

Mine music

Saurav Roy, Subhajit Pal

ESL LAB, Chandannagar CA-5, Salt Lake City, Kolkata – 700064

Abstract : This technical report describes the implementation details of a music recommendation system based on data mining techniques. The system enables the user to receive a playlist depending on his mood and age. Mood and age, these two parameters were considered by the system to recommend a song to a specific user.

The proposed application also uses encryption and decryption technique to prevent outside attacks and secure its dataset from the attackers.

The methodology section of this report describes the step algorithm along with the data flow diagram. We hope that the design and development concept described in this project will aid the developers' community in near future.

Keywords: Python, Data mining, Recommendation system, Encryption Decryption, SQL, Cryptography

I. Introduction

1.1 Python

1.1.1 What is Python? Python is high level, interpreted, pure object oriented widely used scripting language. Like all other scripting languages, Python code resembles pseudo code. Its syntax's rules and elegant design make it readable even among multi-programmer development teams. The syntax of the language is easily understandable, which is really helpful. The idea behind that is to keep you thinking about the business rules of your application and not to spend time trying to figure out what command you should use.

1.1.2 Why Python? Python has clear and simple code structure, and is closer to English than any of the languages we use. Creating Python programming is so straightforward that it's been called "programming at the speed of thought."

The simplicity of Python makes it easy to learn. In addition to the list (dynamic array) data structure, Python provides tuples (immutable lists) and dictionaries (hash tables). Together with the class mechanism, these can be used to quickly build sophisticated data structures for interesting projects. Statements are terminated by end of line, and block structure is indicated by indentation. Python programs look like executable pseudo-code. The absence of type declarations makes for less code and more flexible programming.

1.1.3 Platform Independent Python programs are platform independent, which means that regardless of the operating system you use to create your program, it'll run on any other computer with Python. So if you write some program on your PC, you can e-mail a copy to your friend who runs Linux or to someone who has a Mac and the program will work same.

1.1.4 Object Oriented Programming Just like other programming Languages as C#, Java, Python is also a object-oriented Language. But it works better than them. The codes of Python are smaller and easier than other languages and easy to implement. In Python, using OOP techniques is optional. You have all of OOP's power at your disposal, but you can use it when you need it. Got a short program that doesn't really require OOP? No problem. Got a large project with a team of programmers that demands OOP? That'll work too. Python gives you power and flexibility.

1.1.5 Open Source Python software is free. You can install it on your computer or laptop and never have to pay for it. Python's license also lets you do much more than that. You can copy or modify Python. You can even resell Python if you want (but don't quit your day job just yet). Embracing open-source ideals like this is part of what makes Python so popular and successful. Python is a programming language that lets you work more quickly and integrate your systems more effectively. You can learn to use Python and see almost immediate gains in productivity and lower maintenance costs.

1.1.6 Python is self-adjustable Python provides full dynamic runtime type checking and bounds checking on array subscripts⁴. Python employs garbage collection so there is no problem with Dangling pointers or memory leaks. It is impossible for user code in Python to produce a segmentation violation. In this respect Python is similar to Java, and both are much safer than C++.

1.2 Data Mining Data mining is the process of analysing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analysing data. It allows users to analyse data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.

1.2.1 Data mining consists of five major elements

- Extract, transform, and load transaction data onto the data warehouse system.
- Store and manage the data in a multidimensional database system.
- Provide data access to business analysts and information technology professionals.
- Analyses the data by application software.
- Present the data in a useful format, such as a graph or table.

1.3 SQL SQL stands for Structured Query Language. SQL is used to communicate with a database. According to ANSI (American National Standards Institute), it is the standard language for relational database management systems. SQL statements are used to perform tasks such as update data on a database, or retrieve data from a database. Some common relational database management systems that use SQL are: Oracle, Sybase, Microsoft SQL Server, Access, Ingres, etc. Although most database systems use SQL, most of them also have their own additional proprietary extensions that are usually only used on their system. However, the standard SQL commands such as "Select", "Insert", "Update", "Delete", "Create", and "Drop" can be used to accomplish almost everything that one needs to do with a database.

