FILE ZIPPER SOFTWARE

Group Members

Name: Imran khan lashari (13567)

Shumaila Farooq (13890)

Course: DSA lab

Project report

Under the supervision of

Sir Muhammad Amin Qureshi

Ma'am Dr Ayesha siddiqa

Date	Signature
30 jan 23	ImranKhanL

File zipper using Huffman coding

Introduction to huffman coding algorthim.

Huffman Coding is a technique of compressing data to reduce its size without losing any of the details. It was first developed by David Huffman.

Huffman Coding is generally useful to compress the data in which there are frequently occurring characters.

How Huffman Coding works?

Suppose the string below is to be sent over a network.

Initial string

Each character occupies 8 bits. There are a total of 15 characters in the above string. Thus, a total of 8 * 15 = 120 bits are required to send this string.

Using the Huffman Coding technique, we can compress the string to a smaller size.

Huffman coding first creates a tree using the frequencies of the character and then generates code for each character.

Once the data is encoded, it has to be decoded. Decoding is done using the same tree.

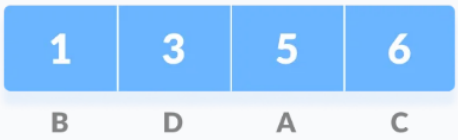
Huffman Coding prevents any ambiguity in the decoding process using the concept of **prefix code** ie. a code associated with a character should not be present in the prefix of any other code. The tree created above helps in maintaining the property.

Huffman coding is done with the help of the following steps.

1. Calculate the frequency of each character in the string

freequency of string

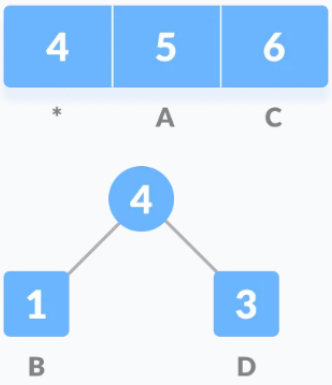
2. Sort the characters in increasing order of the frequency. These are stored in a priority queue Q.



Characters sorted according to the frequency

3. Make each unique character as a leaf node.

4. Create an empty node z. Assign the minimum frequency to the left child of z and assign the second minimum frequency to the right child of z. Set the value of the z as the sum of the above two minimum frequencies.

