



IMTISAL KHAN
01-134131-035
ASAD ULLAH KHAN
01-134131-015

Terrorism Portal - Pakistan

Bachelor of Science in Computer Science

Supervisor: Dr. Arif Ur Rahman

Department of Computer Science
Bahria University, Islamabad

December 2016

Certificate

We accept the work contained in the report titled “Terrorism Portal Pakistan”, written by Mr. Imtisal Khan and Mr. Asad Ullah Khan as a confirmation to the required standard for the partial fulfillment of the degree of Bachelor of Science in Computer Science.

Approved by . . . :

Supervisor: Dr. Arif Ur Rahman (Assistant Professor)

Internal Examiner: Syed Khushal Shah (Assistant Professor)

External Examiner: Dr. Muhammad Sajjad Khan (Assistant Professor)

Project Coordinator: Dr. Arif Ur Rahman (Assistant Professor)

Head of the Department: Dr. Faisal Bashir (Associate Professor)

December 13th, 2016

Abstract

In this project firstly the user gives the requirements, crawler will start crawling with the help of base URL and the requirement which the user has given, as a result the crawler will extract and parse the information related to the user requirement and store it in database. Then the system will extract the parsed data from the Database which exactly matches with the user query and a full-fledged report is generated in the form of Graph/Charts and in description which contain the answer of the user's query.

Acknowledgments

First and foremost, this project would not have been possible without Allah Almighty's help. Very special thanks to our beloved Prophet Muhammad (PBUH) who brought us the real source of knowledge the Holy Quran.

We are profoundly grateful to our Supervisor Mr. Arif Ur Rahman whose help and encouragement helped us during our project.

The constant support provided to us by our families including specially our parents, who gave us the confidence to complete our Project.

IMTISAL KHAN & ASAD ULLAH KHAN
Islamabad, Pakistan

December 2016

“Put your heart, mind, and soul into even your smallest acts. This is the secret of success”

Swami Sivananda

Contents

Abstract	i
1 Introduction	1
1.1 Problem statement	2
1.2 Scope of new system	2
1.3 Proposed solution	2
1.4 System Overview	2
1.5 Methodology	2
1.5.1 Data gathering	3
1.5.2 Database management	3
1.5.3 Report generation	3
1.6 Tools and technologies	3
1.7 System Requirement	4
1.7.1 Hardware Requirements	4
1.7.2 Software Requirements	4
2 Literature Review	5
2.1 Functional Requirement	5
2.2 Non Functional Requirements	6
3 Requirement Specifications	7
3.1 EXISTING SYSTEM	7
3.2 Drawbacks	7
3.2.1 No Terrorism Crawler for Pakistan	7
3.2.2 Lack of data repository	8
3.2.3 Difficulty in analyzing terrorism activities	8
3.2.4 Data Duplication	8
3.3 Proposed System	8
3.3.1 Terrorism Portal for Pakistan	8
3.3.2 Data Repository	9
3.3.3 Analyzing the terrorist activities	9
3.4 Requirement Specifications	9
3.5 Functional Requirement	9
3.6 Non Functional Requirements	10
3.7 Usecase Diagram	10

4	Design	11
4.1	System Architecture	11
4.1.1	Data Scraping	12
4.1.2	Data Manipulation	12
4.1.3	Database Management	12
4.1.4	Terrorism Portal	12
4.2	Design Constraints	12
4.2.1	System Dependencies	12
4.2.2	System Assumptions	12
4.3	Design Methodology	13
4.3.1	Hardware & Software Requirements	13
4.4	High Level Design	13
4.4.1	Activity Diagram	13
4.4.2	Context Diagram	14
4.4.3	Data Flow Diagram	14
4.4.4	State Diagram	14
4.4.5	Package Diagram	15
4.5	Low Level Design	15
5	System Implementation	21
5.1	System Architecture	21
5.2	Tools and Technology Used	21
5.3	Development Environment & Languages Used	21
5.3.1	Eclipse	22
5.3.2	JAVA	22
5.3.3	XAMPP	22
5.3.4	HTML5	22
6	System Testing and Evaluation	25
6.1	Graphical user interface testing	25
6.1.1	Main Interface	25
6.1.2	Result Output	25
6.1.3	Graph Output	27
6.2	Usability testing	27
6.3	Performance testing	28
6.4	Compatibility Testing	28
6.5	Installation testing	28
6.6	Test Cases	29
6.6.1	Test Case1: Get requirements	29
6.6.2	Test Case2: Reading Web-pages	29
6.6.3	Test Case 3: Data Extraction	29
6.6.4	Test Case4: Saving Extracted Results	29
6.6.5	Test Case5: View Results	30
7	Conclusion	35
7.1	Good Features	35
	References	37

List of Figures

1.1	Methodology	3
3.1	Use Case Diagram	10
4.1	Project module	11
4.2	Activity Diagram	14
4.3	Context Diagram	15
4.4	Data Flow Diagram	16
4.5	State Diagram	17
4.6	Package Diagram	18
4.7	Architecture Diagram	19
5.1	System Architecture	23
6.1	Main Interface	26
6.2	Result Output	27
6.3	Graph Output	28
6.4	Get Requirements	29
6.5	Reading Web Pages	30
6.6	Data Extraction	31
6.7	Save Result	32
6.8	Show result	33

Chapter 1

Introduction

Computer Science (CS) has played an important role in our modern world. The role of information technology continues due to which communication and our life have become much easier. CS deals with software and hardware that are used to manage information these can be used for storing data, retrieving data and retrieving the required information.

We are living in the environment in which the things are changing in every second and more advancements are coming in every field. Everything is on our screens, just on one click we can get our required information. The reason behind this is information technology, due to information technology everything is on our finger tips and we are able to connect with the whole world. CS has evolved different fields of life.

By using the advancements of CS, students and teachers, are more connected in and out of the classroom. Information technology provides teachers an endless choice of multimedia, software, applications and devices with which to create more exciting, interactive lessons. Online courses and web-based training provided the facility to the students to learn new courses according to their own schedule.

CS provided the benefits to the business world by allowing the organizations to work more efficiently and to maximize productivity. A major advantage of CS in the field of business is e-commerce. E-commerce are expanding locally and globally without the barriers of distance, time and language.

Web crawler is a computer program used by search engine to search URLs of the required data from different websites or internet pages. Search engines like Google and Yahoo, use Web crawler to gather data which is available on different Web pages. So the basic purpose of using Web crawler is to search the specific data from different Websites.

1.1 Problem statement

There is lack of a data repository that may contain data for different terrorism related activities in Pakistan. So analyzing such activities is also difficult. There is a need of getting data from the different sources from the net and populate it in a database that can be used for analyzing the data.

1.2 Scope of new system

The scope of this system is that just on one click we can retrieve all the information related to all the terrorist acts. The reason involved behind that activity, and the nominated terrorists.

1.3 Proposed solution

We will design and Implement a Web crawler that will bring the information from the sites that may contain data about different terrorism activities in Pakistan. Then that system can be used for analysis purpose.

Web crawler used by search engine and as it extract the required information, so our Web crawler will extract all the information which contain the data related to the terrorism. After having all the data related to terrorism we will provide the data management system to analyze the terrorism activities. Our system will provide the useful data that will be further used for analysis.

1.4 System Overview

Our system consists of three modules that are:

- Data Gathering
- Database Management
- Report generation for decision making

1.5 Methodology

Methodology of our system as shown in Figure 1.1.

