Visiting Patterns and Personality of Foursquare Users

Martin J. Chorley, Gualtiero. B. Colombo, Stuart. M. Allen, Roger. M. Whitaker

School of Computer Science & Informatics

Cardiff University

Cardiff, UK, CF24 3AA

{m.j.chorley, g.b.colombo, stuart.m.allen, r.m.whitaker}@cs.cardiff.ac.uk

Abstract—The study of personality can now provide us with greater insights into human behaviour, when combined with data from social and mobile networks. In this paper we study the relationship between the types of locations that Foursquare users visit and their personality. Using the five-factor personality model, personality profiles were obtained for 173 participating Foursquare users whose activities were analysed, representing a total of 474,539 check-ins across the globe. In this paper we present initial results concerning the number and types of venues and the relationship with personality traits. Particular correlations are identified concerning the openness, conscientiousness and neuroticism of participants and their levels of Foursquare activity. Extracting different categories of venues reveals variation in the correlation strength of personality traits of the users. As far as we are aware this is the first study of its type, linking human personality with spatial behaviour in mobile social networks.

I. INTRODUCTION

Mobile social networks such as Foursquare[1] are now allowing us to represent human behaviour and mobility at a much more granular level than ever before. Over a sustained period of usage, Foursquare users build up a profile of places that they visit which are locally identifiable and also broadly categorised. Accessing this through APIs allows an opportunity to carry out novel research in this area that was not previously possible. In particular it permits us to explore, for the first time, potentially hidden relationships between the human activity and the human disposition. In this paper, we present initial results from an application which combines fine-grained user location data from the mobile social network Foursquare with knowledge of the user's five-factor personality profiles. This application allows us to discover previously unknown relationships between location visiting patterns and personality, providing indicators that can improve our understanding of human behaviour and be used to improve personalisation in future applications. Using personality as a potential signal in future recommendation systems could bring improvements to such systems, basing recommendations for future destinations not just on *where* a user has been previously but also including aspects of *why* they have been there. To the best of our knowledge, this is the first such study examining the relationship between personality and place to a fine grained (venue) level.

A. Five-Factor Personality Profiles

The five-factor personality model [1] is a recognised and established model capable of reliably describing facets of an individual's personality [2] and its use in personality research

has been growing in recent years [3]. The model attempts to describe the human personality in terms of five-factors:

- **Openness** encompasses traits such as originality, curiosity, spontaneity and imaginative. A high Openness score may indicate an artistic nature, with a desire to increase the breadth and depth of ideas, views and experiences encountered. Low openness scores tend to indicate a more conservative or conventional attitude.

- **Conscientiousness** is related to items such as organisation, resourcefulness and perseverance. A high score in Conscientiousness can indicate a focused, organised behaviour, while a low score can indicate a more easy-going, relaxed approach to life.

- **Extraversion** is associated with sociability, assertiveness, being outgoing. High scorers tend to be very social with large groups of people, while low scorers are more likely to be more reserved and introverted.

- **Agreeableness** covers terms such as trusting, cooperative, helpful, altruistic. A high agreeableness score can indicate someone who is able to work well with others, while a low score can indicate someone who is less cooperative and trusting.

- **Neuroticism** relates to items including impulsiveness and emotional stability, also covering negative emotion expression. Neuroticism scores are high in those prone to experiencing stress, or those who can be moody or tense, while low scores indicate someone who is more relaxed and self-confident.

This five-factor model allows us to capture individual personality in an easily computable and comparable form, making it highly suited as an input to current and future personalisation algorithms.

B. Foursquare

Foursquare is a location-based social network, focused on user activity involving places or 'venues'. A large crowd-sourced database of venues is maintained by the service, and data about these venues is exposed to users through the website and mobile applications. While visiting a location, users of the service can choose to 'check in' to a venue to record their visit in their own personal history and alert their friends to their whereabouts. Users are motivated to continue checking in to win points, badges and mayorships of venues, so competing with their friends and other users (so-called 'gamification' of the service). Data from the service is exposed through a REST

---

[1] http://www.foursquare.com

based API, allowing application developers to access details of venues and, with the correct authentication, user activity.

