



# The Open Transportation Journal

Content list available at: <https://opentransportationjournal.com>



## RESEARCH ARTICLE

# Crowdsourcing as a Social Interaction Tool to Stimulate Sustainable Transportation Mode Use

Shailesh Chandra<sup>1,\*</sup> and Vinay Kumar<sup>1</sup>

<sup>1</sup>Department of Civil Engineering and Construction Engineering Management, California State University, Long Beach, 90803 California, U.S.A.

### Abstract:

#### Background:

In this paper, a crowd-based social interaction framework is developed to assess the potential increase in the use of sustainable transportation modes – such as walking, bicycling and public transit.

#### Methods:

The empirical data were used to validate mode shift behaviors for 77 participants from California State University Long Beach. Data collection spanned over two phases, Phase I followed by Phase II. Each study phase lasted a month. Participants used one of the four modes – personal car, walking, bicycling and public transit - to arrive at the university campus. During Phase I, a control group was created, and individual mode choice of participants were obtained. Individual participants in Phase II were assigned short-encrypted distinct names and were asked to post a daily comment on the quality of experience using the mode that was used to arrive at the campus. The participants were asked to post the comments over a “Twitter” page that was used as the crowdsourcing platform for this study. The encrypted name masked the individual identity of the user. Analysis at the end of Phase II showed that there was an overall mode-shift of almost 19% of personal car users to other sustainable modes of walking, bicycling and transit.

#### Results:

Results show very important policy implications of using crowdsourcing as a social interaction tool to influence mode choice behavior of commuters, especially among college students and young adults.

#### Conclusion:

A crowd-based social interaction framework is developed to assess potential increase in the use of sustainable transportation modes – such as walking, bicycling, and public transit. Results showed that providing advanced information on traffic and parking problems can result in a mode shift to active transportation modes.

**Keywords:** Crowdsourcing, Car, Bicycling, Mode choice, Parking, Public transit, Mode shift.

### Article History

Received: February 20, 2020

Revised: April 27, 2020

Accepted: May 08, 2020

## 1. INTRODUCTION

Crowdsourcing is emerging as a significant social interaction tool to provide possible solutions to problems that are traditionally expensive to solve individually [1,2]. Crowdsourcing refers to the technique of gathering opinions and information from the crowd [3]. Therefore, when used effectively, crowdsourcing can use the public’s intelligence and skills to solve complex issues [4]. The collection of information through crowdsourcing is often facilitated by

social media platforms such as Twitter, Facebook, *etc* [5]. An additional utility of crowdsourcing can be found in the use of disaster relief operations in which traditional relief methods are unavailable or not viable [2].

Crowdsourcing can be used to determine human behavior on transportation choice decision making. In transportation, crowdsourcing, when used as social media, can be used to obtain real-time conditions of nearby public transit and rail lines, traffic delays, and parking conditions [6]. Inherent complexities involved in capturing human behavior make surveys and interviews the best approach for understanding mode choice. However, most of the survey-based research

\* Address correspondence to the author at the California State University, Long Beach, California, U.S.A, 90803; Tel: +1-562-985-4654; fax: +1-562-985-2380; E-mail address: [shailesh.chandra@csulb.edu](mailto:shailesh.chandra@csulb.edu)

outcomes are limited to providing insights at the individual level, which may not be uniform across all potential transport users. Although the information obtained through surveys is very useful for understanding transportation problems that exist locally or within the area surveyed, the policy implications and impacts can only be empirically justified. Thus, a suitable choice modeling technique can be deployed to understand the determinants that govern complex human mode choice among a set of available options and variables. To simplify the analysis, most of the research in this field can be further subdivided into qualitative and quantitative studies. An in-depth understanding of both qualitative and quantitative aspects of people’s perceptions, attitudes, and behaviors toward choosing between a private car over public transport helps make policy decisions justifiable. While qualitative methods directly allow assessment and explanations of an individual transport user’s behavior and attitudes, quantitative methods draw implications and attitudes from a traveler’s mode choice based on statistical data analyses [7].

Example applications of crowdsourcing consist of individuals able to express their thoughts and concerns on a topic over a common platform while also contributing to a diverse group of people different ideas in hopes of solving one or more specific problems [8]. Fig. 1 shows the visual example

of how basic crowdsourcing is theorized to work, as well as some of the main steps that are typically involved in crowdsourcing [6]. This method of seeking a solution allows for a larger range of possible solutions being observed, while also increasing the participation of the public on a project. This method of participation assists many urban planners that tend to have difficulty in obtaining public involvement [9]. Furthermore, using crowdsourcing, one can make use of the information needed for planning decisions [10].

Crowdsourcing system types are divided into three types - based on participation expertise or no expertise and based on time and location (Fig. 2). Participant expertise is important as this elicits a response from actual transport mode users, while general participation is useful when expertise is not an issue; however, an informative idea is formed about the utility of a system using a general participation of non-experts. A crowdsourcing system can further be classified based on whether the information that is sought at the same place or if the participants are at different places. Audience-centric (audience played games) and geocentric systems (route choice data) are the same place sub-systems, while event-centric (event-based) and global systems (such as Wikipedia) are different places sub-systems.

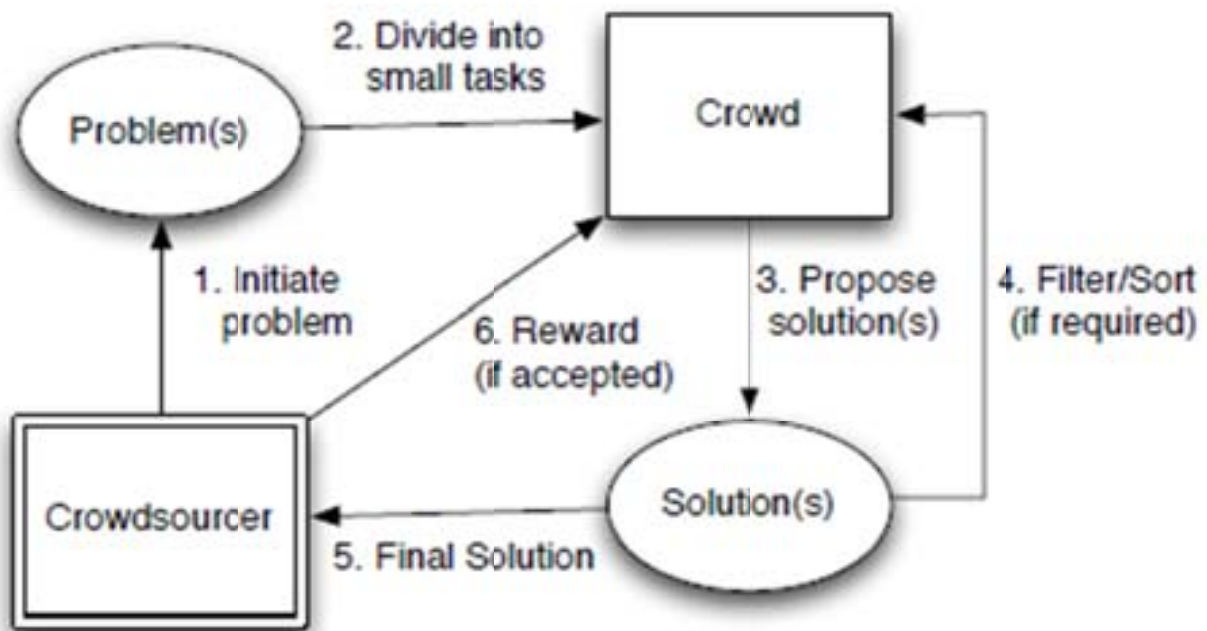


Fig. (1). Overview of Main Steps in Crowdsourcing [3].

