International Journal of Soft Computing and Engineering (IJSCE)
ISSN: 2231-2307, Volume-2, Issue-6, January 2013

# Determining Keirsey Temperament Class of a Person Based on his GPS Data

**Shreyash Srivastava, Shaurya Ahuja, Shivam Tyagi**

*Abstract— **The smart phone market has steadily gone up in the recent years. The technologies have taken a further leap and immensely improved. The Global Positioning System (GPS) is one location acquisition technology that is present in almost every smart phone that we use. The dataset obtained from the location acquisition technologies can in some way be used to determine various useful results and conclusions from the movement patterns of a person. The places a person visits and the frequency of visits, to some extent describe his personality. This paper presents one such innovative idea to determine the Keirsey Temperament class of a person by processing his collected GPS data.***

***Index Terms—Global Positioning System, location acquisition, temperament, data mining, reverse geocoding.***

## I. INTRODUCTION

Global Positioning System works for finding the geographical coordinates of a location accurately and within quick time. This location information can be used in uncountable ways to derive conclusions which can find their application in modern fields of research. With the advancements in such location acquisition technologies, it has become increasingly simple to save the location histories of a person. These location histories can be used to determine various character traits of a person. In this paper, we propose a method to determine the Keirsey temperament class of a person by analyzing his GPS dataset.

### A. *Keirsey's Temperament Classification*

Temperament is a configuration of observable personality traits, such as the habits of communication, patterns of action, and sets of characteristic attitudes, values, and talents. It also encompasses personal needs, the kinds of contributions that individuals make in the workplace, and the roles they play in society.

The temperament classification [1] was done by Dr. David Keirsey based on the study of temperaments by Hippocrates and Plato. The names of the classes were the same as suggested by Plato. The classification of temperaments is further sub-classified into sixteen character types based on introversion versus extraversion, thinking versus feeling, sensation versus intuition, and the coexistence of principal and auxiliary functions which were described by Isabel Briggs Myers and Katharine Cook Briggs called the Myers-Briggs Type Indicator. The temperament classes are as follows:

***A.1. The Guardians:*** Those belonging to this type are concrete and cooperative. They are generally concerned with responsibility and duty and seek security and belonging,. Logistics is their greatest strength. These people usually excel at organizing, facilitating, checking, and supporting.

***A.2. The Artisans:*** Those fitting in this category are concrete and pragmatic. Seeking stimulation and virtuosity, they are concerned with making an impact. Tactics usually is their greatest strength. They excel at troubleshooting, agility, and the manipulation of tools, instruments, and equipment.

***A.3. The Rationalists:*** These are abstract and pragmatic people. In pursuit of their search for mastery and self-control, they are highly concerned with their own knowledge and competence. Their greatest strength is strategy. These people outdo others in any kind of logical investigation such as engineering, conceptualizing, theorizing, and coordinating.

***A.4. The Idealists:*** Those belonging to this type are abstract and cooperative. In their quest for meaning and significance, their major concerns are personal growth and finding their own unique identity. Their greatest strength is diplomacy. They excel at clarifying, individualizing, unifying, and inspiring.

In this paper, we propose a method to determine the temperament class from one of the above to which a particular person may belong. This is performed by performing a set of operations on the GPS trajectory of the user as collected from his GPS receiver. The process starts by first calculating the stay points from the trajectory and then determining the category (a hotel, an office, a pub, etc.) of the stay point. A sequence of these categories is then worked upon by using a rating algorithm. This gives the Keirsey temperament class to which the person may belong.

The organization of the paper is as follows. Section 2 describes the related work that has been done. Section 3 starts by defining some of the basic definitions that have been used throughout the paper and then goes on to describe the approach that has been used. Section 4 describes the results obtained. The conclusion and future work form the section 5 of the paper.

## II. RELATED WORK

Various personality classifications have been proposed till date dating back to 400 BC. Table below shows these classifications and the period when they were popular.

