Open Source Intelligence from Online Social Networks: Identifying Insiders

Miltos Kandias, Vasilis Stavrou

September 2014
Open Source Intelligence from Online Social Networks: Identifying Insiders

Miltos Kandias, Vasilis Stavrou

Information Security & Critical Infrastructure Protection Laboratory Dept. of Informatics, Athens University of Economics & Business, Greece
Outline

- Online Social Networks (OSN)
- Open Source Intelligence (OSINT)
- The insider threat
- Collecting and analyzing data from OSN
- Behavior prediction capabilities
  - **Case 1**: Success story - Insider detection and narcissism
  - **Case 2**: Success story - Predicting delinquent behavior
  - **Case 3**: Success story - Detecting stress levels
  - **Case 4**: Horror story - Revealing political beliefs
- Ethical and legal issues
- Conclusions
Online Social Networks

- OSN and Web 2.0 enable users to add online content.
- Content can be crawled and utilized for:
  - personalized advertising,
  - personalized content promotion and
  - user/usage profiling
- Can content be crawled and utilized for:
  - User behavior prediction?
  - User psychosocial characteristics extraction?
  - Proactive cyber defense?
What happens online in 60 sec

Source: http://socialmediatoday.com/
Open Source Intelligence (OSINT)

• Open Source Intelligence is produced from publicly available information, which is:
  – collected, exploited and disseminated in a **timely** manner,
  – offered to an **appropriate** audience and
  – used for the purpose of addressing a specific **intelligence requirement**.

• Publicly available information refers to (not only):
  – traditional media (e.g. television, newspapers, radio),
  – web-based communities (e.g. social networking sites, blogs),
  – public data (e.g. government reports, official data, hearings) and
  – amateur observation and reporting (e.g. amateur spotters, radio monitors).
A generic model for predicting threats

**Data:** Directly available information
- email
- Inter/Intranet traffic
- Remote access traffic
- Social media
- Geospatial data
- Calendar & local documents

**Observation:** Inference from data that reflects a specific state
- Web sites
- Instant messaging
- File size
- HR/performance information
- Instant scripts
- Location
- Authentication attempts

**Indicators:** Action/event as evidence of precursor to inferred behavior
- Psychosocial characteristics
- Disregard for policies
- Unauthorized access attempts
- Data harvesting
- Suspicious communications

**Behavior:** Sequence of actions associated with a specific purpose

Source: US Pacific Northwest National Laboratory
Insider Threat

- The insider threat is a severe problem in cyber/corporate security, which originates from persons who:
  - are legitimately given access rights to information systems,
  - misuse privileges and
  - violate security policy.

Data can leak to outsider and have an impact:

- Data:
  - Trade secrets
  - Account numbers
  - Social Security Numbers
  - Intellectual property

- Network storage
  - Shared folders
  - Removable devices
  - Transmitted data

- Competitor
  - Regulator
  - Unauthorized personnel
  - Press or media

- Company defamation
  - Monetary expenses
  - Legal liabilities
  - Asset loss
  - Customer relations
  - End business
# Insider Threat: When is its impact high?

<table>
<thead>
<tr>
<th>Technical literacy</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal process knowledge</td>
<td>High</td>
<td>Highest impact</td>
</tr>
<tr>
<td>Low</td>
<td>Potentially significant impact</td>
<td>Insignificant impact</td>
</tr>
</tbody>
</table>

Insignificant (though demonized) impact

Source: Gartner Group, Report 5605
Insider Threat severity

- Insider threat is recognized as one of the most important security issues for 2014:
  - The case of former US Government contractor E. Snowden is casting a shadow over 2014.
- Insiders consist the top source of data breaches.
- Priority to protect the network from the insider threat.

Sources:
ZDNet Asia IT Priorities Survey 2008/09
Insider Threat severity

Source:
2011 Cyber Security Watch Survey: How Bad Is the Threat?, CERT, Carnegie Mellon University, USA
Threat parameters

• We have a threat when:
  – At least one attacker is adequately motivated.
  – Opportunity to unleash attack exists.
  – At least one vulnerability exists.
  – Attacker is skilled enough.

