

Study of Cross-Site Scripting Attacks and Their Countermeasures

Gurvinder Kaur
University Institute of Engineering and Technology
Kurukshetra University
India

Abstract: - In present-day time, most of the associations are making use of web services for improved services to their clients. With the upturn in count of web users, there is a considerable hike in the web attacks. Thus, security becomes the dominant matter in web applications. The disparate kind of vulnerabilities resulted in the disparate types of attacks. The attackers may take benefit of these vulnerabilities and can misuse the data in the database. Study indicates that more than 80% of the web applications are vulnerable to cross-site scripting (XSS) attacks. XSS is one of the fatal attacks & it has been practiced over the maximum number of well-known search engines and social sites. In this paper, we have considered XSS attacks, its types and different methods employed to resist these attacks with their corresponding limitations. Additionally, we have discussed the proposed approach for countering XSS attack and how this approach is superior to others.

Keywords: - Cross-Site Scripting (XSS), Malicious Injection, Web Security, and Web Application Attacks.

1. INTRODUCTION

With the everywhere-ness of information superhighway, i.e. Internet, organizations are serving people with their business on web. However, as the owners of the business emphasize greater on their business logic they do not get concerned about the vulnerabilities and security hazards inclined to their websites. Web Security describes the guidelines used to block threats to diminish the web attacks. An attack may be feasible due to the existence of vary types of flaws and bugs in the coding. As per Ponemon Institute Life Threat Intelligence Impact Report 2013 if the actionable intelligence about cyber attacks is available only 60 seconds before then the average cost of exploit could be reduced to 40 percent [1]. That is if we have an appropriate method to handle an attack at the very first step then the cost of the damage caused due to that attack can be diminished largely.

The inaccurate authorization and sanitization of data given by web server has brought in the accountability for XSS attacks. It is the attack on the secrecy of customer of a specific website by approving injection of inputs containing HTML tags and JavaScript code. As per OWASP (Open Web Application Security Project) 2013 release cross-site scripting is one of the major attacks performed [2]. Cenzic Application Vulnerability Trends Report 2013 confers that among the top 10 attacks 26% comprises of XSS attacks only [3].

The rest of the paper is organized as follows:

Section II discusses the web application architecture. Section III discusses the XSS attacks and its types in detail. Section IV provides the survey explored with their relative weaknesses. Section V discusses the proposed approach and how it would be better. Finally, section VI concludes the paper.

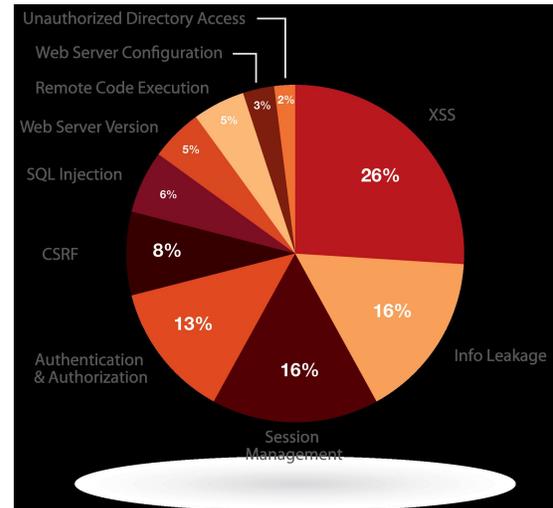


Fig1: Web Application Security Vulnerability Population [3]

2. WEB APPLICATION ARCHITECTURE

As XSS attacks occur over the application layer so it is important to know how the web application works over the internet. Web Application has three-layered architecture as shown in Fig 2.

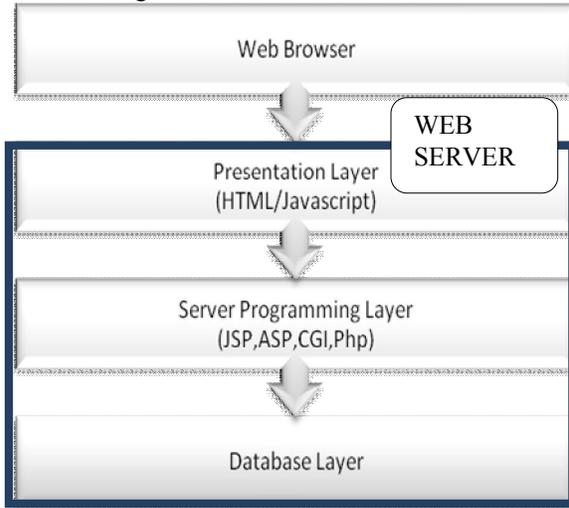


Fig 2: Web Application Architecture [4]

- 1.) *Presentation Layer*: This tier accepts input from end user and display output to user. It functions as a graphical user interface (GUI). It literally attach with the client.
- 2.) *Server Programming Layer*: This layer is located between presentation layer and database layer. Data processing is handled in this layer and it can be programmed in any of server scripting languages like JSP, PHP and ASP etc.
- 3.) *Database Layer*: This tier stores and manages all the delicate data of web application. This layer is responsible for access of authenticated users and rejection of malicious users.

3. CROSS-SITE SCRIPTING ATTACKS

In this section, XSS attacks & its types are discussed in detail. An XSS attack is one of the most common web application attacks that are used by hackers to sneak into web applications. In XSS, attacker embeds malicious script into a website. Whenever a user browser run this code the attacker can shape the browser to do whatever it wants .XSS attacks occur whenever an application takes un-trusted data and sends it to web browser without proper validation and sanitization[5]. So in XSS attacks three parties

are involved- the attacker, the client and the website. In XSS attacks, the attacker insert malicious scripts to target websites for session hijacking, cookie stealing, and malicious redirection. This attack arises, as the web server does not appropriately assure that generated pages are properly encoded to avert the inadvertent execution of scripts and when input is not justified to prevent malicious HTML from being displayed to the users.

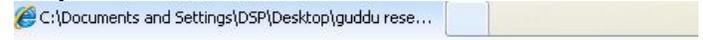
Example:

```
<% out.println(" welcome " + request.getParameter("name")); %> (Example of poorly –written code on Web server- saving it as test.jsp)
```

Case 1: Normal User

```
<HTML>  
<BODY>  
Welcome Stefan  
</BODY>  
</HTML>
```

Output:



Welcome Stefan

Fig 3: Response of a normal HTML code

Case2: Attacker

```
<HTML>  
<BODY>  
Welcome <script>alert ("Attacked") </script>  
</BODY>  
</HTML>
```



Welcome



Fig 4: Output after code inserted by attacker

The example above shows a web site having XSS vulnerability. In Fig 3 a normal user when enters his name it is displayed. In Fig 4, an attacker misleads the application server by entering the JavaScript in place of name. This is quite a simple example.

There are primarily three types of cross-site scripting attacks as follows:

1. *Stored or Persistent attacks*- This attack appears when the malevolent code is sent to a web application where they are stored permanently on the mark server. When a user requests for the stored information this script is executed in its context. Examples for raider favorite targets include web mail messages, forum, comment field, visitor log etc. this attack occurs in the following manner:

-Malicious script is inserted into a web form by the attacker, which is then stored by the server in the database.

-When a request is send to the infected page by the user, the script is executed in the user's browser and the user's cookies are passed to the attacker.

2. *Reflected or Non – persistent attacks*- In this the malicious code is breed into a URL as a value of parameter, as a HTTP query in such a way that the rebounded content consists of unprocessed script. This attack occurs in a following way:

-Attacker examines a website and finds vulnerability on a web page.

-Attacker may embed a URL to exploit the weakness and may send an email to the user captivating the user to click on a link for the URL under false charade

-The URL will point to the normal website, but it will contain attacker's malicious code that the site will reflect

-User will visit the URL provided by the attacker while log into the given website

-The malicious script is executed on the user's browser as if it came directly from the web server.

-By this way, an attacker can steal sensitive information through user's cookies [6].

3. *DOM (Document Object Model)-based XSS attacks*- This exists within a site and can be used in a reflected manner. In this case, the malicious data exists solely in the browser and is not sent to the server. A brief example of a DOM-based XSS attacks would be a modifying the web history of a user.

4. RELATED WORK

Johns *et al.* has proposed a passive detection system to identify successful XSS attacks [5]. It uses two different approaches based on generic observations of XSS attacks and web applications. In this reflected attack is detected by a request/response matching which is based on the direct relationship between the input data and the injected scripts. In this the input parameters and the scripts found in final HTML is converted into a non-ambiguous representation by removing all encodings and the appropriate matching is done by constructing a DFA for each of the input parameter. For stored attacks, it adopts a generic XSS detection using a list of known scripts in which they used a training based XSS detector in which list of all outgoing script is matched up with the detector's known list. The weakness of this system is that it uses different implementation schemes for the two types of XSS that increases the overhead. It just detects the already existing attacks and false positives are there.

A static analysis for finding XSS vulnerabilities is demonstrated by Wassermann & Su [7] that straightforwardly addresses weak or absent input validation. The approach integrates work on tainted information flow with string analysis. The proposal has two parts: (1) an adapted string analysis to track untrusted substring values, and (2) a check for untrusted scripts based on formal language techniques. String-taint analysis not only represents the set of string values a program may create, it also defines the formal language representation with labels that indicate which substrings come from untrusted sources. The second phase of the method enforces the policy that generated web pages include no untrusted scripts. It has many disadvantages like the tool produces false positives and it failed to resolve certain alias relationships between variables whose values are used for dynamic features. It failed to detect the DOM-based XSS. The string analysis-based tool could not handle arbitrarily complex and dynamic code.

Wurzinger *et al.* [8] introduced a tool known as SWAP (Secure Web Application Proxy), a server-side solution for discovering and preventing cross-site scripting attacks. SWAP contain a reverse proxy that intercepts all HTML responses, as well as a make use of modified Web browser to detect script content. SWAP contains a JavaScript detection component,

which is able to determine whether script content is present or not, a reverse proxy, which block all HTML responses from the server and subjects them to analysis by the JavaScript detection component and a set of scripts to axiomatically encode/decode scripts/ script IDs. SWAP introduces a performance overhead. It cannot guard counter to other kinds of objectionable content, such as static links pointing to sites including malicious scripts

In this paper a result for unit testing and action-level security of Struts Web applications is demonstrated by Wu *et al.* [9] by experimenting the applications from model-view-controller (MVC) respectively, and safe Struts applications with the help of different access control implementations. The aim of using struts is to neatly separate the model (application logic that communicates with a database) from the view (HTML pages shown to the client) and the controller (instance that passes info between view and model). JUnit tests the model, and StrutsTestCase does the testing of controller, while HttpUnit does the testing of view. The action level security solution comprises of four stages: access control in actions and JSPs, by extending the request processor, access control by servlet filtering, & WEB-INF. The solution of Struts Web application unit testing and action-level security just extended the general Web unit testing methods. The performance of application is degraded.

Galan *et al.* [10] suggested a multi-agent system for the automated scanning of web sites to disclose the existence of XSS vulnerabilities exploitable by a stored-XSS attack. The set of agent's part of the proposed architecture and the operation of the scanner are as follows. A webpage parser agent crawl the web application from which information about the different web forms found is used to build a repository of potential injection points (Injection point repository). A script injector agent reads the list of injection points recognized by the parser agent and also makes a selection of vectors attacks from the Attack vector repository. The desired set of attack vectors is launched against each of the potential points of attack of the application. A list of the performed attacks is stored in a Performed attack list. The verifier agent gets the list of the attacks to be verified and crawl the web application looking for each of the attacks. A report about the results of the scanning process

is elaborated and stored. This approach has not better performance & accuracy.

The approach by Putthacharoen & Bunyatnparat [11] aims to change the cookies so that they would become impractical for XSS attacks. This technique is called “Dynamic Cookie Rewriting” enforced in a web proxy where it will automatically put in place of the cookies with the randomized value before sending the cookie to the browser. In this way browser will keep the randomized value rather of original value sent by the web server. At the web server, end the return cookie from the browser over rewritten to its original form at the web proxy before being dispatched to the web server. So in case if XSS attacks swipe the cookies from the browser's database, the cookies cannot be used by the attacker to imitate the users. Four domains are kept to identify the cookie that is Name, Domain, Path, and Port. The initial value of the cookie and the randomized value are also kept in the same table. This table is reserved at the web proxy database server. The web proxy server will use this information to rewrite back cookies. The drawback of this approach is the compatibility problem that occurs while implementing the proxy server and the single point failure issue

In [12] a technique is proposed which is invoked when user injects code in the field of web application by V. & Selvekumar. It uses the complete HTML parser and JavaScript Tester to detect the presence of JavaScript for filtering it out. The user created HTML content is passed to the XSS sanitizer and the static tags are checked. The static tags are retained while rests of tags are filtered out. Even static tags contain dynamic content, which are filtered out by JavaScript Tester. After filtering HTML, content is converted into DOM. It includes parse tree generator at client side browser to reduce the anomalous behavior of browsers. It is restricted to server side only and browser source need to be modified for obtaining results.

Choudhary & Dhore [13] proposed code injection detection tool based on a Proxy Agent, which classifies the request as scripted request, or query based request. There are two modules: Query Detector and Script Detector. The HTTP request coming from client side is first send to the CIDT within which the request is passed to both modules one by one. Firstly, the Query detector validates the request and the query is rejected if any invalid character is found. Only

the valid requests are passed to the next module, the Script detector that also filters the request for invalid tags and encodes it before forwarding it to the web server. The disadvantage of this approach is it requires more time to response that is the delay time is more.

Matsuda & Koizumi [14] suggested a detection algorithm against cross-site scripting attacks by extracting an attack feature of XSS by considering the appearance position and frequency of symbols. It learns the attack features from given attack samples. In this three modules are presented. The first one is the classification module in which sample of 32 characters is gathered based upon the characters that occur frequent in attacks. The second module will calculate the important degree of characters. The final module will detect the attack-by-attack feature value & threshold that is taken as 15 for the proper detection of the attack in proposed approach. The main disadvantage is it is calculation based and it does not tackle the new attacks effectively.

Elhakeem & Barry [15] broach on the issues surrounding cross-site scripting attacks and providing a simple and useful security model to protect websites from such attacks using ZEND framework application. The security model is based on a chain of levels and is built using a combination of tools. It is divided into four levels as: Security Awareness, Server Security, Client Security, & Design Guidelines. The framework described by them is Zend Framework (ZF), which is an open source framework for developing web applications and services with PHP. This loosely coupled architecture allows developers to use components individually and offers a robust Model View Controller (MVC) implementation. The MVC paradigm breaks the application's interface, into three parts which are: Model: The model part of the application is the part that is concerned with the specifics of the data to be displayed, View: The view consists of bits of the application that are concerned with the display to the user, & Controller: The controller ties together the specifics of the model and the view to ensure that the correct data is displayed on the page. It accepts input from the user and instructs the model and view port to perform actions based on that input. It requires lot of tools to be combined so compatibility issues are there.

5. PROPOSED APPROACH

Cross-Site Scripting is one of the most dangerous and the common attacks found over the web applications. This survey presents study of the ongoing techniques against XSS attacks. These techniques suffer from the following weaknesses:

- Built-in limitations
- Partial implementations
- Complicated framework
- Developer's ability
- Run-time overhead
- False positives and false negatives
- Insecure channel between the web server and web browser
- Response delay
- Additional infrastructure
- Cost of deployment
- Don't prevent DOM based attacks

Our proposed system will try to remove almost all of these weaknesses. It will include a two-tier approach- one for detecting persistent and non-persistent XSS attacks and second for prevention of DOM based XSS attacks. For the first tier we will implement our logic of script guard in the controller part of MVC2 architecture of server. The controller receives all requests from the clients & forwards those requests to the respective pages as per request. The controller receives parameters sent by the client and scans these parameter values for suspected XSS attacks. These values are matched with sets of expressions where every expression match means an attack. In case of an attack, the requests do not go beyond the controller and the client is redirected back to the page where he requested. For the second tier that is for prevention of DOM based attacks we will have a small JavaScript code (DOM attack detector script) which is sent to the client with every response. This code acts only at client side and will prevent any sort of DOM based XSS attacks. Thus, our proposed work will detect all types of XSS attacks. Even it will have a little performance overhead but it will have a minimum response delay. There is no need of additional infrastructure and have not a complex framework.

6. CONCLUSION

XSS attack is one of the most common and dangerous web application attacks that can reveal information about a user or company profile. This paper presented what XSS attacks are, what are there types, the previous approaches for prevention of these attacks with there limitations. Then we showed our proposed approach and how it is better.

Many industries are employing web services for their benefits on the World Wide Web but for relieving themselves from the additional cost, they do not go for the security of the websites they created. Eventually it harms the users and company too. With the expansion of web applications, it is urgency to have an comprehensive and coherent structure for the prevention of unified XSS and other important web application attacks.

7. REFERENCES

- [1] <http://www.ponemon.org/blog/live-threat-intelligence-impact-report-2013-1>
- [2] https://www.owasp.org/index.php/Category:OWASP_Top_Ten_Project
- [3] <http://info.cenzic.com/rs/cenzic/images/Cenzic-Application-Vulnerability-Trends-Report-2013.pdf>
- [4] Jeom-Goo Kim, "Injection Attack Detection using the Removal of SQL Query Attribute Values", IEEE, pp. 1-7, 2011.
- [5] Martin Johns, Bjorne Englemann, Joachimm Posegga, "XSSDS: Server-side Detection of Cross-site Scripting Attacks", Annual Computer Security Applications Conference, IEEE, pp. 335-344, 2008.
- [6] http://en.wikipedia.org/wiki/Cross-site_scripting
- [7] Gary Wassermann, Zhendong Su, "Static detection of cross-site scripting vulnerabilities", ACM/IEEE 30th International Conference on Software Engineering, ICSE '08. pp. 171-180, 2008.
- [8] Peter Wurzinger, Christian Platzer, Christian Ludl, Engin Kirda, and Christopher Kruegel, "SWAP: Mitigating XSS Attacks using a Reverse Proxy", ICSE Workshop on Software Engineering for Secure Systems, IEEE, pp. 33-39, 2009.
- [9] Qinglin Wu, Yanzhong Hu, Yan Wang, "Unit Testing and Action-Level Security Solution of Struts Web Applications Based on MVC", International Conference on Biomedical Engineering and Computer Science, IEEE, pp. 1-4, 2010.
- [10] E. Galan, A. Alcaide, A. Orfila, J. Blasco, "A Multi-agent Scanner to Detect Stored-XSS Vulnerabilities", International Conference for Internet Technology and Secured Transactions (ICITST), IEEE, pp. 1-6, 2010.
- [11] Rattipong Putthacharoen, Pratheep Bunyatneparat, "Protecting Cookies from Cross Site Script Attacks Using Dynamic Cookies Rewriting Technique", 13th International Conference on Advanced Communication Technology (ICACT), IEEE, pp. 1090-1094, Feb 2011.
- [12] Sharath Chandra V., S. Selvekumar, "BIXAN: Browser Independent XSS Sanitizer for Prevention of XSS Attacks", ACM SIGSOFT, Volume 36 Number 5, September 2011.
- [13] Atul S. Choudhary And M.L Dhore, "CIDT: Detection Of Malicious Code Injection Attacks On Web Application", International Journal Of Computing Applications Volume-52-N0.2, pp. 19-25., August 2012.
- [14] Takeshi Matsuda, Daiki Koizumi, "Cross Site Scripting Attacks Detection Algorithm Based on the Appearance Position of Characters", 5th International Conference on Communications, Computers and Applications, IEEE, pp. 65-70, October 2012.
- [15] Yousra Faisal Gad Mahgoup Elhakeem , Bazara I. A. Barry, "Developing a Security Model to Protect Websites from Cross-site Scripting Attacks Using Zend Framework Application", International Conference on Computing, Electrical and Electronics Engineering (ICCEEE), pp. 624-629, August 2013.

Location Based Encryption-Decryption Approach for Data Security

Borse Manoj V.
Sandip Institute of
Engineering and
Management
Nasik, India

Bhandure Harshad D.
Sandip Institute of
Engineering and
Management
Nasik, India

Patil Dhiraj M.
Sandip Institute of
Engineering and
Management
Nasik, India

Bhad Pratik B.
Sandip Institute of
Engineering and
Management
Nasik, India

Abstract: Data security is an important task in today's life. Data security can be done using GPS device. Among computer user mostly use data in electronic format. How to provide a security for data is important. In this paper, we propose a Location Based Data-Security System to secure data by applying Encryption-Algorithm and co-ordinate using GPS device. Encryption means of efficient secure integer comparison. The encryption technology cannot restrict the location of data decryption. In order to meet the demand of a location-dependent approach location-dependent data encryption algorithm is needed. A target latitude/longitude co-ordinate is determined firstly. The co-ordinate is incorporated with a random key for data encryption. The receiver can only decrypt the cipher text when the co-ordinate acquired from GPS receiver is matched with the target co-ordinate. GPS-based encryption is an innovative technique that uses GPS-technology to encode location information into the encryption keys to provide location based security. GPS-based encryption adds another layer of security on top of existing encryption methods by restricting the decryption of a message to a particular location. Our experimental results not only validate the effectiveness of our scheme, but also demonstrate that the proposed integer comparison scheme performs better than previous bitwise comparison scheme.

Keywords: encryption; decryption; security; GPS technology; location.

1. INTRODUCTION

Most of the data encryption techniques are location-independent. They cannot restrict the location of clients for data decryption. In proposed system, a novel location-dependent approach is used for incorporating location information into data transmission.

It is important to provide a secure and convenient data transmission. We propose a location-dependent approach for better data security. The client put the coordinates manually in application for data encryption. Then our application create a encrypted file and then we send that encrypted file using e-mail or by any external device to our destination .The client only decrypt the cipher text when the coordinate acquired from GPS receiver matches with the target coordinate. According to our discussion, the approach can meet the confidentiality, authentication, simplicity and practicability of security issues. As a result, the proposed approach can meet the demand for personal and industrial data security.

2. PROPOSED WORK

Enhancing the security is the prime aspect of the proposed system. By adding the location based services with the encryption process one can make the data more secure.

System consists of following components:

1. Login and Registration
2. Encryption
3. GPS Interfacing and Location Matching
4. Decryption

2.1 Login and Registration:

Login and registration module provide user the access rights to interact with the system. Registration contains some basic

details regarding to username, password and email id. Login uses username and password to allow the user to pass in to the system.

For storing the details, we use SQL server 2005. For username and password separate table is maintained. Tables are handled by administrator.

2.2 Encryption:

The process of converting the plaintext to human non understandable form, so that if the data is obtained by third party person then they will not able to understand or retrieve it.

For this purpose, we use various algorithms like M. Aikawa et al. proposed a light-weight encryption algorithm for the copyright protection. T. Jamil proposed an enhanced algorithm for the typical DES algorithm, called AES (Advanced Encryption Standard). J. Jiang proposed a parallel processing algorithm for the RSA. S. Lian et al. proposed a fast video encryption scheme based on chaos. M. McLoone and J. V. McCanny designed a hardware circuit for DES based on the FPGA technique. M. Shaar et al. proposed a new data encryption algorithm, called HHEA. M. E. Smid and D. K. Branstad analyzed the past and future of DES algorithm. Y P. Zhang et al. proposed a stream cipher algorithm with respect to the traditional block-based cipher approaches [2].

Location co-ordinates are used as a 'key' for encrypting the contents.

2.3 GPS Interfacing and Location Matching:

Global Positioning System satellites broadcast signals from space that are used by GPS receivers to provide current location by making use of longitude and latitude.

The interfaced GPS device will appear as virtual serial port on PC to which one can communicate through our designed software which can transmit receive by this serial port like HyperTerminal or custom made software.

Location matching is the key process for successful decryption of data. The co-ordinates fetched by GPS must be matched with the co-ordinates which were entered while encrypting the data. As current location retrieved by GPS device will not be exactly same every time due to weather conditions, etc. Tolerance distance (TD) important role in rounding up or down the co-ordinate values at certain extent.

2.4 Decryption:

The location co-ordinates which were used as key while encryption must be matched with co-ordinates values fetched by GPS device at receiver side. If this condition is satisfied then only user can decrypt the data otherwise encrypted file will be discarded from the system automatically.

3. SYSTEM ARCHITECTURE

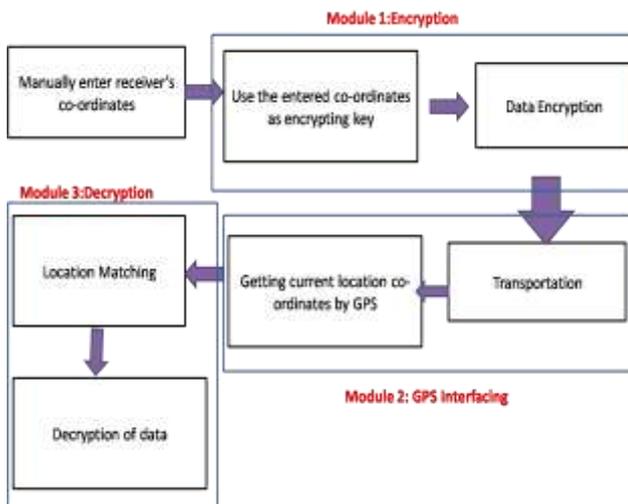


Figure. 1 System Modules

The figure shows the overall flow of the proposed system.

4. APPLICATIONS

- Military- In military this technology can be used to keep the data secured from the attackers during wars.
- Banks- This technology can also be used in banking for the purpose of money transaction.
- Individual use- It can also be used to store one's confidential data. For e.g.: for business purpose.
- Multinational Industries-In Industries important data can be secure by using this technology.
- College-In college's important data can be secure by using this technology. For e.g. Question paper.

