A Reputation-based Shared Transport System. A Case Study of E-collaboration in the City of Ambato

Ana Sánchez-Paredes, Félix Fernández-Peña, Pilar Urrutia-Urrutia, Hernán Naranjo-Ávalos
Facultad de Ingeniería en Sistemas, Electrónica e Industrial
Universidad Técnica de Ambato
Ambato, Ecuador
{asanchez6019, fo.fernandez, elsapurrutia, hf.naranjo}@uta.edu.ec

***Abstract*—The governments of cities like Ambato face difficult geographic conditions for establishing a clean and efficient transportation system. This paper shows the results of 1) a theoretical study that measured the expected incidence of a reputation-based shared transportation system for e-collaboration and social welfare in a medium-size community and 2) an experimental evaluation of the effectiveness of a value-based reputation system with malicious user detection in such a scenario of e-collaboration. The study was conducted based on economic savings. The theoretical study involved the participation of 185 young citizens and 160 in the experimental evaluation. The theoretical results show that e-collaboration is meant to be successful in a shared transportation system in Ambato providing that trust is guaranteed. The experiment results show that the proposed model made it possible to detect 85% of malicious service consumers and that the satisfaction with the recommendations of the system is pretty high. These results are encouraging for using the proposed method in the implementation of a shared transport system in medium-size cities. Therefore, it is assumed that the obtained results are extrapolates to the context of a generic e-collaborative system.**

***Index Terms*—e-collaboration, green government, values-based reputation system, malicious user detection.**

I. Introduction

Road transportation emissions and traffic jams are a problem that cities with a tendency toward population growth must mitigate. The situation in the city of Ambato gets worse because of its irregular topography [1]. Also, 34.45% of the interviewed population in Ambato use private vehicles as their means of transportation and 54.86% of them drive alone, according to statistics from the National Institute of Statistics and Census (INEC) in 2016 [2]. This medium-sized city has the worst situation in Ecuador in comparison to Quito, Guayaquil and Cuenca which are bigger cities but the people share more. The INEC statistics for this factor show values of 45.72%, 46.27%, 30.51% for these other cities, respectively [2].

In the last decade, the means of transportation have evolved, including e-collaborative transportation. The use of this technology allows a user to request, enjoy and evaluate a transportation service through a mobile device [3]. The most famous transportation apps in app stores are Uber with a hundred million downloads, Ola Cabs with fifty million downloads, Cabify with five million downloads, as well as Yaxi and TaxiBeat with a million downloads each. Only with these apps, there are more than 150 million users of e-collaborative transportation around the world, demonstrating the acceptance of this kind of service. The use of interactive maps in these applications enriches the user experience by rendering routes on the map and showing the actual position of vehicles in service.

The successful implementation of a shared transportation application in Ambato is expected to give similar benefits to those witnessed in other cities of a similar size around the world like those described by Cheyne et al. [4]. However, the development of this service in this place faces the challenge of any medium-size city in the Ecuadorean highlands - since the citizens belong to a traditional society, they resist change.

The aim of this study is to explore the possibilities of a reputation-based shared transportation system in the case of this scenario from a theoretical perspective and practice. The system is expected to recommend services taking into account the profile of each user while decreasing uncertainty in the e-collaborative environment. The outlined method reputation conducts, as a starting point, a survey with young local citizens and it is based on the state-of-the-art in recommendation systems that was found in the literature. The proposal was validated with a practical case study using data from a metaphoric scenario already accepted by potential users of a shared transportation system.

This paper is structured as follows. In Section II, related works are analyzed. Next, the proposed method is outlined for adapting an implementation of a Values-Based Reputation System (VRS) to a case study of interest. In Section IV, the design of the experiment that was carried out is described while its results are discussed in Section V. Finally, the conclusions of the case study and future work are depicted.

978-1-5386-2521-7/18/$31.00

II. RELATED WORK

According to Jnanamurthy et al. [5], a reputation system is a tool to facilitate trust between entities, as it increases the efficiency and effectiveness of online services and communities. Moreover, reputation systems have been implemented in diverse e-collaborative initiatives [6], [7], social networks [8]–[10], and e-commerce [11]–[14]. In a growing community of any of these platforms, the preservation of trust demands reputation systems to be constantly optimized, so that the community can be protected from the various types of fraud and unfair competition [6], [13].

