Using ICT for Reliable Identification and Authentication of Medicine

By
Saif ur Rehman
2007-NUST-MS-PhD IT-02

Supervisor
Dr. Raihan ur Rasool
Department of Computing

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science in Information Technology (MS IT)

In
School of Electrical Engineering and Computer Science,
National University of Sciences and Technology (NUST),
Islamabad, Pakistan.
(March 2011)

Abstract

The production and marketing of counterfeit medicine is a health related issue that has been growing despite efforts by government, pharmaceuticals and international organizations. Due to weak legislature and law enforcement issues, the problem continues to threat the health and lives of millions of people world over. Recently, there have been efforts to create medicine tracking and authentication mechanisms to restrain the counterfeit medicine sales and growth. In this thesis we review and analyze these existing solutions and identify their weaknesses and drawbacks. Different aspects these of authentication mechanisms are discussed and a criteria is determined for a practical solution. None of the existing solutions have all the desired features. The solutions that are simple to use are prone to attacks and counterfeiting whereas solutions that are secure are too complicated to implement in real world and may not provide end user level verification. The desirable solution would be scalable so that its implementation can be extended over a large geographic region under the same servers. It should also be easy to use and useable by medicine users. The security of the system is the most important factor and it should be resistant to counterfeiting and attacks. The architecture should be such that different functional components can perform independent of each other.

In this thesis we propose a solution that meets these criteria of scalability, usability and reliability. The proposed solution is based on existing infrastructure of GSM networks and does not require any specialized equipment. The proposed solution aims at enabling the medicine consumer to verify the medicine using a simple camera phone. Machine readable 2-D Data matrix is used for conveying the verification code. This feature saves the user’s time and effort of typing in a long code and allows for the code to be long and complex enough to deter counterfeiters from mounting any successful attacks. A mathematical

proof is given to show the security of the verification codes, and an intensive security analysis is carried out for the proposed solution. Different possible attacks such as Brute Force, DoS, DDoS, Man in the middle and Spoofing are considered and their countering techniques are discussed. The results show that the proposed solution is extremely scalable and can potentially support billions of registered products. The solution also meets the most stringent security requirements and is as secure against different attacks as international cryptographic standards. A skeletal prototype of the proposed framework has been developed to demonstrate the usability of the system from a common user's perspective. Practical testing has been carried out to determine the proper dimensions of the data matrix to be used. It is determined that QR codes of medium size, ranging from 5 to 8 cm square have good readability for a lengthy character string comprising of more than 60 characters from the basic SMS character set.

In the end, suggestions for implementation, future directions and further possible uses of the proposed solution are given. The proposed solution has numerous applications in the domain of authentication and product verification. The character set using the full length of a single SMS is complex enough to provide billions of unique identifiers for each individual in the entire population of planet earth. The limits on scalability in this regard are virtually non-existent if the processing is distributed over cloud servers.

Certificate of Originality

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any degree or diploma at NUST SEECS or at any other educational institute, except where due acknowledgement has been made in the thesis. Any contribution made to the research by others, with whom I have worked at NUST SEECS or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except for the assistance from others in the project's design and conception or in style, presentation and linguistics which has been acknowledged.

Author Name: Saif ur Rehman

Signature:___________________.

Acknowledgements

First and most of all, I am extremely thankful to Almighty Allah for blessing me with faculties, opportunities and the support that enabled me to complete my work, for protecting me from mistakes and pitfalls and for correcting my ideas and direction in research as well as other areas of my life. I hope and pray for his everlasting guidance to guide me throughout my life and his blessings upon me to grow evermore.

I would like to thank my family, especially my parents, for their prayers, understanding and support, and my siblings for being confident in my abilities when I myself doubted them. I am also very thankful to my uncle Dr Abdul Ghafoor and his family for bearing my presence and always being the gracious hosts while I stayed with them in Islamabad.

I am thankful to my thesis committee members for their valuable suggestions and guidance, their sense of practicality and keen observations improved many aspects of this thesis research. Here, I would also like to thank my friends Ejaz, Sardar and Saeed who are not only great company but also good council and a motivation as well.

Finally, I would like to thank my thesis supervisor Dr Raihan ur Rasool, without whose support and guidance I would never have completed this work. Dr Raihan has been a great source of inspiration during this thesis work, his concern for his student, his dedication, focus and the amazing ability to handle numerous responsibilities have never ceased to amaze me. I am very thankful to him for his guidance, patience and understanding, not only as a teacher and supervisor but also as a person with an open mind and a forgiving nature. I thank him for his time and efforts that led to this thesis culminating in its current form, and apologize for any weakness and drawback that might remain owing to my faults.

Saif ur Rehman

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Motivation

1. Counterfeit Medicine

The word “counterfeit” is defined in several ways by different dictionaries. The definition given below represents the sense in which the word is used for counterfeit medicine.

> ***“Counterfeit:*** *Fabricated in imitation of something else, with a view to defraud by passing the false copy for genuine or original; as, counterfeit antiques; counterfeit coin.”*

Different countries and dictionaries of law have several diverse and different definitions for the term ***“Counterfeit Medicine”*** or ***“Counterfeit Drugs”.*** In order to create a uniform criteria for whether to label a drug as counterfeit or not, the definition provided by World Health Organization (WHO) is used.

WHO defines counterfeit medicine as:

> *“A counterfeit medicine is one which is deliberately and fraudulently mislabeled with respect to Identity and/or source. Counterfeiting can apply to both branded and generic products*

and counterfeit products may include products with the correct ingredients or with the wrong ingredients, without active ingredients, with insufficient active ingredients or with fake packaging." [1]

Counterfeit medicine manufacturing and trade is illegal world over but it is still a growing business of international scale. Counterfeit drug manufacturing and sales are carried out at different levels, ranging from cottage industries to organized cartels that can sell this counterfeit medicine over the internet and buy alliances of authorities in poor countries.

Counterfeiters can and do market replicas of any and almost all types of medicine, ranging from food and health supplements and weight loss products to vaccines and life saving drugs.

1.1 Counterfeit medicine: a growing global problem

Counterfeit medicine causes numerous deaths and diseases world wide. While accurate statistics are difficult to obtain, incidents of counterfeit medicine leading to deaths and illnesses are regularly reported in media and over the internet. The ability to purchase medicine over the internet has widened the scope of such incidents and now reports of counterfeit medicine related incidents are regularly emerging from developed countries as well as developing and poor countries. For example in USA between Jan-June 2008, as many as 81 people were reported dead and hundreds suffering from allergic reaction due to counterfeit "Heparin" [16], a multi purpose anticoagulant used for dissolving blood clots inside veins, arteries, lungs and is also used in surgeries and dialysis. At least another 68 deaths from the same drug have been reported from other countries.

Over the years, counterfeit medicine's hold and share in the global market has grown continuously and significantly. A recent survey conducted by Pfizer[2] shows that up to 10 billion Euros were spent on counterfeit medicine within a year (statistics for Europe).

So far the poor and developing countries have been the main target of counterfeit medicine, but now the growing trend of ordering cheaper prescription medicine over the internet is steadily increasing the ratio of counterfeit medicine in developed countries as well. While some of these counterfeit medicines contain the active ingredients, many may contain inadequate or even potentially harmful ingredients.

1.1.1 Factors contributing to growth of this problem

There are several factors that contribute to the growth and penetration of counterfeit medicine.

In poor and developing countries, a lack of regulatory bodies and control systems is one of the major primary causes. Where a regulatory procedure and regulating body is present, the enforcing of these procedures is either insufficient or neglected altogether due to several reasons.

Developed countries were the least effected by this problem until the recent past, but now they are also showing an alarming rate of increased incidents of counterfeit drug sales. This trend is partly a result of purchasing medicine over the internet[3] by apparently legitimate online stores offering discounts on popular drugs and supplements. Another issue that borders on the “grey market” is laws regarding pharmaceutical manufacturing license of drugs that are not under patent anymore, some countries grant exclusive rights to a particular manufacturer/importer whereas others can grant license to multiple manufacturers / importers. In countries where a pharmaceutical manufacturer/importer has monopoly over sales of a particular brand, the prices are much higher than those where there are multiple manufacturers in the market. This can lead to smuggling of genuinely produced medicine to such countries, and counterfeiters can join in the profit raking by injecting their own products.

The question as to **why it is easy to counterfeit drugs** is answered by the fact that unlike most products such as electronics or mechanical equipment, or consumables such as edibles, the quality, working and effectiveness of a drug can not be observed with accuracy by a common user or observer or even a professional such as a doctor or pharmacist. This is due to the fact that appearance of a medicine does not indicate its composition or content and several other factors can effect and contribute to the health of a patient therefore the effectiveness of even an authentic drug may not be visible immediately. This property of medicine makes it hard to determine the fidelity of a medicine from its appearance and usage.

From the point of view of a counterfeiter, the motivational factor is obviously financial gains. Medicine has high weight/size to price ratio. The price of genuine medicine is high due to manufacturing cost of active ingredients, compensation for expenses incurred during the development of the drug, taxes and profit for the manufacturer. As counterfeiters are saved many of these expenses, counterfeiting medicine gives them a huge profit margin and smaller risk factor compared to other high profit illegal ventures. Manufacturing fake medicine costs very little and the fake products are mostly sold at a so called "discounted" price that gives the counterfeiter good profit and the original manufacturer's product a tough competition in the market. Marketing these products over the internet has greatly increased the target customers and decreased the risks involved for counterfeit manufacturers.

On the public level, there is a serious need of awareness and reliable methods for identification of original medicine, while chemical analysis and different spectroscopy techniques are used for identification of several medicines, such tests usually are for statistical reasons, conducted by organization and government bodies. Such tests can not address the severity of the issue or screen medicine entering the market through illegal channels.

The following reasons are listed by WHO[10] as contribution factors towards encouraging counterfeiting of medicine.

- High value in relation to size
- Lack of appropriate medicine legislation
- Absent or weak National Medicine Regulatory Authorities NMRA in some countries
- Lack of political will and interest on part of Governments
- Weak enforcement of NMRA
- Corruption and conflict of interests
- Shortage of medicine
- Inappropriate use of medicine
- High price of medicine
- Price differentials
- Lack of control over export of medicine
- Trade through free ports, free-trade zones and many intermediaries

1.2 Scope of the problem

The problem of counterfeit medicine is a global issue; more or less all countries are affected by this to some extent. Predominantly, the poor and developing countries of Africa and Asia are major targets due to lack of a stable regulatory structure in these countries. It is also shown by various reports that counterfeit drugs are manufactured on a large scale in developing countries specially India[4, 5] followed by Egypt and China[5]. These drugs are then supplied to various countries world over.

