User Recognition Using Cognitive Psychology Based Behavior Modeling in Online Social Networks

A. Saleema(✉) and Sabu M. Thampi

Indian Institute of Information Technology and Management-Kerala, Kazhakkoottam, India
{saleema.res17,sabu.thampi}@iiitmk.ac.in

Abstract. Computer-mediated social contexts, specifically Social Media has been identified as an appropriate platform to assess the human personality as well as behavior as it involves massive interactions and communications. Social Behavioral Biometrics is a very recently introduced research area that attempts to extract unique behavioral features from the social interactions of individuals, which are powerful enough to be used as biometric identifiers. This research investigates the feasibility of generating biometric templates from user profiles, which satisfies the basic requirements of a biometric identifier namely distinctiveness and stability. This work proposes a user identification as well as user behavior modeling approach using some novel concepts on the grounds of certain unassailable theories in Human Psychology. The authors have performed an empirical study on Online Social Network data by extracting certain psychometric properties of temperament from user profiles in order to generate unique and stable user behavior patterns. Based on the Cognitive Affective System theory of Human Personality which has been proved indubitable years ago in offline social contexts, we have created Situation-Behavior profiles of a sample set of Twitter users and analyzed that whether the theory befits the OSN data. The Big Five behaviors of users are measured by content analysis and are analyzed in different situations like Political, Personal, Entertainment, Sports and Technical. The results of the empirical study reveal that the Situation-Behavior profiles are pertinent enough to be used as biometric patterns with high stability and distinguishability.

1 Introduction

Nowadays, Social Media has become a standard way of life and thereby the identity of human beings is extending beyond the real world to an equivalent virtual world. User profiles in typical online social networks such as Facebook, LinkedIn, Twitter, Instagram, etc assigns them a virtual identity through which they perform massive interactions and communications. As evidenced by the ComScore report, approximately 1 of every 5 min spent online by internet users is for social networking sites [1]. Among these, Twitter-the rapidly evolving microblogging

S. M. Thampi et al. (Eds.): SIRS 2019, CCIS 1209, pp. 130–149, 2020.
https://doi.org/10.1007/978-981-15-4828-4_12

platform launched in 2006, is the most propitious one utilized for research due to its tight limits and restrictions in posting.

In the real world, a person can be identified using their Physiological as well as Behavioral traits which are categorized under the two classes of Biometrics. Physiological Biometrics make use of the static traits of the human body that are not subjected to change over aging, which includes the fingerprint, retina, iris, finger vein, face, palm, etc. Behavioral Biometrics, on the other hand, corresponds to the unique individual behavioral patterns which span from the walking style, typing style, speech, handwriting, etc to the most recent ones using social interactions and communications.

As far as computer-mediated social contexts are concerned, person authentication relies upon the username and password method, which doesn't guarantee adequate security, especially when we log into various social networking sites from different locations such as libraries or cafes. Also, the username password method is associated with security questions, which are quite painstaking to the users especially because of the need for memorizing the circumstantial questions and their answers. Most importantly, password-based authentication do not ensure a strong identity check, as it can be stolen/hacked. That means the security has highly relied upon the strength and confidentiality of the passwords.

The majority of people nowadays use social networks through mobile phones or other smart device applications, where login credentials are not verified every time. So in case of a device steal/loss or sometimes in cases of death, the social media profiles of the person can be misoperated by others. Therefore, the need for new static as well as continuous authentication techniques in computer-mediated social contexts is being critical to ensuring authenticity, in such an era of unfettered social media usage.

Social Behavioral Biometrics(SBB) is such an authentication approach that is defined as the identification of an actor/person/avatar based on their social interactions and communication in different social settings [2]. Social settings for studying SBB is described as the environments and their properties, over which the communication and interaction between actors take place. SBB is a most recently introduced area that seeks immediate research attention. The scope of social behavioral biometrics resides in continuous authentication in cyberspace as well as online social networking platforms. In today's scenario where millions of people are being connected to the OSNs in a day to day basis, the compromise of one user's account could pose malicious security threats to many of the user's acquaintances. While users of SNSs are authenticated continuously based on social behavioral biometric features, if the trust level of the user drops significantly, the system can take necessary actions such as forbidding the user from using the service, generating alarms, etc. This application would facilitate anomaly, fraud, or intrusion detection as well. The social behavioral features extracted from social networking sites contain predominant information about the user and there comes the possibility of automatically generating security questions which would increase the security to the users' account as well as lessen the strain of setting security questions and answers. The scope and relevance of the Social Behavioral Biometric(SBB) feature extend to the integration of SBB with Multimodal/Multifactor authentication systems. The combination of SBB with other behavioral biometric methods opens a new door of research since it improves the uniqueness and enhances identification or verification rates of users in the virtual world. SBB features can be used for author identification in cases of law enforcement and forensic applications of the investigations of cybercrimes. The scope of social behavioral biometrics is not confined to person authentication but spans to risk analysis, situation analysis, anomaly detection, access control, customer profiling, etc [2].

The massive data on online social networks have already been explored by researchers in the past decade for new application fields such as sentiment analysis, stock market, election results prediction, education tools, opinion mining, etc. A multitude of works has been carried out utilizing the concepts of personality to detect the factors influencing the use of social networks, the addictive tendencies towards social networking sites, the way how people behave online, etc. Most of the research work investigated how the five factors of personality are correlated to the social networking sites' duration and frequency of usage, sharing of information, the number of friends, etc. Also, using the big five model as a basis, researchers have identified how the different traits are related to different types of users namely listeners, influencers, popular, etc. Beyond these, personality-based research using online social network data is proliferating in diverse directions such as the development of friend suggestion methods, targeted marketing, automatic website morphing using cognitive styles, authorship attribution, etc.

