An Analysis of the Relationship between Human Personality and Favored Location

Ha Yoon Song, Eun Byul Lee

Department of Computer Engineering,
Hongik University, Seoul, Korea
Email: hayoon@hongik.ac.kr, quftmxk@gmail.com

***Abstract*—It is a long-held belief that human personality affects the preferred locations of humans. We used statistical regression analysis in this research to examine the relationship between human personality and favored locations. The personality data is represented in accordance with the Big Five Factor (BFF) and location data was acquired from smartphone apps used by the volunteers. Instead of using raw location data, a total of 11,154 location datum was collected and then subject to further, more detailed categorization. The categorized location data were then classified into a higher-level aggregation. Finally, the relationship and the statistical results from regression analysis are presented, alongside categorized and classified location data. A summary of the relationship is provided, whereas a Coefficient of Determinant (CoD) is used for the results of the regression analysis.**

Keywords–Personality; Favorite location; Regression analysis, Personality-location relationship, Big Five Factors.

I. INTRODUCTION

The actions of humans can be predicted by time, personality factors, occupation, gender, and so on. Among these factors, human personality factors are typically fundamental in the prediction of human behavior, especially location preference. In this research, the effect that personality factors exert on a human's preferred location was examined. The five-factor model (FFM), a representative theoretical framework for understanding human personality, is comprised of the Big Five Factor (BFF), which can be obtained by using the Big Five Inventory (BFI) questionnaire set [1]. For this research, the BFFs of the 14 volunteer participants were obtained anonymously.

Nowadays, smartphones can be used to easily collect location or position data. Numerous smartphone functions including Global Positioning System (GPS), GLONASS, and indoor positioning systems are used for positioning, and humans can use smartphone apps to easily check-in wherever they are located. The most popular location apps are Foursquare and Swarm. This location data can be categorized into higher-level categories using the categories-id; for example, Vietnamese restaurant and Japanese restaurant can be categorized in the higher-level category of Restaurant, while Coffee house and Milk tea shop can be categorized in Beverage store. These categories can be identified by the categories-id. However, for the purpose of detailed research, lower-level categories are also used in this paper.

Regression analysis [2] was a helpful tool for examining the relationship between BFF and the categorized locations. Collected data sets were formed as variables for the regression analysis. In this research, an examination of the relationship between human personality and location preference was conducted. Apart from other research, regression analysis on personality data and location data were utilized. Mobile services which are strongly dependent on location can utilize the benefit of this research. For example, the result of this research will be applied for recommendation system which recommends a specific travelers attraction to a person with certain personality. A possible scenario is that a travel recommendation system based on our approach, which can match the traveler's personality with a property of travelers' attractions and can recommend a specific travelers attraction which fits the traveler's personality.

This paper is organized as follows. Section II will discuss about related works on location and personality. In Section III, personality data, location data, and the categories of locations are described as parts of the preparation for the regression analysis. Section IV will analyze the results of the regression analysis and will present examples of the way in which the BFF affects locations. Section V will present the conclusion and possible future research topics.

II. RELATED WORKS

There has been little research into how our personality relates to the categories of place that we visit. There are mainly two reasons. The first one is related with personal information protection, meaning peoples do not like to totally reveal personal information and only volunteers help to collect personality data without any legal problem. The second one is that the use of mobile device is not general at all, even though the smartphones are widely spread. This restricts the collection of location data. Therefore, only volunteers with mobile devices can provide their personality data and location data. Despite of these barriers of data collection, several related researches have found in last few years.

The relationship between social network and human personality are conducted in [3]. The effect of human personality on Facebook usage is researched. Self-report of participants are used to extract object criteria and measurements from Facebook data, and the result shows that strong relationship between human personality and actions on Facebook.

Location Based Social Networks (LBSN) are being started in this research area nowadays. One of the research [4] is mostly related to our research. Personality factors and foursquare check-ins are investigated and utilized, and the relationship is founded. Conscientiousness, Openness, and Neuroticism are found related to specific locations in this research.

 ISBN: 978-1-61208-428-2

As a result of this research, the combination of personality factors and LBSN can explain the human personality.