There are many parameters and approaches for reputation systems to obtain the best results in different scenarios of interest. The parameters and approaches such as Gains (P1), Behavior of users (P2), QoS Parameter (P3), Statistical computes (P4), Vote-based (P5), Service Feedback (P6), Sybil attackers (P7), Trust degree (P8) and Scarcity feedback (P9) are part of the suggestions from various authors, as shown in Table 1.

TABLE I
PARAMETERS AND APPROACHES TO COMPUTE REPUTATION

Work	P1	P2	P3	P4	P5	P6	P7	P8	P9
Y. Wang [12]	X	X			X				
W. Zhang [10]	X		X	X	X	X			
M. Pouryazdan [9]				X	X				
X. Zhou [8]			X	X	X	X			X
S. Cuomo [6]				X	X	X	X	X	
A. Panagopoulos [13]	X	X			X		X		
F.-H. Hsu [15]		X			X		X		

According to the literature review that was carried out, 75% of the authors propose statistic measures like *mean*, *median* and *mode* for measuring reputation. In all of these studies, the use of a vote-based feedback mechanism has been taken into account. In the particular case of the use of the Page Rank algorithm, by Zhang et al. [10], the number of interactions of *service consumers* with provided services are taken into account. Zhou et al. implemented a hybrid model which includes implicit and explicit rating. Their proposal is based on the *Elo Algorithm* and it is focused on solving the rating scarcity problem [8].

Panagopoulos et al. proposed something similar but it is called public and private reputation using the feedback mechanism of eBay [13]. Hsu et al. and Zhou et al. consider that user preferences are important to confirm results at the time of ranking for each entity [8], [15]. Hsu et al. and Jnanamurthy et al. determine users' similarities to provide a more accurate reputation calculation [5], [15].

Jnanamurthy used Cosine similarity for detecting malicious users in his proposal of a system for quantifying the reputation of *service providers* [5]. However, Hsu et al. findings show that VRS offers better results than Cosine Similarity [15].

As a result of the literature review, the VRS algorithm was chosen as a robust candidate for implementing a shared transportation system in Ambato. This decision is based on its capacity to resist attacks from malicious users and its effectiveness to give recommendations based on users' preferences [15]. Nevertheless, there are malicious user behaviours that the VRS algorithm is unable to detect. This study is focused on filling in the gap between VRS and malicious user behavior detection.

III. OUTLINE OF METHOD

Previous to the definition of the proposed method, a survey was applied in order to uncover the factors that determine the acceptance of a shared transportation system in Ambato. Survey items were mostly about which factors determine the acceptability of a shared transportation system, by *service consumers* and which would be the most acceptable fee for the service among drivers and passengers. The value of a Cronbach Alpha of 0.843 in a pilot test confirmed the internal stability of the questionnaire leading to a continuation of the analysis of the data collected.

185 young local citizens participated in the theoretical study. 98% of the respondents are under 25 years old and 55% are men. The results showed that *trust* and *safety* are the most important factors, higher than time needed to travel, fee, time spent waiting and profit, as shown in figure 1 and figure 2. Thus, the goal of this study was set to the detection of malicious *service consumers* in a shared transportation system that implements VRS.

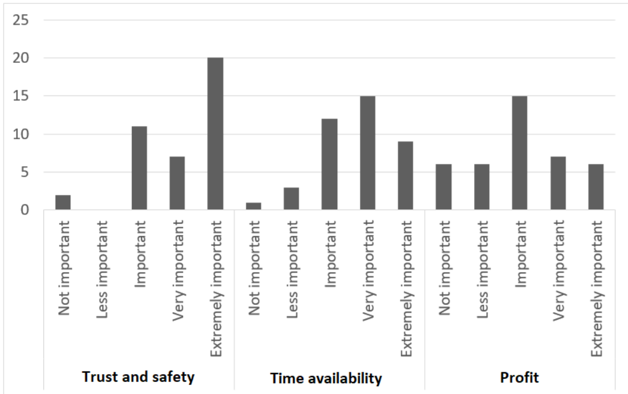

Fig. 1. Acceptance factors for drivers

The reputation should be calculated based on honest *service consumers*'s feedback for a positive collaborative community [5], [15] and this model proposes to ensure this. The procedure of the proposed reputation system is depicted in figure 3 and explained below. Its architecture is graphically represented in figure 4.