C. Foursquare Personality Experiment

We have created an online experiment which administers the Big-Five Inventory personality test [3], [4] to visitors while simultaneously accessing their location visiting profile from Foursquare. Once they have completed the personality test, users are able to examine how their personality profile relates to the average personality profile of all other users at the venues they have visited. We are able to 'tag' each venue a user has visited with their personality profile, allowing us to analyse the geographic distribution of personality profiles to a fine-grained level.

II. RELATED WORK

The MyPersonality[2] experiment is a large scale research project using Facebook as a test bed for conducting personality research. Through a Facebook application, personality tests have been administered to a large group of users and their data collected along with features of their Facebook profiles. Much analyis has been focused on this data. In [5] it is found that Extraversion scores are correlated with the size of a users social network and a model is described that can predict the number of friends using just Age and Extraversion scores. In [6] the Facebook activity (number of friends, number of likes, number of photo uploads etc) of users is correlated with their personality scores, and it is found that it is possible to predict some personality traits from Facebook profile features. Meanwhile, [7] covers similar work studying Facebook and personality, with a far smaller sample size, but still finding a correlation between Extraversion and social network size, while then going on to analyse correlations between personality and not just Facebook profile features, but also the linguistic features of language used on Facebook, finding a number of correlations. These features are then also used to predict personality profiles from the data available in a Facebook profile.

Similar work has been carried out on Twitter. [8] examines the network structure of Twitter users, classifying users based on their number of followers, the number of people they are following, how many lists they appear on, or their influence (measured using the analytics service 'Klout'). Correlations are then found between these clusters and personality profiles, and these correlations allow personality profiles to be predicted based on this network structure. A similar study also looks at Twitter users [9], this time analysing the linguistic features of users tweets, finding many correlations between these features and personality profiles. Again, these features are used to predict personality successfully.

A body of work exists looking at the modelling and categorising of human mobility and visiting patterns. Human movement shows a high degree of temporal and spatial regularity, with each individual presenting a characteristic travel distance and a significant probability to return to a restricted number of highly frequented locations [10], [11], [12]. It has also been shown that it is possible to analyse mobility information for similarity and statistical characteristics that can be used to classify users according to their movements [13], [14] and predict aspects of their social or mobile interactions [15].

Previous online experiments have seen high dropout rates compared to lab based studies [16]. In this work, 88% of participants who started participating in the experiment completed the personality test, although we have no measure of the number of people who accessed the website but did not decide to take part. [17] examines the quality of data received from internet based surveys, finding that common preconceptions about the validity of such data are not necessarily true, and those that do present issues can be mitigated by researchers. Indeed, they find that the results of internet based surveys are at "of least as good quality as those provided by traditional paper-and-pencil methods" [17]. Similarly, [18] examines the data provided by internet based personality research, and also concludes that such data can be entirely valid.

III. HYPOTHESIS

It has been seen in previous work that correlations exist between activity on and use of various social networks and five-factor personality profiles. In this experiment we therefore seek to answer the following questions related to mobile social networks:

- Is there a relationship between the number of venues a user visits and their personality profile?

- Is there a relationship between the number of checkins a user makes and their personality profile?

- Which of the five personality factors most closely correlate to user activity on Foursquare?

- Do these relationships exist at a finer grained level, when venues are grouped by category features?

Of the five-factors, it is suggested that Openness would be most closely related to the use of Foursquare. Users with a high Openness score are more likely to seek out new experiences, and to try many different things. It could reasonably be assumed that they would therefore have checked in to a higher number of venues overall. Similarly, users high in Extraversion are likely to be more sociable and outgoing, and it is reasonable to assume they would therefore have higher numbers of checkins in venues associated with social situations (bars, clubs, cafes etc).

