Online Social Networks as Sensors in Smart Environments

Ricardo Barbosa
CIICESI, ESTGF.IPP School of Technology and Management of Felgueiras – Polytechnic of Porto
Felgueiras, Portugal
8090242@estgf.ipp.pt

Ricardo Santos
CIICESI, ESTGF.IPP School of Technology and Management of Felgueiras – Polytechnic of Porto
Felgueiras, Portugal
rjs@estgf.ipp.pt

***Abstract*—Smart environments need to successfully identify their user needs, personality and behaviour in order to prepare appropriate interactions when facing different types of events. Personality is strongly associated with human behaviour and can be predicted with relatively accuracy by analysing online social profiles. Due to a constant presence in the lives of their users, online social networks have a decidedly strong social impact leading to a blur between offline and virtual life as well as the concept of digital identity and provide a rich source of data that can be used to profile and understand personal preferences and interactions of an individual. Users reveal a lot about themselves and share their personal preferences in a freely way and this paper exploits a new approach for smart environments to correctly identify their user's personality, using online social network as a sensor.**

Keywords*—*Online social networks, Personality profiling, Smart environments, Big Five model

I. Introduction

The immerse world of data that we live on today is full of possibilities and realisations of scenarios that were only possible on sci-fy movies or dreams. Ambient Intelligence and Smart Environments are impulsed by ubiquitous computing and take advantage of the ease of collecting data from numerous devices in order to produce tasks such as optimisation of energy consumption [1], [2], recognition of human activity and preferences [3], [4], aid the elderly or persons with health problems [5], or even increase the lifestyle of blind people [6]. Although this type of implementations require physical sensors to capture data, there is another way to gather data without the need of a physical sensor.

With online social networks being an important aspect of our society, many people shift themselves into the virtual world making their virtual profile a mirror of their true identity. It is unrealistic to say that this is a passenger trend when the numbers speak for themselves, the numbers of September of 2016 gave the market lead to Facebook with an astounding number of 1712 million active users, being the first social network to surpass the mark of one billion registered accounts, followed by WhatsApp and Facebook Messenger with 1000 million active users each, QQ (Chinese social media network) with 899 million and a little bellow we can find social networks like Instagram and Twitter with 500 and 313 million active users respectively [7].

User modelling is essential in the context of adaptive systems, the ability to profile someone can contribute to know the potential needs in different contexts [8] and it is being beneficial for many activities on a daily basis such as customer support, recommendation of services and products and job applications [9]. Similarly, smart environments eventually will involve a crucial step of implicit (mental) or explicit (trough a user profile analysis) modelling of the user personality [10].

In order to successfully profile an user trough his online social network activity, it is vital to understand the correlation between his personality traits and behaviour. Due to the strong impact of personality in human behaviour [11], personality is reflected and has impact on a users online profile or activity [12]. Personality is a way for humans to describe themselves and others, and for decades, psychology researchers have worked to understand personality in a systematic way. After extensive work to develop and validate a widely accepted personality model, researchers have shown connections between general personality traits and many types of behaviour [11].

This paper is structured as follows. Section two presents a brief definition of personality and its direct relation and influence in human behaviour and explains the model used for classifying the traits of personality, known as the Big Five model [13]. Section three correlates the behaviour of an individual with his online social network profile and identifies the motivations for the usage of this type of social media as long as an understanding of personality with the scope of online social behaviour, and presents an overview for personality profiling through text and some characteristics that can be found on text and that can be used to predict the personality of an individual including some limitations related to the task of personality profiling through online social networks. Section four contains the overview of our idea for using online social networks as sensors in smart environments and examples of its application. The fifth and last section is related to our vision of the concept and future work.

II. Behaviour and Personality

According to Poria et al. [9] since the times of Aristotle that is discussed the existence of various personality types and its connection with the different patterns of human behaviour.

The definition of personality is still ambiguous and psychologists have their own way to describe it, but most consider personality as a dynamic organisation, inside the person, of

psychophysical systems that create the characteristics patterns of behaviour, thoughts and feelings of an individual [14]. D. Markovikj et al. [15] consider personality as a key component to identify a profile, and a uniquely identifier for each one of us which affects a lot of aspects of human behaviour, mental processes and affective reactions. All of the existing personalised systems used in domains such as e-learning, information filtering, collaboration or even e-commerce could greatly benefit from a user interface that adapts the interaction (like motivational strategies, presentation styles or recommendations) according to user's personality.

