Stress level detection via OSN usage pattern and chronicity analysis: An OSINT threat intelligence module

Miltiadis KANDIAS, **Dimitris GRITZALIS, Vasilis STAVROU**, **Kostas NIKOLOULIS**

Information Security & Critical Infrastructure Protection (INFOSEC) Laboratory
Dept. of Informatics, Athens University of Economics & Business (AUEB)
76 Patission Ave., Athens GR-10434 Greece
{ kandiasm@aueb.gr, dgrit@aueb.gr, stavrouv@aueb.gr, knikoloulis@gmail.com }

Abstract. Online Social Networks (OSN) are not only a popular communication and entertainment platform but also a means of self-representation. In this paper, we adopt an interdisciplinary approach combining Open Source Intelligence (OSINT) and user-generated content classification techniques with a user-driven stress test as applied to a Greek community of OSN users. The main goal of the paper is to study the chronicity of the stress level users experience, as depicted by OSN user generated content. In order to achieve that, we investigate whether collected data are able to facilitate the process of stress level detection. To this end, we perform unsupervised flat data classification of the user-generated content and formulate two working clusters which classify usage patterns that depict medium-to-low and medium-to-high stress levels respectively. To address the main goal of the paper, we divide user-generated content into chronologically defined sub-periods in order to study potential usage fluctuations over time. To this extent, we follow a process that includes (a) content classification into predefined categories of interest, (b) usage pattern metrics extraction and (c) metrics and clusters utilisation towards usage pattern fluctuation detection both through the prism of users' usual usage pattern and its correlation to the depicted stress level. Such an approach enables detection of time periods when usage pattern deviates from the usual and correlates such deviations to user experienced stress level. Finally, we highlight and comment on the emerging ethical issues regarding the classification of OSN user-generated content.

Keywords: Online Social Networks (OSN), Open Source Intelligence (OSINT), Privacy, Usage Pattern Deviation, Stress Detection, Insider Threat, Threat Intelligence.

1 Introduction

Modern mediated means of communication have been radicalised by the introduction of Web 2.0 (O'Reilly, 2009), where users are not mere observers. Instead, they are able to contribute and thus become content generators. As a result, another aspect of users' behaviour and personality is manifested within the context of Online Social Networks. Initial debate on whether users present their real self when being online or not has concluded to differentiated results; Amichai-Hamburger and Vinitzky indicated that users tend to transfer their offline behaviour online (Amichai-Hamburger and Vinitzky, 2010) while Tufekci (2013) introduces some conceptual issues which question that. These results may be used to further study human behaviour and personality through the prism of OSN usage patterns and what they depict.

These studies have introduced and widened the area of psychosocial characteristics detection and extraction, making it possible to examine users' OSN behaviour (Rogers, Smoak and Liu, 2006). This area has been further developed by the use of Open Source Intelligence (Best, 2008) which is defined as *"data produced from publicly available information that is collected, exploited, and disseminated in a timely manner to an appropriate audience for the purpose of addressing a specific intelligence requirement"* by the US Dept. of Defense. Further research has indicated that a variety of psychosocial characteristics (e.g., predisposition towards law enforcement, narcissism, various personality traits etc.) may be extracted via Open Source Intelligence (OSINT), produced in the context of OSN and by further data mining (Kandias et al., 2013a) or graph theoretic processing (Kandias et al., 2013b). Regarding possible applications of the results, research indicates that when a person experiences stress, she is more vulnerable to fall prey to third parties, overcome moral inhibitions, or manifest deviant behaviour (Greitzer et al., 2012; Shaw et al., 1998; FBI, 2012). These research papers indicate that stress can be an important indicator when analysing potential insiders, in advanced Threat Intelligence feeds and in forensics analysis of past incidents.

Our major contribution relies on the development of a method that extract results and statistics over the stress level that usage patterns depict through the prism of the "chronicity" of OSN usage. The term chronicity refers to the usage deviations that users may exhibit in the context of OSN over time. To this extent, we selected a Greek community of OSN users (after accepting their informed consent) and classified its members to detect their stress levels over time.

Initially, we conducted unsupervised flat data classification to explore the dataset and formed two working clusters that detect overall usage patterns. The first cluster included mainly users with medium-to-high and high stress level, according to the stress test, while the second cluster included mainly users with low and medium-to-low stress level. In this process, the results of the user-driven stress test were utilised only to calculate the amount of users who were classified in each cluster. One could argue that these results could be linked to the Big-5 trait of neuroticism (Hayes and Joseph, 2003), but analysing this is beyond the scope of this paper and will be addressed in our future work.

In order to study deviations that appeared in the users' OSN usage pattern we utilised machine learning techniques so as to categorise user content into predefined categories of interest (sports, music, politics and miscellaneous). We also examined usage patterns over time in order to form clusters of similar characteristics. For doing so, we developed a series of usage chronicity metrics and fed them to a classification/polling scheme, similar to other relevant classification methods (Veeramachaneni and Arnaldo, 2016), that decides over the usage fluctuations of the OSN. Forming clusters of similar (OSN-usage) time periods enables detection and matching of usage patterns and time fractions. We tested several time-windows, varying from one day to one month, and observed that the time-span of one week suits better the purposes of our research. As a result, rarely observed usage patterns are considered potential deviations from usual, which may mean that the user has experienced real-time events that have consequently affected the OSN usage pattern. We have formed and evaluated eight usage pattern clusters that share similar characteristics and correlated them with users' stress levels.

These methods inherently include a series of ethical issues along with conflicts between stakeholders. Regarding the ethical issues, the methods may interfere with user privacy and lead to a social threat. Regarding the conflicts, it is understandable that law enforcement, crime prevention agencies or even employers or HR managers may request to "connect the dots" and combine information about potential suspects or employees respectively. At the same time, citizens/employees demand that their privacy be respected. Thus, processing user-generated content could be ethically acceptable solely under strict terms and conditions and given the informed consent of the user.

The paper is organised as follows: in section 2 we review the existing literature. In section 3 we describe the dataset collection methodology. In section 4 we describe statistical and content analysis of the dataset. In section 5 we describe the flat data classification method. In section 6 we describe the chronicity analysis and quote our results. In section 7 we discuss the ethical issues. Section 8 concludes the research results.

2 Related work and motivation

The advent of Web 2.0 has contributed to the transformation of the average user from a passive reader into a content contributor. Web 2.0 and OSN have, in particular, become a valuable source of personal data, which are available for crawling and processing, allowing for the extraction of a variety of evaluations through them. Data Science was applied to Psychology issues only recently. The existing literature includes studies that focus on the area of personality, psychopathology and behaviour in relation to OSN usage.

2.1 Psychology-focused studies

Regarding personality traits and OSN usage, research highlights that in the context of OSN, users reveal their actual personality rather than self-idealisation (Back et al., 2010). This is further supported by another research that highlights a strong connection between personality and Facebook usage (Amichai-Hamburger and Vinitzky, 2010), supporting that users tend to transfer their offline behaviour online. To this extent, Rosenberg and Egbert (2011) investigates the utility of personality traits and secondary goals as predictors of self-presentation tactics employed by Facebook users. Furthermore, the Big-5 model (Hayes and Joseph, 2003) is studied through the prism of OSN usage. Ross et al. (2009) investigates how the 5-Factor Model of personality relates to Facebook use. They demonstrate a series of correlations between the Big-5 elements and the Facebook usage (e.g. individuals high on the trait of extraversion belong to more Facebook groups but do not have significantly more friends). Similar results are shown by Bachrach et al. (2012) about how users' activity on Facebook relates to their personality, as measured by the standard Big-5 Model, where they attempt to predict personality traits on the basis of Facebook usage patterns. Another research examines the influence of all Big-5 personality traits on Facebook usage and the interactions of traits in this context based on Torgersen's typological approach (La Sala, Skues and Grant, 2014). Further research based on the Big-5 personality traits and the use of OSN reveals differential relationships between them and highlights the ability to extract correlations (Hughes, Rowe, Batey and Lee, 2012; Quercia et al., 2012; Correa et al., 2010; Ryan and Xenos, 2011). In these studies, personality traits are studied (openness, conscientiousness, extraversion, agreeableness, neuroticism) and correlated to a series of Facebook usage metrics, such as usage intensity, post length, time of engagement, number of posts per day, etc. These publications show that the higher the openness score is, the higher the number of likes and posts per day are. On the other hand, the higher the conscientiousness, the lower the number of likes and posts per day. Furthermore, extraversion is directly connected to the number of likes, posts per day and group associations. However, there are studies such as Tufekci (2013) which have introduced some conceptual issues which question the aforementioned.