Even Though the human behavioral patterns have been studied for individual behavior analysis over the past decade, only a few researchers considered the potential of OSN data as social behavioral biometric templates/features. The previous works on SBB explored frequency-based and knowledge-based social behavioral biometric features, extracted from online social networking sites. More specifically, they have created distinct biometric profiles of users by utilizing the reply/retweet networks, hashtag networks, URL, friendship network, etc as features. The major challenges in creating such biometric profiles include the difficulty in extracting features from huge random datasets, selection of small consistent features for identification and adaptation to the change in behavior over time. Although these methods could achieve reasonable recognition rates, these biometric profiles were not sufficiently unique to identify a person with high precision. However, in combination with other modalities, identification systems with high accuracy and precision were obtained. Also, the spatial and temporal information collected from social networking sites while combined with these SBB profiles resulted in high recognition accuracy. So, there is a big motive in developing social behavioral biometric algorithms to efficiently authenticate individuals from the online social network data, that provide high uniqueness and accuracy.

We have performed an extensive study on psychology and its applications in online social networks and discovered that there is substantial scope in exploring cognitive psychology as well as human psychological theories of Disposition in order to generate SBB features. Since the traditional cognitive psychological personality theories based on temperament have been well investigated and proved to be irrefutable, the use of these aspects in the context of online social networks will be highly admissible with appreciable reliability and accuracy, as the conversations/interactions in OSN possess much resemblance to the real-world situations. Our paper gives a direction to move a first step towards utilizing these aspects for person authentication, providing high stability and uniqueness.

Besides person identification, a good deal of applications is expedient by integrating the social network analysis with cognitive psychology. According to our analysis, the utilization of the social media platform for person identification along with context awareness and cognitive psychology will actualize the future generation intelligent automated biometric systems, that are expected to mimic the human brain. Figure 1 is an illustration of such an intelligent system that will automate the process of user recognition, user change detection, behavior analysis, rumor detection, targeting and recommendation systems, etc.

Our research work aims to collaborate with cognitive psychology and online social network analysis in order to extract person distinctive properties. The notion behind this collaboration is the clear and solid descriptions of the Dispositions and Temperament aspects of personality in human psychology. The term personality itself is defined in psychology as a combination of characteristics and qualities that form an individual's distinctive character. Temperament refers to the qualities distinguishing a person while Disposition implies the customary

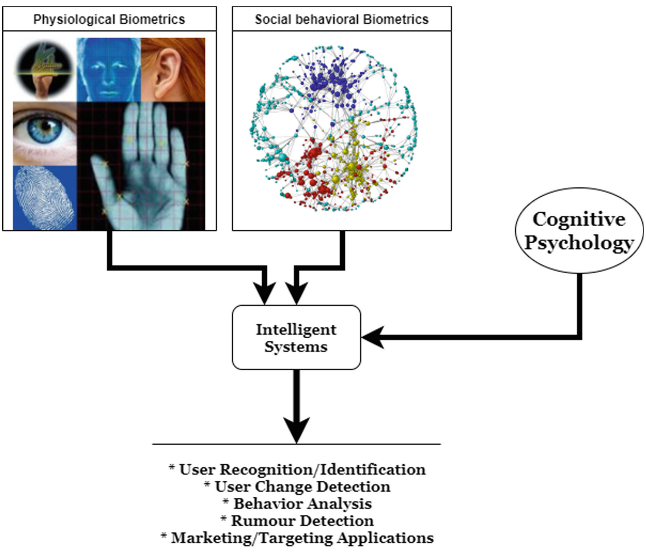

Fig. 1. The depiction of a next-generation Intelligent System that automatically recognizes users, profile changes, rumors, etc.

moods and attitude towards the life around one. Our research work employs the conception of personality in terms of behavioral dispositions or traits that predispose individuals to engage in relevant behaviors.

The key contributions of this research work are as follows:

1. Our work is the first to collaborate the concept of cognitive psychology with online social network analysis in order to check the feasibility of generating biometric templates.
2. We have performed an empirical study on social network data based on the Cognitive Affective System Theory of Personality to extract Situation-Behavior Profiles of users. Two hypotheses related to the properties of stability and distinctiveness are formulated and tested on the OSN data.
3. This work proposes a behavior modeling technique that makes use of emotion extraction based content analysis and mapping the emotions to the PleasureArousal-Dominance temperament space. Further, it employs the evaluation of Big Five traits from the PAD space.
4. We have compiled a Twitter user dataset with user tweets from two separate periods and annotated with situations such as Personal, Political, Entertainment, Technical, etc to suit our research problem.

The organization of the remaining sections in the paper are as follows: Sect. 2 examines some of the relevant works in the area of social behavioral biometrics, personality prediction based research works in social networks, modeling of psychological theories in social media and some of the computational models of personality from psychology. In Sect. 3, we state our hypotheses, maps the CAPS theory to the social networks and elaborate our proposed approach for modeling human behavior and creating the Behavioral Signatures. In Sect. 4, we brief the experimental setup in which we outline the details of dataset collection, summarize the findings based on the preliminary analysis of the dataset and discuss the results of our empirical study. Section 5 concludes the work and specifies the limitations and future works.

2 Related Works

This section reviews some research works on social behavioral biometrics, related works on behavioral modeling using psychological theories in social media, some relevant works on behavior and personality predictions on social media and some significant computation models of personality from the literature. The concept of SBB was introduced in [2] and it defines SBB as the identification of an actor based on the communications and social interactions in different social settings. The interaction between actors, the communication manner, etc is considered while generating SBB. It was the first study on social behavioral biometrics which deals with human social interactions from the aspect of identifying a person. The authors have created reply networks, retweet networks, URL networks, etc and applied as the features for person identification.

In [3], the authors have attempted to fuse the social behavioral data with other physiological biometrics using confidence-based decision fusion to develop a multimodal approach, being the first work of decision fusion for multimodal biometrics using social network data. Later in 2017, they have attempted to recognize users in a computer-mediated social context [4]. They have created a new framework by identifying several social biometric features, proposed a new matching method and evaluated three basic biometric properties namely accuracy, uniqueness and stability. The above-discussed ones are purely based on the extraction of social media profile features but not related to psychological concepts. Now we discuss some of the approaches which model the traditional theories of human psychology in OSNs, but not in an authentication perspective.