• Given sufficient motive, time and budget every system is vulnerable.
Threat parameters

A threat consists of:

- Motive
- Opportunity
- Vulnerability
- Skills
Malevolent user characteristics

• Malevolent user needs:
  – **Opportunity** to unleash prepared attack.
  – **Motive** to attack.
  – Ability to **overcome inhibitions**.
  – **Appropriate stimulation** and impulse.

• Under certain circumstances every user is vulnerable to diverge towards delinquency.
Malevolent user characteristics

Malevolent user requirem’s:

- Opportunity
- Motive
- Ability to overcome inhibitions
- Stimulus/impulse
Personal factors (Shaw)

• Insiders’ characteristics:
  – Inward turning,
  – resistance to the fulfillment of individual will,
  – turn to ICT to change moods (anxiety, depression etc.),
  – lack a strong and stiff code of ethics,
  – lack faithful adherence to government, leader or cause,
  – narcissism,
  – lack capacity to recognize emotions and
  – negative predisposition towards authorities, laws, government, state.

Personal factors (Shaw)

- Introversion
- Social and personal frustrations
- Computer dependency
- Ethical “flexibility”
- Reduced loyalty
- Entitlement – Narcissism
- Lack of empathy
- Predisposition towards law enforcement
According to the FBI most insiders share:

- desire to possess wealth,
- vanity and narcissism - anger and revenge syndrome,
- problematic attitude towards coworkers,
- divided loyalty,
- urge to feel sudden quiver of excitement or emotion,
- ability to overcome inhibitions,
- narcissistic behavior,
- Establishment in the favor or good graces of others,
- self destructive tendencies and
- problems with relatives.

Personal factors (FBI)

- Greed/financial need
- Anger/Revenge
- Problems at work
- Ideology/Identification
- Divided loyalty
- Adventure/Thrill
- Vulnerability to blackmail
- Ego/self-image (Narcissism)
  - Ingratiation
  - Compulsive and destructive behavior
- Family problems
The threat

Motive
• Opportunity
• Vulnerability
• Skills

Threat consists of:

Malevolent user requirem’s:

Opportunity
• Motive
• Ability to overcome inhibitions
• Stimulus/impulse

Introversion
• Social and personal frustrations
• Computer dependency
• Ethical “flexibility”
• Reduced loyalty

Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Personal factors (Shaw)

Personal factors (FBI)

Greed/financial need
• Anger/Revenge
• Problems at work
• Ideology/Identification
• Divided loyalty
• Adventure/Thrill
• Vulnerability to blackmail

Ego/self-image (Narcissism)
• Ingratiation
• Compulsive and destructive behavior
• Family problems

Reduced loyalty
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Ethical “flexibility”
• Reduced loyalty
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement

Lack of empathy
• Predisposition towards law enforcement

Predisposition towards law enforcement
• Entitlement-Narcissism
• Lack of empathy
• Predisposition towards law enforcement
Delinquent behavior prediction theories

- **General Deterrence Theory (GDT):** Person commits crime if expected benefit outweighs cost of action.
- **Social Bond Theory (SBT):** Person commits crime if social bonds of attachment, commitment, involvement and belief are weak.
- **Social Learning Theory (SLT):** Person commits crime if associates with delinquent peers.
- **Theory of Planned Behavior (TPB):** Person’s intention (attitude, subjective norms, perceived behavioral control) towards crime is a key factor in predicting his behavior.
- **Situational Crime Prevention (SCP):** Crime occurs when both motive and opportunity exist.
Case 1
Scope: Insider threat prediction based on Narcissism

<table>
<thead>
<tr>
<th>OSINT</th>
<th>OSN: Twitter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Tools used for the analysis

<table>
<thead>
<tr>
<th>Science</th>
<th>Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computing</td>
<td>Graph Theory</td>
</tr>
<tr>
<td>Sociology</td>
<td>Theory of Planned Behavior</td>
</tr>
<tr>
<td></td>
<td>Social Learning Theory</td>
</tr>
</tbody>
</table>
Case 1: Insider threat prediction based on Narcissism

- Individuals tend to transfer offline behavior online.
- Convicted insiders do share this personality trait (narcissism).
- Utilize graph theoretic tools to perform analysis.
- Detection via social media popularity and usage intensity.
- Trait of narcissism relates to delinquent behavior via:
  - sense of entitlement,
  - lack of empathy,
  - anger and “revenge” syndrome and
  - inflated self-image.

