The Smart Phone and New Opportunities for Near Field Communication (NFC)

Roy Want
Google Research
1600 Amphitheatre Parkway
Mountain View
CA 94043

Bill N. Schilit
Google Research
1600 Amphitheatre Parkway
Mountain View
CA 94043

Abstract: **This paper provides an overview of Near Field Communication (NFC) technology. A focus is the integration of NFC with Smart Phone applications, and some of its expected use cases, including smart posters and electronic payments. It also considers future implications, opportunities, and challenges that we might expect from NFC in the coming years.**

1. Introduction

Near Field Communication (NFC) is a technology that allows proximate, contactless communication between two computers, or a computer and a passive electronic tag (no battery). It is a standard established under the NFC Forum [1], which was founded in 2004 by Nokia, Sony, and Philips. The primary goal of the standard is to promote data sharing between mobile computers, such Smart Phones, and network services.

By design, NFC has a nominal short communication range of about 5cm (10cm max) enabling both purposeful data transfer and physical security. For example, the short-range characteristic ensures that two devices only connect if they are deliberately brought in close vicinity to each with their antennas aligned, thus mitigating accidental connection. It also provides physical security as a remote eavesdropper cannot overhear the communication, or trick a user into connecting with an unseen device.

By comparison, longer-range wireless technologies require the user to initiate a discovery process, which takes some time, and then generates a list of devices that a user must select from to establish a connection. Furthermore, device names in the list are not always as clear, or distinct, as they should be e.g. many original WiFi access points had the same station name (SSID) based on the manufacture's default, such as 'linksys', which resulted in a confusing user experience.

It should be pointed out that NFC has a relatively low data-rate of up to 424kbps, and is not intended to replace modern WLAN standards at 100Mbps+, but instead support a set of use cases that would be difficult using existing technologies. These include electronic business-card exchange, smart posters, and electronic payments. In addition, NFC can be used to augment existing WiFi and Bluetooth radio standards, by enabling more intuitive connections using an NFC data exchange protocol to directly transfer MAC addresses and authentication tokens, thus avoiding a discovery and pairing process.

The paper is organized as follows: In section 2, we discuss the principles of Near Field Communication, and in section 3 we describe NFC tags and their integration with semiconductor memory. In section 4, we examine the characteristics of the data exchange format for NFC messages, and in section 5 touch on the main NFC protocols in use. Section 6 describes some novel capabilities that are made possible using NFC with modern smartphones, while in 7 we reflect on the challenges that remain for the technology. Finally, in section 8 we conclude.

2. Principles of NFC Communication

NFC is a subset of the group of technologies broadly referred to as Radio Frequency Identification (RFID) [2]. These include non-contact automatic identification systems based on a variety of standards that work at different ranges and frequencies. Each can be classified into Lower Frequency (LF) 125-135KHz, High Frequency (HF) 10-100MHz, or Ultra High Frequency (UHF) 900-3GHz. NFC operates at 13.56MHz in the HF band based on the ISO/IEC 14443 standard for physical layer communication. It is part of the band reserved for Industrial, Scientific and Medical (ISM) applications in the US, and for this reason it is used by most HF RFID standards. The transport mechanism used by both LF and HF systems is near field coupling; whereas UHF systems are based on far field communication, or electromagnetic propagation and capture.

In near field systems, data and energy are transferred from the transmitter to the receiver using the principle of magnetic induction. It is much like the operation of a household AC transformer having a primary coil, the transmitter, and a secondary coil, the receiver. An alternating voltage applied to the primary, results in an alternating magnetic field. This in turn induces an alternating voltage in the secondary. If a load is applied to the secondary, the resulting alternating current generates its own magnetic field that apposes the magnetic field in the primary, resulting in an increased current in the primary to compensate, thus energy is both transferred and conserved in the process. With an NFC communication system the primary and secondary coils are effectively moved apart, but the magnetic field extends beyond the confines of the primary coil, becoming weaker with distance. In fact, the field strength decreases at distance **r** by a factor of $1/\mathbf{r}^3$ measured along the center axis of a circular coil, and thus the range at which it can effectively induce a voltage at the secondary is very constrained. To design an effective communication link, the size of the transmit and receive coil area will also define the amount of power that can be captured; and the sensitivity and power needs of the receiver also play a role in building a functional system. In practice for pocket-sized smart phones, and a 2.5cm diameter passive NFC tag, the usable range works out as 2.5-5cm, based on today's technologies.