5. CONCLUSION

Location's latitude/longitude co-ordinates plays vital role in the formation of encrypted data along with decryption process.

The proposed approach can be extended to the other application domain e.g. Authorization of software. If the system software is authorized within a pre-defined area, such as for particular organization the execution of the software may achieve the location check based on proposed approach. Decryption process is carried out when the authorized user is located in specified area.

This approach can be used for mobile applications such as in Smartphone.

6. REFERENCES

- [1] Swapna B Sasi, Betsy K Abraham, Jinil James, Riya Jose "Location Based Encryption using Message Authentication Code in Mobile Networks", In IJCAT International Journal of Computing and Technology Volume 1, Issue 1, February 2014
- [2] H.Liao, P.Lee, Y.Chao, C.Chen, "A Location-Dependent Data Encryption Approach for Enhancing Mobile Information System Security", In The 9th International Conference on Advanced Communicate Technology, pp. 625-626, Feb. 2007.
- [3] L. Scott, D. Denning, "A Location Based Encryption Technique and Some of Its Applications", Proceedings of ION NTM 2003.
- [4] V. Rajeswari, V. Murali, A.V.S. Anil, "A Navel Approach to Identify Geo-Encryption with GPS and Different Parameters (Locations And Time)", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (4), 2012.

Tagging based Efficient Web Video Event Categorization

A.P.V.Raghavendra
V.S.B Engineering College
Karur, India

Abstract: Web video categorization is one of the emerging research fields in the computer vision domain due to its massive volume growth in the internet which demands to discover events. Due to insufficient, noisy information and large intra class disparity makes it more daunting task to recognize the events. Most of the recent works focus on constrained (fixed camera, known environment) videos with supervised labelling to categorize the web videos. In this paper, we propose the subject based Part-Of- Speech (POS) Tagging technique with the assist of Named Entity Recognition (NER) and WordNet is applied on YouTube video titles to discover the events based on the subject, not on the objects visualized in the videos. Unsupervised learning method is used on high level features (titles) because of incoming videos are not known and large intra-class variations. For the experiment, we have chosen topics from Google Zeitgeist and downloaded the related videos from YouTube. A novel conclusion is derived from the experimental result that use of low level features will lead to a poor classification in discovering intra class events based on the subject of the videos.

Keywords: video categorization, Natural Language Processing, Parts Of Speech Tagging, Named Entity Recognition, WordNet

1. INTRODUCTION

Computer vision is one of the vast and important domain in which video event classification becomes more important than ever in nowadays. Video event classification is important because of the number of volumes growing at exponential rate in the internet and moreover replication of the identical video exist in different video sites calls a need to research and mine the efficient and effective events. According to YouTube [1], "100 hours of video are uploaded to YouTube every minute" shows the importance level of event recognition. In this paper, we have mined only the high level features to find subject based events and also proved that low level features will not be useful for classifying the intra-class events.

Lemmatization is the algorithmic process of determining lemma for a given word. In linguistics [3], "lemmatization is the process of grouping together the different inflected forms of a word so they can be analysed as a single term". Lemmatization is closely related to stemming. The difference is that a stemmer operates on a single word without knowledge of the context, and therefore cannot discriminate between words which have different meanings depending on part of speech.

Parts-Of-Speech (POS) Tagging is useful for identifying a word used in the corpus corresponding to a part of speech. In this work, we have used 36 types of POS tags to determine to each and every word. Named Entity Recognition (NER)[4] is the boost up for the POS tagging process because it is not able to identify the named words (names, places, songs, organizations, countries, etc). NER is used to tag the word with named entities. WordNet [5] is used to minimize the number of repeated terms to get an efficient mining result.

There are two novel conclusion found out in this work. First, the low level features are not efficient for classifying the intra class events for that SIFT (Scalable Invariant Feature Transformation) is used. Second, two level hierarchy events are represented such as for example Level 1 "blackberry" Level 2 "blackberry torch" which clearly shows that level 2 is

the specified part of level 1. The rest of the paper is organized as follows. Section 2 gives a brief overview of related works. Section 3 gives the detail of the proposed framework for unconstrained video classification. Section 4 presents experimental results. Finally section 5 concludes the work.

2. RELATED WORKS

Supervised learning method is used to classify the videos by means of training set. Some of the works are either limited to some specific domains (e.g. movies [12, 13], TV videos [14, 15, 16] etc.) or focus on certain predefined content such as human face [17, 18] and human activities [19]. Nowadays, the challenging task is to categorize the videos by using unsupervised labelling techniques.

In recent trends, web video event classifications are done by using high level and low level features. In [6], association Rule mining is applied on the titles and descriptions of the video to mine the events and the find events are used as the label for classifying the videos by using statistics and distribution characteristics. Video taxonomic classification systems are presented in [7, 8], with more than one thousand categories in consideration. However, in [7, 8], the taxonomic-structured category labels are predefined by domain experts. Also those categories can include anything and do not necessarily correspond to events.

3. PROPOSED FRAMEWORK

In the proposed framework, POS tagging is mainly applied on the YouTube titles by considering descriptions as a noisy data based on the analysis. The following diagram depicts the architecture of the proposed framework.

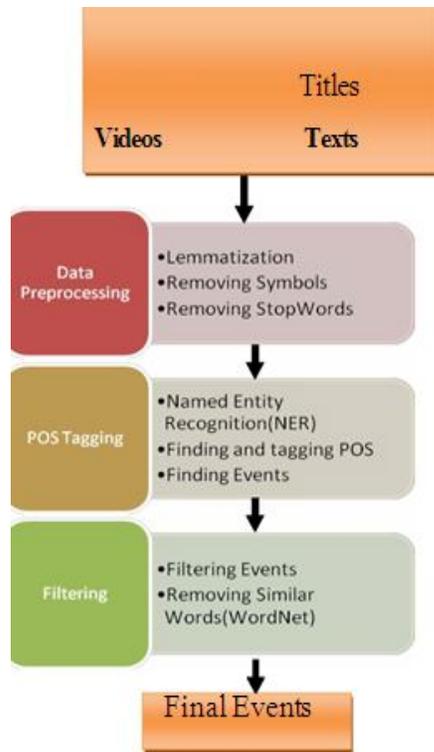


Figure 1. Architecture of proposed framework

3.1 Data Preprocessing

3.1.1 Lemmatization

Lemmatization or Stemming is the process of reducing inflectional forms and sometimes derivationally related forms of a word to a common base form. In stemming, the effective and popular algorithm is Porter's algorithm. But, there is small different between lemmatization and stemming is that stemming refers to a crude heuristic process that chops off the ends of words in the hope of achieving this goal correctly most of the time, and often includes the removal of derivational affixes. *Lemmatization* usually refers to doing things properly with the use of a vocabulary and morphological analysis of words, normally aiming to remove inflectional endings only and to return the base or dictionary form of a word, which is known as the *lemma*.
 Examples:

sinking -> sink+ ing -> sink
 Received -> receive
 Went -> go

3.1.2 Removing Symbols

3.2.2 Finding and tagging POS

Parts Of Speech (POS) tagging is the problem of assigning each word in a sentence the part of speech that it assumes in that sentence. In our work, there are 36 tags are covered to find and tag the words [2,10].

In the event recognition, symbols are not important to mine the events. Hence, we remove all the symbols (!@#\$%^&*()-_+=;:"<>!.^[]{})) by using the tokenization concept.

3.1.3 Removing stopwords

Stopwords is a set of words which is not constructive in the process of mining the events from the corpus text. The list of words that are not to be added is called a stop list. Stop list is used as a training set to remove all the stopwords. In order to save both space and time, these words are dropped at indexing time and then ignored at search time. We have used the stop list from [9] which is recommended and we have added own stopwords based on the analysis

3.2 POS tagging

3.2.1 Trained Entity Recognition

Named Entity Recognition (NER) is typically viewed as a sequential problem. Named Entities is the single term in the text corpus belonging to predefined classes such as person, location, nationalities, organizations, etc. The effective NER can be built by using heavy trained data. We have collected maximum data to make it as well trained set. The list of training data is given below as:

Table 1. Training set used in NER

List of topics	Number of words
Numbers, temporal words, currencies, measurements	1103
Names	649
Places	20226
Jobs	10,020 1,064
Titles	915
Location	952,674
Artworks	128,193
Competition	129,388
Films	146,129
Object name	148,125
Songs	212,851
People	2,319,335
Total	4,814,852

The SNOW (Sparse Network Of Linear Separators) [11] utilizes the Winnow learning algorithm is used to tag the each word in the sentences.

Training of the SNOW tagger network proceeds as

follows: Each word in a sentence produces an example. Given a sentence, features are computed with respect to each word thereby producing positive examples for the part of speech the word labelled with, negative examples for the other parts of speech. In testing, this process is repeated, producing a test example for each word in a sentence.

Once an example is produced, it is then presented to the networks. Each of the sub-networks is evaluated and we select the one with the highest level of activation among the separators corresponding to the possible tags for the current tags. After every prediction, the tag output by the SNOW tagger for a word is used for labelling the word in the test data. Therefore, the features of the following words will depend on the output tags of the preceding words.

3.2.3 Finding events

Out of all the tags, we are mainly concentrating on noun tags such as NN (Noun Singular), NNP (Noun Plural), and NNPS (Proper Noun Singular) to find the events based on the subjects.

Examples:

Emerging **India** vs sinking **Pakistan**
Blackberry torch 9800 for sale

3.3 Filtering

3.3.1 Filtering events

It is the process of removing similar events for finding at the first instances. It not only filters the event but also gather the events which are similar to each other and represent that event with all the related videos under particular event.

3.3.2 Removing words

It is the intelligent process of filtering the events. WordNet is a lexical database for the English language. It groups English words (nouns, verbs, adverbs, adjectives) into set of synonyms called as synsets, provides short, general definitions and records the various semantic relations between these synonym sets. Thus the final events will be displayed along with the related videos

4. EXPERIMENTAL RESULTS

To evaluate the effectiveness of proposed work, we have collected and used our own dataset. Since all the existing dataset is using the videos which is constrained and only used to find particular set of events. Thus, we have collected the top ten events happened in India, 2013 by getting the result from Google Zeitgeist.

Table 2. Dataset used

Events	Number of videos
PM candidature of Narendra Modi	200
Blackberry sale	200
Dravid retirement	200
Air india news	200
Indian economy	200
Laptop distribution scheme	200
Karnataka election results	200

The above events are searched in the YouTube and gathered all the live videos along with its title. This live dataset is useful to test the level of effectiveness in the categorization process.

The outcome of the result consists of list of events in which the hierarchy based on the number of occurrence occurred in the final event result. The higher occurrence event is considered as a Level 1 hierarchy and level 2 hierarchy i.e sub event will be the lower number of occurrence.

In the proposed work, low level features are not included due to the poor classification result in the intra class event categorization process. The derived conclusion is represented as below:

Table 3. Derived conclusion

Parameters	Subject based classification	Object based classification
Events	Yes	Yes
Intraclass events	No	Yes

Yes – low level features can be used
 No- Low level features cannot be used

The above table shows the clear cut picture of the usage of low level features for classification. In order to prove the above conclusion, the sample proof is given below as:

Event name= Arsenal under 18



Event name = arsenal



Figure 2: Contradiction in classification for intra class event classification using SIFT

The above diagrams shows that the both features are merely same to each other. Features are identified by using SIFT (Scalable Invariant Feature Transform) technique which is most preferred technique used in the low level feature extraction. Thus, when the events are finding out by basing on the subjects, intra event classification contradiction will occur and lead to the poor classification if the low level features are used for classification.

5. CONCLUSION

The fast growth of the volume of videos in the internet becomes exponential which needs an urgent call to research in the video categorization to find out the effective major events. Web video categorization becomes more challenging by considering the unconstrained videos with the situation of not using low level features during the intra class events makes it more complicated. In the proposed work, three eyes namely POS Tagging, NER and WordNet enhance

process and detect the useful events from the titles and also the novel conclusion is derived that low level feature is not useful for classifying the intra class events based on the subject.

6. REFERENCES

- [1]<http://www.youtube.com/yt/press/statistics.html>
- [2]<http://nlp.stanford.edu/software/corenlp.shtml>
- [3]Collins English Dictionary, entry for "lemmatise"
- [4]L. Ratnov and D. Roth, Design Challenges and Misconceptions in Named Entity Recognition. CoNLL (2009)
- [5]G. A. Miller. Wordnet: A lexical database for english. (11):39-41.
- [6]Chengde Zhang, Xiao Wu, Mei-Ling Shyu and QiangPeng, " Adaptive Association Rule Mining for Web Video Event Classification", 2013 IEEE 14th International Conference on Information Reuse and Integration (IRI), page 618-625.
- [7] Y. Song, M. Zhao, J. Yagnik, and X. Wu. Taxonomic classification for web-based videos. In CVPR, 2010.
- [8] Z. Wang, M. Zhao, Y. Song, S. Kumar, and B. Li. Youtube-cat: Learning to categorize wild web videos. In CVPR, 2010.
- [9] <http://www.ranks.nl/resources/stopwords.html>
- [10]<http://cs.nyu.edu/grishman/jet/guide/PennPOS.html>
- [11]Roth and D. Zelenko, Part of Speech Tagging Using a Network of Linear Separators. Coling-Acl, The 17th International Conference on Computational Linguistics (1998) pp. 1136—1142
- [12]O. Duchenne, I. Laptev, J. Sivic, F. Bach, and J. Ponce. Automatic annotation of human actions in video. In Proc. of ICCV, 2009.
- [13]Laptev, M. Marszalek, C. Schmid, and B. Rozenfeld. Learning realistic human actions from movies. In Proc. of CVPR, 2008
- [14]M. Everingham, J. Sivic, and A. Zisserman. Hello! my name is... buffy automatic naming of characters in tv video. In Proc. of BMVC, 2006.
- [15]F. Smeaton, P. Over, and W. Kraaij. Evaluation campaigns and trecvid. In Proc. of ACM Workshop on Multimedia Information Retrieval, 2006
- [16]J. Yang, R. Yan, and A. G. Hauptmann. Cross-domain video concept detection using adaptive svms. In Proc. of ACM MM, 2007.
- [17] M. E. Sargin, H. Aradhye, P. J. Moreno, and M. Zhao. Audiovisual celebrity recognition in unconstrained web videos. In Proc. of ICASSP, 2009.

[18] J. Liu, J. Luo, and M. Shah. Recognizing realistic actions from videos. In Proc. of CVPR, 2009.

[19] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, "A novel ultrathin elevated channel low-temperature poly-Si TFT," IEEE Electron Device Lett., vol. 20, pp. 569-571, Nov. 1999

Assessment of the Adoption of Mobile Parking Management System in the Parking Industry in Nairobi County: A Case of Lulu East Africa

Jades Kalunda Muema
Information Technology
Dedan Kimathi University of
Technology
Kenya

Benjamin Kyambo
Faculty of Information Sciences
Kisii University
Kenya

Benjamin Kyambo
School of Information Sciences
KCA University
Kenya

Caroline Kirichu
Faculty of Information Sciences
Kisii University
Kenya

Kennedy Senagi
Information Technology
Dedan Kimathi University of Technology
Kenya

Abstract In the parking industry, savings can be made by providing alternative payment options for customers, such as cashless parking schemes, and providing the parking workforce with more sophisticated equipment. This technological change is expected to contribute to the development of more flexible, convenient and efficient parking services, increases revenue and customer satisfaction. This study assesses the adoption of mobile parking management information systems in the parking industry in Nairobi County with a special focus on Lulu East Africa. The researcher adopted a descriptive research design and the primary data was collected using questionnaires from the sample population of 60 comprising of workers and clients. This data was analyzed using SPSS version 21 and presented using distribution tables and graphs. The research findings indicated that Nairobi county and the parking industry were generally ready to adopt the mobile parking management system whose success is subject to a detailed feasibility study, although as with any technological adoption it is bound to face some barriers which can be overcome.

KEYWORDS: Adoption, mobile parking management system

1. INTRODUCTION

This sections presents the research background, problem statement, research objectives, significance of the research and the conceptual frame work.

1.1. Research Background

Often city officials presume that the providing of more parking spaces for the citizens means meeting the mobility needs. On the contrary, every car that is on the road needs a place to be parked: it is a key issue in almost all urban areas. (Bayless and Neelakantan, 2012). Evidence suggests that savings can be made by providing alternative payment options for customers, and providing the parking workforce with more sophisticated equipment. This technological change is expected to contribute to the development of more flexible and convenient parking services, raising public perceptions and expectations of the industry, whilst driving parking employer demand for the specialist skills to install, use and maintain equipment. (Walker, 2011).

Traffic flow, allocation and availability of parking space within the streets of Nairobi, is a major concern to for every motorist. Parking is managed by the former City Council of Nairobi and this exercise is currently a nightmare for the officials of the county government. (Kinyanjui, 2010).

Lulu East Africa Ltd, an expansive yard located on NSSF land on Kenyatta Avenue. The company has various approaches of car park management and allocation depending on the various client needs. Lulu East Africa seeks to improve its service delivery to gain competitive advantage, improve revenue collection and customer satisfaction. In this regard the management seeks to adopt technology to streamline its processes and provide convenient ways of engaging with the clients. Mobile management system is an option for them especially considering the wide use of mobile payment by majority of its target market, availability of the necessary technology and the convenience it brings. (Lulu East Africa, 2014).

1.2. Statement of the Problem

It is estimated that nearly 30% of urban congestion is created by drivers cruising for parking. Finding a parking space in most metropolitan areas, especially during the rush hours, is difficult for drivers. The difficulty arises from not knowing where the available spaces may be at that time; even if known, many vehicles may pursue very limited parking spaces to cause serious traffic congestion. (Wang and Wenbo, 2013).

Modern services are designed to get drivers to their ultimate destination without searching and the uncertainty related to cost, travel time, payment, and other practical considerations. In exchange for this convenience, smart parking service providers take advantage of consumers' willingness to pay a marginal mark-up over existing parking fees. It is this need coupled with the extensive use of mobile devices and the convenience of mobile payment that has given rise to the need for inclusion on mobile devices in the parking industry to improve performance and gain a competitive edge. (BPA, 2012; Wang and Wenbo, 2013). In line with this, there arises a challenge in the parking industry as far as adoption of mobile technology adoption is concerned since in Kenya it is rather new with limited studies and information especially in the case of Kenya; hence treated with caution. This study seeks to fill this gap and will present a critical discussion on the dynamics of the adoption of mobile parking management system and identifying potential gaps for future research.

1.3. Research Objectives

The main objective of this study is to assess the adoption of mobile parking management system in Nairobi County with a special focus on Lulu East Africa; its specific objectives are:

- i. To examine the current state of management of parking at Lulu East Africa
- ii. To investigate the successful adoption of mobile parking management information at Lulu East Africa.
- iii. To conduct a pre-feasibility analysis on the adoption of mobile parking management system at Lulu East Africa

1.4. Significance of the Study

The study will be significant to Lulu East Africa Ltd in providing valuable information in their efforts to adopt technology in their operations that will go a long way in not only helping them in decision making but also in anticipating the challenges that come with the adoption and recommendations of the measures on how to counter them. Investors in the parking industry will stand to benefit from this study to make informed decisions on mobile technology adoption and also benefit from its features. This study and consequent successful adoption of mobile parking management system will help make parking a non-event for drivers where it will be so easy and painless that it no longer becomes an issue. The study will add to the existing body of knowledge and it will increase awareness on not only the variables that impact successful adoption but also barriers to successful adoption of mobile management systems and their counter measures. In this regard further insights in this issue are of great importance to scholars in the parking industry and information technology in general. The stakeholders might use the research findings to evaluate the adoption of mobile technology and the extent to which it affects the parking industry in advent of improving its performance.

1.5. Conceptual Framework

Perceived usefulness refers to the degree to which a person believes that using a particular system would enhance his or her job performance. Perceived ease of use refers to the degree to which a person believes that using a particular system would be free of physical and mental effort. Social influence is defined as the degree to which an Individual perceives that important others believe he or she should use the new system. Facilitating conditions refer to the degree to which an individual believes that an organizational and technical infrastructure exists to support technology use.

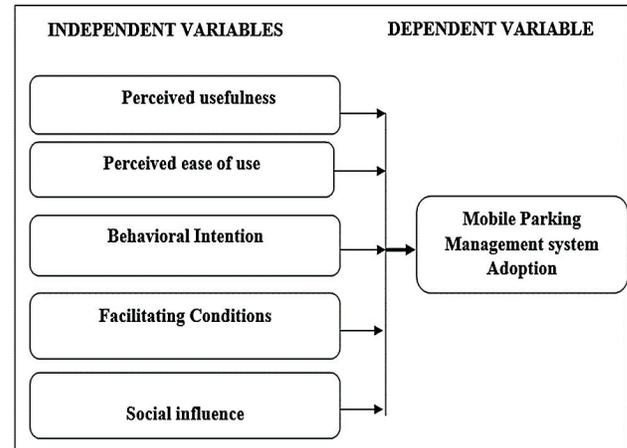


Figure 1: Conceptual framework

2. LITERATURE REVIEW

2.1. Features of Mobile Parking Information System

Mobile payments; this refers to a variety of financial transactions initiated with a mobile device.

Real-Time Parking Information; Mobile applications on smartphones are a significant area of growth for the parking industry, with an increasing number of customers relying on apps to get information on a car park's precise location, opening hours, tariff and facilities. (IPI, 2012; Flood et al, 2013). Automatic Number Plate Recognition (ANPR); the use of cameras to take timed photos of the vehicle (with close ups of the number plate) entering and leaving the car park. These photos are then used to calculate how long a vehicle has been parked.

2.2. Mobile Parking Management System Adoption Feasibility

During feasibility analysis we should consider that systems are adopted in businesses for two primary reasons: to take advantage of business opportunities and to solve business problems. A pre-feasibility study (PFS) is broadly defined as preparatory studies enabling funders to undertake a

successful feasibility study for a particular investment opportunity. (Hart and Gregor 2010).

This analysis can be categorised into technical, economic and operational feasibility. Technical feasibility which is carried out to determine whether the company has the capability, in terms of software, hardware, personnel and expertise, to handle the completion of the proposed solution. (Blanchard and Fabrycky, 2010). Economic feasibility; O'Brien and Marakas (2011), indicate that the purpose of the economic feasibility assessment is to determine the positive economic benefits to the organization that the proposed system will provide. It includes quantification and identification of all the benefits expected. This assessment typically involves a cost/benefits analysis. Operation feasibility; is a measure of how well a proposed system solves the problems, and takes advantage of the opportunities identified during scope definition and how it satisfies the requirements identified in the requirements analysis. It focuses on the willingness and ability of the management, employees, customers, suppliers, and others to operate, use, and support the proposed solution. (Hart and Gregor 2010).

2.3. Theoretical Framework

There have been several theoretical models, primarily developed from theories in psychology and sociology employed to explain technology acceptance and use. They include:

2.3.1. Technology Acceptance Model (TAM)

Technology Acceptance Model has been developed by Davis (1989) is one of the most popular research models to predict use and acceptance of information systems and technology by individual users. According to TAM individuals' intention to use technology determines the actual use of the application and attitudes toward technology affect the intention. Perceived usefulness and perceived ease of use factors are affected by various external variables such as level of education, gender, or organizational features such as training in computer use. (Davis et al., 1989; Davis and Venkatesh, 2004; Venkatesh et al., 2012).

2.3.2. Theory of Reasoned Actions (TRA)

Fishbein and Ajzen (1975) developed TRA to define the links between the beliefs, attitudes, norms, intentions, and behaviours of individuals. The theory assumes that a person's behaviour is determined by the person's behavioural intention to perform it, and the intention itself is determined by the person's attitudes and his or her subjective norms towards the behaviour. The subjective norm refers to "the person's perception that most people who are important to him think he should or should not perform the behaviour in question" (Fishbein and Ajzen, 1975, 302).

2.3.3. Unified Theory of Acceptance and Use of Technology (UTAUT)

The purpose of UTAUT is to explain a user's intentions to use ICT and the subsequent user behaviour. The model considers four constructs as direct determinants of user acceptance and usage behaviour, namely performance expectancy, effort expectancy, social influence, and facilitating conditions. The authors stated that UTAUT provides a tool for managers to assess the likelihood of success of technology introductions and to understand the drivers of acceptance in order to design interventions, which include, e.g., training or marketing. UTAUT focuses on users who may be less willing to adopt and use new systems. (Venkatesh et al., 2003; Yu, 2012).

3. RESEARCH METHODOLOGY

This study will adopt a descriptive research design. A sample size of 60 respondents composed of staff and customers of Lulu East Africa were selected. The researcher used questionnaires that were formulated based on the research objectives to collect primary data that was later analysed using SPSS. Content analysis was used to analyse the data while descriptive analysis such as mean, frequencies and percentages were used to analyse the data. Also the inferential statistics such as correlation models and Chi square will be used. Tables and charts were used for data presentation.

4. DATA ANALYSIS AND PRESENTATIONS

As indicated in table 4.1, the respondents were split between the customers and Lulu East Africa staff. The customers made up 50% of the respondents, normal employees 23%, 4 supervisors, 1 manager and two directors.

Table 1: Organizational position

	Frequency	Percentage
Director	2	3.3
Manager	1	1.7
Supervisor	4	6.7
Employee	23	38.3
Customer	30	50
Total	60	100.00

4.1. Access to Mobile Devices

According to the survey, 86.7% of the respondents owned a phone, tablet and or a PDA; all of them had access to conventional forms of mobile payment while 90% had access to alternative forms of mobile payment such as mobile banking as indicated in figure 1. This is key to the mobile paring adoption since the entire concept is based on mobile device with mobile payment being one of the key functionalities of the system.