Due to different definitions of counterfeit medicine and uncertainty about the percentage of counterfeit detection, it is difficult to estimate the market size of counterfeit medicine. Different sources quote it to be up to 70 - 75 billion USD[6, 7] globally. Pfizer, a leading pharmaceutical estimated in a recent study across fourteen European countries[9] that counterfeit medicine worth more than 10 billion Euros are purchased every year in these countries.

Different kind of medicine are counterfeited and marketed, some of them include life saving drugs. The image below[8] shows some statistics from WHO for the years 1999 to 2002.

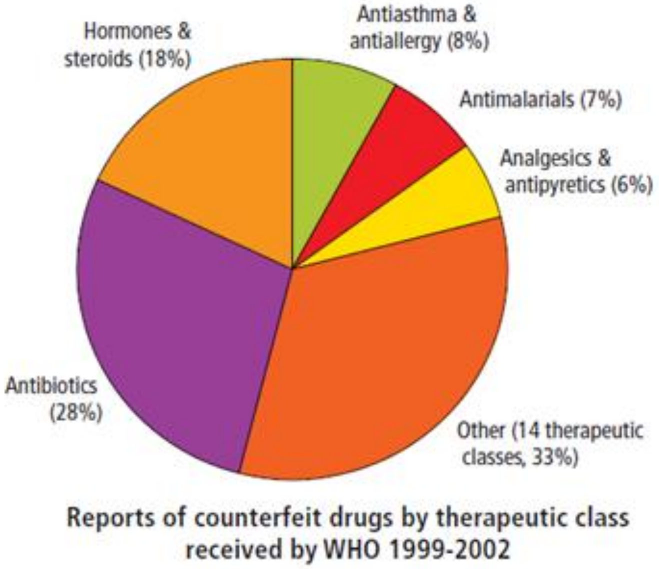


Figure 1.1 **Reference: Ref[8]**

The table given below shows some of the recent incidents mentioned in a report by WHO[17]

Counterfeit medicine	Country/Year	Report
Anti-diabetic traditional medicine (used to lower	China, 2009	Contained six times the normal dose of glibenclamide

Counterfeit medicine	Country/Year	Report
blood sugar)		(two people died, nine people hospitalized)[1]
Metakelfin (antimalarial)	United Republic of Tanzania, 2009	Discovered in 40 pharmacies: lacked sufficient active ingredient[2]
Viagra & Cialis (for erectile dysfunction)	Thailand, 2008	Smuggled into Thailand from an unknown source in an unknown country[3]
Xenical (for fighting obesity)	United States of America, 2007	Contained no active ingredient and sold via Internet sites operated outside the USA[4]
Zyprexa (for treating bipolar disorder and schizophrenia)	United Kingdom, 2007	Detected in the legal supply chain: lacked sufficient active ingredient[5]
Lipitor (for lowering cholesterol)	United Kingdom, 2006	Detected in the legal supply chain: lacked sufficient active ingredient[6]

Table 1.1. recent incidents mentioned in a report by WHO

1.3 Control and mitigation efforts

World Health Organization urges all member countries to create well defined medicine regulatory laws and create National Medicine Regulatory Authority/Agency that will implement these regulatory laws. Unfortunately, this is easier said than done, especially in countries with weak governments who are struggling with issues such as wars, rebellions, law and order situations and keeping possession of power. Any country with weak law enforcement agencies are a target for counterfeiters.

United Kingdom’s Medicine and Healthcare products regulatory Agency developed an anti-counterfeiting strategy[11] to combat the increasing risk of counterfeit medicine in the country.

European commission is working on creating legislation[12] that will enforce medicine supply chain monitoring at all levels. A pilot project for verification has also been carried out at Stockholm, which will be discussed in detail in the existing solutions review section.

The state of California is creating a medicine pedigree law which will force pharmaceuticals to provide all information regarding a drug and will monitor all transactions, also a certificate which will certify the accuracy of the product will be obtained from the responsible party i-e the product owner or manufacturer.

Different approaches including new regulatory laws, strategies and new technological aids have been developed and further are being developed to create means which would make marketing of counterfeiting difficult if not impossible. These techniques are discussed in details in the literature review chapter.

1.3.1 Limiting Factors

Limiting factors in anti-counterfeiting efforts are many and may vary depending upon the location. In poor and developing countries, the major factors are:

- lack on interest on part of Government
- absence of proper regulatory laws and implementing bodies
- corruption
- lack of awareness in public

In developed countries like Europe, UK and USA, the major factors are

- Online purchases of cheaper medicine, these medicines are usually discounted. The sellers can justify the lower prices to unaware users with such terms as “Generic medicine”, cheap labor and tax laws in countries of manufacture. WHO reports[17] that over 50% of online retailers who do not provide a physical address are involved in selling counterfeit medicine.

1.4 Motivation

Counterfeit medicine sale is a global problem and affects the lives of hundreds of thousands of people. However, very little attention has been paid to conducting research for finding a practical solution to this problem by Information Technology researchers. Most of the research conducted for identification of counterfeit medicine is done in the domain of chemistry and pharmaceuticals.

The rapid growth, advancement and penetration of information technology especially that of communication technology such as GSM networks and mobile internet presents opportunities that were unimaginable before. In today’s communication age, mobile phones and internet are almost in every one’s reach.

In this thesis research, it is aspired to use the aforementioned well known technological tools for developing a solution that can be used by every one for convenient and reliable detection of counterfeit medicine and authentication of genuine medicine.

1.5 Contribution

In this thesis the author proposes a medicine verification mechanism that makes consumer level verification possible and convenient. The proposed mechanism can also be used for verifying the source of other products as well.

Machine readable verification codes are used for authentication of medicine. The consumer can use a simple java enabled camera phone to acquire the verification code. The verification code is designed to be resistant against potential attacks such as brute force attack. The architecture allows for scalability so that an implementation can be extended over a large geographical location. Unlike other similar proposed solutions, only the manufacturers', automated verification center's and consumer's participation is required for successful verification of any medicine.

1.5.1 Problem Statement

The problem statement of this research thesis is

> ***"There is no medicine verification system working on SMS based verification of machine readable data-matrix such that it is easy to use, scalable, secure and easy to implement"***

1.6 Thesis Organization

In chapter 2 of this thesis various proposed and implemented solutions are reviewed along with related research. The potential weaknesses and shortcomings of these solutions are identified and finally the criteria for desired solution are created.

In chapter 3 the proposed solution is presented along with the security overview, which is the main focus of this thesis research work. A detailed description of the architecture and its different module is also presented in this chapter.

In chapter 4 the security of proposed solution is evaluated mathematically and it is shown that verification identifier generation by counterfeiters is financially and temporally infeasible. A comparison between different solutions in terms of features is also presented to highlight the advantages of the proposed solution.

the results of experiments are shown. The experiments are conducted using existing software tools and the purpose of the experiments is to gauge the usability of the solution for a common user.

In chapter 6 the conclusion of the research and future research possibilities are indicated.

Chapter 2

Literature Review

The problem of counterfeit medicine is almost as hold as medicine itself because of its profitable nature. Several different techniques and strategies have been tried by manufacturers, copyright holders and governments in the past and they keep coming up with new ones. Globalization and fast and reliable transportation and communication have enabled manufacturers to distribute their products world wide and earn more profits along with benefiting the health of people. Counterfeit manufacturers are also using the same resources to distribute their fake products in the same markets, making it difficult for the users to distinguish between fake and original products.

Counterfeit sales are a reason of concern for both government and pharmaceuticals. Government suffer by facing loss to its public's health, damage to its own reputation and loss of revenue generated by taxes paid by licensed manufacturers. Manufacturers suffer loss by losing market share to counterfeit products and potentially the loss of faith and trust on part of consumers, who might come to regard the medicine as ineffective.

To discourage counterfeit sales and manufacture, the governments pass laws and develop regulatory mechanisms such as supply chain monitoring. The pharmaceuticals at their end develop special markings and seals to safeguard against counterfeiters. While government policies and regulatory mechanisms are beyond the scope of this thesis, technological tools used by governments and the techniques used by pharmaceuticals are relevant and are discussed in detail.

2. Existing solutions

In this section, different solutions created for authentication, identification or verification of medicine are described. These solutions include those that are based on information technology and others based on hard to copy identification techniques such as seals and marks etc.

ICT based solutions focus on verification from a central repository or database which contains the verification data pertaining to a particular package of medicine.

Non ICT solutions focus on development of special identification markers that should be impossible or at least very difficult for the counterfeiters to copy accurately. There are various such options but most popular ones are holograms, UV inks, color shifting inks and seals on bottles and packaging.

2.1 General solutions

Holograms

Commonly used holograms are actually multilayer image that changes its color/image depending upon the viewing angle. Holograms that are commonly used are two or three layered. The layers are placed on top of one another and the image seen changes with the change in viewing angle. However these holograms, commonly used for product identification are not holograms in the true sense of definition as they do not create a 3-D image by dispersing light.

There are advanced hologram generation techniques such as Electron Beam Lithography. These holograms are very difficult to copy and are usually used identification. These holograms can have a resolution of up to 12000 dpi.

Special Inks

Ultraviolet /Invisible Ink

There are various types of invisible inks, know from many decades, even centuries. Usually they are some organic material that changes color upon heating.

Ultraviolet inks are invisible in normal light but appear when viewed under an ultraviolet lamp.

Color changing inks

These inks appear to be of different colors depending upon the angle of viewing. While the color remains the same, from different angles, different shades are visible due to the dispersion of light at different angles.

Magnetic Inks

Magnetic inks are easier for machines to read and are usually used in baking sectors to reduce reading errors.

Thermochromatic Inks

These inks appear/disappear, change color when stimulated with heat such as rubbing with finger tips.

2.2 ICT based solutions

2.2.1 Authentication codes

Authentication codes refer to a string of characters, mostly numeric and sometimes alphanumeric, usually 6 to 16 characters long. Authentication codes are widely used in various scenarios, such as calling cards, adding mobile phone debit, gift cards, software registration, and ATM pin codes etc.

mPedigree is an organization in Ghana which is using authentication code printed on medicine package for verification of medicine. The code is send via an SMS to a verification service which verifies the code from its database of valid authentication codes. There are some weaknesses in this solution from the security point of view; they are described in a later section.

2.2.2 Machine readable data

Machine readability gained mass acceptance and implementation with the invention of barcodes, and since mid 1970's barcodes have been used commercially. Since advent in Information technology, the need to put more data in a smaller space arose. To address this need, 2-D barcodes, also know as data matrix has been developed.

