**DATA STRUCTURE**

**TERM PROJECT**

**TWEET**

Student Number : 2014190729

Name : 이정현

1. What data structure you chose and why?

이번 프로젝트에서 가장 핵심적인 data structure는 graph입니다. 그 이유는 단지 배열로만 진행하려고 하였으나 교수님께서 주신 sample 크기가 너무 커서 graph로 했습니다. Vertex는 두 가지 종류가 있습니다. 첫번째는 사용자의 정보가 저장되어있는 vertices입니다. 사용자의 정보는 “user.txt”에 나와있는 정보들을 저장하는 방식으로 진행됩니다. 그리고 다른 하나는 words입니다. Words에는 tweet되어진 단어들의 정보를 저장합니다. 이는 “word.txt”에 있는 정보들을 저장하는 것입니다.

이렇게 vertices와 words를 저장하고 나서 이제 edge를 추가해야 합니다. 참고로 필자가 이번 텀 프로젝트에서 구현한 graph는 directed graph입니다. Vertices의 경우에는 edge 여부를 “friend.txt”에 근거하여 만들었습니다. 그리고 나중의 shortest path를 구하기 위해서 각각의 edge에 weight도 추가했습니다. Weight는 교수님께서 주신 프로젝트 명세에도 적혀 있듯이, direct 당하는 user의 친구 수를 기준으로 했습니다. 그 다음으로 words는 그 단어를 사용하는 사용자들을 알아야하기 때문에 vertices 와 연결을 시켰습니다. Words의 경우에는 edge에 weight가 없는 directed graph입니다.

Edge를 만들 때 그냥 vertex끼리만 이어지는 것이 아닌, adj라는 구조체를 사용했습니다. 그 이유는 정보를 수정할 때도 용이하고 edge에 대한 정보를 더 저장할 수 있기 때문입니다.

이렇게 모든 정보를 저장하는 Data Structure를 만들고 나서, 교수님이 요구하신 문제에 대해 해당 data를 찾아서, 혹은 수정하여 결과를 출력하도록 했습니다.

2. What is your expected performance?

제 term projrct는 일단 프로젝트 명세에 있는 모든 기능들을 다 구현했습니다. Data를 delete할 경우 다른 것들을 수행할 때 그 결과값이 다르게 출력 되도록 했습니다.

3. How would you improve the system in the future?

이번 텀프로젝트를 진행하면서 가장 혼란스러웠던 부분은 메모리 관리였습니다. 이전의 프로젝트들은 알고리즘을 어떻게 더 효율적으로 구성할 수 있을까에 관해서만 생각을 하다가 메모리 때문에 막히는 경험은 처음이엿습니다. 평상시에 data를 다룰 때는 수행시간의 차이를 못 느꼈는데, 이번에는 확실히 느꼈습니다. 다음번에 사회로 나가서는 더더욱 큰 데이터를 만지고 다룰 수 있으면 좋겟습니다.