Narcissistic behavior detection

Study: Motive, ego/self-image, entitlement

Means: Usage Intensity, Influence valuation, Klout score
Focus on a Greek Twitter community:
- Context sensitive research.
- Utilize ethnological features rooted in locality.
- Extract and analyze results.

Analysis of content and measures of user influence and usage intensity.

User categories: follower, following and retweeter.

Graph:
- Each user is a node.
- Every interaction is a directed edge.

41.818 fully crawled users (personal and statistical data)
- Name, ID, personal description, URL, language, geolocation, profile state, lists, # of following/followers, tweets, # of favorites, # of mentions, # of retweets.
• **Strongly connected components:**
  – There exists 1 large component (153.121 nodes connected to each other) and several smaller ones

• **Node Loneliness:**
  – 99% of users connected to someone

• **Small World Phenomenon:**
  – Every user lies <6 hops away from anyone
Graph Theoretical approach

- **Indegree Distribution:**
  - # of users following each user
  - Average 13.2 followers/user
Graph Theoretical approach

- **Outdegree Distribution:**
  - # of users each user follows
  - Average 11 followers/user
Graph Theoretical approach

- **Usage Intensity Distribution:**
  Weighted aggregation of {# of followers, # of followings, tweets, retweets, mentions, favorites, lists}

![Graph showing usage intensity distribution with an important cluster of users highlighted.](image-url)
Narcissism detection

- Majority of users make limited use of Twitter.
  - A lot of “normally” active users and very few “popular” users.
  - Users classified into 4 categories, on the basis of specific metrics (influence valuation, Klout score, usage valuation).

- Above a threshold:
  - User becomes quite influential/perform intense medium use.
  - User get a “mass-media & persona” status.

The excessive use of Twitter by persons who are not mass-media or personas connects to narcissism and identifies narcissists, i.e. persons who - inter alia - tend to turn insiders

<table>
<thead>
<tr>
<th>Category</th>
<th>Influence valuation</th>
<th>Klout score</th>
<th>Usage valuation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loners</td>
<td>0 - 90</td>
<td>3.55 - 11.07</td>
<td>0 - 500</td>
</tr>
<tr>
<td>Individuals</td>
<td>90 - 283</td>
<td>11.07 - 26.0</td>
<td>500 – 4.500</td>
</tr>
<tr>
<td>Known users</td>
<td>283 – 1.011</td>
<td>26.0 - 50.0</td>
<td>4.500 – 21.000</td>
</tr>
</tbody>
</table>
In a nutshell

- **Twitter Users**
- **Content generation**
- **Twitter**
- **Crawling & storing**
- **Content Aggregator**
- **Usage intensity valuation**
- **Indegree/Outdegree aggregator**
- **Influence valuation**
- **User classification according to categories**

**List of categories:**
- Loner
- Individual
- Known User
- Mass media & personas

**Klout score server**

**Collector**

**Klout score queries**
## Case 2

**Scope:** Revealing negative attitude towards law enforcement

<table>
<thead>
<tr>
<th>OSINT</th>
<th>OSN: YouTube</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tools used for the analysis</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Science</strong></td>
<td><strong>Theory</strong></td>
</tr>
<tr>
<td>Computing</td>
<td>Machine Learning</td>
</tr>
<tr>
<td></td>
<td>Data Mining</td>
</tr>
<tr>
<td>Sociology</td>
<td>Social Learning Theory</td>
</tr>
</tbody>
</table>
Case 2: Revealing negative attitude towards law enforcement

**Law enforcement predisposition**

**Study**: Motive, anger, frustrations, predisposition towards law enforcement

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

• Individuals tend to transfer offline behavior online.
• Extract results about users’ negative attitude towards law enforcement and authorities.
• Trait of negative attitude towards law enforcement is connected to delinquent behavior via:
  – sense of entitlement,
  – lack of empathy,
  – *anger and revenge syndrome* and
  – inflated self-image.
Dataset: General parameters

- 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) and
  - comment-related information (comment content, # of likes, # of dislikes).