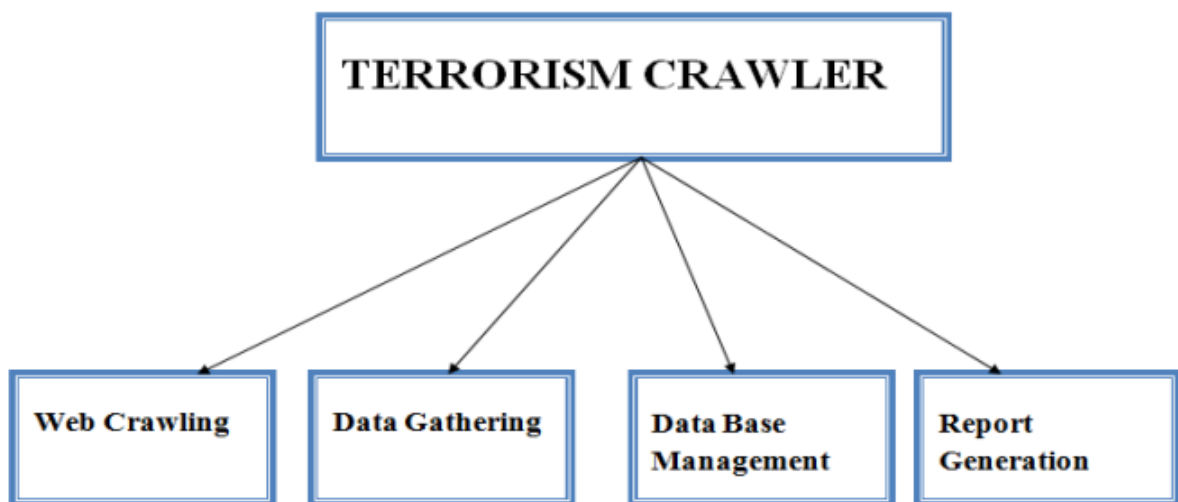


Figure 1.1: Methodology

1.5.1 Data gathering

Data gathering is the process in which Jsoup library extracts the data from the given websites and store it in database.

1.5.2 Database management

Database management is used for creating a database that will keep the record of all the data fetched by the crawler.

1.5.3 Report generation

Report generation helps to reply different queries and requests of relevant data.

1.6 Tools and technologies

Following are the tools and technologies used to build the system:

- Eclipse Luna (4.4) with JDT Java 7.0 or above.
- Database Server: MySQL
- Browser: Internet Explorer 9 and above, latest versions of Mozilla, Safari, Opera & Google Chrome (for Client Side browsing)
- Programming Language, Basic: Java, HTML5.

1.7 System Requirement

The system requirements are as under

1.7.1 Hardware Requirements

- Processor: Pentium 4 or above
- Ram: 1 GB
- Disk Space: 40GB

1.7.2 Software Requirements

- Windows
- Java
- XAMPP(MySQL, Apache, TomCat)
- Eclipse

Chapter 2

Literature Review

System analysis is one of the phases in developing software. It is the analysis of the components of the system how they interact with each other and work. In system analysis we add, modify, and delete the components toward achievement of our goal. System analysis describes the early phases of system development. System Analysis helps making better decision and adopt better course of action. System analysis is the process which involves first investigation of the system second identifying the problem and lastly using the information and the recommend the solution of the problem for improvement. This chapter includes functional non functional requirement, use case analysis.

2.1 Functional Requirement

Functional requirements get the projected behavior of the system, which is expressed as the tasks services or functions the system is required to do. Functional requirements describe the basic functionality of the system. The functional requirements of terrorism portal pakistan are as follow:

Load website: The website is loaded by the system into the browser.

Search: If the user wants to search something related to terrorism in Pakistan then the user must go to search tab and select the requirement from the drop down list.

Requirements: The user select the requirements from the drop down lists and click the search button.

Reading Web Pages: The crawler Reads the Web Pages and will extract the relevant

data.

Data Extraction: Crawler extracts the relevant description.

Saving Extracted Result: The extracted data is saved in the database.

Result: When user select the given options the result will displayed in Graph along with Description.

2.2 Non Functional Requirements

Non-functional requirement is a requirement that specifies the process that can be used to judge the operation of a system. Nonfunctional requirements define how easy the software is to use? how quickly it executes? how reliable it is? and how well it behaves when unexpected conditions arise.

Availability: A system's availability is the amount of time the system is available for the use. This system will be available all time for the user.

Efficiency: This system is very efficient.

Flexibility: If the organization wants to increase or change the functionality of the software after it is developed, that should be planned from the beginning; it depend on the choices made during the design, development and testing, of the system. This system is very much flexible.

Performance: The performance depends upon the characteristics of time. The performance of this system is good.

Reliability: Reliability specifies the capability of the software to maintain its performance and perform all its functions properly. This system is quite reliable.

Usability: Usability is the ease-of-use. This system is user friendly.

Chapter 3

Requirement Specifications

3.1 EXISTING SYSTEM

Most terror related information is sparsely available on internet. While sites like South Asia Terrorism Portal, Global Terrorism Database, and Wikipedia do provide an interface to view available information, trying to search through that information for an incident happening in a certain city in a certain month in Pakistan is rather impossible. For finding such information, one needs to traverse through loads of information that is both inconvenient and at times wastes too much time and energy [1]. Also important here is to note the fact that these sites are not dedicated/exclusive to/for terror related incidents in Pakistan and hence at times the information available is not credible, complete and lacks references which puts question marks on its authenticity [2].

3.2 Drawbacks

Following are the Drawbacks:

3.2.1 No Terrorism Crawler for Pakistan

There is no terrorism crawler for Pakistan so the information can be gathered through radio, television and through internet. The information which user gets through radio is useless in this manner that user can only listen it and after that the one who listens about the activity have the information about it the information cannot be shared among millions. The same is the case with the information user get from the television we cannot store the information somewhere, where it is visible to everyone. On the internet the information is there for everyone and the information is accessible by anybody but the flaw is that when

the user searches for the information of any terrorist activity then the irrelevant material also comes.

3.2.2 Lack of data repository

There is a lack of data repository which only stores the information related to the terrorist activities held in Pakistan. There is no database maintained which keeps the record of all the terrorist activities in Pakistan.

3.2.3 Difficulty in analyzing terrorism activities

As there is no proper database maintained which keeps the record of all the terrorist activities held in Pakistan so analyzing the terrorism activity is also difficult.

3.2.4 Data Duplication

When user searches on the internet there is a lot of duplication of information which is retrieved and creates problem for the user and it may consume the memory and it is also wastage of time.

3.3 Proposed System

While there are plethora of sites available all over the internet, none of them is dedicated for information pertaining to Pakistan. Hence creating a dire need for a portal that provides information that is credible, authentic, and complete and is relevant to Pakistan.

3.3.1 Terrorism Portal for Pakistan

Our proposed Terrorism Portal will provide a centralized data repository for all Terror related incidents, dating back to 2002. After properly analyzing different sites, we have shortlisted following forums that we will use to crawl/scrap data from:

- Wikipedia
- South Asia Terrorism Portal
- Global Terrorism Database

Our decision to choose these sites from a variety of available options was governed largely by the factor of sites credibility, data authenticity, and completion that these sites provide and others don't.

3.3.2 Data Repository

Our System will scrap these sites using a highly intelligent algorithm that will ensure there is no data duplicity. Once data scraping is complete, the system will manipulate/format data to find common attributes and to formalize similar attributes before storing them in our centralized database.

3.3.3 Analyzing the terrorist activities

Centralized data repository will allow journalists, media houses, civilian forces, researchers, students and general public to track terror related incidents, match patterns and do other useful/relevant activities using this data.