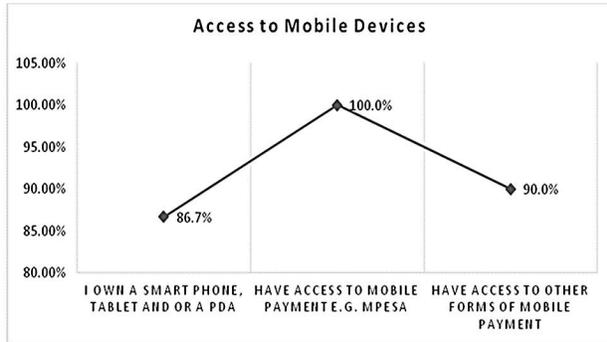


Figure 1: Access to mobile devices

4.2. Current State of Management of Parking

Table 2: Process automation

	Frequency	Percentage
Some activities are automated and others are manual	47	78.3
Don't Know	13	21.7
Total	60	100.00

The survey determined as shown in table 2 above that majority of the respondents 78.3% had some knowledge on the operation of Lulu East Africa, and that their activities were a combination of manual and computerized.

4.3. Desirable Features in Parking Management System

The findings established that provision of real time parking information, according to the survey was the most desirable feature of the parking management system according to 83.3% of the respondents, this was followed by mobile payment with 71.7%. Automatic Number Plate recognition was the least important feature to the respondents with only 43% feeling that this was an important feature. A considerable number of respondents 68.3% were of the opinion that booking and reservation was also a feature they would like integrated in the mobile parking management system as shown in figure 2.

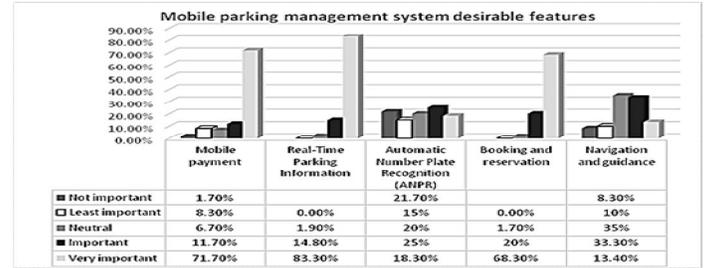
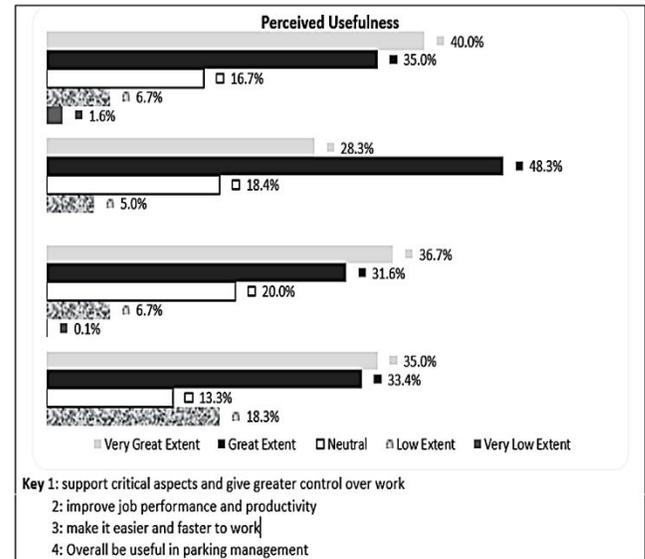


Figure 2: Desirable features in mobile parking management system

4.3.1. Perceived Usefulness

The survey response indicated that majority of the respondents felt that the mobile parking management system would to a great extent be overall useful in parking management, improvement of job performance was second. In essence, there was a consensus that the mobile parking management system would impact all aspects of parking management as indicated in figure 3 below.

Figure 3: Impact of adoption of mobile parking management system



4.3.2. Perceived Ease of Use

In regards to the respondents perception on the ease of use of the new mobile system, the survey showed that 63.3% of the respondents felt that the mobile parking management system would be both easy for them to learn and operate and easy for them to get it to do what they wanted to do. Also, a significant number of respondents 63.4% were of the opinion that generally they would find the system easy to use as indicated in figure 4. This response could be attributed to the fact that majority of them have access to mobile devices and mobile forms of payment and as such have experience with them. The constant 3% who were reluctant about the system could be

deduced to be the ones without access to smart devices and lower education levels.

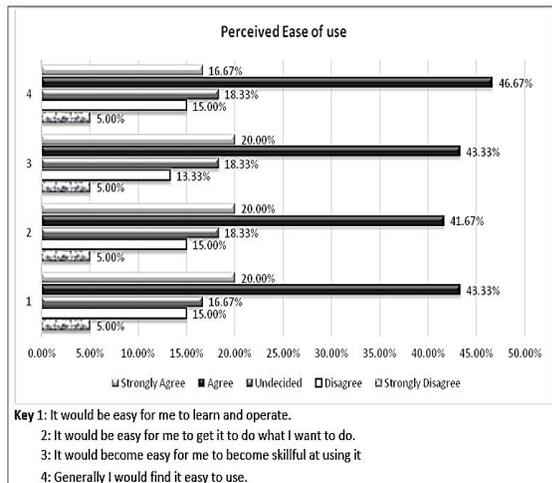


Figure 4: Perceived ease of use

4.3.3. Behavioural Intention

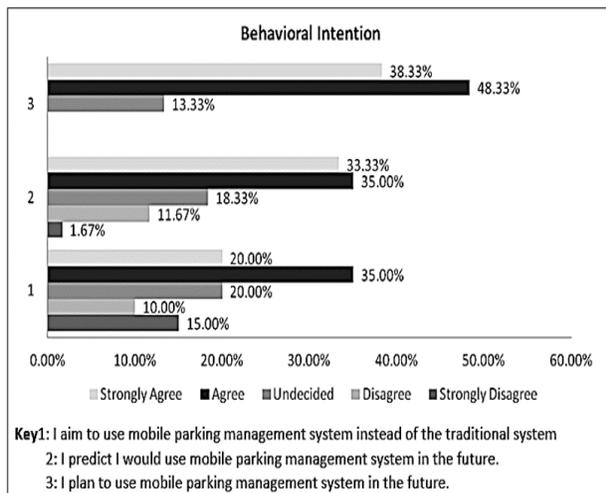


Figure 5: Behavioural Intention

The survey shows that generally the respondents were intended or were willing to adopt mobile parking management system (and any mobile technology) at some time in future (86.7%); 88.3% predicted they would use mobile parking management system in future while 55% aim to use mobile parking management system instead of the traditional system as shown in figure 5.

4.3.4. Social Influence

In regards to social influence of the respondents 65% agreed that that management system as depicted in figure 6.

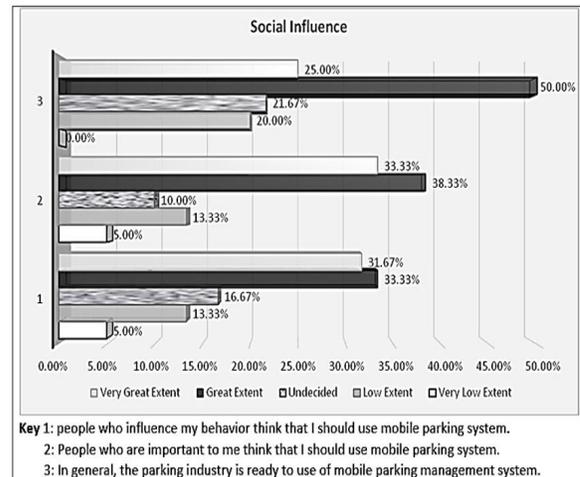


Figure 6: Social Influence

4.3.5. Facilitating Conditions

According to the survey majority of the respondents 91.7%, agreed that their living and working environment supports them to use mobile parking management system and also that using mobile parking management system is compatible with their life. In regards to availability of sufficient technological infrastructure to support mobile parking management system, 71% of the respondents felt that this was the case; however, as indicated in figure 7 the respondents were not so sure about the availability of help when they get problems in using mobile parking management system as only 40% agreed to this.

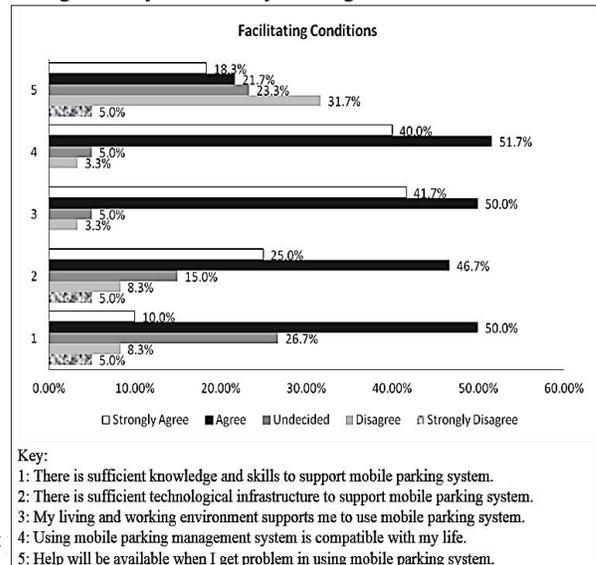


Figure 7: Facilitating conditions

4.4. Pre-feasibility Study

4.4.1. Technical Feasibility

Lulu East Africa according to the survey has the technical capability to facilitate to fruition the adoption of mobile parking management system. According to 83% of the respondents they are able to acquire and or lease all the hardware and network resources required. All the respondents agreed that the company was able to restructure their organization and its processes to accommodate the changes that come with the adoption of the new system and get the necessary technical support required for the system. As indicated in figure 8, 97% of the staff felt that software resources such as mobile apps would also be acquired as required by the mobile management system.

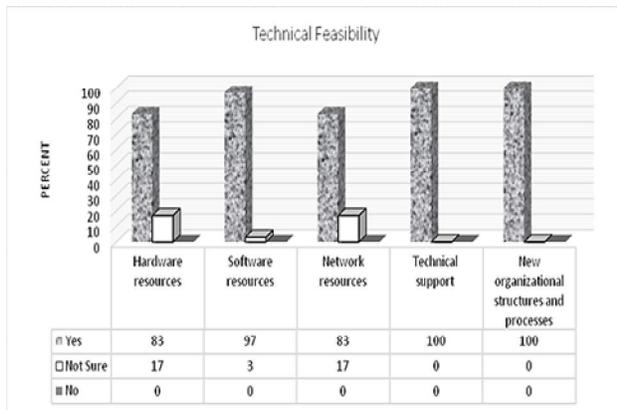


Figure 8: Technical feasibility

4.4.2. Economic Feasibility

The survey indicated as shown in figure 9 that majority of the respondents 90% agreed that adopting the system is a good business decision after a cost benefit analysis; 65% felt that the company stood to gain competitive advantage hence increase revenue as a result of the mobile management system adoption while 64% were of the opinion that generally benefits of the mobile system will exceed the costs of adopting and operating it. The initial investment on the system adoption would be high but there was expectation of Return on Investment (ROI) in reasonable time according to 58% of the respondents.

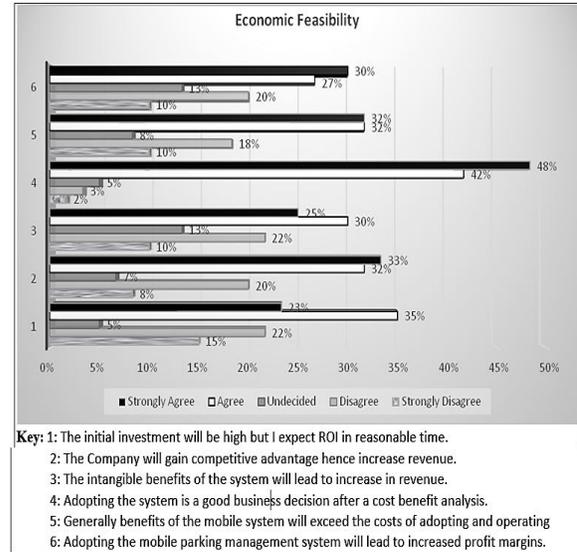


Figure 9: Economic feasibility

4.4.3. Functional/Operational Feasibility

Majority of the respondents 88 and 87% respectively as shown in figure 4.11 were of the opinion that adopting mobile parking management system will maximize efficiency of parking payment system and that of parking facilities and services; 81% agreed that the adoption would improve customer experience and satisfaction. However a lower percentage (46%) were convinced that the mobile parking system would will share real time parking information with other agencies and stakeholders.

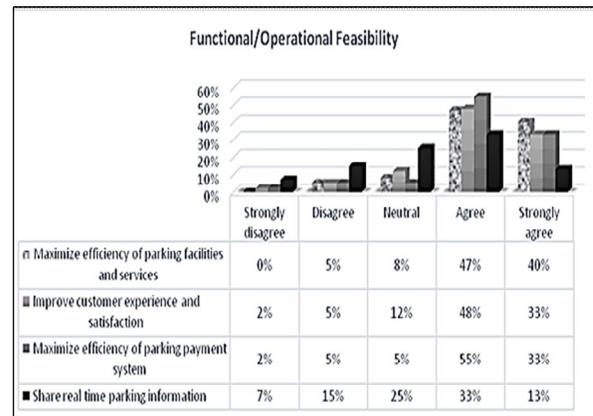


Figure 4.11: Functional/Operational feasibility

5. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1. Summary of the Findings

5.1.1. Demographic information

The respondents consisted of Lulu East Africa and their clients. The staff was composed of two directors, a manager, four supervisors, other employees and the rest were other employees. Access to mobile devices is important a crucial factor in the implementation of the mobile parking management system not only for Lulu East Africa but also the employees. This is however not an issue in this study since majority of the respondents not only owned smart phones, tablets and or PDAs, but also access to conventional modes of mobile payment (such as Mpesa, Orange money and Airtel money) and other alternative forms of mobile payment.

5.1.2. Current State of Management of Parking at Lulu East Africa

Before improving the service delivery and the way parking management is done, it is important to analyse the current state of affairs. In regard the study determined that daily activities were conducted using a combination of comprised (automated) and manual methods. Notably, some clients had no idea which methods were used. Generally the satisfaction level of the state of affairs at Lulu were below average from both the clients and employees perspective with parking allocation and booking rating highest followed by payment processing. Communication with clients was rated poorly. It is worth noting here that majority of the respondents were neutral about the most of the categories of service delivery implying that changes needed to be done to ensure client satisfaction, gain a competitive edge and improve service delivery.

5.1.3. Analysis of Adoption of Mobile Parking Management System

Mobile parking management systems desirable features ensure efficiency and effectiveness and include; provision of real time parking information and mobile payment services respectively were the most sought after functionalities by majority of the respondents. In terms of perceived usefulness, the mobile parking management system would make it easier and faster to work, improve job performance and productivity, support critical aspects and give greater control over work. The system is perceived to be both easy to learn and operate, and easy for them to get it to do what they want to do. This perception could be attributed to the accessibility to mobile devices by the majority and the wide use of mobile payment.

There was a general intention, willingness and ability to use mobile parking management system in future in fact more than half of the respondents aim to replace the traditional system with the new management system. In terms of facilitating conditions that is, the individuals belief that an organizational and technical infrastructure exists to support technology use; the living and working environment supports the use of mobile parking management system and also using mobile parking management system is compatible with the life of majority of the respondents; there is also sufficient technological infrastructure to support mobile parking management system however the

respondents were sceptical about the availability of help when they get problems in using mobile parking management system.

5.1.4. Pre-feasibility Study

This was the initial feasibility study conducted to as a guide to determine whether or not the pursuit for the subject technological adoption was worth further exploration. It also provides a basis for a detailed feasibility study. Lulu East Africa according to the survey has the technical capability to facilitate to fruition the adoption of mobile parking management system. It was established by the study that adopting the system would be a good business decision after a cost benefit analysis ad that that the company stood to gain competitive advantage leading to increased revenue as a result of the mobile management system adoption. The research established that adopting mobile parking management system will maximize efficiency of parking payment system and that of parking facilities and services and improve customer experience and satisfaction.

5.2. Conclusions

The management of parking at Lulu East Africa is currently a combination of manual and computerised, predominantly manual. Most of its publics are not satisfied with the service delivery and job performance. There is a mutual agreement that the management could use some improvements. There are numerous desired features of the mobile parking system that the respondents would like to be implemented including mobile payment, booking and reservation, and provision of real-time parking information. Generally Lulu East Africa is ready to adopt the mobile parking management system and its clients are willing and able to follow suit; with the right implementation approach and manipulation of the adoption variables as informed by this research there are high chances of success. A proper detailed personalised feasibility study however needs to be conducted by the respective parking service provider to ensure the fine details are considered to make informed decision.

5.3. Recommendations

Based on the research findings the researcher recommends that Lulu East Africa adopts the mobile parking management system with all the desired features. It is however imperative that they conduct an elaborate and detailed feasibility study so as to be able to make informed decisions. The new system and its features should be made public (consider marketing), the respondents concern for support in case they need help should also be factored and catered for. There would be need for training and support for both clients and staff especially in the initial stages of adoption.

Considering the fact that the respondents exhibited intention to use the new technology and the availability of facilitating environment, proactive efforts should be made

to influence this factors such as proper marketing, education training and awareness. There is need to work on the weakness brought forth by this study in service delivery such as communicating with clients, and navigation and guidance. There will also be need for Lulu to work with stakeholders and partners to such as mobile network service providers and regulatory bodies ensure success and compliance.

5.4. Suggestions for Further Research

Further research is needed not only in the parking industry but also mobile technology especially with the wide use of mobile devices in the country. More studies need to be conducted on the adoption of an integrated mobile parking information system that works for both the private and public sector.

6. REFERENCES

1. Abu-dalbouh H., (2013). A questionnaire approach based on the technology acceptance model for mobile Tracking on patient progress applications. *Journal of Computer Science* 9 (6): 763-770, 2013 ISSN: 1549-3636. Science Publications
2. Andrews, Robert J. (2011) IBISWorld Industry Report 81293: *Parking Lots & Garages in the US*. Santa Monica: IBISWorld.
3. Ajzen, I. (1991). *The Theory of Planned Behaviour*, Organizational Behaviour and Human Decision Processes 50, pp. 179–211.
4. Ajzen I. (2010). *The Theory of Planned Behaviour*, <http://www.people.umass.edu/ajzen/tpb.html> Accessed August 2014.
5. Ajzen I. and M. Fishbein (1980) *Understanding Attitudes and Predicting Social Behaviour* Englewood Cliffs, NJ: Prentice-Hall, Inc.
6. Babbie E., (2012). *The Practice of Social Research* (13th edition). Belmont, Calif; Wadsworth Cengage Learning.
7. Bayless S. H. and Neelakantan R. (2012). Smart Parking and the Connected Consumer Opportunities for Facility Operators and Municipalities. The Intelligent Transportation Society of America (ITS America) research.
8. Blanchard S. and Fabrycky W. (2010). *Systems Engineering & Analysis*. 5th ed Prentice Hall International Series in Industrial & Systems Engineering.
9. British Parking Association, (BPA). (2012). *Technology in the Parking industry*. Parking practice notes Revised March 2012. British Parking Association.. West Sussex
10. Cooper R., Schindler P. (2011). *Business Research Methods*. McGraw-Hill/Irwin
11. Chwelos, P. et al. (2001). *Research Report: Empirical Test of an EDI Adoption Model*. Information Systems Research (12)3, pp. 304–321.
12. Davis F. (1989). *Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology*. MIS Quarterly (13)3, pp. 319–342.
13. Davis F. et al (1989). *User Acceptance of Computer Technology: A Comparison of Two Theoretical Models*, Management Science (35)8, pp.982–1003.
14. Deci, E. and Ryan M. (2000). *The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior*, Psychological Inquiry (11)4, pp. 227–268.
15. Ervasti, M. and Helaakoski, H. (2010). *Case study of application-based mobile service acceptance and development in Finland*. Int. J. Information Technology and Management, 9 (3), 243-259
16. Flood et al, (2013). *Trends in Mobile Payments in Developing and Advanced Economies*. Reserve bank of Australia. Australia.
17. Griffin, R. and Pustay, M. (2010). *International Business*. New Jersey: Pearson Education Ltd.
18. Hart N. and, Gregor S. (2010). *Information Systems Foundations: Theory Building in Information Systems*. The Australian National University (ANU). ANU E Press
19. International Parking Institute- IPI. (2011) *Emerging Trends in Parking: Report on a survey conducted by the International Parking Institute*. 2011. (Accessed August, 2014).
20. Turner M., et al. (2010). *Does the Technology Acceptance Model Predict Actual Use? A Systematic Literature Review*. Information and Software Technology 52, pp. 463–479.
21. Kinyanjui E., (2010). *Mobile Phone-Based Parking System*. School of Computing and Informatics. University of Nairobi.
22. Korpelainen E., (2011). *Theories of ICT System Implementation and Adoption – A Critical*

Review Working paper. Aalto University, Helsinki.

23. O'Brien A., and Marakas M. (2011). *Developing Business/IT Solutions In Management Information Systems* (pp. 488-489). New York, NY: McGraw-Hill/Irwin.
24. Pénicaud C (2013). *State of the Industry: Results from the 2012 Global Mobile Money Adoption Survey*, GSMA. Available at <http://www.gsma.com/mobilefordevelopment/wp-content/uploads/2013/02/MMU_State_of_industry.pdf>.
25. Robson C., (2011). *Real World Research* (3rd edition). John Wiley & Sons. London.
26. Silverman, D. (2011). *Interpreting qualitative data*. London: SAGE Publications Ltd.
27. Walker A., (2011). *Practical benefits of a hosted parking management solution*. Presentation at the International Parking Institute's Conference & Expo.
28. Wang A. and Wenbo, H., (2013). *Parking Management Best Practices*. Chicago: American Planning Association.
29. Yousafzai S. et al. (2007). *Technology Acceptance: A Meta-Analysis of the TAM: Part 1*. *Journal of Modelling in Management* (2) 3, pp. 251–280.
30. Yu C., (2012). *Factors affecting individuals to adopt mobile banking: empirical evidence from the UTAUT model*. *Journal of Electronic Commerce Research*, VOL 13, NO 2, 2012.
31. Zhu K., et al. (2004). *Information Technology Payoff in EBusiness Environments: An International Perspective on Value Creation of E-Business in the Financial Services Industry.*, *Journal of Management Information Systems* (21)1, pp.17–54.

A Metamodel and Graphical Syntax for NS-2 Programing

Elham Azadi Marand

Department of Computer Engineering,
Shabestar Branch, Islamic Azad University,
Shabestar, Iran.

Moharram Challenger

Department of Computer Engineering,
Shabestar Branch, Islamic Azad University,
Shabestar, Iran.

Abstract: One of the most important issues, around the world, which manufacturers pay special attention to is to promote their activities in order to be able to give high quality products or services. Perhaps the first advice to achieve this goal is simulation idea. Therefore, simulation software packages with different properties have been made available. One of the most applicable simulators is NS-2 which suffers from internal complexity. On the other hand, Domain Specific Modeling Languages can make an abstraction level by which we can overcome the complexity of NS-2, increase the production speed, and promote efficiency. So, in this paper, we introduce a Domain Specific Metamodel for NS-2. This new metamodel paves the ways for introducing abstract syntax and modeling language syntax. In addition, created syntax in Domain Specific Modeling Language is supported by a graphical modeling tool.

Keywords: Metamodel, Graphical Concrete Syntax, Model Driven Engineering, Network Simulation, NS-2 Programing.

1. INTRODUCTION

From physics to biology, to weather forecast, to prediction of a new processor design, researchers make models or, in fact, simulate the aspects of the science, variety of phenomena or theoretical scenario which they cannot analyze, produce or scientifically observe. Therefore, researchers rely highly on simulators. This is also generalized to computer networks. Network simulation, as an important technology in modern age, is a technique which simulates network behavior with calculating pro actions among existent networks, and uses mathematical formulas, receives observations and uses variety of tools such as NS-2 [6,7], NS-3[8], OMNET++ [9], JIST [5], PeerSim [10] and OPNET [4] which may help simulation and designing process [4,5].

NS-2 is one of the most important simulators. Development of NS began in 1989. From the time on, it has gradually been improved. NS-1, under supporting of American Defend Minister – central research and development administration, introduced in 1995 by Berkley Laboratory. In 1996, the second version of NS which was originated over Keshav, S. early work, introduced in university of California. It had plenty of major architectural changing, and later on, was known as NS-2. The development of NS-2 was supported by DARPA VINT Project from 1997 till 2000, and SAMANDARPA and NSF CONSER from 2000 to 2004. Nowadays, NS-2 is one of the most popular simulation tools in source free networks which are based on objective and discrete event simulator. NS-2 also supports TCP simulator, navigators and multi-purpose protocols in different networks (wired and wireless). They can be used for distributer and parallel simulators as well. NS-2 simulator has been designed and performed over C++ programmer and OTCL manual programmer.

Some of main advantages of NS-2 are as below:

- Most of the protocols have been done on it in advance.
- It has countless of models in hand.
- It is very popular and supported by many communities.
- It is capable to be used in parallel and wireless simulators.
- It is source less and free.

Despite the advantages, NS-2 has also some defects. It takes a long time to get familiar with it, and its source code

and instructions are mall documented, and its documentations are not easy for amateurs, difficulty in rapid evaluation of a simple idea, and being risky in manual code of Tel are examples of NS-2 problems [11,12,6,7]. Considering the latest problems, NS-2 needs to reduce its complexity.

Model-driven engineering (MDE) [14] is one way of promoting abstract level in software development the aim of which is changing programing focus from code to model. In this method, models are created with required details; then, codes are generated automatically. In some cases, full models with executable tasks can be made. This increases productivity, reuse (through reusing standardized model), simplifying design process, and promoting team work capability on the system. It must be considered that a modeling paradigm is effective when its model is understandable from a user's view that is familiar with that domain, enabling the model to act as a basis for executing systems [15].