- A basic anonymization layer added to the collected data:
  - MD5 hashes instead of usernames.
Small World Phenomenon:
- Every user of the community is 6 hops away from everyone else.
Graph Theory and Content Analysis

• **Indegree Distribution:**
  – Presentation of statistical distribution of incoming edges per node.
• **Outdegree Distribution:**
  - Presentation of statistical distribution of outgoing edges per node.
Graph Theory and Content Analysis

• Tag Cloud:
  – Axis of content of the collected data via tag cloud analysis.
Graph Theory and Content Analysis

- **YouTube’s nature:**
  - Popular social medium, emotional-driven responses, audio-visual stimuli, alleged anonymity, users interact with each other, contains political content.
How was the analysis performed?

We assessed a YouTube user negative attitude towards law enforcement by classifying YouTube content using (a) Machine Learning and (b) Flat Data

• Machine Learning
  – Examined YouTube’s videos via their comments
  – Performed comment classification as text classification
  – Content analysis based on comments and videos (via their comments)

• Flat Data
  – An assumption-free method
  – An easy-to-scale method
How was the analysis performed?

<table>
<thead>
<tr>
<th>Approach</th>
<th>Metrics</th>
<th>Machine Learning</th>
<th>Flat Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classifier</td>
<td>Logistic Regression</td>
<td>Naïve Bayes</td>
<td></td>
</tr>
<tr>
<td>Classes</td>
<td>P</td>
<td>N</td>
<td>P</td>
</tr>
<tr>
<td>Precision</td>
<td>86</td>
<td>76</td>
<td>72</td>
</tr>
<tr>
<td>Recall</td>
<td>74</td>
<td>88</td>
<td>92</td>
</tr>
<tr>
<td>F-Score</td>
<td>80</td>
<td>81</td>
<td>81</td>
</tr>
<tr>
<td>Accuracy</td>
<td>81</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Both approaches achieve similar results
- Flat data behaves slightly more efficiently (better f-score)
- Flat data performs faster
- Flat data verifies the results obtained by Machine Learning
• 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 are written in Greek – greeklish comments exist.

• Training sets (greeklish, greek) were merged - One classifier was trained.

• Two categories of content were defined:
  – Users with a **negative** attitude (**P**redisposed negatively (**P**)).
  – Users with a **not negative** attitude (**N**ot-predisposed negatively (**N**)).
Machine Learning (2)

- **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**

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classes</td>
<td>NBM</td>
</tr>
<tr>
<td>Precision</td>
<td>P</td>
</tr>
<tr>
<td>Recall</td>
<td>P</td>
</tr>
<tr>
<td>F-Score</td>
<td>P</td>
</tr>
<tr>
<td>Accuracy</td>
<td>P</td>
</tr>
</tbody>
</table>

- **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 by all data.
• **Video** classification:
  – Examination of a video on the basis of its comments.
  – Voter process to determine category classification.

• **(Video) Lists** classification:
  – Voter process to determine category classification (same threshold).

• Conclusions about **user behavior**:  
  – If there is at least one category P attribute then the user is classified into category P.
Example of conclusion extraction (1/2)

- Each comment falls into a category (P or N) based on the classifier’s prediction.
- Each video falls into a category based on its comments.

<table>
<thead>
<tr>
<th>Comment</th>
<th>Classifier’s output</th>
<th>Likes</th>
<th>Dislikes</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>P</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>#2</td>
<td>P</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>#3</td>
<td>N</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>#4</td>
<td>P</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>#5</td>
<td>N</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>#6</td>
<td>P</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

Video “Example”

Only comments #2 and #4 will be fed to the voter (if N, then ignore. If no likes and at least 1 dislike, then ignore).

Video contains (at least) 2 negatively predisposed comments. Thus, it falls into category P.

- The voter decides on the basis of the number of P comments (category P).
- Comments with only dislikes and no likes are excluded.
- Same method applies to list of videos (instead of comments), i.e. user’s uploaded videos, favourite videos, and playlists.
To decide over the user’s behaviour the following parameters are examined.

- The voter decides based on a vector of the above 4 attributes.
- Security Officer may determine threshold cut-off, as well as minimum number of category P comments taken into consideration by voter.
- False positive case: User is classified as P although she is not
  - Such cases mean that there might be an indication that further examination is needed so as to decide whether a user shares this psychosocial characteristic or not.
Flat Data (1)

• Addressing the problem from a different perspective:
  – assumption-free and easy-to-scale method,
  – verify (or not) the results of the Machine Learning approach,
  – machine trained by a set of users of categories P and N.

• 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 user test set (Sociologist served as field expert).