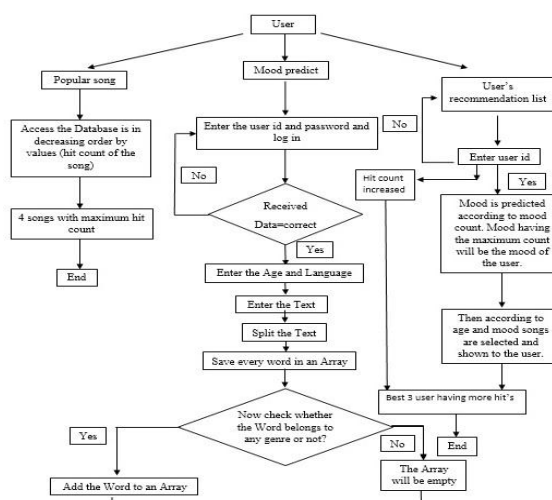
1.4 Encryption and Decryption Encryption is the conversion of data into a form, called a cipher text (it is an encrypted text), that cannot be easily understood by unauthorized people. Decryption is the process of converting encrypted data back into its original form, so it can be understood.

1.4.1 Uses Encryption/decryption is especially important in wireless communications. This is because wireless circuits are easier to tap than their hard-wired counterparts. Nevertheless, encryption/decryption is a good idea when carrying out any kind of sensitive transaction, such as a credit-card purchase online, or the discussion of a company secret between different departments in the organization. The stronger the cipher -- that is, the harder it is for unauthorized people to break it -- the better, in general. However, as the strength of encryption/decryption increases, so does the cost.

II. Motivation

In present era music player is very common. We use it while travelling and also for pass time. Many types of music players are available in the market. The competition is very high. To stand in the market we need to implement new key feature in the music player. So we use data mining in this player so that it can recommend music to the user according their best interest which is tracked by the player. If the user is new, we recommended him the music belongs to their category.

III. Flow Chart



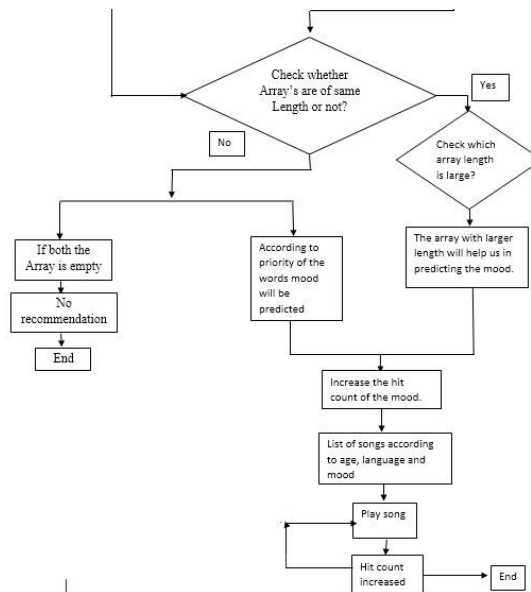


Fig. 3.1 user block

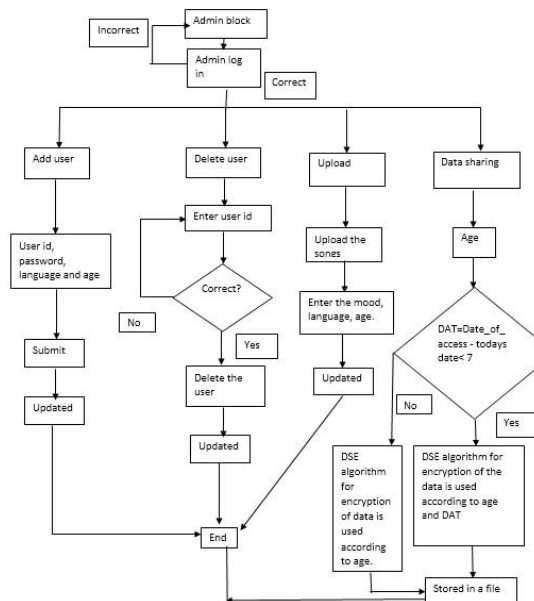


Fig. 3.2 admin block

Step algorithm

Step 1

Starts the program a window opens with two options “admin” and “user” as shown in Fig. 3.3.

