EagleEye: A Crowdsourcing social media crime related App

Sibelius Seraphini, Harold R. Austin, Xiaohu Nian, Samresh Jyotishi

CSI 431 - Data Mining

Introduction

- Social Media (Twitter, Facebook, Youtube, and the blogosphere)

Motivation

- Can we extract crime information from the social media?

Applications

- Crowdsourcing crime visualization

Applications

- Crowdsourcing crime visualization
  - Crime information from the social media

Applications

- Crowdsourcing crime visualization
  - Crime information from the social media
  - Crime information directly from the users

Applications

- Crowdsourcing crime visualization
  - Crime information from the social media
  - Crime information directly from the users
- Safety Routing

Applications

- Crowdsourcing crime visualization
  - Crime information from the social media
  - Crime information directly from the users
- Safety Routing
  - Based on crime information provide a safety routing for one place to another

Applications

- Crowdsourcing crime visualization
  - Crime information from the social media
  - Crime information directly from the users
- Safety Routing
  - Based on crime information provide a safety routing for one place to another
- Crime Alert

Applications

- Crowdsourcing crime visualization
  - Crime information from the social media
  - Crime information directly from the users
- Safety Routing
  - Based on crime information provide a safety routing for one place to another
- Crime Alert
  - Panic Button alerts a new crime based on GPS information

Applications

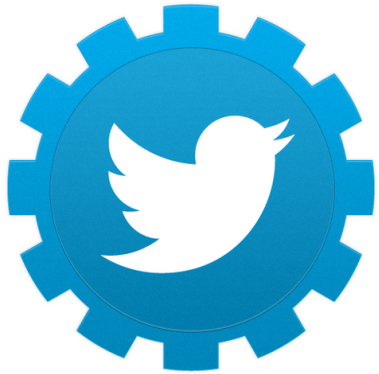
- Crowdsourcing crime visualization
  - Crime information from the social media
  - Crime information directly from the users
- Safety Routing
  - Based on crime information provide a safety routing for one place to another
- Crime Alert
  - Panic Button alerts a new crime based on GPS information
  - A espace safety route is provided to the user

Applications

- Crowdsourcing crime visualization
  - Crime information from the social media
  - Crime information directly from the users
- Safety Routing
  - Based on crime information provide a safety routing for one place to another
- Crime Alert
  - Panic Button alerts a new crime based on GPS information
  - A espace safety route is provided to the user
  - The police can be notified

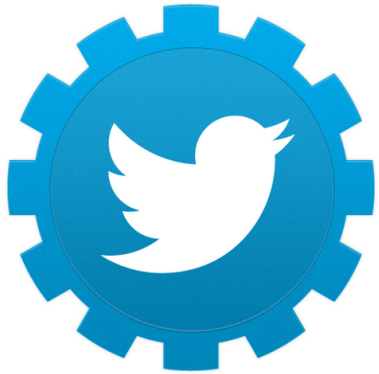
Social Media Crime Information Extraction

1. Crawl tweets from Twitter



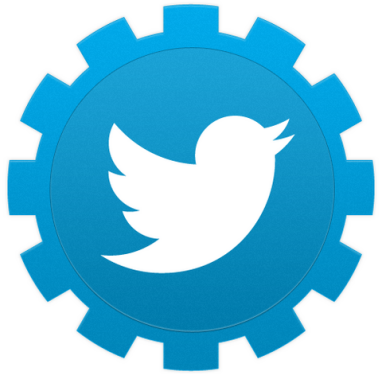
Social Media Crime Information Extraction

1. Crawl tweets from Twitter
2. Classify tweets in related to crime or not



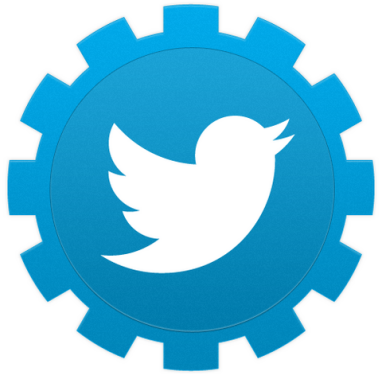
Social Media Crime Information Extraction

1. Crawl tweets from Twitter
2. Classify tweets in related to crime or not
3. Determine the type of crime



Social Media Crime Information Extraction

1. Crawl tweets from Twitter
2. Classify tweets in related to crime or not
3. Determine the type of crime
4. Extract Location of tweet



Crawl tweets from Twitter

- Streaming API
  - New York city region
  - Keywords related to crime

Classification of tweets

- Training Data
  - 14094 tweets manually labeled (5837 positives, 8257 negatives)
  - Types of Crimes:
    - Robbery, Shooting, Theft, Assault
    - Burglary, Vandalism, Arson, Drug Possession

Classification of tweets

- Training Data
  - 14094 tweets manually labeled (5837 positives, 8257 negatives)
  - Types of Crimes:
    - Robbery, Shooting, Theft, Assault
    - Burglary, Vandalism, Arson, Drug Possession
- Classifier of crime related Statistics
  - Precision: 0.98 (+/- 0.01)
  - Recall: 0.98 (+/- 0.01)
  - F1 Score: 0.98 (+/- 0.00)

Classification of tweets

- Training Data
  - 14094 tweets manually labeled (5837 positives, 8257 negatives)
  - Types of Crimes:
    - Robbery, Shooting, Theft, Assault
    - Burglary, Vandalism, Arson, Drug Possession
- Classifier of crime related Statistics
  - Precision: 0.98 (+/- 0.01)
  - Recall: 0.98 (+/- 0.01)
  - F1 Score: 0.98 (+/- 0.00)
- Classifier of type of crimes Statistics
  - Precision: 0.92 (+/- 0.01)
  - Recall: 0.93 (+/- 0.01)
  - F1 Score: 0.93 (+/- 0.01)

Extract Location of Tweet's text

- Preprocessing

Extract Location of Tweet's text

- Preprocessing
  - Lowercase

Extract Location of Tweet's text

- Preprocessing
  - Lowercase
  - Remove punctuation

Extract Location of Tweet's text

- Preprocessing
  - Lowercase
  - Remove punctuation
  - Apply alias dictionary

Extract Location of Tweet's text

- Preprocessing
  - Lowercase
  - Remove punctuation
  - Apply alias dictionary
- Match street addresses

Extract Location of Tweet's text

- Preprocessing
  - Lowercase
  - Remove punctuation
  - Apply alias dictionary
- Match street addresses
- Match state names

Questions ?