Table 1: Personality Classification History

| | Date | Type1 | Type2 | Type3 | Type4 |
|---|---|---|---|---|---|
| Emped-ocles | 400 BC | Fire | Air | Water | Earth |
| Hippo-crates | 460 BC | Yellow Bile | Blood | Phleg-m | Black Bile |
| Galen | 131 | Choler-ic | Sangui-ne | Phleg-matic | Melan -choly |
| DISC | 1920s | Domi-na nt | Influe-ntial | Steady | Cauti-ous |
| Jungi-a | 1940s | NF | SP | NT | SJ |

**Manuscript received on January, 2013.**

**Shreyash Srivastava**, Department of Computer Science, Indian Institute of Technology Roorkee, Roorkee, India.
**Shaurya Ahuja**, Department of Computer Science, College of Engineering Roorkee, Roorkee, India.
**Shivam Tyagi**, Department of Computer Science, College of Engineering Roorkee, Roorkee, India.

| | | | | | |
|---|---|---|---|---|---|
| n | | | | | |
| Keirsey | 1960s | Idealist | Artisan | Ratio-nal | Guard-ian |

One of the most influential works was by Carl Jung [5] which gave the Jungian classification. Another classification related to Jungian classification is Myers-Briggs Type Indicator (MBTI). Myers-Briggs [2][3] type indicator classifies the population into 16 types based on the following criteria:

- Source and direction of energy: Introverts and Extraverts
- The perceiving function: Sensing and iNtuition
- The decision function: Feeling and Thinking
- Relating to Outer World: Judging and Perceiving

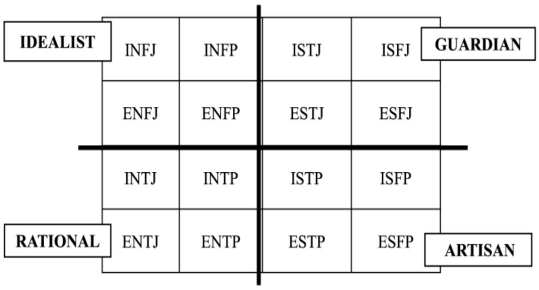
The combinations of these criteria make a total of 16 personality types. The percentage breakdown [15] of population into these types is shown in the fig. 1.

The relation between the Myers-Briggs Type Indicators (MBTI) and Keirsey Temperament Classes is also shown in figure 2.

| Population Breakdown | | | |
|---|---|---|---|
| ISTJ 11.6% | ISFJ 13.8% | INFJ 1.5% | INTJ 2.1 % |
| ISTP 5.4% | ISFP 8.8% | INFP 4.4% | INTP 3.3% |
| ESTP 4.3% | ESFP 8.5% | ENFP 8.1% | ENTP 3.2% |
| ESTJ 8.7% | ESFJ 12.3% | ENFJ 2.4% | ENTJ 1.8% |

Fig. 1: MBTI Population Breakdown Source: New World Encyclopedia- http://www.newworldencyclopedia.org/entry/Myers-BriggsTypeIndicator

Therefore, we can have a percentage breakdown of world population into the four temperament classes. From the above two figures, we can have an estimate of percentage of population that belongs to each temperament class. This breakdown can be shown as below in fig. 3.




**Source:** Adapted from Isachsen and Berens (1988)

Fig. 2: Relation between Keirsey Temperament Class and Myers-Briggs Type Indicator

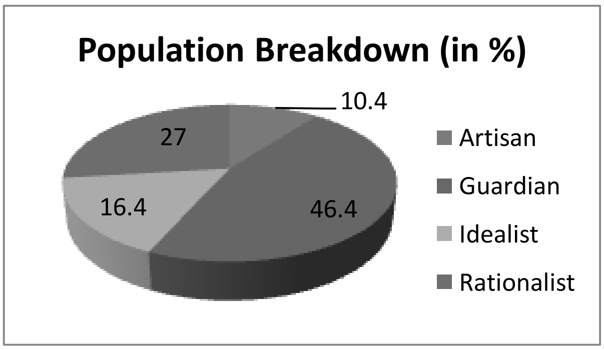



Fig. 3: Population Breakdown for Keirsey Temperament Class

Apart from personality classifications, much work has been done to find various patterns and conclusions from GPS trajectories of users. The work of Yu Zheng et. al. [6][7][8] has been incredible. Much work has been done in finding significant places from the GPS trajectories. Spatial coordinates have been used to group the locations and in our previous work we have tried to perform temporal grouping of GPS data [14].