Step 2

Admin part

When admin is selected another window opens where administrator is allowed to login as shown in Fig. 3.10. If administrator userid and password is correct then another window pops up:-“adduser”, “delete user”, “upload”, “datasharing”.

Step 3

If “adduser” is selected a window pops up as shown in Fig 3.11. Where we have to enter the new userid, password, name and age after that a new row is added in the user table.

Step 4

If “delete user” is selected a window pops up which asks for the userid .if a correct userid is chosen the following row is deleted as shown in Fig. 3.13 and Fig. 3.14.

Step 5

If “upload” is selected a window pops up to search for the songs you want to upload. After the song is selected a window pops up asks for mood,language,age after entering all the data that the songs is updated in the mood tables shown in Fig 3.16 and Fig. 3.17.

Step 6

If “datasharing” is selected a window pops up asking for age. The songs listened by any user within a week and belongs to that particular age according to it the data of the database is selected and encrypted and saved into a file and shared as shown in Fig. 3.15. After then exit.

Step 7

If “user option” is selected a window pops up with three options “popular songs”, “mood predict”, “user_recomendation_list” as shown in Fig 3.6.

Step 8

If “popular songs” is selected the song’s with maximum “values” or hit count is selected. Top 4 songs are shown in Fig. 3.7.

Step 9

When “mood predict” is selected a window pop’s up asking for username and password .if the data is correct another window pop’s up asking for the text and language. From the user database age of the user is taken. The text entered is split up into words and saved into an array. And from this array we checked each and every word one by one. If the word belongs to any genre add that word to any specified array for that genre. Else the array of that genre will be empty.

Now check the length of each and every specified array for the genre. The array with maximum length will give us the mood. If the lengths are equal then the priority of word will decide the mood. From the mood, language and age the list of songs will be shown. A text box is provided in the list so that we can play a particular song as shown in Fig 3.8.

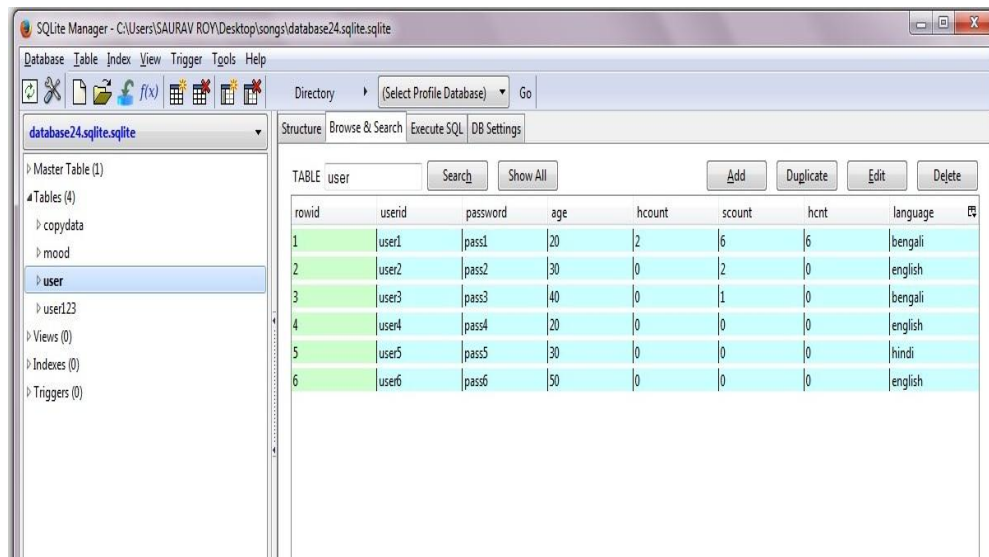
Step 10

User part

If “user_recomendation_list” is selected. A window pop’s up asking for userid. If it is valid then the age and language will be taken from the user database. From the mood count we can decide the mood. The moodcount with maximum value will be treated as the mood. And from that mood, language, and age, recommendation list is shown in Fig. 3.9. After that hit count of user is increased.

The user database are arranged in decreasing order of the user hit count so that every time top 3 user can be selected.