<table>
<thead>
<tr>
<th>Naïve Bayes metrics</th>
<th>P</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classes</td>
<td>72</td>
<td>93</td>
</tr>
<tr>
<td>Precision</td>
<td>92</td>
<td>73</td>
</tr>
<tr>
<td>Recall</td>
<td>81</td>
<td>82</td>
</tr>
<tr>
<td>F-Score</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>81</td>
<td></td>
</tr>
</tbody>
</table>
Flat Data (2)

• Connection between users of category P and confidence of accuracy of comments belonging to category P.

**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).

1721 users are (almost certainly) negatively predisposed towards law enforcement.
Case 3

Scope: Detecting stress level usage patterns (overall and over time)

<table>
<thead>
<tr>
<th>OSINT</th>
<th>OSN: Facebook</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tools used for the analysis</td>
<td></td>
</tr>
<tr>
<td>Science</td>
<td>Theory</td>
</tr>
<tr>
<td>Computing</td>
<td>Machine Learning</td>
</tr>
<tr>
<td></td>
<td>Data Mining</td>
</tr>
<tr>
<td>Sociology</td>
<td>Social Learning Theory</td>
</tr>
</tbody>
</table>
Case 3: Detecting stress level usage pattern (overall and over time)

**Stress level detection**

**Study:** User’s overall and over time stress level

**Means:** Machine Learning, flat data classification, chronicity analysis.

- Individuals tend to transfer offline behavior online.
- Extract results about usage pattern depicted stress level.
- Analyze each user under the prism of stress level both overall and over time (chronicity analysis).
- High stress has been found to:
  - Make individuals vulnerable to fall prey to third parties.
  - Overcome moral inhibitions.
- Analysis is based on Social Learning Theory and stress correlations are based on Beck’s Anxiety Inventory stress test.
Dataset: General parameters

- Crawled Facebook & created dataset solely by **Greek** users.
- Users offered informed consent.
- Utilized Facebook’s Graph API:
  - Only publicly available data collected.
  - De facto respect of users’ privacy settings.
- Collected data were classified into four categories:
  - User information (friends list and profile description),
  - user-generated content (statuses, comments and links),
  - groups of interest (communities, events and activities) and
  - interests (music, actors, sports, books etc.).
- A basic anonymization layer added to the collected data:
  - MD5 hashes instead of usernames.
- Opt-out ability integrated, delete all user data upon selection.
- Dataset statistical analysis proved its efficiency and absence of bias.
Dataset: General parameters

- 405 users
- 12,346 user groups
- 98,256 liked objects
- 171,054 statuses
- 250,027 comments
Flat classification (overall indicators)

- Goal: extract correlations between usage patterns and users who share same stress valuation (according to BAI test).
- Transformed relational database into a single tuple record containing solely users’ comments and statuses.
- Flat data tuple subjected to stemming process.
- EM algorithm produced 3 clusters:
  - Cluster 0 has too few users.
  - Cluster 1 includes users with high and medium-to-high stress score.
  - Cluster 2 includes users with low and medium-to-low stress score.
Chronicity analysis (indicators over time)

• Goal: **detect differentiations** of OSN usage patterns over **time related** to depicted stress level.

• Split users’ usage pattern into time periods (from one day to one month).
  – Time period of one week produced best results.

• Chronicity analysis system consists of 2 modules:
  – Preprocessing data module (responsible for the processing of input data).
  – Usage pattern analysis module (responsible for analyzing usage patterns based on a set of metrics).

• Usage pattern fluctuations depict differentiated medium usage.
# Chronicity analysis steps

**Step 1:** Classify user generated content into 4 predefined categories ('S' stands for sports, 'M' for music, 'P' for politics and 'Mi' for miscellaneous).

<table>
<thead>
<tr>
<th>Classifier</th>
<th>NBM</th>
<th>SVM</th>
<th>MLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classes</td>
<td>S</td>
<td>M</td>
<td>P</td>
</tr>
<tr>
<td>Precision</td>
<td>71</td>
<td>92</td>
<td>79</td>
</tr>
<tr>
<td>Recall</td>
<td>77</td>
<td>86</td>
<td>85</td>
</tr>
<tr>
<td>F-Score</td>
<td>74</td>
<td>89</td>
<td>81</td>
</tr>
<tr>
<td>Accuracy</td>
<td>79</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chronicity analysis steps

**Step 2:** Calculate following metrics for each user and time period (metrics developed on an ad-hoc basis according to our observations).