Not many research works have been done to model social media using theories of human psychology and that's why it remains a fertile research field. Significant work was done by Tyshchuk [5] in which the well-known Theory of Planned Behavior is applied. According to the theory of planned behavior [6], an individualś intention to perform a behavior is driven by factors like subjective norms, attitude towards the behavior and perceived behavioral control. In this work, the authors have identified features from the social media profiles that can be well mapped to the factors of the theory and applied TPB to predict their behavioral intents. They have concentrated on the behavior that happens in response to a trigger, specifically an event. A set of elements are selected as features and they were measured by using natural language processing and social network analysis techniques. Another significant work was done by Tripathi [7] which used the Big Five Model as the basis.

Bachrach et al. [8] have presented a work based on examining the correlation between user's personality and the features of online profiles such as the density, size, etc. of friendship networks, number of events, group memberships, number of tags, etc. They could show significant relationships between user personality and the SNS features and showed how multivariate regression is utilized for the prediction of the personality traits of an individual user, given their Facebook profile. Significant work to understand the information sharing behavior using socio-cognitive theory was done in [9].

Bandura's socio-cognitive theory [10] is a well-accepted theory in psychology and is an extension of his own social learning theory [11]. The socio-cognitive theory can be stated as the fact that people do not learn solely by their success or failure, but it is depended upon the replication of the action of others. Bandura has applied SCT to analyze the influence and diffusion of new behaviors through Social networks in [12]. Most of the papers in the literature are focused on personality predictions and behavior predictions which can be applied to the recommendation likes systems.

A considerable amount of research works have been found in the literature which relates to the study of behavior and personality prediction in social media, identifying the psychological connection between personality and social media usage, etc. The majority of the reported works explored the Big Five model of personality which is a strong consensus on the basic traits of human beings,

Table 1. Related Works on Behavior and Personality predictions in social media

Author and Year	Purpose	Social Network Type	Features used
Ross et al. 2009 [13]	Study the influence of personality and competency factors related to facebook usage	Facebook	Attitude, Online sociability
Hamburger et al. 2010 [14]	Examines the connection between personality and faceboook usage	Facebook	No: of friends The user bio Group, admin
Wilson et al. 2010 [15]	Identify psychosocial characters of young people's use of social networks	Myspace, Facebook	Time spend, Attitude
Quercia et al. 2011 [16]	Predict user personality and find the similarity and differences between different types of users	Twitter	Followers, Following, Listed counts
Golbeck et al. 2011 [17]	Predicting Big Five traits from profile information	Twitter	Tweet analysis
Buettner 2016 [18]	Product recommendation system from personality analysis	Xing	No: of contacts, Groups, Page views, Work experience
Pornsakulvanich 2017 [19]	To find the contributing factors to online social support satisfaction and frequency	Facebook, Instagram, Line	Usage duration, Attitude, Social influence, Friends etc
Tedesse et al. 2018 [20]	Predicting the personality traits	Facebook	Content analysis, networksize, betweenness, density, brokerage, transitivity

Table 2. Computational models of personality

Model type	Author and year
Neural Network Model	Pozanski and Thagard 2005 [21]
Neural Network Model	Quek and Moskowitz 2007 [22]
Knowledge and Appraisal Model	Cervone [23]
Personality Systems Interaction Model	Kuhl 2000 [24]
Cognitive Affective System Theory	Mishel and Schoda 1995 [25]

initially evidenced by Fiske in 1949. Table 1 summarizes some of the recent relevant works in this area, all of which have used the Big Five model as the basis.

Now we would like to appraise some of the computation models of personality, that can be explored in the future, on the grounds of social network data. Table 2 delineates several well regarded computational models of personality which will benefit in the study of applying psychological aspects to social media.

From the above-reviewed literature, it is evident that not a single work made an attempt to model any human psychological theories to extract person distinctive properties with a view of authentication.

3 Proposed Approach

We have performed an empirical study on online social network profiles on the grounds of the Cognitive Affective Personality System Theory. Also, we have used the concept of the Pleasure-Arousal-Dominance temperament model and the big five personality traits from psychology. The proposed approach is outlined in three subsections namely the background theory, formulation of hypotheses and the methodology of behavior extraction.

3.1 Background Theory and Formulation of Hypotheses

3.1.1 Cognitive Affective Personality System

The theory accounts for the differences in individuals in predictable patterns of variability across situations. ie If situation A, then she behaves X, but if situation B, then she behaves Y. Also, it states that the average levels of behavior are the behavioral signatures of the same fundamental personality system. The reason for this is illustrated by a framework termed as the Cognitive Affective Personality System.

The Situation-Behavior Profiles (if-then profiles) are proved to be intra-individually stable by an empirical study. Also, these profiles are considered as behavioral signatures of a person, which means that these are unique for each person. Since the properties of uniqueness and stability are satisfied in their studies according to the theory, we have considered the possibility of generating

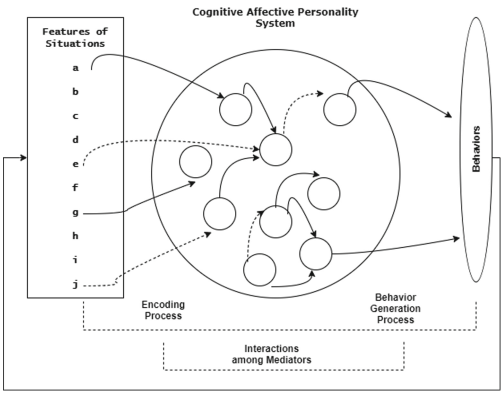

Fig. 2. The simplified illustration of the cognitions and effects as mediating processes in a Cognitive Affective Personality System.

Table 3. Cognitive affective units in the personality system

Encodings: Categories (constructs) for the self, people, events, and situations (external and internal)
Expectancies and Beliefs: About the social world, about outcomes for behavior in particular situations, about self- efficacy
Affects: Feelings, emotions, and affective responses (including physiological reactions)
Goals and Values: Desirable outcomes and affective states; aversive outcomes and affective states; goals, values, and life projects
Competencies and Self-regulatory Plans: Potential behaviors and scripts that one can do and plans for organizing action and affecting outcomes and one's own behavior and internal states

Note: From [25] in 1995

a biometric template from these Situation-Behavior profiles, on the background of social networks.