IV. DATA COLLECTION EXPERIMENT

Data used in this paper was collected through the 'Foursquare Personality Experiment'[3], a web-based application designed and built by the authors. Upon accessing the website users are asked to login using their Foursquare account (utilising the OAuth protocol so that personal login information is not revealed to the experiment). Once the application is authorised, the user is directed to a 44-item Five Factor model personality test [3], [4] which they are instructed to complete, by selecting how well they believe a statement matches their characteristics, for each of 44 statements. While the test is administered the application accesses the list of places the user has visited and checked in to on Foursquare. Once the
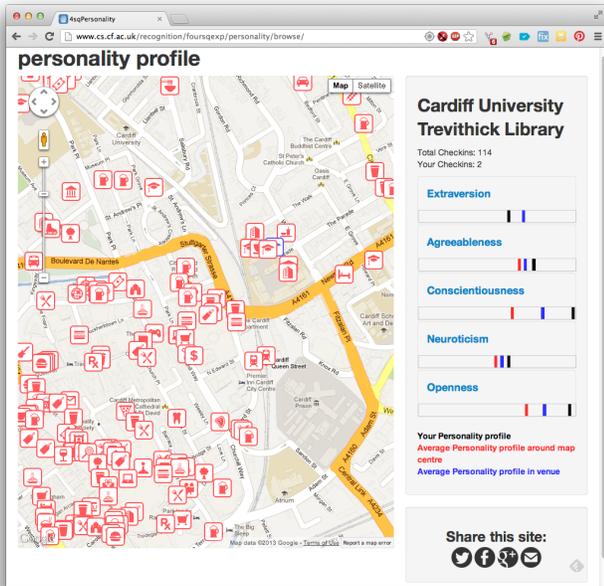
---

[2]http://mypersonality.org

[3]http://www.cs.cf.ac.uk/recognition/foursqexp

Fig. 1: Foursquare Experiment Results Page

Fig. 2: Locations of Foursquare venues visited by participants in the study

test has been completed, the user is presented with their personality profile, along with a map-based interface showing the venues they have previously checked in to. Selecting a venue on the map reveals the average personality recorded in that venue, allowing the user to make a comparison between their personality profile and the average of the other people who also visit the venue (see Figure 1). In this way, the users derive some utility from the experiment beyond just their scientific contribution.

The data accessed from a user's Foursquare profile is limited to that available from the `/users/venuehistory` API endpoint, which provides a list of the venues that a user has visited and the number of times they have been there. Individual checkin histories were not accessed, so although we know where a user has been and how many times, we do not know exactly *when* a user has visited each venue, thus no examination of checkin frequency is possible. This reduces the privacy risk for the user in taking part in the experiment. Although we do access user profile information initially, this is not stored. As such, demographic information is not available about the users of the experiment. The demographic information captured by Foursquare itself is limited, and would therefore be of little use to the experiment. Multiple studies looking at demographics and personality of social network users have already been carried out [5], [6], hence a demographic analysis of users would add little that is not already known.

V. RESULTS

The Foursquare Personality Experiment was launched on November 19th 2012. The results in this section are based on an analysis of the first 194 participants in the study.
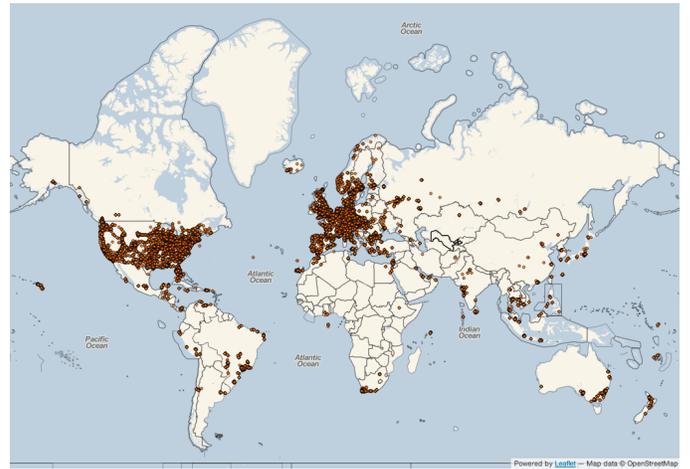
A. Participant Discussion

The users participating in the study are a self-selecting group comprised of people who both use Foursquare and are willing to take part in a personality-based experiment. The group is necessarily restricted to those with internet access, and also to those who own or have access to smartphones, as the Foursquare application is predominantly smartphone based, particularly with regards to checking in at locations.