If user id is not correct it will give us an empty list.



rowid	userid	password	age	hcount	scount	hcmt	language
1	user1	pass1	20	2	6	6	bengali
2	user2	pass2	30	0	2	0	english
3	user3	pass3	40	0	1	0	bengali
4	user4	pass4	20	0	0	0	english
5	user5	pass5	30	0	0	0	hindi
6	user6	pass6	50	0	0	0	english

Fig. 3.3 showing the user table in the database

rowid	mood	language	songs	value	age	time_of_access
1	sad	bengali	sad1.mp3	3	20	2014-01-24
2	sad	bengali	sad2.mp3	4	20	2014-01-29
3	sad	bengali	sad3.mp3	5	20	2014-02-18
4	sad	bengali	sad4.mp3	1	20	2014-01-29
5	sad	bengali	sad5.mp3	5	20	2014-02-12
6	happy	bengali	hpy1.mp3	3	20	2014-02-12
7	happy	bengali	hpy2.mp3	1	20	2014-01-29
8	happy	bengali	hpy3.mp3	2	20	2014-01-29
9	happy	bengali	hpy4.mp3	3	20	2014-02-12
10	happy	bengali	hpy5.mp3	1	20	2014-01-29
11	sad	bengali	sad1.mp3	1	40	2014-02-12
12	sad	bengali	sad2.mp3	4	40	2014-01-29
13	sad	bengali	sad3.mp3	5	40	2014-02-18
14	sad	bengali	sad4.mp3	1	40	2014-01-29
15	sad	bengali	sad5.mp3	5	40	2014-02-12
16	happy	bengali	hpy1.mp3	3	40	2014-02-12
17	happy	bengali	hpy2.mp3	1	40	2014-01-29
18	happy	bengali	hpy3.mp3	2	40	2014-01-29
19	happy	bengali	hpy4.mp3	3	40	2014-02-12
20	crying	bengali	03_Ami Eto Je.mp3	0	40	2014-02-08
21	happy	bengali	[[Doridro.com]05.Na...]	0	30	2014-02-08