In their work, J. W. Pennebaker and L. A. Ki [16] explain that we all know that people differ in the way they act, talk and write. A simple posture can tell a lot about some individual and even when the content of a message is the same, individuals express themselves verbally with their own distinctive styles. The same characteristics are also valid for written language, which is also unique from person to person. S. Adali and J. Golbeck [12] referred a really important fact, a person behaviour is not simply a function of their personality traits, as an example, an aggressive person will behave aggressively in certain situations. The situational cues lead to activation of personality traits which then lead to a behavioural expression. The authors also define personality as an important trait that moderates peoples behaviour and interactions with one another. In essence, personal tendencies are shaped further through social interactions where individuals in a social network act similarly, sometimes referred to as normative (or normal) behaviour. Furthermore, research has shown that distinctive characteristics of ones personality are more likely to manifest themselves in situations that satisfy individuals basic psychological needs.

A. Personality Classification

The ability to predict personality has implications in many areas, existing research has shown connections between personality traits and success in both professional and personal relationships [11].

Although it is possible to predict Internet usage by cognitive style, personality was chosen over those individual differences and psychologists have reached a consensus that the domain of personality can best be described by the Big Five dimensions [17]. Several well studied personality models have been proposed, however the Big Five model, introduced by Norman in 1963 [18] and matured by Goldberg [13], was established as the most popular one and is currently the most widespread and generally accepted model of personality [11], [15], [19]. The five dimensions can be described as the following:

- Openness to Experience: curious, intelligent, imaginative. High scorers tend to be artistic and sophisticated in taste and appreciate diverse views, ideas, and experiences (insightful vs unimaginative).
- Conscientiousness: responsible, organised, persevering. Conscientious individuals are extremely reliable and tend to be high achievers, hard workers, and planners (organised vs careless).
- Extroversion: outgoing, assertive. Friendly and energetic, extroverts draw inspiration from social situations (sociable vs shy).
- Agreeableness: cooperative, helpful, affection. People who score high in agreeableness are peace-keepers who are generally optimistic and trusting of others (friendly vs uncooperative).
- Neuroticism: anxious, insecure, sensitive. Neurotics are moody, tense, and easily tipped into experiencing negative emotions (calm vs insecure).

The work of U. Gupta and N. Chatterjee [20] demonstrates the importance of Big Five traits in identification of human behaviour related traits, trough psychological experiments, such as deception, job performance, individual preferences, and other aspects.

III. Online Social Behaviour

With the social media impact on the daily life of a large part of people in the world, social media is a place where users present themselves to the word where new clues and aspects about an individual personality emerge from little aspects that most individuals do not notice at all. Due to a constant presence in the lives of their users, online social networks have a decidedly strong social impact leading to a blur between offline and virtual life as well as the concept of digital identity [7]. The motivations for the usage of an online social network differ from user to user, some focus on broadcasting information about themselves while others are more interested in passively consuming information produced by others [21]. This usage is primarily motivated by two basic social needs, the need to belong (necessity to affiliate with others and gain social acceptance) and the need for self-presentation (the continuous process of impression management) [22].

Moosavi [23] classifies online social networks as a good index to predict potential actions of users, however a lot of rich information is encoded in the content of those interactions. Since, generally, a social network is shown in a graph and defined as a network of interactions and relationships, where the nodes consist of actors and the edges consist of the relationships or interactions between these actors, is worth analysing the interactions between people and determining structural patterns present on them. According to S. Adali and J. Golbeck [12], people reveal their personality traits through their use of online social networks, who can be predicted with a relatively high accuracy by analysing public data that people liberally share online. We can say that online social networks are a mirror where users reveal a lot about themselves both in the way they share and how they share-it.

In this large amount of everyday constant data there are many options and behaviours to analyse in order to profile the personality of an user. According to S. Adali and J. Golbeck [12], people reveal their personality traits through their use of social media in the case of Facebook and on Twitter, where personality traits can be predicted with a relatively high accuracy by analysing public data that people share online.