A. Information gathering

Similar to TripAdviser [15] and others, the reputation system that has been proposed implements a five-star mechanism for voting for service quality. In mathematical notation, this means a five-level cardinal rating metric where each rating level is associated with a quantitative value. The values are defined in a Likert scale as

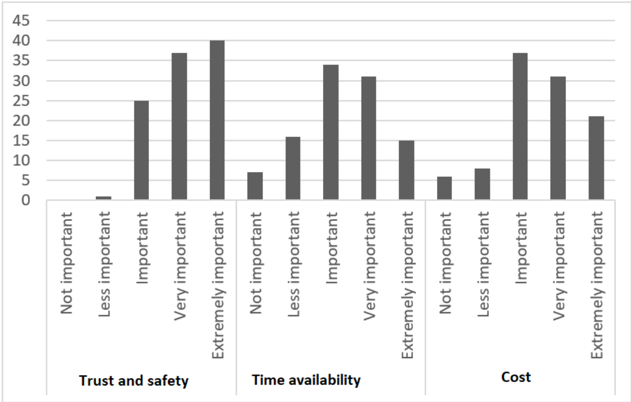

Fig. 2. Acceptance factors for passengers

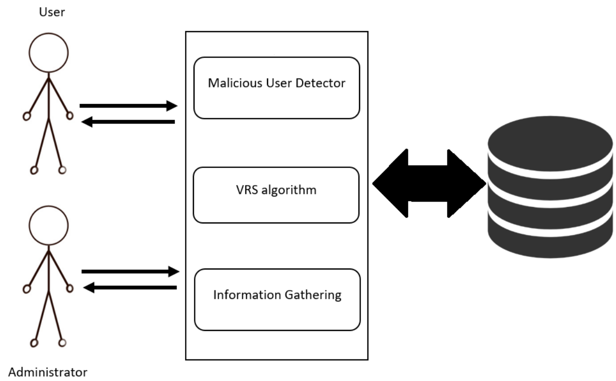

Fig. 4. Software architecture of the proposal

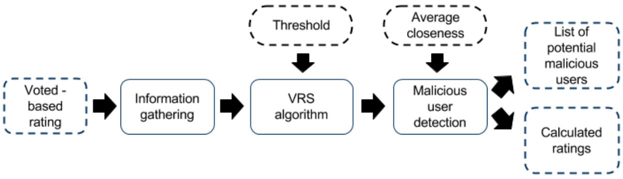

Fig. 3. Procedure to be followed in this method

in the set $\{1 = Poor, 2 = Fair, 3 = Good, 4 = Verygood, 5 = Excellent\}$.

B. VRS definition

In VRS [15], terms such as *service*, *service provider*, *service consumer*, *entity*, and *score* are fundamental. In this study, the definition of such terms was adapted to a shared transportation system as follows:

Routes (R). Set of *n transportation services* connecting origins to destinations. Each route R_β connects a specific origin to a specific destination.

Drivers (D). Set of *g service providers* in the context of a shared transportation system. Each driver D_α provides service in a set of routes $R' \subset R$.

Entities (E). The set of elements $E_{\alpha\beta}$, each representing a driver D_α serving a route $R_\beta : \alpha \in [1, g]$, $\beta \in [1, n]$.

Passengers (P). Set of *q service consumers* who make use of entities $E' \subseteq E$ or require to do so.

Scores (S). Set of ratings of *service consumers* concerning the entities. Each rating is a function $S(P\gamma, E_{\alpha\beta})$ which evaluates for each pair of *service consumer* and entity meeting that $S(P\gamma, E_{\alpha\beta}) \in [0, 5]$.