3.4 Requirement Specifications

Requirement Specification includes functional and non-functional requirements of the system, which are as follows:

3.5 Functional Requirement

Functional requirements get the projected behavior of the system, which is expressed as the tasks services or functions the system is required to do. Functional requirements describe the basic functionality of the system [3].The functional requirements of terrorism portal pakistan are as follow:

Data Scraping: The system is using Jsoup library to scrap data from the sites.

Data Manipulation: Scrapped data will be manipulated by performing data parsing, data conversations and proper formatting shall be applied on that data to avoid data duplications and to refine data in a way to find common attributes among scraped data coming from different site.

Database Management: Properly formatted data will be stored in a centralized data repository that will act as the backbone for our terrorism portal.

Terrorism Portal: A modern day website that will provide advanced level data searching facility to its users. Search criteria will use fields such as Year, City, and Incident Type etc. to provide relevant users flexibility to narrow down target results in an intelligent and efficient way. The same information can be utilized by relevant persons, agencies etc. to form patterns and draw conclusions.

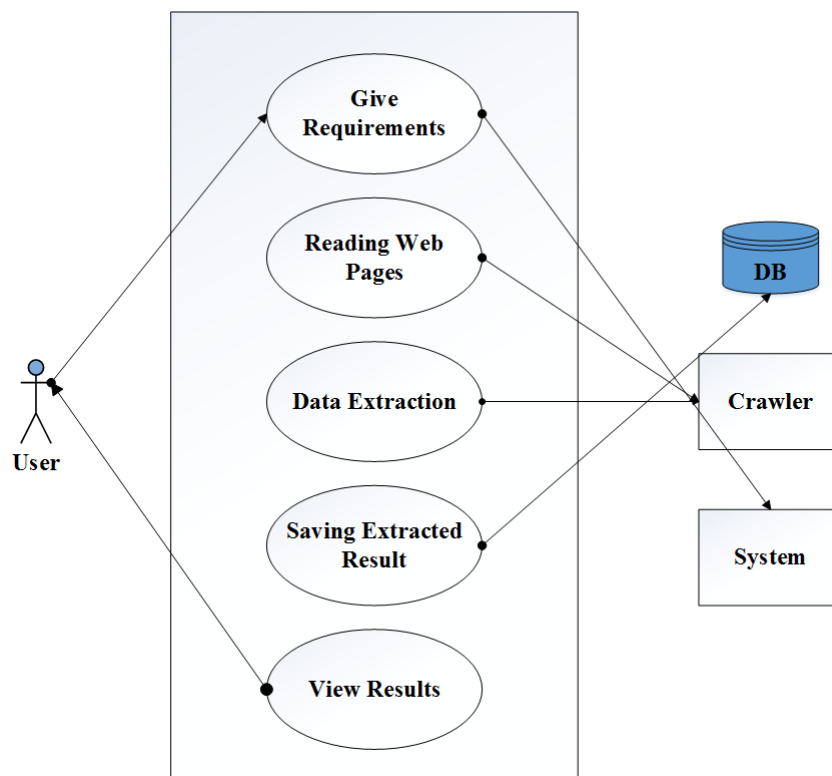


Figure 3.1: Use Case Diagram

3.6 Non Functional Requirements

Availability: Terrorism Portal, once it is housed on a server shall ensure 24/7 availability to its users.

Performance: Terrorism Portal will not take more than 5 seconds to load search/query results, no matter how detailed the query be.

Flexibility: Terrorism Portal is flexible enough to accommodate all changes in requirements, pertaining to source sites and design/orientation of scraped information.

Usability: This system is user friendly. User will be able to learn and understand the application without any problem.

3.7 Usecase Diagram

Usecase Diagram shows the interaction between user and system. In this system user give input to the system and system performs accordingly, use case is as shown in Figure 3.1.

Chapter 4

Design

Systems design is the process of defining the architecture, components, modules, interfaces, and data for a system to satisfy specified requirements.

4.1 System Architecture

Our proposed system consists of following modules. Figure 4.1 presents the overview of the system.

- Data Scraping
- Data Parsing
- Database Management
- Website / Portal (for search and other features)

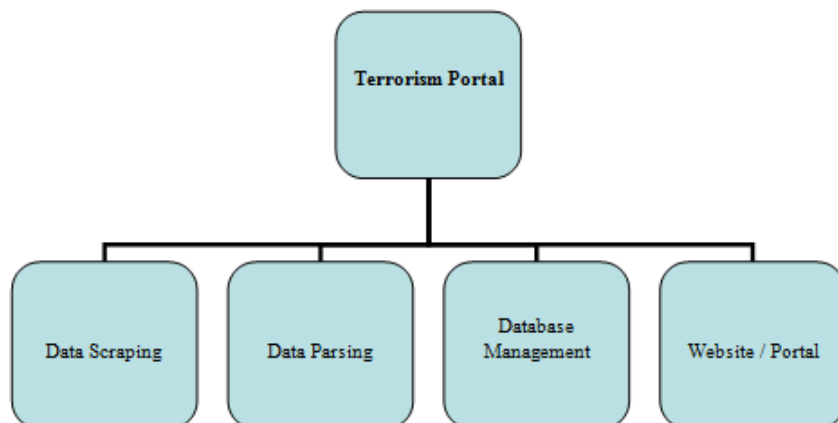


Figure 4.1: Project module

Description of identified modules is described below:

4.1.1 Data Scraping

The system, using a highly sophisticated algorithm will scrap selected sites for relevant data.

4.1.2 Data Manipulation

Data scrapped in first step will be passed onto data parsing module that will perform data conversions and proper formatting shall be applied on that data to avoid data duplications and to refine data in a way to find common attributes among scrapped data coming from different site.

4.1.3 Database Management

Properly formatted data from second module will be stored in a centralized data repository to ensure issues like data redundancy, data duplication, and data manipulation are handled coherently from a single source.

4.1.4 Terrorism Portal

A user-friendly website will allow people from different fields advance level data searching facility. Search criteria will use fields such as Year, City, etc. to provide flexible search facility in an intelligent and efficient way. This information can later be utilized to form patterns and draw conclusions.

4.2 Design Constraints

Following are the design constraints of the system:

4.2.1 System Dependencies

Since our terrorism portal is heavily reliant on existing systems for data gathering, few compromises have been made with respect to data formatting and conversions to ensure data coherence and same data definitions e.g. fields can be used for information coming from all nominated sources for data scraping.

4.2.2 System Assumptions

In absence of RSS feeds, we have to rely on traditional data scraping methods for data generation that relies on generated HTML of relevant pages. Hence it is assumed that HTML design of target pages won't change during the period of data gathering, change in design structure will result in change in back end system accordingly.

4.3 Design Methodology

Design methodology is consist of hardware and software requirements which are as follows:

4.3.1 Hardware & Software Requirements

- Platform: Windows
- Programming Language: Java EE.
- Programming Language, Others: HTML5, CSS, JavaScript.
- Tools: Eclipse, XAMPP
- Database Server: MySQL
- Other Libraries: jsoup for HTML based data scraping
- Browser: Internet Explorer 9 and above, latest versions of Mozilla, Safari, Opera & Google Chrome (for Client Side browsing)
- Other Hardware Requirements:
- Processor: Pentium 4 & above,
- Minimum Ram: 1 GB

4.4 High Level Design

High Level Design is the overall system design covering the system architecture. It describes the relation between various modules and functions of the system. Following are the designs which describe our system functionality.