Regression analysis on personality data and location data can be rarely found as shown in [5] in order to predict the future human location with five volunteers' data and Back Propagation Network (BPN). An opposite of this research can be found in [6] which deduces personality factors form location data of volunteers. BFFs can be deduced from sets of location data also by BPN. Apart from the other researches, our research is a direct match between BFF (personality) and location categories. Personality data and location data collected by volunteers themselves are used with minimal treatment of the raw data.

In the next sections we will present our method to reveal the relationship between personality and location preference.

III. Personality and Location Data Preparation

Two different assortments of location data will be addressed in this section. The first method is a classification of location data. From the location data source, several information such as frequency of visit, name of location, category-id will be extracted. The second method is the categorization of location data. From the location data source, several information such as frequency of visit, name of location, name of location categories will be extracted. Note that categorization of location data is a more ramification of location data comparing to classification of location data.

Total fourteen volunteers provided their personality and location information. All volunteers are anonymous for personal information protection and called by their numbers. Most of the volunteers are university students and are in their twenties. University students tend to go out lively and visits university, restaurant, and other places frequently. BFF of each volunteer is identified by BFI. Due to the nature of volunteers that they are university students, places in university campus are taken special care and categorized as detail as possible.

TABLE I. BFF OF VOLUNTEERS

	O	*C*	*E*	*A*	*N*
Volunteer 1	3.30000	3.88889	3.25000	3.66667	2.62500
Volunteer 2	3.60000	3.33333	2.75000	3.22222	2.75000
Volunteer 3	2.70000	3.22222	3.25000	2.66667	2.75000
Volunteer 4	4.33333	3.12500	2.25000	3.20000	2.88889
Volunteer 5	4.20000	4.33333	3.50000	3.55556	2.62500
Volunteer 6	4.00000	3.66667	4.00000	3.88889	2.75000
Volunteer 7	3.50000	3.77778	3.37500	3.22222	3.00000
Volunteer 8	2.20000	3.44444	3.00000	3.11111	2.62500
Volunteer 9	2.60000	2.77778	3.37500	3.11111	2.62500
Volunteer 10	3.30000	2.88889	3.12500	3.11111	3.25000
Volunteer 11	3.40000	3.22222	3.37500	3.33333	3.12500
Volunteer 12	3.10000	3.66667	3.37500	3.22222	3.50000
Volunteer 13	3.40000	3.55556	3.62500	2.88889	2.50000
Volunteer 14	3.80000	4.00000	3.12500	3.77778	2.25000

A. *Personality data*

Research has been conducted on the various theories that seek to represent human personality. Among these theories, McCrae and Costa's [7] research on BFF is one of the foremost studies on personality representation, whereby the five factors are Openness (O), Conscientiousness (C), Extraversion (E), Agreeableness (A), and Neuroticism (N). This distinguished Five Factor Model (FFM) is composed with BFF which can be obtained by Big Five Inventory (BFI) [8]. In order to obtain BFF of volunteers, volunteers answer the questionnaire set called BFI which is usually composed of 44 questions and the answers are selections of a degree from one to five. The BFI is designed to calculate every five factors of BFF from a combination of degrees of the 44 answers. The degree of each five factor is designed to be represented from one to five. A major merit of BFF is that the factors can be quantized and it reflects the concept of normal distribution, i.e., each quantitative factor is in a numerical form that stands for positions of normal distribution. Due to the properties of BFF, corresponding numerical data can be utilized by any algorithm and, of course, it can be applied to regression analysis as shown in this research. It is notable that each factor of BFF is mutually orthogonal, which stands that factors do not interfere each other and thus can be used as independent variables in regression analysis.

TABLE II. SAMPLE LOCATION DATA OF VOLUNTEER 1

CoV	*Location name*	*categories-id*
75	Hongik Univ.T-819	4bf58dd8d48988d19e941735
4	Starbucks	4bf58dd8d48988d1e0931735
3	CGV	4bf58dd8d48988d180941735
2	Hongik Univ.	4bf58dd8d48988d1a8941735
1	Everland	4bf58dd8d48988d182941735

Table I shows BFFs for 14 volunteers, and the values can be used to easily interpret the personality of a human. For example, in Table I, Volunteer 4 has the highest Openness rating among the 14 volunteers, implying that this volunteer is highly creative in the arts, highly intellectual, curious, and adventurous. The other BFF ratings of volunteer 4 imply planning skills, propensity for action, and diligence in regard to work. Volunteer 6 has the highest Extraversion rating, implying an energetic, verbose, and positive-minded personality.

