Social Media Intelligence in Practice: The NEREUS Experimental Platform

Dimitris Gritzalis & Vasilis Stavrou
Information Security & Critical Infrastructure Protection Laboratory
Dept. of Informatics | Athens University of Economics & Business
Presentation outline

⇒ Web 2.0 and Online Social Networks

⇒ Open Source and Social Media Intelligence

⇒ The NEREUS Framework

⇒ SOCMINT and behavior prediction capabilities

⇒ Conclusions
Web 2.0 and Online Social Networks (OSN)

Source: http://socialmediatoday.com/
Open Source Intelligence (OSINT) is produced from publicly available information, which is:

- Collected, exploited and disseminated in a timely manner
- Offered to an appropriate audience
- Used for the purpose of addressing a specific intelligence requirement

Publicly available information refers to (not only):

- Traditional media (e.g. television, newspapers, radio, magazines)
- Web-based communities (e.g. social networking sites, blogs)
- Public data (e.g. government reports, official data, public hearings)
- Amateur observation/reporting (e.g. amateur spotters, radio monitors)

Social Media Intelligence (SOCMINT) is produced from Online Social Networks and the Web 2.0
Revealing attitude towards law enforcement/infringement

<table>
<thead>
<tr>
<th>OSINT</th>
<th>OSN: YouTube</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means utilized for the analysis</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Science</th>
<th>Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computing</td>
<td>Machine Learning</td>
</tr>
<tr>
<td>Social Learning Theory</td>
<td></td>
</tr>
</tbody>
</table>

Applications:
(a). Assist in detecting attitude towards law enforcement/infringement
(b). Assist in detecting deviant behavior of minors
NEREUS: Architecture in a nutshell

*Flat data path*

- YouTube User
- Flat data transformation
- User classifier (naïve bayes)
- Naïve Bayes metrics
  - Classes: P: 72, N: 93
  - Precision: 72, Recall: 92, F-Score: 81, Accuracy: 81

*Comments classification path*

- Categories
  - Negatively Predisposed (P)
  - Not negatively predisposed (N)

- Video, uploads, lists & favorites classifier
- User classifier (Voter System)
- Comments results
- Storage
- Data preprocessing
- Comment classifier (LR)
- YouTube Crawler
- Anonymization layer
- YES
- Researchers’ compliance with ethical standards
  - YES
  - Legal Expert
  - Critical infrastructures
    - National security
    - Public interest

*Legend*

<table>
<thead>
<tr>
<th>Web 2.0 Medium:</th>
<th>YouTube</th>
<th>YouTube</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain Expert:</td>
<td>Sociologist</td>
<td>Political Scientist</td>
</tr>
</tbody>
</table>

Information Security & Critical Infrastructure Protection Laboratory

Ver. 2.11.14, 24.11.2014
The utmost importance of the social context

**Authoritarian Regimes**

Revealing personal attitude towards law enforcement/infringement will be used by the Regime against resisting pro-civic rights movements.

Pro-civic rights movements should prevent such platforms from being used by the Regime, using any available means.

**Democratic States**

Revealing personal attitude towards law enforcement/infringement may be used to protect Democracy from its opponents.

Democratic States may resist to social changes supported by, for example, grassroots political rights movements.

Democratic States may make use of such intrusive platforms, provided they are put under strict democratic control.
Revealing attitude towards law enforcement/infringement

Attitude towards law infringement

Study: Motive, anger, frustrations, predisposition towards law enforcement/infringement

Means: Machine Learning, comment classification, flat data classification.

- Individuals tend to **transfer online** their offline behavior
- Identify users’ **attitude towards law enforcement/infringement**
- Assist in detecting **delinquent behavior**
- Assist in predicting **deviant behavior of minors**
Dataset description

- Crawled YouTube and created dataset consists solely of Greek users.
- Utilized YouTube REST-based API (developers.google.com/youtube/):
  - Only publicly available data collected
  - Quote limitations (posed by YouTube) were respected
- Collected data were classified into three categories:
  - User-related information (profile, uploaded videos, subscriptions, favorite videos, playlists)
  - Video-related information (license, # of likes, # of dislikes, category, tags)
  - Comment-related information (comment content, # of likes, # of dislikes)

- A basic anonymisation layer added to the collected data:
  - MD5 hashes instead of usernames
• Comment classified into categories of interest:
  – Process performed as text classification
  – Machine trained with text examples and the category each one belongs to
  – Excessive support by field expert (Sociologist)

• Test set used to evaluate efficiency of resulting classifier:
  – Contains pre-labeled data fed to machine, labeled by field expert
  – Check if initial assigned label is equal to predicted one
  – Testing set labels assigned by field expert

• Most comments written in Greek/greeklish

• Conversion of greeklish text to Greek

• Categories of content defined:
  – Users with a negative attitude towards law enforcement
    (Predisposed negatively (P))
  – Users with a not negative attitude towards law enforcement
    (Not-predisposed negatively (N))
Comment classification using:
- Naïve Bayes (NB)
- Support Vector Machines (SVM)
- Logistic Regression (LR)

Classifiers efficiency comparison:
- Metrics (on % basis): Precision, Recall, F-Score, Accuracy

Logistic Regression algorithm:
- LR classifies a comment with 81% accuracy

Precision: Measures the classifier exactness. Higher and lower precision means less and more false positive classifications, respectively.

Recall: Measures the classifier completeness. Higher and lower recall means less and more false negative classifications, respectively.

F-Score: Weighted harmonic mean of both metrics.

Accuracy: No. of correct classifications performed by the classifier. Equals to the quotient of good classifications performed by all data.
Analysis based on flat data

- Addressing the problem from a different perspective:
  - Connection between users of category P and confidence of accuracy of comments belonging to category P.
  - Verify (or not) the results of the Machine Learning approach based on flat data (Flat Data).

Blue: Users of category P classified on the basis of the comment-oriented tuple (Flat Data).
Red: Users of category P classified on the basis of their comments only (Machine Learning).

- Data transformation:
  - User represented by a tuple (username, content of comment, video ID the comment refers to, country, age, genre, # of subscribers, # of video views).

- Machine trained by a set of users of categories P and N:

1721 users are (almost certainly) negatively predisposed towards law enforcement/infringement.
Selected observations

- **6%** of comments (among 2,000,000 collected) express negative attitude towards respecting the law (i.e., positive to law infringement)
- **3.5%** of videos (among 200,000 collected) classified into a specific category of interest
- **14%** of users (among 13,000 collected) express negative attitude towards respecting the law (i.e., positive to law infringement)

Ability to assist in predicting delinquent behaviour of minors
- Violent behaviour
- Cyber bullying
- Emotional or sexual harassment
General conclusions

✓ SOCMINT can transform into intelligence the vast amount of data produced by Web 2.0.
✓ SOCMINT is an intrusive technology and could put in danger civic rights.
✓ SOCMINT utilization is not - and should not be considered as - a solely technical issue.
✓ SOCMINT could assist in predicting attitude towards law infringement.
✓ SOCMINT could assist in predicting delinquent behavior of minors.
References