There are various types of 2-d Barcodes including Aztec code, Maxi code, QR code and SPARQ code. These are designed for different purpose and have different size VS data capacities. Machine readability reduces human effort and error thereby increasing accuracy and speed of processing.

Data matrix has been used in a pilot project carried out by EFPIA[13, 14] in Stockholm for verification of medicine at supply chain and point of dispense. User level verification is not considered in this solution.

2.2.3 Encrypted data for authentication

Encryption has long been used for protecting data from unauthorized access and tampering. Asymmetric encryption can enable users to decrypt the encrypted data using the public key and verify encrypted information as genuinely generated from assumed source.

National Agency of Food and Drug Administration Control of Nigeria (NAFDAC) is using verification scratch cards within medicine packing printed with encrypted symbols coved by stretchable material, which can be scratched off and the symbols can be sent for verification via an SMS[15].

2.2.4 RFID identification

RFID identification is used for identification and verification of various products and personnel in industry. However, the comparatively higher cost of between 0.30 to 0.15 USD for a single RFID chip makes it infeasible for verification of disposable products such as medicine. The RFID readers are also expensive, priced in the range of 500 USD and upwards, which makes them infeasible for use in poor and developing countries.

2.3 Critical review of existing solutions

A critical review of the aforementioned technique follows. Keeping in view the fact that counterfeit medicine industry has billions of dollars a year market; the security weaknesses of the existing solutions are highlighted.

The purpose of all verification and authentication techniques is to thwart copying and faking attempts by counterfeiters, if that purpose fails or is prone or susceptible to failure, the solution becomes unreliable.

2.3.1 Problems in General solutions

General solutions such as different types of inks, seals and holograms have various drawbacks which are described below.

- Copying is possible: While there are very few manufacturers with equipment advanced enough to develop real holograms, many can create inferior copies that can confuse or trick an ordinary user. In case of UV and color shifting inks, the technology is widely available and poses not serious problem to counterfeiters.
- Difficult to differentiate between real and fake: For an inexperienced user, it is difficult to tell apart the real seal or hologram from the counterfeit ones.
- Holograms and seals are small in size due to price considerations, a hologram or seal which is large enough to facilitate identification can impact the price of a product.

2.3.2 Problems in existing ICT based solutions

Existing solutions are developed and implemented at local scales so far, and have serious potential issues regarding scaling of the solution.

For example when multiple manufacturers from various parts of the world would supply a medicine to some area, how will the verification and authentication codes be managed?

Apart from the above mentioned potential issues, existing solutions using random identifiers have the following security problems

- Solutions using identifiers for verification either encrypted or random are vulnerable to reuse of a genuine authentication code.

- Solutions that monitor verification code reuse are susceptible to brute force attack which can result in verification of genuine codes, thereby resulting in declaration of a genuine medicine as fake when a user/seller tries to verify the same code.

2.4 Possible fixes for existing solutions

The above mentioned security problems can be fixed using the following techniques

- Limiting the number of verifications from one source/user per unit time.
- Verifying a code only once, thereafter putting it in a “pre verified” list of codes
- Charging a small amount when verifying a code
- Increasing the length of verification code, each additional symbol exponentially increase the difficulty for a brute force attack

2.5 Potential problems despite fixes

While the above mentioned fixes seem to solve the issues in existing solutions, they are only improving the existing strength but not fool proofing it. The potential problems that can remain or arise after these fixers are given below

- Users can type only a short number of random symbols without making mistakes, longer strings will increase the chances of errors by users.
- Charging for verification may discourage sellers and some users from verifying a medicine.
- Even a longer verification code will be prone to attack by bots.

2.6 Existing Research work

Very little research is known to have been carried out in order to create a medicine verification mechanism using information technology tools. In this section we review a paper[18]by M Paik published in 2009, apparently the only published research attempting to create a solution to the counterfeiting problems based on mobile communication technology.

2.6.1 Overview

In the system proposed in their paper, the authors propose

- The use of 2-d barcode for assigning unique identifiers to different levels of granularity of a shipment i-e starting from cartons, dozens, individual units and blisters.
- All non-consumer participants, here onwards called **User** (i-e suppliers, distributors, retailers) are also assigned a photo ID card each with a unique 2-dbarcode on each from the **Regulatory Body RB**.
- The user registers with the RB and chooses a password to communicate with it.
- The system tracks the movement of all shipments from start to end i-e starting from supplier to individual retailer.

2.6.2 Working of the system at supplier-distributor-retailer level

The system is based on a series of transactions between user and RB. Each user has a camera phone with the custom application (Epothecary) installed. The following steps as shown in figure 2.1 take place every time the medicine moves from one user to another.

- The user enters his password for authentication which is compared to a hash of it stored in the application.

- In case of correct password, the purchaser scans his ID barcode, the ID barcode of the seller, the ID barcode of the item to be purchased.
- An SMS is sent to RB with an initialization vector IV
- RB checks the sender number in the list of valid users, in case of valid user, it replies with the same IV.
- The user then sends the scanned information in encrypted form. The password and IV is used in the encryption process.
- The RB checks seller's ID against the item for sale's ID.
- Some optional features such as location estimation using GSM or additional verification might be done.
- If the item for sale's ID is registered against Seller's ID, the RB registers sends to the buyer the tag keys of subsidiary units (smaller parts of a shipment)
- Upon receiving the subsidiary units, the buyer scans then subsidiary units and their keys are compared to those received from RB
- If the keys match, an ACK is sent to the RB. If the keys do not match, a NAK is sent instead.
- Upon receiving ACK the RB updates the registry and tags the buyer as the owner of purchased item and its subsidiary units. RB records the transaction and sends and ACK to the buyer's number confirming successful completion of the transaction.
- In case of NAK the product ID is tagged for investigation and further transaction from the product ID or any of its subsidiaries generates alerts.

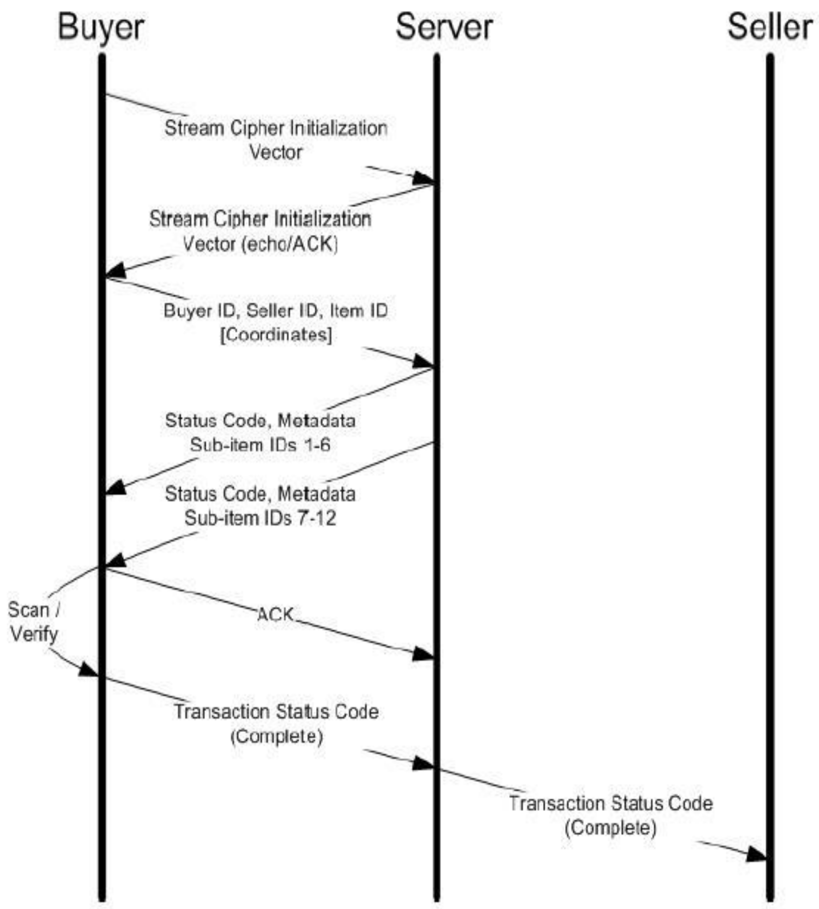

Buyer
Server
Seller
Stream Cipher Initialization Vector
Stream Cipher Initialization Vector (echo/ACK)
Buyer ID, Seller ID, Item ID [Coordinates]
Status Code, Metadata Sub-item IDs 1-6
Status Code, Metadata Sub-item IDs 7-12
Scan / Verify
ACK
Transaction Status Code (Complete)
Transaction Status Code (Complete)

Figure 2.1 : Protocol message flow for a successful transaction between users. Source: Reference[18]

2.6.3 Working of the system at Retailer-consumer level

The steps of sale to a consumer are slightly different than the interactions between users. These steps are as follows

- The user chooses an option that indicates a sale rather than a purchase.

- Scans his own ID tag and those of the item/s to be sold.
- The RB checks the ownership of the items to be sold and once verified, marks the units as sold.
- RB sends a summary to the seller along with a random 8 digit reference number.
- The purchaser can send this number to a public number of RB to receive the transaction metadata.

2.6.4 Assumptions

The authors make the following assumptions

- The RB is a trusted party
- The GSM is not secure
- ID tags are resistant to counterfeiting
- Attacks on system possible from within (registered parties) and outside (non-registered parties)

2.7 Analysis of the Epothecary Drug Authentication system

In this section we will highlight the pros and cons of the Epothecary Drug Authentication system.

2.7.1 Pros of Epothecary Drug Authentication system

- The system takes in to account different possible attacks such as Brute force attack, Denial of Service (DOS) attack, Spoofing attacks, identity theft and cloning.
- The proposed solution counters these attacks based on its architecture, technology being used and multiple factors participating in authentication and completion of a transaction.

- The system maintains a registry of authorized vendors and distributors, thereby holding them liable for the transactions they make.
- The end user has the facility of verifying his purchase for a limited time. This can help in circumventing any trickery on part of the retailer.
- The system is of relatively lower cost compared to other options such as RFID
- The system does not require any specialized training or equipment
- Machine reading the tags saves time that would be taken by manual entry.

2.7.2 Cons of Epothecary Drug Authentication system

- Multiple number of SMS transactions between a user and RB are required per purchase. In case of GSM network congestion, the SMS can possibly get delayed for several hours.
- All SMS should successfully reach for the transaction to complete.
- Lets assume probability of error in SMS or failure to reach destination is given by '*p'*
- As shown in figure 2.1, for a successful transaction, at least 7 messages are sent. This increases the probability of failure to complete transactions to '7p'. This means that probability of a transaction failure is 7 times the probability of an SMS error or failure in the GSM network
- Some transactions might require dozens of tag scans, incase of even a single scanning error, the whole process might have to be repeated.
- Suppliers at higher level sell and purchase thousands of consignments on daily bases.