If two active devices are in communication, bi-directional data transfer needs to occur. This can be achieved by amplitude modulating the primary signal resulting in a corresponding amplitude modulation of the induced voltage of the signal in the secondary, and the receiver can demodulate the output to recover the digital signal with a simple rectifier circuit. The return path is achieved by a technique called *load modulation*. In this case the receiver applies a variable load to the receiving coil, usually a transistor placed across the windings, that when turned on increases the coils load, and hence results in a greater current flow.

Thus a digital signal applied to this transistor will result in a varying secondary coil current, and a sympathetic increase in current at the primary coil. This small signal can then be used to recover the secondary's signal by measuring the voltage drop across a resistor placed in series with the primary coil. Thus two-way communication can be achieved between primary and secondary. The high layers of the NFC standard define the format of data packets exchanged in this way, and the protocol for two-way interaction.

If, however, the primary device is active but the secondary is passive (e.g. an NFC tag), the passive receiver also needs to derive its power from the energizing signal of the primary. This is achieved by rectification of the induced voltage at the secondary feeding into a capacitor reservoir that accumulates charge until its potential reaches the same as the peak induced-voltage. The output can then be regulated down to a lower stable voltage source that supplies the device's electronics. For an NFC tag one of the primary limitations on range is the amount of energy that can be captured in the capacitor reservoir at the receiver. Potentially, two active devices could communicate at greater range, as they do not have this power limitation. However, the use model does not require this capability, and the systems are designed for similar range for the passive and active case.

3. NFC Tag Integration with Read/Write Memory

At the simplest level, when an NFC tag is read, it will return a unique identification number with nominally 128 bits of data. However, many NFC tags also incorporate electrically erasable and programmable memory (EEPROM) that can be written to, and read from, by an NFC reader (see Figure 1, an example tag). Such tags typically allow the storage of between 64 bytes and 4K bytes of user data. Memory tags provide a user with a unique capability to associate data with a place, which can be retrieved with the simple tap of a read. For example, a world-wide-web based Unique Resource Locator (URL) can be stored in a tag allowing it to reference a web page stored anywhere in the Internet; see section 4. As a result, NFC may be one of the first widely deployed standards that can be used to support the Internet of Things, or Tangible Computing, two concepts written about by the research community since the early 1990's [3, 4, 5 6]. These visions describe how physical objects can be linked to information, or control mechanisms, provided by available cloud services.

Figure 1: A passive NFC tag with 2K bytes of memory.

NFC tags can be written multiple times, or just once, and then locked down. This puts stored data under the control of the tag owner, which is important as a tag may be used for business applications, and should not be easily modifiable by the general public. Another feature provided by some tag manufacturers is a 'kill-switch' that enables a tag to be disabled. When an NFC tag is used by a vendor to label a product, it is often desirable to disabled it at the point of sale to prevent it being read and identified by a third party at a later time. This is because active product tags might be a personal security concern for people who have purchased, and then carry around, high-value items such as small state-of-the-art consumer electronics. For NFC, with a typical max read range of 5cm, this is not as much of a problem as UHF RFID tags, as promoted by EPC Global, which can by design be read 3-6m away. However, with specialized readers it is possible to extend an NFC tag's read range to 1m, which may be enough to automatically interrogate the contents of a bag as it is carried through a doorway, and would be another tool thieves might use to find victims to rob.

4. NDEF data exchange format

The NFC forum has defined a basic format for messages that can be exchanged between active NFC capable computers, or from a computer to a passive NFC Tag. The format is called

NFC Data Exchange Format (NDEF), and defines the fundamental unit to transfer data using NFC between devices.

The protocol allows any number of NDEF messages to be stored in a tag up to its memory capacity. Each NDEF message can be any length since it is composed of an arbitrary number of NDEF records. Each NDEF record contains a length field that defines the storage requirement, and a record type field (RTD) that identifies characteristics of the data it contains. There are several defined RTDs in the standard that include: Text, Universal Resource Identifier (URI), Generic Control, Signature and Smart Poster.