The above BFF of Volunteers data set will be used for this research henceforth.

TABLE III. SAMPLE CLASSIFICATION OF LOCATIONS OF VOLUNTEER 1

categories-id	*Name*	*ICC*
4bf58dd8d48988d16d941735	Beverage Store	5622
4bf58dd8d48988d110941735	General Restaurant	5611
4bf58dd8d48988d19e941735	Institutions of Education	8530
4bf58dd8d48988d1e2931735	Museum and Historic Site Managing	9022
4bf58dd8d48988d1d0941735	Other Restaurant	5619
4bf58dd8d48988d182941735	Theme Park Operations	9121

B. *Classification of location data*

Location data, as different from positioning data, represent conceptual locations such as home, restaurant, school, and so on. Positioning data is typically numerical data, including latitudinal and longitudinal pairs. The location data in this research were collected using the smartphone apps named Foursquare and Swarm. Volunteers checked-in on their preferred app whenever they visited a meaningful location. The app then collated the location data based on the source data provided by the volunteers, including count of visit (CoV), the location's name, and the categories-id. Table II shows a sample of location data from volunteer 1 as an example of the data representation of this study.

In Table II, both the formal and numerical forms of the location information are presented; for example, two visits

TABLE IV. COUNT OF VISIT (CoV) TO THE LOCATIONS AND ICC OF VOLUNTEER 1

ICC	*8530*	*5611*	*5619*	*5291*	*5622*	*5621*	*5914*	*9900*	*9022*	*4712*	*9112*	*9121*
CoV	30	22	13	6	6	5	3	3	2	1	1	1

TABLE V. VOLUNTEERS' LOCATION DATA

Volunteer	*FI*	*RB*	*TI*	*Restaurant*	*Bar*	*BS*	*Bank*	*PAI*
1	0.15008	0.04384	0.04384	0.27487	0.02529	0.10118	0.00337	0.00000
2	0.16484	0.06154	0.01319	0.32527	0.02637	0.09890	0.00439	0.00219
3	0.03030	0.09848	0.05303	0.63636	0.02272	0.02272	0.00000	0.00000
4	0.29339	0.01222	0.02119	0.18744	0.00570	0.16299	0.00081	0.00081
5	0.30158	0.00000	0.02116	0.19047	0.01587	0.04232	0.00000	0.00000
6	0.12931	0.02586	0.02874	0.28735	0.02298	0.10344	0.00000	0.00000
7	0.13500	0.00500	0.07000	0.18500	0.01500	0.01000	0.00000	0.00000
8	0.27003	0.04748	0.01187	0.16617	0.01186	0.04154	0.00593	0.00593
9	0.24915	0.20478	0.02048	0.15017	0.00682	0.07508	0.01023	0.00341
10	0.18143	0.02109	0.08439	0.22784	0.03375	0.10126	0.00000	0.00000
11	0.26042	0.01042	0.01042	0.22916	0.03125	0.03819	0.00000	0.00000
12	0.20952	0.11429	0.04286	0.09523	0.00476	0.04285	0.00952	0.00476
13	0.18705	0.03597	0.18705	0.17266	0.01438	0.01438	0.00000	0.02877
14	0.03904	0.14571	0.28619	0.15262	0.04303	0.02951	0.00399	0.00860
Volunteer	*LS*	*IE*	*Hospital*	*CH*	*LM*	*SF*	*TP*	*BB*
1	0.00337	0.30860	0.00337	0.01349	0.01855	0.00505	0.00505	0.00000
2	0.00000	0.27033	0.00879	0.00879	0.01098	0.00219	0.00219	0.00000
3	0.00000	0.09848	0.02272	0.01515	0.00000	0.00000	0.00000	0.00000
4	0.00000	0.28280	0.00244	0.00733	0.00570	0.00733	0.00489	0.00489
5	0.00000	0.39153	0.00000	0.01587	0.01587	0.00529	0.00000	0.00000
6	0.00287	0.37931	0.00862	0.00000	0.00862	0.00287	0.00000	0.00000
7	0.00000	0.57000	0.01000	0.00000	0.00000	0.00000	0.00000	0.00000
8	0.00000	0.41246	0.00296	0.00890	0.00890	0.00000	0.00593	0.00000
9	0.00000	0.19453	0.01706	0.00682	0.01706	0.03413	0.00000	0.01023
10	0.00000	0.23628	0.00843	0.01687	0.02109	0.06751	0.00000	0.00000
11	0.00000	0.29513	0.01736	0.05556	0.00347	0.03819	0.01041	0.00000
12	0.00000	0.42381	0.00476	0.02857	0.01428	0.00000	0.00476	0.00000
13	0.00000	0.28777	0.00719	0.01438	0.03597	0.00000	0.00719	0.00719
14	0.00292	0.07254	0.06947	0.04980	0.04857	0.02797	0.01644	0.00353