Research has also focused on psychopathology issues and the way they are related to OSN usage. Depression seems to gain more attention with Pantic et al. (Pantic et al., 2012) studying it through the prism of the time spent using OSNs. Skues et al. (Skues et al., 2014) try to detect associations between depression and Facebook use. Jelenchick, Eickhoff and Moreno (Jelenchick, Eickhoff and Moreno, 2013) do not find evidence supporting a relationship between OSN use and clinical depression. Another attempt on a broader spectrum of disorders (schizoid, narcissistic, antisocial, compulsive, paranoid and histrionic personality disorders and major depression, dysthymia and bipolar-mania mood disorders) checks into whether the use of specific technologies or media, technology-related anxieties and technology-related attitudes would predict clinical symptoms of disorders (Rosen et al., 2013). Finally, researchers have examined if criteria regarding users' profiles could discriminate among individuals who were higher and lower in social anxiety (Fernandez, Levinson and Rodebaugh, 2012).

Another group of studies focuses on users' behaviour, their personality dependencies and OSN usage. (Alloway, Runac, Quershi and Kemp, 2014) examine the relationship among use of Facebook, empathy, and narcissism in adults. Narcissism has been studied by Bergman et al. (Bergman et al., 2011) who examine the link between narcissism and both OSN activities and motivation behind them. Carpenter (Carpenter, 2012) measures self-promoting/narcissistic Facebook actions and anti-social behaviour. Clayton et al. (Clayton et al., 2013) examine the relationships between loneliness, anxiousness, alcohol and marijuana use in the prediction of college freshman students' connections with others on Facebook, as well as their emotional connectedness to Facebook. College student performance and quality of life, in conjunction with Facebook usage, is studied by Kabre and Brown (Kabre and Brown, 2011).

2.2 Computer-science-focused studies

Regarding the research carried out by computer scientists, Kandias et al. (2010) propose a combination of technical and psychological approaches towards detecting malevolent insiders, while Greitzer et al. (2012) took into consideration the psychosocial perspective of an insider. Greitzer et al. (2012) developed a psychosocial model to assess employees' behaviour that could lead to insider abuse with an increased risk. According to that research, stress has been found to be a very useful indicator regarding insider threat manifestation. This result has been supported by other researchers, too (Shaw et. Al, 1998, FBI, 2012). As a results, stress level could be consider as an Indicator of Compromise in cyber fraud involving malevolent insiders. Brdiczka et al. (2012) propose an approach that combines Structural Anomaly Detection from social and information networks and psychological profiling of individuals to identify threats. Personal factors that may increase the likelihood of someone developing malevolent behaviour is proposed by FBI (2012) and Greitzer and Frincke (2010).

Opinion mining, sentiment analysis, and relational or flat data classification techniques (Witten and Frank, 2005) are computational techniques used in social computing (King et al., 2009). Social computing is a computing paradigm that involves multidisciplinary approach in analysing and modeling social behaviour on different media and platforms to produce intelligence and interactive platform results. One may collect and process the available data so as to draw conclusions about a user's mood (Choudhury and Counts, 2012). Choudhury and Counts explore ways in which expressions of human moods can be measured, inferred and made visible/exposed from social media activity. As a result, user and usage profiling and conclusion extraction from content processing are more feasible and valuable than ever. Researchers have examined the psychosocial traits described by Shaw and other researchers (Shaw et al., 1998; Greitzer and Fincke, 2010), indicating that such characteristics can be extracted via social media. To this extent, conclusions over traits, such as narcissism (Kandias et al., 2013a) or predisposition towards law enforcement (Kandias et al., 2013c; Kandias et al., 2013d) have been extracted via Twitter and YouTube, demonstrating the capability of online monitoring of users' malevolent behaviour.

Automated user profiling can lead to accurate prediction of personal information, such as ethnicity, religious or political views (Kosinski et al., 2013; Kandias et al., 2013b). Jakobsson et al., as well as Ratkiewicz (Jakobsson et al., 2008; Jakobsson and Ratkiewicz, 2006) have also carried out research on a realistic social media environments regarding online fraud experiments, such as phishing, with respect to users' privacy. Such approaches are proposed as a reliable way to estimate the success rate of an attack in the real world and as a means to raise user awareness on the potential threat. The researchers note the social threat that emerges from automated user and usage profiling (Mitrou et al., 2014).

Further information on the detection of usage pattern deviation in social media is feasible. Usage patterns tend to vary during different time periods or during particular life events. Such changes have also been examined by Facebook's Data Science Team. During 2014, the team showed alteration of timeline posts before and after the initialisation of a relationship (Facebook Data Science Team, 2014b). Similar behaviour was identified during the NFL season (Facebook Data Science Team, 2014a). The researchers examine the emotion expressed in anonymous status updates to provide an intuitive and useful measure of how fans feel about their teams. Also, the sentiment seemed to vary from negative to positive depending on their favorite team's results.

Based on the above, researchers have focused on psychosocial characteristics and their correlations to OSN behaviour and usage patterns, thus forming an interdisciplinary area of research that combines data science and psychology elements. To this extent, we focus on detecting potential correlations between OSN usage patterns and the stress level experienced by OSN users. Furthermore and to the best of our knowledge, there has been no publication on stress, personality or usage pattern deviations over time.

3 Background

According to Greitzer and Fricke (2010), an individual is stressed when she appears to be under physical, mental or emotional strain or tension that she finds difficult to handle. In order to draw conclusions about the user's stress levels, we used the Beck Anxiety Inventory (BAI) questionnaire (Beck et al., 1988), a widely used and acknowledged test.

Our research uses the Beck's stress questionnaire into a Facebook application, usable solely by people who have given their informed consent. By transferring Beck's questionnaire to an online form we are able to examine stress along with the user-generated content and the usage pattern exhibited in Facebook.

The issue of informed consent and whether it introduces a significant consent bias is a very interesting and diachronic issue in social studies which has been heavily researched upon since the 1960s. One of the most recent studies (Rothstein & Shoeben, 2013) conclude that the amount of consent bias is overstated, commonly known statistical methods can account for consent bias and *"any residual effects of consent bias are below an acceptable level of imprecision"*. As a result, the consent bias is deemed as a sensible social cost for researchers to conduct research within ethically and legally accepted limits. Regarding the applicability of informed consent in the context of big data and modern analytics Ioannidis (2013) contributes by highlighting the importance of informed consent and sheds light in other emerging issues.

3.1 Beck's anxiety inventory

Beck's anxiety inventory is a broadly used and acknowledged test that measures an individual's stress level. It consists of 21 questions that refer to common symptoms of stress and asks the responder to indicate how much she has been bothered by those symptoms during the past four weeks. The responder is asked to answer to these questions according to a Likert scale of 4 possible answers. At the end of the test the answers are aggregated and the respondent is classified into one of three categories, namely (a) low stress level, (b) medium stress level and (c) high stress level. Furthermore, the results of this test are utilised as ground truth in order to connect stress level with OSN usage patterns.

As almost every psychometric tool, Beck's Anxiety Inventory has been thoroughly researched upon and a lot of interesting results have been extracted. Since stress and anxiety are part of all DSM manuals, including the last DSM 5 (American Psychiatric Association, 2013), psychometric tests that attempt on quantifying stress and anxiety have a clear clinical and research targeting. BAI tests has been reported to have some correlation with depression and depict different results for different populations. Regarding the test's correlation with depression and BDI, researchers have claimed that this happens due to the extensive comorbidity of stress and depression (Maruish, 2004). Nonetheless, BAI test remains one of the most highly utilised measure both in research and clinical practice (Maruish, 2004). Furthermore, it is a self-report instrument which is brief and self-explanatory, which makes it ideal for self-assessments such as the one used in this research. Yet another significant advantage this test provides us with is its fundamental characteristic of depicting users' stress level on a time-stamp basis. This facilitates the chronicity analysis module (described in Section 6). In conclusion, this research inherits both strengths and limitations of the BAI stress test, the psychometric parameters are beyond the scope of this paper, however further analysis of its applicability and correlation to BDI based on user-generated content is part of our future work.

3.2 Facebook

Facebook is an online social networking service published in 2004. Its users are able to create online profiles, add friends to their networks, communicate by messaging each other and receive notifications about their friends' activity. It's one of the most popular OSNs and its users are willing to share a vast amount of their personal information and time too.