To work with Domain-specific Languages (DSL) [1,16] a programmer designs his/her own program in the language specific for the applied domain and the language produces the architectural code for him/her. In this way, instead of dealing with coding details, the programmer works with a model of the program. Also, instead of working with a general purpose language (GPL), he/she works with a specific language of the domain; thus, key words, relations, and concepts will be very close to regarded application's terms. One of DSL types is Domain-specific Modeling Language (DSML) [20, 23], in which graphical elements and models are used for designing programs, giving a better insight about software elements and relations. This methodology is applied in many other cases. In this paper, NS-2 programs were selected as domain to apply DSML-based environments for increasing abstract level and reducing production complexities of NS-2 programs.

For this reason, we have drawn NS-2 concepts and their relations in the format of a metamodel which shows an abstract and concrete syntax which pave a way for generating a domain specific language.

The rest of the paper as below: Section 2 discusses metamodel and abstract syntax. Section 3 discusses graphical concrete syntax, Section 4 talks about sample model and its explanation, Section 5 explains related work, and Section 6 concludes the paper.

2. METAMODEL

For many years, metamodels have been made along side with different application domains, and Objective Management Group (OMG) has played important role in utilizing them in a standardized way. With beginning of Model Driven Architecture (MDA) and increasing the need for standardization, the amount of metamodel applications has been highly increased. Whenever there is a need for defining a piece of language, abstract syntax comes on the scene and appears in Ecore format. In fact, an abstract syntax of a language explains key concepts of that language and shows how these concepts combine to make new models. In addition, abstract syntax includes definitions and relations among meta-elements. These relations exist among concepts and the way they combine with each other [20]. It can be said that, abstract syntax of a language is independent from semantic and concrete syntax and works with structure of the concepts, regardless the meaning or usage of those concepts [17]. In this part of the paper, we have tried to deal with the topology and NAM of NS-2 programs and draw on their relations among metamodels which show the abstract syntax. To have a clear understanding and efficient use, we draw it under several meta-elements as in figures 1 and 2, and will explain them later on.

2.1 Topology Viewpoint

Before any activity, there should be a simulator object for NS-2 program in every Tel document. In this case, it is necessary to have a simulator meta-element in order to install simulator object for ns variable. Since the formation of a network is by nodes and relations among them, there is a *hasNode* and *hasLink* relation among simulator meta-element, and Node and Link meta-elements. Node meta-element is anything that connects to network. According to this definition, computers, cell phones, printers and routers are all defined as nodes in the

program. Name property differentiates nodes from each other. To send a package, there should be a tool in every network formation, so every node has countless of tools. In proposed metamodel, we supposed *connectAgent* relation between Node and Agent meta-elements. Agent meta-element is a protocol which runs inside nodes, and determines nodes' behavior in sending and receiving packages. We have given a name for each agent to differentiate tools from each other. We also defined the sort of protocol in metamodel which is shown by TCP and UDP meta-elements connected to Agent meta-element with inherit relation. TCP is a reliable dynamic protocol in controlling crowd. TCP also gets benefit from corrected packages produced in destination part and checks whether these packages have appropriately received or not. Lost packages are interpreted as crowd. In these meta-elements senders and receivers should be given. AgentTCP and AgentTCPSink with *hasAgentTCP* and *hasAgentTCPSink* relation are meta-elements which have been connected with TCP and show receiver and sender agents. Both of these meta-elements have connected to Node meta-element with *attachAgent* link and show which node is receiver and which one is sender.

UDP meta-element is a part of the Internet protocol. By using this meta-element, data are sent as datagram format. Like TCP, UDP also has sender and receiver agents named AgentUDP and AgentNull. Both of these meta-elements connect UDP meta-element through *hasAgentUDP* and *hasAgentNull* and show receivers and senders. These meta-elements connect Node through *attachAgent* and show which node is sender and which one is receiver.

As it was mentioned earlier, the time we define nodes in simulator program, we need to define a connector object for them as well. In this case the existence of Link meta-element is necessary. There should be a *connecting* link between Link and Node meta-elements to connect them. Name property is supposed for each link to differentiate them.

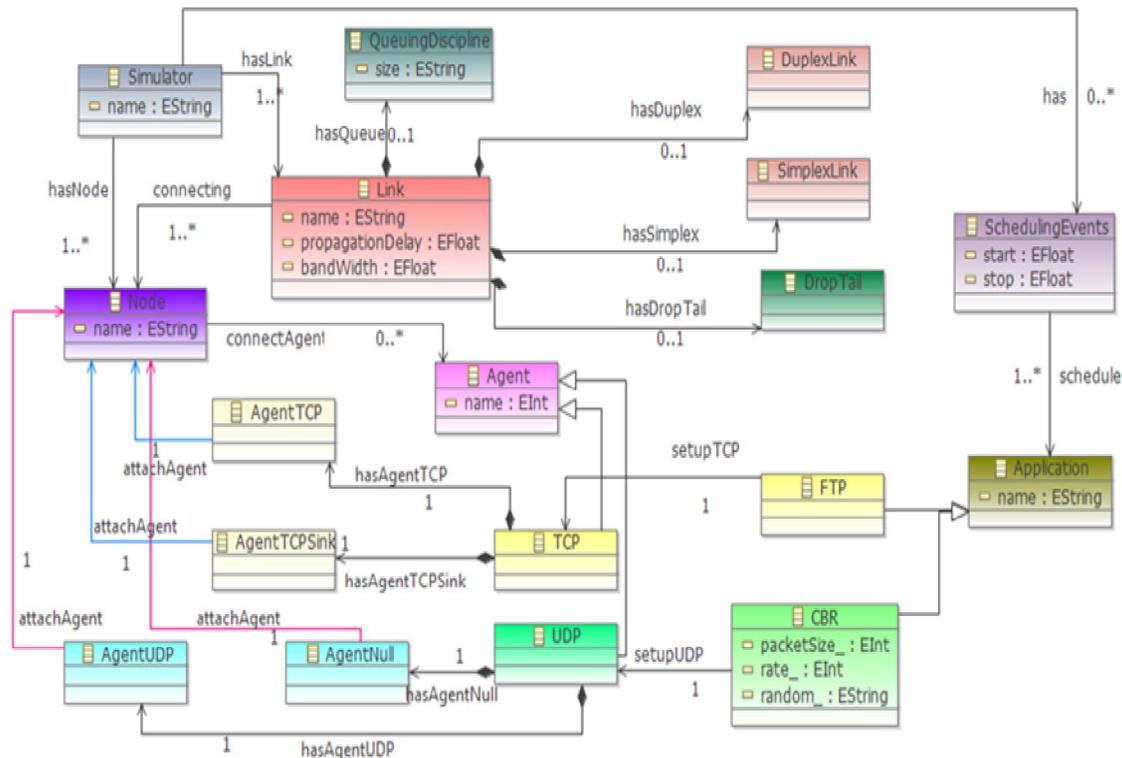


Figure 1. Metamodel of Topology Viewpoint

Each link also has own band width property as bandwidth and distribution delay as propagationDelay which are shown in proposed metamodel. Any link may have duplex or simplex relation with nodes. For this reason, DuplexLink and SimplexLink meta-elements have been connected to Link meta-element using *hasDuplex* and *hasSimplex* relations. It should be said that in NS-2 programs, queues are performed as a part of links; therefore, there is a *hasQueue* relation between Link and QueueingDiscipline meta-elements. QueueingDiscipline meta-element sends and receives buffers in each node. On the other hand, size property has been defined to show the length of the queue and limit the size of the link.

We also can define the type of the queue for every Link meta-element. In proposed metamodel, we supposed DropTail meta-element whose structure enables it to drop the latest package by the time the queue capacity finishes. In addition to all items mentioned above, every simulator has an event timer which schedules the simulations. In Proposed metamodel, we name it as schedulingEvents and suppose *has* link for it which includes *start* and *stops* points to show commence and end points. Since NS is a discrete event simulator, TCLs are being defined by the time events happen. In this case, this meta-element determines start and stop points of applications and connects to *SchedulingEvents* through schedule meta-element. It should be noted that *Application* meta-element includes FTP meta-element (this is the first attempt to make standards in file transportation in networks based on TCP protocols) and CBR meta-element (this meta-element sends bits with fixed rate and has PacketSize, rate and random properties). In these meta-elements, inherit concept is attendant, in a way that, both FTP and CBR meta-elements acquire Application meta-element characteristics. This issue is important because it supports hierocracy concepts.

It should be said that, FTP meta-element, uniquely, uses TCP protocol (it never uses UDP protocol). There is *setupTCP* between TCP elements that is responsible for scheduling.

2.2 NAM Viewpoint

As it was mentioned before, to make programs in NS-2 there should be an object simulator for manual TCL. In this case, the existence of a simulator meta-element in turning an object to ns is necessary. Since NS-2 simulator gets help of NAM in graphical simulation, supposing a NAM named meta-element makes a graphical picture out of network topology.

Away from all these items, every network formation is made up of nodes, links and events that cause sending and receiving process. In suggestive meta-elements, we have supposed (Agent, Link and node) meta-elements as concepts for NS-2, and connected these meta-elements to simulator meta-element through a link.

To distinguish Node meta-elements from each other, we supposed a *name* property. In addition, if in NS-2 programs the place and shape of the nodes have not been statically defined, in that case, we can use Position and Shape meta-elements to decide on the location and shape of the nodes. We can also decide on node color using ColorNode meta-element. To mark nodes to make difference among them, we can use Markup meta-element, and to eliminate the marks we can use MarkDelete meta-elements. We can also use AddLabel meta-element to put label on the nodes. All these meta-elements connect to Node meta-element through separate links.

To distinguish Link meta-elements from each other, we supposed a *name* property. To color Link meta-element, we can use ColorLink meta-element. Like Node and Link meta-elements, to distinguish Agent meta-elements from each other, we supposed a name property. We can also use ColorFlow meta-element to give color to flows.

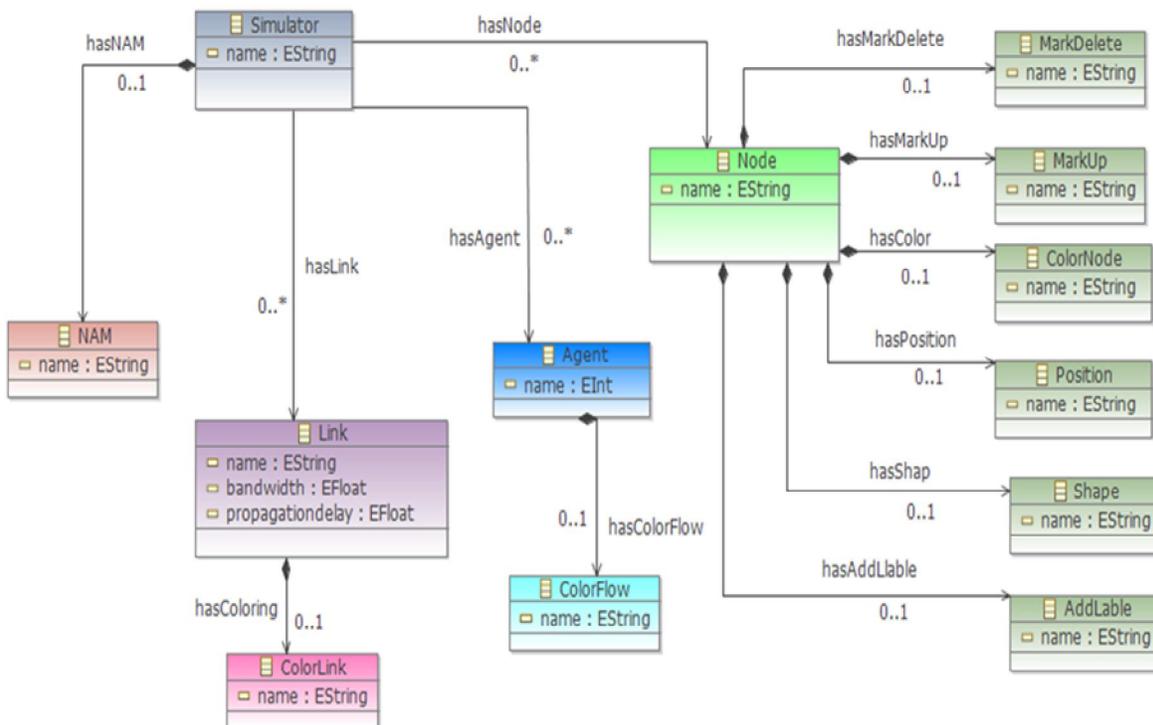


Figure 2. Metamodel of NAM Viewpoint

3. GRAPHICAL CONCRETE SYNTAX

All languages provide notations that describe representation and building blocks of programs in those languages. These notations are known as concrete syntax of those languages. Concrete syntax has basically two main types: visual and text syntax [13], used by different languages [17]. This paper suggests a graphic concrete syntax for NS-2 programing.

To achieve concrete syntax, first a metamodel is defined for NS-2 programs in Ecore. Then to check its models' validity, we produce an EMF based tool and chose a graphical symbol for each concept and its relations. Table 1 and Table 2 show graphical symbols for concrete syntax in Topology Viewpoint, and NAM Viewpoint for NS-2 programs.

TABLE 1. SOME OF THE CONCEPTS AND NOTATIONS FOR THE CONCRETE SYNTAX OF TOPOLOGY VIEWPOINT

Concept	Symbol	Concept	Symbol
Simulator		Node	
Link		Application	
DuplexLink		TCP	
SimplexLink		UDP	
QueuingDiscipline		FTP	
Agent		CBR	
DropTail		AgentTCP	
CBQ		AgentTCPSink	
RED		AgentUDP	
SchedulingEvents		AgentNull	

TABLE 2. SOME OF THE CONCEPTS AND NOTATIONS FOR THE CONCRETE SYNTAX OF NAM VIEWPOINT

Concept	Symbol	Concept	Symbol
Simulator		Node	
NAM		Shape	
Link		AddLabel	
ColorLink		Position	
Monitor		ColorNode	
Agent		MarkUp	
ColorFlow		MarkDelete	

After selecting graphical symbols, eclipse modeling framework was used for relating domain concept, Ecore format, and their symbols. Resultant structure is a graphic editor developed according to concrete syntax of NS-2 programs.

4. CASE STUDY

Since there are plenty of samples in NS-2 programs, in this chapter, we have selected simulation scenario in formation dynamic networks using nodes and relations among them, and the way we send information from a node to node, and the observing way of a queue and running simulation *nam* scripts, and since it is one of the most important issues in network formation and is popular in computer world, and more importantly, it has enough complexity in code generation.

We have drawn the methodology over these samples in figures 3 and 4 with graphical editor. With a quick view on research sources about dynamic network scenario, we can easily understand that circular topology of this case study has made more nodes than the others which can hold them in a Tel presentation. With using ring links with reciprocal connection among nodes in DropTail type, we can connect circular topology nodes in a way that connects from one node to the other, except the latest node that joins the first node. To transfer information among nodes, there is a factor defined to recognize sender and receivers in the nodes. It also helps transfer information from one node to the other in a way that data traffic uses the shortest way. It should be said that in topology scenario for transferring information, we have used UDP.

In addition, in this study case, there is the possibility for visualizing with the help of *nam* program which also helps color nodes, shape them, color links among nodes, mark nodes, add labels to nodes, and color the flow of information transferring among nodes [22].

Needless to say that, illustrated models with this tool will help change models to executable codes, as the next work in this paper, in NS-2 program.

As it can be seen in this case study scenario of Figure 3, circular topology formation includes seven nodes (N0, N1, N2, N3, N4, N5 and N6).

To connect all the nodes to each other, except the last node that connects to the first one, we have used reciprocal links with band width of 1 mega bytes per second and 10 mille second of delay. Every node also gets benefit of a queue with DropTail (dropping from the head part) property.

A UDP protocol agent is connected to N0. This agent is in relation with Null server in N3. N3 is connected with UDP agent that is CBR traffic generator, as well. CBR traffic generator has been formed in a way that can make 500 kilo bytes per second with the rate of 0.005. CBR generator starts in 0.5 and stops after 4.5 seconds after simulation.

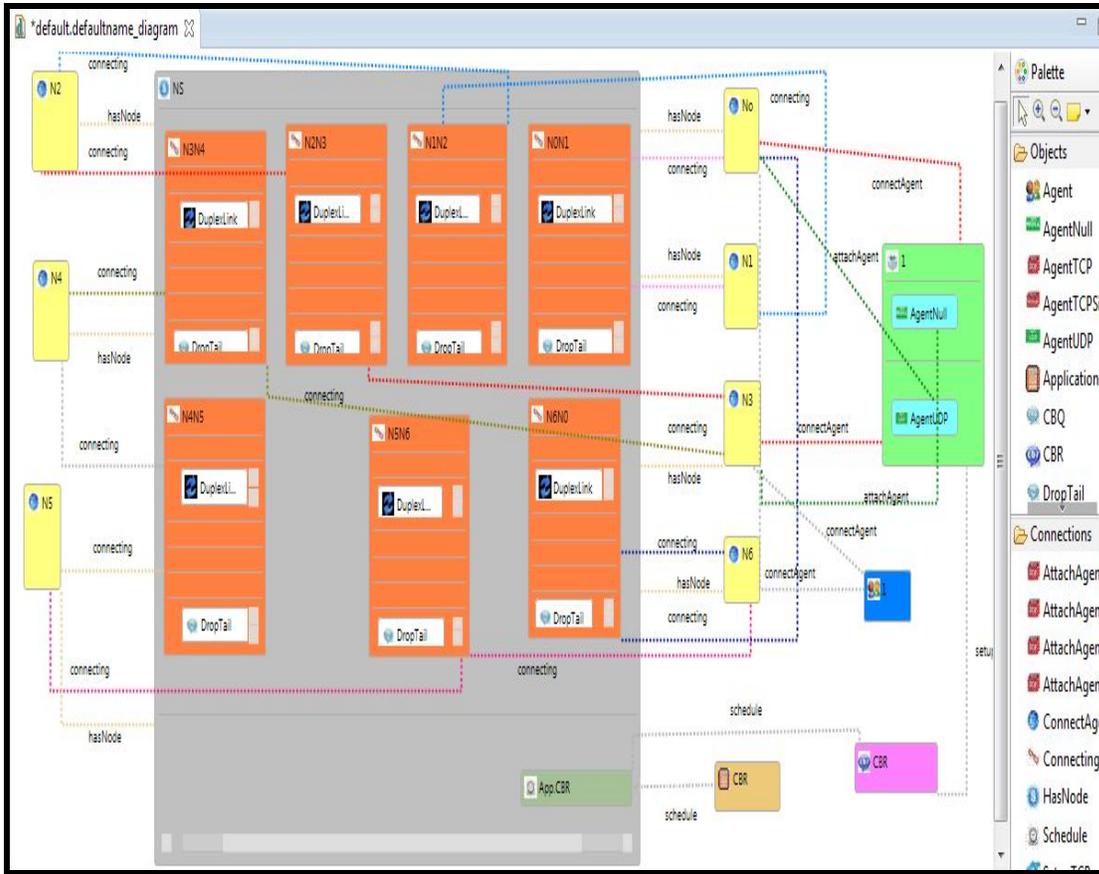


Figure 3. Topology viewpoint modeling



Figure 4. NAM viewpoint modeling

In this scenario of Figure 4, NS simulator has NAM which supplies the visual picture of network topology. In addition, it also has seven nodes and relative links among them, and a factor that causes data be sent from sender to receiver node. By using *nam* program, we can specify separate property for each node, link and agent. In this sample, we have supposed the circle as the shape and red as the color of the nodes. To differentiate N0 from N3, by the time a node starts sending a package, circumference of the node changes into blue color and a label appears saying the active node's name, and when the sending activity finishes, there will be a label saying that the node is inactive. In addition, for all links among nodes, we have specified green color. Finally, flow number 1 is highlighted with blue color.

5. LITERATURE REVIEW

Evaluation of application and network potential efficiency are gotten through simulation experiments. This needs appropriate environment with several simulators and result analyzers tools. We evaluated and compared them in [2, 3]. One of the network simulators is NS-3 simulator which formally, started in mid 2006, but its first version appeared in 2008. It has appeared in library and dynamically or statically joins C++. These libraries start NS-3 and define simulation topology. Because of NS-3 architecture hardware simulators can work with all types of protocols. It also needs a few codes in protocol simulation location to perform it in real world. Unlike NS-2 that only works with some selective protocols, NS-3, because of its architecture can work with all types of protocols. It also needs a few codes in protocol simulation location to perform it in real world. Despite of the advantages, NS-3 has some disadvantages as well. This is a software simulator that is mainly based on GNU GNV2, and generally used for instructional and research purposes. Experiments also showed that this simulator can control large packages; however, it is problematic with small packages. On the other hand, there are not many models accessible for NS-3. Unlike DSMLs, there is no GUI for making typologies, let alone its picturing is temporary [8].

OMNET++ is another simulation tool which is known as a free multi-purpose broken simulation tool, and is used mainly for university and non-business purposes. Architecture based on objective structure of OMNET++ makes it possible to simulate in all systems. Since OMNET++ is just a multi-purpose simulator, it can only supply the core and programmer connector for API application to make a simulator. This simulator does not make elements to be used in particular networks such as computer networks. With a deep looking at OMNET++ structure, it can easily be understood that everything in OMNET++ appears as a module. Modules are written in C++, and form the real behavior of the system. Simulation activity is so complex in OMNET++. A person needs to learn NED language to understand interactive modules, INET frames, NED files, and ini files. OMNET++ is also supported by GUI and is performable on GUI as well. The GUI simulator of OMNET++ only shows the performance of simulation rather than making a scenario. In comparison with Domain Specific Modeling Languages, we can make several formation scenarios graphical formats [9].

To compensate the disadvantages of simulators, there is an attempt to have system developments over formal designing languages such as SDL, and supporting them in creation, holding, validation and performing happens automatically. For this reason, researchers introduced NS+SDL simulator which is developed version of NS and is a combination of NS2 simulator and SDL language design.

NS+SDL enables developer to use SDL designing plan as basic common codes in generating codes for simulators. From this basic code, C code is generated automatically using SDL-to-C which expected to reflect the results of system behavior application evaluation. In addition, when NS2 combines with a proper environmental SDL package which is fully supported by NS+SDL, all the system nodes may use their formation and Log files. With this shape of support, the attempt to simulate system application will decrease, and NS+SDL simulator can survive as a valuable ring in tools' cycle. A sample work has been done in Humboldt University of Berlin in which integrative tool environment of SDL has been created with compiler parameters in C++ destination language and JAVA. The problem here is that, since compiler parameters such as code generators suffer from lack of constancy and enough documents, the project hasn't been certified. On the other hand, self-contained systems on SDL can be simulated. For instance, it is impossible to simulate and control several SDL systems with an external simulator or with simulation parameters from variety of different sources. Recently, an application environment for simulation based on SDL set has been created in Aocher University. This environment is made of up SDL application evaluation tool of class library (SPEETCL), SDL-to-C compiler for code generation and graphical interactive simulation tool (GIST). SDL set of planning property is used as a common basic code in code generation and simulation. The problem is that, since compilers that are used in code generation and simulation are different from each other, the validity of results drops down [18].

In a research [25] subjected as SDL code generator for network simulation, Brumbulli and Fischer introduced new techniques for combining and changing SDL codes. The aim was to generate automatic C++ codes to produce modules for NS-3 simulator. Since new module integrates with simulator, the simulator stays unchanged. In this trend, system developing process uses SDL designing language and RTDS tools to make an integrative environment for validation, holding and automatic performing for the system. On the other hand, these tools disable to evaluate application behavior of network and distribution systems.

One of the tools in showing and evaluating simulation products is NAM network animator which supplies a visual picture of network topology. This tool enables operators to generate NS-2 codes with operator graphical link and based on Tel. NS-2 cable manual simulators have 5 main elements: nodes, links, queues, packages and agents. NAM, visually explains nodes with defined shapes in directed files and shows cable links among connected nodes as lines between them. It also gives cable packages as blocks which run in link routs. During replaying active information and packages are shown on monitoring window of NAM. NAM first was made to cable networks. It can be used to show nodes' movement rather than wireless systems. On the other word, there is no visualize in transporting and commuting in produced routs during NAM usage for NS-2 wireless scenarios [19].

Another simulator is Pyviz, which explains visual aspects, nodes moving, links among nodes and data transferring. It has a mediator interactive program. After NS-3010 was published, Pyviz was made as a separate connector from NS-3. This visualizing tool is a part of NS-3 source code that mainly is written by Python language; however, supports C++ network simulator. The motivation beyond Pyviz production was completely operator oriented and sociological. Like other network simulators, visualizing tools are needed in evaluation of its validity. Pyviz has no information supporting to tolerate delays in network simulation. A defect for Pyviz in

moving nodes in visualizing is that nodes pay no attention to moving area of Pyviz so they disorder program while it is running [21]. Another tool in showing simulation results is Scalar that can be used in producing scalar files. It enables the extraction of a single text which can be read by MATLAB, and spread sheets such as Excel and open Office Calc. All these external applications are strongly able to statistical analyze and visualize. On the other hand, OMNET++ cannot improve its capacity. On the contrary, OMNET++ mainly focuses on real simulators and pays no attention to the simulation results [19]. All the existent simulators have their own defect in Domain Specific Modeling Language. A Domain Specific Modeling Language not only is able to model a graphical program and control it, but also can change that model into other peer models and improve language semantics, generate performable codes on popular languages and improve efficiency. In this case the speed of code generation increases and the risk of error arising decreases, as a result, overall efficiency will increase.