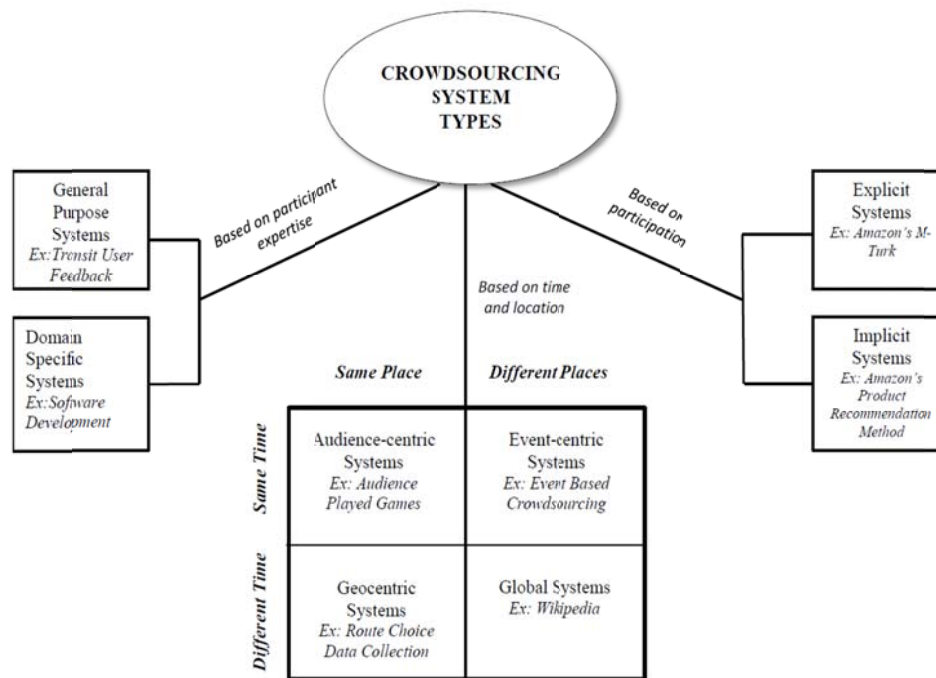


Fig. (2). Crowdsourcing systems and sub-systems [6, 11 - 13].

Cities tend to be perfect environments for crowdsourcing – since access to useful digital tools and individuals willing to share data are readily available [10]. The current usage of smartphones has increased to such an extent that by using all of the basic applications normally available in typical smartphones, the opportunity to obtain beneficial data and fill previous data gaps would increase substantially [1]. Some of the possible data which would be possible to obtain through the use of crowdsourcing would include travel behaviors of users, the current physical environment of the transportation system, improvement opportunities, public perceptions of new infrastructure projects, bicycle and foot trips, transportation demands within cities, cycling safety and routes, and contextual geographic information about current events in social media [6, 10, 14 - 17]. Other possible applications of crowdsourcing include smart parking, ridership data, transit troubleshooting, road condition monitoring, and assessment, urban traffic planning and management, and many other issues involving big data [10, 15, 18].

In order to fully understand mode shifts, it is also important to consider behavioral aspects of users in mode choice decisions [19, 20] This is in conjunction with a variety of other explanatory variables which constitute psychological, socioeconomic, and demographic factors in modeling [21, 22]. The inclusion of psychological factors pertaining to human behavior clearly makes the mode-choice modeling complex [23], primarily to improve the structure of the utility function which leads to improvement in the goodness of fit with the psychological factors in the mode-choice analyses [23, 24]. Proxies or dummy variables have often been used to include behavioral aspects of individuals involved in mode choice

decisions [25]. Osman Idris *et al.* [23] addressed this behavioral complexity by employing a “multivariate statistical modeling approach to investigate the causal relationships between the underlying psychological aspects affecting mode choices such as habit, attitude and affective factors” and then consolidating the approach with the Theory of Interpersonal Behavior by Triandis [26]. Determinants that are often causal for mode switch between private cars and public transport are dependent on the mode available, such as light rail [27], walking/bicycling [28] and bus [29]. Thus, in summary, extensive studies have been carried out to identify and determine factors that cause an individual’s preference for a private car over public transportation (and *vice versa*).

In this research, a mathematical model is developed using crowdsourcing as a social interaction tool to assess use of sustainable transportation modes such as walking, bicycling and public transit. The popularity of a mode is first modelled by gathering and evaluating opinions received from participants on a series of transportation modes. There are various ways of obtaining opinions such as those gathered through traditional blogs and/or using social media platforms (such as Twitter, Facebook *etc.*). In both these data gathering methods, there is an open mechanism of knowing opinions of others about a mode before presenting one’s own views for a mode choice. This enables others to gauge the most popular and commonly used mode to travel to a destination. Thus, a participant can access comments/opinions posted by others on the platform, arranged in chronological order. In a way, the process amounts to assisting an individual’s decision-making based on collective intelligence of the crowd [30].

Opinions shared socially on a common platform reflect one's preference from among the series of modes. In this process, there is a potential that a person's opinion might become influenced by a participant's prior experience or perception of the use of the mode and opinions of others on a mode. This illustrates decision-making skills of favoring or not favoring the use of a mode – using crowdsourcing in a social interaction setting.

## 2. MATERIALS AND METHODS

An analytical framework is developed for evaluating the popularity level of a mode. The popularity is based on positive or negative posts of the mode on a Twitter platform. Further, mode choice behavioral impacts are observed on participating individuals who read those posts. The following notations have been used for the actual mathematical formulation described below.

*X<sub>m</sub> Group Typology:* The individuals belonging to this group use mode *m* and post only positive comments on mode *m*. These individuals are rigid in posting their positive comments on mode *m*, irrespective of any negative opinions they read or notice about the mode posting on the Twitter page.

*Y<sub>m</sub> Group Typology:* The individuals belonging to this group are flexible in choosing a mode and post comments (positive or negative) based on the crowd's opinion noted through posts on the Twitter page. *N* = Number of participants posting comments,  $N = (X_m + Y_m)$ .

$\lambda$  = constant used as a buffer by the participants/individuals in the scoring mechanism.

$\alpha$  = the ratio of individuals in *Group X<sub>m</sub>* to the *Total number of participants N*, i.e.  $\alpha = \left(\frac{X_m}{N}\right) = \left(\frac{X_m}{X_m + Y_m}\right)$ . A very high value of  $\alpha$  above 0.5 would mean majority of the individuals are in favor of mode *m*. The fraction  $\alpha$  also indicates the ratio of positive comments from all the comments posted on mode *m*. Therefore,  $(1 - \alpha)$  is the fraction of posted comments that have a negative connotation on mode *m*.