In order to explain how and why these behavioral signatures are exhibited, the authors incorporated the role of situations, events or contexts into the conception of personality. Figure 2 shows a combined view of the personality system. Each individual is characterized in terms of two things as follows

(a) The cognitions and affects.
(b) The distinctive way of organization of the interrelations among cognitions and effects and psychological features of situations.

When certain situations are encountered by an individual, some particular situational features are activated and it inturn activates a set of cognitions and affects which are characteristic in nature. So, as the figure indicates, within an individual a characteristic set of interrelationships among the cognitive and affective units further activates certain units and finally activates some behavior. The types of cognitions and affects are described as shown in Table 3 and the structure of CAPS is depicted in Fig. 2.

3.1.2 Formulation of Hypothesis

Based on the conception of behavioral signatures of disposition, we have formulated two hypotheses on the grounds of online social networks.

Hypothesis (H1): People in online social networks show considerable variability in similar behaviors say extroversion, agreeableness, conscientiousness, emotional stability and sophistication across various situations like political, entertainment, technical, personal, sports, etc □

Hypothesis (H2): People tend to show high consistency in those behaviors within the same classes of situations and their Situation-Behavior profiles are stable and distinctive enough to be utilized as a biometric template. □

Traditional researchers in human personality commented that the basic coherence in the underlying personality (Disposition) results in cross-situational consistency. But later on, as research embellished, researchers unveiled the role of situations and theorized about the low cross-situational consistency in behavior. Mishel and Shoda have later proved that individuals' behavior shows less cross-situational consistency and the behavior patterns across situations are intra-individually stable. They have provided empirical evidence by analyzing the if....then...profiles of a set of individuals. The if...then..profiles are termed as Situation-Behavior profiles and are considered as signatures of Personality. In their study, the situations were defined in nominal terms as places and activities in the setting, ie arithmetic tests, dining halls, woodworking activities, school playgrounds, etc. They have selected some dimensions of behavior such as verbal aggression, friendliness, withdrawal, prosocial behavior, etc in a study conducted in a residential camp setting. For each person, the Situation-Behavior profiles are examined and assessed the stability and distinctiveness.

In our study, we have mapped the entire entities from the real world to a virtual cyber world. The conviction behind this mapping is that the interactions, interrelationships, communications, etc in online social networks possess much resemblance with the real face to face conversations. Therefore we have reformulated the hypotheses in the context of the online social networks as in hypotheses 1 and 2.

Hypothesis 1 corresponds to the behavior variability across situations more specifically the if-then relations, which are to be proved intra-individually stable. Hypothesis 2 corresponds to the stability and distinctive property, ie the if..then..profiles are stable as well as unique for each individual.

3.2 Mapping of CAPS to the Context of Online Social Networks

The real-life situations like a family party, group meeting, worksite, etc, when mapped to the online social network becomes political, technical, entertainment, personal, etc according to the posts. Especially on twitter, there will be hashtags or URLs associated with the majority of the posts. In the real-world, the behaviors taken for the study were friendliness, verbal aggressiveness, anxious, being silent, etc, but in OSN we have considered the big five traits proposed by J.M.Digman. Table 4 shows the mapping of the situations and behaviors from the real world to the online social network.

In the study conducted with students in the residential camps, they have identified the situational features as the valence of the interaction which can be either positive or negative, the type of person(whether adult counselor or child peer), etc. Here we have identified the situational features like whose tweet, the no: of replies, it has obtained so far, the no: of retweets, whether it is associated with a hashtag or a URL, which all people replied for it, etc. The organization of the cognitions and affects shown inside the circle is the same for the OSN

Table 4. Mapping of Situations and Behaviors from real world to Online Social Network

Real life contexts	OSN contexts	Real life behaviors	Behaviors in OSN
Family Party	Political	Verbal Agressiveness	Extroversion
Group Meeting	Technical	Friendliness	Conscientiousness
Working Site	Personal	Anxious	Emotional Stability
Working in an urgent project	Disasters	Tease and make fun of	Agreeableness
Individual Assignmen	Entertainment	Surf web	Sohistication
Trying to get a date	Business	Be Silent	
	Sports	Explore environment	

context also, as it purely represents the underlying cognitive structure of the personality system.

The whole system of CAPS in the context of OSN is shown in Fig. 3 and can be described as follows. The figure shows how behavior signatures are produced by a personality system characterized by the available cognitive-affective units, organized in a distinct way of inter-relationships. When a situation is encountered, say a tweet related to a political party, some situational features say whose tweet, which all friends replied will stimulate certain cognitions and affects inside that person. Since the way of organization is different for different individuals, the stimulations or activations of the units will vary. Within any individual, a rich set of relationships exists between the units and according to them, they will activate or deactivate further units and ultimately behavior is generated. The situational features that causes stimulation of these units also will be different for different individuals. For example, some people will reply if their friends are

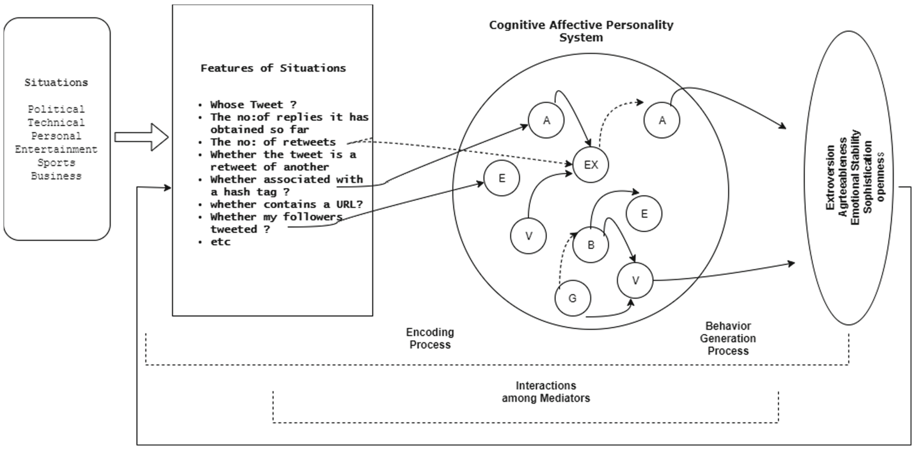

Fig. 3. The illustration of how behavioral signatures are generated, with the help of CAPS theory in the Online Social Network context

replied, some others will retweet or reply if there is a presence of a hashtag or a URL since these things represent clear shreds of evidence of the subject in hand. The mediating units are activated in response to some situational features and are deactivated with respect to some situations. ie The relation may be positive (represented by the solid line) or negative (represented by the dashed line).