As Facebook does not permit direct crawling of its users' data, we developed a Facebook application in order to get access only to data the user allowed us to, the Facebook Application named "Stress Calculator" was active between September 2013 and September 2014. The data we have access to consists of (a) user information (friends list and profile description), (b) user-generated content (statuses, comments and links), (c) groups of interest (communities, events and activities) and (d) likings (music, actors, sports, books; content that the user has liked and is displayed in her profile). Provided the user's informed consent, we stored her results and the user-generated data she gave us access to. The developed application has an opt-out ability integrated which deletes all user data upon selection, thus implementing the right to be forgotten (Rosen, 2012). If the user has set privacy settings that do not allow access to parts of information we want to crawl, Facebook does not give us access to them. For example, if the user has not made her friend-list visible to other users, we don't have access either, thus her personal privacy settings prevail.

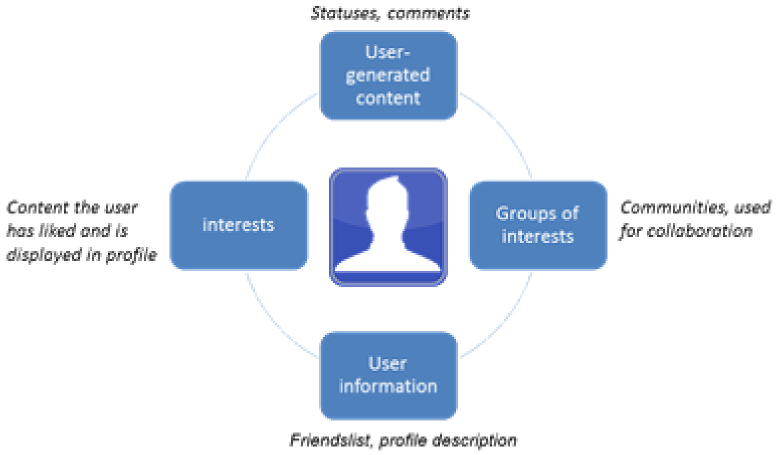

Fig. 1. Accessed user profile information

3.3 Data Crawling

In order to create our dataset, we used a Facebook application that would allow us to gather the required data after the completion of the questionnaire by the user and given her informed consent. The application uses Facebook's API to get access to users' data and to simplify the gathering procedure. The application development is based on Facebook's SDK that provides access to Facebook Graph API (https:// developers.facebook.com/docs/ graph-api) and allows the developer to manage the interaction between Facebook and the application.

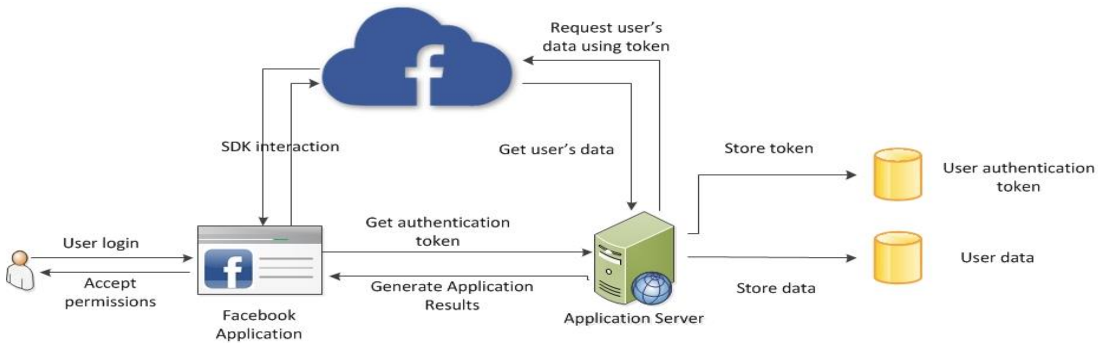

Fig. 2. Data crawling architecture

The process of collecting users' data consists of three phases. In the first phase, the user connects to her Facebook account and then installs the application for the first time. At this stage, the user accepts the permissions required and explicitly allows access to her data. In the second phase, an authentication token is generated by Facebook which allows our application to have access to users' data using Facebook API. This token is used every time our application performs a request to Facebook for specific information about a user. Access is allowed only to data that the user has chosen to give. If an API call requests data that are not permitted by the token, a notification error alert will occur and no data will be sent as a reply. The third phase involves the use of the authentication token to gather users' data from Facebook. For each user that has installed the application, completed the questionnaire and given their informed consent we perform API requests with the authentication token to access their data and store them in a database for later processing.

We have also added an anonymisation layer to the collected data. More specifically, usernames have been replaced with MD5 hashes so as to eliminate possible correlations between collected data and real life users. Each user is processed as a hash value, so it is hardly feasible for the results to be reversed. Consequently, single real life users cannot be detected. The collected dataset includes (a) 405 fully crawled users, (b) 12.346 user groups, (c) 98.256 liked objects, (d) 171.054 statuses and (e) 250.027 comments. The 51 users who did not offer informed consent, didn't fully fill the questionnaire or withdrew Facebook access rights for the application are excluded from the study. A brief description of user demographics including ages, gender and average statuses per user and per day are quoted in Table 1, further demographics analysis is beyond the scope of this paper.

Table 1. User Demographics.

Demographics		
Age	13-17	5%
	18-24	37%
	25-34	38%
	35-44	14%
	45-54	6%
Gender	Male	52%
	Female	48%
Average statuses per user	338 statuses	
Average statuses per user/day	0.7 post/day	

4 Results from Descriptive Statistics

Upon completing the data crawling process we performed data analysis of the dataset in order to evaluate its statistical validity. Data analysis was performed by using IBM SPSS Statistics (ver. 20). Furthermore, we conducted content analysis of the gathered user-generated content in order to detect users' major axes of interest (sports, music, politics, miscellaneous).

4.1 Statistical data analysis

In order to statistically anlyse the sample we conducted sample descriptive statistics with means comparisons among the data mining groups (53 users belonging to low stress category - Category0, 186 to medium stress category - Category1 and 166 to high stress category - Category2, making a sum of 405 users). Furthermore, we performed factor analysis with extraction method principle components and varimax rotation method. This served as an interdependency technique to find the latent factors that account for patterns of collinearity among the available metric variables.

Additionally, in order to calculate the determinate of the matrix of the sums of products and cross-products from which the intercorrelation matrix is derived, we utilised the Bartlett's Test of Sphericity. The null hypothesis is that the intercorrelation matrix is derived from a population in which the variables are non-collinear (namely an identity matrix) and that the non-zero correlations in the sample matrix are due to sampling error. Statistical decision for the Bartlett's Test of Sphericity from the calculated Chi-Square=3968.521 with $p=0.000<0.001$ was that the sample intercorrelation matrix did not come from a population in which the intercorrelation matrix is an identity matrix and that the non-zero correlations in the sample matrix are not due to sampling error.

The Kaiser-Mayer-Olkin (KMO) Measure of Sampling Adequacy (which indicates whether the sample size is adequate for performing factor analysis and varies from 0 to 1.0) was 0.840, higher than the recommended level of 0.6 (Hair et al., 2009). Variable-factor correlations may highlight the underlying structure. For this reason we have checked the Absolute Factor Loadings that allow for the quick interpretation of the correlation structure. According to our findings, it was revealed that Weighted OutDegree, StronglyConnected ID, Modularity Class and stress_score are the most important independent variables in our dataset. The discriminant analysis begins with the purpose of the statistical separation between two or more groups of cases. These "groups" are defined by the specific situation of the investigation (namely the available data mining clusters). The discriminant analysis was conducted by means of a stepwise selection process with the use of the Wilks' lambda criterion in order to extract Fisher's linear discriminant functions.

4.2 Content analysis

For a better understanding of the content of the collected data, we visualised it in order to identify the basic axes of content. Fig. 3 represents the categories of most common Facebook objects that are liked by our collected users.

Fig. 3. Tag cloud of top categories of liked objects.

As one may notice, the major interests of the users refer to music, entertainment, sports, politics, traveling, technology, etc. Furthermore, we examined the content of Facebook posts and comments in order to detect common interest axes with regard to categories of liked objects. The major axes that both categories of content seem to converge at are music (27% of posts), politics (14% of posts) and sports (17% of posts). Following, smaller subsets are detected to involve shopping (less than 1% of posts), leisure (2% of posts), cars (2% of posts), cooking and dinning (slightly more than 2% of posts), fashion (3% of posts) and the rest of posts were unclassified. Due to the small amount of the content in these categories, we decided to form a broader category (i.e. miscellaneous) that contains the former. For future work, we intend on forming more categories further analysing the dataset.

5 Exploratory unsupervised flat classification

In the initial phase of our research, we explored the collected dataset and we were not aware of any underlying characteristics. To this end, we decided to conduct unsupervised learning and enquire if stress characteristics could be depicted in the dataset. This module does not serve as input to the chronicity module even though it could be linked to the Big-5 trait of neuroticism; alternatively, we could have conducted supervised learning based on the ground truth of the stress test results, however these approaches are beyond the scope of this paper and will be addressed in our future work.