According to the opinion of Hsu et al., $S(P_p, E_{ab})$ is generally correlated with the set of all $S(P_p, E_{de})$ where $a \neq d \mid\mid b \neq e$., which means that the rating of *service consumers* is correlated with the other ratings of the same *service consumers* [15]. For this reason, Hsu et al. have proposed to take into account ratings of other *service consumers* for making a recommendation only if the other *service consumer* has a certain level of similarity with the *service consumer* for whom the recommendation is made [15]. Steps to estimate a non-explicitly declared $S'(P_p, E_{ab})$ include [15]:

FIRST STEP: to determine the set of *service consumers* P' that have rated entity E_{ab} and, at least, one entity of those rated by P_p. An element of the set P' is represented as $P's$ where $s \in [1, q]$, $s \neq p$.

SECOND STEP: To quantify the difference between each P'_s and the *service consumer* requesting a recommendation (P_p) as the average value $\bar{H}_s$ of the actual difference of explicitly rated entities, as formalized in equation 1.

$$H = |S(P_p, E_{ij}) - S(P's, E_{ij})| \\ where S(P_p, E_{ij}) \neq 0; S(P's, E_{ij}) \neq 0 \tag{1}$$

The difference is computed when both, $S(P_p, E_{ij})$ and $S(P's, E_{ij})$, take value. The average of the computed differences is the metric for quantifying the difference between two *customer service*.

THIRD STEP: To determine the set P'' of *service consumers* who have significant similarities in behaviour to P_p; this set is formed by the *service consumers* for whom $\bar{H}$ value is less than or equal to the threshold.

FOURTH STEP: To calculate $S'(P_p, E_{ab})$ taking into account a weight value w as shown in equation 2.

$$S'(P_p, E_{ab}) = \frac{\sum_{i=1}^{k} w(P''_i) S(P''_i, E_{ab})}{\sum_{i=1}^{k} w(P''_i)} \tag{2}$$

$$w(P''_i) = \frac{m}{\bar{H} + m} \tag{3}$$

$w(P''_i)$ weighs the relevance of $\bar{H}$ in terms of the amount m of rated entities in common while k is the dimension of the set P''. The greater the number of rated entities in common is, the greater the influence on the rating will be to estimate a rating.

C. Definition of Threshold value

In this proposal, the threshold value was set based on an experimental calculation with an available dataset. The calculation involved a random 10% of the entities that were not rated by at least one *service consumer*. 10% of the population was used for guaranteeing the results to be statistically representative. The threshold T was calculated as depicted in equation 4.

$$T = \frac{\sum_{i=1}^{u} \sum_{j=1}^{n} diff(P_i, P_j)}{u * (n-1)} \tag{4}$$

A potential T value was calculated for each $S'(P'_P, E_{\alpha\beta})$ that belongs to the chosen 10% of the population. In each case, u refers to the dimension of the set S' and n is the amount of *service consumers* that rated $E_{\alpha\beta}$. The following was the obtained result.

1.01	1	1	0.99	0.99
0.99	0.99	0.99	0.99	0.99
0.99	0.99	0.99	0.99	0.99
0.99	0.99	0.99	0.99	0.99

Based on these calculations, the average of the threshold resulting from random sampling has led us to establish a T value at 0.99.

D. Malicious User Detector

Based on the criterion of Jnanamurthy et al. [5], it is assumed that malicious *service consumers*' ratings tend to be closer to the threshold more than those from honest *service consumers*. To demonstrate this criterion, a 10% of the total number of users was introduced in the system, assigning low ratings (from 1 to 3) to some randomly selected entities. This election was made thinking that malicious users seek to discredit the reputation of entities. Using equation 1, the difference value was calculated and compared with the difference value of honest *service consumers*. Five malicious *service consumers* were automatically discarded in step three of VRS. The values of the eleven *service consumers* who were not discarded are compared with the values of honest *service consumers* in figure 5.