In order to understand personality with the scope of online social behaviour, in their work S. Adali and J. Golbeck [12] analysed various behaviours of individuals in their social

group. Some actions were considerate in the following main groups:

- Network Bandwidth (NET): the amount of overall activity and size of social network, the distribution of activity over time and how long they have been using the online social network.
- Message Content (MSG): the type of messages sent, whether they contain URLs (or other types of links) and whether they are forwarded.
- Pair Behaviour (PAIR): their behaviour towards their friends and followers.
- Reciprocity of actions (REC): to which degree their actions are reciprocated by their friends.
- Informativeness (INF): How informative are various behaviour features across all the friends.
- Homophily (HOM): All the previous features computed for the persons friends to understand her social circle.

A study made by K. Moore and J. C. McElroy [17] developed some interesting results. They found a significant positive relationship between gender and a number of variables of interest where was possible to find that women spend more time on Facebook, had a greater number of friends, posted more photographs and did more postings about themselves, when compared to men. Although in terms of frequency, women visit their Facebook less frequently than men do. In terms of personality traits, this study found that more extroverted people have more Facebook friends and report less regret over Facebook content, however, extroversion was not significantly related to time spent, number of photographs or the number of wall postings (either about themselves or others). The high-scorers in agreeableness expressed a greater levels of regret about inappropriate content they may have posted and, surprisingly, they did a greater number of postings about themselves than did the low-scorers. Conscientiousness was not related to time spent, frequency of use, number of friends or number of photographs, and people with high-score in conscientiousness made significantly fewer wall postings, and expressed more regret than did people with low-score. Emotional stability was not significantly related to actual number of friends or photographs, or to the number of wall posting, it was positively related to both how frequently they use Facebook to keep up with others and regret. For last, openness proved to have no significant effect on either Facebook usage or content.

A. Personality Through Text

People may try to pretend being someone that they're not, however, the personality traits are so strong that they left some piece of themselves on every sentence they use. Some works approach this situation by content-based and style-based features [10] or by function words that may seem worthless but can actually tell a lot about someone [24]. Similarly to the human behaviour, the way we write can also be directly related with our personality. R. Pitcher [25] came to the conclusion that when analysing a text, either quantitatively or qualitatively, a lot of valuable and useful data is thrown away by ignoring words and phrases that are used figuratively, such as metaphors, exaggerations and pictorials.

A study by J. Oberlander and A. J. Gil [26], similar to the work of J. Shen et al. [27], focus on email communications, and by analysing the corpus of email messages they classify each user using the Big Five model. They expect that extrovert individuals use more positive emotional language (warmth, assertiveness, positive emotions), use more social language (gregariousness) and produce more complex or extended utterances, reflecting their tendency to dominate interactions (assertiveness, excitement-seeking). In the other hand, they expect that neurotic individuals use more negative emotional language (anxiety, angry hostility, depression), use more self-oriented language (self-consciousness, vulnerability) and produce more emphatic utterances (angry hostility, impulsiveness).

Some characteristics mentioned by D.N.Chin and W.R. Wright [28] related to social media characteristics that are likely to affect personality include:

- Word length of entries
- Number of entries/author
- Author identification
- Spelling and grammar errors
- Topic bias
- Time-period bias
- Author self-selection bias
- Legal access and privacy restrictions
- Unusual syntax, usage, abbreviations

There is also the fact that different populations might tend to write about different topics as well as to express themselves differently about the same topic. One simple communication task like e-mailing a friend about recent activities, is likely to be accomplished differently by two people. Some differences depend on their recent experiences, age, geographic location, past experiences or on what they think interests the recipient while others might depend on their character or personality [26].

B. Limitations

According to the work of R. Wald et al. [21] users of online social networks are becoming more aware of the information they post, and more concerned about how this information can be used to identify them. Most concerns related to social networking posts focus on raw demographic information or specific offensively posts, for example, religious or political affiliation may impact on how an employer views a potential employee, the same way as an inflammatory post or inappropriate photograph.

It is important to notice, as pointed on D. N. Chin and W. R. Wright work [28], that the different social media outlets each have different characteristics that will likely affect their effectiveness for personality profiling and, for that reason, it is doubtful that any single personality classifier will provide the best results for all social media. For example the limitation on number of characters used, such as the one found on

Twitter, can lead to unusual grammatical usage which can affect personality profiling trough text analysis.