6. CONCLUSION

In this paper, NS-2 programs are selected as regarded domain for developing a DSML. Considering simulation concepts and relations, abstract syntax was formally defined as a metamodel proportional to language complexities. Then, the concrete syntax was prepared in the form of appropriate notations and a graphic modeling editor supported by a graphical tool. Also, an instance model was represented by this tool to illustrate the use of proposed methodology.

As our future work, we are going to focus on the formal semantic representations for NS-2 programs in the form of constraints [24]. To identify language statements and their meanings considering language, semantic is a must. Otherwise, some imaginations may form about language, leading to its incorrect application. As another future work, to have the new language applicable, we aim to use the models as input for code generation. This will be realized in the form of operational semantics and model transformation. Finally, integrated tools will be developed for supporting these features of NS-2 programs in target language.

ACKNOWLEDGMENT

The authors would like to acknowledge Dr. Ali Akbar Dadjouyan for his helpful comments and suggestions during the preparation of this study regarding NS-2 programming. Also, Elham Azadi Marand would like to acknowledge Moharram Challenger, her M.Sc supervisor, for all his guidance and support during her thesis study.

REFERENCES

- [1] Azadi, E., and M.Challenger, 2013. Assessment and Comparison of designing methods, Implementation and required tools for Domain Specific Languages (DSLs) development. First National Conference on Advances in computer science and information retrieval approaches. Young Researchers Club, Islamic Azad University. Roudsar and Amlash, Iran, pp. 1-8. (in Persian).
- [2] Azadi, E., E. Azadi and M. Challenger, 2014. Assessment and Comparison of Network Simulation Tools and Languages. 8th Symposium on Advances in Science & Technology (8th SASTech 2014), Mashhad, Iran, pp. 1-8, (in Persian).
- [3] Azadi, E., E. Azadi and M. Challenger, 2014. Assessment and Comparison of Topology generation tool and View and analyze the results of network simulation. 8th Symposium on Advances in Science & Technology (8thSASTech 2014), Mashhad, Iran, pp. 1-8. (in Persian).
- [4] Lacage, M. and T.R. Henderson, 2006. Yet Another Network Simulator. In Proceedings of the Seventh ACM Symposium on Modeling. pp. 1-8.
- [5] Barr, R., V. Haas and R. Renesse, 2004. JiST: An efficient approach to simulation using virtual machines. Software practice and Experience. pp.1-7.
- [6] Siraj, S., A.K. Gupta and R. Badgujar, 2012. Network Simulation Tools Survey. International Journal of Advanced Research in Computer and Communication Engineering. Vol. 1, Issue 4, pp. 1-10.
- [7] Naicken, S., A. Basu, B. Livingston and S. Rodhetbhai, 2006. A Survey of Peer-to-Peer Network Simulators. Proceedings of the Seventh Annual Postgraduate Symposium, Liverpool, UK. Pp. 1-8.
- [8] Thomas, C. 2011. NS-3 Simulation Of Wimax Networks. MSc Thesis. Department of Computer Science & Engineering - Washington University in St. Louis. pp. 1-60.
- [9] Hogie, L. 2007. Mobile Ad Hoc Networks: Modelling, Simulation and Broadcast-based Applications. Ph.D thesis in Computer Science. University of Luxembourg.
- [10] Montresor, A. and M. Jelasity. 2009. PeerSim: a scalable P2P simulator. In Proc. of the 9th Int. Conference on Peer-to-Peer (P2P'09). pp. 1-2.
- [11] Ezreik, A. and A. Gheryani, 2012. Design and Simulation of Wireless Network using NS-2. 2nd International Conference on Computer Science and Information Technology (ICCSIT'2012). Pp. 157-161.
- [12] Weingartner, E., H. Lehn and K. Wehrle, 2009. A Performance Comparison of Recent Network Simulators. in: Proceedings of the IEEE International Conference on Communications 2009 (ICC 2009), Dresden, Germany, IEEE, pp:1 – 5.
- [13] Demirkol, S., M. Challenger, S. Getir, T. Kosar, G. Kardas and M. Mernik, 2013. A DSL for the Development of Software Agents Working in the Semantic Web Environment. International Journal of Computer Science and Information Systems (ComSIS). (Accepted) DOI: 10.2298/CSIS121105044D.
- [14] Kent, S. 2002. Model Driven Engineering. Lecture Notes in Computer Science Volume 2335, pp 286-29.
- [15] Schmidt, D.S. 2006. Model-driven engineering. IEEE Computer Society, pp. 25-31.
- [16] Voelter, M., S. Benz, C. Dietrich, B. Engelman, M. Helander, L. Kats, E. Visser and G. Wachsmuth, 2014. DSL Engineering. 2010-1013, Available online at: <http://dslbook.org>.
- [17] Clark, T., P. Sammut and J. Willans, 2008. Applied Meta Modeling A Foundation For Language Driven Development. available online at: [http://eprints.mdx.ac.uk/6060/1/Clark-Applied_Metamodelling_\(Second_Edition\)\[1\].pdf](http://eprints.mdx.ac.uk/6060/1/Clark-Applied_Metamodelling_(Second_Edition)[1].pdf) pp. 1-224.
- [18] Kuhn, T., A. Gerdaly, R. Gotzhein and F. Rothl, 2005. ns+SDL – The Network Simulator for SDL Systems. A. Prinz, R. Reed, and J. Reed (Eds.): SDL 2005, LNCS 3530, pp. 103–116.
- [19] Reineck, K.M. 2008. Evaluation and Comparison of Network Simulation Tools. Master Thesis, University of Applied Sciences Bonn Rhein-Sieg, Department of Computer Science.
- [20] Challenger, M., S. Getir, S. Demirkol and G. Kardas, 2011. A Domain Specific Metamodel for Semantic Web enabled Multi-agent Systems. Lecture Notes in Business Information Processing, vol. 83, pp. 177-186. Multi-agent Systems", Lecture Notes in Business Information Processing, Vol. 83, pp. 177-186, 2011.
- [21] Kurkowski, S., T. Camp and M. Colagrosso, 2004. A Visualization and Animation Tool for NS-2 Wireless Simulations: iNSpect. Dept of Mathematical and Computer Sciences.

- [22] Ekram, H. and I. Teerawat, 2009. Introduction to Network Simulator NS2. Springer New York Dordrecht Heidelberg London. ISBN: 978-1-4614-1405-6.
- [23] Challenger, M., Demirkol S., Getir S., Kosar T., Kardas G., Mernik M, 2014. On the use of a Domain-Specific Modeling Language in the Development of Multiagent Systems. International Journal of Engineering Applications of Artificial Intelligence, Vol. 28, pp. 111–141.
- [24] Getir, S., M. Challenger and G. Kardas, 2014. The Formal Semantics of a Domain-specific Modeling Language for Semantic Web enabled Multi-agent Systems. International Journal of Cooperative Information Systems, vol. 23, issue 3, pp. 1450005.
- [25] Brumbulli, M and J. Fisher, 2011. SDL Code Generation for Network Simulators. Springer-Verlag Berlin Heidelberg 2005, pp. 103–116.

A Comparative Analysis of Slicing for Structured Programs

LipikaJhaK.S.Patnaik

Department of Computer Science Department of Computer Science
and engineering, and engineering,
Birla Institute of Technology, Birla Institute of Technology,
Mesra, Ranchi, India Mesra, Ranchi, India

Abstract: Program Slicing is a method for automatically decomposing programs by analyzing their data flow and control flow. Slicing reduces the program to a minimal form called “slice” which still produces that behavior. Program slice singles out all statements that may have affected the value of a given variable at a specific program point. Slicing is useful in program debugging, program maintenance and other applications that involve understanding program behavior. In this paper we have discuss the static and dynamic slicing and its comparison by taking number of examples.

Keywords: Slicing techniques, control and data dependence, data flow equation, control flow graph, program dependence graph

1. INTRODUCTION

Program slicing is one of the techniques of program analysis. It is an alternative approach to develop reusable components from existing software. To extract reusable functions from ill-structured programs we need a decomposition method which is able to group nonsequential sets of statement. Program is decomposed based on program analysis. A program is analyzed using data flow and control flow. The decomposed program is called slice which is obtained by iteratively solving data flow equations based on a program flow graph. Program analysis uses program statement dependence information (i.e. data and control dependence) to identify parts of a program that influence or are influenced by a variable at particular point of interest is called the slicing criterion.

Slicing Criterion:

$$C = (n, V)$$

where n is a statement in program P and V is a variable in P . A slice S consists of all statements in program P that may affect the value of variable V at some point n .

Program slicing describes a mechanism or tool which allows the automatic generation of a slice. All statements affecting or affected by the variables V in n mentioned in the slicing criterion becomes a part of the slice.

This paper gives the detail description of slicing techniques and comparison of different slicing. The paper is organized as follows. Section 2 defines some common slicing techniques. Section 3 defines static and dynamic slicing. Section 4 defines comparison of different slicing, section 5 presents conclusion and finally acknowledgement and references.

2. SLICING TECHNIQUE

The original concept of a program slice was introduced by Weiser[1]. He claims that a slice corresponds to the mental abstractions that

people make when they are debugging a program. A slice S consists of all statements in program P that may affect the value of variable V at some point n.

Variables V at statements n can be affected by statements because:

- Statements which control the execution of n (Control Dependence)
- Statements which uses the V at n (Data Dependence)

The goal of slicing is to create a subprogram of the program (by eliminating some statements), such that the projection and the original program compute the same values for all variables in V at point n.

The computation of a slicing is done using data dependence and control dependence. Data dependence and control dependence are defined in terms of the CFG of a program. A control flow graph is a graph which is interpreted as a program procedure. The nodes of a graph represent program statements and edges represent control flow transfers between statements.

A Control Flow Graph for program P is a graph in which each node is associated with a statement from P and the edges represent the flow of control in P. With each node n (i.e., each statement in the program) associate two sets: USE(n), the set of variables whose values are used at n, and DEF(n), the set of variables whose values are defined at n.

There are three main techniques to compute slice:

2.1 Data dependence (DD) and control dependence (CD)

Data dependence and control dependence are defined in terms of the CFG of program. A statement j is data dependent on statement i if a value computed at i is used at j in some program

execution. A data dependence may be defined as; there exists a variable x such that

- (i) $x \in \text{DEF}(i)$ and $x \in \text{USE}(j)$
- or
- $x \in \text{DEF}(i)$ and $x \in \text{DEF}(j)$
- (ii) There exists a path from i to j without intervening definitions of x.

Control dependence information identifies the conditionals node that may affect execution of a node in the slice. In a CFG, a node j post-dominates a node i if and only if j is a member of any path from i to Stop. Node i is control-dependent on j in program P if

1. there exists a path P from i to j such that j post dominates every node in P.
2. i is not post dominated by j.

$$S_c = \{m \mid n . n \in C \text{ and } m \rightarrow^* n\}.$$

Or

$$S_c = DD \cup CD$$

2.2 Data Flow Equation

According to Weiser the slicing is computed by iteratively solving data flow equation. He defines slice as an executable program that is obtained from the original program by deleting zero or more statement. At least one slice exists for any criterion for any program that is the program itself.

The set of relevant variable $R_c^0(i)$ with respect to slicing criterion $C=(p,V)$ is:

1. $R_c^0(i)=V$ when $i=p$
2. $R_c^0(i)= (R_c^0(j)-\text{DEF}(i)) \cup (\text{USE}(i) \text{ if } R_c^0(j) \cap \text{DEF}(i) \neq \emptyset)$

The set of relevant statements to C denoted S_c^0 , is defined as:

$$S_c^0 = \{i \mid \text{Def}(i) \cap R_c^0(j) \neq \emptyset, i \rightarrow^{\text{CFG}} j\}$$

$$S_c^0 = \{3,4,9\}$$

The set of conditional statements which control the execution of the statements in S_c^0 , denoted B_c^0 is defined as:

$$B_c^0 = \{b \in G \mid \text{Infl}(b) \cap S_c^0 \neq \emptyset\}$$

$$\text{Infl}(5) = \{6,7,8\}$$

The sliced program S_c is defined recursively on the set of variables and statements which have either direct or indirect influence on V. Starting from zero, the superscripts represent the level of recursion.

$$R_c^{i+1}(n) = R_c^i(n) \cup_{b \in B_c^i} R_{b, U(b)}^o(n)$$

$$B_c^{i+1} = \{b \in G \mid \text{INFL}(b) \cap S_c^{i+1} \neq \emptyset\}$$

$$S_c^{i+1} = \{n \in G \mid \text{DEF}(n) \cap R_c^{i+1}(j) \neq \emptyset\} \cup B_c^i$$

The sliced program includes the conditional statements with an indirect influence on a slice, the control variables which are evaluated in the logical expression, and the statements which influence the control variables.

2.3 Program dependence graph

Program dependence graph is defined in terms of a program's control flow graph. The PDG includes the same set of vertices as the CFG, excluding the EXIT vertex. The edges of the PDG represent the control and flow dependence induced by the CFG.

The extraction of slices is based on data dependence and control dependence. A slice is directly obtained by a linear time walk backwards from some point in the graph, visiting all predecessors.

3. TYPES OF SLICING

3.1 Static Slicing

It includes all the statements that affect variable v or affected by the variable at the point of interest (i.e., at the statement x). It is computed by finding consecutive sets of indirectly relevant statements, according to data and control dependencies. Static slicing criterion consists of a pair (n, V) where n is point of interest and V is a variable in a program based on which program will be sliced.

3.2 Dynamic Slicing

A dynamic program slice includes all statements that affect the value of the variable occurrence for the given program inputs, not all statements that did affect its value. Dynamic slicing criterion consist of a triple (n, V, I) where I is an input to the program. In static slicing, only statically available information is used for computing slices.

4. COMPARISON USING DATA FLOW EQUATION

Example 1: Let us consider a program which computes the sum and product of first n numbers, using a single loop.

```
void main()
1.  {int n;
2.  cin >> n;
3.  if (n > 0)
4.  int i = 1;
5.  int sum = 0;
6.  int product = 1;
7.  int k;
8.  while (i <= n)
9.  {cin >> k;
10. sum = sum + k;
11. product = product * k;
12. i = i + 1;}
13. cout << sum;
14. cout << product;
}
```

Static slicing: Slicing criterion C is $(14, \text{product})$

Statement	USE()	DEF()	R_c^0	R_c^1
1		n	\emptyset	\emptyset
2		n	\emptyset	\emptyset
3	N		\emptyset	N
4		i	\emptyset	i, n
5		sum	\emptyset	i, n
6		product	\emptyset	product, i, n
7		k	product	product, i, n
8	i, n		product	product, i, n
9		k	product	product, i, n
10	sum, k	sum	product, k	product, i, n, k
11	product, i	product	product, k	product, i, n, k
12	I	i	product	product, i, n
13	Sum		product	Product
14	Product		product	Product

$$S_c^0 = \{6, 9, 11\}$$

$B_C^0 = \{3, 8\}$
 $R_C^1(n) = R_C^0(n) \cup_{b \in \langle 3, 8 \rangle} R_{b, \cup(b)}^0(n)$
 $R_{3,n}^0(3) = n$ $R_{3,n}^0(14) = n$ $R_{3,n}^0(13) = n$
 $R_{3,n}^0(12) = n$ $R_{3,n}^0(11) = n$
 $R_{3,n}^0(10) = n$ $R_{3,n}^0(9) = n$ $R_{3,n}^0(8) = n$
 $R_{3,n}^0(7) = n$ $R_{3,n}^0(6) = n$ $R_{3,n}^0(5) = n$
 $R_{3,n}^0(4) = n$ $R_{3,n}^0(2) = \emptyset$ $R_{3,n}^0(1) = \emptyset$
 $R_{8, \langle i, n \rangle}^0(8) = i, n$ $R_{8, \langle i, n \rangle}^0(12) = i, n$
 $R_{8, \langle i, n \rangle}^0(11) = i, n$ $R_{8, \langle i, n \rangle}^0(10) = i, n$
 $R_{8, \langle i, n \rangle}^0(9) = i, n$ $R_{8, \langle i, n \rangle}^0(7) = i, n$
 $R_{8, \langle i, n \rangle}^0(6) = i, n$ $R_{8, \langle i, n \rangle}^0(5) = i, n$
 $R_{8, \langle i, n \rangle}^0(4) = i$, $R_{8, \langle i, n \rangle}^0(3) = \emptyset$
 $R_{8, \langle i, n \rangle}^0(2) = \emptyset$ $R_{8, \langle i, n \rangle}^0(1) = \emptyset$
 $S_C^1 = \{12, 11, 9, 6, 4, 2\}$
 $B_C^1 = \{b \in G \mid \text{INFL}(b) \cap S_C^{i+1} \neq \emptyset\}$
 $B_C^1 = \{3, 8\}$
 $S_C^{i+1} = \{n \in G \mid \text{DEF}(n) \cap R_C^{i+1}(j) \neq \emptyset\} \cup B_C^i$
 $S_C^1 = \{2, 3, 4, 6, 8, 9, 11, 12\}$

Dynamic slicing: Slicing criterion
C: (14, product, n=2)

```

void main ()
1. {int n;
2. cin >> n;
3. if (n > 0)
4. int i = 1;
5. int sum = 1;
6. int product = 1;
7. int k;
81. i <= n
91. cin >> k;
101. sum = sum + k;
111. product = product * k;
121. i = i + 1;}
82. i <= n
92. cin >> k;
102. sum = sum + k;
112. product = product * k;
122. i = i + 1;
83. i <= n
13. sum = sum + k;
14. cout << product;
}

```

$DU: \{(2, 3), (2, 8^1), (2, 8^2), (2, 8^3), (4, 8^1), (4, 12^1), (12^1, 8^2), (12^1, 12^2), (12^2, 8^3), (6, 11^1), (11^1, 11^2), (11^2, 14), (9^1, 11^1), (9^2, 11^2)\}$

TC:
 $\{3, 4\}, \{3, 5\}, \{3, 6\}, \{3, 7\}, \{3, 8^1\}, \{(8^1, 9^1), (8^1, 11^1), (8^1, 12^1), (8^2, 9^2), (8^2, 11^2), (8^2, 12^2)\}$
 IR: $\{(8^1, 8^2), (8^1, 8^3), (8^2, 8^1), (8^2, 8^3), (8^3, 8^1), (8^3, 8^2)\}$

$S^0 = \{11^2\}$
 $A^1 = \{11^1, 9^2, 8^2\}$
 $S^1 = \{11^2, 11^1, 9^2, 8^2\}$
 $A^2 = \{6, 9^1, 8^1, 2, 8^3, 12^1\}$
 $S^2 = \{11^2, 11^1, 9^2, 8^2, 6, 9^1, 8^1, 2, 8^3, 12^1\}$
 $A^3 = \{4, 12^1, 12^2\}$
 $S^3 = \{11^2, 11^1, 9^2, 8^2, 6, 9^1, 8^1, 2, 8^3, 4, 12^1, 12^2\}$

Example 2: Let us consider a program which computes the sum and product of first n numbers, using for loop.

```

void main()
1. {int n;
2. cin >> n;
3. if (n > 0)
4. int sum = 0;
5. int product = 1;
6. int k;
7. for (i = 1; i <= n; i++)
8. {cin >> k;
9. sum = sum + k;
10. product = product * k;
}
11. cout << sum;
12. cout << product;
}

```

Static slicing: Slicing criterion C is (12, product)

Statement	USE()	DEF()	R_C^0	R_C^1
1		n	\emptyset	I
2		n	\emptyset	I
3	N		\emptyset	i, n
4		sum	\emptyset	i, n
5		product	\emptyset	i, n
6		k	Product	product, i, n
7	i, n	i	Product	product, i, n
8		k	Product	product, i, n

				,
9	sum,k	sum	product, k	product,i,n ,k
10	product, k	product	product, k	product,i,n ,k
11	Sum		product	Product
12	Product		product	Product

$$S_C^0 = \{11,6,4\}$$

$$B_C^0 = \{7,8,9\}$$

$$R_{7,n}^0(3)=n \quad R_{7,n}^0(11)=\text{product},n$$

$$R_{7,n}^0(10)=\text{product},k,n$$

$$R_{7,n}^0(9)=\text{product},k,n \quad R_{7,n}^0(8)=\text{product},n$$

$$R_{7,n}^0(7)=\text{product},n \quad R_{7,n}^0(6)=\text{product},n$$

$$R_{7,n}^0(5)=n$$

$$R_{3,n}^0(4)=n \quad R_{7,n}^0(2)=\emptyset \quad R_{3,n}^0(1)=\emptyset$$

$$R_{7,<i,n>}^0(7)=i,n \quad R_{7,<i,n>}^0(11)=i,n$$

$$R_{7,<i,n>}^0(10)=i,n \quad R_{7,<i,n>}^0(9)=i,n$$

$$R_{7,<i,n>}^0(8)=i,n \quad R_{7,<i,n>}^0(6)=i,n \quad R_{7,<i,n>}^0(5)=i,n$$

$$R_{7,<i,n>}^0(4)=i,n \quad R_{7,<i,n>}^0(3)=i,n$$

$$R_{7,<i,n>}^0(2)=i \quad R_{7,<i,n>}^0(1)=i$$

$$S_C^1 = \{10,8,7,5,2\}$$

$$B_C^1 = \{3,7\}$$

$$S_C = \{2,3,5,7,8,10\}$$

Dynamic slicing: Slicing criterion

C:(12,product,n=2)

void main ()

1. {int n;
2. cin>>n;
3. if (n>0)
4. int sum =1;
5. int product=1;
6. int k;
- 7¹. i=1; i<=n
- 8¹. cin>>k;
- 9¹. sum=sum+k;
- 10¹. product=product*k;
- 7². i<=n
- 8². cin>>k;
- 9². sum=sum+k;
- 10². product=product*k;
- 7³. i<=n
11. cout<< sum;
12. cout<<product;
- }

DU: $\{(2,3),(2,7^1),(2,7^2),(2,7^3),(4,9^1),(9^1,9^2),$
 $(9^2,11),(5,10^1),(10^1,10^2),(10^2,12),(8^1,9^1),(8^1,10^1),$
 $(8^2,9^2),(8^2,10^2)\}$

TC:
 $\{(3,4),(3,5),(3,6),(3,7^1),(7^1,8^1),(7^1,9^1),(7^1,10^1),$
 $(7^2,8^2),(7^2,9^2),(7^2,10^2)\}$

IR: $\{(7^1,7^2),(7^1,7^3),(7^2,7^1),(7^2,7^3),(7^3,7^1),$
 $(7^3,7^2)\}$

$$S^0 = \{10^2\}$$

$$A^1 = \{10^1, 8^2, 7^2\}$$

$$S^1 = \{10^2, 10^1, 8^2, 7^2\}$$

$$A^2 = \{5, 8^1, 7^1, 2, 7^3, 10^1\}$$

$$S^2 = \{10^2, 10^1, 8^2, 7^2, 5, 7^1, 8^1, 2, 7^3\}$$

$$A^3 = \{3\}$$

$$S^3 = \{10^2, 10^1, 8^2, 7^2, 5, 7^1, 8^1, 2, 7^3, 3\}$$

5. CONCLUSION

The concept of program slicing was originally defined by Weiser as a method for decomposing a program into pieces, i.e., slices. The static slice produced by his method is of the partially equivalent program type. His solution produces a static program slice based on solving data flow equations iteratively on the flow graph of the program. Here in this paper we have used the same concept for computing the static and dynamic slicing. We have taken the number of example and analyzed the slicing. As the number of control statement increases in a program it becomes difficult to compute the control dependent statement and then the relevant variable. In for loop we can either consider the different statement for the three section or we can consider a single statement. The output will be same the only difference is that it will be easy to compute from the different statement.

6. ACKNOWLEDGEMENT

I wish to convey my sincere gratitude and appreciation to each and every person who helped me in writing this paper. I am grateful to my institution, Birla Institute of Technology and my colleagues. I would especially like to thank Dr. K. S. Patnaik, my guide for his advice and guidance.

7. REFERENCES

- [1] M. Weiser, “Program Slicing,” IEEE Transactions on Software Engineering, Vol. 16, No. 5, 1984, pp. 498-509.
- [2] F. Tip, “A Survey of Program Slicing Techniques,” Journal of Programming Languages, Vol. 3, No. 3, 1995, pp. 121-189
- [3] D. Binkley and K. B. Gallagher, “Program Slicing,” Advances in Computers, Vol. 43, 1996, pp. 1-50.
- [4] B. Korel and J. Laski. Dynamic program slicing. *Information Processing Letters*, 29(3):155–163, 1988.
- [5] M. Kamkar, “An overview and comparative classification of program slicing techniques,” J. Systems Software, vol. 31, pp. 197–214, 1995.

VHDL Implementation of FPGA Based OFDM Modem for Wireless Applications

Ch. Goutham Raj
Joginapally.B.R.Engg
College
Yenkappally Moinabad,
Rangareddy, TS, India

G.Krishnaiah
Joginapally.B.R.Engg
College
Yenkappally, Moinabad,
Rangareddy, TS, India

D. N. Rao
Joginapally.B.R.Engg
College
Yenkappally, Moinabad,
Rangareddy, TS, India

Abstract: The new mobile technologies trying to give broadband over wireless channel allowing the user to have bandwidth connectivity even inside moving vehicle. The metropolitan broadband wireless networks require a non-line-of-sight (NLOS) capability, and the scheme Orthogonal Frequency Division Multiplex (OFDM) becomes essential to overcome the effects of multipath fading. Orthogonal Frequency Division Multiplexing (OFDM) has become very popular, allowing high speed wireless communications. OFDM could be considered either a modulation or multiplexing technique, and its hierarchy corresponds to the physical and medium access layer. A basic OFDM modulator system consists of a QAM or PSK modulator, a serial to parallel, and an IFFT module. The iterative nature of the IFFT and its computational order makes OFDM ideal for a dedicated architecture outside or parallel to the main processor. The VHDL implementation allows the design to be extended for either FPGA or ASIC implementation, which suits more for the Software Defined Radio (SDR) design methodology. In this project the OFDM modulator and demodulator will be implemented with full digital techniques. VHDL will be used for RTL description and FPGA synthesis tools will be used for performance analysis of the proposed core. Modelsim Xilinx Edition will be used for functional simulation and verification of results. Xilinx ISE will be used for synthesis. The Xilinx's chipscope tool will be used for verifying the results on Spartan 3E 3S500EFG320-4 FPGA.