- Given the complexity of the process of handing over even a single carton (involving dozens of scans)and keeping in view the fact that to scan a clearly captured tag, a mobile phone takes at least 2 to 3 seconds, the minimum time taken by a single transaction is calculated as follows. The numbers in brackets indicate seconds.

  Starting application, entering password (3-5), initial tag scans 3 * (3-5), sms send-receive time2* (3-4), scanning of subsidiary tags 12 * (3-5), sms send- time 2(3-5)

  These numbers sums up to a total of 60-98 seconds. This is the absolutely minimum time assuming no human error and network latency. Realistically, it normally takes a person from 5 to 10 seconds to scan a 2-d barcode tag.

- Assuming the impossibly optimum 60 seconds required to complete a transaction, if a supplier supplies/sales 1000 cartons every day, it would take him 1000 minutes or 16.667 hours of continuous non-stop work. In the duration of which he will have to scan instances of 2000 user ID card tags and his phone will have to process 7000 SMS messages.
- As we can clearly see from the above calculation, the system proposed in the discussed paper is theoretically sound but practically very unrealistic and un-implementable beyond a very small scale.

Number of Shipment	Number of SMS	Number of minimum Tag Scans	Time taken in hours
100	700	300	1.666666667
1000	7000	3000	16.66666667
1440	10080	4320	24

Table 2.1. The number of shipments handled by a single user

2.8 Conclusion of review

In the light of above discussion, we reach the conclusion that a simple yet reliable medicine verification system needs to be developed. The system should have all the security features described in the discussed paper. At the same time, it should avoid the complexity that renders the solution proposed in the discussed paper impractical to use and implement.

Chapter 3

Proposed Solution

The objective of this thesis research is to develop a solution for medicine verification that is practical in nature. As discussed in previous chapters, several solutions have already been proposed but they were either too complex to implement and use or their security or authentication was not reliable. We strive to remove these draw backs with our proposed solution.

3.1 Framework overview

The proposed solution would use existing cellular infrastructure for reliable verification of medicine. The cellular network has seen a rapid growth in the past decade and the availability of vast GSM coverage and the decreasing prices of services and devices make this technology an ideal choice for this purpose. For some years now, every mobile phone from mid-range price and onwards is equipped with a built in camera and with time, the quality of the cameras keep increasing with respect to the unit's price. Most of the camera phones are using an operating system based on java, which allows subsequent addition of applications to the phone. This enables us to develop a custom application, designed for the sole purpose of medicine verification. Keeping these points in view, we propose to use the camera phones in the proposed solution.

The usage of camera phones will decrease the time required to enter the medicine verification identifier value. To deter brute force attacks on verification system, we propose to use a considerably long string of characters and a large character set. The probability of a human error increases with the increase in length of the string, especially with similar looking characters, whereas machine readable data is much safer in this regard due to checksums and other error correction and reduction techniques.

The medicine packaging would be labeled with a QR code containing the identifier. The QR code can be covered with scratch-able material or kept inside the packaging to protect from verification without purchase. QR code is a type of 2D data matrix which is gaining popularity over other data matrix due to its 360 degree readability feature, which saves the users from the necessity of aligning their phones accurately. The Data matrix technology is gaining wide range acceptance and newer phones such as various smart phones are preloaded with a barcode reader.

As it can be seen, the proposed solution is fairly simple for the end users. The security of this solution primarily lies in the length of the identifier and the architecture at the back end.

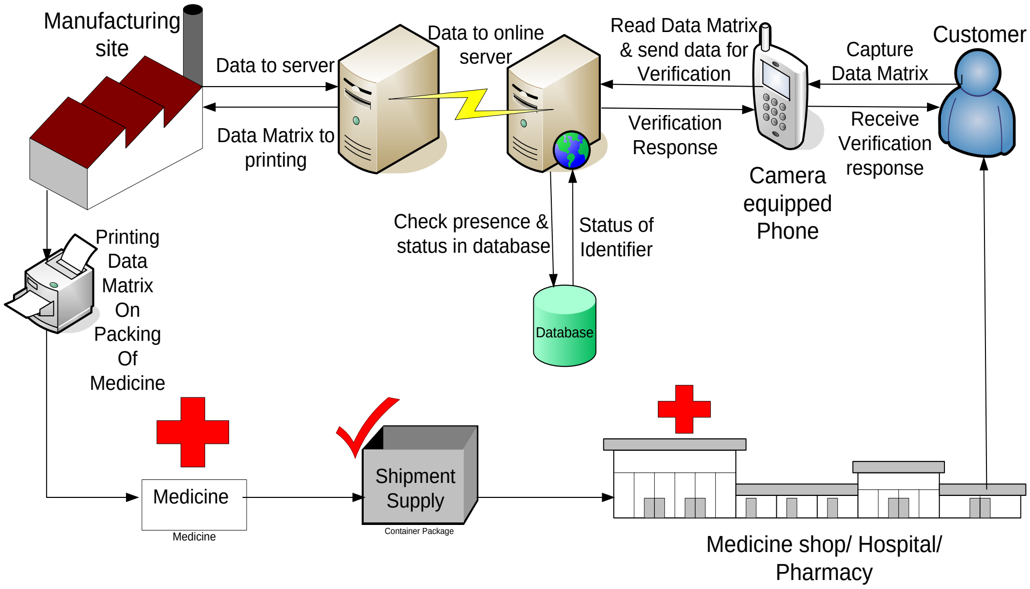


Figure 3.1 Components participating in the system

A camera equipped mobile phone will be used to snap shot the data matrix using a data matrix reader specially developed for this purpose. Many data matrix reading applications are available over the internet for different formats of data matrix. These readers read a data matrix and copy the data; usually a URL into the mobile's browser address bar to save the user from typing in long URL and avoiding a typographical error thereby ensuring that user's reaching the desired web site. The application developed for medicine verification will read the data and paste the unique identifier code into a text message i-e SMS, the SMS will then be send to the Central Verification Server (CVS). The CVS will verify the identifier code, log the verification event and respond to the user via SMS.

3.2 Architecture

The over all architecture of the proposed solution consists of different components or modules that work independent of each other. The proposed solution intends to automate the process of verification code generation and its association with medicine. The human involvement would be on administrative level and no human intervention would be required in the working of the system. This is especially important because human involvement in routine processes can create a bottleneck and affect the scalability of a solution.

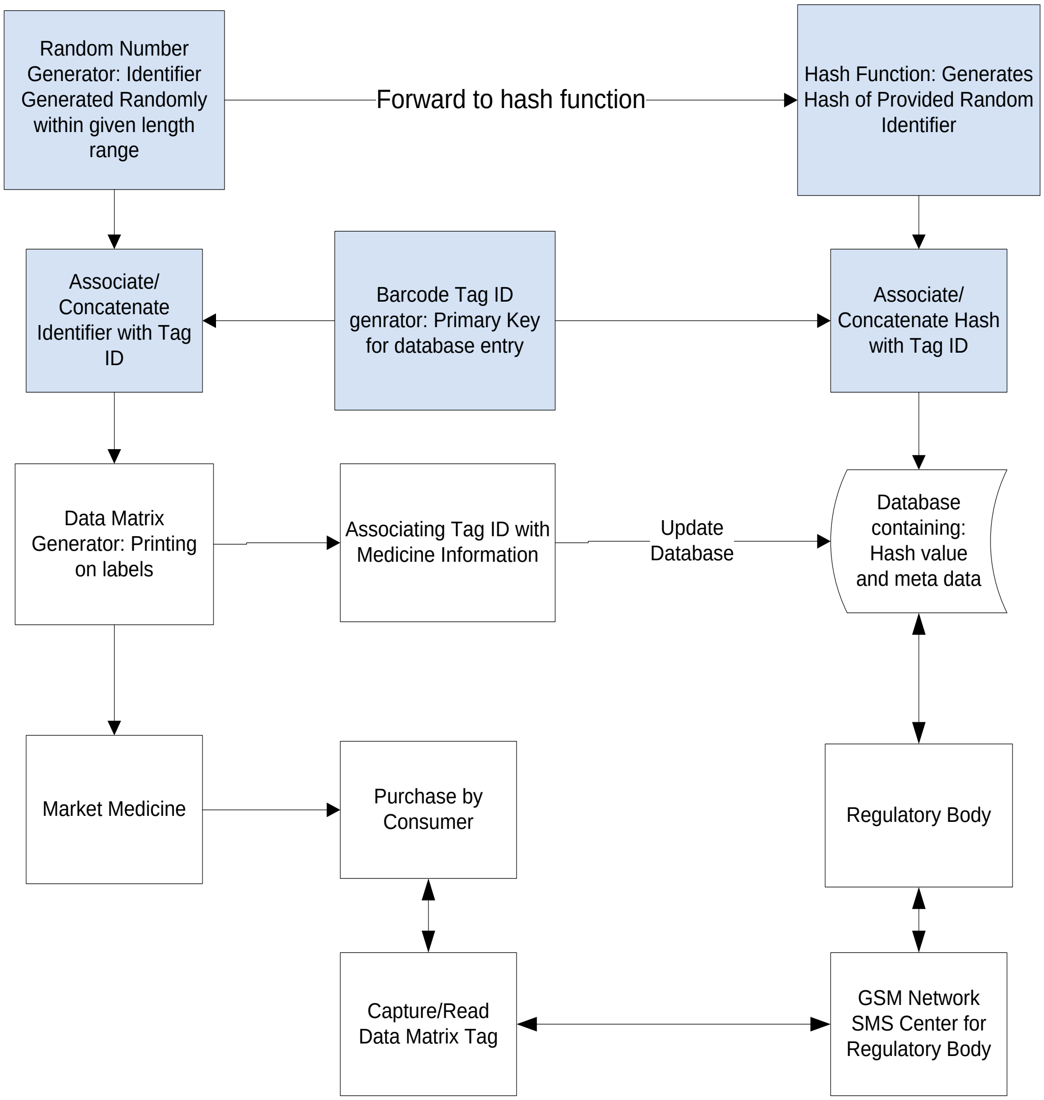

Random Number Generator: Identifier Generated Randomly within given length range
Forward to hash function
Hash Function: Generates Hash of Provided Random Identifier
Associate/ Concatenate Identifier with Tag ID
Barcode Tag ID genrator: Primary Key for database entry
Associate/ Concatenate Hash with Tag ID
Data Matrix Generator: Printing on labels
Associating Tag ID with Medicine Information
Update Database
Database containing: Hash value and meta data
Market Medicine
Purchase by Consumer
Regulatory Body
Capture/Read Data Matrix Tag
GSM Network SMS Center for Regulatory Body

Figure 3.2. This Diagram represents the overall architecture and the communication between different modules.