IV. Online Social Networks as Sensors

At the end of the 20th century, M. Wiser et al. work [29] described the existence of a new field of computer science created by ubiquitous computing, a field with a vision of a physical world filled with sensors, actuators, displays, and other computational elements, embedded on the objects of the daily life and connected through a continuous network. A vision of what we can recognise today as a smart environment.

But it is possible to use an online social network as a sensor in a smart environment? Is there enough quality data available to adjust an environment based on his inhabitant preferences?

For the usage of an online social network like Facebook it is required, at minimum, a name, gender and a date of birth. Among these required basic fields, users can add basic facts about themselves such as home town, contact information, personal interests, job information and even a profile photograph. An online social network profile can provide insights about preferences, behaviour and emotions that can be captured in order to adapt and shift the environment around a person, or group of persons.

When talking about personality classification based on online social networks, Ghavami et al. [18] work affirms that it is possible to avoid having the standard test scores in order to identify user personality by finding relationships between user's behaviour and personality, or even connection between a user's network properties and personality. For instance D. Chapsk [30] assert that data disclosed on the popular online social network Facebook, can be used to make probabilistic inferences about an individual's socio-economic status, cognitive ability, life outcomes, cultural preferences, developed behaviours, average income, educational attainment, family size or even what genres of movies he is interested in.

For example, if someone publicly announces on his online profile "I'm disappointed that i missed the show last night", through text analysis it is possible to understand the negative emotion towards missing an event (in this case a show) in a specific time interval (last night). Theoretically speaking it shouldn't be hard to cross this reference with other data available, like TV guide schedules, newspapers, between other sources (since 'show' is ambiguous and can be related to a numerous of events), and find possible results for the show that were missed. However, due to the generality of an online social network be showed as a graph and defined as a network of interactions and relationships [23] and the fact that people express their preferences and make their online profile a mirror of their true identity, some network relationships can provide clues about certain topics and, in this particular example, a relationship with a TV show can be the 'missing show' referenced in the publication.

But how can we take advantage of this? The usage of online social networks as sensors can gather information that would be missed otherwise and takes advantage of the tendency of people and organisations to shift themselves into this virtual world. In the example described before, the knowledge that a person is expressing a negative emotion towards missing a specific TV show, can be passed to an actuator in order to play that TV show next time that person turns on the television. This can extend to numerous areas and applications, in a similar way to the example, if a friend creates a publication on your page saying "Hey, you should listen The Beatles - Let it Be" and you reacted positively that suggestion, that song should be automatically added to your music playlist. The personality profiling trough online social networks can be helpful to the personalisation of an environment in function of his inhabitant, and the detection of humour shifts trough text analysis can be used to adapt the environment to match the current state of emotion.

As an ambitious example of this application can be described on a context of a social event. Let's say that a room contains five people, with data present on their online profiles, the room should be able adapt his characteristics based on the common preferences between the five. The room music playlist or even the content playing on TV, for example, should be adapted and adjusted based on their common preferences in order to contribute to a more harmonious environment.

V. The Concept and Future Work

Due to online social networks being a part of the daily life of numerous people around the world, theoretically, there is no need to teach them how to use them on a basic level. Combined with the fact that there is a history of usage, online social networks contain an immediate significant amount of data ready to be analysed.

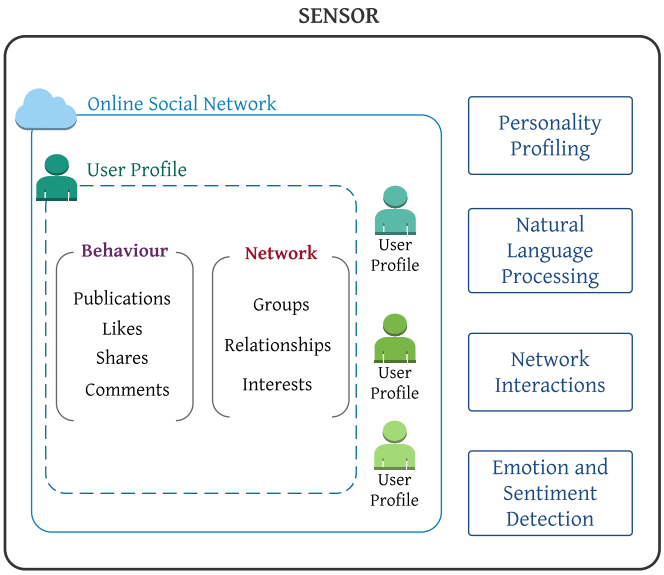

Fig. 1: An online social network based sensor that uses online user behaviour and network references to profile an individual