The experiment was promoted online through the social networks Twitter and Facebook, and also through mentions on websites whose focus is news about Foursquare. As a result, the people likely to have heard about the experiment (and therefore the pool from which the participants are drawn) are likely to be tech-savvy users of social network sites.

The location of all venues visited by participants in the study is shown in Figure 2. As can be seen, the venues are spread globally, although with a particular focus on the United States and Western Europe.

It is clear that the population of users of the experiment is not likely to be representative of the global population as a whole. Results from this group may not therefore be applicable to the general population.

B. Dataset Statistics

The application was logged into by 194 Foursquare users in the period up to March 2013. Of these, 173 users completed the personality test and submitted their answers. The average number of venues visited by a user was approximately 707, while the average number of checkins per user was 2743.

TABLE I: Average Personality Profile

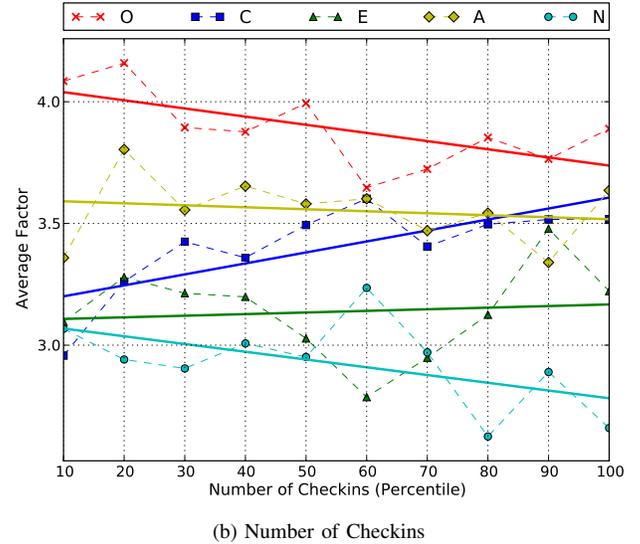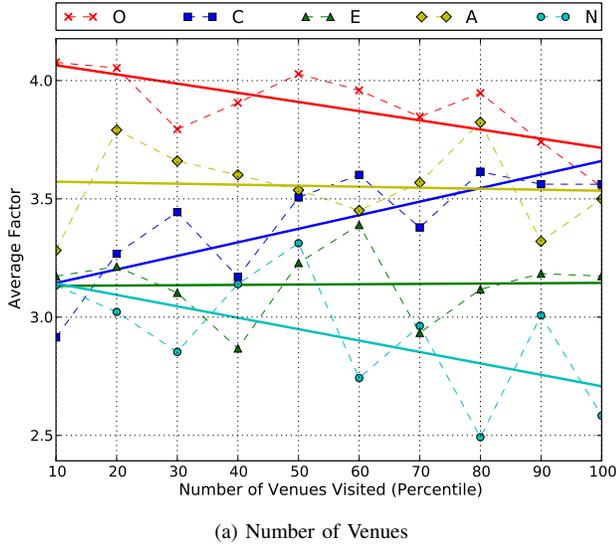| Factor	Average	St.Dev
Openness	3.880347	0.614155
Conscientiousness	3.423250	0.657760
Extraversion	3.137283	0.837681
Agreeableness	3.547848	0.644414
Neuroticism	2.907514	0.733519

(a) Number of Venues

(b) Number of Checkins

Fig. 3: Relationship between personality factors and Number of unique Venues visited, and between personality factors and Number of Checkins

The average personality profile recorded by the application is given in Table I. These values are comparable to other published personality profile norms [19], [7], presenting similar averages and standard deviations. While this is indicative that there is nothing necessarily drastically different between Foursquare users and other populations studied in relation to the five-factor model, it does not indicate that our results are applicable to the general population, for the reasons discussed above.

C. Correlation with Location Data

Pearson correlation tests have been performed on features of the users location visit data, examining the relationship between the number of unique venues visited per user, the number of checkins made, and their five-factor personality profiles.

1) Number of Unique Venues: Pearson correlation coefficients for the number of unique venues visited are given in Table II. Statistically significant values ($p < 0.05$) are shown in bold.