3.2.1 Random Verification code generation

The random verifier generator would generate total random verification codes in a given range of string length. These codes will be generated upon request from a registered pharmaceutical manufacturer, for a registered and licensed product. The identification numbers of the manufacturer and product will be concatenated with the random verification code in a predetermined format. Then the record will be associated with a sequential Identifier or ID which will be used as a primary key and the data will then be passed on to two different components, the server side database and the bar code generating queue.

3.2.2 Server Side database

The server side database is the main database which will be accessed during the process of verification. The database will contain all the verification codes and the event logs of verification instances. Rather than using a single table, multiple tables with a proper normalized structure will be used to ensure efficiency and security. The verification code generator will have a "Write only" access to this database. Once a record is submitted in to it, it can neither be modified nor removed. To further enhance the security of the data and protect against stolen data, the original verification codes will not be stored in the database. The verification code will be hashed using a one way hash function such as SHA and the hash value will be stored against the ID of its verification code. Upon receiving a verification request, the verification module will, among other tasks, hash the value of the verification code before querying the database.

3.2.3 Data Matrix Generation

The data-matrix can be generated either at the central server or it can be generated at the printing location. The benefit of generating Data Matrix at server is that it would save the software requirements at the printing location. However, this option will lead to higher communication bandwidth requirements and problems with scalability. Generating the Data Matrix at printing site would distribute the processing cost, reduce the bandwidth requirement and would only require a one time setup. The server can send the verification codes to the printing location using a secure channel or over an unsecure channel in the form encrypted data. Even if the security of the printing side is not good and some codes are stolen, they can not be used for generating other codes and if the theft is reported; the relevant codes can be pre-blocked to avoid sales of illegal medicine.

3.2.4 Verification Service

This service will receive the authentication request from the GSM network interface. The Verification Service module will have read only access to the stored hash values. However, before accessing the actual values, two other checks will be performed. The received data will be parsed so as to separate the manufacturer ID, The product ID and the verification code. First the manufacturer ID will be matched, if it is present in the registered manufacturer’s list, then the product ID will be checked in the products licensee to that manufacturer, when that is also verified, the verification code will be hashed and matched. This process has dual benefits; first of all, if any product by any manufacturer is to be banned, it can be done in the licensed products table. Also, this process will save processing time of searching a much larger record set in case of problem with the ID.

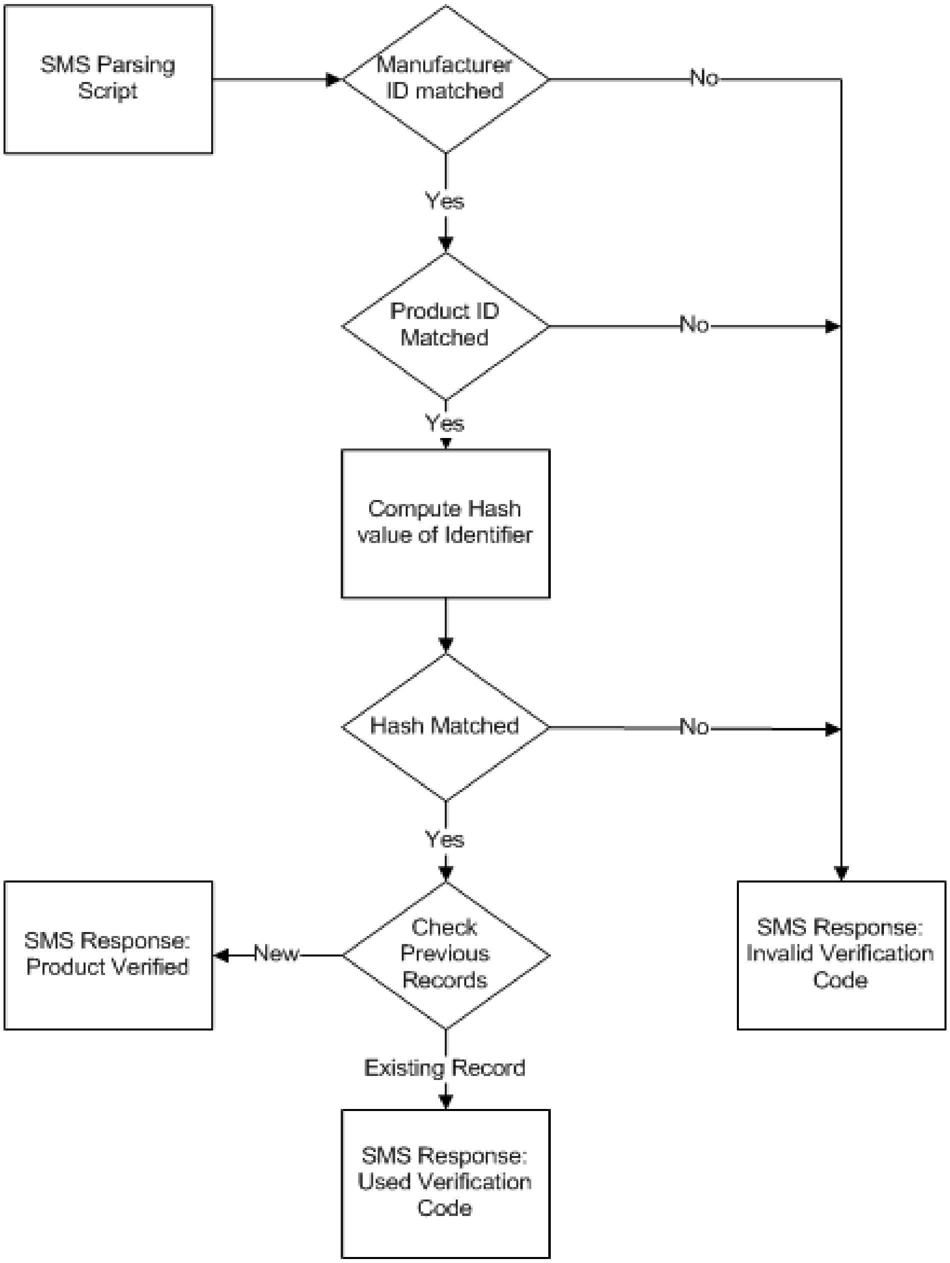

SMS Parsing Script
Manufacturer ID matched
No
Yes
Product ID Matched
No
Yes
Compute Hash value of Identifier
Hash Matched
No
Yes
SMS Response: Product Verified
New
Check Previous Records
SMS Response: Invalid Verification Code
Existing Record
SMS Response: Used Verification Code

Figure 3.3. This diagram represents the verification service process flow.

3.2.5 The GSM network Interface

The SMS sent by the user will be routed through the GSM network to the Central Verification Server. For this purpose an SMS Center (SMSC) will be setup at the GSM network Interface module. It is desirable to assign a special short number to the CVS SMSC which would be uniform across different GSM networks. While a GSM modem can be used to send/receive SMS messages, they are a big scalability issue. Also the dependence on a GSM cell causes geographically fixed location; this can make attacks easier on the server. Therefore, the SMS for the CVS will be forwarded using a secure network connection to the CVS. The SVS should reside on a cloud server, or at least an online server. The SMSC can be configuring to reply using the same GSM network. A script running at the SMSC will parse the incoming text and extract data of interest. This data can be logged if needed, and passed on to the verification module.

Chapter 4

Security of the Proposed Framework

Introduction

When a verification or authentication technique is being developed, the foremost importance is given to its security. If an authentication system can be bypassed, faked, or even attacked to reduce its efficiency or accuracy, that technique is not acceptable. Different techniques have been used for securing data from unauthorized access and temperance, such as passwords, cryptography, steganography and hash functions. Different techniques have been developed to address different needs. In general, a technique is considered secure if it is temporally and financially infeasible for the attackers to mount. A temporally secure technique can withstand attack to such time until the data is not secret any more, or loses its usefulness, such as corporate decisions, military strategies and other data of such nature. A financially infeasible attack means that the cost of attack is more than the expected gain. In our proposed framework, the security is based on temporal infeasibility of attacks. However, adding financial infeasibility is also very simple; it is not added because it is not required.

4.1 Temporal infeasibility

The proposed solution uses random string of characters as verification codes. Due to the random nature of the verification codes, they are naturally secure from different analysis attacks that are mounted on underlying algorithms. The only remaining attack is brute force attack, to try all the possible combinations in order to get the working ones. This attack is rendered useless by using a long string of characters. The mathematical verification proving the infeasibility of the attack follows.

4.1.1 Mathematical proof

The probability for a brute force attack is given by

$$\mathbf{p = n/m} \qquad \mathbf{\ldots . (1)}$$

Where ‘p’ is probability of correct guess, ‘n’ is total number of valid codes and ‘m’ is total possible codes. In this case, for constant value of ‘n’ the higher the number of ‘m’, the lower the value of ‘p’

As it can be seen from equation 1; for a fixed number of selected codes, a larger pool size will make it harder to guess any selected codes.

The very first mobile phone that was text capable use a character set of 127 characters. Each character 7 bits in length. A single text message consists of 160 characters in size. So the total number of bits comes to.

$$160 \times 7 = 1120$$

The total possible combinations in a string of binary numbers is given by 2^n

Putting 1120 as ‘n’ we get

$$2^{1120} \sim 1.42 \times 10^{337}$$

This gives the maximum possible number of identifiers that can be sent in a single SMS using the full character set. This is a huge number and there is no need to use the full character set or all the 1120 bits. This calculation is to demonstrate the potential complexity within a single SMS. A string of selected characters and of sufficient length can be used for generating the random identifier.

The security of Advance Encryption Standard (AES) at its highest is 2^{256} and is considered secure for all type of data. Breaking (AES) will take longer than the age of universe if state of the art computer processors are used. This level of complexity can be obtained in our solution by using 37 characters of SMS using the full set of 127 characters.

This number is large enough to counter brute force attacks. Due to random selection of identifiers, different analysis based attacks which are commonly used on encryption algorithms are not possible. A longer string is advisable to compensate for a large of identifiers might be present in the database. To calculate the length of string to be used, equation 1 can be used. The length of identifier string should be decided based on the expected number of records in the database.

4.2 Security against various attacks

Apart from the brute force attack that can be mounted to generate working codes, attacks can be mounted to steal working codes from the database, denial of service attacks, attempts of false verification etc. The system is secure against these attacks as explained below in detail.