were made to a location named Hongik Univ.T-819, which has a predefined categories-id of 4bf58dd8d48988dla8941735.

According to the Standard Industrial Classification, higher categories can be created by merging subcategories using the categories-id; for example, Vietnamese restaurant and Japanese restaurant can be merged into the higher category of Restaurant. In another case, Coffee house and Milk tea shop can be merged into Beverage Store, which is a higher level category also classified in the Standard Industrial Classification. Using Standard Industrial Classification prepared by the National Tax Service and the Industry Classification Code (ICC), locations can be classified into higher-level ICC categories. Table III shows several samples of the location classifications using the data from Volunteer 1. Table III shows the categories-id, the names of the classified categories, and the ICC; for example, Hongik Univ. T-819's categories-id is 4bf58dd8d48988d19e941735, and after it was merged into Institutions of Education, a higher classification, the corresponding ICC became 8530. Similarly, categories-id 4bf58dd8d48988d182941735 was merged into Theme Park Operations with an ICC of 9121.

Table IV shows extracts of CoV values to each ICC; for example, one visit was made to the ICC location of 4712 and 22 visits were made to the ICC location of 5611. Table II, Table III, and Table IV show the sequence of the location-classification process for Volunteer 1.

It was necessary to arrange all of the data from the 14 volunteers in the same way. Table V summarizes the classification results of every volunteer. The meanings of the location-classification acronyms in Table V are:

- FI = Foreign Institutions
- RB = Retail Business
- TI = Travel Industry
- BS = Beverage Store
- PAI = Public Administration Institutions
- LS = Leasing Service
- IE = Institutions of Education
- CH = Concert Hall
- LM = Library, Museum
- SF = Sports Facilities
- TP = Theme Park
- BB = Beauty Business

Every number in Table V represents the ratio of visits according to (1), as follows:

$$VisitingRatio(loc) = \frac{count\_of\_visit\_to(loc)}{total\_count\_of\_visit} \quad (1)$$

where *loc* is one of the locations in Table V.

For example, the CoV of Volunteer 1 to Foreign Institute is 89, while the total CoV is 593; therefore, visiting ratio (FI) = 0.15008. As shown in Table V, Volunteer 1 has a visiting ratio of 0.15008 for Foreign Institute. In the case of Concert Hall, Table V shows zero values for Volunteer 6 and Volunteer 7, while other volunteers have a non-zero visiting ratio, meaning that more than one visit was made.

 ISBN: 978-1-61208-428-2

C. Categorization of location data

From the source check-in data, each location category was parsed, extracting the CoV, the location name, and the categories-name. Table VI shows several of the extracted samples, including one example where 58 visits were made to a location named Hongik Univ.T, which is categorized as University Buildings.