3.3 Methodology

The Big Five Behaviors in OSN can be measured in various ways. In [7], two methods, content-based analysis, and linkage data-based analysis are discussed for evaluating the behavior. The huge amount of contents such as text, image, video, etc are used in content analysis while the network data which includes the relationship among the entities are used for linkage data-based analysis. Factors like no: of friends, whether profile photo is displayed, whether user biodata is displayed etc helps in determining which category of behavior the users belong to. Several deep learning models are available for detecting the big five traits directly from the text. We have performed a content analysis by predicting the emotion probabilities from the tweet text instead of the big five traits since we were interested to analyze the emotional dispositions of individuals. The proposed process of extraction of big five behavior from tweets is as explained below.

Initially, the tweets in two separate periods of the sample set of users are collected and labeled with situations such as technical, entertainment, political, personal, sports, etc. The tweets are preprocessed by removing the URLs, Mentions, Hashtags, Twitter-reserved words, punctuations, single letter words, blank spaces, stopwords, profane words, and numbers.

The processed tweets are then given to a trained Recurrent Neural Network model presented in [26] for predicting emotions. Most of the emotion recognition models for twitter use lexicons and bag-of-word models. We have taken the Recurrent Neural Network model for consideration since it proved to be more efficient than the state-of-the-art.

The Ekman's emotions are then mapped to the Pleasure-Arousal-Dominance (PAD) temperament space using a transformation matrix. Several mapping techniques have been proposed in the literature for transforming the Ekman's emotions to the PAD space, among which the one illustrated in [27] seems the most appealing. The mapping matrix presented presented in [27] is as follows
PAD[Happy, sad, Angry, Scared, Disgusted, Surprised, 1]=

$$\begin{bmatrix} 0.46 & -0.30 & -0.29 & -0.19 & -0.14 & 0.24 & 0.52 \\ 0.07 & -0.11 & 0.19 & 0.14 & -0.08 & 0.15 & 0.53 \\ 0.19 & -0.18 & -0.02 & -0.10 & -0.02 & 0.08 & 0.50 \end{bmatrix} \tag{1}$$

The mapping of Ekman's emotions to the Pleasure-Arousal-Dominance temperament space can be defined mathematically as the matrix multiplication of the mapping matrix in Eq. 1, with the emotion vector obtained from the pre-trained model.

We have chosen the PAD temperament space since these are characteristic emotional predispositions that are stable over the lifetime of an individual. The three measures Pleasure, Arousal and Dominance are proved to be orthogonal. The trait Pleasure refers to the relative predominance of positive versus negative affective states across a life sample of situations. The scale Arousability is the measure of how easily a person is aroused and returned by certain stimuli. The third one is the Dominance-Submissiveness scale that refers to a person's characteristic feelings of influence and control over his life circumstances versus feelings of being controlled and influenced by others or events.

Then, as the last step, we have computed the big five Traits of behavior using the PAD space with the linear mapping equations proposed in [28], as shown below.

$$Extraversion = 0.23P + 0.12A + 0.82D \tag{2}$$

$$Agreeableness = 0.83P + 0.19A - 0.21D \tag{3}$$

$$Conscientiousness = 0.32P + 0.30D \tag{4}$$

$$EmotionalStability = 0.57P - 0.65A \tag{5}$$

$$Sophistication = 0.33A + 0.67D \tag{6}$$

The transformation from PAD space to the big five space can be expressed as $f(P, A, D) : R^3 \rightarrow R^5$, where R^3 represents the PAD space and R^5 Five space. If V is the PAD vector and W is the Big Five Traits vector, this linear transformation can be represented using matrix multiplication as Five space.

$$\overrightarrow{W} = A\overrightarrow{V} \tag{7}$$

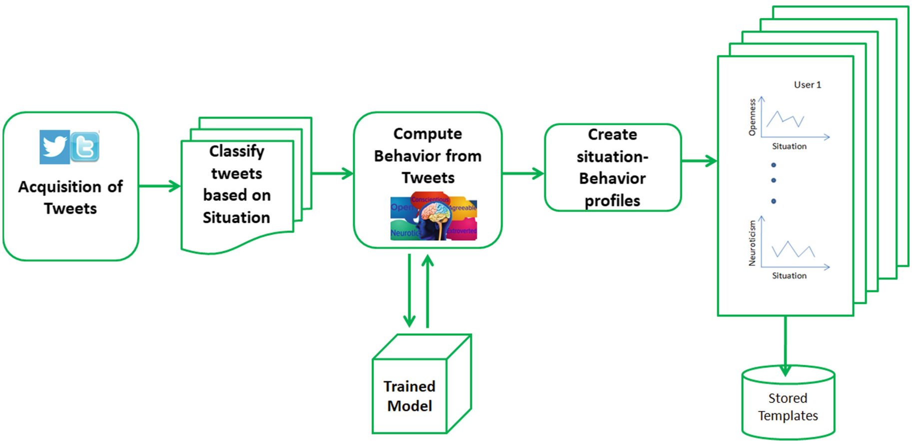

Fig. 4. The illustration of extracting the Situation-Behavior Profiles

$$whereA = \begin{bmatrix} 0.23 & 0.12 & 0.82 \\ 0.83 & 0.19 & -0.21 \\ 0.32 & 0 & 0.30 \\ 0.57 & -0.65 & 0 \\ 0 & 0.33 & 0.67 \end{bmatrix} \tag{8}$$

The whole process of Situation-Behavior Profile extraction is depicted in Fig. 4.

4 Experimental Results

This section provides the experimental results of our evaluation of the Hypothesis 1 and 2 on a sample set of twitter users. All the experiments including the implementation of our methodology for Situation-Behavior Profile extraction are carried out in a 3.50 GHz Intel(R)Xenon(R) CPU with 64GB RAM. Python 3.7 is used for the implementation of emotion and behavior extraction.