4.4.1 Activity Diagram

This is the activity diagram of the system in which the user starts the application and select the requirements and press submit button if the there is some error the system goes to the main interface and process starts again if there is no error the crawler crawls and scrap all the data and save it in the data base and display accurate and relevant result. Figure 4.2 shows the architecture design of the system.

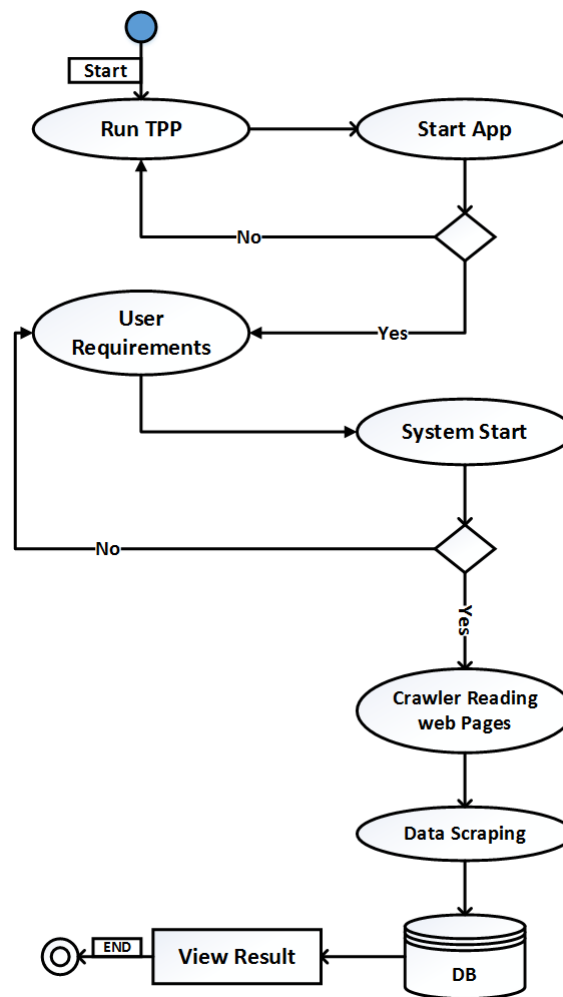


Figure 4.2: Activity Diagram

4.4.2 Context Diagram

This is the context diagram of the system which shows the boundary between the systems. User selects the requirements and Developer maintain the Database of the system. Figure 4.3 shows the Context diagram of the system.

4.4.3 Data Flow Diagram

This is the data flow diagram of the system which describes the graphical representation of the "flow" of data through an information system. Figure 4.4 shows the Data Flow Diagram of the system.

4.4.4 State Diagram

This is the state diagram of the system which shows the behavior of the system. In this diagram user run the app and select the desirable option , then the crawler will move on

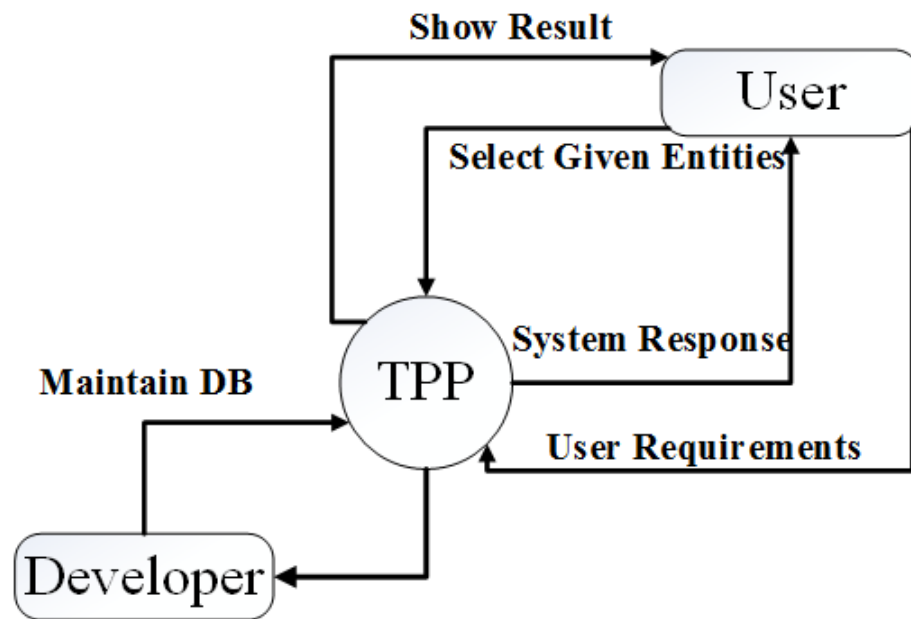


Figure 4.3: Context Diagram

and crawl the high rated websites and will scrap the accurate and relevant data. When the required data is extracted it will store in database and display results. At the end system will terminated. Figure 4.5 shows the State Diagram of the system.

4.4.5 Package Diagram

This is the package diagram of the system which shows the dependencies between the packages. In this diagram there is the windows operating system at the user end which shows the interaction among back end processes. Figure 4.6 shows the State Diagram of the system.

4.5 Low Level Design

Low-level design is a component-level design process that follows a step-by-step refinement process. Low level design of this application is shown in Figure 4.7.