$f_{Y_m}$  = average value of the quantified positive perception for mode *m* by *Y<sub>m</sub>* individuals

$f_{X_m}$  = average value of the quantified negative perception for mode *m* by *X<sub>m</sub>* individuals

A perception score is developed in this research based on a group's capacity and player's fitness theory adopted by Guazzini *et al.* [31] to solve an increasingly challenging task. In theory proposed by Guazzini *et al.*, for a given group, an integer parameter called the capacity is introduced, which is incremented successively by an integer if a task is solved. The basis was that there is an incremental nature of human advances, for which there is evidence of superlinear behavior through the chain of fitness gains. Thus, in this research, there are two mechanisms underlying the superlinear behavior - individual scheme related to the skills of providing positive/negative comments on the performance of a transportation mode in a given time and the accumulated knowledge developed from all the comments over the history.

In developing the formulation for the perception score, the

impact on the individual scheme and the capacity developed from the *Y<sub>m</sub>* target group is modeled for impact from comments from all the *N* individuals. Mode choice of individuals from the *X<sub>m</sub>* group is assumed to be rigid, irrespective of positive or negative opinions/comments posted by the participants.

Consider an incremental increase in capacity of an individual modeled for an initial time interval,  $t = 0$  and there is incremental integer increase of the time interval,  $t = 1, 2, 3, \dots$ , and so on. Within each time interval, all positive and negative comments from *N* individuals about a mode are posted and eventually read by all participating individuals. Within each known time interval, a set of new positive and negative comments from all the *N* individuals are posted on the Twitter page. Each time interval is assumed to be large enough to accommodate comments from all the *N* individuals. The comments are assumed to impact ridership and usage of mode *m* for the *N* individuals. A high number of positive comments for a mode would indicate a higher number of usage and popularity of the mode. After a sufficient number of positive and negative comments have been received, no more changes in mode choice or usage of a mode occur. This means perception is developed among all *N* individuals after some initial corrections or adjustments in their comments with respect to their mode choice.

In this paper, the individuals from the group are treated as the focus group, as it is assumed that any change in mode shift will occur from among this group in a social set-up influenced due to postings on a mode use experience crowdsourced over the Twitter page. Therefore, the goal is to evaluate if any individual from *Y<sub>m</sub>* would switch to use mode *m* because of being influenced due to the positive posts on the page. Alternatively, individuals from *Y<sub>m</sub>* might discontinue using mode *m* after being influenced by the negative posts by any of their peers from the *Y<sub>m</sub>* group.

If positive perception of mode *m*,  $f_{Y_m} > f_{X_m}$  negative perception of mode *m*, over a period of time, *Y<sub>m</sub>* individuals using mode *m* will increase.

Within a crowdsourcing set-up a certain group of individuals notice social behavior of their peers and are influenced from growing popularity of an object, commodity, *etc.* and thus, begin to align their perception/opinion/actions aligned with the majority of the crowd with time [6,32]. This basic idea is exploited in this research to build a framework of allocating quantitative scores to the choice of a mode by an individual, every time the individual from *X<sub>m</sub>* (or, *Y<sub>m</sub>*) sees a positive (or, negative) post about the mode. A positive post consolidates faith of the individuals who are already in favor of mode *m*, while a negative post may partially deteriorate the favorable opinions the individuals might have on mode *m*.

The first individual from the *Y<sub>m</sub>* group develops the intelligence capacity or perception score equal  $\alpha \times f_{Y_m}$  to which accounts for the positive perception of the individual for mode *m*. The perception could be developed before or after the posts on the Twitter page is read. Consequently, the quantitative perception equal to  $(1 - \alpha)(f_{X_m})$  accounts for the perception of the same first individual for the negative comments. The second individual from the *Y<sub>m</sub>* group builds a quantitative

perception of value equal to  $\alpha (f_{Y_{m,p}} + 1)$  and  $(1 - \alpha) (f_{Y_{m,n}} + 1)$  for positive and negative posts, respectively. This is because the post from the first individual from  $Y_m$  has influenced the perception of the second individual. Similarly, the  $n^{\text{th}}$  individual from  $Y_m$  builds a positive perception score equal to  $\alpha \{f_{Y_{m,p}} + (n - 1)\}$  for mode  $m$  for the positive posts, and for the

negative posts, a score of  $(1 - \alpha) \{f_{Y_{m,n}} + (n - 1)\}$  is assigned against mode  $m$ .

Therefore, the expression for the perception score,  $S_{m,n}$ , developed by an individual from  $Y_m$  individuals who is at the  $n^{\text{th}}$  observing position of posts for mode  $m$  is expressed as:

$$S_{m,n} = \begin{cases} f_{Y_{m,p}} + (n-1) & \text{if score is given for positive perception of mode } m \text{ by an individual from } Y_m \text{ at } n^{\text{th}} \text{ position} \\ f_{Y_{m,n}} + (n-1) & \text{if score is given for negative perception of mode } m \text{ by an individual from } Y_m \text{ at } n^{\text{th}} \text{ position} \end{cases} \quad (1)$$

Parameters  $f_{Y_{m,p}}$  and  $f_{Y_{m,n}}$  which are quantified perceptions of mode  $m$  from individuals belonging to  $Y_m$  group are assumed to be constant when equilibrium is reached and no more mode shift occurs irrespective of positive or negative posts from the  $N$  individuals. At the end of each time-period, the summation

of total score (termed as the crowd-based perception score,  $\Phi_m$ ) is obtained for the mode  $m$ .  $\Phi_m$  is a simple summation of  $S_{m,n}$  with  $n = 1, 2, \dots, Y_m$ , calculated across all the individuals in the group. The value of  $\Phi_m$  serves as the proxy of the overall popularity of mode  $m$ .

$$\begin{aligned} \Phi_m &= \alpha (f_{Y_{m,p}}) - (1 - \alpha) (f_{Y_{m,n}}) + \alpha (f_{Y_{m,p}} + 1) - (1 - \alpha) (f_{Y_{m,n}} + 1) + K + \alpha (f_{Y_{m,p}} + Y_m - 1) - (1 - \alpha) (f_{Y_{m,n}} + Y_m - 1) \\ &= \alpha Y_m f_{Y_{m,p}} - Y_m f_{Y_{m,n}} + \alpha Y_m f_{Y_{m,p}} - \frac{(Y_m - 1) Y_m}{2} + \alpha Y_m^2 - \alpha Y_m \end{aligned} \quad (2)$$