Keywords: OFDM, FFT, IFFT, VHDL, FPGA.

1. INTRODUCTION

The telecommunications industry faces the problem of providing telephone services to rural areas, where the customer base is small, but the cost of installing a wired phone network is very high. One method of reducing the high infrastructure cost of a wired system is to use a fixed wireless radio network [8]. The problem with this is that for rural and urban areas, large cell sizes are required to obtain sufficient coverage. This results in problems caused by large signal path loss and long delay times in multipath signal propagation.

Currently Global System for Mobile Telecommunications (GSM) technology is being applied to fixed wireless phone systems in rural areas or Australia. However, GSM uses Time Division Multiple Access (TDMA), which has a high symbol rate leading to problems with multipath causing inter-symbol interference.

Several techniques are under consideration for the next generation of digital phone systems, with the aim of improving cell capacity, multipath immunity, and flexibility. These include Code Division Multiple Access (CDMA) and Coded Orthogonal Frequency Division Multiplexing (COFDM). Both these techniques could be applied to providing a fixed wireless system for rural areas. However, each technique has different properties, making it more suited for specific applications.

COFDM is currently being used in several new

radio broadcast systems including the proposal for high definition digital television Digital Video Broadcasting (DVB) and Digital Audio Broadcasting (DAB)[1]. However little research has been done into the use of COFDM as a transmission [8],[6] method for mobile telecommunications systems.

With CDMA systems, all users transmit in the same frequency band using specialized codes as a basis of channelization. The transmitted information is spread in bandwidth by multiplying it by a wide bandwidth pseudo random sequence. Both the base station and the mobile station know these random codes that are used to modulate the data sent, allowing it to de-scramble the received signal.

OFDM/COFDM [2] allows many users to transmit in an allocated band, by subdividing the available bandwidth into many narrow bandwidth carriers. Each user is allocated several carriers in which to transmit their data. The transmission is generated in such a way that the carriers used are orthogonal to one another, thus allowing them to be packed together much closer than standard frequency division multiplexing (FDM). This leads to OFDM/COFDM [2] providing a high spectral efficiency. Orthogonal Frequency Division Multiplexing (OFDM)[9] is a multicarrier transmission technique, which divides the available spectrum into many carriers, each one being modulated by a low rate data stream. OFDM is similar to FDMA in that the multiple user access is achieved by subdividing the available bandwidth into multiple channels,

which are then allocated to users. However, OFDM uses the spectrum much more efficiently by spacing the channels much closer together. This is achieved by making all the carriers orthogonal to one another, preventing interference between the closely spaced carriers. With the rapid growth of digital communication in recent years, the need for high-speed data transmission has been increased. The mobile telecommunications industry faces the problem of providing the technology that be able to support a variety of services ranging from voice communication with a bit rate of a few kbps to wireless multimedia in which bit rate up to 2 Mbps. Many systems have been proposed and OFDM system has gained much attention for different reasons. Although OFDM was first developed in the 1960s, only in recent years, it has been recognized as an outstanding method for high-speed cellular data communication where its implementation relies on very high-speed digital signal processing. This method has only recently become available with reasonable prices versus performance of hardware implementation.

Since OFDM is carried out in the digital domain, there are several methods to implement the system. One of the methods to implement the system is using Field-Programmable Gate Array (FPGA) [3]. This hardware is programmable and the designer has full control over the actual design implementation without the need (and delay) for any physical IC fabrication facility. An FPGA combines the speed, power, and density attributes of an ASIC with the programmability of a general purpose processor will give advantages to the OFDM system. An FPGA could be reprogrammed for new functions by a base station to meet future needs particularly when new design is going to fabricate into chip. This will be the best choice for OFDM implementation since it gives flexibility to the program design besides the low cost hardware component compared to others.

2. FPGA BASED VLSI DESIGN

Recent strides in programmable logic density, speed and hardware description languages (HDL's) have empowered the engineer with the ability to implement digital signal processing (DSP) functionality within programmable logic devices (PLDs or FPGAs)[2]. In this chapter the VLSI design methods on FPGAs using VHDL are discussed. The FPGA architectures basics are also discussed. The following section presents fundamentals of VHDL and synthesis issues related to it.

2.1 VHDL

VHDL is a high level description language for system and circuit design. The language supports various levels of abstraction. In contrast to regular netlist formats that supports only structural description and a boolean entry system that supports only dataflow behavior, VHDL supports a wide range of description styles. These include structural descriptions, dataflow descriptions and behavioral descriptions.

The structural and dataflow descriptions show a concurrent behavior. That is, all statements are executed concurrently, and the order of the statements is not

relevant. On the other hand, behavioral descriptions are executed sequentially in processes, procedures and functions in VHDL. The behavioral descriptions resemble high-level programming languages.

VHDL allows a mixture of various levels of design entry abstraction. Precision RTL Synthesis Synthesizes will accept all levels of abstraction, and minimize the amount of logic needed, resulting in a final netlist description in the technology of your choice. The Top-Down Design Flow is shown in Figure 1.

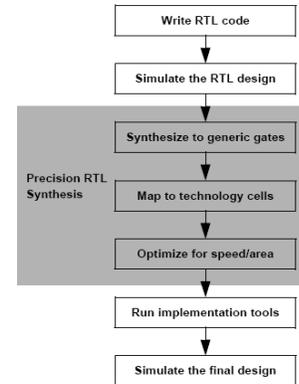


Figure 1. Top-Down Design Flow with Precision RTL Synthesis

2.2 VHDL and Synthesis

VHDL is fully simulatable, but not fully synthesizable. There are several VHDL constructs that do not have valid representation in a digital circuit. Other constructs do, in theory, have a representation in a digital circuit, but cannot be reproduced with guaranteed accuracy. Delay time modeling in VHDL is an example. State-of-the-art synthesis algorithms can optimize Register Transfer Level (RTL) [3] circuit descriptions and target a specific technology. Scheduling and allocation algorithms, which perform circuit optimization at a very high and abstract level, are not yet robust enough for general circuit applications. Therefore, the result of synthesizing a VHDL description depends on the style of VHDL that is used.

3. DESIGN OF OFDM MODEM

3.1 Block Diagram of OFDM MODEM:

The OFDM Modem [4] which consists of shows the configuration for a basic OFDM transmitter and receiver.

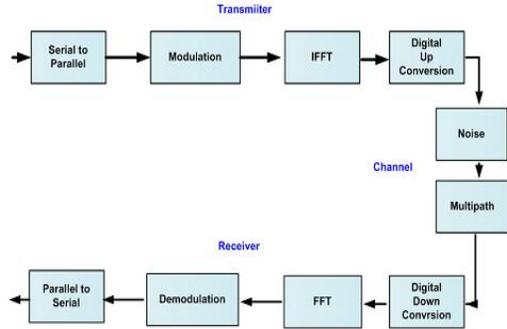


Figure 2: Block Diagram of OFDM MODEM

The **OFDM transmitter** which consists of

1. Clock Distributor
2. Input data sampler
3. Symbol Mapper
4. Modulation
5. Serial to Parallel
6. IFFT

The **OFDM receiver** which consists of

1. Clock Distributor
2. FFT
3. Parallel to Serial
4. Demodulation
5. IQ extractor
6. Symbol demapper
7. Output Data

The signal generated is at base-band and so to generate an RF signal the signal must be filtered and mixed to the desired transmission frequency [9].

To generate OFDM successfully the relationship between all the carriers must be carefully controlled to maintain the Orthogonality of the carriers. For this reason, OFDM is generated by firstly choosing the spectrum required, based on the input data, and modulation scheme used. Each carrier to be produced is assigned some data to transmit. The required amplitude and phase of the carrier is then calculated based on the modulation scheme (typically differential BPSK, QPSK, or QAM) [8],[2]. The required spectrum is then converted back to its time domain signal using an Inverse Fourier Transform. In most applications, an Inverse Fast Fourier Transform (IFFT) is used. The IFFT performs the transformation very efficiently, and provides a simple way of ensuring the carrier signals produced are orthogonal.

The Fast Fourier Transform (FFT) transforms a cyclic time domain signal into its equivalent frequency spectrum. This is done by finding the equivalent waveform, generated by a sum of orthogonal sinusoidal components. The amplitude and phase of the sinusoidal components represent the frequency spectrum of the time domain signal. The IFFT performs the reverse process, transforming a spectrum (amplitude and phase of each component) into a time domain signal. An IFFT converts a number of complex data points, of length that is a power of 2, into the time domain signal of the same number of points. Each data point in frequency spectrum used for an FFT or IFFT [4],[5] is called a bin.

The orthogonal carriers required for the OFDM [6] signal can be easily generated by setting the amplitude and phase

of each frequency bin, then performing the IFFT. Since each bin of an IFFT [2] corresponds to the amplitude and phase of a set of orthogonal sinusoids, the reverse process guarantees that the carriers generated are orthogonal.

3.2 Design of *Ofdm_tx*:

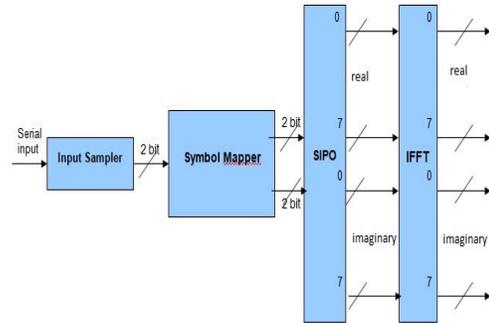


Figure 3 : Block Diagram Ofdm_tx

- 1) **Input Sampler IQ Gen:** This block samples the serial input and generates 2 bit IQ output.
- 2) **Symbol Mapper:** This block maps the input I, Q to the corresponding to the real part and imaginary part of the constellation symbols.

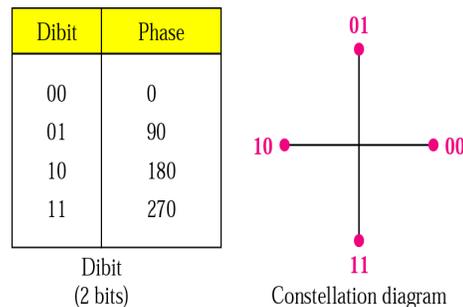


Figure 4: Constellation Diagram

- 3) **SIPO:** This block converts the serial input to the parallel output. This block is used in OFDM TX [2], to convert serial input to parallel output. This block's output is given to the input of IFFT.

3.2 Design of *Ofdm_rx*:

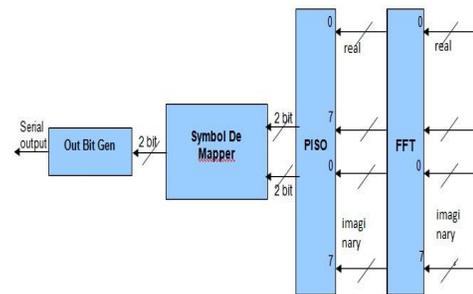


Figure 5: Block diagram of Ofdm_rx

1) **PISO**: This block converts the parallel input to the serial output. This block is used in OFDM RX [7], to convert parallel input to serial output. This block's output is given to the Symbol_de_mapper.

2) **Symbol_de_mapper**: This block maps, the Real and imaginary parts of the serial out from PISO, to the IQ corresponding to the real part and imaginary part of the constellation symbols. It extracts the IQ values from the serial out of PISO.

3) **Out_bit_gen**: This block takes 2 bit IQ s from Symbol_de_mapper and generates output bits.

4) **Clk distr**: This is the clock distributor block, which generates two enable signals

- en_div_2 and en_div_16.
- en_div_2 is divided by 2 of input clock.
- en_div_16 divided by 16 of input clock.

4. RESULTS AND DISCUSSIONS

The following chapter consists of all the Software and Hardware results observed in the project. The Results include snapshots of Top module with the inputs, outputs and intermediate Waveforms.

4.1 Top ofdm_tx_and_rx:

OFDM complete transmitter and receiver output.

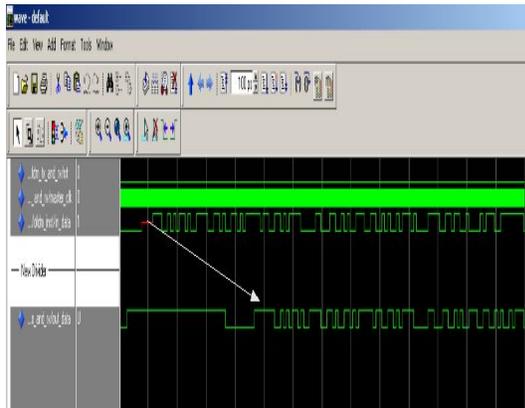


Figure 6: Simulation Result of Top ofdm_tx_rx

OFDM TX:

It consists of all the above sub blocks as components. The simulation results of above module is shown below.

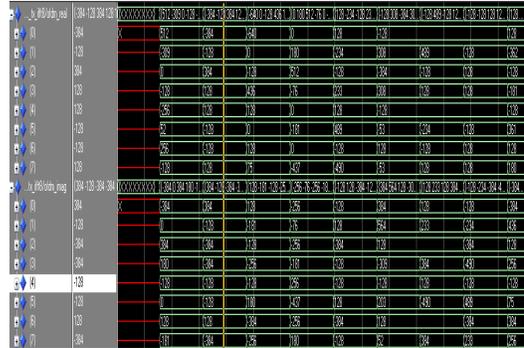


Figure 7: Simulation Result of Top ofdm_tx

OFDM_RX:

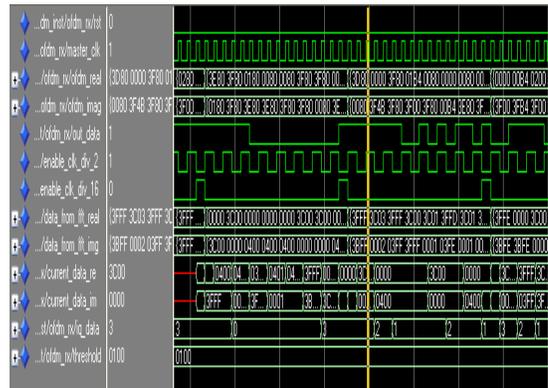


Figure 8: Simulation Result of Top ofdm_rx

4.2 Chipscope result:

Chipscope is an embedded software based logic analyzer. By inserting an “integrated Controller core” (icon) and “integrated logic analyzer” (ila) into your design and connecting them properly, you can monitor any or all of the signals in your design. Chipscope provides you with a convenient software based interface for controlling the “integrated logic analyzer,” including setting the triggering options and viewing the waveforms. Users can place the ICON, ILA, VIO, and ATC2 cores (collectively called the Chipscope Pro cores) into their design by generating the cores with the Core Generator and instantiating them into the HDL source code. We can also insert the ICON, ILA, and ATC2 cores directly into the synthesized design net list using the

Core Inserter tool. The design is then placed and routed using the ISE 9.2i implementation tools. Next, we download the bit stream into the device under test and analyze the design with the Analyzer software.

The tools design flow merges easily with any standard FPGA design flow that uses a standard HDL synthesis tool and the ISE 9.2i implementation tools.

Chipscope pro analyzer we are using here.

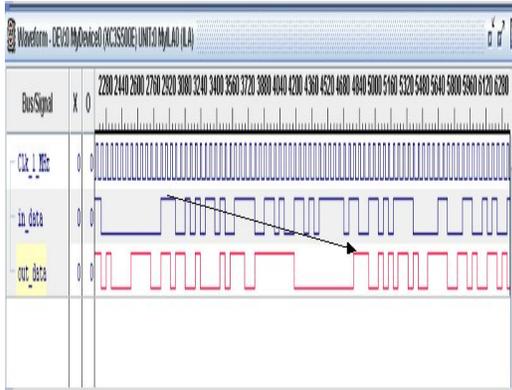


Figure 9 : ChipScope Result of Top Ofdm_tx_rx

5. CONCLUSION

In this paper, OFDM MODEM has been studied and implemented for Modulator and Demodulator and Its applications have been extended from high frequency radio communications to telephone networks, digital audio broadcasting and terrestrial broadcasting of digital television. The advantages of OFDM, especially in the multipath propagation, interference and fading environment, make the technology a promising alternative in digital communications including mobile multimedia. Therefore this design can be applied to real-time signal processing system, which completes the main computing modules in the OFDM for multi services.

The capability of designing and implementing an OFDM MODEM is presented in this work, the design considered using a pure VHDL with the aid of IPs to implement the IFFT and clock Synthesis Function, from the Mapping results the design can be easily fit into Xilinx FPGA XC3S500E.

6. REFERENCES

- [1] Lenin Gopal*, Daniel Wong Sing Tze, Nur Zawanah Ishak Design of an FPGA-Based OFDM Transceiver for DVB-T Standard, School of Engineering and Science, Curtin University Sarawak Malaysia Miri, Sarawak, Malaysia
- [2] On the design of an FPGA-Based OFDM modulator for IEEE 802.16-2004, Joaquín García, René Cumplido, Department of Computer Science, INAOE, Puebla, Mexico, joaquinr@ccc.inaoep.mx, rcumplido@inaoep.mx
- [3] E. Lawrey, Multiuser OFDM, in Proc. International Symposium on Signal Processing Applications '99 vol2, 1999, pp.761.764.
- [4] Ahmad sghaier, Shawki Areibi and Bob Dony, School of Engineering, Guelph, Guelph ON Canada N1G2W1, A Pipeline Implementation Of OFDM Transmission On Reconfigurable Platforms
- [5] Joaquín García, René Cumplido Department of Computer Science, INAOE, Puebla, Mexico On the design of an FPGA-Based OFDM modulator for

IEEE 802.16-2004, joaquinr@ccc.inaoep.mx, rcumplido@inaoep.mx

- [6] Fundamentals of Wireless communication by David Tse and Pramod Vishwanath.
- [7] Downlink BER Simulation for IEEE 802.16e OFDM-PHY by Pangan Tingt, Yung-Yih Jian, Chia-Tsung Lin, Yu-Tao Hsieh, Chi-Fang Li, Tsan-Huang Chen, Shuenn Gi Wei Chao-Kai J-H Lee, Ping Chuang, Wen, Racy Cheng IEEE 2005
- [8] "Air interface for fixed Broadband wireless Access systems" IEEE standard 802.16-2004, 2004
- [9] Weinstein, S. and Ebert, P.; "Data Transmission by Frequency-Division Multiplexing Using the Discrete Fourier Transform" IEEE Trans. on Commun. vol. 19 Issue: 5, pp. 628 –634, Oct. 1971

A Review Study on Secure Authentication in Mobile System

Ahitha.H.S
Ponjesly College of Engineering
Nagercoil, TamilNadu
India

Deepa.A.J
Ponjesly College of Engineering
Nagercoil, TamilNadu
India

Abstract: In mobile environment, the users communicate with each other about location based information and services with the help of network nodes. But providing security to data has become a difficult task. Currently available network security mechanisms are insufficient, but there is a greater demand for flexible, scalable security mechanisms. Mobile security is of vital importance but can't be inbuilt. In this work, proposing two techniques for authenticating short encrypted messages that helps to preserve the confidentiality and integrity of the communicated messages. The messages to be authenticated must also be encrypted using the secure authentication codes that are more efficient than message authentication codes. The key idea in this technique is to provide the security using Encryption Algorithm. Security model should adjust with various capability requirements and security requirements of a mobile system This paper provides a survey on security in mobile systems

Keywords: Authentication, Message Authentication code, Computational security, Unconditional Security.

1. INTRODUCTION

Mobile computing is human-computer interaction by which a computer is expected to be transported during usual usage. It includes mobile communication that leads to several communication issues. To resolve this issue the proposed work has two techniques, in the first technique message to be authenticated is also encrypted with any secure encryption algorithm to attach a short random string in the authentication process. Since the random strings are independent for different operations the authentication algorithm will be faster and more efficient without the difficulty to manage one-time keys. In the second technique we make more assumption to use an encryption algorithm in block cipher based to improve the computational efficiency.

As networking technology become common place and essential to everyday life, companies, organizations and individuals are increasingly depending on electronic means to process information and provide important services in order to take advantage of ambient brainpower in PCEs Pervasive computing environments. PCEs with their interconnected devices and abundant services promise great combination of digital infrastructure into many aspects of our life. Traditional authentication which focuses on identity authentication may fail to work in PCEs, to a certain extent because it conflicts with the goal of user privacy protection because the assurance achieved by entity authentication will be of diminishing value [10].

Preserving the integrity of the messages that are exchanged over the public channels is the traditional goal in cryptography that is mainly done using message authentication codes (MAC) for the only purpose of preserving message integrity. Based on security, message can be classified into two as unconditionally or computationally secure. Unconditionally secure MACs provide message integrity against forgers with the unlimited computational power. But computationally secure MACs are only secure when forgers have limited computational power. Computationally secure MACs are further classified as block cipher, cryptographic hash function, or universal hash

function. Phillip Rogaway [9] suggested a Block-Cipher Mode of Operation for Efficient Authenticated Encryption

More security issues are to be concerned with the fast transmission of digital information over wireless channels, the security issues include the spread of viruses and malicious software. There are about 200 mobile viruses or a malware program that causes problems to the systems based on F-Secure [1]. One of the major issues in mobile system is the low computational power. Many organizations are interested in deploying mobile application to improve efficiency and allow the capabilities. Further, the sections are preplanned as follows, Section 2 explains several cryptographic terms, Section 3 includes several Authentication methods, Section 4 Security Model that describes an application about deploying of (Radio Frequency Identification) RFID system and a cryptography based security model Section 5 Benefits of secure authentication, Section 6 Conclusion.

2. CRYPTOGRAPHIC TERMS

Basic security terms used in cryptography is the plaintext, cipher text, Encryption and Decryption. Plaintext is an ordinary form of message that is known to a viewer. Cipher text is the result obtained after performing Encryption on the plaintext. Encryption is done to hide the meaning of the message from everyone than the legal users. Decryption is inverse of Encryption. Strength of the scheme depends on the secrecy of the keys used. William Stallings [2] described in detail about the commonly used security terminologies. Security protocols identify the security objectives with the use of cryptographic algorithm. The main security objectives are as follows:

2.1 Authentication: The process by which the system checks the identity of the user who wish to access the information.

2.2 Confidentiality: The secrecy of the communicated data should be maintained. No one other than the legitimate user should know the content of the data.

2.3 Integrity: This is to check the originality of the message. It allows the receiver to verify that the message received was not altered during transfer.

The security objectives help us to provide trust on web. Trust is another feature of security coin that is related with both authentication and authorization. The algorithms are of two types as Symmetric algorithms and Asymmetric algorithms. Symmetric algorithms uses same key for both encryption and decryption. It is mainly used for providing confidentiality. An asymmetric algorithm uses different keys for encryption and decryption as public key and private key respectively. It is mainly used for authentication and non-repudiation [2].

3. AUTHENTICATION TYPES

Authentication is a process by which the system checks the identity of the user who needs to access it. It is one of the most essential security primitive. Authentication mainly based on the three factors as knowledge, possession and Attribute. The most common authentication methods are as follows:

3.1 Passwords: The most popular authentication scheme is password that is used for multiple services. This is a straight forward method that provides sufficient security. Subjects should be aware about the length of their password and security. The drawback is that multiple password is hard to remember.

3.1.1 Single Factor Authentication: In single factor authentication user can use only one factor for authentication as the basic user name/password. The password may be textual, graphical Password or PIN. Harsh Kumar Sarohi et al [3] proposed graphical authentication method here a password consist of sequence of one or more images, with the help of mouse events like click, drag etc the user can input their password.

3.1.2 Multiple Factor Authentication: ShindeSwapnil et al [4] proposed multiple factor authentication user that can use multiple factors for authentication such as what the user knows (password or PIN), what the user has (smart card), what the user is (Biometric authentication). Multi-factor authentication is a type of strong authentication.

3.2 Tokens: Tokens are supposed to have high security and usability. To many service are grouped into the device and hence the token generators are included in the device. Every subject should use tokens in any of the form either as paper tables or key generators.

3.3 Biometrics: Fingerprint is the most chosen biometric authentication method. Users like the feature of quick learning, but it may lead to the risk of data loss. Biometric factors are unable to change and the loss of private data is bad and it should be avoided.

4. SECURITY MODEL IN MOBILE SYSTEM

4.1 RFID (Radio frequency Identification)

RFID is the wireless non-contact use of radio frequency with electromagnetic fields to transfer data. The

main use is to automatically identify and track the tags that are attached to objects. The tags contain information that are stored electronically. RFID in (figure 1) is similar to barcodes, it uses an electronic chip that is fixed on a product or an artifact. The information could be read, recorded or rewritten.

RFID System can be classified by the type of tag and Readers. A radio frequency identification system use tags that are attached to the objects that are to be identified. Tags are similar to labels. Two-way radio transmitter-receivers are called as readers. The readers send the signal to the tag and read its response. RFID tags contain two parts as an integrated circuit and an antenna used for storing and processing information and for receiving and transmitting the signals respectively. Tag receives the message and then responds with its identification and other information. The Tag information is stored in a non-volatile memory. Readers are also known as interrogators. Reader transmits an programmed radio signal to interview the tags.