Fig. 3.4 showing the mood table in the database

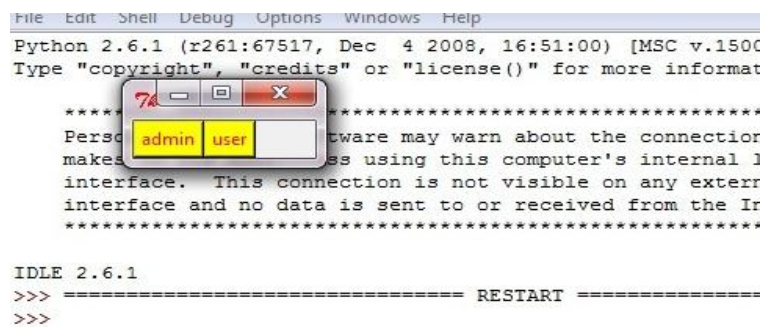


Fig. 3.5 starting window

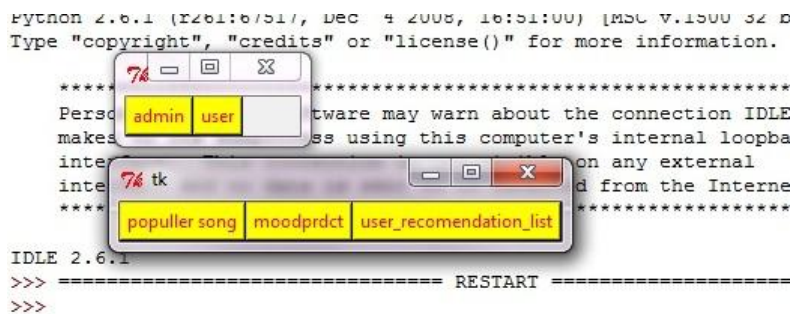


Fig. 3.6 After selecting user next window pops up with three options

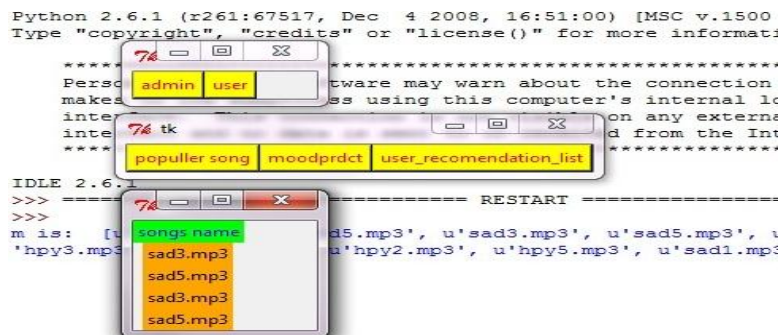


Fig. 3.7 After selecting popular songs a window pops up with songs

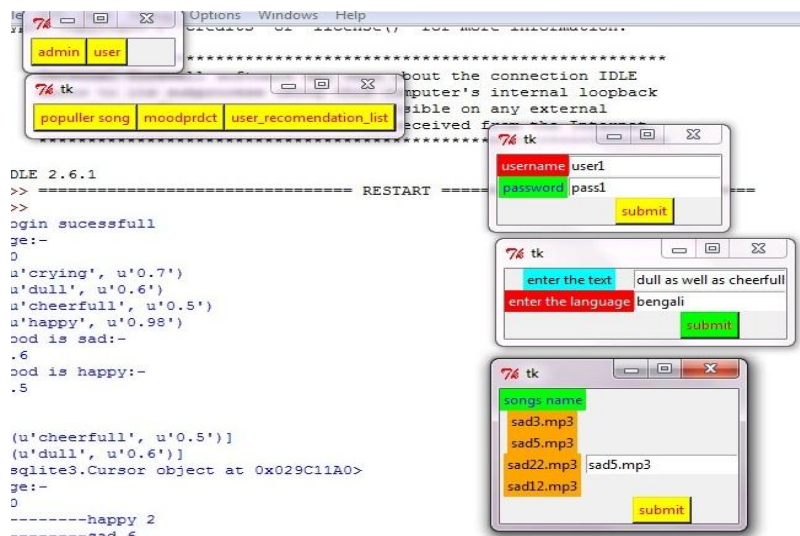


Fig. 3.8 When moodprcdt option is selected

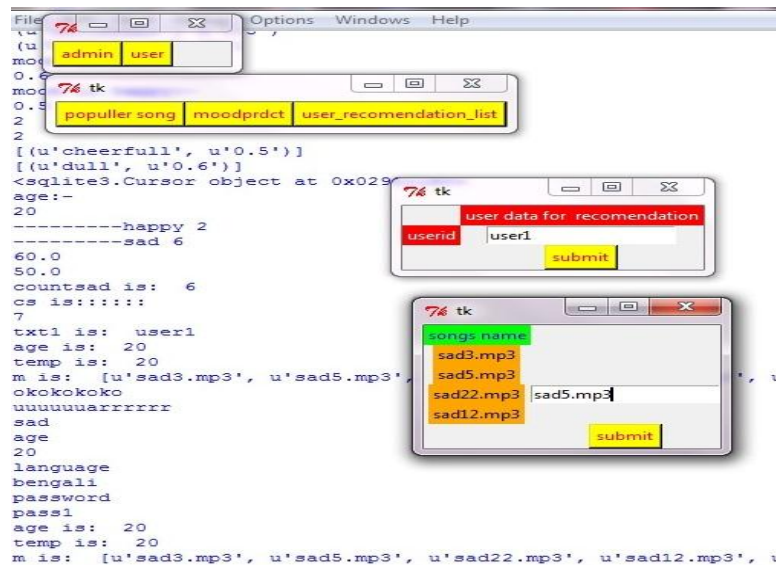


Fig. 3.9 When user_recomendation_list

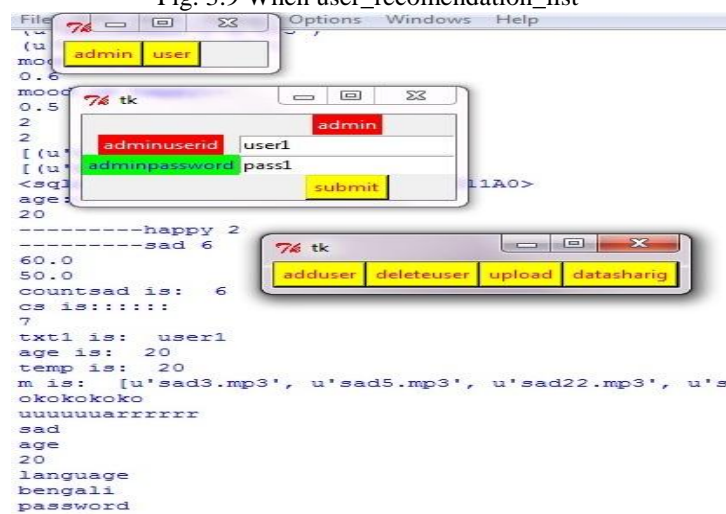


Fig. 3.10 When admin is selected and correct userid and password is entered

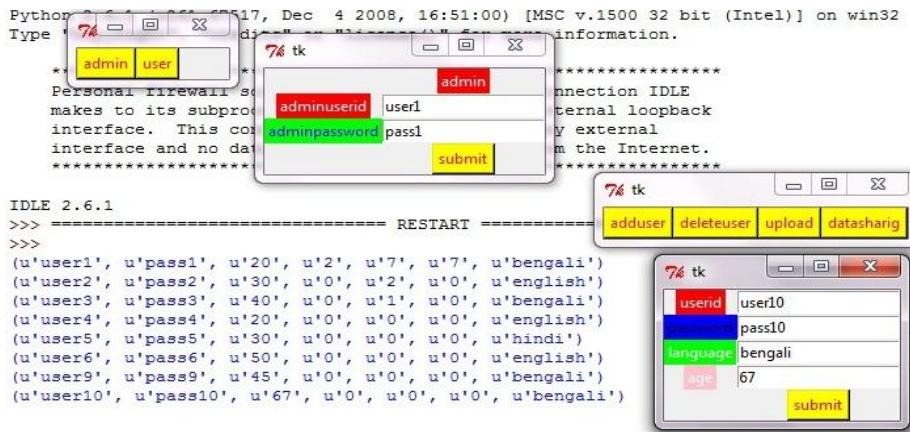


Fig. 3.11 When adduser is selected