Instead of considering each location with the standards of location classification, the CoV was aggregated for each categories-name as shown in Table VII. Table VII shows CoV according to non-overlapped categories-name. Also, Table VII has more detailed information for Volunteer 1 than Table III; for example, bus terminals are clearly separated from train stations, i.e., categorized rather than classified. An interpretation of Table VII shows that Volunteer 1 visited Home 80 times, University Buildings 76 times, and Korean Restaurant 39 times. The sample data shown in Table VI and Table VII was collected only from Volunteer 1. Table VII shows the location data only from volunteer 1, meaning that there is a total of 14 similar tables for each volunteer.

In sum, 96 categories-names were derived from the data of the 14 volunteers with some overlapping of categories-names. After the Standard Industrial Classification and ICC, location categories with a total CoV less than 10 for all volunteers were merged into Etc. category. The final data set consists of 55 categories-names of the 14 volunteers and provided dependent variables for the regression analysis.

TABLE VI. SAMPLE OF CATEGORIZED DATA OF VOLUNTEER 1

CoV	*Location name*	*Categories-name*
58	Hongik Univ.T	University Buildings
5	CGV Ori	Theater
4	Songtan yeongbillu	Chinese Restaurant
3	Shinsegae Department Store	Department Store
3	TongtongPig	BBQ shop

TABLE VII. VOLUNTEER 1: REPRESENTATIVE CATEGORIES-NAME AND CORRESPONDING COUNT OF VISIT

Categories-name	*CoV*	*Categories-name*	*CoV*
University Lab	83	Language School	4
Home	80	Train Station	4
University Buildings	76	Art Museum	4
Cafe	60	Bank	4
Korean Restaurant	39	Fast Food Shop	4
etc.	39	Bus Stop	3
University Library	21	Cosmetics Shop	3
Dessert Shop	17	Sushi Shop	3
Outdoor Activity	15	Subway	3
Bar	12	Grocery Store	3
Chinese Restaurant	11	Shopping Centre	3
Snack Shop	11	Clothing Store	2
Noodles Restaurant	8	American Restaurant	2
Fried Chicken Shops	8	Hospital	2
BBQ Shop	8	Convenience Store	2
Japan Restaurant	8	Museum	1
Italy Restaurant	6	Gymnasium	1
Bakery	6	Salad Section	1
Theater	6	Electronics Stores	1
Airport	5	University Stores	1
Rest Area	5	Asian Restaurant	1
Department Store	5	Pizzeria	1
Bus Terminal	5	Seafood Shoppe	1

IV. Relationship Result Analysis

In this section, two sorts of location data set from Section III will be regarded as two sets of dependent variables. Corresponding results will be addressed.

A. Regression analysis on location classifications

A basic regression analysis was conducted to further develop the examination. Each volunteers' BFF values, as shown in Table I, was regarded as an independent variable. The location information, as shown in Table V, was regarded as a dependent variable. The regression analysis showed the effect of personality on location preference. The basic result is shown in Table VIII using Coefficient of Determination (CoD).

Table VIII contains the following acronyms:

- CoD = Coefficient of Determination
- IE = Institutions of Education
- PAI = Public Administration Institutions

The relationship between each variable can be represented by CoD. In Table IX, the range of CoD and the interpretations of relationships are presented. In cases where the CoD value is higher than 0.4, the relationship was considered to be high. According to this criteria, a case with a CoD value higher than 0.4 was regarded as one in which the personality and location are meaningfully related.

TABLE VIII. RESULTS OF REGRESSION ANALYSIS FOR CLASSIFIED LOCATIONS

Location Classification	*CoD*	*Location Classification*	*CoD*
Leasing Service	0.71391	Library, Museum	0.31678
IE	0.64162	Retail Business	0.26152
Restaurant	0.55895	Travel Industry	0.25720
Bank	0.52616	Foreign Institutions	0.23086
Beauty Business	0.49252	Theme Park	0.22845
PAI	0.45437	Bar	0.17817
Sports Facilities	0.39169	Beverage Store	0.15551
Hospital	0.32293	Art Viewing Facilities	0.09206

TABLE IX. RELATIONSHIP DETERMINED BY COEFFICIENT OF DETERMINATION

Range of CoD	*Interpretation*
0.9-1.0	Very High
0.7-0.9	High
0.4-0.7	Relatively High
0.2-0.4	Low
0.0-0.2	Rarely Related

Table VIII shows location classifications and the corresponding CoD. The location classifications of Restaurant, Bank, Public Administration Institutions, Leasing Service, Institutions of Education, and Beauty Business show CoD values higher than 0.4, meaning that only 6 classifications showed a meaningful relationship between personality and location.

