

# CS 6140 Data Mining Project Proposal

## How ISIS uses twitter

### 1 Team Members

- Yulan Wang (u1264235)
- Guineng zheng (u1006611)
- Yu-ching Hsu (u1264420)

### 2 Data Set

#### Twitter

We plan to use the dataset collected from Twitter, which contains over 17,000 tweets from 100+ pro-ISIS fanboys from all over the world since the November 2005 Paris Attacks.

We got the dataset for our project from Kaggle open datasets site-  
<https://www.kaggle.com/fifthtribe/how-isis-uses-twitter>

### 3 Data Structure to Mine

We plan to do **Social Network Cluster Analysis** for the pro-ISIS twitter users. The main techniques would be Clustering, and several clustering methods will be used in our project.

### 4 Interestingness of the problem

Nowadays, we are in an information age where the network changes the modes of our life as well as the terrorist world. In this project, we try to use data mining techniques to analyze the social network of pro-ISIS that may help us know better the activity of ISIS and do some security work.

### 5 What we could learn

- How to apply data mining techniques in practice problem
- How to do data analysis and visualization
- How to analyze the network of social contact