreover, in Europe and Asia, the deployment of NFC payment terminal continues to grow[19]. These migrations will certainly boost the adoption of contactless payments.

The different roles of the stakeholders in the mobile payment ecosystem was dictated by the control of the financial credentials of consumers. With a SIM-centric approach, the market could not evolve if not all the stakeholder agreed on a win-win business model. Today, with the emergence of varied SE architectures, we can hope that consumers will even be able to choose from a selection of mobile payment platforms in the near future. Competition between the available wallets along multi-homing possibilities will definitely instigate the growth of mobile payments.

This short paper offers a primary study on how technology and strategy are inter-related in mobile payments. Future research should further investigate the intricate alignment process between technology and strategy. More cases should be studied in order to better understand how different platforms are emerging. Due to the arrival of Apple Pay, we can expect that more mobile payment initiatives based on NFC technology will be launched in the next years. It will be essential to fully measure the impact of a technology choice (i.e., secure element implementation) on the ecosystem.

References