- **Text Type**: A text string encoded in ASCII or Unicode, and a field for the language type in use.

- **URI Type**: A Unique Resource Identifier encoded in ASCII. URI headers are already a well defined set of strings, such as "http://www", "telnet://", "tel:" or "sms:". The URI RTD type incorporates a compression field that allows these prefixes to be compressed into a single byte. Although a tag may contain 2K bytes of memory, it adds to the cost. For large deployments smaller capacity tags are used with only 64 bytes of memory, and therefore the compression byte is important to help represent both the header and the body of the URI, given limited space.

- **Generic Control Type**: This is a generalized mechanism to start an application specified by an encapsulated URI. It also provides information about where to store the received data, and what state should be modified on the target device.

- **Signature Type**: This is a format for digitally signing a set of NDEF records. It specifies the algorithm used to create the signature, and the certificate type/format that is included. A signature can be used to mitigate the problems associated with an unknown URI. If the NDEF message contains a valid signature from a standard certificate authority, a user may be inclined to proceed with the transaction, knowing the tag contains valid data from a known and trusted source.

- **Smart Poster Type**: One of the promising uses of NFC tag technologies is to enable the creation of active printed material, usually referred to as *Smart Posters*. Essentially, printed materials are made smart by placing adhesive NFC tags behind key regions of text, graphics, and images; and may be used with posters, advertising, signage, and bulletins etc. A user with a smart phone can wave, or tap, their phone over these regions to acquire more information. The smart poster RTD is really a meta-type indicating the presence of other embedded records that describe the poster. In this case the Text Type record contains the title of the poster, a URI defines where the webpage extending the poster can be found, and a Generic Control Type contains the recommended action to be performed.

It should be noted that the NFC Forum is actively evaluating and defining new RTDs extending the standard in the future.

5. NFC Protocols currently in use on smart phones

NFC is a base technology that can potentially support a large number of high-level protocols. At this time, in the US market, there are only two NFC capable phones: the Samsung Nexus S, and the Galaxy Nexus. Below is a short summary of two protocols currently in use on these phones, but this is likely to grow as more products are introduced and other standards appear.

NDEF Push Protocol (NPP): The NDEF Push Protocol (NPP) is designed to push an NDEF message from one device to another. It is a key part of the peer-to-peer (P2P) NFC application used for tag reading/writing and vCard exchange. The protocol is built on top of a lower-layer protocol called LLCP, and constitutes a mechanism for one-way transfers of NDEF messages from client to a server. NPP capable devices must support the server, whereas the client is optional.

Google Wallet: Probably one of the most significant driving forces behind NFC is the opportunity to support financial transactions between a smart phone and a point of sale (POS) terminal at a store checkout. Some of the first trials are currently in progress to evaluate the effectiveness of the Google Wallet system in an Android phone. The system is based on MasterCard's PayPass protocol and the POS terminals can be found at multiple vendors. There are three key advantages of the system: 1) your phone never needs to leave your hand, 2) it only works at very close proximity, and 3) if you accidentally tap multiple times, only one payment is recorded.

6. New Opportunities for NFC and Smart Phones

Smart phones are powerful computers in their own right containing a high performance processor (sometimes multi-core), a high-quality display, a multi-point touch screen, several different radio technologies 3G/4G, WiFi-n, Bluetooth, a rich set of sensors, and a camera. One of the opportunities for Smart Phones supporting NFC is to take advantage of the combined capabilities of these technologies to enable novel applications.

Another area that NFC can have real impact is enabling the vision of tangible computing [3, 4, 5, 6], allowing users to control many aspects of their environment by interacting with objects in the physical world. This is a goal often discussed in the ubiquitous and pervasive computing community, but despite a number demonstration systems described in the literature, it has not taken off. However with the rapid adoption of smart phone handsets in recent years [7]; and the confluence of 3G/4G metropolitan networks, NFC hardware, and NFC Forum standards, the time for tangible computing may have finally come.

