Project4 Report

Group member:

Qinxuan Shi   UFID: 8351-8162
Jiahui He   UFID: 7717-1034

How to start: (run in the windows terminal)

1. Start engine:
   cd .\src\
   erl -sname twitterEngine
   c(engine).
   engine:startEngine().
2. Start client:
   cd .\src\
   erl -sname client (to open more nodes you need different node name)
   c(client).
   Client:start().

What is working:

1. Register account and connect to the engine.
2. Disconnect from the engine.
3. Send tweets, including hashtags and mentions.
4. Users can subscribe to other's tweet accounts.
5. Users can retweet an interesting tweet to let his subscribers see.
6. Query tweets by hashtags, mentions and subscribers.

Implementation detail:

The engine.erl works as the twitter engine to provide services to all the clients. The client.erl works as the twitter client, users can enter commands to get different kinds of services.
Firstly, when the client starts, the user is required to enter his user name to register or connect back to the engine. Then the engine will send back the tweets queried by the tweets subscribed to and tweets that mentioned the user, after receiving the queried tweets, client will show all the tweets.

```
PS C:\UF\Fall-2022\DOSP\project4_1> cd .\src\
PS C:\UF\Fall-2022\DOSP\project4_1\src> erl -sname client2
Eshell V13.0.4  (abort with ^G)
(client2@DESKTOP-ERVAM65)1> client:start().
<8931.102.0>
(client2@DESKTOP-ERVAM65)2> Please enter user name to connect>qinxuan
(client2@DESKTOP-ERVAM65)2> Welcome qinxuan!
(client2@DESKTOP-ERVAM65)2> "admin" tweeted <admin0001> : hahaha @qinxuan #qatar !
(client2@DESKTOP-ERVAM65)2> -------------------------------------
```

If the user wants to subscribe to other users, he needs to enter the command "subscribe" first, then enter the name of the user he is going to subscribe to. After the engine receive the message from the client, it will update the list of users that subscribed to.

```
(client1@DESKTOP-ERVAM65)3> Enter>subscribe
(client1@DESKTOP-ERVAM65)3> Please enter the user you want to subscribe>qinxuan
(client1@DESKTOP-ERVAM65)3> subscribed qinxuan!
```

When the user is going to send or retweet, he needs to enter the command "send"/ "retweet" first, then enter the needed content according to the instructions. After the twitter engine receive the data from the client, the engine will extract the list of hashtags and mentions from the content string; then create the tweet id for the new tweet depending on the user name and the number of tweets the user has already sent, in the format of "admin0001". Then the engine will update both the list of tweets and users to store the new tweet and tweet id.

If the user wants to see the tweets sent by other users, he can enter the command "receive" and then follow the instructions to enter the querying conditions. Here we set three conditions: subscribed users, mentions and hashtags, user can choose different kinds of combinations of conditions to get the tweets he wants. The client will show the tweet as well as all the retweet connected to the tweet.

The user only needs to enter the "quit" command and then the client will close and the engine will delete the user from the connecting list.

```
(client2@DESKTOP-ERVAM65)2> Enter>quit
(client2@DESKTOP-ERVAM65)2> user disconnected!
```

The engine will also show the information while the user connects or disconnects.

```
(twitterEngine@DESKTOP-ERVAM65)3> user qinxuan registed
(twitterEngine@DESKTOP-ERVAM65)3> user qinxuan disconnected
```

Results:

The largest number of users we simulated is 2000 users.
Here is the result table for simulating connection:

Number of clients	Connection time (ms)
10	569.71
20	732.33
50	904.58
100	1337.66
200	1442.60
500	1576.27
1000	3215.09
2000	5783.21

Connection time (ms)
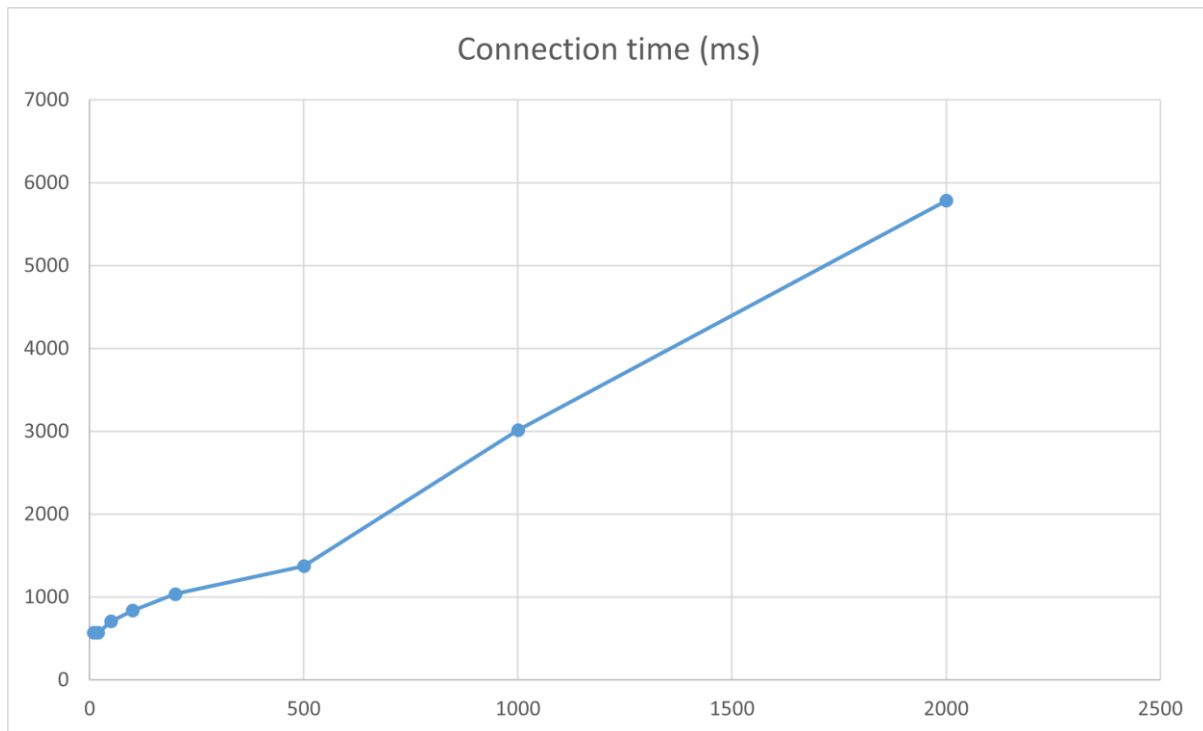
Zipf distribution is used to simulate the way users subscribe each other on Twitter as well as the frequency of users' tweeting and forwarding on Twitter. If the number of users is n, n-1 user subscribe user 1. Similarly, for the second user, (n-1)/2 users subscribe the user, while for the third user, (n-3)/3 users subscribe the third, and so on, for all users in the user list. About the tweet and retweet, the users with the most followers tweeted the most frequently. Set t as the Twitter frequency of the user with the largest number of subscribers, then the Twitter frequency of the nth user with the largest number of followers will be n/t.

Here is the result table for Zipf distribution:

Number of clients	Zipf distribution (ms)
10	21.89
20	47.53
50	127.71
100	357.91
200	875.17
500	1634.35
1000	3374.54
2000	7075.41

Zipf distribution (ms)

Connection time (ms)   Zipf distribution (ms)