<table>
<thead>
<tr>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of posts regarding sports</td>
</tr>
<tr>
<td>Frequency of posts regarding music</td>
</tr>
<tr>
<td>Frequency of posts regarding politics</td>
</tr>
<tr>
<td>Frequency of posts regarding miscellaneous</td>
</tr>
<tr>
<td>Interest shift per interest pair</td>
</tr>
<tr>
<td>Average frequency of posting</td>
</tr>
<tr>
<td>Average frequency of commenting</td>
</tr>
<tr>
<td>Major interests</td>
</tr>
<tr>
<td>Minor interest shift frequency</td>
</tr>
<tr>
<td>Frequency of uploading photos</td>
</tr>
<tr>
<td>CommentedBy ratio</td>
</tr>
<tr>
<td>StatusVarianceFlattened</td>
</tr>
<tr>
<td>CommentVarianceFlattened</td>
</tr>
</tbody>
</table>
Chronicity analysis steps

**Step 3:** Transform metrics results into arithmetic vectors and perform data mining on them using (a) *K*-means, (b) *EM*, and (c) *Canopy* algorithms. Utilize voter to decide fluctuations.
Chronicity analysis results

- Metrics results per detected cluster.

<table>
<thead>
<tr>
<th>Cluster id</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>7%</td>
<td>16%</td>
<td>8%</td>
<td>3%</td>
<td>1%</td>
<td>7%</td>
<td>9%</td>
</tr>
<tr>
<td>TotalComments</td>
<td>3</td>
<td>78</td>
<td>93</td>
<td>5</td>
<td>79</td>
<td>410</td>
<td>44</td>
</tr>
<tr>
<td>TotalPosts</td>
<td>588</td>
<td>227</td>
<td>513</td>
<td>185</td>
<td>704</td>
<td>914</td>
<td>292</td>
</tr>
<tr>
<td>SportsFreq</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>MusicFreq</td>
<td>0.02</td>
<td>0.34</td>
<td>0.61</td>
<td>0.05</td>
<td>0.17</td>
<td>0.43</td>
<td>0.28</td>
</tr>
<tr>
<td>PoliticsFreq</td>
<td>0.00</td>
<td>0.06</td>
<td>0.02</td>
<td>0.00</td>
<td>0.04</td>
<td>0.05</td>
<td>0.15</td>
</tr>
<tr>
<td>MiscellaneousFreq</td>
<td>0.02</td>
<td>0.22</td>
<td>0.09</td>
<td>0.04</td>
<td>0.13</td>
<td>0.18</td>
<td>0.22</td>
</tr>
<tr>
<td>PhotosFreq</td>
<td>0.68</td>
<td>0.08</td>
<td>0.06</td>
<td>0.39</td>
<td>0.40</td>
<td>0.10</td>
<td>0.13</td>
</tr>
<tr>
<td>CommentsFreq</td>
<td>0.05</td>
<td>0.41</td>
<td>0.57</td>
<td>0.03</td>
<td>0.61</td>
<td>2.42</td>
<td>0.26</td>
</tr>
<tr>
<td>StatusesFreq</td>
<td>10.39</td>
<td>1.36</td>
<td>3.32</td>
<td>1.56</td>
<td>6.33</td>
<td>5.29</td>
<td>1.99</td>
</tr>
<tr>
<td>MinorInterestSiftFreq</td>
<td>0.01</td>
<td>0.18</td>
<td>0.13</td>
<td>0.02</td>
<td>0.12</td>
<td>0.16</td>
<td>0.24</td>
</tr>
<tr>
<td>CommentedBy ratio</td>
<td>0.08</td>
<td>1.25</td>
<td>0.73</td>
<td>0.27</td>
<td>0.65</td>
<td>1.52</td>
<td>0.88</td>
</tr>
<tr>
<td>StatusDispersalFlattened</td>
<td>34.89</td>
<td>3.01</td>
<td>4.93</td>
<td>7.85</td>
<td>15.20</td>
<td>7.98</td>
<td>6.28</td>
</tr>
<tr>
<td>CommentDispersalFlattened</td>
<td>0.01</td>
<td>1.52</td>
<td>0.95</td>
<td>0.06</td>
<td>1.08</td>
<td>5.95</td>
<td>0.60</td>
</tr>
</tbody>
</table>
Chronicity analysis results