At the simplest level, NFC tags can be used to trigger actions based on reading a tag's URI, and then executing the encapsulated command. Consider a smart home application that allows control to be achieved through a URL typed into a web browser, in turn displaying a control form. If an NFC tag contains the same URL, with a set of form parameters filled out, it will achieve the same result. For example, executing the URL stored in a tag can turn on a lamp:

www.myhomeserver.com/lamp?state=on

Another tag storing the complementary parameter value will turn the lamp off:

www.myhomeserver.com/lamp?state=off

This means that an on/off light switch can be established in any room of a house by simply mounting two adhesive tags on the wall without the need for additional wiring or remodeling; the only condition is that your phone needs to be with you.

The obvious question comes to mind, why not just run the browser application on the phone with the two lamp URLs bookmarked. This is certainly possible, but requires concentration from the user to activate the application, access the bookmarks menu (or type the URL if you don't have it bookmarked), scroll to the correct item, and then select it. Instead, NFC allows the phone to be used like a wand. We can take advantage of our spatial memory of where certain controls are in a room, just as we do with conventional light switches, and then with imprecise motor control move the phone over the programmed tags. It requires little focus compared to the attention needed to interact with a complex smart phone.

Considering this example is relatively simple; is it possible to use the same principles for a more complex control task? An extension of the previous example would be a lamp dimmer. One approach would be to use the tag to launch either an app or a website with a graphical dimmer control showing on the touch display, perhaps either a linear slider, or a circular wheel. Although this technique is very flexible, it still requires the user to pay attention to what appears on the screen. Another approach is to continue with the tangible computing metaphor, using the phone as a wand. A dimmer function can be achieved by scanning a tag that connects with the dimmer home service, and then uses the output of the on-board accelerometer to measure the phone's tilt angle relative to gravity. Thus a user can dim the lamp by turning their hand clockwise or anti-clockwise – see Figure 2.

To achieve an interactive dimmer function a URL with an updated parameter needs to be sent incrementally for every discrete angle turned, and this increment can also be defined in the tag. The brightness of the dimmer can be then fixed by tapping the phone against the tag a second time.

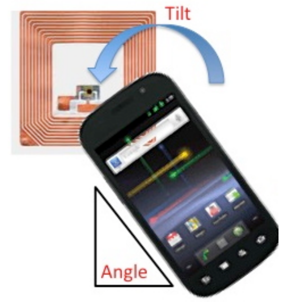

Figure 2: A smart phone reading a tag while being tilted to define an analog parameter that can control a remote device.

Another tangible implementation of the dimmer function can be achieved by emulating a physical linear slider. Instead of using a single NFC tag, a row of tags is mounted on a wall, each containing an increasing dimmer setting from left to right. Once again, the user is able to employ imprecise movements with low cognitive load to control the light setting. It is also possible to determine intermediate positions between the tags by double integrating the output of the accelerometer to track distance moved. The sensor itself has some error so this kind of inertial measurement could rapidly lead to distance errors, but in this case it only needs to be accurate enough to interpolate a value until the next tag has been encountered.

There are many other examples of tangible computing that can be implemented using similar principles to those described above. For example, a smart poster does not need to be limited to a few hot zones with information or ordering capability, the poster itself can become a rich user interface to make selections for a multi-dimensional task. Consider a poster that can serve as a kiosk and allow you to order a sandwich. There are various selections that need to be made, and a set of parameters that can have multiple values. First the type of bread must be selected, then the dressing, next the type of filing, whether cheese should be added, and if the sandwich should be toasted when complete. This could all be achieved by using a smart phone as a wand, touching labels on the poster that establish each parameter of the sandwich, and then allow the user to select the ingredient. There also need to be labels that start a new sandwich, and remove the last ingredient chosen in case of an error, and then a final submission button when the order is ready to be sent to the kitchen. Fortunately, most of the building blocks that enable this capability are already defined by standards. However, in this poster / kiosk example, industry needs a standard way of chaining commands, obtained from a sequence of NFC tags, in order to generate a single composite command that can be executed.

Consider another illustrative example; an electronic calculator. Here each calculator key is represented on a poster with NFC tags mounted under each graphic on the backside of the poster. The idea is to use the phone like a wand to enter the required arithmetic sum, and then on pressing the '=' key, the result can be printed, or called out using text to speech on the phone.