Regarding the questionnaire scores each participant achieved, we observe that 39% of the dataset users belong to the high stress category, 38% to the medium stress category and 23% to the low stress category as depicted in Fig. 4. In order to enable flat classification, we transformed the relational database that stored users' data (gathered by our Facebook application) into a single tuple record containing solely users' comments and statuses. The output of this process is a document containing the total amount of words and expressions used per user. In an effort to achieve better results and reduce the dispersion of different words, each user's flat data tuple was subjected to a stemming process. The produced document was used as input for the EM (Dempster et al., 1977) classification algorithm. Classification data were selected via feature selection based on tf-idf frequencies and term occurrence (greater than 6). This configuration has produced the best results. We avoided using bigrams, trigrams and parts of speech as features because they had been found to decrease the accuracy of the algorithm.

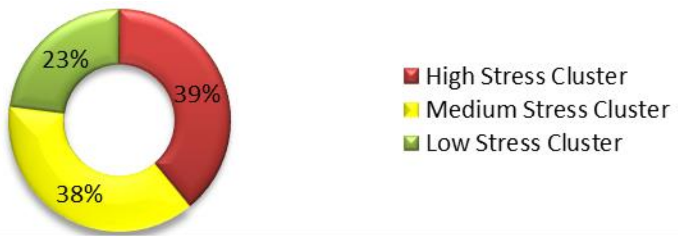

Fig. 4. Questionnaire results clusters.

The purpose of this approach is to perform unsupervised learning and let the machine form clusters of users based on their comments and statuses, along with content-related meta-data, such as number of comments or words in a comment. Thus, we were able to examine potential common user characteristics, the parameter being the result they obtained at the BAI test. Fig. 5 presents the clusters of users produced by the classification algorithm. The number of clusters was automatically generated by the process/machine according to the clustering conducted and is based on the combination of linguistic usage patterns along with content-related metadata.

5.1 Results analysis

As a follow-up to the aforementioned clustering, further examination of the population of each cluster produced by the unsupervised learning procedure was conducted. For each one of the detected clusters, we examine each user with regard to her BAI test score to draw conclusions about each cluster segment.

By comparing the flat classification results (Fig. 5.) we observed that Cluster_0 contains less than 10% of the dataset users. Along with the fact that it contains users from all the spectrum of the BAI test results, the referring cluster is of minor interest. On the contrary, Cluster_2 contains 48% of the dataset users and consists mainly of users who have scored low or medium score in the BAI test. Consequently, this cluster depicts the lower bound of stress score users. Indicatively, 89% of the cluster consists of users with low and the lowest bound of medium stress valuation. Regarding Cluster_1, it contains more than 42% of the dataset's users and its contents are almost complementary to Cluster_2. In both clusters 1 and 2 the percentage of users characterised by medium stress, according to the BAI test, is almost the same, contrarily to the users characterised by low and high stress levels. This is better explained by the fact that users with medium stress valuation belonging to Cluster_1 have scored very close to the upper bound of the category valuation, while those who belong to Cluster_2 have scored closer to the lower bound of the medium stress category. Furthermore, users of each cluster share similar characteristics, such as vocabulary and OSN communication patterns which further fed the chronicity analysis presented in the following section.

The results indicate that the flat data mining process formulated two working clusters able to classify users according to the stress level that their OSN communication patterns indicate. Cluster_1 includes those users who tend to have medium to high stress scores contrarily to Cluster_2 which includes mainly users who tend to have medium to low stress scores. Thus, these results imply that a correlation exists between users' OSN usage patterns and BAI stress scores.

Fig. 5. Flat classification clusters.

6 Chronicity

There have been several studies that classify users into predefined categories according to OSN user-generated data. However, there is a lack in current bibliography regarding the possible differentiations of OSN usage patterns over time.

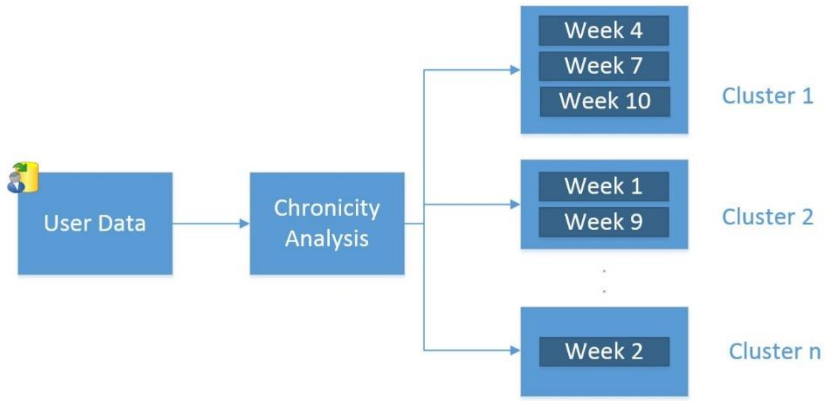

Fig. 6. Chronicity analysis example - Clusters of usage patterns.

Therefore, our research attempts to detect such usage pattern fluctuations of users' OSN user-generated content. To do so, we decided to split users' usage pattern into weekly time periods by using the four weeks prior to the BAI stress test response as a basis for the chronicity analysis module. This is because the BAI test measures stress symptoms as experienced by the respondent the past four weeks. We followed similar approaches (Eagle and Pentland, 2006) and experimented over time periods of 1 day, 2 days, 1 week, 2 weeks, 1 month. The time-span of one week has been found to be the most appropriate, as, according to our observations, users tend to manifest usage patterns based on certain events of their routines, which almost take place on a weekly basis. Additionally, for time periods greater than seven days, potential deviations are harder to detect as they are affected by the overall usage pattern. This is what we define as chronicity of users' OSN usage pattern and it's depicted in Fig. 6. More specifically, we group each user's weakly usage patterns into clusters of similar OSN behavior, so as to detect common usage patterns, periods of deviant OSN usage patterns and detect correlations between usage patterns and stress level experienced by the user (the four weeks prior to the completion of the BAI stress test serve as ground truth).

6.1 Chronicity Analysis

In order to tackle the challenge of studying OSN user-generated content through the prism of chronicity, we developed a system consisting of two major modules. Throughout our process, we were based on the fact that the BAI stress test corresponds to a bounded time period (four weeks prior to undertaking the test). This way we had a ground truth about the time period prior to the completion of the BAI test by the users. The first module (preprocessing data module) is responsible for the processing of input data, i.e. user comments and statuses, in order to be transformed in an appropriate form that maximizes the information gained. Additionally, such processing facilitates the latter content categorisation by using appropriate classifiers. Content categorisation is also conducted by this module. The second module (usage pattern analysis module) is responsible for receiving the preprocessed output from the first module and analysing usage patterns based on a set of metrics (further described in Section 6.3).

The preprocessing data module transforms the input data it receives by setting all letters to lower case, removing stop words and stemming all words. The major problem that had to be tackled in this module was the use of "Greeklish" in users' content. The term "Greeklish" refers to the writing of Greek words by using the Latin alphabet. To overcome this issue we transformed all Greeklish words to Greek ones by using GreeklishToGreek (www.innoetics.com/) web service, provided by Innoetics. The module processes only Greek words and ignores any other language, as it mainly focuses on a Greek community of users. Finally, the preprocessing module uses machine learning techniques to classify the content into the predefined categories of interest. The classification process is discussed in section 6.2.

The usage pattern analysis module processes a number of metrics which are calculated on the basis of the output of the previous module. It aims at representing usage patterns in a quantitative way and thus being able to detect possible deviations in it. The current module creates clusters of the user's OSN usage over time and searches for repetitive patterns of usage. Similar OSN usage patterns are categorised in similar clusters. Thus, clusters containing very few usage patterns indicate that these patterns are divergent ones.

Overall, the whole procedure involves both supervised and unsupervised classification. We apply supervised classification on user content (i.e. comments and statuses) in order to determine the category of interest each piece of information falls into. Following, we apply unsupervised classification on the set consisting of the weekly metrics extracted for each user, to identify usage patterns and potential deviations among these time periods.

6.2 Content Classification

We classify user content (i.e. comments and statuses) into the following categories of interest: (a) sports, (b) music, (c) politics, and (d) miscellaneous. We chose to create these categories based on the observations about the content of the collected dataset. The analysis of the axes of content led us to pick these major categories of interest to classify user content (as described in section 4.3).