TABLE II: Number of Venues Visited Correlation Factors

Factor	Pearson Coefficient (r)	p
Openness	-0.148526	0.051149
Conscientiousness	**0.186298**	0.014123
Extraversion	-0.006376	0.933647
Agreeableness	-0.006241	0.935054
Neuroticism	**-0.175086**	0.021221

There is a weak but statistically significant positive correlation between Conscientiousness and the number of venues visited ($r = 0.186298$, $p = 0.014123$). There is also a weak negative correlation (again statistically significant) between Neuroticism and the number of venues visited ($r = -0.175086$, $p = 0.021221$). This suggests that the higher the number of venues visited, the higher the Conscientiousness

TABLE III: Number of Checkins Correlation Factors

Factor	Pearson Coefficient (r)	p
Openness	-0.122239	0.109120
Conscientiousness	**0.175710**	0.020757
Extraversion	0.007482	0.922175
Agreeableness	-0.025716	0.736989
Neuroticism	-0.083147	0.276781

of the user is likely to be, and the lower the Neuroticism. The relationship between Conscientiousness and number of venues may be due to the fact that in order for someone to have checked in at a larger number of venues they need to be an organised person, used to routinely checking in whenever they visit somewhere. A less organised and focused person may be less likely to remember to check in, or to care about recording their visit somewhere new. The negative correlation with Neuroticism meanwhile may also be expected - a more relaxed personality could be more willing to share location information as they visit places.

It is interesting that there is a weak negative correlation between Openness and the Number of Venues visited ($r = -0.148526$). Although this relationship is not statistically significant ($p = 0.051149$), it is opposite to what may be expected - users with a higher level of Openness would be more expected to seek out new experiences, and should therefore have visited more unique places than those users with a lower Openness score. This does not appear to be the case however, those with a higher Openness score seem to have visited *fewer* unique venues.

2) Number of Checkins: Pearson correlation coefficients for the relationships between the number of checkins and five-factor personality elements are given in Table III. Again, statistically significant ($p < 0.05$) coefficients are presented in bold type.

With this data related to the number of checkins, the

TABLE IV: Number of Unique Venues visited in each root category against Personality factors

Category	Openness		Conscientiousness		Extraversion		Agreeableness		Neuroticism	
	r	p	r	p	r	p	r	p	r	p
Professional & Other Places	-0.110306	0.148529	0.13695	0.072381	0.012519	0.870145	0.035852	0.63958	-0.139589	0.067
Shops & Services	**-0.188371**	0.013067	0.147942	0.05208	-0.056881	0.457279	0.009698	0.899232	-0.106119	0.164656
Residences	-0.075845	0.321305	0.092338	0.226935	0.028077	0.713848	-0.010738	0.888485	-0.022386	0.770024
Food	-0.110188	0.148968	**0.19164**	0.011543	-0.003756	0.960885	-0.030102	0.694213	-0.136136	0.074109
Nightlife Spots	-0.088245	0.248289	**0.187111**	0.0137	0.106631	0.162614	-0.01297	0.865507	-0.133626	0.079653
Travel & Transport	-0.132521	0.082197	0.09578	0.210008	-0.029837	0.696771	-0.047004	0.539146	**-0.180142**	0.017709
Arts & Entertainment	-0.138031	0.070137	0.138534	0.06911	-0.00652	0.932155	0.031503	0.680741	**-0.212213**	0.005063
Colleges & Universities	-0.005717	0.940497	**0.229986**	0.002335	-0.063614	0.405698	-0.011591	0.879691	0.01531	0.84154
Outdoors & Recreation	-0.085075	0.265755	0.118977	0.118973	0.038028	0.619373	0.028142	0.713216	**-0.200974**	0.008017

correlation coefficients are weaker than those seen when examining the number of unique venues visited, and only the relationship with Conscientiousness remains statistically significant. This could be an important factor when considering the relationship between a user's personality and their activity on Foursquare. There is a clear relationship between the length of time an individual has been using Foursquare and the total number of checkins they have made. Assuming they continue to use the service consistently, the total number of checkins made will steadily increase as time goes on. However, the number of unique venues visited may not change as time increases. Once a user has checked in to all the venues they regularly go to, or are willing to reveal publicly that they go to, the number of venues may remain constant (or close to constant), if their individual characteristics are such that they do not regularly visit new places or areas. Conversely, if their individual personality is such that they undertake a lot of travel or exploration of new places, the unique number of venues may continue to climb as time increases.