- Visual representation of users belonging to each cluster.
- Clusters 0 and 3 contain mainly users classified in high stress category.
- Clusters 1 and 7 contain many users classified in medium or low stress category.
- In cluster 0, users post mainly photos.
- In cluster 3 users post photos, discuss about music, whereas a small fraction of the content is referring to miscellaneous information.
- Clusters 1 and 7 refer mainly to music and miscellaneous content and also contain limited content referring to sports.
# Case 4

**Scope:** Identifying Political Beliefs

<table>
<thead>
<tr>
<th>OSINT</th>
<th>OSN: YouTube</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tools used for the analysis</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Science</strong></td>
<td><strong>Theory</strong></td>
</tr>
<tr>
<td>Computing</td>
<td>Machine Learning</td>
</tr>
<tr>
<td></td>
<td>Data Mining</td>
</tr>
<tr>
<td>Political Sociology</td>
<td></td>
</tr>
</tbody>
</table>
Case 4: **Horror story** – Identifying Political Beliefs

**Divided loyalty**

- Study: Motive, ideology, divided/reduced loyalty, predisposition towards law enforcement
- **Means:** Machine Learning, Content Analysis, comment classification

- Same YouTube dataset.
- Political beliefs profiling-clustering.
- Three (indicative, local context based) clusters: Radical - Neutral – Conservative.
- Machine Learning and Content Analysis methods used.
- Analysis also based on:
  - Social Learning Theory
  - General Deterrence Theory
Methodology

• Three (indicative) categories: Radical, Neutral, Conservative:
  – Assumptions are local-context-dependent (Greece, 2007-12).
  – Test case consists of an indicative subset of the local community.
  – Analysis reflects the current local political scene.

• Defined (indicative) classes:
  – Radical political affiliation: center-left, left, far-left.
  – Neutral political affiliation: neutral or non-specified political affiliation disclosed.
  – Conservative political affiliation: center-right, right, far-right.

• Comments classification:
  – Comments classification performed as text classification.
  – Machine trained with text examples and the category each one belongs to.
  – Assistance of field expert (Sociologist).
Analysis of results

- **Comment** classification by:
  - Naïve Bayes Multinomial (NBM)
  - Support Vector Machines (SVM)
  - Multinomial Logistic Regression (MLR)

- Each classifier’s **efficiency** was compared by:
  - Metrics (%): Precision, Recall, F-Score, Accuracy

- Multinomial Logistic Regression was chosen:
  - MLR classifies appropriately a comment with 87% accuracy.
  - Use of precision, recall and f-score to further examine classifiers' efficiency.

---

<table>
<thead>
<tr>
<th>Classifier</th>
<th>NBM</th>
<th>SVM</th>
<th>MLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classes</td>
<td>R</td>
<td>N</td>
<td>C</td>
</tr>
<tr>
<td>Precision</td>
<td>65</td>
<td>93</td>
<td>55</td>
</tr>
<tr>
<td>Recall</td>
<td>83</td>
<td>56</td>
<td>85</td>
</tr>
<tr>
<td>F-Score</td>
<td>73</td>
<td>70</td>
<td>60</td>
</tr>
<tr>
<td>Accuracy</td>
<td>68</td>
<td>84</td>
<td>87</td>
</tr>
</tbody>
</table>
Example of conclusions extraction (1/2)

- Each **comment** falls into a category, based on the classifier’s prediction.
- Each **video** falls into a category based on its comments.

<table>
<thead>
<tr>
<th>Comment</th>
<th>Political affiliation</th>
<th>Likes</th>
<th>Dislikes</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>R</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>#2</td>
<td>C</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>#3</td>
<td>R</td>
<td>30</td>
<td>5</td>
</tr>
<tr>
<td>#4</td>
<td>N</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>#5</td>
<td>R</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>150</td>
<td>40</td>
</tr>
</tbody>
</table>

All comments are taken into account.