Friend Recommender System by Zheng et. al.[8] has been a commendable work in this field. Location histories and user experiences have been made used wisely. Movement histories and mobility patterns [9][13] have also been used to derive conclusions of various sorts.

In this paper, we attempt to relate the research in both fields as described above. We attempt to find a way to determine the personality type (Keirsey Temperament Class) of a GPS user based on some common trajectory based methods. This work seems to be one of its kind and could prove beneficial at an organizational level.

## III. APPROACH

### *A. Preliminary Definitions*

The standard terminology has been followed throughout the paper. Given below are a few definitions which have been used frequently.

***Definition 1: GPS User:*** A GPS user is any user, *u*, whose GPS data is being used.

***Definition 2: GPS Point:*** A GPS point is a triplet consisting of latitude (the x coordinate), longitude (the y coordinate) and the timestamp at which the coordinates were recorded. It can generally be denoted by P(x, y, t).

***Definition 3: GPS Trajectory:*** A GPS trajectory is a sequence of GPS points collected between a given interval of time. It can generally be denoted by $traj_k$ to denote the trajectory of $k^{th}$ GPS user. $Traj_k = (p_1, p_2, p_3 \ldots p_n)$.

***Definition 4: Stay Point:*** A stay point is any location coordinate pair where a GPS user stayed for a given duration of time (called time threshold, and denoted by $\theta_t$). To introduce some flexibility, since a user cannot exactly be at the same point throughput the duration, distance threshold ($\theta_d$) is used. So a stay point is taken as the mean of the GPS points, if the user stays a boundary defined by $\theta_d$ and if he stays within the boundary for at least $\theta_t$ amount of time.

### *B. Approach*

In this section, we discuss the procedure adopted by us to determine the Keirsey Temperament class of a user based on his GPS data. We start by processing the trajectory of a GPS user, calculating stay points from it and then finally rating the stay points based on their category to find the temperament class of the person.

**Step 1: Calculating Stay Points**

Calculation of stay points is in no way a new thing and methods exist to compute the GPS points from a given GPS trajectory. For our purpose, we have used the algorithm StayPointDetection (Traj, $\theta_d$, $\theta_t$) proposed in [8] which takes a GPS trajectory, the distance threshold, $\theta_d$ and time threshold $\theta_t$ as its input and returns all the stay points present in the trajectory Traj according to the given thresholds.

**Step 2: Determining the Category of Stay Points**

After the first step, we have a set of stay points characterized by a latitude and longitude value. As these coordinate values were not sufficient for our task to be fulfilled, we used reverse geocoding techniques to find out the category of every stay point. The category could be anything ranging from a hotel, a restaurant, an office building, a pub, a stadium and so on. Geocoding generally refers to finding the location coordinated from a specified address or a landmark. The converse of this, i.e. finding the address, category or a place name, is referred to as reverse geocoding.

**Step 3: Calculating the Percentage of Occurrence of a Category Type.**

Once the categories of all the stay points determined in step 1 have been found, the next step is to calculate the percentage of times a given category of stay point has been found.

**Step 4: Rating Algorithm**

This is the final and the main step where we try to determine the Keirsey Temperament class based on the data retrieved so far. We have the percentage of times a particular category type has been visited by the user. Using this, we aim to find one of the four temperament classes (The Guardian, The Rationalist, The Idealist, The Artisan) to which the user may belong. We try to calculate the deviations of this percentage value from a standard matrix which stores the percentage contributions of a category type to each temperament class. The one showing the minimum deviation is returned as the temperament class of the person.

The algorithm takes as input the rating matrix *R*, which contains the values of the weightage in percentage of the contribution of each category being considered towards the personality type. It also takes the matrix *freq* as input, which keeps account of the percentage of times the category of the stay point is visited. The third and the final input is the number of categories being considered, n.

---

**Algorithm Rating (*R*, *freq*, *n*)**

---

**Input:** The rating matrix R, frequency count matrix Freq, number of categories being checked after n.

**Output:** A string denoting the personality type.

1. **for** i = 1 to 4 **do**
2. **for** j = 1 to n **do**
3. deviation[i] = deviation[i] + abs(R[i][j]-freq[j])
4. min = deviation[0]
5. index = 0
6. i = 1
7. **for** i = 1 to 4 **do**
8. **if** deviation[i] < min **then**
9. min = deviation[i]
10. index = i
11. **if** index = 0 **then** return "Artisan"
12. **if** index = 1 **then** return "Idealist"
13. **if** index = 2 **then** return "Guardian"
14. **if** index = 3 **then** return "Rationalist"

---

The rating matrix stores the estimates of the percentage contribution of each category towards each of the temperament class. *R* is a 4 x n matrix, where 4 is the number of Keirsey's temperament classes and n is the number of categories of stay points under consideration.

$$\begin{array}{lc} & \begin{array}{ccccc} \textbf{Cat 1} & \textbf{Cat 2} & \textbf{Cat 3} & \dots & \textbf{Cat n} \end{array} \\ \begin{array}{l} \textbf{Artisan} \\ \textbf{Idealist} \\ \textbf{Guardian} \\ \textbf{Rationalist} \end{array} & \left( \begin{array}{ccccc} x_{11} & x_{12} & x_{13} & \dots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2n} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3n} \\ x_{41} & x_{42} & x_{43} & \dots & x_{4n} \end{array} \right) \end{array}$$

Fig. 4: Rating Matrix, R

The values $x_{ij}$ in the matrix represent the percentage contribution of the $Cat_i$ towards the $j^{th}$ temperament class, where $1 \le I \le n$ and $1 \le j \le 4$, where j = 1 implies the Artisan, j = 2 represents the Idealist, j = 3 represents the Guardian and j = 4 represents the Rationalist. The sum of the values in each row of the matrix should be approximately 100.

The second input that we have is the calculated percentage of all stay points which belong to a particular stay point category. It is a vector (single dimensional matrix) f*req* of size 1 x n, where n is the third input of the algorithm - the number of categories under consideration.

The working of the algorithm is explained in the next few lines. In the lines 1 – 3, we try to calculate the deviation of calculated values in vector f*req* from the standard values in *R*. We store these values in the matrix called *deviation* which is a column matrix. Its dimensions are 4 x 1. The value in $j_{th}$ row of the deviation matrix denotes the total deviation for the temperament class j. In the loop from 7 to 10, we find the index in the deviation matrix which has the minimum value. This is the index of the temperament class closest to the user.

## IV. RESULTS AND DISCUSSION

In this section, we describe the experiments done by us to work out the above described approach. We begin by describing the GPS data used in the process.

The GPS trajectory data was collected by Microsoft Research Asia [11]. The data consists of the trajectories of 165 users, collected over a period of two years (from April 2007 to August 2009). Each record in the dataset was composed of following fields:

- Latitude (in decimal degrees)
- Longitude (in decimal degrees)
- Code : 1, if a break in the track line, otherwise 0
- Altitude (in feet), -777 if not valid
- Date -see Date Format below, if blank a pre-set date will be used
- Date (string)
- Time (string)

*Example records –*
27.350436, 153.055540, 1, -777, 36169.6307194, 09-Jan-99, 3:08:14
27.348610, 153.055867, 0, -777, 36169.6307194, 09-Jan-99, 3:08:14

Not all the trajectories were used. Some of the trajectories were discarded siting the inconsistency of data. A total of 150 trajectories were selected to be used in the experiment.

To start with, we calculated the stay points from the GPS trajectories. The threshold values used are as under:

- Distance Threshold: 100 meters
- Time Threshold: 45 minutes

Using these threshold values, a total of 19232 stay points for 150 users were calculated by processing all the trajectories. For calculating the stay points, we used the algorithm described in section III, subsection B, step 1.