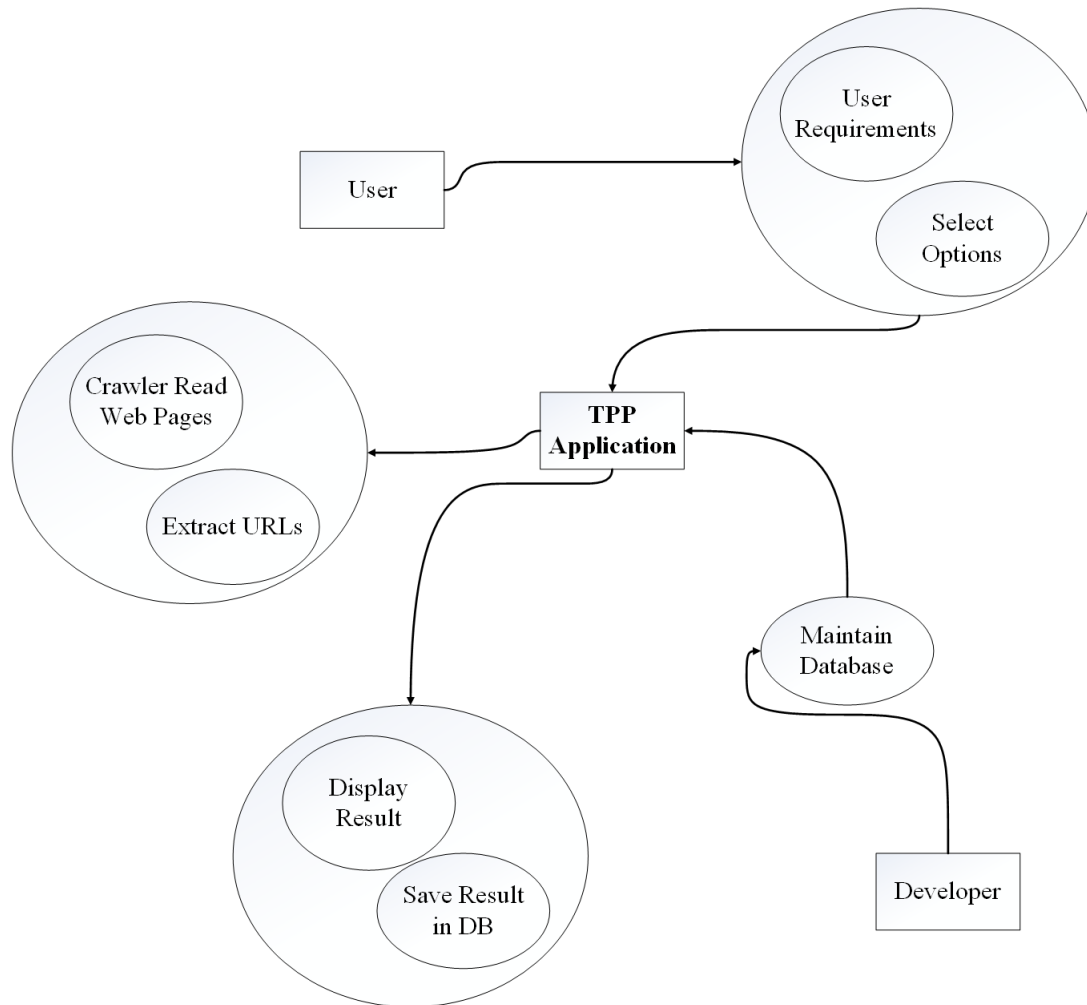


Figure 4.4: Data Flow Diagram

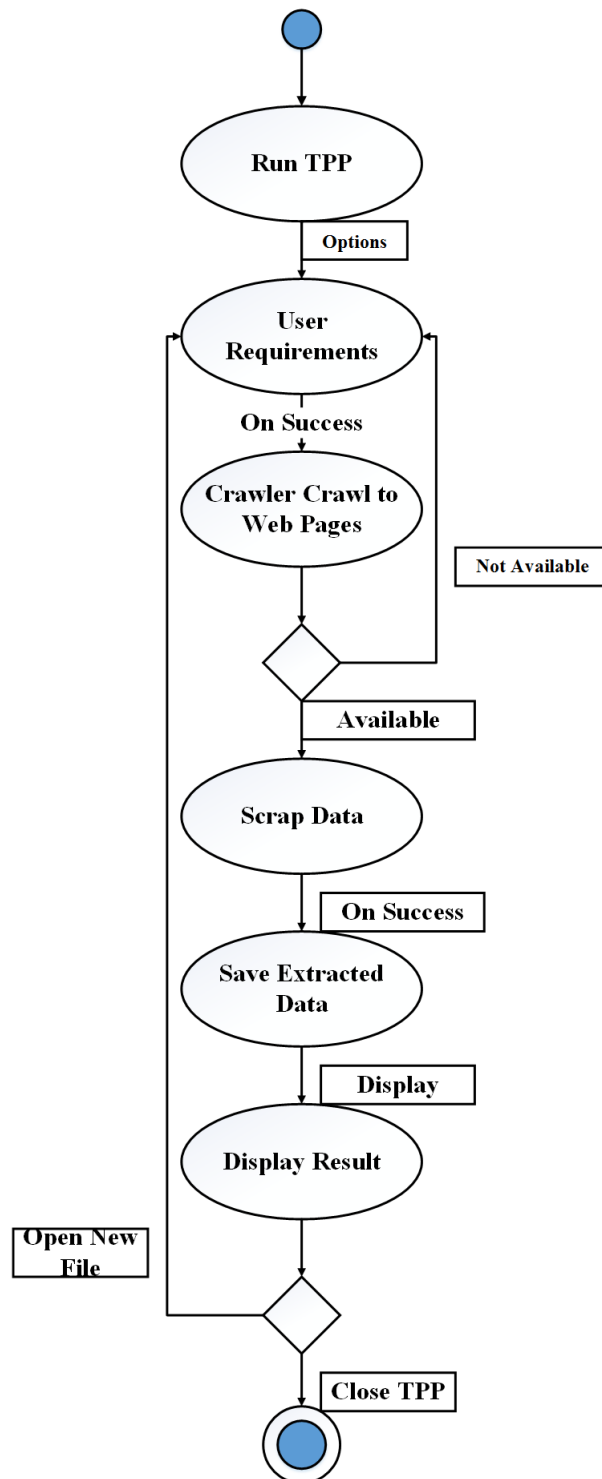


Figure 4.5: State Diagram

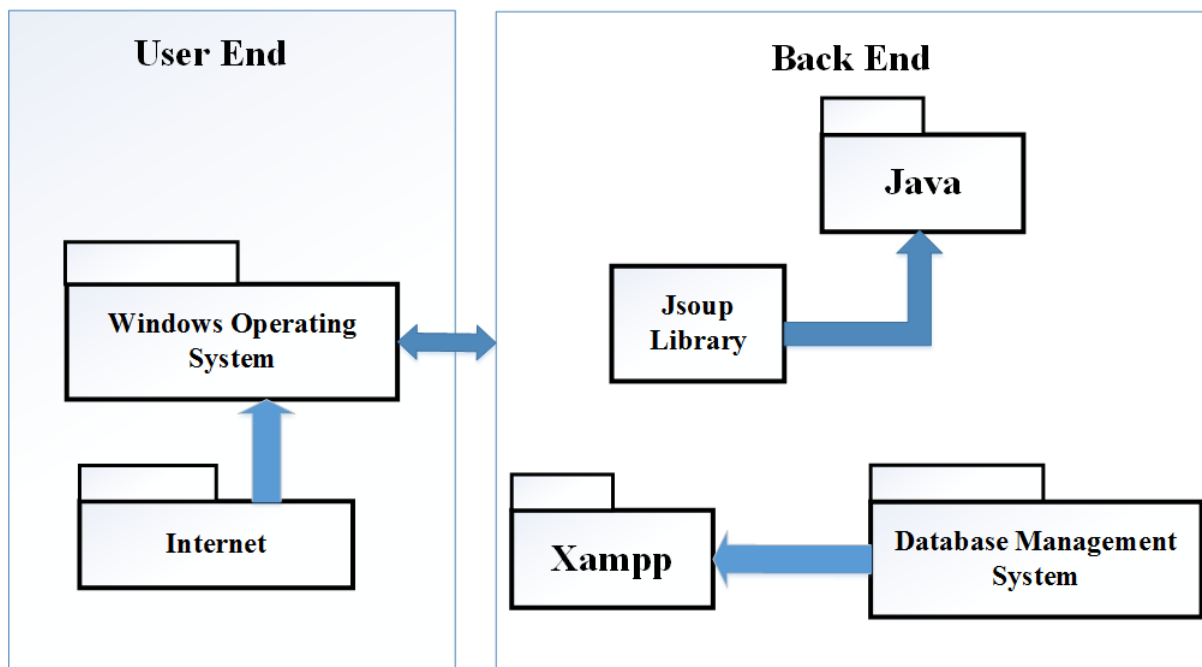


Figure 4.6: Package Diagram

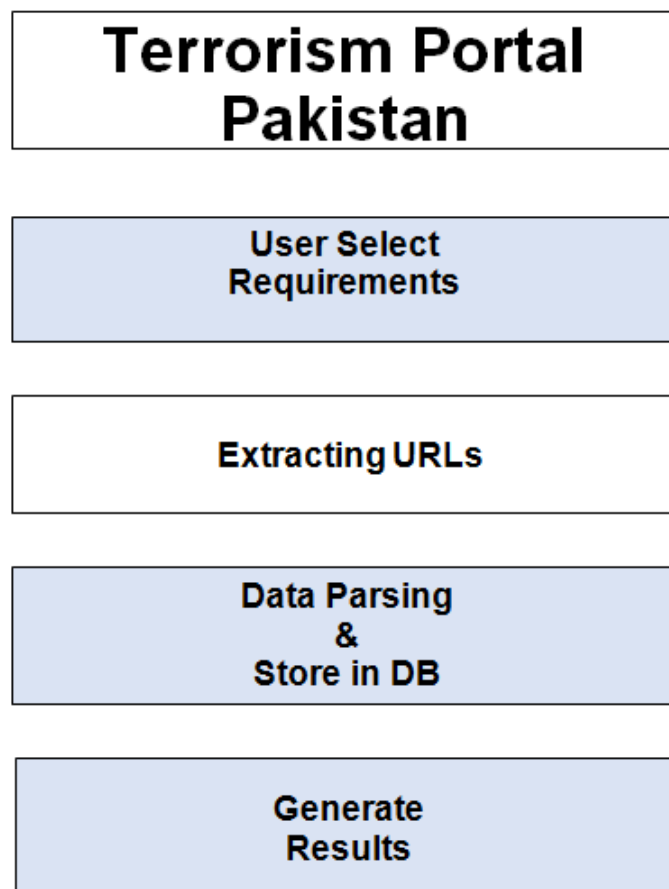


Figure 4.7: Architecture Diagram

Chapter 5

System Implementation

Implementation is the phase where we put the ideas into work. It is a systematic way to effectively transform our design into actual working model. Implementation is the phase where we actually write the code for the project. Implementation can also be considered as the final step of moving the solution from development status to production status. There are many ways to implement applications but we have to choose the way that best fits our design. So here is the Introduction to tools and reasons for choosing these tools.

5.1 System Architecture

This System Architecture Describes System internal components, Functionality of the components, Communication between the components shown in Figure 5.1.

5.2 Tools and Technology Used

Selection of appropriate tools plays a vital role in developing an application. Choosing suitable tools will save time, cost, and effort. The tools used in the project include Eclipse, XAMPP (Apache, MySql, Tomcat), Java, HTML5, Java Script and CSS.

5.3 Development Environment & Languages Used

Several tools involved in the Development Environment. Java language is used in this Web-application. Java is powerful type-safe and simple language. Eclipse is used to run the java code for the development of this application. The other languages like HTML5, Java Script and CSS are used at the front end to build the interface of the system.

5.3.1 Eclipse

Eclipse is a Java-based open source platform that allows a software developer to create a customized development environment. It is very helpful in development. It saves time, it prompts when you are writing code, so it helps in saving your typing effort as well.

5.3.2 JAVA