rowid	userid	password	age	hcount	scount	hcnt	language
1	user1	pass1	20	2	7	7	bengali
2	user2	pass2	30	0	2	0	english
3	user3	pass3	40	0	1	0	bengali
4	user4	pass4	20	0	0	0	english
5	user5	pass5	30	0	0	0	hindi
6	user6	pass6	50	0	0	0	english
7	user9	pass9	45	0	0	0	bengali
8	user10	pass10	67	0	0	0	bengali

Fig. 3.12 After adduser the database is updated

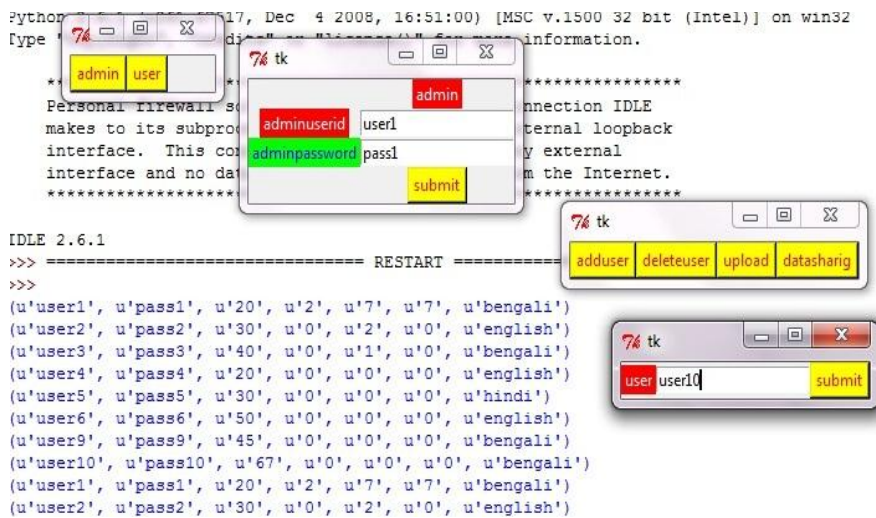


Fig. 3.13 When deleteuser is selected

rowid	userid	password	age	hcount	scount	hcnt	language
1	user1	pass1	20	2	7	7	bengali
2	user2	pass2	30	0	2	0	english
3	user3	pass3	40	0	1	0	bengali
4	user4	pass4	20	0	0	0	english
5	user5	pass5	30	0	0	0	hindi
6	user6	pass6	50	0	0	0	english
7	user9	pass9	45	0	0	0	bengali

Fig. 3.14 Database of user after deleteuser is selected

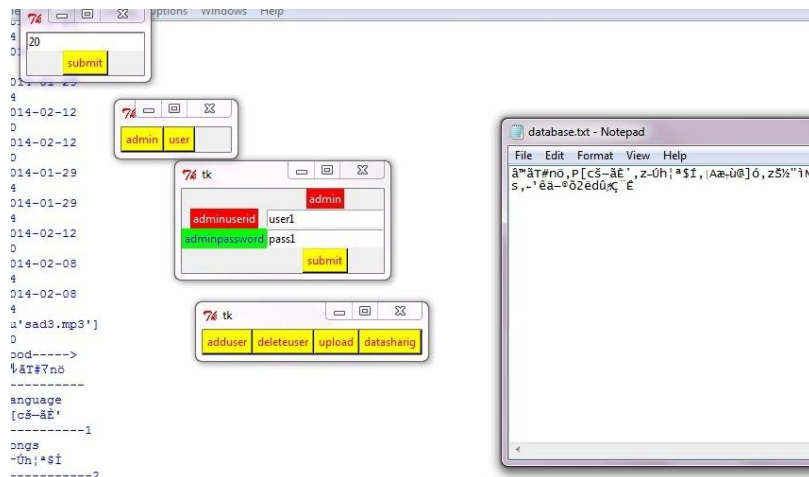


Fig. 3.15 When datasharing is selected the encrypted data's are saved into the text

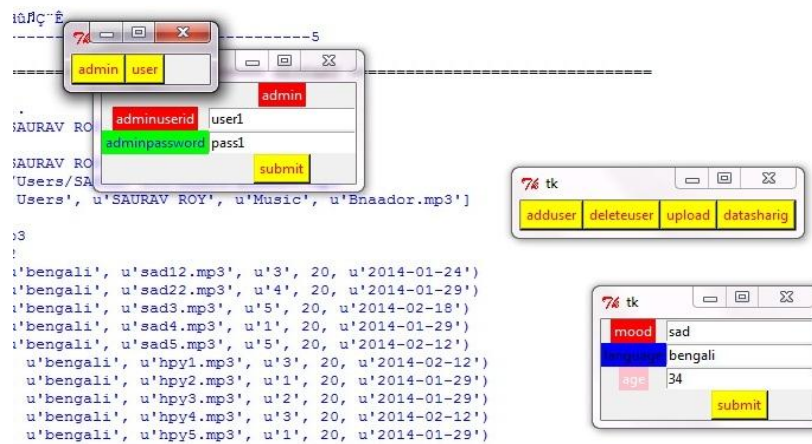


Fig. 3.16 When songs are uploaded

6	happy	bengali	hpy1.mp3	3	40	2014-02-12
7	happy	bengali	hpy2.mp3	1	40	2014-01-29
8	happy	bengali	hpy3.mp3	2	40	2014-01-29
9	happy	bengali	hpy4.mp3	3	40	2014-02-12
10	crying	bengali	03. Ami Eto Je.mp3	0	40	2014-02-08
11	happy	benagli	[Doridro.com]05.Na Je...	0	30	2014-02-08
12	sad	bengali	Bnaador.mp3	0	34	2014-02-22

Fig. 3.17 After upload the database is updated

IV. Conclusion And Future Work

The methodology section describes the implementation of the system and the result section shows that the proposed system works as per our expectation. We haven't used any security mechanisms to store our data secretly. We can import security features in order to prevent outside attacks to this system. We hope that the design and implementation described in this report will aid the developers community in near future.

References

A Method for Bypassing Keystroke Recognition Based Security System Using Social Engineering Page number: 87-93