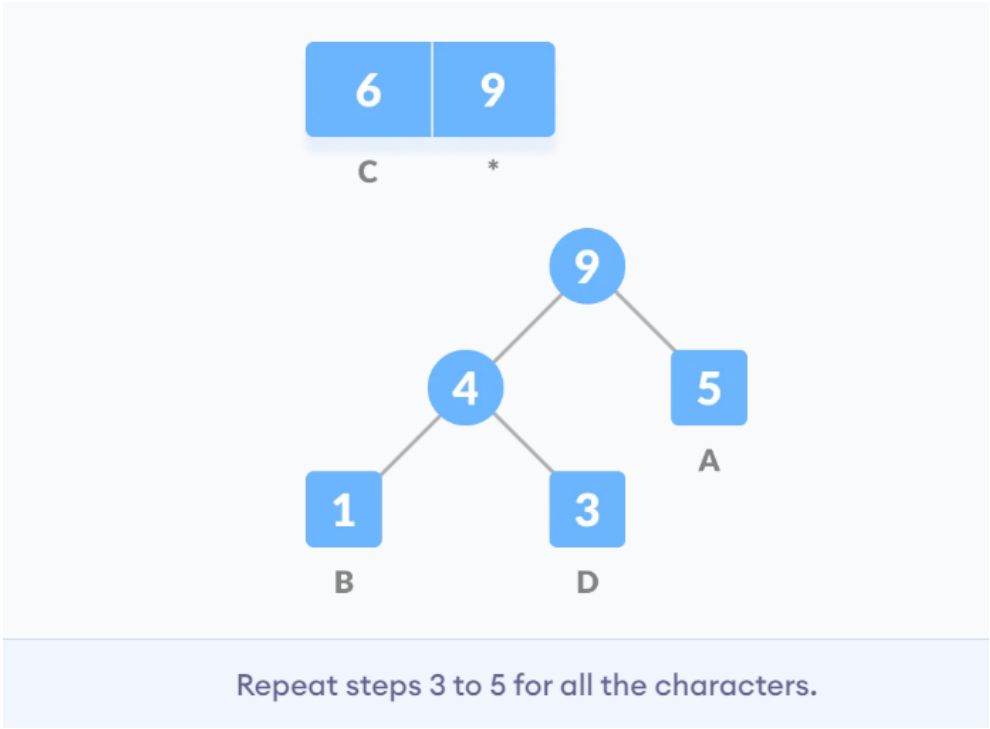


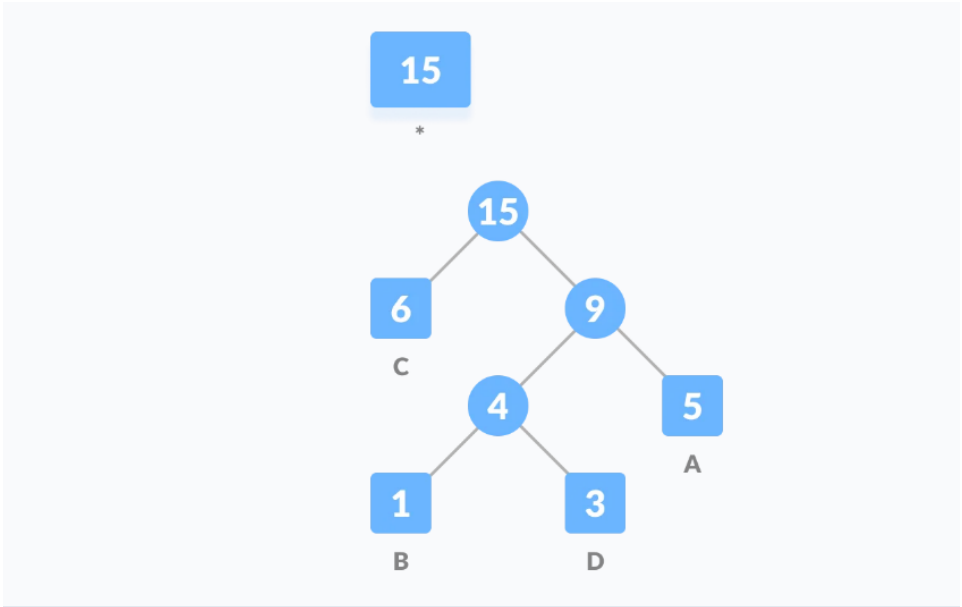
Getting the sum of the least numbers

5. Remove these two minimum frequencies from Q and add the sum into the list of frequencies (* denote the internal nodes in the figure above).

6. Insert node z into the tree.

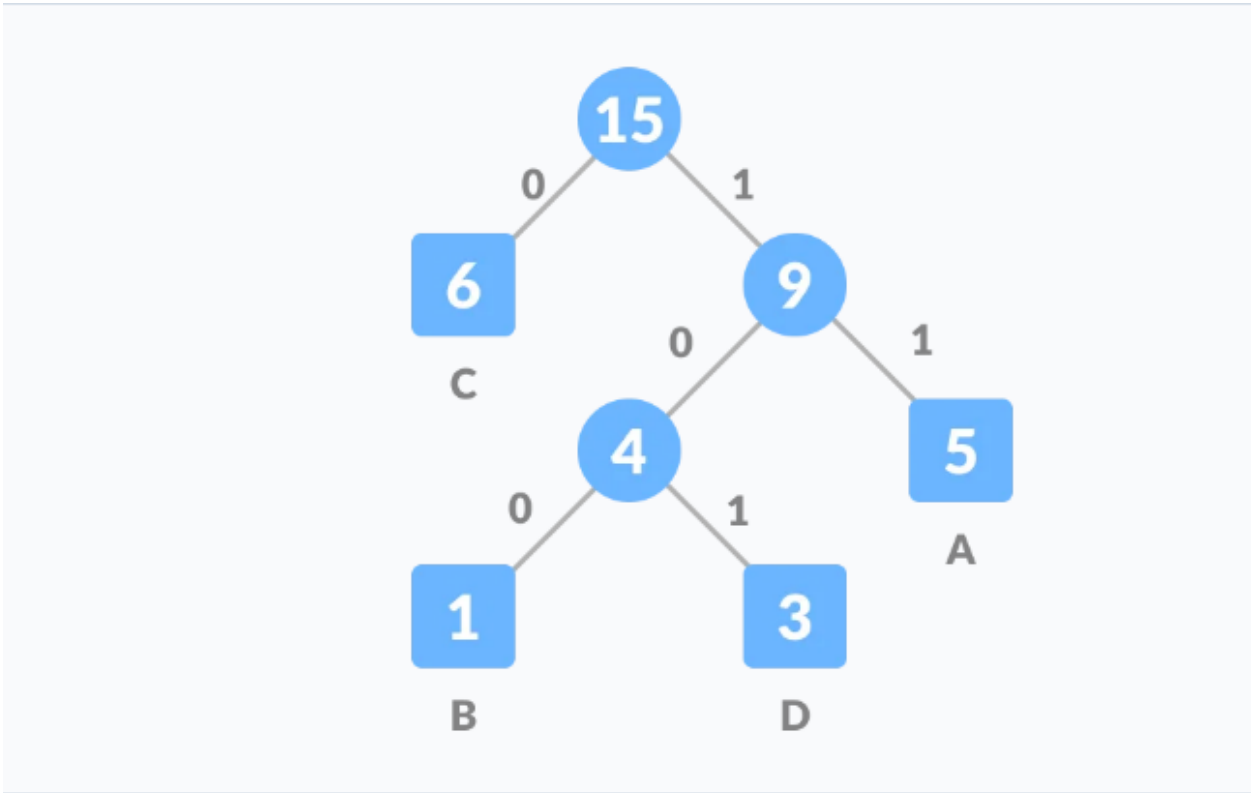
7. Repeat steps 3 to 5 for all the characters.