User-generated content is categorised by using text classification (Sebastiani, 2002) techniques and machine learning. The first step of the process is to train a classifier to be able to classify user comments and statuses into one of the predefined categories of interest (sports, music, politics and miscellaneous). Text classification aims at training a system to decide the category in which a text falls into.

The machine is trained by having text examples as input as well as the category the examples belong to. Label assignment requires the assistance of an expert who can distinguish and justify the categories each text belongs to. We consulted a domain expert (i.e. Sociologist) who could assign and justify the chosen labels on the training sets. Thus, we created a reliable classification mechanism.

We performed comment classification by using: (a) Naïve Bayes Multinomial (Mc Callum and Nigam, 1998) (NBM), (b) Support Vector Machines (Joachims, 1998) (SVM) and (c) Multinomial Logistic Regression (Anderson and Blair, 1982) (MLR), so as to compare the results and pick the most efficient classifier. We compared each classifier's efficiency based on the metrics of precision, recall, f-measure and accuracy (Manning, 2008). Accuracy measures the number of correct classifications performed by the classifier. Precision measures the classifier's exactness. Higher and lower precision means less and more false positive classifications (the comment is incorrectly classified to a specific category) respectively. Recall measures the classifier's completeness. Higher and lower recall means less and more false negative classifications (the content is not assigned as related to a category, although it should be) respectively.

Precision and recall are increased at the expense of each other. That's the reason why they are combined to produce the f-score metric which is the weighted harmonic mean of both metrics.

We formed our training dataset by using the statuses and comments gathered by users. It comprises 275 sports, 301 music, 889 politics and 700 miscellaneous texts. Each text feature was subjected to stemming and stopwords removal. The classifier uses stemmed word features and neither n-grams were used nor parts of speech, as they decreased classifier's efficiency.

Table 2 presents each classifier's efficiency based on accuracy, precision, recall and f-score metrics, which are proper metrics to evaluate each classifier. The algorithms are compared based on 10-fold cross-validation (Witten and Eibe, 2005) in order to detect the most efficient one. Regarding the classes, 'S' stands for sports, 'M' for music, 'P' for politics and 'Mi' for miscellaneous.

The three algorithms achieve similar results regarding the chosen metrics. Naïve Bayes Multinomial and Multinomial Logistic Regression are characterised by less than 70% values for precision and recall, while Support Vector Machines is not. As a result, we decided to pick Support Vector Machines because of the better f-score value achieved for most categories and because all values for precision, recall and f-score are greater than 70%.

Table 2. Metrics comparison of classification algorithms.

	Metrics											
Classifier	**NBM**				**SVM**				**MLR**			
Classes	S	M	P	Mi	S	M	P	Mi	S	M	P	Mi
Precision	71	92	79	74	79	97	87	70	89	96	85	68
Recall	77	86	85	67	72	89	75	88	72	89	75	86
F-Score	74	89	81	70	75	93	81	78	79	93	80	76
Accuracy	79				81				80			

Finally, we created an additional classifier to detect aggressive and offensive content in user comments and statuses. To this end, we asked a domain expert to locate such content and populate an appropriate training set, including offensive jargon and vocabulary. The dataset comprises 300 aggressive and 320 non-aggressive texts. The classification scheme is applied to user content alongside with the above-mentioned categories of interest, as aggressive content can be expressed in each one of these categories. The classification algorithm used is Naïve Bayes and the metrics presented in Table 3 are evaluated based on 10 fold Cross Validation.

Table 3. Metrics of aggressive content classification.

	Metrics	
Classifier	**NB**	
Classes	Aggressive	Non-aggressive
Precision	81	83
Recall	82	81
F-Score	81.5	82
Accuracy	83	

6.3 Chronicity metrics

Usage chronicity is calculated via a set of ad hoc metrics, which are presented in Table 4. They focus on the following areas: (a) user interests, (b) usage patterns over time, (c) multimedia usage and (d) aggressive language. These areas cover an important range of usage patterns which they represent in a quantitative way so as to detect deviations in usage patterns over time. We focus on detecting fluctuations of usage patterns by examining user's overall OSN usage pattern, clustering similar time periods and spotting the diverging ones in which the usage pattern deviates significantly. Usage chronicity metrics are extracted by users' meta-data (namely posting and being-online time, content category etc.) rather than the text content itself. By categorising text content via the classification schemes developed, we are able to get meta-data regarding the content and decompose it to the appropriate metrics. Metrics regarding content classification are based on the analysis presented in section 4.3, which contains the major axes of content detected by analysing our dataset.

Table 4. Chronicity analysis metrics.

Frequency of posts regarding sports
Frequency of posts regarding music
Frequency of posts regarding politics
Frequency of posts regarding miscellaneous
Interest Shift per interest pair
Average frequency of posting
Average frequency of commenting
Major interests
Minor interest shift frequency
Frequency of aggressive comments
Frequency of uploading photos
CommentedBy ratio
StatusVarianceFlattened
CommentVarianceFlattened

The metrics depicted in Table 4 refer to a set of usage pattern characteristics and were extracted automatically by our chronicity analysis and usage pattern classifiers: (a) frequency of posts regarding sports/music/politics/ miscellaneous refers to the percentage of the user's total number of posts regarding the user's corresponding field of interest per time period; (b) interest shift per interest pair is the total number of changes between two different types of interest per time period; (c) average frequency of posting/commenting refers to the average number of user posts to her own wall or comments to other users per time period; (d) major interests are the type of interests that have a very high frequency of occurrence in a user's comments and posts. As a result, occurrences of posts referring to major interests are less likely to contribute to the appearance of a fluctuation in the usage pattern; (e) minor interest shift frequency is the type of interests that have a low frequency of occurrence in a user's comments and posts. As a result, occurrences of this type of posts are more likely to depict a usage pattern fluctuation since a user does not usually have an interest in this category of topics; (f) frequency of aggressive comments is the ratio between the number of comments and posts that contain offensive content and the total number of comments and posts per time period per user; (g) frequency of uploading photos is the ratio between the number of posts that contain links to photographs and the total sum of posts per period; (h) commentedBy ratio is the inverse ratio between the number of posts in a time period and the comments that refer to these posts by themselves or other users; (i) dispersal of user posts (StatusVarianceFlattened) is a criterion with which the volatility of the usage pattern is measured. The dispersal of publications is calculated by summing all of the posts in each subset and finding their dispersion. Large dispersal means that the user does not have a specific time usage pattern as the number of posts per time period varies. On the other hand, low dispersion means that the user's time usage pattern is constant and even small fluctuations are indicators of changes in the usage pattern; (j) dispersal of user comments (CommentVarianceFlattened) functions in a similar way with the above mentioned Status

VarianceFlattened metric but focuses on the commenting usage pattern, which may vary greatly from the posting usage pattern.

6.4 Decision process

To draw conclusions about a user's content, we follow the procedure depicted in Fig. 7. The aim is to detect time periods in which usage patterns fluctuate significantly. These time periods are indicators of deviating usage patterns within the OSN.

At the first stage, user's content is processed in order to remove the noise from the data. Then the content is processed by the classification schema (i.e. Support Vector Machines classifier) where the machine decides over the category that each piece of content falls into. Processed data along with the category of each classified instance serve as input to the chronicity analysis module.

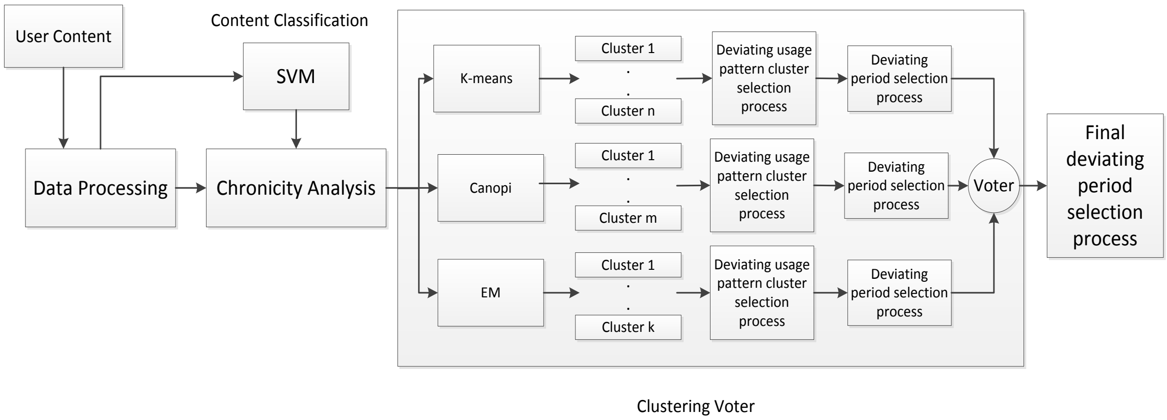

Fig. 7. Chronicity decision process.