This can be achieved by establishing a URI type that is not executed at the time it is read, but instead tells the reader it is part of a chain of data that will later be resolved into a command. To enable the chaining process, we also need a URI that will begin and end a chain. Using the calculator as an example, the following URIs can be used in sequence to calculate the expression 6*7.

nfc:chain?begin=1&cmd=calculator

This URL begins a new chain, cancelling any prior chains or partial chains that may have been stored. This is the URI programmed into the tag under the clear key "C". The tag also tells the system that a new chain is about to start for the "calculator" function. The command name can be displayed on the screen (or spoken) so that the user is reassured the calculator function is in progress.

nfc:chain?param=6

The next entry is the URI under the '6' key, and as a result, the character '6' is appended to the empty chain string. Again to reassure the user they are making progress, the "6" can be shown on the display next to the string 'calculator'

nfc:chain?param=*

Now the second character is appended to the chain string resulting in an intermediate value of "6*", then

nfc:chain?param=7

resulting in a chain string of "6*7", also displayed next to the word calculator.

nfc:chain?end=1&cmd=calculator

Finally, when the phone scans the "=" key on the poster, it can determine the chain has been completed, and that the expression contained in the string can be executed; the command key/value is repeated in the URI to confirm the desired operation. The data accumulated is then repackaged into another final URI:

nfc:calculator?input=6*7

This string is fed back into the URI evaluator, and assuming calculator is determined to be a valid function for this device, it can be executed with the key/value parameters that follow. In this case, there is a single parameter input = "6*7" which is passed to an arithmetic string evaluator producing the result "42". The calculator function could then clear the phone's screen, and print the word "calculator" followed by the complete expression "6*7=42", or just announce the result audibly.

It is also possible that in the middle of a chain, the user may make a mistake choosing the next character of the expression, or the phone may not be positioned correctly and read the wrong tag. In which case, a 'delete' function is also needed. Although this could also be achieved with a soft button on the phone, keeping with the wand metaphor a poster key with the "DEL" graphic can also be provided, served by an NFC tag containing the URI below:

nfc:chain&delete=1

If this is read at any stage during the formation of the chain, the last entry that was added to the chain needs to be deleted. Although, in this example, each parameter is a single character, in other situations parameters may contain multiple symbols represented by whole words. For example, a scientific calculator may have employed SIN, COS, and TAN functions, in which case the delete function needs to remove the last set of symbols added. In practice this means that a chain of parameters is best implemented as a FIFO stack enabling a user to delete an arbitrary number of parameters added in reverse order, and then continue to build the chain up again from that point.

Furthermore, if a chain has been executed there is no need to clear it at that point in time, only when a new chain is started. In the case of the calculator, this allows intermediate results to be calculated and displayed, and then the output to be included as part of the next calculation. Thus in the "6*7" calculation after the '=' tag is read, a user can continue to read the '/' and '2' key, and then selecting the '=' key again, obtains the result "6*7/2=21".

7. Challenges for NFC Adoption

Although there is a rich set of tangible computing applications waiting to be written that take advantage of unique properties of NFC tags, there are also a number of problems and concerns that arise. Here we consider three of the main issues.

First, tangibility gives rise to uncertainty about what is tagged and what is not. The virtue of hiding a tag behind a poster or graphic leads to invisibility, and as a result a user may well be uncertain where to scan their smart phone, and what the consequences of performing a valid scan will be. Many of these issues can be resolved by using conventions, or graphical clues, to guide the user. For example, the world-wide-web used the underline style to indicate that text was hyperlinked, and clicking an underlined word would take you to a related webpage. In practice, many other features of a page can be hyperlinked, such as pictures and sub-regions of images, and users have become accustomed to moving their mouse over a page to find the embedded links. It is possible that tangible computing may follow a similar path, starting out with well-defined graphics describing their function, but as many people become accustom to using it, allowing more subtle indications. The NFC Forum has already defined a symbol shown in Figure 3 to alert users that a computer or poster is an NFC active device, but other conventions may need to be put in place to support tangible computing.