Repeat steps 3 to 5 for all the characters.

8. For each non-leaf node, assign 0 to the left edge and 1 to the right edge.



Assign 0 to the left edge and 1 to the right edge

For sending the above string over a network, we have to send the tree as well as the above compressed-code. The total size is given by the table below.

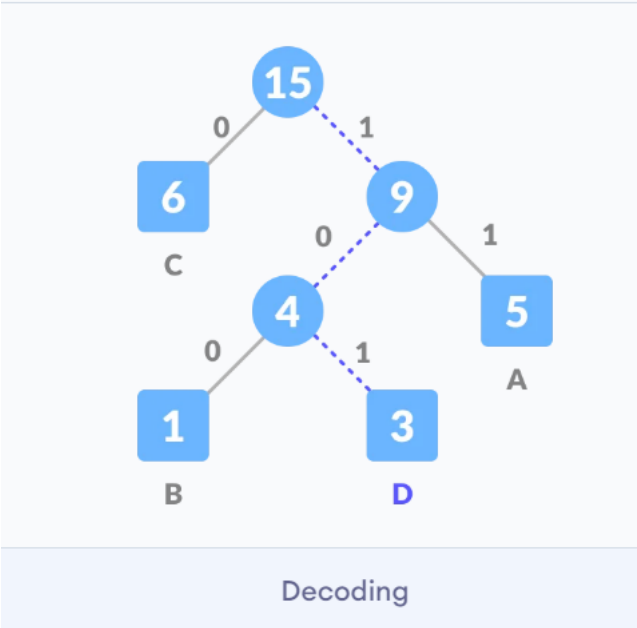
Character	Frequency	Code	Size
A	5	11	5*2 = 10
B	1	100	1*3 = 3
C	6	0	6*1 = 6
D	3	101	3*3 = 9
4 * 8 = 32 bits	15 bits		28 bits

Without encoding, the total size of the string was 120 bits. After encoding the size is reduced to 32 + 15 + 28 = 75.

Decoding the code

For decoding the code, we can take the code and traverse through the tree to find the character.

Let 101 is to be decoded, we can traverse from the root as in the figure below.



Decoding

Huffman Coding Algorithm

create a priority queue Q consisting of each unique character.
sort then in ascending order of their frequencies.
for all the unique characters:
create a newNode
extract minimum value from Q and assign it to leftChild of newNode
extract minimum value from Q and assign it to rightChild of newNode
calculate the sum of these two minimum values and assign it to the
value of newNode
insert this newNode into the tree
return rootNode

File Compression

What Does File Compression Mean?

File compression is a data compression method in which the logical size of a file is reduced to save disk space for easier and faster transmission over a network or the Internet. It enables the creation of a version of one or more files with the same data at a size substantially smaller than the original file. File compression is also known as file zipping.

File compression is enabled through a file or data compression software that creates a compressed version of each processed file. Typically, file compression works by scanning an entire file, identifying similar or repetitive data and patterns and replacing duplicates with a unique identifier. This identifier is usually much smaller in size than the original word and consumes less space. Thus, the size of the compressed file is considerably smaller.

ADVANTAGES OF DATA COMPRESSION:

The main advantages of compression are **reductions in storage hardware, data transmission time, and communication bandwidth**. This can result in significant cost savings. Compressed files require significantly less storage capacity than uncompressed files, meaning a significant decrease in expenses for storage.