The chronicity analysis module transforms the information into arithmetic vectors based on the metrics described in section 6.3 and performs clustering in order to detect similar time periods of specific usage patterns. Clustering is performed by using (a) K-means (Hartigan and Wong, 1979), (b) EM (Dempster et al., 1977), and (c) Canopy (McCallum et al., 2000). Each algorithm performs a classification process to the same input data and produces a number of clusters which contain time periods in which similar usage patterns are observed. The next step of the process includes selecting the clusters that are likely to contain similar usage patterns which are not often displayed by the user. The sensitivity of this selection process is set manually. As soon as the selection is completed, the time periods are temporarily saved as possible periods of deviant usage patterns. The final step of the process involves a voting procedure. Each detected time period is compared to the possible deviant periods that are created by the three classification algorithms. If at least 2 of the 3 clusters have noted a period as possibly deviant then this particular time period is classified as fluctuating from the major usage pattern. This process is repeated for all user periods in order to produce the final results.

The selection of combining three separate clustering approaches in an ensemble was made to provide a weighted output with regard to usage pattern fluctuations, as it has been used in other studies with similar requirements (Veeramachaneni and Arnaldo, 2016). Each algorithm was tested separately and responded differently regarding the sensitivity parameter. Thus EM and K- means algorithms were found to be more sensitive when very small changes occurred, while Canopy was found to detect significant changes in usage patterns. Consequently, having three algorithms to vote for the result makes it more accurate, since two or more classifiers should detect that the period under examination is a deviant one. The beneficial effect of this ensemble is that the extracted result becomes more accurate and two (or more) algorithms have detected a deviant period. In this way, potential false positives that could occur due to the use of a single clustering algorithm are avoided.

One may evaluate the aforementioned classification schema based on entropy, ground truth or observation by using a domain expert (Liu, 2007). Applying entropy would be difficult for the evaluation of the schema due to the size of the word vector. Confirmed deviant time periods are required to apply ground truth, which is not available in our case. Therefore, the most accurate approach to evaluate the classification schema is via observation of the results and confirmation of the deviant periods by the domain expert. The domain expert examined a vast amount of deviant time periods and confirmed the validity of the classification results.

6.5 Chronicity results

Based on the chronicity analysis performed on the collected content, we were able to categorise usage patterns into seven clusters, regarding the metrics vector that characterises each time period. Table 5 represents the mean values for each metric per week and the population percentage that belongs to each cluster. The number of clusters depicting usage patterns was automatically generated by the process/machine according to the clustering conducted. To form these clusters we used EM algorithm having as input all users' processed data and by using the decision schema described previously. The output was seven clusters of similar usage patterns.

Table 5. Chronicity analysis metrics.

Cluster id	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Population	7%	16%	8%	3%	4%	1%	7%	9%
TotalComments	3	78	93	5	50	79	410	44
TotalPosts	588	227	513	185	631	704	914	292
SportsFreq	0.00	0.01	0.00	0.00	0.01	0.00	0.02	0.03
MusicFreq	0.02	0.34	0.61	0.05	0.13	0.17	0.43	0.28
PoliticsFreq	0.00	0.06	0.02	0.00	0.02	0.04	0.05	0.15
MiscellaneousFreq	0.02	0.22	0.09	0.04	0.11	0.13	0.18	0.22
PhotosFreq	0.68	0.08	0.06	0.39	0.21	0.40	0.10	0.13
CommentsFreq	0.05	0.41	0.57	0.03	0.38	0.61	2.42	0.26
StatusesFreq	10.39	1.36	3.32	1.56	5.48	6.33	5.29	1.99
MinorInterestSift_Freq	0.01	0.18	0.13	0.02	0.09	0.12	0.16	0.24
CommentedBy ratio	0.08	1.25	0.73	0.27	0.45	0.65	1.52	0.88
StatusDispersalFlattened	34.89	3.01	4.93	7.85	14.39	15.20	7.98	6.28
CommentDispersalFlattened	0.01	1.52	0.95	0.06	1.02	1.08	5.95	0.60

During the classification process, the metrics of statuses and comments dispersal affected negatively the clustering due to their very high values with regard to the other metrics. In order to cope with this problem the status and comment dispersal of each user were divided by their sum of comments in order to achieve lower and more "flattened" values, but also to preserve the comparative analogies between the different users. As a result, the comparison metrics of Status dispersal and Comment dispersal were replaced with Status and Comment Dispersal Flattened.

The above-mentioned metrics analysis is summarised in Fig. 8. Clusters 0 and 3 contain the users who were classified in the high stress category according to the BAI test, while clusters 1 and 7 contain many users classified in category 2 of the BAI test. Regarding usage patterns in these clusters, users tend to upload more photos with higher frequency than the other users. In cluster 0, users post mainly photos and in cluster 3 users post photos, discuss music, whereas a small fraction of the content refers to miscellaneous information. Clusters 1 and 7 refer mainly to music and miscellaneous content and also include limited content referring to sports. Therefore, when the OSN usage pattern of a user falls into clusters 0 or 3, one could claim there is a higher possibility that the user suffers from higher stress levels for this period of time. Accordingly, if a user's usage pattern falls into cluster 1 or 7, there is a higher possibility that the user feels less stressed. Consequently, usage pattern deviations from one cluster to another could indicate a period of more or less stress according to the usage pattern depicted by the user.

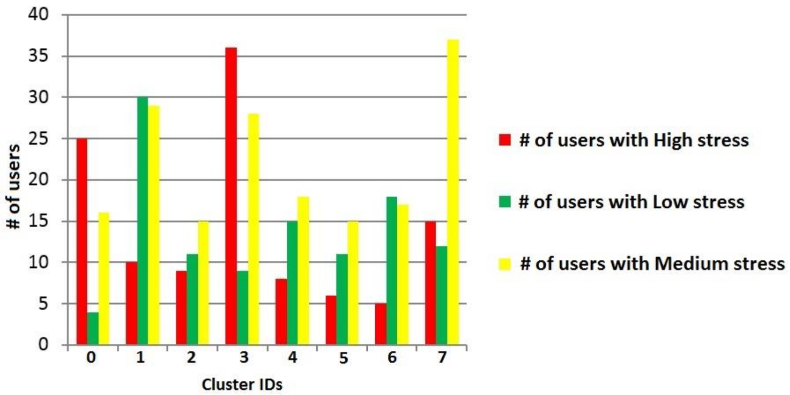

Fig. 8. Relation between clusters and stress.

7 Ethical issues

Methods that include mining of OSN user-generated content and user/usage profiling include inherent ethical controversies concerning a series of issues. These issues include human rights, the predictability of human behaviour, democratic boundaries and threats, such as mass social exclusion, prejudice and discrimination against certain minorities, etc. (Mitrou et al., 2014). Through the prism of the possible application of the above-mentioned method, other ethical issues and controversies rise as well.

The limits between the private and the public sphere or even between personal and professional life seem blurred before the advances of "boundary-crossing technologies" (Abril et al., 2012), the ongoing transformations of workplace demands and the radical penetration of OSNs as a means of communication. These changes are apparent in the working place and the recruiting procedures, where certain concerns have been raised about the reasons why job applicants have been rejected (Schermer, 2011). These reasons are attributed, amongst others, to lifestyle characteristics of the job seekers extracted from online communities and OSNs (Broughton et al., 2009).

Methods like those proposed in this paper pose or even introduce a series of risks, namely workplace discriminations, the phenomenon of the "well-adjusted employee" (Simitis, 1999), chilling effects on employees' personality and freedom of speech or even disturbance of the relationship of mutual trust and confidence between the employer and the employee. Yet another threat that emerges is the fact that users' personal life is processed and mined via a process that may lead to user/usage classification outside of the initial context. This decontextualization of the OSN user-generated content pertains to the over-simplification of social relations and the wide dissemination of information (Dumortier, 2010). Additionally, in previous research of ours (Kandias et al., 2013a; Mitrou et al. 2014) we referred to another issue that emerges, namely the possibility of collecting inaccurate or mistimed information which could reflect a different life-phase of the person. This point of view differentiates vastly if the results of this paper are taken into consideration. The ability to study usage pattern chronicity and extract results from that, highlights the ability to use big data in order to study each user through the prism of her fluctuations over time.