It may be meaningful to see the effect of the subcategories of the location classifications. The next subsection will show the results for location categories.

B. Regression analysis on location categories

The regression analysis on location categories found a higher number of meaningful relationships. BFF values, as shown in Table I, were also used as an independent variable, while the total 55 location categories were used as a dependent variable. The regression analysis showed that 31 location categories had a CoD higher than 0.4, indicating a meaningful relationship between personality and location.

 ISBN: 978-1-61208-428-2

Table X shows a selection of meaningful results from the regression analysis and, in addition to the CoD, a coefficient for Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (O, C, E, A, N) was used.

The coefficient of O, C, E, A, N shows the degree of personality's effect on corresponding location categories. First, the absolute value of each coefficient stands for the strength of affection: positive coefficient means positive effect, while negative coefficient means negative effect.

The criteria of each coefficient are determined to be relative when the absolute value of the coefficient is more than 0.01, and strongly related when the absolute value coefficient is more than 0.1. Table X contains a column titled RELATION which represents the effect of personality on the location categories. The RELATION column shows the final results of this research. As a factor of meaningfulness, the plus sign (+) stands for the positive relation to the corresponding location, and the minus sign (-) stands for the negative relation to the corresponding location. For the strong positive relation, (++) is used, while the strong negative relation is indicated with the use of (- -). In addition, p-value must be considered. Once p-value is greater than 0.05 for a relation, the personality factor will be statistically disregarded and will not be presented in RELATION column.

As an example, University Buildings have an O, C, E, A, N coefficient of -0.13274, 0.07956, 0.01043, 0.00477, 0.21439, respectively. Openness and Neuroticism, which are also two factors of BFF, show an absolute value more than 0.1; therefore, - -O and ++N show that Openness strongly affects the location of University Buildings positively and Neuroticism strongly affects University Buildings negatively. Conscientiousness and Extraversion have coefficient values more than 0.01 but lower than 0.1; therefore, +C and +E are also applicable to the location category of University Building. Overall, the corresponding relationship column may contain - -O, +C, +E, and ++N. However, +C and +E must be excluded since their corresponding p-values are greater than 0.05.

Some of the RELATION columns are remained blank because the corresponding relation is negligible, whereby the coefficients of O, C, E, A, N are less than 0.01. The corresponding locations may be related to BFF since CoD is bigger than 0.4; however, the corresponding coefficient for each BFF is less than 0.01. For example, in the case of Museum, the coefficients for O, C, E, A, N are $0.00250, 0.0008, 0.00402$, -0.00828, -0.00541, respectively. Even though Agreeableness has the largest absolute coefficient value, implying a negative relationship, it is not shown.

In terms of restaurants, categories such as Korean Restaurant, Asian Restaurant, and Japanese Restaurant sometimes showed a strong relationship with BFF, while American Restaurant and Italian Restaurant showed a more minor relationship with BFF. This phenomenon might have been affected by CoV, which suggests that more data needs to be collected.

As shown in Table VII, the highest CoVs can be seen for Home and University, which is unsurprising considering that most of the volunteers were students at the time of the study. However, in Table X, University-related locations are present, while a Home location is not. The CoD for Home is less than 0.4, so maybe the location Home has a neutral relationship with human personality. Such a large CoV for University, University Buildings, University Library, University Lab, and University Stores should be categorized and more closely investigated.

Table X has the highest CoD for University Buildings of 0.87732, and O and N are RELATED for University Buildings. A similar BFF pattern could possibly be found for workers' offices. In the case of Fast Food Shop, which has a CoD of 0.75860, negative Extraversion have relations but they are not strong.