4.2.1 Brute Force Attack

The brute force attack is temporally infeasible as explained in the mathematical proof section. However, in case of string length restrictions, for instance if plain characters have to be used instead of Data Matrix, the Brute force attack can be made financially infeasible by charging a small amount of money for each verification attempt. Similarly, a log can be maintained in which recent numbers requesting verification can be paused for a few seconds. This will make brute force attack temporally infeasible but processing overhead will be incurred.

4.2.2 Denial of Service Attack (DoS, DDoS)

Denial of service attack is a popular attack usually launched on websites and web based services. It overburdens and chokes the processor and memory by sending multiple requests from fabricated fake addresses. A large number of active threads/ connections slow down the system and actual users can not get services. CVS however receives all its authentication requests via GSM network and it is not publically accessible via an IP address. The communication with remote modules such as data matrix generating modules should be done via VPN connections rather than using internet. The absence of access to CVS from internet makes is secure against DoS and DDoS type of attacks.

4.2.3 Spoofing attacks & man in the middle attack

Spoofing attacks are usually used for authenticating with a service. The attacker impersonates a legitimate user by using his/her IP or MAC address. In this case, the attacker can possibly try to spoof a GSM user's SIM. SIM spoofing is difficult in it's self, but it is not impossible. However, in this particular scenario it is futile because the attacker has no way of stopping the genuine user from receiving the response. Even if an SMSC is used to spoof many mobile numbers, the reply from CVS will reach the original owner of the SIM and not the SMSC generating the traffic. Similarly man in the middle attack is not practical because the attacker does not have the ability stop a GSM user connecting to its network. Incase a jammer is used, the attacker's own signal would also be jammed out and the attack would not take place.

4.2.4 Attack on communication between modules

The counterfeiters might try to listen to the communication between manufacturers and the CVS. Standard protocols and procedures are already in existence to secure communication between client and servers, such as used in online shopping and bank transactions. A standard secure connection would be used between CVS and manufacturer to avoid eavesdropping and data capture.

4.2.5 Attack on stored data

In case counterfeiters gain read access to the CVS database, they still won't be able to generate working Data Matrix because the verification code is not stored in any database at any point, only its hashed value is stored and generating a value for a hash is infeasible temporally, in this case also financially, because if a value is generated for a valid hash, it will only be useable for a single medicine package.

4.2.6 Copying working codes

Any verification code can only be used once, subsequent attempts at verification return negative response and therefore multiple copies of a working code can not be verified. The codes printed on medicine packing are not to be openly visible; they would either be covered by a removable coating or placed within the packing. This will ensure that medicine authentication request is sent by the end user and not any random person having access to the medicine packing.

4.3 Advantages over other solutions

The proposed solution has various advantages over the reviewed solutions. These are listed as follows.

4.3.1 User Level Verification

User level verification is a requirement of importance in developing and poor countries because of high levels of corruption and lack of regulatory measures and their weak implementation.

4.3.2 Ease of Use

The verification mechanism is fairly simple and easy for a common user. The Data matrix reader can be sent to the user's phone upon request, similar to the internet settings and apps being sent by the GSM network. The installation of Data Matrix reader is a one time process, subsequent verifications only require the user to run the application, snap a picture and send an SMS. The verification response is sent to the same number automatically.

4.3.2 Reliable Results

The results are reliable if the GSM network is reliable. That means, as long as the GSM network sends the message to the CVS, the response is reliable because different attacks described earlier are not possible or countered.

4.3.3 Scalable

The proposed solution is very scalable because there is on human involvement on the operational level and every module is designed to work automatically. Compared to other solutions that require human participation in the working steps, the proposed solution is much faster. Different functions are performed by independent modules that can be hosted on separate servers and any module that is over loaded can easily be replicated for load balancing. Hosting the CVS on a cloud server will give it hardware and geographical independency.

	General Solutions	mPedigree	EFPIA	NAFDEC	Epothecary	Proposed Solution
User level Verification	✓	✓		✓	✓	✓
Ease of Use		✓	✓	✓	✓	✓
Reliable Results				✓	✓	✓
Security against Attacks					✓	✓
Secure against compromised Database						✓
Scalable	✓		✓	✓		✓

Table 4.1. Comparison of different solutions based on required features

	General Solutions	mPedigree	EFPIA	NAFDEC	Epothecary	Proposed Solution
User level Verification	✓	✓	✗	✓	✓	✓
Ease of Use		✓		✓	✓	✓
Reliable Results	✗	✗		✓	✓	✓
Security against Attacks				✗	✓	✓
Secure against compromised Database					✗	✓
Scalable						✓
Solution/s Containing all Required features						✓

Table 4.2. Elimination of solutions based on absence of required features

4.4 Comparison between reviewed and proposed solution based on Scalability and Resilience against brute force attacks

The given comparison between these reviewed solutions and proposed solution is based on their capacity, potential for scalability and resistance against brute force attacks for a given number of attempts. The trends that emerge from the analysis are discussed in detail.

The table below shows the values used for generating the graphs. It should be noted that all the values for the proposed solution presented in the table and graphs are based on an identifier string length of **48 characters.**

	Total pool size	**Useable identifiers**	**Brute force attempts needed for a single match in** **One million codes**	**Brute force attempts needed for a single match in** **ten million codes**
mPedigree	100000000	340	100	10
NAFDEC	2176782336	7401.059942	2176.782336	217.6782336
Proposed Solution	9.6068E+100	3.26631E+95	9.6068E+94	9.6068E+93

Table 4.3 Comparison between reviewed and proposed solutions

The values of useable keys are derived using the 6 sigma standard of quality. The six sigma standard was developed by Motorola and several other companies are now following this standard. The six sigma standard dictates that the probability of error should be in the 6th movement around the mean on a standard normal bell curve. For simplicity and easy understanding, numerical values are also given. The six sigma standard requires the number of errors or defective products to be less than 0.00034%. In this evaluation, the error or defect refers to a false negative response i-e when a verification code is genuine but fails to verify. It is assumed that a brute force attack can be mounted for verification of working codes so that later on legitimate codes will return negative response and thereby undermining the system's reliability.

The number of brute force attempts needed for a single correct verification depends upon the ratio between the total pool size 'P' and the number of selected identifiers 'n'. The number or required brute force attempts is calculated by 'P/n'

The graphs given below represent these values.

4.4.1 Results based on Six sigma standard

The figure 4.1 shows the values of useable keys in different solutions if used in compliance with the six sigma standard. These values are same as the number of useable codes because if more codes are used, the probability of success for a brute force attempt goes higher than the permissible error rate.

The number of useable identifiers is derived by the following formula.

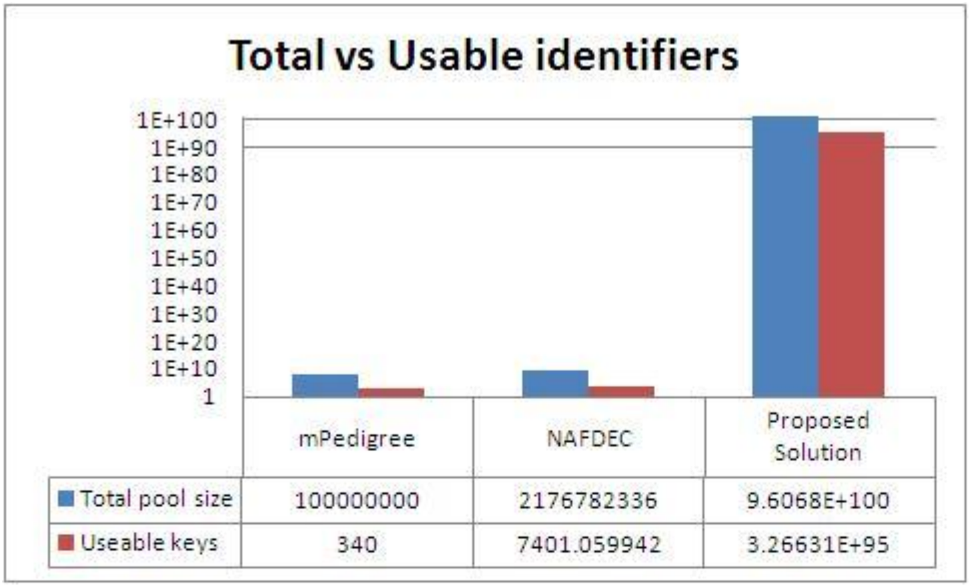
Let the pool size = x

Let the required success rate = n

The permissible error rate = 1-n

The maximum number of permissible errors in the pool size is given by

$$X * (1-n)$$



	mPedigree	NAFDEC	Proposed Solution
🟦 Total pool size	100000000	2176782336	9.6068E+100
🟥 Useable keys	340	7401.059942	3.26631E+95

Figure 4.1: Values of total possible identifiers & useable identifiers

4.4.2 Results based on number of identifiers used

The figure 4.2 shows the number of brute force attempts required for one success. Here an inverse linear relation between the number of used identifiers and brute force attempts is seen. The proposed solution is not included in this graph for a better comparison between the reviewed solutions.

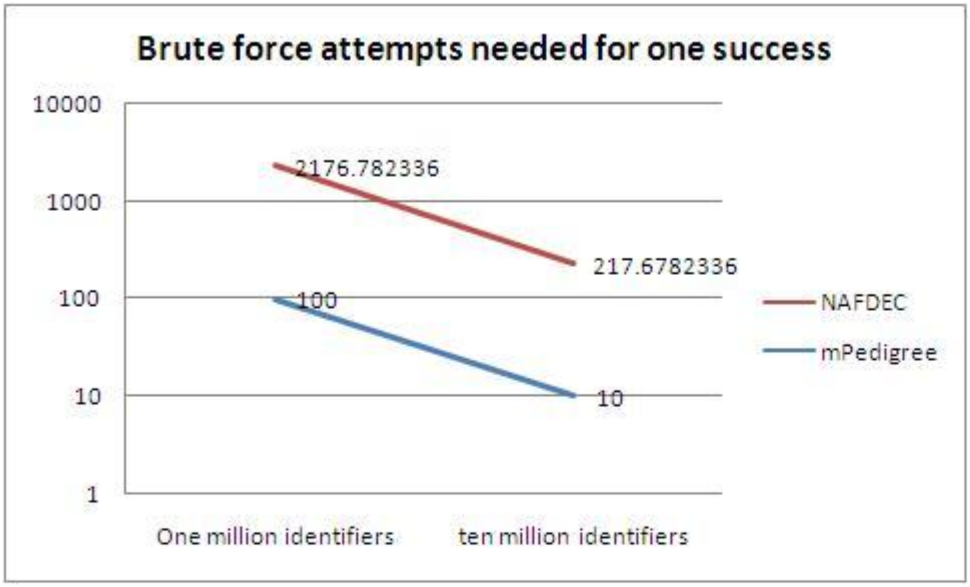


Figure 4.2: The number of brute force attacks needed for one success

Figure 3 represents the same comparison as figure 2 with the inclusion of data for the proposed solution. It is clearly observed that the proposed solution is much more secure against brute force attacks as compared to the other reviewed solutions; however the inverse linear relationship between the number of used identifiers and required brute force attempts remains the same.