Apart from the private sector and the corporate milieu, governments have also shown interest in the development and use of such technologies. The case of US Government's PRISM program that involves the US NSA collecting and analysing foreign communications collected from a range of sources, including OSNs (Cumbley and Church, 2013), has even lead to the cancellation of the US-EU Safe Harbor. The interest of governments in OSN user-generated content relies on the ability OSINT offers to discover or infer previously unknown facts, user-/usage patterns and correlations (Rubinstein, 2013). Methodologies resembling those presented in this paper raise significantly the threat posed on society by such types of surveillance. The ability to extract results about the behaviour and the unique characteristics of individuals along with the possibility to classify users into predefined categories according to specific metrics infringes a series of fundamental rights. Additionally, correlating every-day activities, psychosocial characteristics (such as stress levels, predispositions or political beliefs) and usage patterns introduces alternative social stratifications that could lead to stigmatisation or exclusion of individuals belonging to certain categories or social groups.

8 Conclusions

In this paper, we adopted an interdisciplinary approach to detect the stress levels depicted by the OSN usage patterns. We performed our research on a Facebook dataset collected by users who offered their informed consent, and performed statistical and content analysis on it. In particular, we described a method of unsupervised flat data classification of the overall user-generated content which was used to explore our dataset. We aggregated the sum of users' comments and statuses into a single document, preprocessed the included data using a Greeklish to Greek converter and a stemmer and classified them by using the EM classification algorithm. Apparently, this approach is able to classify users into two major clusters based on users' overall OSN usage pattern; the first cluster includes users with low and medium-to-low stress according to the user driven stress test, while the second cluster includes mainly users with high and medium-to-high stress.

Our major contribution relies on the development of a method of usage pattern analysis through the prism of fluctuations over time. To this end, we followed a process which begins with the classification of the user generated content into four predefined categories, namely sports, music, politics and miscellaneous. We achieved that via text classification on the content by using the SVM algorithm. Afterwards, we defined a series of chronicity metrics, which were determined either manually from our observations on the dataset or in an automated manner by the flat data classification and the correlations detected by the chronicity analysis method per se. These metrics refer to several usage pattern characteristics and are utilised in order to facilitate the process of fluctuation detection of the usage patterns that are being analysed. Then the results of these metrics were transformed into arithmetic vectors and clustered by K-means, EM and Canopy algorithms in order to detect similar time periods of specific usage patterns. The above-mentioned process was repeated for several time spans varying from one day to one month in order to empirically detect the most useful one. In this process we used a fundamental characteristic of the BAI stress test, namely its ability to depict stress level regarding a specific period of time. According to our results, the time span of one week fits better in order to extract useful results of usage chronicity through the prism of deviation from the usual usage pattern.

Taking into consideration the produced results, we argue that several ethical issues arise along with a conflict of interests between parties over the usability and applicability of the proposed methods. Broad application of such methods by employees or governments could lead to several civil rights violations affecting both public and private life of the employees/citizens. Under this point of view, it is clear that such methods could be implemented solely under strict terms and conditions and provided the explicit informed consent of the user. Such methods could be used in the context of critical infrastructures, where security requirements are accountable for nationwide or even global well-being and human lives. Thus, the recruitment and monitoring process of high profile individuals in key positions could be updated with psychosocial Indicators of Compromise. Similar methods based on psychometric evaluations have been successfully used by the US Air Force for the pilots' recruitment (Boyd, Patterson and Thompson, 2005). Furthermore, such approaches have been used by attackers during advanced spear-phishing attacks, thus, the proposed method could be used as a Threat Intelligence module producing equivalent Indicators of Compromise or as part of a forensics analysis in order to thoroughly analyse past incidents.

For future work, we plan on further studying the flat data classification results and its relation to the Big-5 trait of neuroticism, conduct supervised learning based on the ground truth of the stress test results and extend the content categories in order to be able to detect even more delicate behaviour fluctuations. Furthermore, we plan on extending and meta-training our metrics set in order to re-evaluate and strengthen metrics validity and applicability and compare the results between users of different ethnicities and heterogeneous societal and cultural characteristics. In terms of our method's applicability we intend on making it less intrusive by providing solely abstract Indicators of Compromise and threat intelligence both for predicting future cyber-attacks and forensically analysing past incidents.

Acknowledgements. The authors would like to thank Innoetics (www.innoetics.com) for providing the Greeklish-to-Greek transformation software, as well as Themis Markopoulos for developing the Facebook application.

References

Patricia Sánchez Abril, Avner Levin and Alissa Del Riego. 2012. Blurred Boundaries: Social Media Privacy and the Twenty First Century Employee. American Business Law Journal, 49, 1, 63-124.

Tracy Alloway, Rachel Runac, Mueez Quershi and George Kemp. 2014. Is Facebook Linked to Selfishness? Investigating the Relationships among Social Media Use, Empathy, and Narcissism. Social Networking, 2014.

American Psychiatric Association. 2013. DSM 5. American Psychiatric Association.

Yair Amichai-Hamburger and Gideon Vinitzky. 2010. Social network use and personality. Computers in Human Behavior, 26, 6, 1289-1295.

JA Anderson and Val Blair. (1982). Penalized maximum likelihood estimation in logistic regression and discrimination. Biometrika, 69, 1, 123-136.

Yoram Bachrach, Michal Kosinski, Thore Graepel, Pushmeet Kohli and David Stillwell. 2012. Personality and patterns of Facebook usage. In proc of the 3rd annual ACM web science conference, 24-32.

Mitja Back, Juliane Stopfer, Simine Vazire, Sam Gaddis, Stefan Schmukle, Boris Egloff, and Samuel Gosling. 2010. Facebook profiles reflect actual personality, not self-idealization. Psychological Science.

Aaron Beck, Norman Epstein, Gary Brown and Robert Steer. 1988. An inventory for measuring clinical anxiety: psychometric properties. Journal of consulting and clinical psychology, 56, 6, 893.

Shawn Bergman, Matthew Fearrington, Shaun Davenport, and Jacqueline Bergman. 2011. Millennials, narcissism, and social networking: What narcissists do on social networking sites and why. Personality and Individual Differences, 50, 5, 706-711.

Clive Best. 2008. Open source intelligence. Mining massive data sets for security: advances in data mining, search, social networks and text mining and their applications for security. IOS Press, 331-344.

James E. Boyd, John C. Patterson, and Bill T. Thompson. 2005. Psychological test profiles of USAF pilots before training vs. type aircraft flown. Aviation, space, and environmental medicine, 76, 5, 463-468.

Oliver Brdiczka, Juan Liu, Bob Price, Jianqiang Shen, Abhijit Patil, Richard Chow, Evgeniy Bart, Nicholas Ducheneaut. 2012. Proactive insider threat detection through graph learning and psychological context. 2012 IEEE Symposium on. IEEE Security and Privacy Workshops, 142-149.

APA Andrea Broughton, Tom Higgins, Ben Hicks and Annette Cox. 2009. Workplaces and Social Networking - The Implications for Employment Relations. Institute for Employment Studies,UK.

Christopher Carpenter. 2012. Narcissism on Facebook: Self-promotional and anti-social behavior. Personality and individual differences, 52, 4, 482-486.

Munmun De Choudhury and Scott Counts. 2012. The nature of emotional expression in social media: measurement, inference and utility, In Proceedings of the Human Computer Interaction Consortium Workshop.

Russell Clayton, Randal Osborne, Brian Miller and Crystal Oberle. 2013. Loneliness, anxiousness, and substance use as predictors of Facebook use. Computers in Human Behavior, 29, 3, 687-693.

Teresa Correa, Amber Willard Hinsley and Homero Gil De Zuniga. 2010. Who interacts on the Web?: The intersection of users' personality and social media use. Computers in Human Behavior, 26, 2, 247-253.

Richard Cumbley and Peter Church. 2013. Is "Big Data" creepy?. Computer Law & Security Review, 29, 5, 601-609.

Arthur Dempster, Nan Laird and Donald Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B (Methodological), 1-38.