Considering  $Y_m \gg 1$ ,

$$\begin{aligned} \Phi_m &\approx \alpha (N - X) f_{Y_{m,p}} - (N - X) f_{Y_{m,n}} \\ &\quad + \alpha (N - X) f_{Y_{m,p}} - \frac{(N - X)^2}{2} + \alpha (N - X)^2 - \alpha (N - X) \\ &= \alpha N (1 - \alpha) f_{Y_{m,p}} - N (1 - \alpha) f_{Y_{m,n}} \\ &\quad + \alpha N (1 - \alpha) f_{Y_{m,p}} - \frac{N^2 (1 - \alpha)^2}{2} + \alpha N^2 (1 - \alpha)^2 - \alpha N (1 - \alpha) \\ \Rightarrow \Phi_m &\approx N \left\{ N \alpha^3 + \left( -f_{Y_{m,p}} - f_{Y_{m,n}} - \frac{5N}{2} + 1 \right) \alpha^2 + \left( f_{Y_{m,p}} + 2f_{Y_{m,n}} + 2N - 1 \right) \alpha - \left( f_{Y_{m,n}} + \frac{N}{2} \right) \right\} \end{aligned} \quad (3)$$

Thus, Eq. (3) yields a crowd-based perception score  $\Phi_m$  which is a cubic equation and depends on factors  $f_{Y_{m,p}}$ ,  $f_{Y_{m,n}}$  and variable  $\alpha$  (which is the ratio of number of positive posts by  $X_m$  individuals in favor of mode  $m$  over all the number of positive and negative posts posted by  $N$  individuals). The value of the perception score will be negative if  $Y_m$  individuals hold a very low positive perception and a high negative perception of the mode  $m$ . Thus, the value of  $\alpha$  at which the score  $\Phi_m$  attains a positive value would indicate that  $Y_m$  individuals might be shifting their opinion in favor of using mode  $m$ .

### 3. RESULTS

The application of the perception score is being illustrated using real life data collection and analysis. The goal of the application example is to understand at what instance a mode can potentially become popular among individuals. The empirical exercise is carried out to assess any mode shift

observed among four different transportation modes - car, transit bus, bicycling and walking - for 77 students from California State University, Long Beach (CSULB). These four modes are commonly used by CSULB students to arrive at the campus. The data collection was carried out spanning over a two-month period divided into two phases of one month each - Phase I and Phase II. Only those participants were included in this study who could easily choose more than one mode to arrive at the campus. Furthermore, participating students were required to be on campus at least once a week from Monday to Friday, own or can use a smartphone, and be able to either use personal car, public transport, walk, or bike to arrive at the campus.

The weather throughout the data collection effort was always perfect for using all four modes of transport. Therefore, the weather did not influence the mode choice of the participants. Phase I was carried out for the month of October of year 2018 and Phase II was carried out during November

and December of 2018. For both the phases, data collection was made only for the weekday travels of the participants to the CSULB campus. During Phase I, the mode choice of each participant was collected every Tuesday, Thursday, and Friday through emails and was not shared among the participants.

A random number of students were selected from the college of engineering at CSULB. To ensure maximum random sample of participants that were selected, no two participants belonged to the same level of study (freshman, sophomore, junior or senior). In addition, it was ensured that the usual arrival times to campus of the participants belonging to the same major differed by at least two to three hours for a given weekday. There was an approximately equal number of male and female students participating in this study.

During Phase II, each participant was provided with a short-encrypted name (with random four-letter first and also four-letter last name) to be used to post any tweets on a Twitter account created and managed specifically for this research. This was done to mask and protect the privacy of the participants from knowing each other while making any posts on the Twitter account. All unrelated, unwanted or differing posts other than the ten mentioned below, were deleted from the Twitter page as soon as they were found out by the researchers. A warning was also issued to the violator *via* email against any frivolous and unrelated postings other than those ten phrases. The researchers served as the administrator and the owner of the Twitter account. The students were asked to post short phrases on the controlled Twitter account for the mode used to reach the campus on weekdays (Monday-Friday) along with a score rating the mode. The students were provided these ten phrases, as mentioned below:

1. “Heavy traffic to campus” – and any similar phrase that would indicate the traffic scenarios is not conducive to driving
2. “Light traffic to campus” – and any similar phrase that would indicate the traffic is conducive to driving
3. “Difficulty in finding parking” – and any similar phrase that would indicate that the parking lots on campus are full
4. “Found parking” – and any similar phrase that would indicate that there were parking spaces available on campus
5. “Enjoyed bus ride” – and any similar phrase that would indicate easy access to transit bus to campus
6. “Bus ride was rough” – and any similar phrase that would indicate easy access to transit bus to campus
7. “Enjoyed biking” – and any similar phrase that would indicate that biking experience to the campus was a great

experience

8. “Biking was rough” – and any similar phrase which would indicate that biking experience to the campus was not a good experience

9. “Enjoyed walking” – and any similar phrase that would indicate that walking to campus was a good experience

10. “Walking woes” – and any similar phrase that would indicate that walking experience to campus was not a good experience

For the above ten phrases, positive and negative interpretations are tabulated in Table 1. Each phrase had at least one positive interpretation.

The ten phrases were selected such that they also included parking problems for participants driving to school, as expressed using phrase 4 “Found parking”. These phrases provided positive and negative experiences about a mode by a participant under Phase II. Thus, it was anticipated that this new information on a mode being used in real time would then affect the mode choice of other participants.

The data collected for Phase II occurred mostly during the month of November, along with the first week of December, which was done to offset the one week break for the Thanksgiving holiday.

Before the commencement of Phase I, participants were required to provide their positive and negative perception values for each day of the week through email. The positive perception accounted with  $f_{y_{sp}}$  and the negative perception with  $f_{y_{ns}}$ . Thus, a participant was asked to provide a score between lowest possible score of 0 to maximum possible score of 10 as a means for estimating  $f_{y_{sp}}$  and  $f_{y_{ns}}$ . The average values for  $f_{y_{sp}}$  and  $f_{y_{ns}}$  are shown in Table 2 in the beginning of Phase I. It was observed that for the two modes – car and walking – the positive perception of using these modes was higher than the negative perception of using the modes. The information is compiled in Table 2.

### 3.1. Empirical Results

The mode choice of participants was obtained under Phase I and Phase II and compiled for each duration of the phase, as shown in Table 3. The data in Table 3 under Phase I shows the mode split of the participants without any influence of crowdsourcing or outside knowledge of mode choice of other participants. The mode choice of car has the highest percentage use in while the bus was the second most used mode of transportation by the participants to arrive at CSULB campus during Phase I.

**Table 1. Mapping of phrases.**

Phrase No.	Phrase	Transit Bus	Bicycling	Walking	Car
1	Heavy traffic to campus	<i>P</i>	<i>P</i>	<i>P</i>	<i>N</i>
2	Light traffic to campus	<i>N</i>	<i>N</i>	<i>N</i>	<i>P</i>
3	Difficulty in finding parking	<i>P</i>	<i>P</i>	<i>P</i>	<i>N</i>
4	Found parking	<i>N</i>	<i>N</i>	<i>N</i>	<i>P</i>
5	Enjoyed bus ride	<i>P</i>	<i>N</i>	<i>N</i>	<i>N</i>
6	Bus ride was rough	<i>N</i>	<i>P</i>	<i>P</i>	<i>P</i>

(Table 1) cont....