A dataset of 50 sample users from two non-overlapping periods(sessions) of the time was collected for the analysis in the following manner. Initially, a user with a high engagement timeline is selected as a seed. From his acquaintances, another 2 prolific users were selected and continued this manner until we got 50 prolific users.

The dataset was collected using 'vicinitas', which is an online tool that helps track and analyze real-time and historical tweets. For each user, we have collected information such as the tweet Identifier, the tweet text, screen name, the

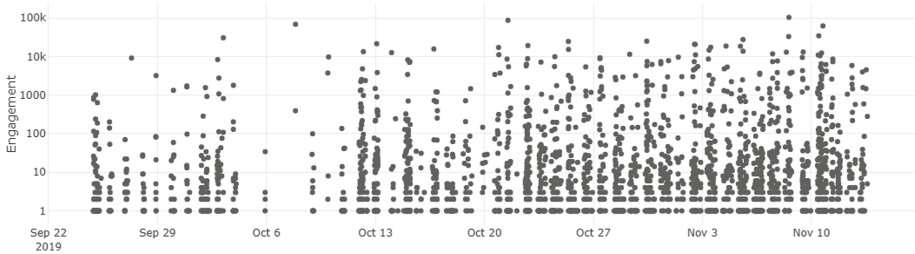

Fig. 5. The engagement timeline of a particular user during a particular session, analyzed using vicinitas tool

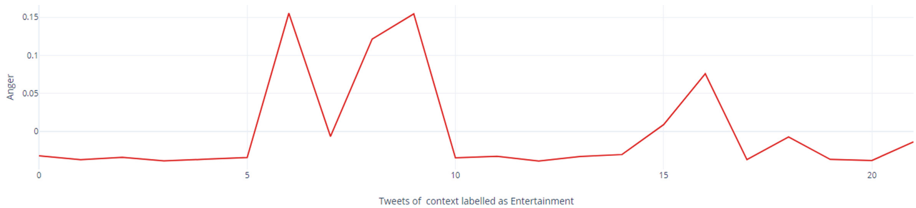

Fig. 6. Graph plotting the standard deviation of the emotion 'Anger' across the tweets of same context 'Entertainment' of User #1

Table 5. Sample tweets annotated with Situations

Tweet	Situation
Heading to Hamburg to meet old friends and new ideas	Entertainment
I hope that's the case. But Android seemed to have confirmed it	Technical
Seriously impressed with Deepak Chahar's bowling, he can strike a few with the bat too	Sports
Delhi's air strays deeper into 'severe' zone, likely to dip to 'severe+' by Thursday	General
Again Joh** baffles his opponents. The video is a piece of clever political communication	Political
Top films for me Super30	Entertainment
I need to work harder?	Personal
Captain Rohit. I shudder to think had he not been leading today	Sports

date and time of the tweet, Tweet type, the no: of retweets, etc. The users we have selected are those having an average of 25 tweets per day. We could analyze the engagement timeline day-wise and hour-wise, the percentage of tweet types, the percentage of rich media, etc using this tool. The engagement timeline of a particular user at a particular session is shown in Fig. 5. The tweet type represents whether the tweet is a reply, retweet, mention or a normal tweet. We have done pre-processing of the dataset by removing elements such as URL, mentions,

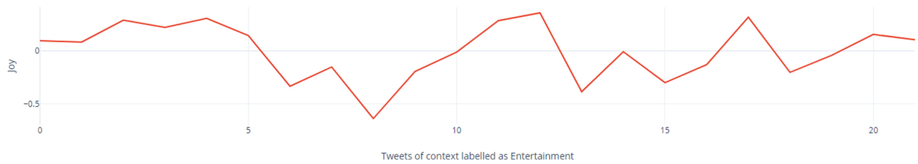

Fig. 7. Graph plotting the standard deviation of the emotion 'Joy' across the tweets of same context 'Entertainment' of User #1

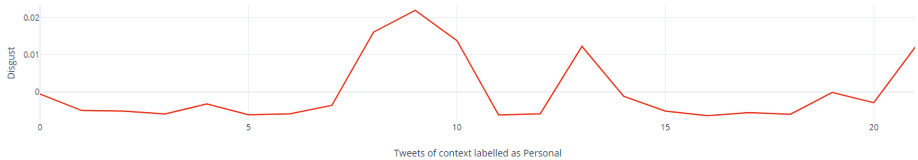

Fig. 8. Graph plotting the standard deviation of the emotion 'Disgust' across the tweets of same context 'Personal' of User #1

hashtags, twitter reserved words, punctuation, single letter words, blank spaces, stopwords, and numbers, as these will not contain any necessary information about the users' characteristics. Also, we have filtered the tweet lists by removing the retweets since it may be written by another user and will not benefit our content analysis.

After filtering and pre-processing, we have annotated the tweets of each user with the 'situation' field, which represents which context the tweet can be related to. This was a time-consuming task, which required a lot of manual effort and considered as the major limitation of this study. However, automatic annotation of the context/situation of a post can be made practical in the near future, utilizing the well advanced natural language processing and artificial intelligence techniques, which will ease the work of all the researchers related to OSN. A sample set of tweets with the annotated situations are shown in Table 5. The number of tweets for each situation varied for different users. So the probability of each tweet to belong to a particular situation differed for every user. The interest or enthusiasm for each user is different for different situations and therefore some people may not respond at all in particular situations, say sports, politics, etc.

We have evaluated hypothesis 1, by plotting each of the big five behaviors with respect to various situations like entertainment, technical, personal, sports, etc and found to be true for each of the users in our dataset. The individuals have shown ample variations in the conditional probability of exhibiting behaviors across various situations. These plots are termed as Situation-Behavior profiles or if..then.. relations and it confirms the intra-individual behavior variability across situations. The Situation-Behavior profiles of two randomly selected users are shown in Fig. 11.