Java is a set of several computer software products and specifications from Sun Micro systems (which has since merged with Oracle Corporation), that together provide a system for developing application software and deploying it in a cross-platform computing environment. Java is used in a wide variety of computing platforms from embedded devices and mobile phones on the low end, to enterprise servers and supercomputers on the high end. Writing in the Java programming language is the primary way to produce code that will be deployed as Java byte code. There are, however, byte code compilers available for other languages such as Ada, JavaScript, Python, and Ruby. Several new languages have been designed to run natively on the Java Virtual Machine, such as Scala, Clojure and Groovy. Java's syntax is mostly like C, C++ but object-oriented features are modeled. Java eliminates certain low-level constructs such as pointers and has a very simple memory model where every object is allocated on the heap and all variables of object types are references. Memory management is handled through integrated automatic garbage collection performed by the JVM. Java is almost same as C Sharp.

5.3.3 XAMPP

In XAMPP MYSQL is used for the database, where the parsed data from the defined websites is stored in database. Tomcat is used as a local host server [4].

5.3.4 HTML5

HTML5 is the latest version of Hypertext Markup Language, the code that describes web pages. It's actually three kinds of code: HTML, which provides the structure; Cascading Style Sheets (CSS), which take care of presentation; and JavaScript, which makes things happen. The HTML5 is working at the front end of the application for the interface development.

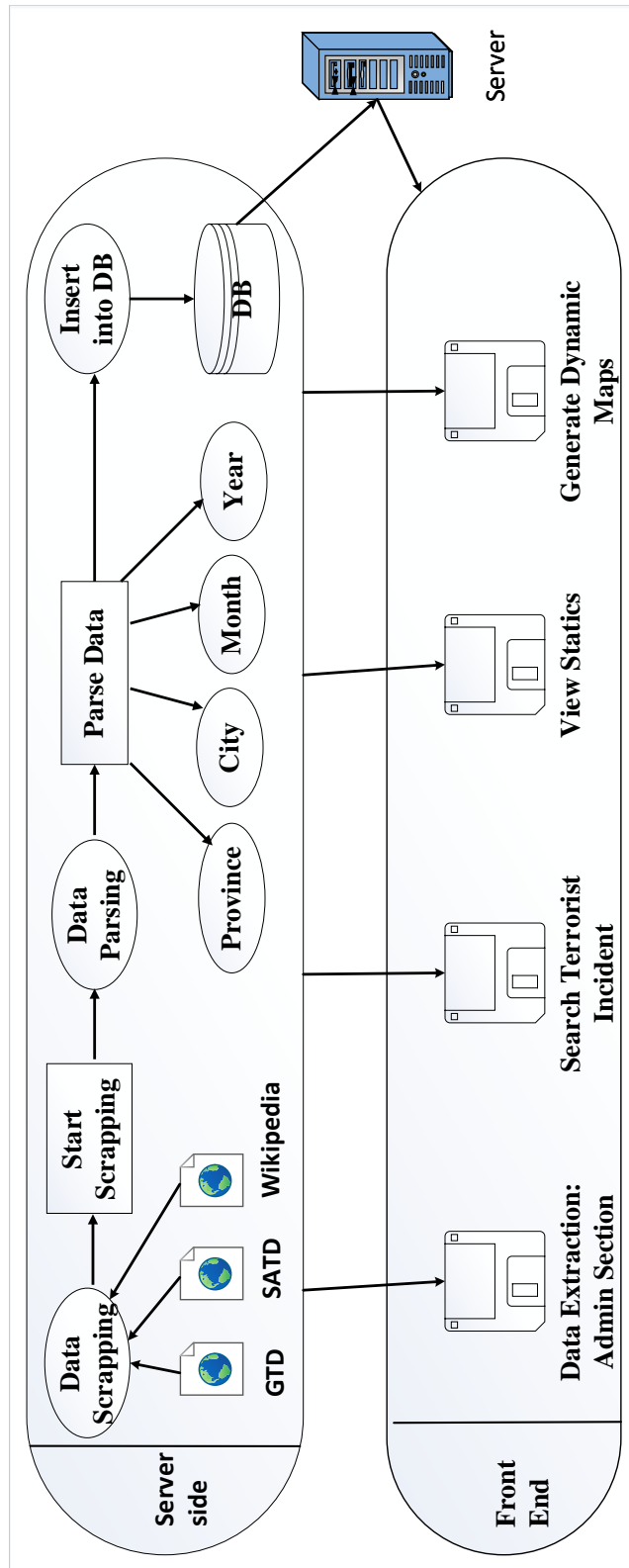


Figure 5.1: System Architecture

Chapter 6

System Testing and Evaluation

Testing is the investigation conducted to find the errors and the flaws in the system so that the errors can be removed. Testing is done to make an error free system. Testing happens every time the system is modified. Software testing is a basic part of the software development life cycle. Software testing is the process of validating and verifying that the system,

- Meets the requirements
- Works as it is expected
- Satisfies the needs of stakeholders

6.1 Graphical user interface testing

Our project interface is user friendly. The user can easily interact with the system and the results shown by the system is in the form of graphs and charts. The user can easily understand and can further analyze the result.