List of Advantages:

- Less disk space (more data in reality)
- Faster writing and reading
- Faster file transfer.
- Variable dynamic range.
- Byte order independent.

Data structure used:

- Prority Queue
- Tree
- Huffman coding algorthim
- Arrays

Tools used for development process.

- Eclipse
- Windows Builder

Desging UI/Ux:

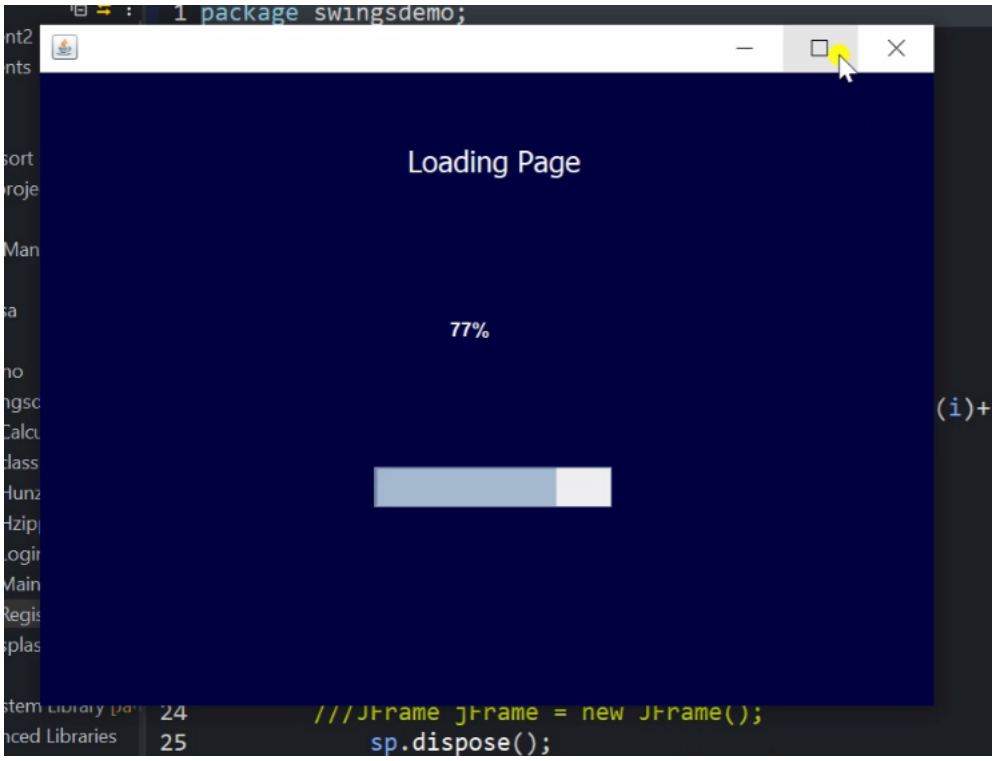
Git url;

Source Code:

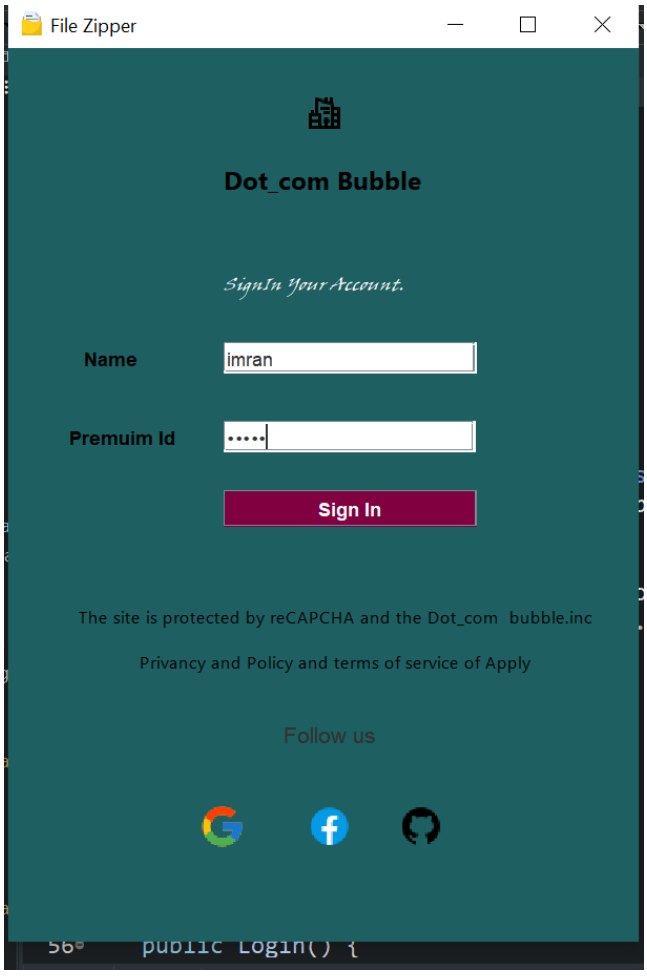
github available

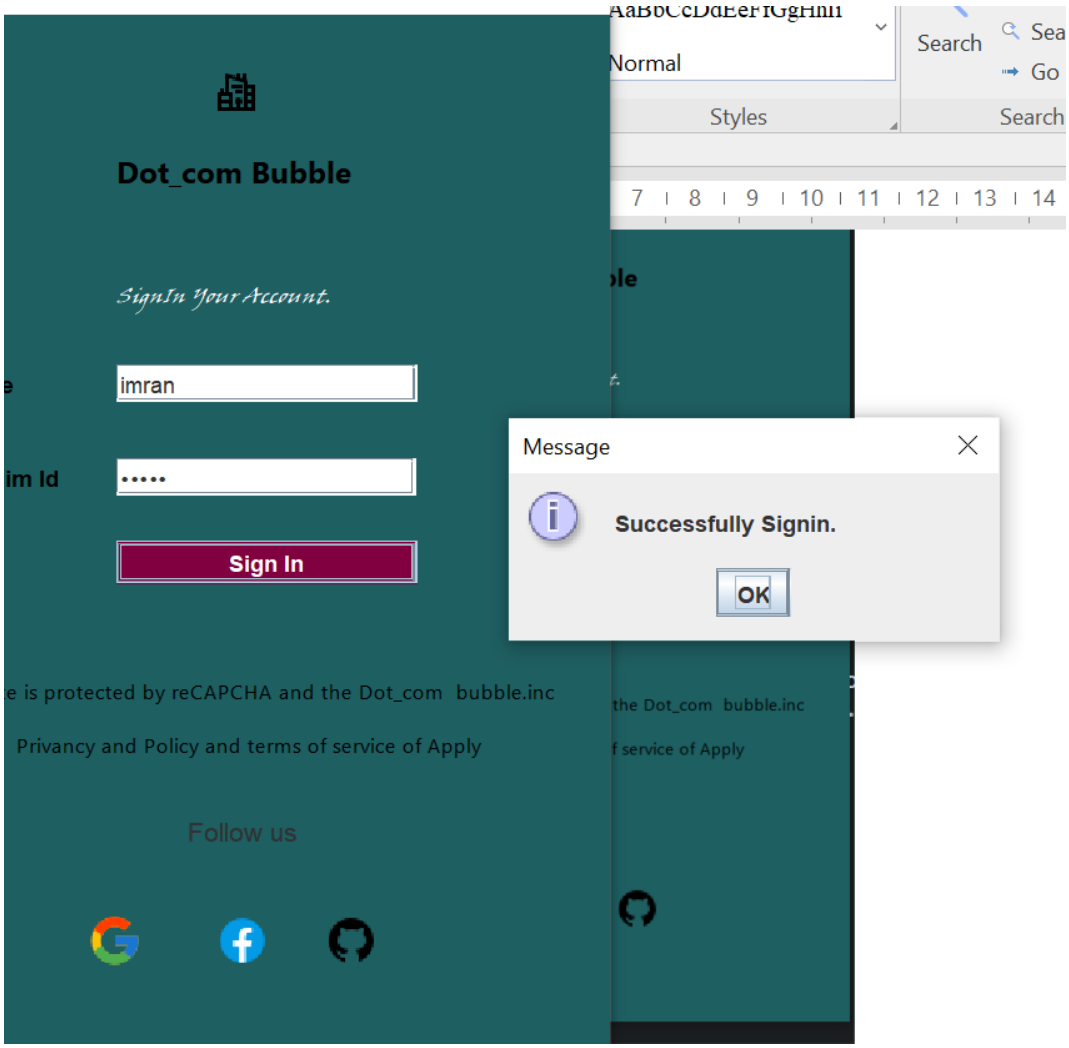
Output:

Splash screen

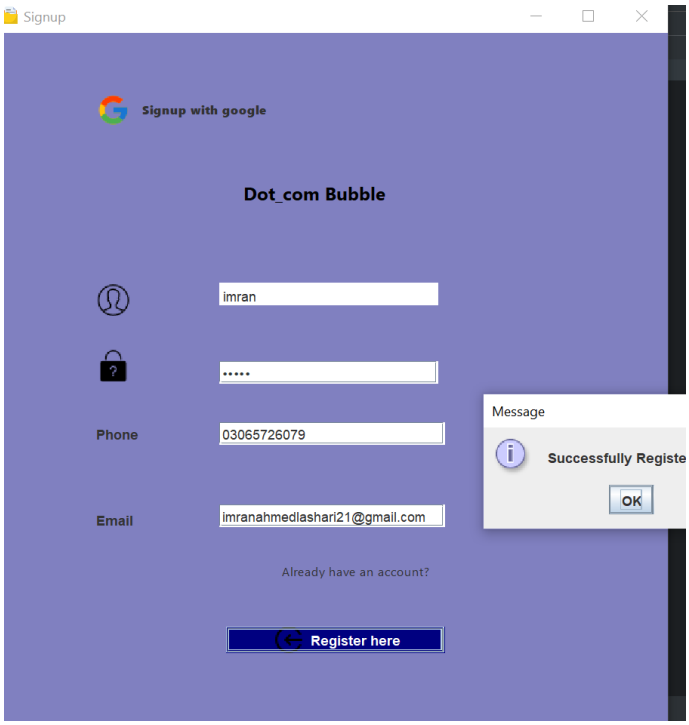


login Page

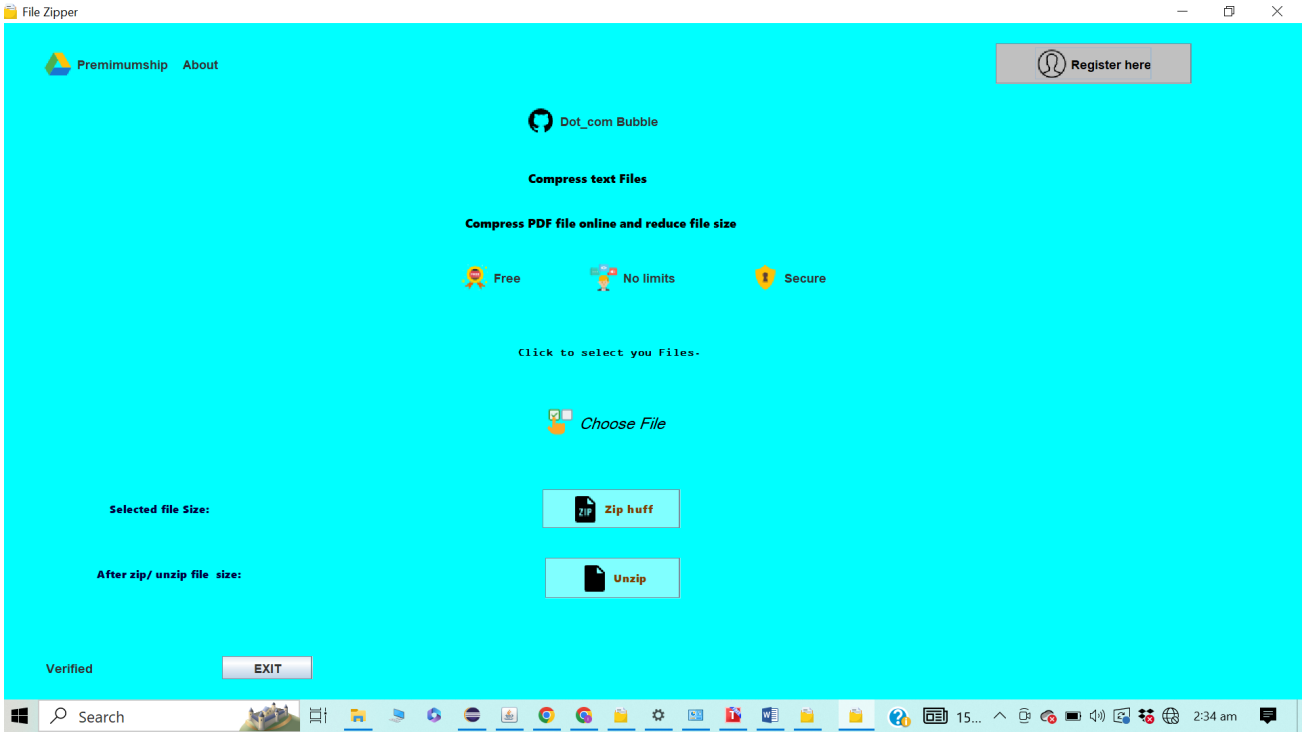


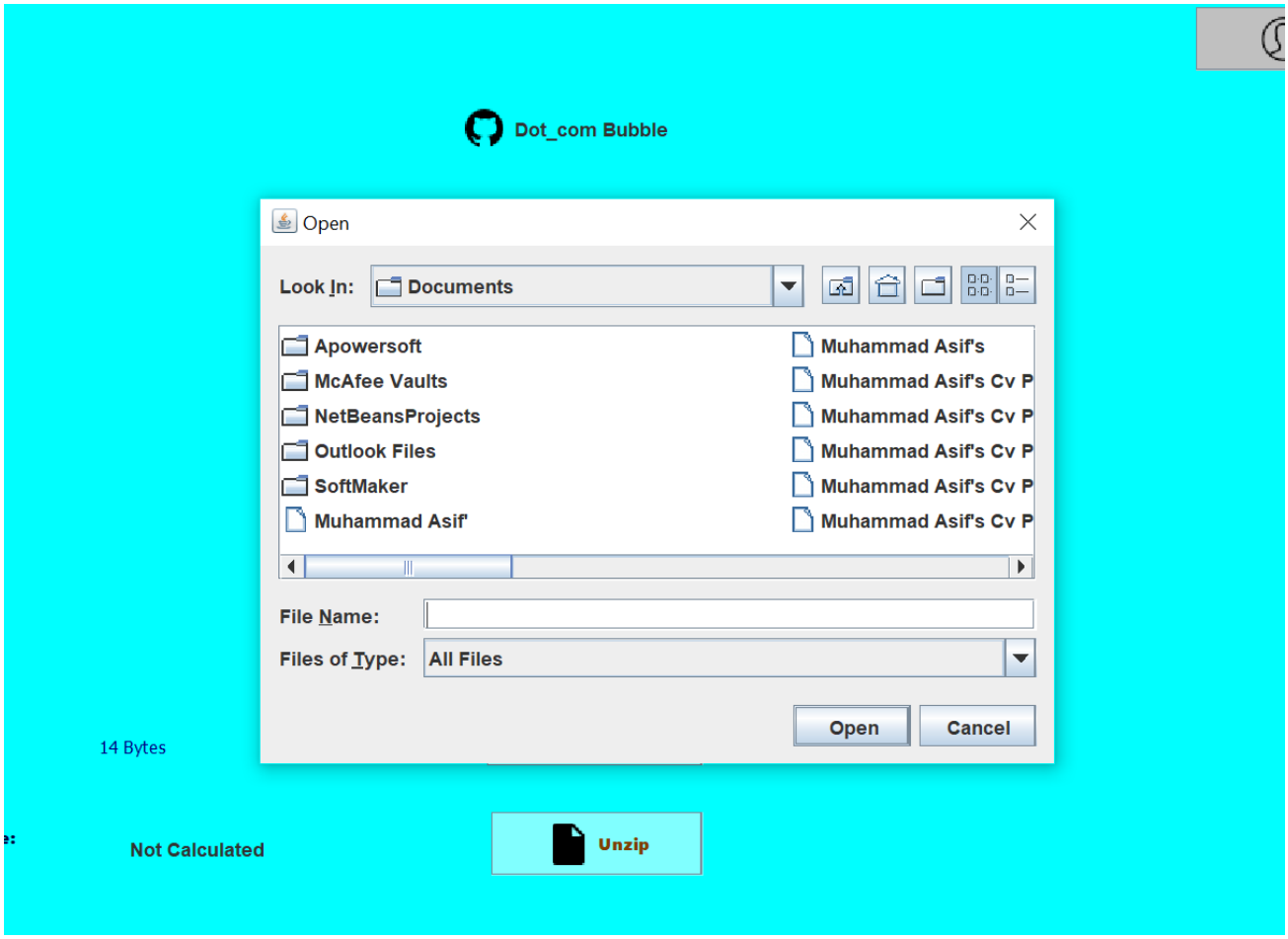

Normal
Styles
Search
Search
Dot_com Bubble
SignIn Your Account.
imran
Sign In
Message
Successfully Signin.
OK
e is protected by reCAPCHA and the Dot_com bubble.inc
Privacy and Policy and terms of service of Apply
Follow us