From the results in Table X, Openness, Extraversion and Neuroticism are affective factors for location visit while Conscientiousness is the negligible factor for location visit. Maybe the more data will reveal more relationship between BFF and locations since it cannot be guaranteed that our data covers all possible locations or covers all possible combinations of personality factors.

V. Conclusion

Regression analysis is usually an acceptable tool to use when deducing statistical fitness in a linear relationship between independent and dependent variables. BFF values were used as independent variables in our research, while categorized and classified location data were regarded as dependent variables. The result of the regression analysis is meaningful and, for several locations, the effect of personality was revealed. It is meaningful that we used a scientific method, instead of relying on conjecture, to identify the effect of personality on several specific locations. We also identified an aspect of the functional relationship between human personality and location preference.

Our research can be applied to many of related areas in order to increase imposed value of each area, such as enhancement of mobility model [9][10], personal travel recommendation system [11], mobile computing [12] and so on. The key is personalized future Internet, especially for location based social network related area.

An implication of this relationship emerged after the basic stage of our research, when we had obtained 2,842 location datum. Eventually, the total of 11,154 location datum was used to provide more meaningful research results. It is possible that, if more volunteer participants check-in at more locations, further insight into the personality-location relationship will be gained.

However, our research is likely to have biased results in two reasons. Even though we collected enough number of locations visited, only 14 volunteers are participating and we have a small set of BFFs comparing to the size of location data. It is an unavoidable bias and will be solved with more volunteers are involved as we continue our research. Sometimes the bias can be intentional supposing the travel recommendation system which we mentioned as an application, where travelers can be considered mostly. The second bias is that most of our volunteers are university students, which leads to lots of check-ins to university buildings. Due to this bias, however, we obtained meaningful results on the relationship between personality and university buildings. Since we continue collecting more BFFs, we expect that more distinct relationship can be found.

The more proper method of our research may be stepwise regression. Independent application of independent variable from BFF by stepwise regression will give more clear result

 ISBN: 978-1-61208-428-2

TABLE X. THE EFFECT OF PERSONALITY TO LOCATIONS WITH COEFFICOENT OF DETERMINATION

Location	*CoV*	*CoD*	O	C	E	A	N	*RELATION*
University Buildings	741	0.87732	-0.13274	0.07956	0.01043	0.00477	0.21439	- -O, ++N
Fast Food Shop	55	0.75860	0.00467	-0.01153	-0.02299	-0.00340	-0.00628	-E
Korean Restaurant	690	0.72413	0.08739	-0.03577	-0.03540	-0.02518	-0.02823	+O
University Library	104	0.66507	-0.02990	0.02783	0.00124	-0.02349	0.01764	-O
Asian Restaurant	50	0.64682	0.01263	-0.00389	-0.02293	-0.00772	0.00011	-E
Seafood Shoppe	31	0.64301	0.00038	0.00223	-0.00106	0.00057	-0.00158	
Sushi Shop	31	0.64136	0.00074	-0.00010	-0.00411	0.00259	-0.00146	
Cosmetics Shop	14	0.63816	-0.00118	-0.00579	0.00513	-0.00517	-0.00851	
Japan Restaurant	126	0.60027	0.00224	-0.01117	-0.01006	0.00342	0.00876	
Pizzeria	10	0.59626	-2.00E-05	-0.00668	-0.00069	0.00297	-0.00254	
Hotel	86	0.57299	0.00773	-0.00104	0.00437	-0.01647	-0.01546	-N
Language School	34	0.57297	-0.01789	-0.01246	-0.00626	0.01823	-0.00212	-O
Bank	42	0.56085	-0.00595	0.00112	-0.00181	0.00366	0.00048	
BBQ Shop	209	0.54467	0.02409	0.00055	0.00260	-0.02265	-0.00241	+O
Snack Shop	79	0.53491	-0.00561	-0.00752	-0.00728	0.01758	0.00432	+A
University Lab	731	0.52897	0.12886	0.08422	0.10599	-0.07828	-0.00644	
Electronics Stores	18	0.52038	0.00352	0.00499	-0.00526	-0.00484	0.00448	
Beauty Salon	24	0.51725	0.00083	-0.00273	0.00226	-0.00261	-0.00528	
Department Store	35	0.51175	-0.00407	-0.00902	-0.00134	0.00555	-0.00569	
Airport	103	0.51101	0.00615	0.00204	0.00314	-0.01302	-0.01522	-N
Rest Area	81	0.50978	0.00911	-0.00300	0.01236	-0.01191	-0.01270	
Outlet Store	11	0.48241	-0.00200	-0.00883	0.00328	0.00314	-0.00608	
Chicken Shops	90	0.48083	0.00429	-0.00713	-0.00655	0.00452	0.00644	
Public Institution	53	0.47681	-0.00358	0.00170	-0.00223	0.00301	-0.00098	
Museum	16	0.47463	0.00250	0.00081	0.00402	-0.00828	-0.00541	
University Stores	14	0.45565	-0.00135	-0.00779	0.00133	0.00766	-0.00356	
Pharmacy	53	0.44116	-9.90E-05	-0.00812	-0.00120	0.00504	-0.00706	
American Restaurant	48	0.43308	0.00320	-0.00392	-0.00216	0.00355	0.00161	
Italy Restaurant	61	0.43121	-0.00162	-0.00147	0.00241	0.00909	0.00281	
Salad Section	14	0.42998	-0.00087	0.00070	-0.00059	0.00158	0.00343	
Convenience Store	279	0.42796	-0.00051	0.00194	-0.01514	0.01295	-0.00944	