Phrase No.	Phrase	Transit Bus	Bicycling	Walking	Car
7	Enjoyed biking	N	P	N	N
8	Biking was rough	P	N	P	P
9	Walked to campus	N	N	P	N
10	Walking woes	P	P	N	P

P = positive; N = negative

Table 2. Parameter values obtained in the beginning of Phase I.

Mode (m)	Phase I	
	$f_{Y_{m,p}}$	$f_{Y_{m,n}}$
Car	25	5
Transit Bus	12	24
Bicycling	15	17
Walking	16	7

Table 3. Comparing Phase I and II mode use.

Mode (m)	Phase I		Crowd-based Perception Score ( $\Phi_m$ )	Crowdsourcing Opinions on Modes	Phase II	
	Value of $\alpha$	Total Number of Users			Total Number of Users	Percentage of Users
Transit Bus	0.23	18	-505		18	0.23
Bicycling	0.04	3	-834		3	0.04
Walking	0.04	3	-570		13	0.17
Car	0.69	53	163		43	0.55

The chart in Fig. 3 shows the variation of crowd-based perception score,  $\Phi_m$ , versus percentage of users for the four modes. It is noted that as the value of  $\Phi_m$  increases, there is also an increase in the number of users for the modes. Car has the largest value for  $\Phi_m$ , with the highest percentage of users, while bicycling has the smallest  $\Phi_m$  with the lowest percentage of users. Thus, this empirical analysis serves as a validation for the crowd-based perception score developed in this research. City planners and stakeholders can improve ridership and user frequencies of sustainable transportation modes (such as transit, bicycling etc.) by utilizing the crowdsourcing technique of soliciting public opinions, which are transparent and are on specific infrastructure facilities. Furthermore, using the crowd-

based perception score developed in this research, the future popularity of the modes can also be assessed.

4. DISCUSSION

The crowd-based perception score  $\Phi_m$  can be used to make decisions regarding the popularity of a mode. Higher the value of  $\Phi_m$  the higher the popularity of the mode  $m$  is among  $Y_m$  individuals. With a varying value for  $\alpha$ , a closed-form expression for optimal  $\Phi_m$  can be obtained using 1<sup>st</sup> and 2<sup>nd</sup> derivatives of  $\Phi_m$ .

The 1<sup>st</sup> derivative of  $\Phi_m$  with respect to  $\alpha$  is expressed as,

$$\frac{d\Phi_m}{d\alpha} = N \left\{ 3N\alpha^2 + 2 \left( -f_{Y_{m,p}} - 2\lambda - f_{Y_{m,n}} - \frac{5N}{2} + 1 \right) \alpha + \left( f_{Y_{m,p}} + 3\lambda + 2f_{Y_{m,n}} + 2N - 1 \right) \right\}$$

For  $\frac{d\Phi_m}{d\alpha} = 0$ ,

$$\Rightarrow \alpha = \frac{\left( f_{Y_{m,p}} + 2\lambda + f_{Y_{m,n}} + \frac{5N}{2} - 1 \right) - \sqrt{\left( f_{Y_{m,p}} + 2\lambda + f_{Y_{m,n}} + \frac{5N}{2} - 1 \right)^2 - 3N \times \left( f_{Y_{m,p}} + 3\lambda + 2f_{Y_{m,n}} + 2N - 1 \right)}}{3N}$$

(4)



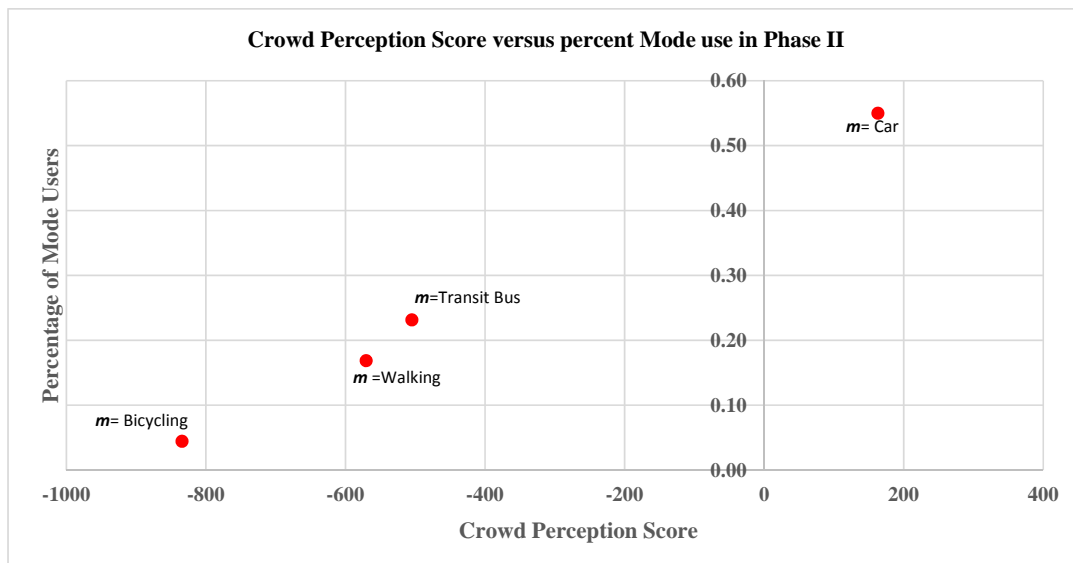


Fig. (3). Variation of percentage mode users with crowd perception score for Phase II.

Although two roots are possible for the quadratic expression resulting from the 1<sup>st</sup> derivative, the other root does not maximize the perception score as shown in the Appendix. The second derivative of  $\Phi_m$  with respect to  $\alpha$  gives,

$$\frac{d^2\Phi_m}{d\alpha^2} = N \left\{ 6N\alpha + 2 \left( -f_{Y_{m,p}} - 2\lambda - f_{Y_{m,n}} - \frac{5N}{2} + 1 \right) \right\} \quad (5)$$

As evident from Eq. (4), in order to maximize the crowd-based perception score  $\Phi_m$  for given mode  $m$ , it is observed  $0 \leq \alpha \leq 1$  that and the second derivative of  $\Phi_m$  with respect to  $\alpha$  is a negative number. This gives a very useful information on developing an understanding of the impact of  $\alpha$  on popularizing a mode based on the technique of crowdsourcing. A summary of information presented in Table 4 provides conditions on  $\alpha$  to maximize  $\Phi_m$  for a given mode  $m$ . Detailed derivations for the summary in Table 4 are shown in the Appendix.

Table 4. Compilation of maximum scores obtained for a mode (see Appendix for derivations).

Value of	Inequality that needs to be satisfied	Maximum Value of Crowd-based Perception Score,
$\frac{l - \sqrt{l^2 - 3N \times m}}{3N}$	$l^2 \geq 3N \times m$	$\left( \frac{l^3 + (l^2 - 3mN)(2\sqrt{l^2 - 3N \times m} - 3l)}{27N} \right) - N \left( \lambda + f_{Y_{m,n}} + \frac{N}{2} \right)$
where $l = \left( f_{Y_{m,p}} + 2\lambda + f_{Y_{m,n}} + \frac{5N}{2} - 1 \right)$ and $m = \left( f_{Y_{m,p}} + 3\lambda + 2f_{Y_{m,n}} + 2N - 1 \right)$		

Note that the above optimization exercise can work only

with a set of individuals from  $Y_m$  that tend to follow the crowd in decision-making. In Table 4, variables  $f_{Y_{m,p}}$ ,  $f_{Y_{m,n}}$  and  $\alpha$  are considered as components of crowdsourcing. Approximate values of and can be determined using survey findings before individual scores from  $X_m$  and  $Y_m$  are sought on the perceptions of mode  $m$ . Perceptions about the use of a mode can be negative or positive. Every individual in  $X_m$  provides an initial perception rating for a mode from 0 to  $N$ , and the higher the rank more inclined is the individual to use the mode. The average of those ranks across all  $X_m$  (or  $Y_m$ ) individuals gives the average range  $\Phi_m$  for  $f_{Y_{m,p}}$  (or  $f_{Y_{m,n}}$ ) which is  $(, N)$ . The analytical model developed in this paper can be used to determine if there could be future potential for enhancing the popularity of sustainable transportation modes such as transit, bicycling etc.

4.1. Sensitivity Analysis with Crowdsourcing Components

An example illustration is presented in the chart of (Fig. 4) to show the impact on crowd-based perception score  $\Phi_m$  of a mode  $m$  for assumed values  $\alpha$  varying from 0 to 1 at an interval of 0.02. The chart is for  $N = 100$ , with assumed values of parameters  $f_{Y_{m,p}}$  and  $f_{Y_{m,n}}$  as shown in Table 2. The values of  $f_{X_m}$  show a decreasing trend of the perception of using mode  $m$  (and the values of show an increasing trend towards the perception of not using the same mode  $m$ ) across all four scenarios shown in Table 5. Scenarios A and B show that  $Y_m$  individuals have a much higher negative perception and low positive perception of using the mode, which indicates that the perception scores will be quite low. Scenarios C and D show that  $Y_m$  individuals have a much higher positive perception and low negative perception for not using the mode, which indicates that the perception scores will be quite high for these two scenarios.



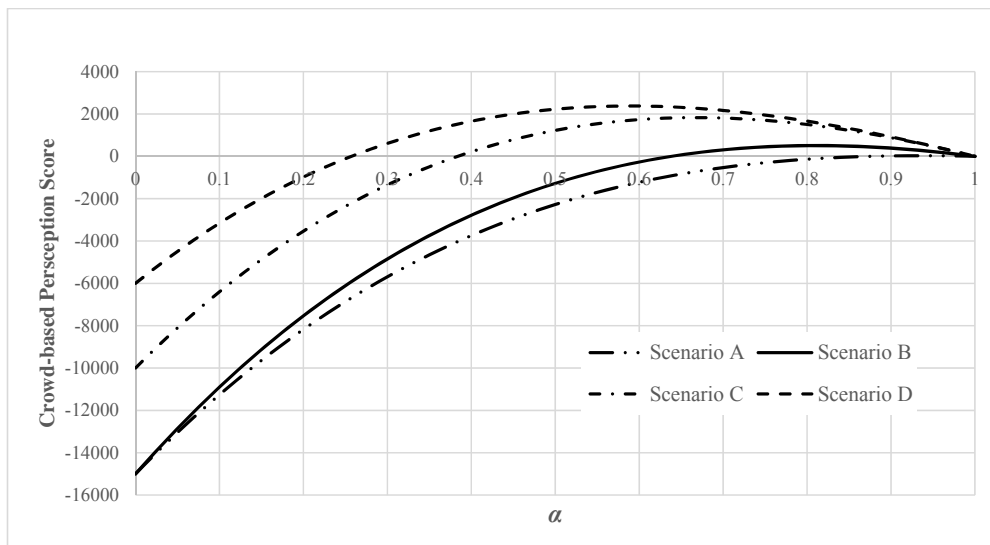


Fig. (4). Variation of the perception score versus  $\alpha$ .

Table 5. Inputs used for sensitivity analysis.

Scenario	Crowdsourcing Parameters		
	$f_{Y_m}$	$f_{X_m}$	$\alpha$
A	10	100	0 to 1 (interval of 0.02)
B	50	100	0 to 1 (interval of 0.02)
C	100	50	0 to 1 (interval of 0.02)
D	100	10	0 to 1 (interval of 0.02)

The observations from the chart in Fig. 4 show the potential of any enhancement in the perception score of mode  $m$  among  $Y_m$  individuals. Perception score is used as a proxy to popularity, attitude or perception of mode  $m$ , which indicates that at  $\alpha = 0.89, 0.63, 0.38$  and  $0.26$ , negative perception of using the mode changes to a positive one for Scenarios, A, B, C and D, respectively.

Scenarios A and B show that with a negative perception of using mode  $m$  (*i.e.*) being the maximum, Scenario B attains a positive perception score at  $\alpha = 0.63$  (with lower  $X_m$ , higher  $Y_m$ ) before Scenario A which attains a positive perception score at  $\alpha = 0.89$  (with higher  $X_m$ , lower  $Y_m$ ). This is due to the positive perception of using mode  $m$  (*i.e.*  $f_{Y_m}$ ) being higher for Scenario B as compared to that of Scenario A. This is expected since a higher positive perception of a mode will accelerate its acceptance among individuals who have an overall negative perception.

The positive perception ( $f_{Y_m}$ ) of using mode  $m$  is the maximum possible for both Scenarios C and D, while the negative perception ( $f_{X_m}$ ) of Scenario D is lower as compared to Scenario C. Thus, a positive value of perception score  $\Phi_m$  for Scenario D is reached even at a lower value of  $\alpha = 0.26$  (with lower  $X_m$ , higher  $Y_m$ ) before Scenario C which has a positive value for the perception score at  $\alpha = 0.38$  (with higher  $X_m$ , lower  $Y_m$ ). The expected score signifying the potential of acceptance and using mode  $m$  is higher in case of Scenario D

for all values of  $\alpha$  when compared to Scenario C.

The sensitive analysis carried out can be extremely useful for transit agencies and managers in understanding perception levels of transit from among potential users who do use transit at all. Thus, the methodology of crowdsourcing illustrated above can be used to successfully assess the level of popularity of transit that can be achieved by increasing its ridership.

CONCLUSION

Crowdsourcing is emerging as a powerful tool in transportation, particularly for travel management and routing decisions. For example, Cyclopath (which is a geo-wiki where bicycle users in Minnesota share a note about bike lane and trail conditions on an editable map) is being used to crowdsource information about missing parts or trails on a lane to fellow bicyclists [33]. Other examples include smartphone-based applications, such as Google Maps, which provide dynamic routes to roadway users by crowdsourcing [34].

This paper provides an application of a model with a crowd-based perception score developed. The perception score is used to study any potential mode-shift behaviour among college students soliciting comments and opinions on mode used to arrive at the CSULB campus. Based on the outcome of this study, albeit for a small sample size, it is shown that the crowd-based perception score can potentially predict the future ridership to a certain extent. The percentage of various transportation mode users with intervention provided using crowdsourcing increased with an increase in the value of the crowd-based perception score.

The findings of this research can be further validated by increasing the participant pool in this crowdsourcing exercise. The results clearly will have some very wide-spread implications beyond the college settings in popularizing mode-shifts to transit and other active transportation modes if appropriate social media and information sharing mechanisms

through crowdsourcing are provided to transport users. However, there are limitations on the use of crowdsourcing as a technique for data collection purposes [35]. In situations when a problem that needs to be addressed is not clearly defined, crowdsourcing may not occur [36]. There may also be issues related to acquiring and integrating unsolicited ideas with crowdsourcing [37]. Therefore, these limitations need to be kept in mind before conducting a full-fledged data collection exercise in the expectation of influencing mode choice.

**CONSENT FOR PUBLICATION**

Not applicable.

**AVAILABILITY OF DATA AND MATERIALS**

The authors confirm that the data supporting the findings of this study are available within the article.

**FUNDING**

None.

**CONFLICT OF INTEREST**

The authors declare no conflict of interest, financial or otherwise.

**ACKNOWLEDGEMENTS**

Declared none.

**APPENDIX**

Optimal crowd-based perception score  $\Phi_m$

A.1. Evaluation for maximum with  $\Phi_m$  with  $0 \leq \alpha \leq 1$

(i) Given,

$$\alpha = \frac{\left(f_{i,c} + 2\lambda + f_{i,c} + \frac{5N}{2} - 1\right) - \sqrt{\left(f_{i,c} + 2\lambda + f_{i,c} + \frac{5N}{2} - 1\right)^2 - 3N \times (f_{i,c} + 3\lambda + 2f_{i,c} + 2N - 1)}}{3N}$$

For  $\alpha \geq 0$ ,  $\Rightarrow \left(f_{i,c} + 2\lambda + f_{i,c} + \frac{5N}{2} - 1\right) - \sqrt{\left(f_{i,c} + 2\lambda + f_{i,c} + \frac{5N}{2} - 1\right)^2 - 3N \times (f_{i,c} + 3\lambda + 2f_{i,c} + 2N - 1)} \geq 0$

$\Rightarrow 3N \times (f_{i,c} + 3\lambda + 2f_{i,c} + 2N - 1) \geq 0$ , which is always true since factors  $f_{i,c}, f_{i,c}, \lambda$  and  $N$  are by design and assumptions all non-negative numbers.

• For  $\alpha \leq 1$ ,

$$\left(f_{i,c} + 2\lambda + f_{i,c} + \frac{5N}{2} - 1\right) - \sqrt{\left(f_{i,c} + 2\lambda + f_{i,c} + \frac{5N}{2} - 1\right)^2 - 3N \times (f_{i,c} + 3\lambda + 2f_{i,c} + 2N - 1)} \leq 3N$$

$\Rightarrow (f_{i,c} + \lambda) \geq 1$ ,

Now using the second derivative,

$$\frac{d^2 \Phi_m}{d\alpha^2} = N \left\{ 6N\alpha + 2 \left( -f_{i,c} - 2\lambda - f_{i,c} - \frac{5N}{2} + 1 \right) \right\}$$

$$\frac{d^2 \Phi_m}{d\alpha^2} = -2N \left\{ 2 \sqrt{\left(f_{i,c} + 2\lambda + f_{i,c} + \frac{5N}{2} - 1\right)^2 - 3N \times (f_{i,c} + 3\lambda + 2f_{i,c} + 2N - 1)} \right\} < 0$$

Therefore, maximum  $\Phi_m$  is possible under this inequality and the other root actually minimizes  $\Phi_m$ , which is not the focus of this study.

and

$$\alpha = \frac{\left(f_{i,c} + 2\lambda + f_{i,c} + \frac{5N}{2} - 1\right) - \sqrt{\left(f_{i,c} + 2\lambda + f_{i,c} + \frac{5N}{2} - 1\right)^2 - 3N \times (f_{i,c} + 3\lambda + 2f_{i,c} + 2N - 1)}}{3N}$$

Assuming,  $I = \left(f_{i,c} + 2\lambda + f_{i,c} + \frac{5N}{2} - 1\right)$  and

$$m = (f_{i,c} + 3\lambda + 2f_{i,c} + 2N - 1)$$

gives,  $\alpha = \frac{I - \sqrt{I^2 - 3N \times m}}{3N}$

Using the expression for  $\Phi_m$  in Eq. (3)

$\Rightarrow \max(\Phi_m)$

$$-N \left\{ N \left( \frac{I - \sqrt{I^2 - 3N \times m}}{3N} \right)^2 - I \left( \frac{I - \sqrt{I^2 - 3N \times m}}{3N} \right) + m \left( \frac{I - \sqrt{I^2 - 3N \times m}}{3N} \right) - \left( \lambda + f_{i,c} + \frac{N}{2} \right) \right\}$$

$$- \left\{ \frac{I^3 + (I^2 - 3mN)(2\sqrt{I^2 - 3N \times m} - 3I)}{27N} - N \left( \lambda + f_{i,c} + \frac{N}{2} \right) \right\}$$

**REFERENCES**

- [1] G. Chatzimilioudis, A. Konstantinidis, C. Laoudias, and D. Zeinalipour-Yazti, "Crowdsourcing with smartphones", *IEEE Internet Comput.*, vol. 16, no. 5, pp. 36-44, 2012. [http://dx.doi.org/10.1109/MIC.2012.70]
- [2] H. Gao, G. Barbier, and R. Goolsby, "Harnessing the crowdsourcing power of social media for disaster relief", *IEEE Intell. Syst.*, vol. 26, no. 3, pp. 10-14, 2011. [http://dx.doi.org/10.1109/MIS.2011.52]
- [3] K. Ali, D. Al-Yaseen, A. Ejaz, T. Javed, and H.S. Hassanein, "Crowdits: Crowdsourcing in intelligent transportation systems," in *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, IEEE, 2012, pp. 3307-3311.
- [4] N. Lewis, "Experiences of upward social comparison in entertainment contexts: Emotions, state self-esteem, and enjoyment", *Soc. Sci. J.*, pp. 1-16, 2020. [http://dx.doi.org/10.1016/j.sosocij.2019.04.011]
- [5] L. Iniguez, J. Jimenez, and S. Chandra, "Investigating impact of crowdsourcing on mode-shift behavior of college students",
- [6] A. Misra, A. Gooze, K. Watkins, M. Asad, and C.A. Le Dantec, "Crowdsourcing and its application to transportation data collection and management", *Transp. Res. Rec.*, vol. 2414, no. 1, pp. 1-8, 2014. [http://dx.doi.org/10.3141/2414-01]
- [7] G. Beirão, and J.S. Cabral, "Understanding attitudes towards public transport and private car: A qualitative study", *Transp. Policy*, vol. 14, no. 6, pp. 478-489, 2007. [http://dx.doi.org/10.1016/j.tranpol.2007.04.009]
- [8] P. Rabinowitz, *Participatory approaches to planning community interventions.*, The Community Tool Box, 2013.
- [9] D.C. Brabham, "Crowdsourcing the public participation process for planning projects", *Plann. Theory*, vol. 8, no. 3, pp. 242-262, 2009. [http://dx.doi.org/10.1177/1473095209104824]
- [10] C. Ferster, T. Nelson, K. Laberee, W. Vanlaar, and M. Winters, "Promoting crowdsourcing for urban research: Cycling safety citizen science in four cities", *Urban Science*, vol. 1, no. 2, p. 21, 2017. [http://dx.doi.org/10.3390/urbansci1020021]
- [11] A. Doan, R. Ramakrishnan, and A.Y. Halevy, "Crowdsourcing Systems on the World-Wide Web", *Commun. ACM*, vol. 54, no. 4, pp. 86-96, 2011. [http://dx.doi.org/10.1145/1924421.1924442]
- [12] A. Steinfeld, J. Zimmerman, and A. Tomasic, Bringing Customers back into transportation: citizen-driven transit service innovation via social computing. *Best practices for transportation agency use of social media.*, CRC Press: Boca Raton, Florida, 2013.
- [13] T. Erickson, "Geocentric Crowdsourcing and Smarter Cities: Enabling Urban Intelligence in Cities and Regions", *A Position Paper for the 1st International Workshop on Ubiquitous Crowdsourcing*, 2010 Copenhagen, DenmarkUbiComp
- [14] O. Contributors, *Open Street Map.*, Geofabrik GmbH: Karlsruhe, Germany, 2017.
- [15] B. Jestic, T. Nelson, and M. Winters, "Mapping ridership using

- crowdsourced cycling data", *J. Transp. Geogr.*, vol. 52, pp. 90-97, 2016.  
[<http://dx.doi.org/10.1016/j.jtrangeo.2016.03.006>]
- [16] T.A. Nelson, T. Denouden, B. Jestico, K. Laberee, and M. Winters, "BikeMaps.org: A global tool for collision and near miss mapping", *Front. Public Health*, vol. 3, p. 53, 2015.  
[<http://dx.doi.org/10.3389/fpubh.2015.00053>] [PMID: 25870852]
- [17] C.A. Le Dantec, M. Asad, A. Misra, and K.E. Watkins, "Planning with crowdsourced data: rhetoric and representation in transportation planning", [<http://dx.doi.org/10.1145/2675133.2675212>]
- [18] Y. Sun, Y. Moshfeghi, and Z. Liu, "Exploiting crowdsourced geographic information and GIS for assessment of air pollution exposure during active travel", *J. Transp. Health*, vol. 6, pp. 93-104, 2017.  
[<http://dx.doi.org/10.1016/j.jth.2017.06.004>]
- [19] T. Gärling, and K.W. Axhausen, "Introduction: Habitual travel choice", *Transportation*, vol. 30, no. 1, pp. 1-11, 2003.  
[<http://dx.doi.org/10.1023/A:1021230223001>]
- [20] R.L. Mackett, "Why do people use their cars for short trips?", *Transportation*, vol. 30, no. 3, pp. 329-349, 2003.  
[<http://dx.doi.org/10.1023/A:1023987812020>]
- [21] T. Morikawa, "Correcting state dependence and serial correlation in the RP/SP combined estimation method", *Transportation*, vol. 21, no. 2, pp. 153-165, 1994.  
[<http://dx.doi.org/10.1007/BF01098790>]
- [22] V. Cantillo, J.D. Ortúzar, and H.C. Williams, "Modeling discrete choices in the presence of inertia and serial correlation", *Transport. Sci.*, vol. 41, no. 2, pp. 195-205, 2007.  
[<http://dx.doi.org/10.1287/trsc.1060.0178>]
- [23] A. Osman Idris, K.M.N. Habib, A. Tudela, and A. Shalaby, "Investigating the effects of psychological factors on commuting mode choice behaviour", *Transp. Plann. Technol.*, vol. 38, no. 3, pp. 265-276, 2015.  
[<http://dx.doi.org/10.1080/03081060.2014.997451>]
- [24] C. Domarchi, A. Tudela, and A. González, "Effect of attitudes, habit and affective appraisal on mode choice: An application to university workers", *Transportation*, vol. 35, no. 5, pp. 585-599, 2008.  
[<http://dx.doi.org/10.1007/s11116-008-9168-6>]
- [25] M.V. Johansson, T. Heldt, and P. Johansson, "The effects of attitudes and personality traits on mode choice", *Transp. Res. Part A Policy Pract.*, vol. 40, no. 6, pp. 507-525, 2006.  
[<http://dx.doi.org/10.1016/j.tra.2005.09.001>]
- [26] H.C. Triandis, "Values, attitudes, and interpersonal behavior," in *Nebraska symposium on motivation.*, University of Nebraska Press, 1979.
- [27] L. Creemers, M. Cools, H. Tormans, P-J. Lateur, D. Janssens, and G. Wets, "Identifying the determinants of light rail mode choice for medium-and long-distance trips: Results from a stated preference study", *Transp. Res. Rec.*, vol. 2275, no. 1, pp. 30-38, 2012.  
[<http://dx.doi.org/10.3141/2275-04>]
- [28] C. L. Antonakos, "Environmental and travel preferences of cyclists",
- [29] K. Chatterjee, "Modelling the dynamics of bus use in a changing travel environment using panel data", *Transportation*, vol. 38, no. 3, pp. 487-509, 2011.  
[<http://dx.doi.org/10.1007/s11116-010-9312-y>]
- [30] D.R. Wilson, *Hand book of Collective Intelligence*  
[<http://dx.doi.org/10.1016/j.soscij.2017.10.004>]
- [31] A. Guazzini, D. Vilone, C. Donati, A. Nardi, and Z. Levnajić, "Modeling crowdsourcing as collective problem solving", *Sci. Rep.*, vol. 5, no. 1, p. 16557, 2015.  
[<http://dx.doi.org/10.1038/srep16557>] [PMID: 26552943]
- [32] X. Hu, "Social drive: a crowdsourcing-based vehicular social networking system for green transportation", *Proceedings of the third ACM international symposium on Design and analysis of intelligent vehicular networks and applications*, 2013pp. 85-92  
[<http://dx.doi.org/10.1145/2512921.2512924>]
- [33] K. Panciera, R. Priedhorsky, T. Erickson, and L. Terveen, "Lurking? Cyclopaths? A quantitative lifecycle analysis of user behavior in a geowiki", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2010pp. 1917-1926
- [34] Google Account Help, Google, Accessed on February 2016  
<https://support.google.com/accounts/answer/3467281?hl=en>[35]
- [35] T. Burger-Helmchen, and J. Pénin, "The limits of crowdsourcing inventive activities: What do transaction cost theory and the evolutionary theories of the firm teach us", *Workshop on Open Source Innovation*, 2010pp. 1-26 Strasbourg, France
- [36] O. Alexy, A. Salter, and P. Criscuolo, "No soliciting: Strategies for managing unsolicited innovative ideas (August 16, 2010)", Available at SSRN 1659749 (2010).