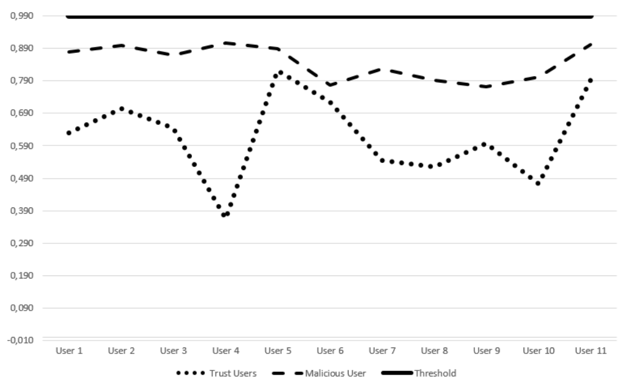

Fig. 5. Difference value between honest and malicious users

When comparing the proximity of the difference value for honest and malicious users against the threshold value, as shown in figure 5, it is obvious that the value for malicious users actually tends to be closer to the threshold. Therefore, the average closeness of the difference value with respect to the threshold could be used for categorizing the behavior of users as malicious or honest. The proposed mathematically formalization to detect malicious users is defined in equation 5.

$$\begin{gathered} differenceNether = \frac{\sum_{a=1}^{n}(1 - diff(P_p, P_a))}{n} \\ where\ diff(P_p, P_a) < Threshold \\ and\ S_{pj} \neq 0; \\ and\ S_{aj} \neq 0; \\ and\ j \in [1, m] \end{gathered} \tag{5}$$

P_p is the *service consumer* that requested a score of a specific entity and n is the number of *service consumers* that rated the entity whose calculated rating was requested. Running calculation with the available dataset, the value 0.23 was determined as the right value for average closeness.

To avoid providing information about the reputation of an entity to malicious *service consumers*, the detection of malicious behaviour is done when a *service consumer* asks to determine the calculated rating of any specific entity.

After running the steps determined by VRS, the differentiation value is calculated for each of these n *service providers* in Formula 5. If $differenceNether$ is less than 0.23, the user is identified as a malicious *service consumer*.

IV. Experimental Settings

An experiment was set in order to measure the efficiency of the proposed reputation system and its ability to detect malicious *service consumers*. The aim of the study is to guarantee a reliable environment based on the honest contributions of the members of the community.

As a none-shared transportation system is used in the city of Ambato, the model was translated to a different case study scenario that made it possible to determine the capacity of detecting malicious *service consumers*. This decision is supported by the need to validate the usability of a shared transportation system prior to the development and/or deployment of this kind of e-collaborative tool.

Considered an economic implementation to have an approximate idea of the validity of the implementation of the reputation system before being applied to a shared transportation system, a system for recommending where to eat was chosen for validating the proposal. In this scenario, the set D was assumed as the set of cafeterias and restaurants around the Universidad Técnica de Ambato and R as the set of meals that those locations provide. The entities E that were evaluated by the *service consumers* refer to those meals they had tried previous to the date when the experiment was conducted. The set of *service consumers* P were students and professors of Computer Science in the university. Those who participated in the experiment provided the vote-based rating feedback of each service provided on each location which was used to give value to the set S.

This scenario was chosen because the involved variables are quite the same and because *service consumers* are potential users of a shared transportation system in Ambato. The values of the set S were collected in an experiment involving 160 participants, 99% of whom were under 25 years old, and 60% were men. A Likert Scale of Values from 1 to 5 was used. Five *service providers* were involved in the experiment with six, twenty, five, twenty-one and four meals that they provide, respectively.

Through the feedback of the participants, the efficiency of the reputation system and the capability to detect malicious users were verified using the threshold value of 0.99 and a borderline of 0.23 that was previously calculated.

To validate the efficiency of the reputation system, the score of foods that had not been tried yet was computed for 10% of participants in the experiment. Afterwards, the food that obtained the highest score for each was recommended and each chosen participant was asked to try this new meal and qualify the service for the first time. The resulting feedback was compared with the estimated rating value by the proposed model in order to compare both ratings and to determine the validity of the proposal. For measuring the capability for detecting malicious *service consumers*, the entry of 10% of malicious users that gave low scores to some services was introduced.

V. Results

A first interesting result of the theoretical study refers to the fee that potential *service consumers* are willing to pay and the fee that potential *service providers* are expecting to charge. In the survey that was conducted, there was a survey item that made it possible to quantify this for *service consumers* and another one to quantify this for *service providers*. The options for potential *service providers* were to offer a free service, to charge the same price as for a bus ticket, to charge slightly higher than for a bus ticket, and to charge a significantly higher price than for a bus ticket.In figure 6, it is obvious that drivers are especially open to accepting the implementation of a free travel mode, and then to a fee equal to that of the bus.