Franck Dumortier. 2010. Facebook and Risks of “De-contextualization” of Information. Data Protection in a Profiled World, Springer, 119-137.
Nathan Eagle and Alex Pentland. 2006. Reality mining: sensing complex social systems. Personal and ubiquitous computing. 10, 4, 255-268.
Facebook Data Science Team. 2014a. The Emotional Highs and Lows of the NFL Season. Facebook.
Facebook Data Science Team. 2014b The Formation of Love. Facebook.
FBI. 2012. The Insider Threat: An introduction to detecting and deterring an insider spy. www.fbi.gov/about-us/ investigate/counterintelligence/ the-insider-threat.
Katya Fernandez, Cheri Levinson and Thomas Rodebaugh. 2012. Profiling Predicting Social Anxiety From Facebook Profiles. Social Psychological and Personality Science, 3, 6, 706-713.
Frank Greitzer, Lars Kangas, Christine Noonan, Angela Dalton and Ryan Hohimer. 2012. Identifying at-risk employees: Modeling psychosocial precursors of potential insider threats. In proceedings of the 45th Hawaii International Conference on System Science, IEEE, 2392-2401.
Joseph Hair, William Black, Barry Babin and Rolph Anderson. 2009. Multivariate data analysis (Vol. 7). Prentice Hall.
John Hartigan and Manchek Wong. 1979. Algorithm AS 136: A k-means clustering algorithm. Applied Statistics, 100-108.
Natalie Hayes and Stephen Joseph. 2003. Big 5 correlates of three measures of subjective well-being. Personality and Individual Differences, 34, 4, 723-727.
Thorsten Joachims. (1998). Text categorization with support vector machines: Learning with many relevant features. Springer.
David John Hughes, Moss Rowe, Mark Batey and Andrew Lee. 2012. A tale of two sites: Twitter vs. Facebook and the personality predictors of social media usage. Computers in Human Behavior, 28, 2, 561-569.
John PA Ioannidis. 2013. Informed consent, big data, and the oxymoron of research that is not research. The American Journal of Bioethics, 13, 4, 40-42.
Markus Jakobsson and Jacob Ratkiewicz. 2006. Designing ethical phishing experiments: a study of (ROT13) rOnl query features. In Proc. of the 15th International Conference on World Wide Web, ACM, 513-522.
Markus Jakobsson, Peter Finn and Nathaniel Johnson. 2008. Why and how to perform fraud experiments. IEEE Security & Privacy, 6, 2, 66-68.
Lauren Jelenchick, Jens Eickhoff and Megan Moreno. 2013. “Facebook depression?” Social networking site use and depression in older adolescents. Journal of Adolescent Health, 52, 1, 128-130.
Faycal Kabre and Ulysses Brown. 2011. The influence of Facebook usage on the academic performance and the quality of life of college students. Journal of Media & Communication Studies, 3, 4, 144-150.
Miltiadis Kandias, Alexios Mylonas, Nikos Virvilis, Marianthi Theoharidou and Dimitrios Gritzalis. 2010. An insider threat prediction model. In Proc. of the 3rd International Conference on Trust, Privacy and Security in Digital Business, Springer, 26-37.
Miltiadis Kandias, Konstantina Galbogini, Lilian Mitrou and Dimitrios Gritzalis. 2013a. Insiders trapped in the mirror reveal themselves in social media. In Proc. of the 7th International Conference on Network and System Security, Springer, 220-235.
Miltiadis Kandias, Lilian Mitrou, Vasilis Stavrou and Dimitrios Gritzalis. 2013b. Which side are you on? A new Panopticon vs. privacy. In Proc. of the 10th International Conference on Security and Cryptography, 98-110.
Miltiadis Kandias, Vasilis Stavrou, Nick Bozovic, Lilian Mitrou and Dimitrios Gritzalis. 2013c. Can we trust this user? Predicting insider's attitude via YouTube usage profiling. In Proc. of the 10th International Conference on Autonomic and Trusted Computing, IEEE, 347-354.
Miltiadis Kandias, Vasilis Stavrou, Nick Bosovic and Dimitrios Gritzalis. 2013d. Proactive insider threat detection through social media: The YouTube case. In Proc. of the 12th ACM Workshop on Privacy in the Electronic Society, ACM, 261-266.
Irwin King, Jiexing Li and Kam Tong Chan. 2009. A brief survey of computational approaches in social computing. In Proc. of the International Joint Conference on Neural Networks, IEEE, 1625-1632.
Michal Kosinski, David Stillwell and Thore Graepel. 2013. Private traits and attributes are predictable from digital records of human behavior. National Academy of Sciences. 110, 15, 5802-5805.
Alexander Liu, Cheryl Martin, Tim Hetherington and Sara Matzner. 2005. A comparison of system call feature representations for insider threat detection. In Proc. of the 6th Annual IEEE SMC Information Assurance Workshop, IEEE, 340–347.
Bing Liu. 2007. Web data mining: exploring hyperlinks, contents, and usage data. Springer.
Peter Malcolm Manning. (2008). Techniques to enhance the accuracy and efficiency of finite-difference modelling for the propagation of elastic waves.

Mark E. Maruish. 2004. The Use of Psychological Testing for Treatment Planning and Outcomes Assessment: Volume 1: General Considerations. Routledge, Mahwah, New Jersey.

Andrew McCallum, Kamal Nigam and Lyle Ungar. 2000. Efficient clustering of high-dimensional data sets with application to reference matching. In Proc. of the 6th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 169-178.

Lilian Mitrou, Miltiadis Kandias, Vasilis Stavrou and Dimitris Gritzalis. 2014. Social media profiling: A Panopticon or Omniopticon tool?. 6th Conference of the Surveillance Studies Network, Spain.

Tim O'reilly, T. 2009. What is web 2.0. O'Reilly Media.

Igor Pantic, Aleksandar Damjanovic, Jovana Todorovic, Dubravka Topalovic, Dragana Bojovic-Jovic, Sinisa Ristic and Sanka Pantic. 2012. Association between online social networking and depression in high school students: behavioral physiology viewpoint. Psychiatria Danubina, 24, 1, 90-93.

Daniele Quercia, Renaud Lambiotte, Daviv Stillwell, Michal Kosinski and Jon Crowcroft. 2012. The personality of popular Facebook users. In Proc. of the ACM 2012 Conference on computer supported cooperative work, 955-964.

Marc Rogers, Natalie D. Smoak and Jia Liu. 2006. Self-reported deviant computer behavior: A big-5, moral choice, and manipulative exploitive behavior analysis. Deviant Behavior, 27, 3, 245-268.

Jeffrey Rosen. 2012. The right to be forgotten. Stanford law review online, 64, 88.

Larry Rosen, Kate Whaling, Sam Rab, Mark Carrier and Nancy Cheever. 2013. Is Facebook creating "iDisorders"? The link between clinical symptoms of psychiatric disorders and technology use, attitudes and anxiety. Computers in Human Behavior, 29, 3, 1243-1254.

Jenny Rosenberg and Nichole Egbert. 2011. Online impression management: personality traits and concerns for secondary goals as predictors of self-presentation tactics on Facebook. Journal of Computer-Mediated Communication, 17, 1, 1-18.

Craid Ross, Emily Orr, Mia Sisic, Jaime Arseneault, Mary Simmering, and Robert Orr. 2009. Personality and motivations associated with Facebook use. Computers in Human Behavior, 25, 2, 578-586.

Mark A. Rothstein, and Abigail B. Shoben. 2013. Does consent bias research? The American Journal of Bioethics, 13, 4, 27-37.

Ira Rubinstein. 2013. Big Data: The End of Privacy or a New Beginning?. International Data Privacy Law, 3, 74.

Tracii Ryan and Sophia Xenos. 2011. Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness and Facebook usage. Computers in Human Behavior, 27, 5, 1658-1664.

Louise La Sala, Jason Skues and Sharon Grant. 2014. Personality Traits and Facebook Use: The Combined/Interactive Effect of Extraversion, Neuroticism and Conscientiousness. Social Networking, 3, 5, 211.

Bart Schermer. 2011. The limits of privacy in automated profiling and data mining. Computer Law & Security Review, 27, 1, 45-52.

Eugene Schultz. 2002. A framework for understanding and predicting insider attacks. Computers & Security, 21, 6, 526–531.

Fabrizio Sebastiani. 2002. Machine learning in automated text categorization. ACM Computing Surveys (CSUR), 34, 1, 1-47.

Eric Shaw, Keven Ruby and Jerrold Post. 1998. The insider threat to information systems: The psychology of the dangerous insider. Security Awareness Bulletin, 2, 98, 1-10.

Spiros Simitis. 1999. Reconsidering the premises of labour law: Prolegomena to an EU regulation on the protection of employees' personal data. European Law Journal, 5, 1, 45-62.

Jason Skues, Robert Banagan and Lisa Wise. 2014. Facebook and Diagnosis of Depression: A Mixed Methods Study. Social Networking.

Zeynep Tufekci. 2013. Big Data: Pitfalls, methods and concepts for an emergent field, SSRN 2229952.

Kalyan Veeramachaneni and Ignacio Arnaldo. 2016. AI2: Training a big data machine to defend. IEEE International Conference on Big Data Security on Cloud (BigDataSecurity 2016). USA.

Ian Witten and Frank Eibe. 2005. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann.