A Fuzzy Logic approach to analyze a Student's Lifestyle

Sourish Ghosh*, Aaditya S. Boob*, Nishant Nikhil*, V. Nayan Raju*, Ankit Kumar*, and S. K. Barai†
*Department of Mathematics
Email: {sourishg, aaditya733, nishantnikhil, v.nayanraju}@iitkgp.ac.in
†Professor, Department of Civil Engineering
Email: skbarai@civil.iitkgp.ernet.in
Indian Institute of Technology, Kharagpur

***Abstract*—A college student's life can be primarily categorized into domains such as education, health, social and other activities which may include daily chores and travelling time. Time management is crucial for every student. A self realisation of one's daily time expenditure in various domains is therefore essential to maximize one's effective output. In this paper we present how an Android application using Fuzzy Logic and Global Positioning System (GPS) analyzes a student's lifestyle and provides recommendations and suggestions based on the results.**

Keywords*—*Fuzzy Logic, GPS, Android Application

I. Introduction

A college student's life is multidimensional. Students are expected to be academically excellent, physically fit and socially active along with managing their daily chores and pursuing their fields of interest. This structure would not only help students in engage all activities but also help them live a balanced life. This practice would eventually help them make better career choices on the basis of their interests. For such a practice one needs to invest a threshold amount of time and effort in all the activities. However only a certain amount of students are involved and excel in such a practice. This paper discusses a novel approach using fuzzy logic to generate an analysis of a student's daily time expenditure in these various categories. Based upon the analysis of the results obtained from the above data appropriate results must be provided on regular basis. This would help the students work in their non-performing fields and maintain a balanced lifestyle.

A. About Fuzzy Logic

Over the past three decades, fuzzy logic is widely used in all problem-solving domains. One of the reasons for such instantaneous growth since its inception is its usability across all sectors be it Dynamic Programming, Process Control or Optimization. Fuzzy logic discards the theory of 'Absolute Truth'and instead proposes a new theory of 'Partial Truth', also referred as degree of membership (suggested by Prof. Zadeh in 1965).

Let $\mathbf{S}$ be a nonempty set, called the *universe set*. Now, consider any crisp set $A \subset \mathbf{S}$. A characteristic function $\boldsymbol{\chi}_A$ is defined as

$$\boldsymbol{\chi}_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{otherwise} \end{cases}$$

A characteristic function assigns value of either 0 or 1 to each element of $\mathbf{S}$. Now consider a fuzzy set $B \subset \mathbf{S}$. A membership function $\boldsymbol{\mu}_B(x)$ is defined as $\boldsymbol{\mu}_B : \mathbf{S} \to [0, 1]$. Unlike the notion of a set in classical set theory where an element either belongs or does not belong to a particular set based on a bivalent condition, in fuzzy set theory an element's belongingness to a particular set is decided using membership function which gives a membership value between 0 and 1.

B. Problem Formulation

The problem can be divided into three major parts:

- **Data Collection**: Using GPS, we first collect where and how long the user spends his/her time and tabulate that data. For this we use the Google Places API.
- **Fuzzification**: Fuzzify the crisp input and calculate the values of corresponding membership functions.
- **Defuzzification**: Set up a fuzzy inference system based on certain rules and then return recommendations and suggestions.

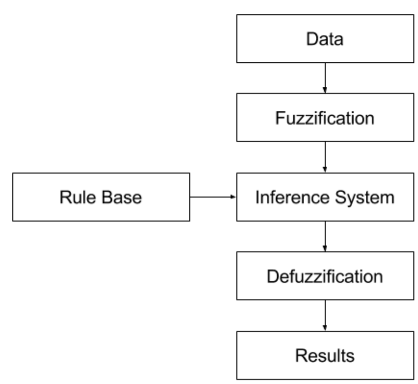

Fig. 1: Fuzzy Logic System

II. Working Principle

A. Data Collection

A college student is carrying his/her smart phone everywhere. Hence using the GPS we can extract his/her position throughout the day. In the application and testing of this paper, the mobile application was developed on Android while the point of interest was extracted using Google Maps API by querying the user's location extracted from GPS. Google Maps

TABLE I: Sample locations and purposes

Location	Purpose
Cafe, Restaurant	Going out with friends and family.
Supermarket, Gas Station	Chores
Gym, Ground, Hospital	Exercise or Health Treatment
Cinema Hall, Spa	Leisure and Relaxation
Bank, Business Associates	Work

API classifies most of the positions into various categories namely `restaurant`, `shopping_mall`, `city_hall` etc. Let us refer to all these categories hence forward as *tags*. Apart from these existing tags, we generate two additional tags namely *home* and *work*. The GPS data for these two additional tags would be user specific. Hence initially every user needs to update their location for these two tags specifically. This step is conducted so as to recognize distinctly one's home and workplace which in further course would generate accurate results. Let consider this example, one might go to a pizza shop to hang out with friends and family. However if someone is working in a pizza shop and the GPS details of the specific pizza shop is not known beforehand, it is very likely that one might consider this entire working time as time utilized for hanging out with friends. However if the person goes to some other pizza shop it is very likely he is going out with friends. To avoid this confusion this initial step has to be carried out.

Let $\mathbf{X}$ be the set of all tags defined as $\mathbf{X} = \{x \mid x \text{ is a tag}\}$. Let T be a function defined as $T : \mathbf{X} \to \mathbb{R}$ where $T(x)$ denotes time spent at a tag $x \in \mathbf{X}$ in minutes. Analysing the way a person lives is governed by many parameters, but in a typical student's life we are mainly concerned about one's health, education, leisure and social life. However a person also invests certain amount of time which fails to fall under this category. A example of this would be travelling time. Now let S, L, H, W, O be subsets of $\mathbf{X}$ defined as

$$
\begin{aligned}
S &= \{x \mid x \in \mathbf{X},\ x = \text{social and } x \neq \text{home, work}\} \\
L &= \{x \mid x \in \mathbf{X},\ x = \text{leisure and } x \neq \text{home, work}\} \\
H &= \{x \mid x \in \mathbf{X},\ x = \text{health and } x \neq \text{home, work}\} \\
W &= \{x \mid x \in \mathbf{X},\ x = \text{work and } x \neq \text{home}\} \\
O &= \{x \mid x \in \mathbf{X} \text{ and } x \notin S \cup L \cup H \cup W\}
\end{aligned}
$$

A *tag* x might belong to one or more of the sets S, L, H, W. For example, a person might visit an Amusement Park. In this case the person's *social* and *leisure* scores are both incremented. Using this categorization technique we can extract one's location and time spent at each *tag* for the entire day. TABLE I lists down some locations and their possible purpose of visits. The locations mentioned below are basically *tags* other than *home* and *work*.

Weighing criteria: For a given purpose, different locations would have different amount of productivity and impact. For example, hospital and gym both fall under the *health* category. However one visits Gym to increase his physical activity and hence visiting Gym has a positive impact on one's health. However one visits a hospital if he/she has fallen sick. Hence, visiting a hospital has a negative impact on one's health. So a monotonous grading for all the locations is bound to fail.

We define a function $Y : \mathbf{X} \to \mathbb{R}$ such that $Y(x)$ for every $x \in \mathbf{X}$ denotes the time spent at the location *tag* x. For example, let $x = \text{gym}$. Say $Y(x) = 50$. This implies a person has spent 50 minutes at a *gym* in the entire day.

We define a function $Z_S : S \to [-100, 100]$ such that $Z_S(x)$ for every $x \in S$ denotes the intensity of the *tag* x with respect to the *social* category. Similarly we define Z_H, Z_L, Z_W, and Z_O for the *health*, *leisure*, *work*, and *other* categories respectively. The range $[-100, 100]$ is chosen for normalization purposes. For example, let $x = \text{gym}$. Say $Z_H(x) = 50 > 0$ as gym has a positive health impact. Let $y = \text{hospital}$, then $Z_H(y) = -20 < 0$ as hospital has a negative health impact. However, $Z_L(x) = Z_L(y) = 0$ as both x and y don't contribute to the *leisure* category. Also note that if a *tag* t belongs to two different categories, then its weightage in both the categories cannot be 0.

For both Y and Z we have excluded the *home* tag as it is a special case. This is explained later.

Assigning weights: One is free to assign the weights independently. However for better results, we assign weights by conducting a survey to understand how appropriately a location *tag* fulfils the purpose of a category. For instance consider the *health* category. In the survey we ask a sample population to rank every $x \in H$ in an order of fulfilment of their positive *health* benefits. Consider the following survey with

$$
\begin{aligned}
H = \{&\text{gym, playground, swimming\_pool, health\_club,} \\
&\text{hospital, pharmacy, physiotherapist, dentist, doctor}\}
\end{aligned}
$$

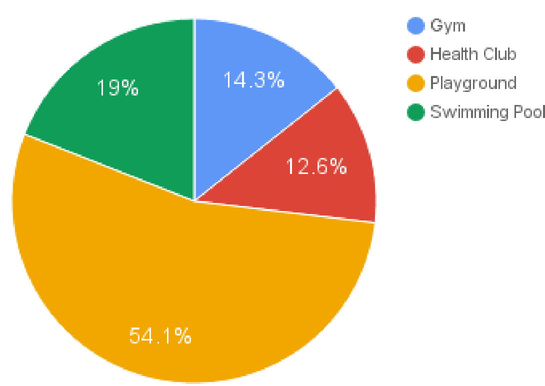

Fig. 2: Survey for *positive* health weights

Figure 2 shows a survey for determining *positive* weights in the *health* category. As 54.1% people taking the survey voted *playground* as their maximum positive benefit from the *health* category, the corresponding weight for $x = \text{playground}$ is computed as $Z_H(x) = \frac{54.1}{100} \times 100 = 54.1$.

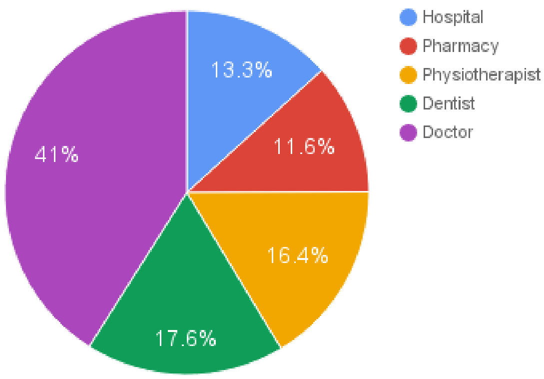

Fig. 3: Survey for *negative* health weights

Figure 3 shows a survey for determining *negative* weights in the *health* category. As 41% people taking the survey voted *doctor* as their maximum non fulfilment from the *health* category, the corresponding weight for $x =$ doctor is computed as $Z_H(x) = -\frac{41}{100} \times 100 = -41$.

Home tag: The time spent at the *home* location might not be entirely used for rest and leisure purpose only. One might practice yoga at one's home and the equivalent time should be added to the *health* category. Let τ denote the total time spent at *home*. And τ_H, τ_W, τ_L, τ_O, τ_S denote the equivalent time in respective categories. This time is taken as user input through the Android application. For better results random push notification system can be used to learn the characteristics of the user. The home tag will be associated with weights ξ_H, ξ_S, ξ_L, ξ_W, ξ_O which denote the intensity of the tags at *home*. For instance, $\xi_W = 30$ and $Z_W(\text{office}) = 50 > 30$ as working at *home* might not be as productive as working at *office*.

B. Fuzzification

Fuzzification of time: Consider a person p. Suppose p visits tags $\{x_1, x_2, \ldots, x_n\}$, with the time spent at these locations denoted by $\{Y(x_1), Y(x_2), \ldots, Y(x_n)\}$. Let K_H, K_L, K_S, K_W, K_O denote the overall time spent in *health*, *leisure*, *social*, *work*, and *other* categories respectively. Then

$$K_H = \sum_{x_i \in H} T(x_i) + \tau_H$$

Similarly, K_W, K_S, K_L, K_O are defined.

We define the following fuzzy sets in the following categories. These sets define the type of lifestyle a person is living in each category. Here *leisure* also includes rest.

$$\begin{aligned} \textit{health} &= \{\text{unfit},\ \text{fit},\ \text{proactive}\} \\ \textit{leisure} &= \{\text{hectic},\ \text{ideal},\ \text{lazy}\} \\ \textit{social} &= \{\text{reserved},\ \text{sociable},\ \text{over\_social}\} \\ \textit{work} &= \{\text{lethargic},\ \text{hard\_working},\ \text{industrious}\} \\ \textit{others} &= \{\text{non\_productive},\ \text{productive}\} \end{aligned}$$

The membership functions for these fuzzy sets are constructed by conducting a survey on a sample population. We will approximate the data from the survey using quantile range and trapezoidal membership functions. However, one can use various other techniques to plot membership functions. For instance, in a sample survey the hours spent by fit students in the *health* category were: 0.45, 1.25, 2, 2.25, 2.5, 2.5, 2.75, 2.75, 3, 4, 4.25. So with respect to the inter quantile range $Q_1 = 2$, $Q_2 = 2.5$, $Q_3 = 3$, $\inf = 0.45$, and $\sup = 4.25$. The trapezoidal membership function for the linguistic term "fit" under the *health* category using these values is shown in Figure 4.

Figure 5, 6, 7, 8, and 9 show the membership functions for each linguistic terms under each *category*.

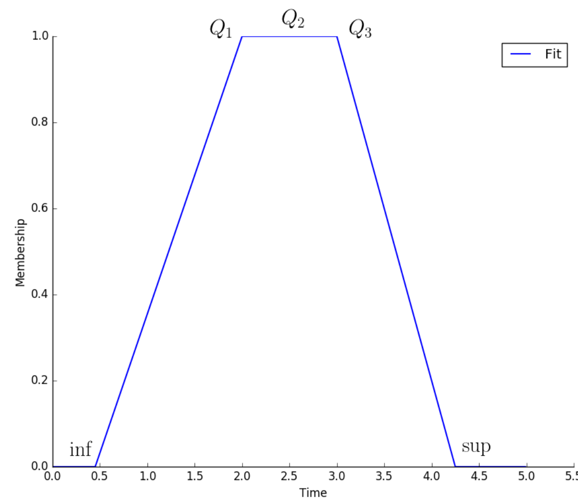

Fig. 4: Membership function for the linguistic term *fit*

Fuzzification of score: Not only the time spent at a location is important but also how the time is spent is important too. This effective utilisation of time is denoted by a *score* M_S, M_L, M_O, M_W, and M_H for the respective categories. The score for the *social* category is calculated as follows

$$M_S = \sum_{x \in S} Y(x) Z_S(x) + \tau_S \xi_S$$

Similarly we define the other scores. The linguistic terms that define the fuzzy score sets for each category remain constant and are "low_score", "ideal_score" and "high_score". But the membership functions for each of these linguistic terms differ from category to category. The shape of these functions are determined by conducting similar surveys on a sample population.

Appendix A
Proof of the First Zonklar Equation

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi

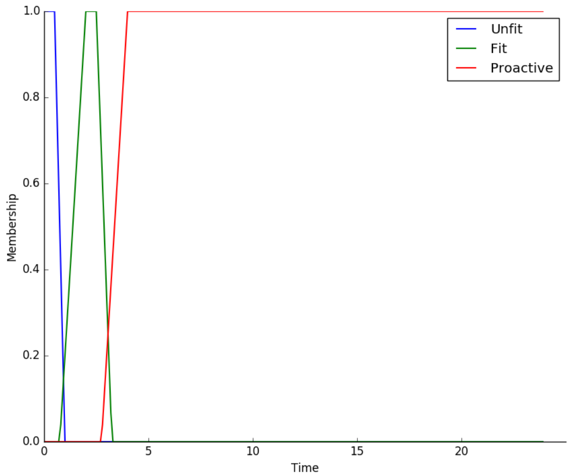

Fig. 5: Membership function for *health* category

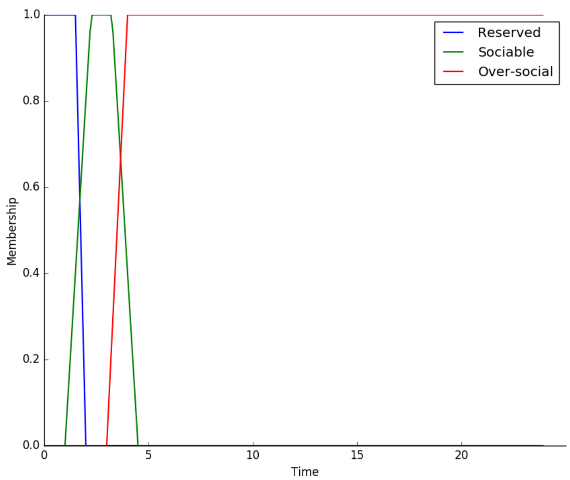

Fig. 6: Membership function for *social* category

sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

ACKNOWLEDGMENT

The authors would like to thank...

REFERENCES

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.

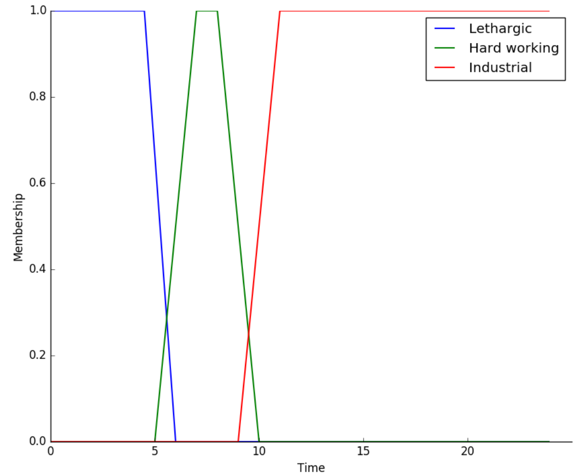

Fig. 7: Membership function for *work* category

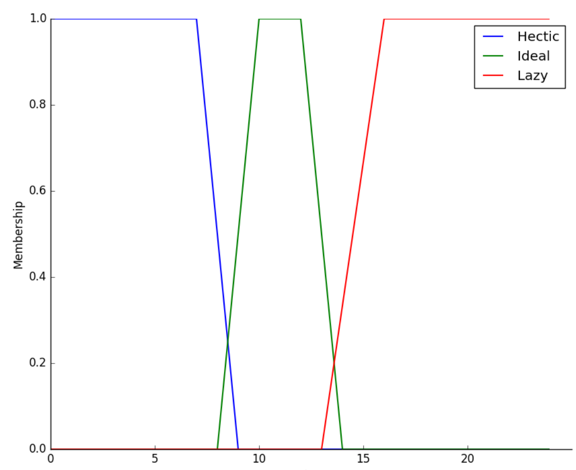

Fig. 8: Membership function for *leisure* category

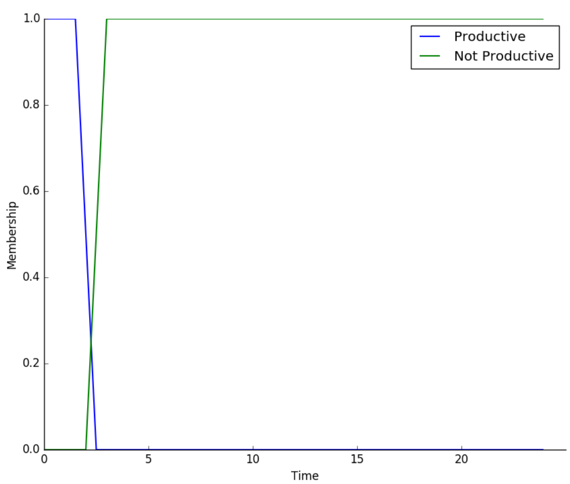

Fig. 9: Membership function for *other* category