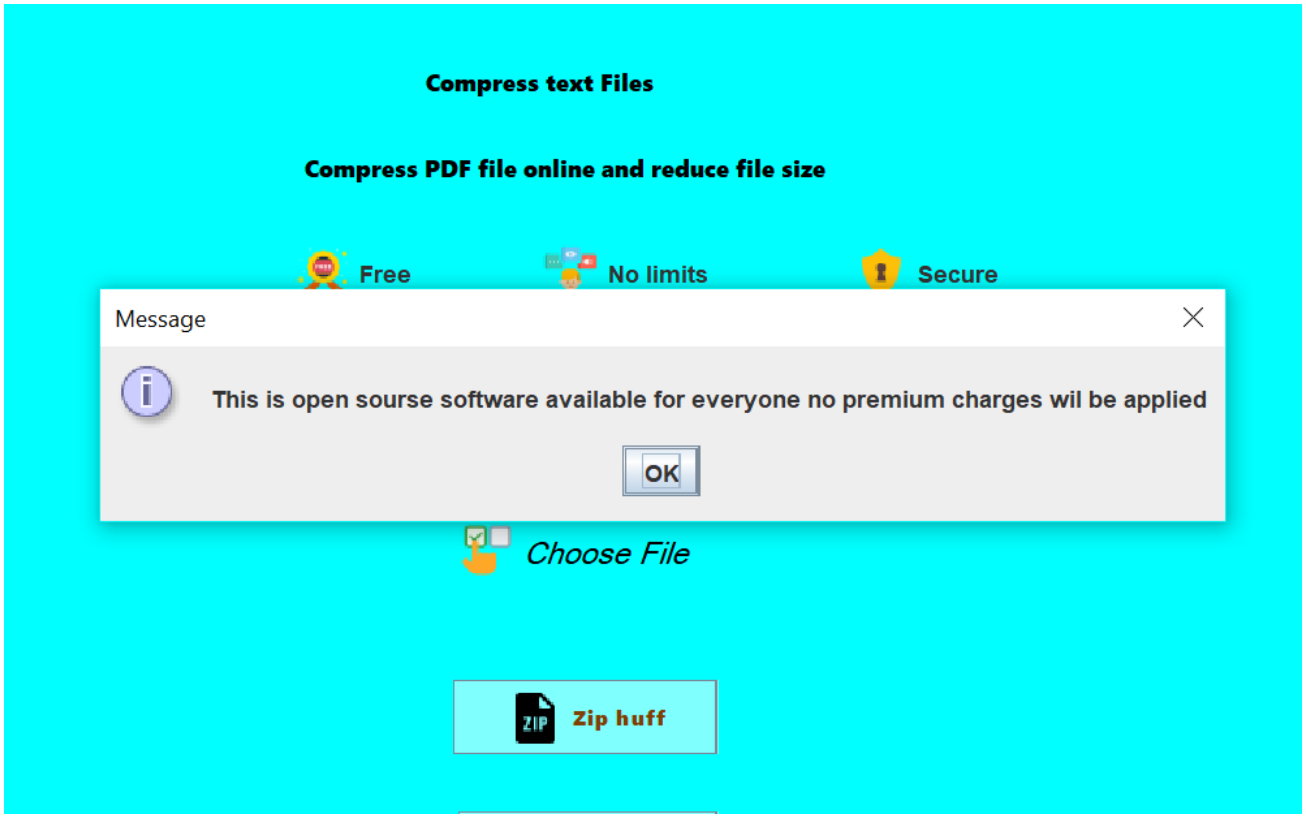
SignUp page


Signup
Signup with google
Dot_com Bubble
imran
Phone
03065726079
Email
imranahmedlashari21@gmail.com
Already have an account?
Register here
Message
OK

Main Menu


File Zipper
Premimumship
About
Register here
Dot_com Bubble
Compress text Files
Compress PDF file online and reduce file size
Free
No limits
Secure
Click to select you Files.
Choose File
Selected file Size:
Zip huff
After zip/ unzip file size:
Unzip
Verified
EXIT
Search


Dot_com Bubble
Open
Look In:
Documents
Apowersoft
McAfee Vaults
NetBeansProjects
Outlook Files
SoftMaker
Muhammad Asif'
Muhammad Asif's
Muhammad Asif's Cv P
Muhammad Asif's Cv P
Muhammad Asif's Cv P
Muhammad Asif's Cv P
Muhammad Asif's Cv P
Muhammad Asif's Cv P
File Name:
Files of Type:
All Files
Open
Cancel
14 Bytes
Not Calculated
Unzip


Compress text Files
Compress PDF file online and reduce file size
Free
No limits
Secure
Message
This is open sourse software available for everyone no premium charges wil be applied
OK
Choose File
Zip huff


Click to select you Files.
Choose File
Selected file Size:
14Bytes
Zip huff
After zip/ unzip file size:
8Bytes
Unzip
Verified
EXIT

Thank you....!