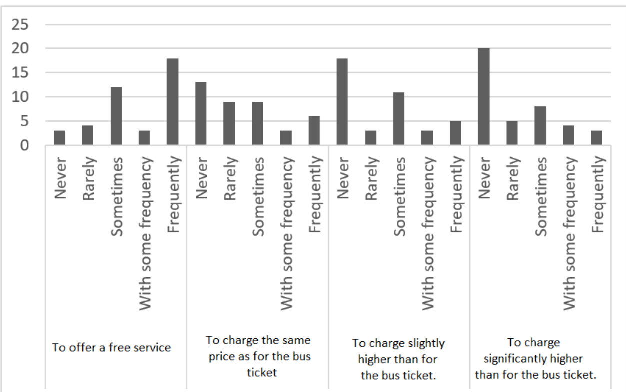

Fig. 6. Acceptance of transport fee by drivers.

In the case of *service consumers*, the options were a free service, to pay the same price as for a bus ticket, to pay a slightly higher price than for a bus ticket and to pay a significantly higher price than for a bus ticket. figure 7 shows that potential passengers are willing to pay the same price of a bus ticket or a slightly higher price than a bus ticket.

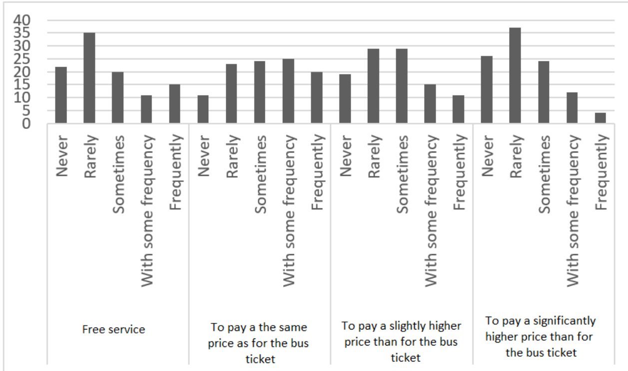

Fig. 7. Acceptance of transport fee by passengers.

As a relevant result of the conducted survey, it seems that, in order to have a successful shared transportation system in Ambato, it is necessary to keep prices around the price of a bus ticket. Furthermore, it was statistically corroborated that there is no significant difference in the importance of safety and confidence when sharing a vehicle between those who travel by car and those who do not. The type of statistical test that was performed was the Mann Whitney test. The resulting value of p-value (sig) bilateral significance was 0.540. In addition, the mean range of this variable in the two groups was 97.49 and 92.92. Therefore, confidence and safety are essential parameters when passengers and drivers decide to use a shared transportation system.

With 10% of malicious users, the system was able to detect 85% of them. Therefore, the ability of the proposed system to classify its users into honest and malicious makes it possible to keep members honest and drive out the malicious ones in order to ensure an environment of trust and safety, thus ensuring the future success of an e-collaborative generic system including the system of shared transportation. On the other hand, through the feedback of the students on the acceptance of the recommended food, it was determined that the reputation system deviates 1.07 from the satisfactory recommendation corresponding to 5 stars, which implies that on average the reputation system that is proposed is giving suggestions that equate to 4 stars, being a result which is considered satisfactory. The whole data of both, the theoretical study and the experimental evaluation are available at: https://soman.uta.edu.ec/trans-rep.

VI. Conclusions

This proposal accepts that the behaviour of malicious users compromises the reliability of any e-collaborative applications such as the shared transportation system. It was actually corroborated that the implementation of a shared transportation system would be feasible to reduce traffic and pollution in the city of Ambato only if trust and safety are guaranteed. Thus, a method for detecting malicious users was defined and validated in the light of this study for a specific study case.

Shared transportation actually promotes a greater social connection through e-collaboration and also a greater awareness of the efficiency in the use of resources that generate pollution. With this study, an attempt was made to overcome the gap between social interactions and a vote-based reputation system. It is expected to improve the users' experience of a shared transportation system while reducing pollution and increasing the efficiency of the transportation system of a medium-sized city, like Ambato.