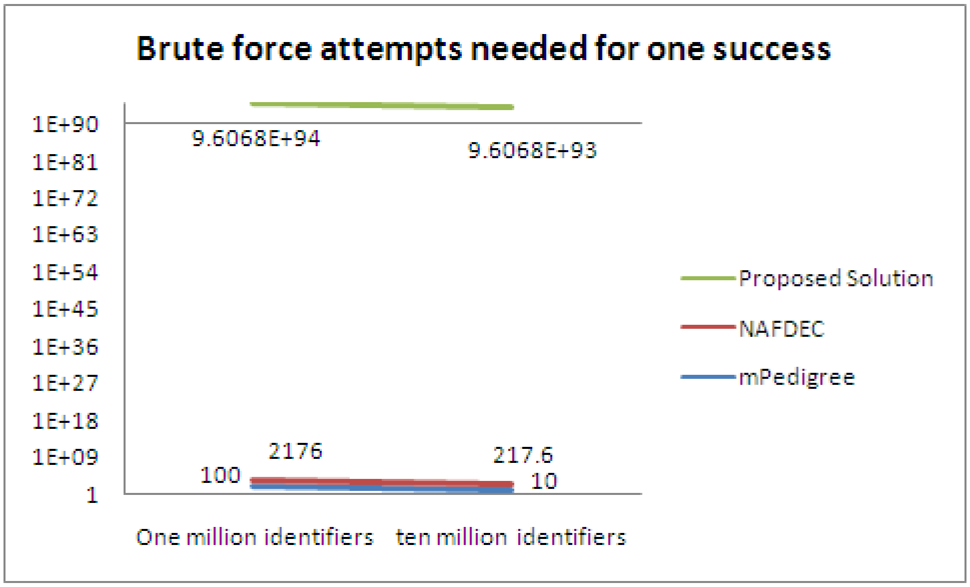


Figure 4.3: Comparison between number of brute-force attempts needed for proposed solution and reviewed solutions

4.5 Conclusion

From the above given tables, it can be clearly seen that while every solution has its advantages and positive points, they all have some draw backs that jeopardize their usefulness. While all features are important, the user level verification is important for countries where medicine regulation laws are weak. Scalability is another important feature if a solution has to gain large scale implementation. Security is the most important feature of all because that is what stops counterfeiters from faking. It is logical to think that a multimillion dollar industry will not quit its efforts to gain high profits and will try to bypass any restricting and controlling mechanisms. Therefore a medicine verification system should be secure enough to withstand and thwart known attacks, scalable to allow for large scale implementation and simple enough to be easily useable.

Chapter 5

Prototype development and testing

Introduction

Alongside theoretical analysis and designing the architecture of the proposed solution, a skeletal prototype is also developed to test and verify the assumptions made regarding user level verification feasibility of the system. The primary goal of prototype development was determining the usability of the solution and the response time of the system. Potential attacks on the data, based on the knowledge of database architecture were also tested.

5.1 Components and Equipment

Two types of components and equipment were involved in testing; that are, user side and server side. The user side simply comprises of a camera phone with a data matrix reader application installed. The server side components are listed and described below.

5.1.1 GSM modem for sending and receiving SMS messages

A GSM modem is required to receive SMS messages on the computer. For the testing purpose, a Sony Ericson K750i GSM phone was used as a GSM modem. The message sending/receiving speed of a phone is considerably less than that of a dedicated GSM modem, however this experiment did not require rapid SMS sending so the phone served as a GSM modem efficiently. The mentioned model of phone charged its battery via the USB data cable and therefore did not require special maintenance in that regard.

5.1.2 SMSC for handling the message reception and sending

Several different SMSC are available in open source and propriety software for different platforms. Different SMSC have different capabilities and features. Some can only handle SMS while others can also handle MMS and other data services. Some SMSC can connect to multiple GSM modems or channels and process their traffic simultaneously while others can support only one at a time. For the prototype testing and development, we used the trial version of NowSMS Lite [19] , which is an SMSC developed for the windows platform and supports the use of GSM phones as GSM modems. This SMSC was selected because of its light weight processing requirements and easy configuration options. SMS using special key words mentioned in the beginning of the body of SMS text are processed or forwarded based on the settings for those particular key words. In testing, all SMS containing verification data started with "ABC" which identified the SMS as a verification request, and a reply was always sent to the sender based on the verification results, all other SMS were logged and no other action was taken.

5.1.3 Parsing script for processing the received SMS

A web server was setup to interact with the SMSC. The SMSC forwarded the SMS to parsing script running on the web server. The script performed the following tasks

- Extract required information from the received data, in this case, the sender number and the body of text.
- Remove the "ABC" key word and generate a query using the body text of SMS.
- Choose an appropriate response to the sender based on the query results.
- Insert verification event record in the database.
- Send the appropriate response to the sender's number via the SMSC.

5.1.4 Database containing verification data

The database containing the verification data was hosted on the same web server as the parsing script. It is possible to host it elsewhere also. The database contained two tables. One with read only access to the parsing script, the other with read and write permissions. The read only table contains the verification values, and its kept read only to avoid any potential changes by manipulation of the script. Every successful verification event is logged by the script in the log table. When a verification request is received, the script checks the value in both tables. If it only exists in the Records table, the user replied with a "successful verification" response, if the verification value does not exist in any table, the user is replied with an "invalid verification code" response, and if the value exists in both the Records table and the Log table, the user is replied with "Used verification code" response.

5.1.5 Data matrix generator

Several data matrix generators are available online. For use in the experimentation, we used an online QR code generating service[20]. The service can generate QR code in different sizes, ranging from small to extra large. The experiments were done with small and medium sized QR codes.

5.1.6 Phones used for Data matrix reading

Two different models were used for capturing & reading data matrix. These were Nokia 6230 and Nokia 6233. These phones are equipped with 2 mega pixel cameras. Their performance was similar to each other in reading data matrix.

5.1.7 Data Matrix Reader application

A data matrix reader application has been developed for Andriod phones and tested successfully using a webcam and Android simulator. However, due to the unavailability of Android phone at the time of testing, the testing was done using a mid range Nokia cell phones. The software used for this purpose was selected from those available over the internet. The software used is called i-nigma reader[21]. This software was selected based on its features such as automatic SMS access and a wide range of support for several different models of mobile phone from different manufacturers.

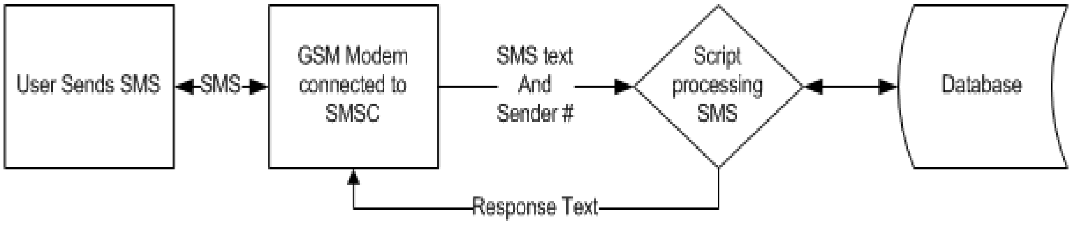
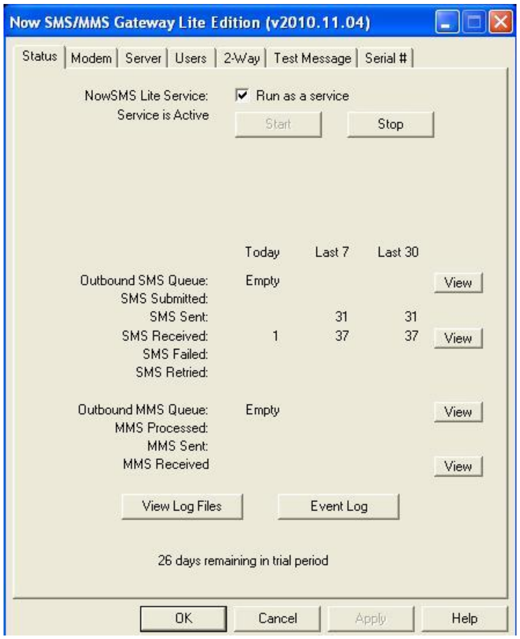


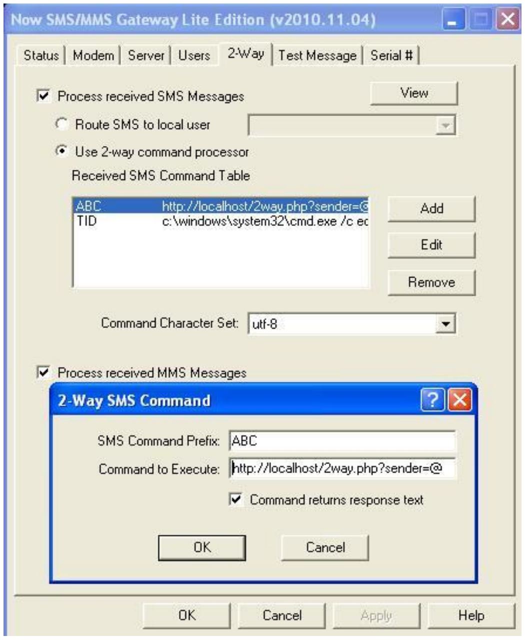
Figure 5.1. This image represents the communication flow of the verification process



NowSMS Lite Service: ☑ Run as a service
Service is Active

	Today	Last 7	Last 30
Outbound SMS Queue:	Empty		
SMS Submitted:			
SMS Sent:		31	31
SMS Received:	1	37	37
SMS Failed:			
SMS Retried:			
Outbound MMS Queue:	Empty		
MMS Processed:			
MMS Sent:			
MMS Received			

26 days remaining in trial period



☑ Process received SMS Messages
○ Route SMS to local user
● Use 2-way command processor

Received SMS Command Table

ABC	http://localhost/2way.php?sender=@
TID	c:\windows\system32\cmd.exe /c ec

Command Character Set: utf-8

☑ Process received MMS Messages

SMS Command Prefix: ABC
Command to Execute: http://localhost/2way.php?sender=@
☑ Command returns response text

Images: 5.2 Shows the status of the SMSC.