This indicates that one of the main factors influencing the number of checkins a user makes in total is more likely to be the length of time they have been using the service, while the number of unique venues visited is quite possibly related both to the length of time using the service combined with the user's individual personality characteristics. This could explain why the correlations with the number of checkins are slightly weaker and less significant than those with the number of venues. It may therefore be more important to consider the unique number of venues visited rather than the number of checkins when relating Foursquare activity to personality.

Despite this, a significant correlation *is* present between Conscientiousness and the number of checkins ($r = 0.175710$, $p = 0.020757$), again hinting that the relationship between this personality factor and the use of Foursquare is related to the individual remembering to check in at multiple venues multiple times as part of a usual routine.

The data for both the number of venues and the number of checkins are represented graphically in Figure 3. Here the number of venues (Figure 3a) and the number of checkins (Figure 3b) are presented in percentiles against the average of each personality factor for that percentile (so each point in the plot represents approximately the same number of personality profiles). The correlations discussed above can clearly be seen in these plots.

D. Top Level Categories

Foursquare venues can be tagged with one or more 'categories' from a finite list provided by Foursquare themselves. These categories are organised as a hierarchy, with nine root categories under which all other categories sit. By separating venues into groups based on these root categories, we can examine the data again to see if any correlations exist between personality factors and the number of venues visited within each root category. This data is shown in Table IV. As can be seen there appear to be many correlations between the number of venues visited and individual personality factors, however these are not all statistically significant ($p < 0.05$). Of note is that all categories again show a weak positive correlation with Conscientiousness, although this is only statistically significant for Food, Nightlife Spots and Colleges & Universities. Besides this already observed relationship, statistically significant negative correlations are seen between Shops & Services and Openness and between Neuroticism and Travel & Transport, Arts & Entertainment and Outdoors & Recreation.

Again, surprisingly, no positive correlation (significant or otherwise) is seen between Openness and any of the root categories. It might have been suspected that of all the categories, it was most likely that people with a high score in Openness may have been to more venues in the Arts & Entertainment categories, as an expression of their artistic nature, but that is not seen to be the case.

There is a weak correlation seen between Extraversion and the number of Nightlife Spots visited ($r = 0.106631$) though this is not statistically significant ($p = 0.162164$). It could be expected that Extraverted people would be more sociable and outgoing and so visit more venues in this category, so it would be expected that this correlation would be stronger, and more significant.

VI. CONCLUSION

We have created an online personality experiment which examines the relationship between Five-Factor personality factors and the number and types of venues visited by Foursquare users. From 173 users, making 474,579 checkins at venues across the globe, we have found correlations between five-factor personality traits and Foursquare checkin behaviour.

There is a positive correlation between Conscientiousness and the number of venues visited, possibly due to the organised routine required to consistently check in at venues. There is also a negative correlation between Neuroticism and the number of venues visited. Although, as might be expected, a positive correlation exists between Openness and the number of venues visited, this is not statistically significant.

Correlations between the number of checkins and five-factor personality traits are weaker and less significant, pos-

sibly due to the number of checkins being less related to individual personality and being more related to the length of time using the service.

Breaking the venues visited down into groups based on their assigned category reveals further correlations with five-factor personality traits.

VII. FUTURE WORK

Much future work can be carried out looking further at the data collected by this experiment. The existence of correlation between five-factor personality profiles and aspects of Foursquare users visiting patterns indicates that it may be possible to predict an individuals' personality from their mobility pattern and/or interaction with Foursquare. Similarly it may be possible to predict the places a user is interested in based on their personality. Using further venue attributes (not just top-level categories) may allow for stronger correlations to be discovered, to allow these predictive models to be created.

These findings could have many uses in the personalisation of future content and place recommendation systems, as once personality can be predicted from a user's visited places a personality test is no longer required in order to use personality as a signal to improve recommendations for future destinations.