Figure 3: The N-Mark symbol designating an NFC device

Another issue relates to the rate at which a smart phone scans for tags or peer devices. In practice, an NFC scan consumes moderate power, and the more frequently it scans, the faster a mobile device's battery will be drained. As a result, most NFC implementations scan every 1-1.5s, which means that rapidly moving an NFC phone over a poster is like to miss tags that were passed over in-between scans. This problem can be solved by adaptively increasing the scan rate as a result of the device's motion. Most smart phones already contain a motion sensor, in the form of an accelerometer, which can be used for this purpose. However, a fixed scanning scheme is adequate for electronic payments, so it depends if the phone manufacturer considers tangible computing use cases to be important enough, and invest the resources to implement adaptive NFC scanning.

One of the biggest concerns users may have with NFC tag based applications is security. The URI contained in a tag could easily point your phone at a web page containing malware, or perform an unexpected operation on your phone, that costs real money, such as a pay-to-call number. There are already examples of nefarious attacks for QRCode applications, in which a fake QRCode sticker is placed over the original [8]. An unsuspecting user will scan the code with their camera phone without being aware of the scam. A criminal might even create a fake site to deliver the malware, and then redirect the user webpage to the intended site, thus further masking the deception. Any tag reader, whether it be a QRCode or NFC tag, needs to have protection in place to warn users about untrusted URLs, or potential access to phone resources that cost money. Ideally all URIs need to be signed to check they do indeed originate from a bone fide source.

Finally, as proximity payment and electronic wallets are one of the important driving forces behind NFC adoption, we also need to overcome the concerns held by the public that contactless payment mechanisms may not be secure [9]. In Europe, and some parts of the US, contactless proximity cards have already been in use, and allow us to understand the concerns that may arise for NFC capable smart phones. Part of the solution seems to be one of education and reassurance that safeguards are in place, or at the very least that the banks and credit agencies will compensate users if they are subject to a malware attack that results in monies being stolen. Adherence to an "RFID Bill of Rights" [10] declared by parties that create NFC services may also be a helpful in reassuring its users.

8. Conclusion

The NFC standard is a novel capability being added to modern smart phones, complementing an assortment of radios, and a rich set of input modalities and sensors. To date, in the US market, only two commercial NFC phones, the Samsung Nexus S, and Galaxy Nexus, have been introduced. Given the NFC Forum was founded in 2004, and in Japan DoCoMo has used phones compatible with the Felica standard for over 10 years (a contactless card embedded in a cell phone), evaluation in the US is long overdue. This paper has provided an overview of the progress made defining the NFC standard, and the opportunity to widely enable the vision of tangible computing, a subject of speculation and research since the early 1990's. The next couple of years will be the proving ground for the technology, and users will have to evaluate the positive value it brings them, relative to the accompanying risks. However, most of the significant barriers preventing NFC applications have now been removed, and with the key ingredients in place, the future looks promising.

References

[1] The NFC Forum, Link: http://www.nfc-forum.org, 2012.

[2] K. Finkelzeller, "*The RFID Handbook*," 2nd ed., John Wiley & Sons, 2003.

[3] R. Want et al., "*Bridging Real and Virtual Worlds with Electronic Tags*," Proc. ACM Conf. Computer-Human Interaction (SIGCHI), ACM Press, 1999, pp. 370–377.

[4] H. Ishii. "*Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms*" Proc. of ACM SIGCHI '97, March 22-27, 1997,

[5] T. Kindberg et al., "*People, Places, Things: Web Presence for the Real World,*" Proc. 3rd. IEEE Workshop Mobile Computing Systems and Applications (WMCSA), IEEE CS Press, 2000, pp. 19–28.

[6] Wellner, P. (1993). "*Interacting with paper on the DigitalDesk*", Communications of the ACM, 36(7), 86-96.

[7] C. Taylor, "*Smart Phone Sales Overtake PCs for the First Time*." 3rd February, 2012. Mashable Tech (Source: Canalys 2012), Link: http://preview.tinyurl.com/6oohcqv

[8] H. Barwick, "*Beware of Malicious QR Codes*", PCWorld Security, 28th Jan 2012, Link: http://preview.tinyurl.com/7a8tzv5

[9] "*Security Fears Delay Switch to Contactless and Mobile Phone Methods of Payment*", Link: http://preview.tinyurl.com/5u6hanb, *Daily Mail*, 18 May 2011.

[10] S. Garfinkel, "*An RFID Bill of Rights*," Technology Rev., Oct. 2002, p. 35.