With the help of data mining techniques and natural language processing, behaviour clues present in publications, likes, shares, or comments (by an user or a group of users) can be used to profile personality accordingly to the Big-Five Model. In a similar way, network connections like groups, relationships, or interests can be used to understand personal preferences such as music, movies, brands, events, or people,

just to name a few. Natural language processing can as well be useful in this context to understand some connections that are expressed in written form, for example "I really like the song Let it Be by The Beatles", that could have not been yet expressed in other way such as a network connection to the band in question.

In the context of smart environments, the combination of behaviour and network characteristics with personality profiling, natural language processing, network interactions and emotion and sentiment detection, results in the general conceptualisation of an online social network based sensor that can be seen in figure 1.

A vital process for the sensor is monitoring the desired profiles of an online social network and gather any updated information related to them. Fortunately this process is simplified since the most popular online social networks offer Application Programming Interfaces capable of extract the desired behaviour and network related content. Relatively to personality profiling, it is based on processing the natural language founded on the online profile combined with machine learning techniques and a best case scenario would include a comparison with the results of a personality questionnaire, especially at the early stages. Network interactions as well as emotion and sentiment detection also benefit from natural language processing and machine learning, the combination of all this processes should be enough for feeding enough quality information to a smart environment in order to know the behaviour and personality characteristics of his inhabitant(s).

VI. Conclusion

It is proved that personality takes a huge role in humans and is a decisive factor that differentiates each of us, in such a way that our actions and patterns of behaviour are strongly connected to our personality traits. Due to this fact, online social network users share a lot of information about themselves and give many clues about their personality traits in some ways that they do not even notice. Smart environments benefit for knowing the personal characteristics of his users and in this paper we exploited the possibility of using online social networks as a sensor for this environments. From demographic information to overall interests and preferences, users liberally express themselves online in a free way, without any restrictions and this valuable information can be automatically fed to a smart environment to create a more personalised and responsive environment for his user. We also presented a concept for this type of approach and outlined a possible path to successfully achieve the desire results.

References

[1] T. G. Stavropoulos, E. Kontopoulos, N. Bassiliades, J. Argyriou, A. Bikakis, D. Vrakas, and I. Vlahavas, "Rule-based approaches for energy savings in an ambient intelligence environment," *Pervasive and Mobile Computing*, vol. 19, p. 123, 2015.

[2] M. Cristani, E. Karafili, and C. Tomazzoli, "Improving energy saving techniques by ambient intelligence scheduling," *Proceedings - International Conference on Advanced Information Networking and Applications, AINA*, vol. 2015-April, p. 324331, 2015.

[3] P. Cottone, G. Lo Re, G. Maida, and M. Morana, "Motion sensors for activity recognition in an ambient-intelligence scenario," *2013 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2013*, no. March, p. 646651, 2013.

[4] P. Cottone, G. Maida, and M. Morana, "User activity recognition via kinect in an ambient intelligence scenario," *IERI Procedia*, vol. 7, p. 4954, 2014.

[5] C. Ramos, "Ambient intelligence a state of the art from artificial intelligence perspective," *Progress in Artificial Intelligence*, vol. 18, p. 285295, 2007.

[6] M. L. Mekhalfi, F. Melgani, A. Zeggada, F. G. B. De Natale, M. A. M. Salem, and A. Khamis, "Recovering the sight to blind people in indoor environments with smart technologies," *Expert Systems with Applications*, vol. 46, p. 129138, 2016.

[7] Statista, "Global social media ranking 2016 | statistic," 2016.

[8] A. Ortigosa, J. I. Quiroga, and R. M. Carro, "Inferring user personality in social networks: A case study in facebook," *2011 11th International Conference on Intelligent Systems Design and Applications*, p. 563568, 2011.