5.3 Shows the configuration setting tab for setting key words and address of the processing script.

5.2 Testing and Results

Several records of random identifiers were inserted into the database; QR codes for these records were generated using the online QR code generator such that the GSM Modem's number was fed automatically to the recipient number upon reading of the QR code. The testing was done using QR codes exceeding hundred characters in length; printed size of these codes was 4.5 x 4.5 cm. The phones used to capture and read the QR codes were Nokia 6233, and Nokia 6200.

Different sizes of QR codes were tested; the readability results are given in the tables below.

Table 5.1 shows the readability results of small size QR codes generated using two different sources. The first source[22] generates smaller QR codes but the readability of these is poor, the second source[23] generated relatively larger QR codes for the same data with good readability results.

Length of Verification code	**Small QR code 1st source**	**Readability**	**Small QR code 2nd source**	**Readability**
Twelve numeric	2.4 cm^2	Poor	4.84 cm^2	Fair
Twelve alphabetic	2.4 cm^2	Poor	4.84 cm^2	Fair
Twelve alphanumeric	2.4 cm^2	poor	4.84 cm^2	Fair
Twelve **Alhpanumeric & special characters**	2.4 cm^2	Poor	4.84 cm^2	Fair

Forty-eight numeric	2.4 cm2	Very poor	5.47 cm^2	Fair
Forty-eight alphabetic	2.4 cm2	Very poor	6.72 cm^2	Fair
Forty-eight alphanumeric	2.4 cm2	Very poor	6.72 cm^2	Fair
Forty-eight alphanumeric and Special characters	2.4 cm2	Very poor	6.72 cm^2	Fair

Table 5.1 Small QR code sizes and readability results

Table 5.2 shows the readability results of medium size QR codes from the same sources. The 2nd source has better results as compared to the first source.

Length of Verification code	**Medium QR code 1st source**	**Readability**	**Medium QR code 2nd source**	**Readability**
Twelve numeric	4.7 cm^2	Fair	5.81 cm^2	Good
Twelve alphabetic	4.7 cm^2	Fair	5.81 cm^2	Good
Twelve alphanumeric	4.7 cm^2	Fair	5.81 cm^2	Good
Twelve	4.7 cm^2	Fair	5.81 cm^2	Good

Alhpanumeric & special characters				
Forty-eight numeric	5.25 cm^2	Fair	6.56 cm^2	Good
Forty-eight alphabetic	5.25 cm^2	Fair	8.06 cm^2	Good
Forty-eight alphanumeric	5.25 cm^2	Fair	8.06 cm^2	Good
Forty-eight alphanumeric and Special characters	5.25 cm^2	Fair	8.06 cm^2	Good

Table 5.2. Medium QR code size and readability results

The following points were observed in the testing of the prototype.

- Installation of the QR code reader took from 2 to 5 minuets, depending upon the speed of the GSM network
- The process of automatic SMS writing always occurred successfully upon successful reading of the QR code.
- The response time from the server was always under 1 Minuit.
- QR code increased in size with the increase in the length of character string.
- Attempts to insert a query targeting the database via SMS failed.
- Very small sized QR codes could not be read by the software. These QR codes were 2.4 x2.4 cm in size.
- Medium sized QR codes were successfully read 4 out of 5 times in good lighting conditions. This was rated as fair performance

- A successfully read QR code was always read accurately, i-e characters were not changed in anyway.
- New users took only a few minutes to understand how to use the system, the time taken to explain the steps to them.

These observations show that larger size of QR code is required for use in medicine verification owing to the long string length of verification code and the over head of the automation process of SMS writing.

The results of these tests are satisfactory because all verifications were successful in the first attempt and the response was timely i-e within one minute. New users took a very short time to understand and use the system with ease. Lighting conditions and a steady hand emerged as important factors in the successful reading of the QR codes.

5.3 Deployment suggestions

The prototype testing shows the ease of use and high rate of success in the verification process. However, the architecture used for prototype testing is not suitable for large scale implementation. The prototype server communicates using a GSM modem, which in tern is dependant upon the GSM service provider's cell for its working. In case of GSM congestion or the lack of service within that particular cell, the whole system will become dysfunctional. Also, the GSM modem has a limited capacity of sending/receiving SMS messages per unit time. A large scale deployment can not be supported by a GSM modem. The use of GMS modem also binds the server to a particular geographical location which in itself can be a security threat to the system and also cause dependence on the geographic resources such as power and internet access.

To avoid these problems and provide ease of scalability, the CVS should be hosted on a cloud server. This CVS should NOT be given a public IP address for security reasons. The GSM network can forward the traffic for CVS to its SMSC using a VPN connection. Alternately, the CVS can use an SMSC of the GSM network and route its traffic to the parsing module from there. A short number which should be uniform across different GSM providers should be assigned to the CVS. This will ensure that counterfeiters will not use a GMS modem based SMSC for authenticating their own products.

5.4 Conclusion

The testing results show the ease of use of the proposed solution. The installation of the required software from the internet is a one time effort and it takes only a few minuets. Once the solution is deployed over a GSM network, the software can be sent to a user’s phone in response to an SMS, similar to the way setting and other applications, which will make the installation even easier. The working steps are few and simple for the average users to follow.

Chapter 6

Conclusion

Counterfeit medicine sales are a global health problem and its severity and the importance of a viable solution can not be overstated. The lack of regulatory mechanisms and laws, and their improper and weak implementations in different countries provide an open market to counterfeiters who put valuable human lives at risk for their profits. A practical and efficient medicine verification mechanism would discourage the sales of counterfeit medicine and discourage the counterfeiters by diminishing their gains. However, in such a case, the counterfeiters are sure to make attempts at compromising and undermining such a solution.

The main objective of this thesis research was to develop a framework for medicine verification that would be secure against different types of attacks. Along with that, it was geared toward a solution that would be user friendly and scalable. In this thesis, we have not addressed the aspect of tracking, which is a focus of different solutions and research. This is so for two reasons, firstly the identification of counterfeit drugs would discourage counterfeiters and can ultimately stop them from manufacturing; secondly this mechanism can work in parallel with any other tracking technique that might be developed subsequently.

Various aspects of developing and finalizing different aspects and modules for the complete project development is a daunting task and would require resources and further research in greater detail. Solutions of this nature are not possible to realize without the proper support of a funding agency. To achieve this end, proposals have been submitted to different research funding organizations. Incase of proper funding, various aspects of this framework can be researched and developed, such as the database at CVS and the proper right assignment to the registered manufacturers, the communication security between the identifier generator and the QR code generator, custom software for different modules such as QR code generators, QR code reading application and SMS handling and parsing scripts.

Furthermore, the same procedure can be applied for the verification of other products to ensure their originality and source. It can safely be said that the advancement in communication technology and the penetration and precipitation of mobile phones to lower income members of society will keep increasing and there by the scope of Data matrix codes and such verification mechanisms.

References

1. World Health Organization (WHO) article on “Counterfeit Medicine” http://www.who.int/medicines/services/counterfeit/overview/en/

2. Hirschler Ben, report on Pfizer survey regarding “counterfeit sales in Europe “ published at Reuters UK website feb 16, 2010 http://uk.reuters.com/article/idUKLDE61E16A20100216?sp=true

3. http://rpsgb.org.uk/pdfs/pr091103.pdf

4. “The Counterfeiting Superhighway” published by European Alliance for Access to Safe Medicines 2008 http://v35.pixelcms.com/ams/assets/312296678531/455_EAASM_counterfeiting%20report_020608.pdf

5. Barnes Kirsty, “New Counterfeit Report Highlights Worrying Trend” published Nov 7, 2007 http://www.outsourcing-pharma.com/Contract-Manufacturing/New-counterfeit-report-highlights-worrying-trends

6. Megget Katrina, “Drug theft costs industry up to $1BN a year” published Oct 2, 2007 http://www.outsourcing-pharma.com/Contract-Manufacturing/Drug-theft-costs-industry-up-to-1bn-a-year

7. Pitts Peter, “Counterfeit Drugs to Reach $75BN by 2010” Published Nov 2005.
http://www.heartland.org/policybot/results/17948/Counterfeit_Drug_Sales_to_Reach_75_Billion_by_2010_Report_Says.html

8. “Counterfeit Drugs Kill” factsheet by WHO and International Medical Products Anti Counterfeiting Taskforce IMPACT, published in 2007
http://www.gphf.org/images/downloads/impactbrochure.pdf

9. Nunwood Survey for Pfizer, “Cracking counterfeit Europe” published November 2009
http://www.pfizer.co.uk/sites/PfizerCoUK/Media/Pages/CrackingCounterfeitEurope.aspx

10. WHO Report “What encourages counterfeiting of Medicine” http://www.who.int/medicines/services/counterfeit/faqs/15/en/index.html

11. Medicine and Healthcare Products Regulatory Agency MHRA (UK) “Anti-Counterfeiting Strategy 2007-2010”
http://www.mhra.gov.uk/home/idcplg?IdcService=GET_FILE&dDocName=CON2033156&RevisionSelectionMethod=Latest

12. Nordwall Label Manufacturer’s report “Solutions to combat counterfeit drugs”
http://www.ngpharma.eu.com/article/Solutions-to-combat-counterfeit-drugs/

13. Tylor Phil, “EFPIA update on Swedish Coding Pilot” Published Oct 23, 2009
http://www.securingpharma.com/40/articles/263.php

14. EFPIA "Coding & Identification of products: Towards safer medicines supply" May 2009
http://www.efpia.org/Content/Default.asp?PageID=566

15. Okoyo Chidi, "NAFDAC to fight drug counterfeiting by SMS" published March 5, 2011
http://dailytimes.com.ng/article/nafdac-fight-drug-counterfeiting-sms

16. Safemedicines.org sponsored "Counterfeit Drug Incident Encyclopedia"
http://www.safemedicines.org/counterfeit-heparin-blamed-for-worldwide-deaths.html

17. Counterfeit Medicine Factsheet by WHO, published Jan 2010
http://www.who.int/mediacentre/factsheets/fs275/en/

18. M Paik, J Chen, L Subramanian "Epothecary: Cost-effective Drug Pedigree Tracking and Authentication Using Mobile Phones" ACM MobiHeld09

19. NowSMS Lite SMSC for use with a single GSM Modem
http://www.nowsms.com/productinfo/nowsms-lite

20. Online QR code generator
http://delivr.com/qr-code-generator

21. Data-matrix/QR code reader free software for mobile download
www.i-nigma.mobi

22. Online QR code generator
http://qrcode.kaywa.com/

23. Online QR Code Generator
http://zxing.appspot.com/generator/