We have analyzed Ekman's emotions extracted using the method prescribed in [26] for the whole tweets and arrived at the finding that during a particular context, the specific behavior of a person remains more or less similar. An

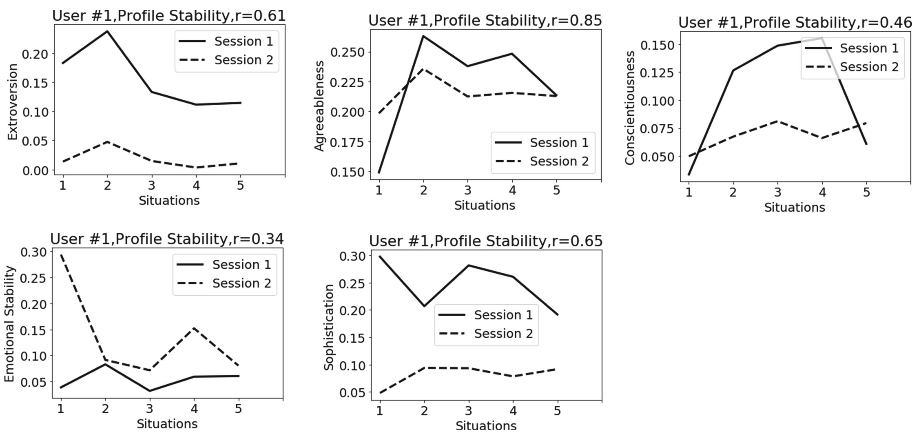

Fig. 9. Situation-behavior Profiles of User #1 at Sessions 1 and 2

example illustration of the emotion 'Anger' extracted from the tweets of the same situation 'Entertainment' during session 1 is shown in Fig. 6. On the vertical axis, the standard deviation of emotion 'Anger' with respect to the normative mean(zero) is shown with respect to the tweets related to entertainment in the horizontal axis. This clearly shows that the values are always nearer to zero, with an overall standard deviation of 0.063. This indisputably indicates the high intra-situational consistency of individuals in expressing emotions. Figure 7 shows the standard deviation of another emotion joy in the same situation entertainment for the same user. Its overall standard deviation is 0.26. The graph plotting Disgust against the tweets of 'Personal' situation is shown in Fig. 8.

About 65% of the If..then Profile shows the stability coefficient greater than 0.5 and when the Agreeableness-Situation profiles alone are considered, around 89% of individuals show stability coefficients greater than 0.65. This strongly recommends Agreeableness as the most stable behavior.

Figure 9 shows the Situation-Behavior Profiles of user #1 for each of the behaviors extroversion, agreeableness, conscientiousness, emotional stability, and sophistication. We have plotted and analyzed the If..then relations of all users in our dataset in two separate sessions producing two profiles for each individual. The profile at session 1, having a duration of 4 months is plotted as a solid line, where the dashed line represents the profile of the same person at session 2 of another 4 months duration. The time period of sessions 1 and 2 are non-overlapping. We have evaluated the stability of the profile using the correlation coefficient, r and found out that the profile at two times shows impressive stability. Here in Fig. 9, the person shows high stability in agreeableness, extroversion, and sophistication while the behaviors conscientiousness, emotional stability is less stable. The user #29 shown in Fig. 10 has exhibited high stability in extroversion, agreeableness, and emotional stability. The traits conscientiousness and sophistication are less stable with r = 0.30 and 0.37 respectively. We have

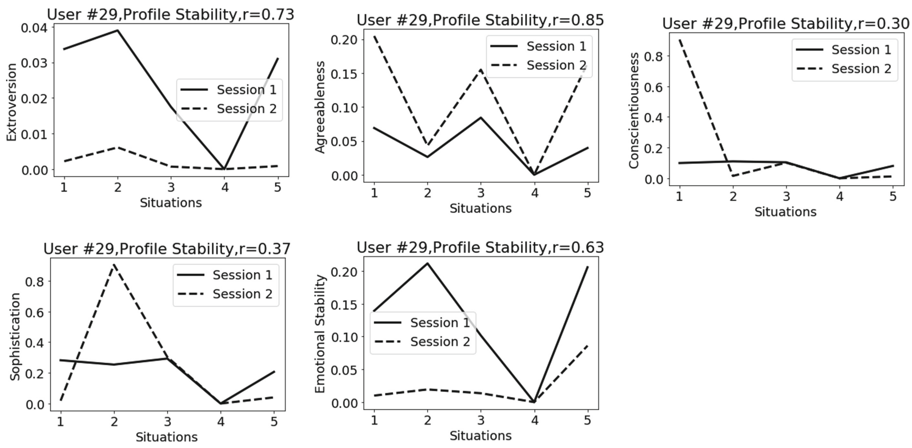

Fig. 10. Situation-behavior Profiles of User #29 at Sessions 1 and 2

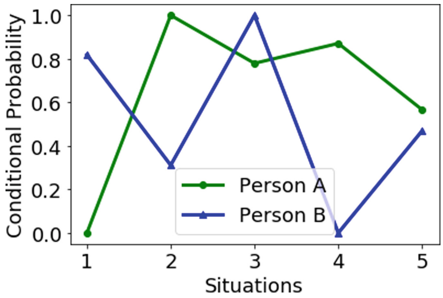

Fig. 11. Intra-individual pattern of Agreeableness of two users plotted across different situations.

analyzed the profile stability of each individual and arrived at the conclusion that the behavior 'Agreeableness' is the one which gives the highest stability across the two time periods and the profile Agreeableness-Situation is the most suitable one to be considered as a behavioral signature that characterizes a person.

Figure 11 shows the intra-individual behavior variability of two persons with the behavior agreeableness plotted over the different situations. The conditional probability of agreeableness is plotted across the situations, since agreeableness was identified as the most intra-individually stable one among the big five, according to our experiments using the profile data.

The Figs. 9, 10 and 11 clearly shows the profile stability and distinguishability property, which are in support of the hypothesis 2. We have examined the stability and distinguishability properties of all the users in our dataset and found complying with hypothesis 2.