[9] S. b. Poria, A. d. Gelbukh, B. Agarwal, E. b. Cambria, and N. Howard, "Common sense knowledge based personality recognition from text," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8266 LNAI, no. PART 2, p. 484496, 2013.

[10] S. Argamon, M. Koppel, J. Pennebaker, and J. Schler, "Automatically profiling the author of an anonymous text," no. Sebastiani 2002, 2009.

[11] J. Golbeck, C. Robles, M. Edmondson, and K. Turner, "Predicting personality from twitter," *Proceedings - 2011 IEEE International Conference on Privacy, Security, Risk and Trust and IEEE International Conference on Social Computing, PASSAT/SocialCom 2011*, p. 149156, 2011.

[12] S. Adali and J. Golbeck, "Predicting personality with social behavior," *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, p. 302309, 2012.

[13] L. R. Goldberg, O. P. John, H. Kaiser, K. Lanning, and D. Peabody, "An alternative description of personality : The big-five factor structure," vol. 59, no. 6, p. 12161229, 1990.

[14] S. Sulaiman, D. R. Awang, K. Amirah, and A. Halim, "A study on the relationship between personality traits and image perceived," 1999.

[15] D. Markovikj, S. Gievska, M. Kosinski, and D. Stillwell, "Mining facebook data for predictive personality modeling," *Proceedings of the 7th international AAAI conference on Weblogs and Social Media (ICWSM 2013), Boston, MA, USA*, p. 2326, 2013.

[16] J. W. Pennebaker and L. A. King, "Linguistic styles: Language use as an individual difference," *Journal of Personality and Social Psychology*, vol. 77, pp. 1296–1312, 1999.

[17] K. Moore and J. C. McElroy, "The influence of personality on facebook usage, wall postings, and regret," *Computers in Human Behavior*, vol. 28, no. 1, p. 267274, 2012.

[18] S. M. Ghavami, "Facebook user s like behavior can reveal personality," p. 79, 2015.

[19] Y. Bachrach, M. Kosinski, T. Graepel, P. Kohli, and D. Stillwell, "Personality and patterns of facebook usage," *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci 12*, p. 2432, 2012.

[20] U. Gupta and N. Chatterjee, "Personality traits identification using rough sets based machine learning," *2013 International Symposium on Computational and Business Intelligence*, p. 182185, 2013.

[21] R. Wald, T. Khoshgoftaar, and C. Sumner, "Machine prediction of personality from facebook profiles," *Proceedings of the 2012 IEEE 13th International Conference on Information Reuse and Integration, IRI 2012*, p. 109115, 2012.

[22] A. Nadkarni and S. G. Hofmann, "Why do people use facebook?," *Personality and individual differences - Elsevier*, vol. 52, no. 3, p. 243249, 2012.

[23] S. A. Moosavi, "Community detection in online social networks using actions of users," *Intelligent Systems (ICIS), 2014 Iranian Conference*, p. 1 7, 2014.

[24] J. W. Pennebaker, "Your use of pronouns reveals your personality," *Harvard Business Review*, Dec 2011.

[25] R. Pitcher, "Revealing the colour and personality in texts : Putting the person back into our results," vol. 19, p. 18, 2014.

[26] J. Oberlander and A. J. Gill, "Language with character: A stratified

corpus comparison of individual differences in e-mail communication," *Discourse Processes*, vol. 42, no. 3, p. 239270, 2006.

[27] J. Shen, O. Brdiczka, and J. Liu, "Understanding email writers: Personality prediction from email messages," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7899 LNCS, p. 318330, 2013.

[28] D. N. Chin and W. R. Wright, "Social media sources for personality profiling," *Proceedings of the 2nd Workshop Emotions and Personality in Personalized Services (EMPIRE 2014)*, 2014.

[29] M. Weiser, R. Gold, J. S. Brown, B. Sprague, and R. Bruce, "The origins of ubiquitous computing research at parc," *IBM Systems Journal*, vol. 38, no. 4, p. 693696, 1999.

[30] D. Chapsky, "Leveraging online social networks and external data sources to predict personality," *Proceedings - 2011 International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2011*, p. 428433, 2011.