Results from this work may be generalised or improved for use in future systems by examining further interactions between people and places, rather than focusing solely on checkins on Foursquare. For instance, a larger population of users participate in online discussion of venues using rating web-sites such as TripAdvisor, Yelp, or Google. These reviews or ratings could be used as a surrogate for checkins in a similar analysis. Further demographic information on users may be available through such websites, allowing conclusions about the applicability of results to the general population to be made.

REFERENCES

[1] L. R. Goldberg, "An Alternative "Description of Personality": the Big-Five Factor Structure." *Journal of Personality and Social Psychology*, vol. 59, no. 6, pp. 1216–29, Dec. 1990.

[2] R. R. McCrae and O. P. John, "An introduction to the five-factor model and its applications." *Journal of Personality*, vol. 60, no. 2, pp. 175–215, Jun. 1992.

[3] O. P. John, L. P. Naumann, and C. J. Soto, "Paradigm shift to the integrative big five trait taxonomy," *Handbook of personality: Theory and research*, vol. 3, pp. 114–158, 2008.

[4] O. P. John, E. M. Donahue, and R. L. Kentle, "The big five inventoryversions 4a and 54," *Berkeley: University of California, Berkeley, Institute of Personality and Social Research*, 1991.

[5] D. Quercia, R. Lambiotte, and D. Stillwell, "The personality of popular facebook users," in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. New York, New York, USA: ACM Press, 2012, pp. 955–964.

[6] 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*, pp. 24–32, 2012.

[7] J. Golbeck, C. Robles, and K. Turner, "Predicting personality with social media," in *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems - CHI EA '11*. New York, New York, USA: ACM Press, 2011, p. 253.

[8] D. Quercia, M. Kosinski, D. Stillwell, and J. Crowcroft, "Our Twitter Profiles, Our Selves: Predicting Personality with Twitter," in *2011 IEEE Third International Conference on Social Computing*, 2011, pp. 180–185.

[9] J. Golbeck, C. Robles, M. Edmondson, and K. Turner, "Predicting Personality from Twitter," in *2011 IEEE Third International Conference on Social Computing*. Ieee, Oct. 2011, pp. 149–156.

[10] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, Jun. 2008.

[11] M. Zignani and S. Gaito, "Extracting human mobility patterns from gps-based traces," in *Proceedings of the the 2nd. IFIP Wireless Days 2009*. Paris, France: IEEE, December 2009.

[12] M. Williams, R. Whitaker, and S. Allen, "Decentralised detection of periodic encounter communities in opportunistic networks," *Ad Hoc Networks*, vol. 10, no. 8, pp. 1544–1556, 2012.

[13] G. Colombo, M. Chorley, M. Williams, S. Allen, and R. Whitaker, "You are where you eat: Foursquare checkins as indicators of human mobility and behaviour," in *2012 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 2012, pp. 217–222.

[14] A. Noulas, S. Scellato, C. Mascolo, and M. Pontil, "Measuring user activity on an online location-based social network," in *Proceedings of NetSciCom 2011*, Shanghai, China, 15, April, 2011.

[15] J. Cranshaw, E. Toch, J. Hong, A. Kittur, and N. Sadeh, "Bridging the gap between physical location and online social networks," in *Proceedings of Ubicomp '10*. New York, NY, USA: ACM, 2010, pp. 119–128.

[16] U. Reips, "Internet-based psychological experimenting: five dos and five don'ts," *Social Science Computer Review*, vol. 20, no. 3, pp. 241–249, Sep. 2002.

[17] S. D. Gosling, S. Vazire, S. Srivastava, and O. P. John, "Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires." *American Psychologist*, vol. 59, no. 2, pp. 93–104, 2004.

[18] T. Buchanan and J. L. Smith, "Using the Internet for psychological research: personality testing on the World Wide Web." *British journal of psychology (London, England : 1953)*, vol. 90 (Pt 1), pp. 125–44, Feb. 1999.

[19] S. Srivastava, O. P. John, S. D. Gosling, J. Potter *et al.*, "Development of personality in early and middle adulthood: set like plaster or persistent change?" *Journal of personality and social psychology*, vol. 84, no. 5, pp. 1041–1053, 2003.