The proposal was experimentally validated. Results in the case study showed that it was possible to detect 85% of malicious users in a scenario with 75% of users being potentially honest with their ratings. This result is encouraging for the implementation of a satisfactory shared transportation system in Ambato based on trust and safety for the citizens. On the other hand, the characteristics of the business logic of a shared transportation system also leads to conclude that this proposal will be applicable in other e-collaboration service environments where there is a ranking of quality of service.

VII. Future Work

This is a work in progress; for this reason it is suggested for future work, to increase the sample of the theoretical and experimental study, in order to corroborate the obtained results in a broader scenario. Additionally for future works, the transportation system would be implemented to carry out the tests with the reputation system validated in this study.

References

[1] Instituto Geográfico Militar, "Carta Topográfica Ambato." [Online]. Available: http://www.igm.gob.ec/work/files/cartabase/enie/ENIEIV_A4.htm

[2] INEC, "Informacion Ambiental en Hogares," Instituto Nacional de Estadisticas y Censos, Tech. Rep., 2016. [Online]. Available: http://www.ecuadorencifras.gob.ec/documentos/web-inec/Encuestas_Ambientales/Hogares/Hogares_2016/TABULADOS_MOD_AMB_2016.xlsx

[3] Forbes, "Uber lanzará servicio de viajes compartidos," 2016. [Online]. Available: https://www.forbes.com.mx/uber-lanzara-servicio-de-viajes-compartidos/

[4] Cheyne and M. Imran, "Shared transport: Reducing energy demand and enhancing transport options for residents of small towns," *Energy Research and Social Science*, vol. 18, pp. 139–150, 2016.

[5] H. K. Jnanamurthy and S. Singh, "Detection and filtering of collaborative malicious users in reputation system using quality repository approach," *Proceedings of the 2013 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2013*, pp. 466–471, 2013.

[6] "IoT-based collaborative reputation system for associating visitors and artworks in a cultural scenario," *Expert Systems with Applications*, vol. 79, pp. 101–111, 2017.

[7] A. A. Kardan, R. Salarmehr, and A. Farshad, "SigmoRep: A robust reputation model for open collaborative environments," *Proceedings - 2014 IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications, TrustCom 2014*, pp. 505–510, 2015.

[8] X. Zhou, D. Lin, and T. Ishida, "Evaluating reputation of web services under rating scarcity," *Proceedings - 2016 IEEE International Conference on Services Computing, SCC 2016*, pp. 211–218, 2016.

[9] M. Pouryazdan, B. Kantarci, T. Soyata, L. Foschini, and H. Song, "Quantifying user reputation scores, data trustworthiness, and user incentives in mobile crowd-sensing," *IEEE Access*, vol. 5, pp. 1382–1397, 2017.

[10] W. Y. Zhang, S. Zhang, and S. S. Guo, "A PageRank-based reputation model for personalised manufacturing service recommendation," *Enterprise Information Systems*, vol. 11, no. 5, pp. 672–693, 2017.

[11] H. Xie and J. C.S. Lui, "Mining Deficiencies of Online Reputation Systems: Methodologies, Experiments and Implications," *IEEE Transactions on Services Computing*, vol. 1374, 2017.

[12] Y. Wang, J. Yang, and L. Qi, "A game-theoretic model for the role of reputation feedback systems in peer-to-peer commerce," *International Journal of Production Economics*, vol. 191, pp. 178–193, 2017.

[13] A. Panagopoulos, E. Koutrouli, and A. Tsalgatidou, "Modeling and Evaluating a Robust Feedback-Based Reputation System for E-Commerce Platforms," *ACM Transactions on the Web*, vol. 11, no. 3, pp. 1–55, 2017.

[14] W. Zheng and L. Jin, "A consumer decision-making model in M-commerce: The role of reputation systems in mobile app purchases," in *Mobile Commerce: Concepts, Methodologies, Tools, and Applications*, 2017, pp. 107–130.

[15] F.-H. Hsu, Y.-L. Hsu, Y.-L. Hwang, L.-H. Chen, C.-S. Wang, C.-K. Tso, S.-C. Liu, P.-C. Lin, and C.-H. Hsu, "VRS: a values-based reputation system for web services," *Security and Communication Networks*, vol. 9, no. 18, pp. 6335–6350, 2016.