The step 2 was to determine the category of each stay point. For this, we needed to perform reverse geocoding. We used the latitude and longitude values of the stay points to determine the categories. For this, we made use of the Wikimapia API [16]. Wikimapia allows users to query the

Wikimapia DB using the API. We queried for the database to find out the category of each stay point using its coordinate values. However, since not every coordinate we used was present in the database, we had to discard a few stay points in some cases, and in others, we had to use the value of a nearby point. It was however ensured that the point being considered is not so far from the one being considered. We checked the stay point for following categories: park, pub, office building and home.

In step 3, process the stay point of each person to find the percentage of stay points belonging to each category type. This gives us the freq matrix discussed in the previous sections.

Finally, in step 4, we use this matrix freq and a rating matrix R shown below to find the deviations. Lastly, we chose the index of the element in the deviation matrix which has minimum value. If the index is 1, we get the person could be an Artisan, if index is 2, we claim he is an Idealist, for index = 3 we say he is a Guardian and for index = 4, we claim he belongs to the class Rationalist. The rating matrix for the four categories used by us is shown below:

$$
\begin{array}{l c}
 & \begin{array}{cccc} \textbf{Cat 1} & \textbf{Cat 2} & \textbf{Cat 3} & \textbf{Cat 4} \end{array} \\
\begin{array}{l} \textbf{Artisan} \\ \textbf{Idealist} \\ \textbf{Guardian} \\ \textbf{Rationalist} \end{array} & \left(\begin{array}{cccc} 15 & 40 & 20 & 12 \\ 10 & 35 & 35 & 17 \\ 30 & 50 & 30 & 10 \\ 10 & 55 & 25 & 7 \end{array}\right)
\end{array}
$$

Fig 5. Rating Matrix Used (All values in %)

After processing a total of 150 trajectories, the distribution of people into the temperament classes was as follows:

- The Artisans: 42 people (28 %)
- The Guardians: 59 people (39.33 %)
- The Idealists: 26 people (17.33 %)
- The Rationalists: 23 people (15.33 %)

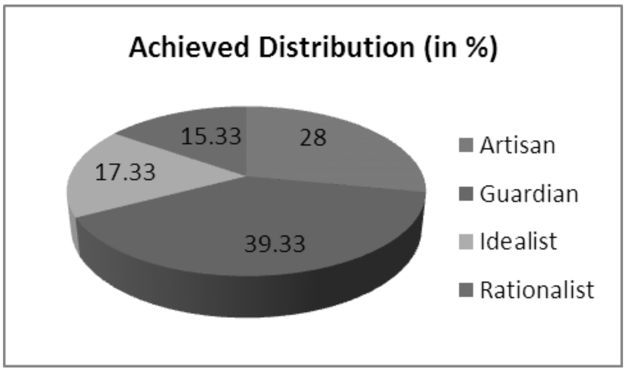



Fig 6. Results Obtained

The results obtained from the experiments and the actual population breakdown as described in Section 1.B, share a common trend. The guardians are the ones who dominate the population in numbers as is visible from the 46.4% in fig.3 and 39.3 % in fig. 6. The share of the idealists in the results is almost same as that in the actual breakdown (17.33% compared to 16.4% in actual). The results obtained for the Artisans and the Rationalist may not be in full agreement but then the data of only a 150 people could not be deemed as sufficient to compare the statistics of the world.

Knowing one's temperament class could prove beneficial in various aspects. Recruitment for jobs, various professions could consider the class of a person to find his suitability for the job. Also it can be used in determining suitable insurance plans for a person. Looking at this point from a different view frame, it would be of utmost importance for a company to develop its products and services so as to attract the customers from all personality classes, or if it wants to cater only one specific class, it would know who her customers are.

## V. CONCLUSION AND FUTURE WORK

In the paper, we attempted to determine the Keirsey Temperament Class of a person using his GPS data. The results obtained were in agreement to the actual stats. However, with greater data set availability the results can be further enhanced. Thus the results can have a lot of practical applications.