about the effect of each personality factors on location preference based on the orthogonality of each five factor.

One more possible research would include temporal information as well as location information. The check-in time and check-in duration will provide more information about the mobility pattern of each volunteer and our understanding of the personality-location relationship will become more precise. The inclusion of other personal factors in future research, including occupation, income, and gender, studies will give keener insight into this relationship.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (NRF-2012R1A2A2A03046473).

REFERENCES

[1] O. P. John and S. Srivastava, "The big five trait taxonomy: History, measurement, and theoretical perspectives," Handbook of personality: Theory and research, vol. 2, no. 1999, 1999, pp. 102–138.

[2] H. Smith, N. R. Draper, and H. Smith, Applied regression analysis. Chapman and Hall., 1981.

[3] Y. Amichai-Hamburger and G. Vinitzky, "Social network use and personality," Computers in human behavior, vol. 26, no. 6, 2010, pp. 1289–1295.

[4] M. J. Chorley, R. M. Whitaker, and S. M. Allen, "Personality and location-based social networks," Computers in Human Behavior, vol. 46, 2015, pp. 45–56.

[5] S. Y. Kim and H. Y. Song, "Predicting human location based on human personality," in Internet of Things, Smart Spaces, and Next Generation Networks and Systems. Springer, 2014, pp. 70–81.

[6] ——, "Determination coefficient analysis between personality and location using regression," in International Conference on Sciences, Engineering and Technology Innovations, 2015, pp. 265–274.

[7] P. T. Costa and R. R. McCrae, "Four ways five factors are basic," Personality and individual differences, vol. 13, no. 6, 1992, pp. 653–665.

[8] L. R. Goldberg, "The structure of phenotypic personality traits." American psychologist, vol. 48, no. 1, 1993, p. 26.

[9] D. Alberg, M. Last, and S. Elnekave, "A spatio-temporal simulation model for movement data generation," in Data Mining Workshops, 2008. ICDMW'08. IEEE International Conference on. IEEE, 2008, pp. 320–325.

[10] D. Guo and W. Cui, "Mining moving object trajectories in location-based services for spatio-temporal database update," in Sixth International Conference on Advanced Optical Materials and Devices. International Society for Optics and Photonics, 2008, pp. 71 432M–71 432M.

[11] Y. Doytsher, B. Galon, and Y. Kanza, "Storing routes in socio-spatial networks and supporting social-based route recommendation," in Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks. ACM, 2011, pp. 49–56.

[12] W.-C. Peng and M.-S. Chen, "Developing data allocation schemes by incremental mining of user moving patterns in a mobile computing system," Knowledge and Data Engineering, IEEE Transactions on, vol. 15, no. 1, 2003, pp. 70–85.

 ISBN: 978-1-61208-428-2