RFID tags are of two types Passive and Active tags. Active tag has a battery attached on-board and it transmits its ID signal periodically. It is also known as Battery assisted passive tags that are activated in the presence of an RFID reader. But passive tags are cheaper and smaller because it has no battery.

Figure 1: RFID SYSTEM ARCHITECTURE

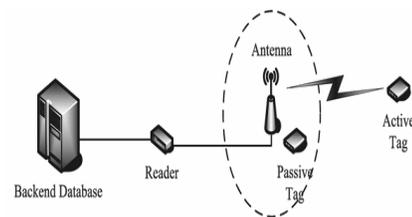


Fig.1 RFID system architecture

RFID Authentication protocols used for secure and low cost RFID Systems [5]. To solve the security problem the low cost RFID system with the authentication protocols were proposed.

- SRAC (Semi-Randomized Access Control)
- ASRAC (Advanced Semi-Randomized Access Control)

4.1.1 SRAC (Semi-Randomized Access Control)

SRAC is designed using a hash function in tags as the security primitives. It resolves not only the security properties such as the tracking problem, cloning problem and denial of service attack but also solves the operational properties as scalability and rareness of IDs. In the randomized access control the tag replies to the reader by sending a message back as response that varies each time. Communication between the Server and Reader is done using strong keys with sufficient resources so they will be secure. But we have to concern more about the communication between reader and tags. The reader will pass the arrived messages to the server.

4.1.2 ASRAC (Advanced Semi-Randomized Access Control)

ASRAC is used mainly for preventing replay attack. Replay attack is a kind of active attack. The ASRAC is designed using a hash function and a random producer for security primitives. SRAC helps on reducing 75% of tag transmission. The advantage is that since both reader and tag confirms the received message using hashed outputs which internally has the generated random numbers, attackers cannot use the past messages.

4.2 Public key Cryptography based security model

Key management is an important issue in public key cryptography. The shared wireless network faces the problem of eavesdropping. R. Shanmugalakshmi [6] proposed that ECC's provide high security with smaller key size than RSA. ECC used in the security development in the field of information security and mobile devices with low computational power. When compared with RSA, hardware implemented ECC has less operating cost. ECC is a talented cryptosystem for next generation and widespread use in devices.

Public key Cryptography is mainly based on the intractability of certain mathematical problems. Elliptic curves have several discrete logarithm based protocol. ECC used for digital signature generation and key exchange mainly used for protecting classified as well as unclassified national security systems and information. In most of the applications the RSA is replaced by ECC.

5. BENEFITS OF SECURE AUTHENTICATION

Most organizations have a major portion of their workers accessing their network from outside the office via mobile devices. Yet user names and passwords are not enough to adequately protect devices against unauthorized access. So Today's enterprise needs an efficient, user-friendly solution to address the security challenges that exist in today's mobile business. Secure authentication that offers a smarter, more flexible alternative to meet the unique needs of business today, including all the economic and business benefits of a hosted solution:

Protection: Validation and Protection Service cuts the risk of unauthorized access, data breaches, and other security problems. An enterprise class security solution and a cloud-based application that meets their cost and reliability needs.

Scalability: Service security is delivered in the cloud, justifying the need for underlying hardware and software resources, enterprises can dial up or dial down their use of the service as their needs change.

Speed: Many times success is defined by being able to move as swiftly as business requires. There is no lag time while new servers, operating systems, and applications are provisioned and installed. Everything is ready to go on demand.

Availability: Validation and ID Protection Service offers carrier-class reliability within the highly secure Symantec global infrastructure, featuring military-grade tier-4 data

centers. The Symantec Internet infrastructure enables and protects up to 30 billion interactions a day, with unmatched scale, interoperability, and security

6. CONCLUSION

As more of their users go mobile and they move critical data and applications into the cloud to achieve cost savings, flexibility, and scalability, enterprises must emphasize security more than ever. Since the numbers of operations are greater for encryption than authentication both in 2G and 3G, throughput for encryption is low compared to authentication as encryption consumes significantly more processing resources compared to authentication. To reduce computational overheads encryption should be used in critical user information only and not for regular traffic flow. Encryption if needed should be combined with authentication. In this case if the message fails authentication, decryption process is saved (not performed). Further the performance analysis determines the cost (in terms of time complexity and throughput). Quantifying the security overhead makes mobile users and mobile network operators aware of the price of added security features and further helps in making optimized security policy configurations. Finally, except for the transformation complexity and the processor capabilities, the real time required for a packet to be protected depends on the overall system load and traffic conditions as well.

7. REFERENCES

- [1] F-secure, "New Century in Mobile Malware", 2006[online].
- [2] William Stallings, "Cryptography and Network Security Principles and Practices, Fourth Edition", 2006
- [3] Shinde Swapnil K et al, "Secure Web Authentication by Multifactor Password a New Approach", International Journal of Software and Web Sciences, 2013.
- [4] Harsh Kumar Sarohi et al, "Graphical Password Authentication Schemes: Current Status and Key Issues", International Journal of Computer Science Issues, March 2013
- [5] Yong ki lee et al, "Secure and Low-cost RFID Authentication Protocols"
- [6] R. Shanmugalakshmi, "Research Issues on Elliptic Curve Cryptography and its Applications", International Journal of Computer Science and Network, June 2009.
- [7] Caimu Tang et al, "An Efficient Mobile Authentication Scheme for wireless networks" IEEE transactions on wireless communications, April 2008
- [8] Kui Ren et al, "A Novel Privacy Preserving Authentication and Access Control Scheme for Pervasive Computing Environments"
- [9] Phillip Rogaway et al, "OCB: A Block-Cipher Mode of Operation for Efficient Authenticated Encryption", August 3 2001.
- [10] S. Creese et al, "Authentication for Pervasive Computing, Security in Pervasive Computing" 2003.

DNA Query Language DNAQL: A Novel Approach

Adekunle Y.A
Department of Computer Science
Babcock University,
Ilishan-Remo, Ogun State
Nigeria

Ogunwobi Z.
Department of Mathematical
Sciences,
Olabisi Onabanjo University,
Ago-Iwoye, Ogun State,
Nigeria

Alao O.D
Department of Computer Science
Babcock University,
Ilishan-Remo, Ogun State
Nigeria

Efuwape B.T
Department of Mathematical
Sciences,
Olabisi Onabanjo University,
Ago-Iwoye, Ogun State,
Nigeria

Ebiesuwa Seun
Department of Computer Science
Babcock University,
Ilishan-Remo, Ogun State
Nigeria

Ainam Jean-Paul
Department of Computer Science
Babcock University,
Ilishan-Remo, Ogun State
Nigeria

Abstract: This paper produces a DNA Query language for querying DNA database in an expressive and concise manner. One of the objectives of this paper was to demonstrate how such a research will be beneficial to biochemistry researchers who are unfamiliar with SQL coding. The paper introduces a new layer between the user application which serves as the interface and the database layer. This layer is then responsible of translating a familiar biochemistry language into a SQL code understandable by the database application. By doing so, the abstraction layer is what is needed to alleviate the use of DNA database by any researcher. Also, a description of common tasks and how they can be translated is given in this paper. Therefore, the novel approach consists of keeping the standard SQL language at the database layer, and yet supplies the same advantages.

Keywords: DNA computing, query language, DNA query language, protein query language and bioinformatics.

1. INTRODUCTION

A Deoxyribonucleic acid (DNA) computer is primarily a nano-computer that makes use of deoxyribonucleic acid to carry out calculations [1].

As a matter of fact, the DNA computers are the next generation microprocessors that make use of the DNA, molecular biology and chemistry instead of the conventional silicon based technologies. DNA computing is a fast developing inter-disciplinary area. The DNA molecules that make up our genes have the ability to perform calculations faster than the world's fastest man-made computers. It has been predicted that DNA might one day be integrated into a computer chip to produce a so-called biochip that will have the capability of pushing the computer even faster.

DNA molecules have already been harnessed to perform complex mathematical calculations. Even while in infancy DNA computers can store billions of time more data than the conventional computers.

A major challenge facing biochemistry and biology researchers is the ability to view relationships among protein data, functions, structures and pathways in a single query or at least in a concise way [2]. For instance, biochemists are performing cutting edge research into carbon-donated hydrogen bonds and their effect on protein structures [3]. In order to do this successfully they need data at the atomic level of the protein to perform calculations. However, no online

database is known to exist that supplies experimental data in an easy-to-use format at the atomic level without parsing the data manually nor do tools exist to facilitate the calculations once data is parsed. To bolster their research chemists have been downloading files from the RCSB in Protein Data Bank (.pdb) format, parsing data manually and loading data into spreadsheets to perform calculations. This approach is wearying and potentially error prone and spreadsheet limitations and other limiting factors obviate the need for a more efficient solution.

2. BACKGROUND

A Deoxyribonucleic acid (DNA) computer is primarily a nano-computer that makes use of deoxyribonucleic acid to carry out calculations [1].

As a matter of fact, the DNA computers are the next generation microprocessors that make use of the DNA, molecular biology and chemistry instead of the conventional silicon based technologies. DNA computing is a fast developing inter-disciplinary area. The DNA molecules that make up our genes have the ability to perform calculations faster than the world's fastest man-made computers. It has been predicted that DNA might one day be integrated into a computer chip to produce a so-called biochip that will have the capability of pushing the computer even faster.

DNA molecules have already been harnessed to perform complex mathematical calculations. Even while in infancy

DNA computers can store billions of times more data than the conventional computers.

A major challenge facing biochemistry and biology researchers is the ability to view relationships among protein data, functions, structures and pathways in a single query or at least in a concise way [2]. For instance, biochemists are performing cutting edge research into carbon-donated hydrogen bonds and their effect on protein structures [3]. In order to do this successfully they need data at the atomic level of the protein to perform calculations. However, no online database is known to exist that supplies experimental data in an easy-to-use format at the atomic level without parsing the data manually nor do tools exist to facilitate the calculations once data is parsed. To bolster their research chemists have been downloading files from the RCSB Protein Data Bank (.pdb) format, parsing data manually and loading data into spreadsheets to perform calculations. This approach is wearying and potentially error prone and spreadsheet limitations and other limiting factors obviate the need for a more efficient solution

3. RELATED WORKS

I-Min A. et al, worked on Advanced Query Mechanisms for Biological Databases. In their paper, they describe generic tools that provide powerful and flexible support for interactively exploring biological database in a uniform and consistent way that is via common data models, formats and notation in the framework of the Object-Protocol Model (OPM). These tools avoid the restriction imposed by traditional fixed-form query interfaces, while providing users with simple and intuitive facilities [4]

Peter Buneman et al [5], worked on Beyond XML Query Languages where they described challenges for Query Languages, the constraints and optimization. In brief, the paper focused on describing the semantics of the underlying semi structured data model, the basic operations on data, the interaction of these operations with constraints, the nature of updates and the problems of generating XML efficiently from existing sources.

Robert Brijder et al, [6] worked on DNA computing and presented a querying language for databases in DNA. In their paper, a set of formal operations on DNA complexes has been defined, much in the spirit of the operations of the relational algebra in the relational data model. The combination of these operations leads to the query language DNAQL.

Sheriff Elfayoumy et al, [7] introduces a Protein Query Language (PQL) for querying protein structures in an expressive yet concise manner, utilizing the work of Patel [8] and Garcia [9]. One of the objectives of their work was to demonstrate how such a language would be beneficial to protein researchers to obtain in-depth protein data from relational database without extensive SQL knowledge. The PQL was an attempt to provide an intuitive declarative language within query application to researchers who are unfamiliar with SQL coding, but their approach still remains primitive and largely procedural. This limits the ease with

which complex queries can be posed and often results in very inefficient query plans.

Sandeep Tata et al, [10] introduced a system called Periscope/SQ based on a well-defined extension of relational algebra – a system that permits declarative and querying on biological sequences. Finally, using a real-world application in eye genetics, they showed how Periscope/SQ can be used to achieve a speedup of two orders of magnitude over existing procedural methods.

Eltabakh et al discussed an extensible database engine for biological databases [12] The proposed engine “extends the functionalities of current DBMSs with (1) annotation and provenance management including storage, indexing, manipulation, and querying of annotation and provenance as first class object in DBMS, (2) local dependency tracking to track the dependencies and derivations among data items, (3) update authorization to support data curation via content-based authorization, in contrast to identity-based authorization, and (4) new access methods and their supporting operators that support pattern matching on various types of compressed biological data type”.

There is a growing and urgent need for declarative and efficient methods for querying biological sequences data [10].

4. METHODOLOGY

4.1 DNA SEQUENCING THEORY

The method of DNA sequencing that will be used in this paper is the enzymatic method which is usually referred to as dideoxy or chain termination sequencing. In this method a short oligonucleotide primer is hybridized to the DNA template that is to be sequenced (Fig.1). A DNA polymerase is then used to initiate DNA synthesis extending from the primer in the 5' to 3' direction. The synthesized DNA is complementary to the template strand of DNA. The reaction contains deoxynucleotides (dNTPs: dATP, dCTP, dGTP, TTP) used by the polymerase to extend the chain. However the reaction also contains a small quantity of dideoxynucleotides (ddNTPs) (Fig.1).

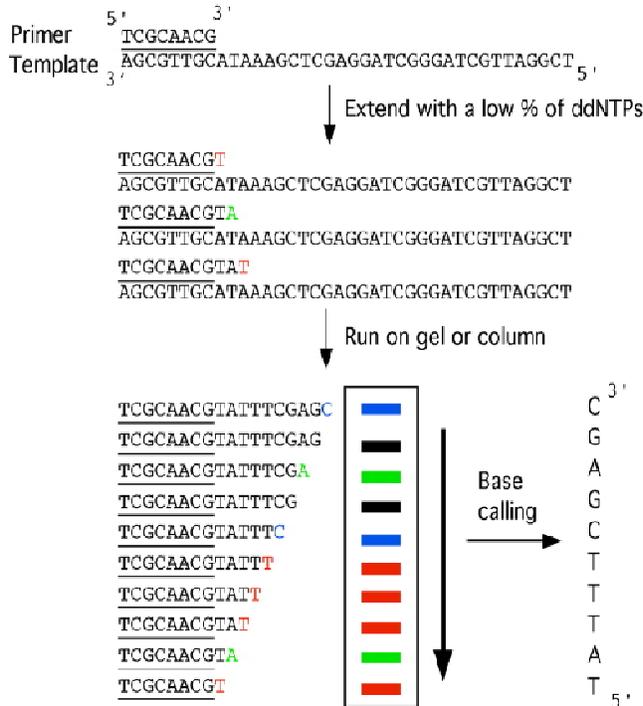


Figure 1: Diagram of DNA sequencing using the chain termination method (Source: Adams G., 2010).

4.2 COMPUTER AS A TOOL FOR DNA SEQUENCE ANALYSIS

In recent time there has been explosion in the number of DNA sequence that has been determined. It is intriguing to know that as DNA sequence information is generated, a problem with the storage and analysis of the vast amounts of information becomes inevitable. This kind of problem is particularly suited to computers. Computers thus serve as tools for handling vast amounts of sequence information produced by molecular biologists.

Computers do not just store sequence information but programs have been written that analyze DNA. For instance, it is important to know where the protein coding sequences are situated on a DNA fragment, what convenient restriction enzyme sites are present in the DNA fragment amongst others. Computers can also be used to determine the similarities between an unknown DNA or protein sequence and a known DNA or protein sequence in the databases.

DNA sequence analysis using computers include searching databases and sequencing databases.

An example of a database entry (dna sequence):
 LOCUS AB231879 1383 bp mRNA linear INV 07-JUN-2006
 DEFINITION Artemia franciscana mRNA for zinc finger protein Af-Zic, complete cds.
 ACCESSION AB231879

VERSION AB231879.1 GI:94966317
 KEYWORDS .
 SOURCE Artemia franciscana
 ORGANISM Artemia franciscana
 Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
 Artemiidae; Artemia.
 REFERENCE 1
 AUTHORS Aruga,J., Kamiya,A., Takahashi,H., Fujimi,T.J., Shimizu,Y.,
 Ohkawa,K., Yazawa,S., Umesono,Y., Noguchi,H., Shimizu,T.,
 Saitou,N., Mikoshiba,K., Sakaki,Y., Agata,K. and Toyoda,A.
 TITLE A wide-range phylogenetic analysis of Zic proteins: Implications
 for correlations between protein structure conservation and body
 plan complexity
 JOURNAL Genomics 87 (6), 783-792 (2006)
 PUBMED 16574373
 REFERENCE 2 (bases 1 to 1383)
 AUTHORS Aruga,J. and Toyoda,A.
 TITLE Direct Submission
 JOURNAL Submitted (10-AUG-2005) Jun Aruga, RIKEN Brain Science Institute,
 Laboratory for Comparative Neurogenesis; 2-1 Hirosawa, Wako-shi,
 Saitama 351-0198, Japan (E-mail:jaruga@brain.riken.jp, URL:http://www.brain.riken.go.jp/labs/lcn/, Tel:81-48-467-9791,
 Fax:81-48-467-9792)
 FEATURES Location/Qualifiers
 source 1..1383
 /organism="Artemia franciscana"
 /mol_type="mRNA"
 /db_xref="taxon:6661"
 gene 1..1383
 /gene="Af-Zic"
 CDS 1..1383
 /gene="Af-Zic"
 /codon_start=1
 /product="zinc finger protein Af-Zic"
 /protein_id="BAE94140.1"
 /db_xref="GI:94966318"
 /translation="MTASLSASVMNPSFIKRESPASATALFVNP QFSAVPNFGFHHVP
 SACATEQSSEMLNPFVDNHLRLNDQSNFQGYHHPHHG
 QIQQHHLGSYAARDFLFRRDM
 GLGMGLEAHHTHAAQHMHMFDPShAAAAAHAMFT
 GFDHNTMRLPTEMYTRDASAA
 QQFHQMGSMAHPASAGAFRLRYMRTPIKQELHCL
 WVDPEQSPKKTCKGKTFGSMHE
 IVTHITVEHVGGPECTNHACFWQGCVRNGRAFKAKYK
 LVNHIRVHTGEKPFPCPPGC
 GKVFARSENKIKHRTHTGEKPFKCEFECDRRFANSS
 DRKKHSHVHTSDKPYNCKVR
 GCDKSYTHPSSLRKHMKVHGKSPPPASSGCDSDENESI
 ADTNSDSAASPPSSHDSSQ
 VQVNHNRPPNHHNLGLGFTNPGHIGDWYVHQSA
 PDM
 VPPATEHSPIGPPMHHPPNSL
 NYFKTELQVN"
 ORIGIN
 1 atgactgcta gtttaagtgc aagcgtgatg aatccaagt ttataagag
 ggaaagtct
 61 gcatcgcta cagccctgtt cgtaccaaac caatttagt cagtgcctaa
 ttttgatt

121 caccatgttc ctagtctgtg tgcaactgag caaagtagtg aatgctgaa
ccctttgtg

4.3 DATABASES

Available databases that can be sequenced are classified into general and organismal specific databases.

General Databases:

- GenBank: DNA sequences (USA database)
- EMBL: DNA sequences (European Molecular Biology Laboratory)
- GenEMBL GenBank and EMBL sequences combined
- DDBJ: DNA sequences (Japan's equivalent of Genbank)
- EST: Expressed Sequence Tags (USA) (DNA sequences)
- STS: Sequence Tagged Sites (USA) (DNA sequences)
- PIR: Protein Identification Resource (protein sequences)
- SwissProt: Protein sequences (Switzerland and EMBL)
- Genpept: Translations of DNA based on authors' information
- PDB: Coordinates for protein 3D structure.

Organismal Specific Databases:

- SGD: Saccharomyces Genomic Database
- YPD: Yeast Protein Database
- WPD: Worm Protein Database
- Wormbase: C. elegans Genome Database
- Sanger: Worm sequence and genomic database
- Flybase: Drosophila sequence and genetic database
- Human: Many

In this paper the GenBank database will be used. The GenBank sequence database is an open access, annotated collection of all publicly available nucleotide sequences and their protein translations. This database is produced and maintained by the National Center for Biotechnology Information (NCBI) as part of the International Nucleotide Sequence Database Collaboration (INSDC). The National Center for Biotechnology Information is a part of the National Institutes of Health in the United States. GenBank and its collaborators receive sequences produced in laboratories throughout the world from more than 100,000 distinct organisms. In the more than 30 years since its establishment, GenBank has become the most important and most influential database for research in almost all biological fields, whose data are accessed and cited by millions of researchers around the world. GenBank continues to grow at an exponential rate, doubling every 18 months. Release 194, produced in February 2013, contained over 150 billion nucleotide bases in more than 162 million sequences. GenBank is built by direct submissions from individual

laboratories, as well as from bulk submissions from large-scale sequencing centers.

4.4 RELATIONAL SCHEMA FOR GenBank

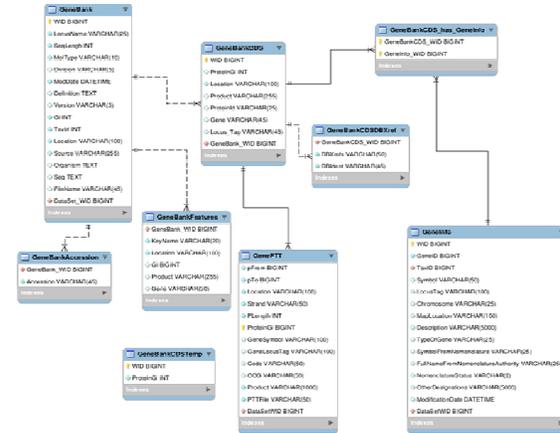


Figure 2: A Framework for Biological Data Integration (Source: JBioWH, 2013)

4.5 DATASET REPOSITORIES

Research chemists around the world do have access to various public DNA data sources, but the access is not designed to support processing and retrieval at the atomic level. Online 'database' supporting biochemistry research include Genbank, EMBL Data Library in the UK, the DNA Bank of Japan (DDBJ), and COLUMBIA [11]. In essence, the only known public access to these databases is via a supplied front-end, and the returned data is formatted for user reading than for storing the data into a databases for further processing and analysis.

5. DNA QUERY LANGUAGE: THE NEW APPROACH

The new approach proposed on this paper is declarative in nature. The new approach provides users with the following feature:

- Users may utilize familiar terms when referring to DNA models and other chemistry terms. The underlying relational model is abstracted from user.
- The ability to use mathematical, Boolean and string functions as part of the language. However, constructs such as conditionals and loping are supported at this time.
- The user shall be able to save DNASQL construct for later utilization.

The approach proposed in this paper consists of defining a layer between the user interface and the DNA database. DNASQL: new approach is an attempt to provide an intuitive declarative language within query application to researchers

who are unfamiliar with SQL coding. Biochemists can use their familiar language and terms within biochemistry domain to query the database. The idea behind is then to define an abstraction layer in charge of translating the familiar biological language in a language comprehensive by the underlying layer; that is the application layer.

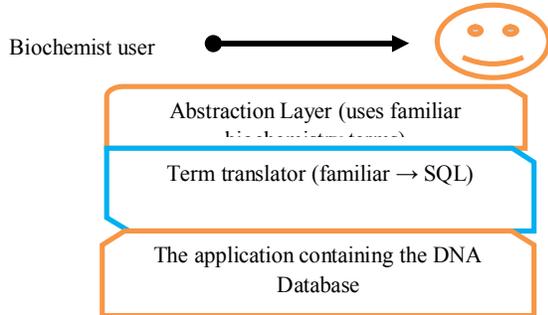


Figure 3: The new approach representation

This paper is concerned about layer 2: the layer between the abstraction layer and the database layer. The abstraction layer provides common tools to the biochemist researcher as simple text editor to input the query while the database layer contains the database itself.

Before describing the SQL language proposed in this paper, let us first describe the following requirement and definitions:

Definition 1: the value of a nucleotide can be obtained by giving its position and is defined as:

Strand (i) where $0 \leq i < n$ with n the length of the DNA strand.

Definition 2: *GetStrand (char) where char belongs to alphabet {A, C, T, G}. The function returns an array of indices where 'char' is found. E.g. for a given strand ACTCAGTA, GetStrand(A) will return Array{0, 4, 7}.*

Definition 3: To obtain a value of nucleotide in a given position, we make use of variable within the function. A variable is defined by preceding the name of the variable by '?'. The function is as follows: *GetStrand(?x, i)* and does the same task as *GetStrand(i)* function but adds the possibility of getting the variable for further usage within the same query. **GetStrand(?x, i)** where x refers to the variable and i the position of the nucleotide within the strand.

Definition 4: To set a nucleotide value in a given position, user can simply make use of *setStrand(x, i)* where x is the value to be inserted and i the position within the strand. This function can easily be used within an update query in a standard SQL language.

Given a table containing record and IDEntry as a primary, the next paragraph describes common tasks used and then translated by the middle layer.

5.1 UPDATE/DELETE/INSERT QUERY

www.ijcat.com

The update statement is used to change **existing** records in a table while delete statement removes one or more records from a table and insert statement adds records to any single table. Either all the rows can be updated, deleted or inserted, or a subset may be chosen using a condition. For the proposed approach, a simple update can be defined as:

```
Modify table
SetStrand(A, 2)
Where IDEntry = 1
```

The SQL equivalent of this query may be:

```
UPDATE table1
SET "get the nucleotide at the second position and then set the value at "= A
Where IDEntry = 1.
```

In this example, the task of getting the nucleotide at the second position may involve another query resulting in a nested query.