6.1.1 Main Interface

This is the Main interface of the project. Which shows that how user will interact to the system, this interface shows the death rate and injured rate to user by default but further, user can search the information by selecting given options as shown in Figure 6.1.

6.1.2 Result Output

This interface shows that user give the input to the system and result is generated in descriptive from database. this result will be relevant to user's query as shown in Figure 6.2.

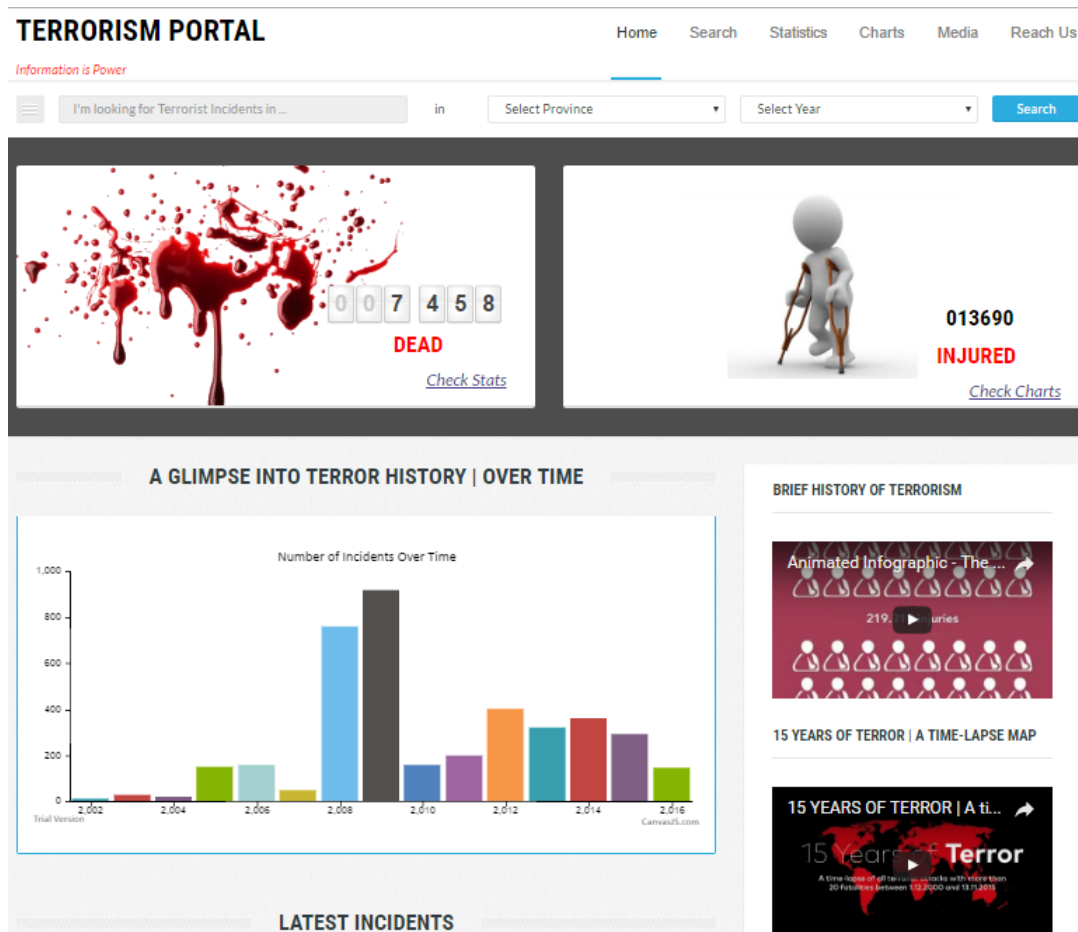



Figure 6.1: Main Interface

TERRORISM PORTAL Home Search Statistics Charts Media Reach Us

Information is Power


SEARCH TERRORIST INCIDENTS 2009 Select Month Punjab Select City Search

TERRORIST INCIDENTS (37)

 **December 15** **Dera Ghazi Khan, Punjab** 2009


A suicide car bomb exploded in a market outside the home of the Punjab Chief Minister's senior adviser, Zulfiqar Khosa in Dera Ghazi Khan District of Punjab, killing 33 people and injured 60 others. Unidentified militants detonated the explosives outside the house of Khosa, who was not in

home at t

 **December 8** **Multan, Punjab** 2009


A group of three Taliban militants launched a gun, rocket and suicide attack on the office of ISI, killing at least 12 people and injuring several others at Multan in Punjab. The blast in Multan destroyed the facades off several buildings in a part of the city largely reserved for Government

and Se

 **December 7** **Lahore, Punjab** 2009

Two bomb blasts killed at least 38 people, and injured more than 100 at the crowded Moon Market in Allama Iqbal area of Lahore in Punjab. The two bombs exploded 30 seconds apart at 8:45 PM (PST). The first blast occurred outside a plaza housing a branch of the Muslim Commercial Bank,

while the othe

 **December 4** **Rawalpindi, Punjab** 2009

40 persons, including 17 children besides serving and retired Army officers and personnel, were killed and over 86 others injured, when a Friday congregation at the Parade Lanes mosque in Rawalpindi was attacked by a group of terrorists. The high number of casualties was caused by

Figure 6.2: Result Output

6.1.3 Graph Output

The result displayed in Graph/charts which is easy to understand by user, when user select the given options according to his requirements and press generate button then the Graph will generated which shows the death rate and injured rate as shown in Figure 6.3.

6.2 Usability testing

This Web application has been designed according to the principles of human computer interaction. Application is easy, efficient to use and learn. All the interfaces are consistent, visible and clear.

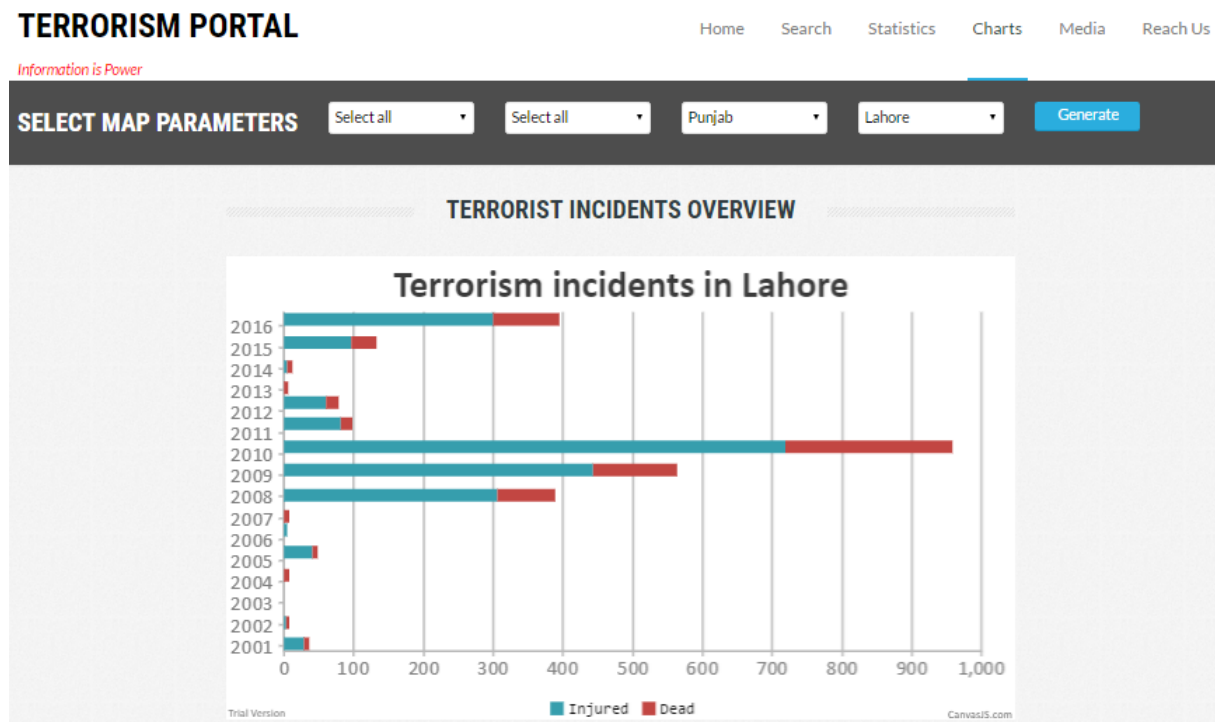


Figure 6.3: Graph Output

6.3 Performance testing

This application is very efficient and user friendly, once the user gives input for desire data related to terrorism activities, the result will be show within minimum time depending on the speed of internet because internet is mandatory to run this application (minimum speed 2Mbps).