5 Conclusion and Future Works

This paper presents a novel approach to social behavioral biometrics by utilizing cognitive psychology in order to develop behavioral signatures of individuals from online social network profiles. To the best of our knowledge, this work is the first to make use of the conceptions of personality dispositions/temperament in a biometric perspective. The notion behind this is the idiosyncrasies and uniqueness in each individual's personality, behind which the cognitive organizations inside a person plays a role. The paper describes how and why the personality and behaviors differ with the help of the Cognitive Affective Personality Systems theory in psychology. We have provided clear descriptions of how the theory can be applied to the online social network with proper mappings and explanations. We have tested two hypotheses related to the stability and distinguishability requirements of biometrics, on a self-compiled Twitter dataset and are proved to be true. The Situation-Behavior profiles of a set of individuals are created and arrived at a conclusion that the profiles are showing considerably good stability and distinguishability to be used as behavioral signatures. Among the five behaviors, the Agreeableness-Situation profile is identified as the most stable and recommended profile.

The annotation of the situation for each individual tweet was a tedious task, which required a lot of manual effort and it is considered as a major limitation of our work. In our work, we have calculated the behavior metrics by content analysis, specifically the text and context of the tweets. In the future, we can utilize a lot of other information like user engagement pattern, temporal data, the network of interactions, etc that may produce better templates/signatures.

Acknowledgements. This research work was funded by the Kerala State Council for Science, Technology and Environment [KSCSTE/5623/2017-FSHP-ES].

References

1. Comscore report. https://wearesocial.com. Accessed 15 Nov 2019
2. Sultana, M., Paul, P.P., Gavrilova, M.: A concept of social behavioral biometrics: motivation, current developments, and future trends. In: 2014 International Conference on Cyberworlds, pp. 271–278. IEEE (2014)
3. Paul, P.P., Gavrilova, M.L., Alhajj, R.: Decision fusion for multimodal biometrics using social network analysis. IEEE Trans. Syst. Man Cybern. Syst. **44**(11), 1522–1533 (2014)
4. Sultana, M., Paul, P.P., Gavrilova, M.L.: User recognition from social behavior in computer-mediated social context. IEEE Trans. Hum. Mach. Syst. **47**(3), 356–367 (2017)
5. Tyshchuk, Y., Wallace, W.A.: Modeling human behavior on social media in response to significant events. IEEE Trans. Comput. Soc. Syst. **5**(2), 444–457 (2018)
6. Ajzen, I.: The theory of planned behavior. Organ. Behav. Hum. Dec. Process. **50**(2), 179–211 (1991)
7. Tripathi, A.K., Hossain, S., Singh, V.K., Atrey, P.K.: Personality prediction with social behavior by analyzing social media data-a survey. University of Winnipeg (2013)
8. Bachrach, Y., Kosinski, M., Graepel, T., Kohli, P., Stillwell, D.: Personality and patterns of Facebook usage. In: Proceedings of the 4th Annual ACM Web Science Conference, pp. 24–32. ACM (2012)
9. Yoon, H.J., Tourassi, G.: Analysis of online social networks to understand information sharing behaviors through social cognitive theory. In: Proceedings of the 2014 Biomedical Sciences and Engineering Conference, pp. 1–4. IEEE (2014)
10. Bandura, A.: Social cognitive theory. In: Annals of Child Development, vol. 6. Six Theories of Child Development (1989)
11. Bandura, A.: Social learning theory of aggression. J. Commun. **28**(3), 12–29 (1978)
12. Bandura, A.: Social cognitive theory of mass communication. In: Media Effects, pp. 110–140 Routledge (2009)
13. Ross, C., Orr, E.S., Sisic, M., Arseneault, J.M., Simmering, M.G., Orr, R.R.: Personality and motivations associated with facebook use. Comput. Hum. Behav. **25**(2), 578–586 (2009)
14. Amichai-Hamburger, Y., Vinitzky, G.: Social network use and personality. Comput. Hum. Behav. **26**(6), 1289–1295 (2010)

15. Wilson, K., Fornasier, S., White, K.M.: Psychological predictors of young Adults' use of social networking sites. Cyberpsychol. Behav. Soc. Network. **13**(2), 173–177 (2010)
16. Quercia, D., Kosinski, M., Stillwell, D., Crowcroft, J.: Our Twitter profiles, our selves: predicting personality with twitter. In: 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing, pp. 180–185. IEEE (2011)
17. Golbeck, J., Robles, C., Edmondson, M., Turner, K.: Predicting personality from Twitter. In: 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing, pp. 149–156. IEEE (2011)
18. Buettner, R.: Predicting user behavior in electronic markets based on personality-mining in large online social networks. Electron. Mark. **27**(3), 247–265 (2016). https://doi.org/10.1007/s12525-016-0228-z
19. Pornsakulvanich, V.: Personality, attitudes, social influences, and social networking site usage predicting online social support. Comput. Hum. Behav. **76**, 255–262 (2017)
20. Tadesse, M.M., Lin, H., Xu, B., Yang, L.: Personality predictions based on user behavior on the facebook social media platform. IEEE Access **6**, 61959–61969 (2018)
21. Thagardz, P.: Changing personalities: towards realistic virtual characters. J. Exp. Theor. Artif. Intell. **17**(3), 221–241 (2005)
22. Quek, M., Moskowitz, D.: Testing neural network models of personality. J. Res. Pers. **41**(3), 700–706 (2007)
23. Cervone, D., et al.: The architecture of personality. Studia Universitatis Babes-Bolyai-Psychologia-Paedagogia **49**(1), 3–44 (2004)
24. Kuhl, J.: A functional-design approach to motivation and self-regulation: the dynamics of personality systems interactions. In: Handbook of self-regulation, pp. 111–169. Elsevier (2000)
25. Mischel, W., Shoda, Y.: A cognitive-affective system theory of personality: reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. Psychol. Rev. **102**(2), 246 (1995)
26. Colneriĉ, N., Demsar, J.: Emotion recognition on Twitter: comparative study and training a unison model. IEEE Trans. Affect. Comput., 1 (2018)
27. Landowska, A.: Towards new mappings between emotion representation models. Appl. Sci. **8**(2), 274 (2018)
28. Mehrabian, A.: Analysis of the big-five personality factors in terms of the pad temperament model. Aust. J. Psychol. **48**(2), 86–92 (1996)