Like means “approve” and dislike means “not approve”.

\[
R = (1+90/150-10/40) + (1+30/150-5/40) + (1+10/150-3/40) = 4.1 \\
C = (1+15/150-20/40) = 0.6 \\
R>C thus the video is classified as R
\]

- Two sums are calculated (i.e. R,C).
- For every comment in a video we calculate: \(1+\{(\text{likes}/(\text{total_likes}))-(\text{dislikes}/(\text{total_dislikes}))\}\).
- The greater sum indicates the category the video falls into.
- Same applies to list of videos, i.e. user’s uploaded videos, favourite videos and playlists.
Example of conclusions extraction (2/2)

- User content is classified. Then, four parameters (comment, upload, favourite, playlist) are calculated by the voter in order to identify user’s beliefs.

- Each parameter gets a weight from the voter (as user comments forms the most direct way to express opinions).

- Search for indications of direct political affiliation expression in the uploaded and favourite videos and in the playlists.

User classification example:

<table>
<thead>
<tr>
<th>User “Example”</th>
<th>Political beliefs</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments</td>
<td>R</td>
<td>3</td>
</tr>
<tr>
<td>Uploaded videos</td>
<td>N</td>
<td>2</td>
</tr>
<tr>
<td>Favorite videos</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>Playlists</td>
<td>C</td>
<td>1</td>
</tr>
</tbody>
</table>

R = 3 + 2 = 5
C = 1

R > C, i.e. the user belongs to category R
2% of **comments** demonstrate political affiliation (0.7% Radical, 1.3% Conservative)

- 2% means that almost **41.000 comments** (of the 2.000.000 collected) include political content.

7% of **videos** classified into a specific category (2% Radical, 5% Conservative)

- 7% means that almost **14.000 videos** (of the 200.000 collected) include political content.

12% of **users** express **Radical** political affiliation and 40% **Conservative** affiliation

- 52% means that **6.760 users** reveal - one way or another - their political beliefs.
Basic observations (2/2)

• Radicals:
  – 20% of their comments includes political position.
  – Prefer Greek alphabet (i.e., 54% comments in Greek, 33% in greeklish, 13% use both).
  – Massively comment on specific videos.
  – Prefer videos with political content (political events, music, incidents of police brutality).
  – Add to their favourites documentaries and political music clips.

• Conservatives:
  – Prefer greeklish in comments (i.e., 55% greeklish, 35% Greek, 10% both).
  – Often share conspiracy-based or videos with nationalistic content.

• Greeklish comments are usually shorter and aggressive.
• Greek comments are usually explanatory, polite and longer.
• The more aggressive a comment - the more misspelled.
• 7% of videos published under Creative Commons license.
  – 55% uploaded by Radicals, 10% by Conservatives, 35% by Neutrals.
OSN data exploitation paths

• **Insider threat prediction:**
  – Adopting Shaw and FBI psychosocial indicators (narcissism, anger or revenge syndrome, etc.).

• **Delinquent behavior prediction:**
  – Analysis of psychosocial characteristics (narcissism, anger or revenge syndromes, etc.).
  – Predisposition analysis (graph theory and content analysis through social learning theory, etc.).

• **Forensics analysis support:**
  – Suspect profiling and analysis (proactive prediction of delinquent behavior, etc.).
Ethical and legal issues

• Users are **not** aware of the actual reach of the information they reveal.
• Some methods used for **OSINT** may:
  – be associated with discrimination,
  – infringe human rights (freedom of speech, conception of identity, privacy, etc.),
  – cause self-censorship and self-oppression and
  – pose a threat of marginalization (employers or rigid micro-societies).
• OSN often offer privacy options which **do not really** help.
• Private profiles are usually **indirectly crawlable**.
• **Laws** may not clearly prohibit the process of data revealing psychosocial characteristics.
• Derogations are often allowed:
  – On a legal manifest of public interest (e.g. critical infrastructures, security officers, etc.).
  – If given an explicit, informed and written consent of the person concerned.
Some conclusions

✓ OSN produce vast amounts of crawlable information and OSINT may transform this information into intelligence.

✓ OSINT can assist in detecting narcissistic behavior, predisposition towards law enforcement, divided political loyalty, etc.

✓ OSINT can be a proactive cyber-defense tool and predict insider threat, predict delinquent behavior, assist in law enforcement.

✓ OSINT may lead to unwanted horror stories.

✓ OSINT intrusive nature dictates limited use, e.g. security officers selection, critical infrastructure protection.
References