6.4 Compatibility Testing

This application is compatible for all hardware platforms and web browsers.

6.5 Installation testing

It is mandatory to install eclipse and XAMPP software to run this web application. The systems having these software will be able to run the application.

6.6 Test Cases

6.6.1 Test Case1: Get requirements

To get requirements from the user and then populate them to the database as shown in Figure 6.4.

Purpose:	To Gather Requirements from user and Populate it in database.							
Pre Condition:	The user must open Website and goes to search tab							
Test Data:	Get Requirements= {valid=requirements submitted invalid=}							
Steps:	<ul style="list-style-type: none"> • User Open the website. • Goes to search tab • Select Requirements. • Submit the Requirements. • Requirements are populated in the database. 							
Test Results:	<table border="1"> <thead> <tr> <th>Basic Flow</th> <th>Expected Result</th> <th>Status</th> </tr> </thead> <tbody> <tr> <td>User Open the search tab to select requirements</td> <td>Requirements are populated in the database</td> <td>Pass</td> </tr> </tbody> </table>		Basic Flow	Expected Result	Status	User Open the search tab to select requirements	Requirements are populated in the database	Pass
Basic Flow	Expected Result	Status						
User Open the search tab to select requirements	Requirements are populated in the database	Pass						

Figure 6.4: Get Requirements

6.6.2 Test Case2: Reading Web-pages

To read web pages line by line and gather relevant data as shown in Figure 6.5 ^{1 2 3}.

6.6.3 Test Case 3: Data Extraction

To find out relevant data according to requirements of the user as shown in Figure 6.6.

6.6.4 Test Case4: Saving Extracted Results

To save extracted results in the database as shown in Figure 6.7.

¹<https://www.start.umd.edu/gtd/> accessed on December 15, 2016

²<http://www.satp.org/> accessed on December 15, 2016

³<https://en.wikipedia.org/wiki/Portal:Terrorism> accessed on December 15, 2016

Purpose:	To gather relevant URL						
Pre-Condition:	Should Load the URL						
Test Data:	Reading Web Pages={valid=read the web pages and get the useful link Invalid=reading the web pages is not Completed						
Steps:	<ul style="list-style-type: none"> • Crawler Read the web pages line by line • It Will skip the Useless Links 						
Test Results:	<table border="1"> <thead> <tr> <th>Basic Flow</th> <th>Expected Result</th> <th>Status</th> </tr> </thead> <tbody> <tr> <td>Crawler will keep reading the Web Pages</td> <td>Maintain the Record of the Extracted Link</td> <td>Pass</td> </tr> </tbody> </table>	Basic Flow	Expected Result	Status	Crawler will keep reading the Web Pages	Maintain the Record of the Extracted Link	Pass
Basic Flow	Expected Result	Status					
Crawler will keep reading the Web Pages	Maintain the Record of the Extracted Link	Pass					

Figure 6.5: Reading Web Pages

6.6.5 Test Case5: View Results

To show results to the user in the Charts/Graph and descriptive form as shown in Figure 6.8.

Purpose:	To find out Relevant Data						
Pre Condition:	Read the Web Pages						
Test Case:	Data Extraction={valid=extract the dataInvalid=no extraction of data						
Steps:	<ul style="list-style-type: none"> • System will extract the data from web pages 						
Test Case:	<table border="1"> <thead> <tr> <th>Basic Flow</th> <th>Expected Result</th> <th>Status</th> </tr> </thead> <tbody> <tr> <td>System will extract the data from web pages</td> <td>Relevant data is found</td> <td>Pass</td> </tr> </tbody> </table>	Basic Flow	Expected Result	Status	System will extract the data from web pages	Relevant data is found	Pass
Basic Flow	Expected Result	Status					
System will extract the data from web pages	Relevant data is found	Pass					

Figure 6.6: Data Extraction

Purpose:	To save all extracted result in Database						
Pre Condition:	Data extraction should be done						
Test Data:	Data extraction={valid=extracted results are saved in Database Invalid=results are not saved						
Steps:	System saves the extracted result in Database						
Test Results:	<table border="1"> <thead> <tr> <th>Basic Flow</th> <th>Expected Results</th> <th>Status</th> </tr> </thead> <tbody> <tr> <td>The Extracted result goes to database</td> <td>Extracted result is populated in database</td> <td>Pass</td> </tr> </tbody> </table>	Basic Flow	Expected Results	Status	The Extracted result goes to database	Extracted result is populated in database	Pass
Basic Flow	Expected Results	Status					
The Extracted result goes to database	Extracted result is populated in database	Pass					

Figure 6.7: Save Result

Purpose:	Show result to the user						
Pre Condition:	Extracted result should be save in database						
Test Data:	View result={valid=results are visible to the user Invalid=results are not visible to the user						
Steps:	<ul style="list-style-type: none"> • After selecting requirements user will wait for result • Result will be displayed in Graphs and in descriptive form 						
Test Results:	<table border="1"> <thead> <tr> <th>Basic Flow</th> <th>Expected Results</th> <th>Status</th> </tr> </thead> <tbody> <tr> <td>User will be able to see result in graphs and along with description</td> <td>System will show the result to user</td> <td>Pass</td> </tr> </tbody> </table>	Basic Flow	Expected Results	Status	User will be able to see result in graphs and along with description	System will show the result to user	Pass
Basic Flow	Expected Results	Status					
User will be able to see result in graphs and along with description	System will show the result to user	Pass					

Figure 6.8: Show result

Chapter 7

Conclusion

The project terrorism Portal Pakistan is very beneficial for media houses, researchers and students as there is no system which brings information about terrorism activities in Pakistan from the internet. This application provides the name of cities, day, month along with its description. Through the URL the description is fetched. The results are populated into the database so a full-fledged database is always there which contain the information of the terrorist activity in Pakistan which is very useful for the analysis purpose. Terrorism crawler is very valuable as on just one button click all the information related to the terrorism activity will be displayed in the form of Graphs and description. There is lack of a data repository that may contain data for different terrorism related activities in Pakistan. So analyzing such activities is also difficult and there is a need of getting data from the different sources from the net and populate it in a database that can be used for analyzing the data.

References

- [1] K.P.S.Gill, 2016. Cited on p. [7](#).
- [2] Global terrorism database, 2016. Cited on p. [7](#).
- [3] Karl E. Wieggers. Software requirements. *Microsoft Press. ISBN 978-0-7356-1879-4*, 2003. Cited on p. [9](#).
- [4] Yasser Ganjisaffar, 2010-2016. Cited on p. [22](#).