## REFERENCES

[1]. David Keirsey and Marilyn Bates, Please Understand Me: Character and Temperament Types, 5th edition, Gnosology Books Ltd., 1984.

[2]. Isabel Briggs Myers and Mary H. McCaulley, Manual: A guide to the development and use of the Myers-Briggs Type Indicator, Consulting Psychologists Press, 1985.

[3]. Myers, Isabel Briggs with Peter B. Myers (1980, 1995). *Gifts Differing: Understanding Personality Type*. Mountain View, CA: Davies-Black Publishing. pp. xi–xii. ISBN 0-89106-074-X

[4]. Olaf Isachsen and Linda V. Berens, Working Together: A Personality Centered Approach to Management, Networld Management Press, 1991.

[5]. Jung, Carl (1976). Campbell, Joseph. ed. *The Portable Jung*. New York, NY: Penguin Books. pp. 178.

[6]. Yu Zheng, Quannan Li, Yukun Chen and Xing Xie. "Understanding Mobility Based on GPS Data". In Proc. Ubicomp 2008. ACM Press: 312-321, Sept. 2008.

[7]. Yu Zheng, Lizhu Zhang, Xing Xie and Wei-Ying Ma. "Mining Interesting Locations and Travel Sequences from GPS Trajectories". In Proc. WWW 2009, April 2009.

[8]. Yu Zheng, Lizhu Zhang, Zhengxin Ma, Xing Xie and Wei-Ying Ma. "Recommending Friends and Locations Based on Individual Location History". In Proc. TWEB. Vol. 5 Issue 1 Article 8, Feb. 2011.

[9]. G. Agamennoni, J. Nieto and E. Nebot. "Mining GPS Data for Extracting Significant Places". In Proc. ICRA '09, IEEE, 855-862, May 2009.

[10]. Kathryn Jo-Ann McPhail, (2002),"The nursing profession, personality types and leadership", Leadership in Health Services, Vol. 15 Iss: 1 pp. 7 - 10

[11]. GeoLife GPS Trajectories (2011), Available: http://research.microsoft.com/en-us/downloads/b16d359d-d164-469e-9fd4-daa38f2b2e13/default.aspx

[12]. Xin Cao, Gao Cong and Christian S. Jensen "Mining Significant Semantic Locations from GPS Data". In Proc. VLDB Endowment, Vol. 3, Issue 1-2. 1009-1020, Sept. 2010.

[13]. Vincent W. Zheng, Yu Zheng, Xing Xie, Qiang Yang. "Collaborative Location and Activity Recommendations with GPS History Data". In Proc. WWW '10.

[14]. S. Srivastava, S. Ahuja, A. Mittal, Determining Most Visited Locations Based on Temporal Grouping of GPS Data, Advances in Intelligent and Soft Computing, Volume 131, 2012, DOI 0.1007/978-81-322-0491-6:

[15]. New World Encyclopedia – http://newworldencyclopedia/entry/Myers-Briggs Type Indicator

[16]. Wikimapia – http://www.wikimapia.org

[17]. Personality Page – http://www.personalitypage.com

**Shreyash Srivastava** is pursuing his post graduate studies at Indian Institute of Technology Roorkee. His international publications include "Determining Most Visited Places Based on Temporal Grouping of GPS Data". He has been working on data mining for past one year. His other interests include mobile computing, big data

computing and machine learning.

**Shaurya Ahuja** completed his graduation from College of Engineering Roorkee, Uttarakhand (India), His previous works include "Determining Most Visited Places Based on Temporal Grouping of GPS Data". His other works include "Network Security Issues Of Smart Cards and its Commercial Aspect" and "Implementation Of Contactless Smart Cards in Industry". His research areas are network security, data mining, mobile and pervasive computing.

**Shivam Tyagi** completed his graduation from College of Engineering Roorkee, Uttarakhand (India). His previous works include that on "Road Traffic Monitoring by Intelligence-Driven Window Based Image Analysis". His research interests include data mining, network security and machine learning.