Same as the update statement, a delete statement can be expressed as:

```
Delete from table1
Where IDEntry = 1
And an insert as
```

```
Insert into table1:
Set setStrand(i)
```

SELECT QUERY

The select statement returns a result set of records from one or more tables or in other words a select retrieves zero or more rows from one or more database tables or database views. In the context of our proposed model, a select statement is defined as:

```
Retrieve from table1
getStrand(?x, i)
where ?x = A
```

6. CONTRIBUTION TO KNOWLEDGE

This paper introduces a layer between the user and the GenBank database. Unlike the conventional approach where the user accesses the database using Structured Query Language (SQL), this approach proposed in this paper is entirely new and different from all other existing works because it permits the user to make use of the language He or She is well acquainted with and the introduced layer now translates this language into the equivalent SQL query thus carrying out the instruction of the user.

7. SUGGESTION FOR FURTHER RESEARCH

In the future, another researcher can go ahead to implement the proposed approach discussed in this paper which allows the user to use any language He or She is familiar with and this user instruction is now translated into the equivalent SQL query by the introduced layer.

8. CONCLUSION

This paper discussed DNA query language making use of a novel approach. The paper aimed at addressing the problem often times faced by people who want to access databases but lack the requisite understanding of the fundamental SQL which is the language GenBank (the database used in this paper) is based on. The proposed approach makes it possible for users with little or no knowledge of the SQL to issue instructions using any language they are familiar with and these instructions get converted into SQL query by the layer introduced in this paper. Hence, access to databases becomes easier, less strenuous and convenient.

9. REFERENCES

- [1]. http://www.webopedia.com/TERM/D/DNA_computer.html Retrieved On Feb 5th, 2010.
- [2]. M.I. Jaya, Z. Zainol, and N.H. Malim (2007) “iProt – A Data Warehouse for Protein Database”; International Conference on Electrical Engineering and Informatics, Institut Teknologi Bandung, Indonesia.
- [3]. K. Compaan, R. Vergenz, P. Von Rague Schleyer, and I. Arreguin (2008) “Carbon-donated Hydrogen Bonding: Electrostatics, Frequency Shifts, Directionality, and Bifurcation”; International Journal of Quantum Chemistry, Vol. 108, No. 15, 2914–2923.
- [4]. I-Min A. Chen, Anthony S. Kosky, Victor M. Markowitz, Ernest Szeto, and Thodoros Topaloglou, “Advanced Query Mechanisms for Biological Databases”, Bioinformatics Systems Division, Gene Logic Inc., 2001.
- [5] Peter Buneman, Alin Deutsch, Wenfei Fan, Hartmut Liefke, Arnadu Sahuguet, Wang-chiew Tan, “Beyond XML Query Languages”, University of Pennsylvania, November 1998.
- [6] Sherif Elfayoumy and Paul Bathen, “Protein Query Language: A Novel Approach”, School of Computing, University of North Florida, Jacksonville, FL, USA.
- [7] J.M. Patel, D.P. Huddler, and L. Hammel. “Declarative and Efficient Querying on Protein Secondary Structures”; Data Mining in Bioinformatics, pp. 243 – 273, 2005.
- [8]M.M. Roldan-Garcia, J.J. Molina-Castro, and J.F. Aldana-Montes. ‘ECQ: A Simple Query Language for Semantic Web’; Processing of DEXA, Turin, Italy, 2008.
- [9] Sandeep Tata, J.M. Patel, J.S. Friedman, A. Swaroop, “Declarative Querying for Biological Sequence Databases”, appears in IEEE ICDE 2006.
- [10] S. Trissl, K. Rother, H. Mueller, T. Steinke, I. Koch, R. Preissner, C. Froemmel, and U. Leser. “Columba: an Integrated Database of Proteins, Structures, and Annotations”; BMC Bioinformatics, Vol. 8, No. 81, 2005.
- [11]M.Y. Eltabakh, M. Ouzzani, M. Aref, “BDBMS- A Database Management System for Biological Data”; Third Biennial Conference on Innovative Data Systems Research (CIDR), 2007.
- [12] M.Y. Eltabakh, M. Ouzzani, M. Aref, A.K. Elmagarmid, Y. Laura-Silva, M.U. Arshad, D. Salt, and I. Baxter. “Managing Biological Data Using BDBMS”; The IEEE 24th International Conference on Data Engineering, 200.

An Adaptative Accident Prevention Technique for Mobile Communications Systems

Banda Srinivas Goud
Vardhaman College of Engineering
Hyderabad,India

Eliyaz Mahammed
Vardhaman College of Engineering
Hyderabad,India

Abstract: This paper introduces how many times our elders told about the ill-effects of using cell phones while driving, but how many of us have taken their advices seriously. Well we think not even thirty percent. Isn't it? But whether we like it or not, it is one of the major reasons for number of accident that are happening nowadays. Even doctors have now recently stated that talking on cell phones while driving is as fatal as driving our car after drinking. It can lead to various disastrous major miss-happenings. Do you know that using a mobile phone while driving can affect the cognitive functions of a person, distract his or her visual concentration and also the speed of processing information Are you feeling Scared after reading this? Well it has been also proved that use of cell phones while driving puts a driver at a significantly higher risk of collision by distracting his or her mind. It hardly matters whether the person makes use of hands free or hand-held phones, there's no escape to it. This deadly combo has significantly increased the risk of accidents in large numbers.

Keywords: Vehicle , GSM, RF MODULE.

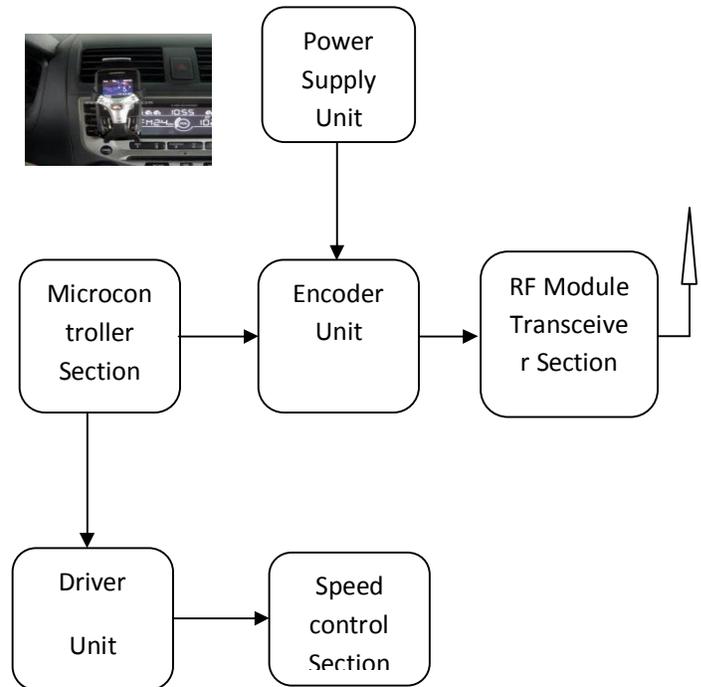
1. INTRODUCTION

In this we Propose a concept for overcome this, it consists of two devices former the vehicle section once we start the ignition the RF signal transfer's from the vehicle to mobile immediately it changes to Driving mode. If mobile receives the call, it attends and gives the voice information, if the call is emergency only means it will inform the driver through ringing.

Certainly there has been large number of figures that shows that people have used their mobile phones just before their accidents. Studies have shown that if people reduce their usage of cell phone while driving, it can cut off the accident rates too.

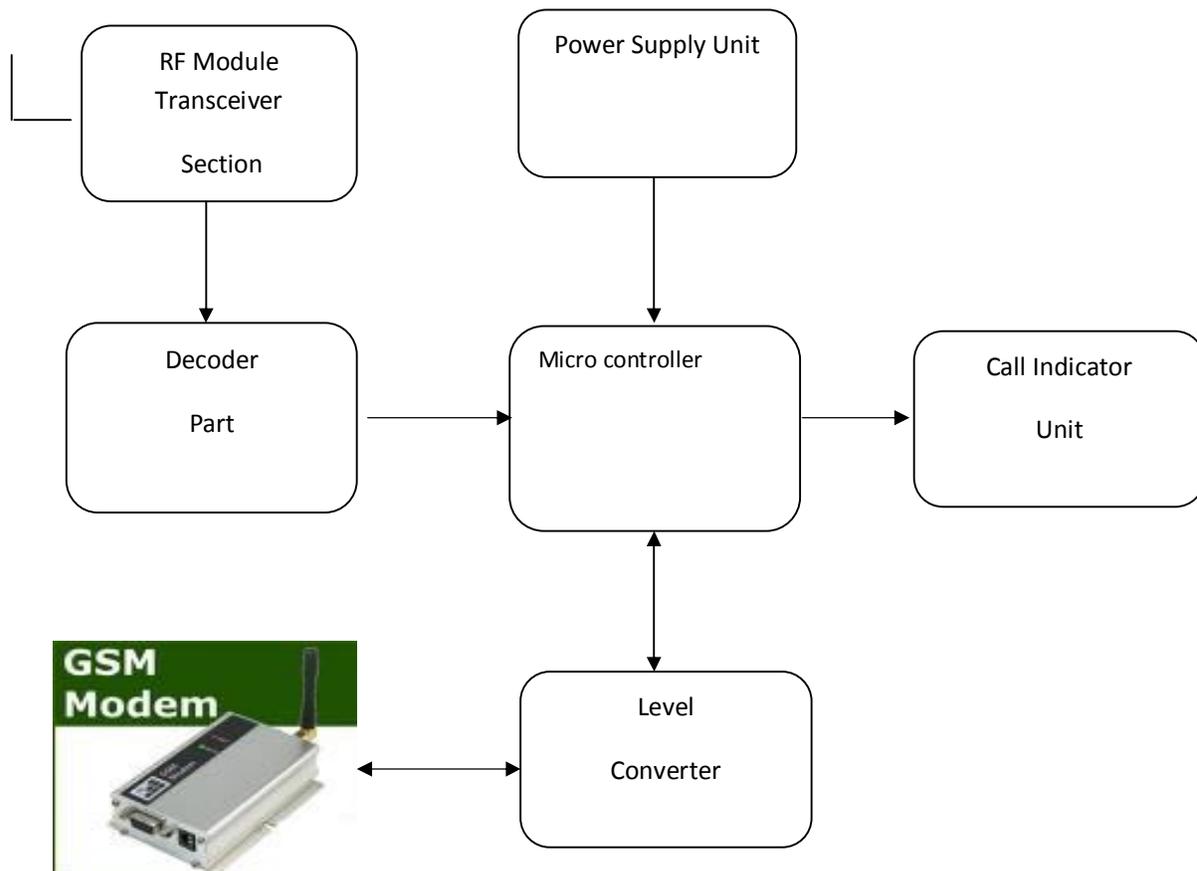
2. BLOCK DIAGRAM

2.1 Dash Board Section:



2.2 Mobile gadget:

phone. folks will check the message and understand once their youngsters



3. DESIGN AND IMPLEMENTATION

Traditionally, the protection of kindergartens is very relied on the human's effort and a spotlight like guardians and academics. However, if there's no active informing service give by kindergartens typically, parents haven't any plan of once and whether or not their children safely arrive in their schoolroom when they're picked up by motor coach. each morning the coed attendance offers the primary hand data of kids safety. however generally it's sophisticated to precisely track the attending since the youngsters arrive during a period of your time within the morning and a few of them return with their folks and a few return by motor coach. So we develop an energetic RFID attending system to overcome the barriers and mistakes of manually taking attendance and mix the wireless GSM message service to produce real time responses to their parents' cellular

square measure safely arrived. On the contrary, a noticing message are going to be broadcasted to administrators and oldsters if the youngsters don't show up during a category on time in order that the adults have the enough time to examine out these specific youngsters and stop accident happens.

4. SYSTEM HARDWARE

4.1. ARM Processor:

The ARM7 family includes the ARM7TDMI, ARM7TDMI-S, ARM720T, and ARM7EJ-S processors. The ARM7TDMI core is the industry's most widely used 32-bit embedded RISC microprocessor solution. Optimized for cost and power-sensitive applications, the ARM7TDMI solution provides the low power consumption, small size, and high

performance needed in portable, embedded applications. The ARM7TDMI core uses a three-stage pipeline to increase the flow of instructions to the processor. This allows multiple simultaneous operations to take place and continuous operation of the processing and memory systems.

4.2 Operating modes

The ARM7TDMI core has seven modes of operation: User mode is the usual program execution state Interrupt (IRQ) mode is used for general purpose interrupt handling Supervisor mode is a protected mode for the operating system Abort mode is entered after a data or instruction pre fetch abort System mode is a privileged user mode for the operating system Undefined mode is entered when an undefined instruction is executed. The interrupt setting of ARM supports the DHLS to response to the interrupt coming from the server section.

4.3 Interrupt controller

The Vectored Interrupt Controller (VIC) accepts all of the interrupt request inputs from the home server section and categorizes them as Fast Interrupt Request (FIQ), vectored Interrupt Request (IRQ), and non-vectored IRQ as defined by programmable settings. So DHLS system can able to separate the command signals and easily will select the priority. The programmable assignment scheme means that priorities of interrupts from the various peripherals can be dynamically assigned and adjusted. Fast interrupt request (FIQ) has the highest priority. If more than one request is assigned to FIQ, the VIC combines the requests to produce the FIQ signal to the ARM processor. The fastest possible FIQ latency is achieved when only one request is classified as FIQ, because then the FIQ service routine does not need to branch into the interrupt service routine but can run from the interrupt vector location. If more than one request is assigned to the FIQ class, the FIQ service routine will read a word from the VIC that identifies which FIQ source(s) is (are) requesting an interrupt.

Vectored IRQs have the middle priority. Sixteen of the interrupt requests can be assigned to this category. Any of the interrupt requests can be assigned to any of the 16 vectored IRQ slots, among which slot 0 has the highest priority and slot 15 has the lowest. Non-vectored IRQs have the lowest priority. The VIC

combines the requests from all the vectored and non-vectored IRQs to produce the IRQ signal to the ARM processor. The IRQ service routine can start by reading a register from the VIC and

jumping there. If any of the vectored IRQs are pending, the VIC provides the address of the highest-priority requesting IRQs service routine, otherwise it provides the address of a default routine that is shared by all the non-vectored IRQs. The default routine can read another VIC register to see what IRQs are active.

4.4 PIC micro controller:

The was originally designed around 1980 by General Instrument as a small, fast, inexpensive embedded microcontroller with strong I/O capabilities. PIC stands for "Peripheral Interface Controller". General Instrument recognized the potential for the little PIC and eventually spun off Microchip, headquartered in Chandler, AZ to fabricate and market the PICmicro.

The PIC has some advantages in many applications over the older chips such as the Intel 8048/8051/8052 and many others. Its unusual architecture is ideally suited for embedded control. Nearly all instructions execute in the same number of clock cycles, which makes timing control much easier. This is a RISC (Reduced Instruction Set Computer) design, with only thirty-odd instructions to remember; its code is extremely efficient, allowing the PIC to run with typically less program memory than its larger competitors.

Very important, though, is the low cost, high available clock speeds, small size, and incredible ease of use of the tiny PIC. For timing-insensitive designs, the oscillator can consist of a cheap RC network. Clock speeds can range from low speed to 20MHz. Versions of the various PIC microcontroller families are available that are equipped with various combinations ROM, EPROM, OTP (One-Time Programmable) EPROM, EEPROM, and FLASH program and data memory. An 18-pin PIC microcontroller typically devotes 13 of those pins to I/O, giving the designer two full 8-bit I/O ports and an interrupt. In many cases, designing with a PIC microcontroller is much simpler and more efficient than using an older, larger embedded microprocessor.

5 ARCHITECTURE OVERVIEW

The PIC uses Harvard architecture, unlike the von Neumann architecture used in most general-purpose processors. The von Neumann architecture uses the same bus for program memory, data memory, I/O, registers, etc. This makes it easy to bring the common bus out to device.

I/O pins for adding memory, but it limits the bus bandwidth that can be used for any one function since the bus is shared. Von Neumann processors are generally micro coded, **CISC** (Complex Instruction Set Computer) designs (though there are, of course, exceptions).

The Harvard architecture uses separate program memory and data memory busses. This makes it easy to design the processor for very efficient use of program memory, since the program memory bus can be of a much different width than the data memory. Instructions usually (always in the case of the PIC) take up only one program memory location, compared to one, two or even three in a typical von Neumann design. Harvard-architecture machines are generally non-micro coded, **RISC** (Reduced Instruction Set Computer) designs (again, exceptions are to be found). One drawback to the Harvard architecture is that it is very difficult to bring the memory address and data busses out to device pins, so adding external program memory is difficult at best. For this reason, most Harvard machines have only internal program memory.

For example, the popular PIC16F84 contains 1K words of FLASH program memory, 68 bytes of data RAM, and 64 bytes of data EEPROM. While this seems like an extremely limited amount of code and data space, the PIC's incredibly compact code makes the most of it. 1024 instruction word memory actually means 1024 instructions, no less. Even immediate-mode instructions, where an operand is part of the instruction itself, takes only one memory location, as do CALL and GOTO instructions. There even exists a single-chip implementation of a TCP/IP stack and HTTP server written for a 16F84.

The PIC is also a non-micro coded design. In larger processors, each binary machine language instruction often is "executed" by a series of microcode steps. While this is a great approach for building large, complex processors with a wide range of instructions, it also leads to great complexity and takes up a lot of real estate. The PIC uses the instruction word itself,

decoded by logic gates as it is read from program memory, to control the flow of data through the chip.

The seemingly odd 14-bit instruction word length is a direct result of the internal architecture of the processor itself. In the case of the 16F84 or 16C711, we need 13 bits just to address all of program memory. In the case of the smaller 16C54 with only 512 words of program memory and 25 bytes of RAM, we can get by with a 12-bit instruction word -- which is exactly what the 16C54 uses. Conversely, with more memory we would use a longer instruction word, like the 16 bits in the 18Cxxx family.

6 DRIVER UNIT

A relay is an electrically operated switch. Current flowing through the coil of the relay creates a magnetic field which attracts a lever and changes the switch contacts. The coil current can be on or off so relays have two switch positions and they are double throw (changeover) switches.

Relays allow one circuit to switch a second circuit which can be completely separate from the first. For example a low voltage battery circuit can use a relay to switch a 230V AC mains circuit. There is no electrical connection inside the relay between the two circuits, the link is magnetic and mechanical.

Relays are very simple devices. There are four major parts in every relay. They are

- Electromagnet
- Armature that can be attracted by the electromagnet
- Spring
- Set of electrical contacts

7 WORKING

When a current flows through the coil, the resulting magnetic field attracts an armature that is mechanically linked to a moving contact. The movement either makes or breaks a connection with a fixed contact. When the current to the coil is switched off, the armature is returned by a force approximately half as strong as the magnetic force to its relaxed position. Usually this is a spring, but gravity is also used commonly in industrial motor starters. Most relays are manufactured to operate quickly. In a low voltage application, this is to reduce noise. In a high voltage or high current application, this is to reduce arcing.

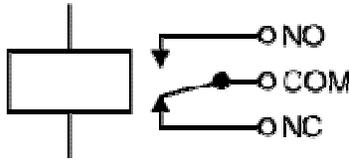


Fig 8.1 Circuit symbol of a relay

The relay's switch connections are usually labelled COM, NC and NO:

- COM = Common, always connect to this, it is the moving part of the switch.
- NC = Normally Closed, COM is connected to this when the relay coil is off.
- NO = Normally Open, COM is connected to this when the relay coil is on.

NOTE: Connect to COM and NO if you want the switched circuit to be on when the relay coil is on. Connect to COM and NC if you want the switched circuit to be on when the relay coil is off.

7.1 Advantages of relays:

Like relays, transistors can be used as an electrically operated switch. For switching small DC currents (< 1A) at low voltage they are usually a better choice than a relay. However transistors cannot switch AC or high voltages (such as mains electricity) and they are not usually a good choice for switching large currents (> 5A). In these cases a relay will be needed. Advantages of relays compared to other switching devices are:

- The complete electrical isolation improves safety by ensuring that high voltages and currents cannot appear where they should not be.
- Relays can switch many contacts at once. Relays come in all shapes and sizes for different applications and they have various switch contact configurations. Double Pole Double Throw (DPDT) relays are common and even 4-pole types are available. You can therefore control several circuits with one relay or use one relay to control the direction of a motor.
- Relays can switch AC and DC, transistors can only switch DC.
- Relays can switch high voltages, transistors cannot.

8 GSM MODEM:

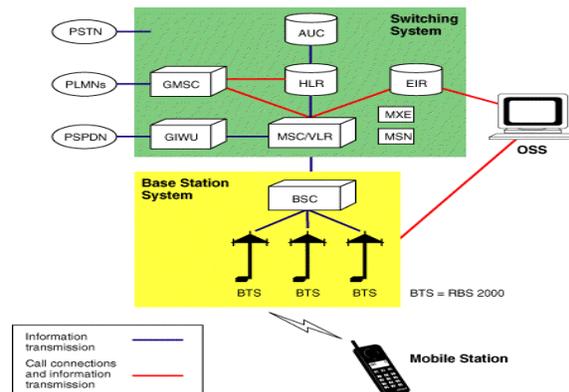
8.1 Definitions:

Global system for mobile communication (GSM) is a globally accepted standard for digital cellular communication. GSM is the name of a standardization group established in 1982 to create a common European mobile telephone standard that would formulate specifications for a pan-European mobile cellular radio system operating at 900 MHz's

8.2 The GSM network:

GSM provides recommendations, not requirements. The GSM specifications define the functions and interface requirements in detail but do not address the hardware. The reason for this is to limit the designers as little as possible but still to make it possible for the operators to buy equipment from different suppliers. The GSM network is divided into three major systems: the switching system (SS), the base station system (BSS), and the operation and support system (OSS). The basic GSM network elements are shown in below

figure



GSM Network Elements

8.3 GSM modem:

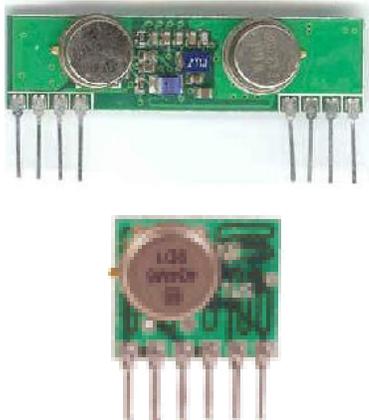
A GSM modem is a wireless modem that works with a GSM wireless network. A wireless modem behaves like a dial-up modem. The main difference between them is that a dial-up modem sends and receives

data through a fixed telephone line while a wireless modem sends and receives data through radio waves.

8.4 RF communication:

Radio Frequency, any frequency within the electromagnetic spectrum associated with radio wave propagation. When an RF current is supplied to an antenna, an electromagnetic field is created that then is able to propagate through space. Many wireless technologies are based on RF field propagation.

Radio Frequency: The 10 kHz to 300 GHz frequency range that can be used for wireless communication. Also used generally to refer to the radio signal generated by the system transmitter, or to energy present from other sources that may be picked up by a wireless receiver.



- Wireless mouse, keyboard
- Wireless data communication
- Alarm and security systems
- Home Automation, Remote control
- Automotive Telemetry
- Intelligent sports equipment
- Handheld terminals, Data loggers
- Industrial telemetry and tele-communications
- In-building environmental monitoring and control
- High-end security and fire alarms

8.5 Transmitter:

The TWS-434 extremely small, and are excellent for applications requiring short-range RF remote controls. The transmitter module is only 1/3 the size of a standard postage stamp, and can easily be placed inside a small plastic enclosure.

TWS-434: The transmitter output is up to 8mW at 433.92MHz with a range of approximately 400 foot (open

area) outdoors. Indoors, the range is approximately 200 foot, and will go through most walls.



Figure – 7.1.2

The TWS-434 transmitter accepts both linear and digital inputs can operate from 1.5 to 12 Volts-DC, and makes building a miniature hand-held RF transmitter very easy. The TWS-434 is approximately 1/3 the size of a standard postage stamp.

9 CONCLUSION

This deadly combo has significantly increased the risk of accidents in large numbers. Certainly there has been large number of figures that shows that people have used their mobile phones just before their accidents. Studies have shown that if people reduce their usage of cell phone while driving, it can cut off the accident rates too. So next time if your cell rings, make sure to answer your phone call only after when you reach your destination and not in-between the way.

10 REFERENCES:

- [1]. A. X. Liu and L. A. Bailey, "PAP: A privacy and authentication protocol for passive RFID tags," *Comput. Commun.*, vol. 32, pp. 1194-1199, 2009.
- [2] .J. A. Wolff, "RFID tags – an intelligent bar code replacement," *IBM Corporation*, 2001.
- [3]. R. Weinstein, "RFID: a technical overview and its application to the enterprise," *IT Professional*, vol. 7, pp. 27 - 33, May-June 2005
- [4] .A. Juels, "RFID Security and Privacy: A Reasearch Survey," *RSA Laboratories*, 28 September 2005.
- [5]. G. Venkataramani and S. Gopalan, "Mobile phone based RFID architecture for secure electronic Payments using RFID credit cards," *Availability, Reliability and Security*, 2007.

ARES 2007. The Second International Conference on, pp. 610-620, 10-13 April 2007.

[6]. Mikko Lehtonen, Thorsten Staake, Florian Michahelles, and E. Fleisch, "From Identification to Authentication—A Review of RFID Product Authentication Techniques," *Printed handout of Workshop on RFID Security - RFIDSec*, Springer, 2006.

[7]. N. W. Lo and K.-H. Yeh, "Novel RFID Authentication Schemes for Security Enhancement and System Efficiency," *Lecture Notes in Computer Science, Secure Data Management*, vol. 4721/2007, pp. 203-212, 2007.

