Smart Resident: A Personalized Transportation Guidance System

Mohammad Amin Kuhail, Bilal Ahmad, Collin Rottinghaus
School of Computing and Engineering
University of Missouri-Kansas City
Kansas City, MO, US
e-mail: {kuhailm, bavwb, cqrfmf}@umkc.edu

***Abstract*—Many people around the globe live in areas that have unhealthy levels of air pollution. Such pollution raises the risks of health problems such as lung cancer and heart disease. A significant contributor to air pollution is the combustion of fossil fuels such as gasoline and diesel to transport people and goods. Using shared and active transportation such as cycling and walking can have great benefits in mobility, clean air, citizen health, and transportation efficiency.**

Driven by the idea of improving air quality, mobility, and the quality of resident lives, we contribute Smart Resident, a transportation guidance system which uses a recommender system and gamification to motivate city residents to move towards active or shared transportation as an alternative to traditional one-passenger vehicles. The system uses a personalized recommender system based on the history of user trips, and user preferences to give recommendations to users for daily transportation. Moreover, the system motivates users towards green means of transportation by showing them impact on health and finances.

Keywords—smart city, recommender system, content-based filtering, recommendation freshness.

I. INTRODUCTION

According to the United States Environmental Protection Agency, the second largest source of air pollution, in the form of carbon emission, is the combustion of fossil fuels such as gasoline and diesel to transport people and goods. It contributes to thirty one percent of the total carbon emission in the US. Unfortunately, air pollution could increase the risk of major health problems such as lung cancer, asthma attacks, heart disease, reproductive problems, and other ailments [1], [2].

Active transportation such as cycling and walking can be a healthy and environmentally-friendly alternative to fossil-fuel transportation. Not only does active transportation facilitate personal mobility, but could also potentially save $24 billion and reduce carbon emission by 11 percent by 2050 [3]. Furthermore, active transportation increases the amount of physical activity of city residents, which contributes to their overall health. Moreover, active transportation encourages residents to be more engaged in community issues, enhance their quality of life, and can potentially improve city safety [4]. Coupled with health and social benefits, active transportation can also help to reduce roadway congestion since it requires less space per traveler than driving.

Given the pervasiveness of information technology in our daily lives, it is becoming increasingly critical to use technology to help city residents to move towards more sustainable and echo-friendly alternatives. One idea that shows potential in that direction is "smart city". A smart city is a city that that uses technology to improve residents' lives.

Inspired by the idea of a smart city, we contribute, Smart Resident, a transportation guidance system which uses a recommender system and gamification to motivate city residents to move towards active or shared transportation as an alternative to traditional one-passenger vehicles.

Traditional Transportation guidance systems such as Google Maps help users to find a few different routes from point A to point B using different means of transportation [5]. Traditional transportation guidance systems are mainly driven by shortest routes. Smart Resident adds a few more features to traditional transportation features: (1) The system uses an evidence-based personalized recommender system to generate the route suggestions. The system takes into consideration the context (e.g. time, traffic, day of week, temperature, and weather condition), user's transportation habits, and preferences. (2) The system gives residents feedback on the positive and negative impact of their transportation decisions on the environment. The objective of this feedback is to motivate users to choose healthier and more economical means of transportation.

The main objective of this work is to shed light on the design principles of Smart Resident, and how we envision it as a framework that can be used to help residents make more environmentally friendly and healthy transportation decisions.

II. RELATED WORK

The related work can be divided into two sections: smart city applications and recommender systems.

A. Smart City Applications

There are many applications that use technology for the benefit of city residents. For instance, some applications collect traffic data to provide residents with information about congestion and road accidents [6]. The information is collected by sensors or reports made by drivers. Other applications allow residents to report issues such as suspicious activity, illegal trash dumping, potholes, defective street lights, broken tiles on sidewalks, and illegal advertising boards [7]. For such applications to be successful, it is imperative that residents participate in them. These applications can be incentivized by extrinsic motivation such as discounts offered by some commercial stores. However, research shows that intrinsic motivation has more long-term success [8]. Hence, rather than relying on extrinsic motivation, our approach capitalizes on residents' intrinsic motivation to help residents move towards more eco-friendly means of transportation. Our approach is to show residents the impact of their transportation decisions on the environment, their budget, and their health.

B. Recommender Systems

A recommender system is an information filtering system that predicts the possibility a user would like an item, and therefore recommends it [9]. Recommender systems typically produce a list of recommendations to users ranked by a relevance score. Recommender systems have been used in various areas including books, music, movies, news, and social tags.

The most popular recommender systems use collaborative filtering or content-based filtering [10]. The collaborative filtering approaches builds a model from decisions made by similar users. Similarity is decided based on common items or attributes that users like. This model is then used to predict items that the user may have an interest in [11]. This approach is powerful since it may predict recommendations that are novel, and not easy to find [12]. However, this approach suffers from the "cold-start problem", not being able to produce meaningful recommendations when there is not a significant user database [13]. The content-based filtering approach uses a series of discrete characteristics of an item in order to recommend additional items with similar properties [14]. This approach is powerful because it does not need data from other users. However, recommendations may not be novel.

Both approaches require a substantial data set to be able to generate meaningful recommendations. Hence, the two approaches are often combined to tackle that issue [15].

Knowledge-based recommender system produce recommendations by inferring user's needs and preferences. Further, they make an expectation about how an item meetings a particular user need, and can therefore reason about the relationship between a need and a recommendation [16].

Context-based recommender systems takes contextual data into consideration such as the change in the interests of users over time [17], location, and environment conditions.

III. DESIGN

Smart Resident is a transportation guidance system which uses a personalized recommender system and feedback to motivate city residents to move towards active or shared transportation,

In designing Smart Resident, we wanted the recommendations to be realistic and behavior changing. Moreover, we wanted the recommendations to give more weight to new observations. In other words, more recent decisions the user makes affect the system's recommendations more than older decisions.

Realistic recommendations are based on the user's preferences and evidence for how the user has been using transportation. Further, the recommendations are based on many different variables that may affect the resident's decision of choosing a specific route or means of transportation. Examples of these variables include weather, distance, road infrastructure, and traffic. *Behavior changing* recommendations attempt to slowly change the user's transportation habits from highly polluting means of transportation (e.g. one-passenger vehicle) towards more active and shared transportation.

In the following subsections, we will introduce the different components of Smart Resident: system architecture, route calculator, recommender system, and feedback provider.

A. System Architecture

Figure 1 shows the system architecture of Smart Resident. The user provides the preferences to the recommender system. The preferences are elicited through a series of questions about when the user is likely to use a certain means of transportation such as a bicycle or car. The user also informs the system about the trip they want to make. The route calculator receives the origin point (point A) and the destination point (point B), and generates several possible routes by different means of transportation. When calculating the routes, the route calculator takes into account traffic, air quality, and distance. The recommender system receives the routes and ranks them based on how likely the user will choose them. The feedback provider gives feedback on the routes in terms of the cost, carbon emission, and calories burnt. The user chooses a route, and the decision goes back to the recommender system to take it into consideration for future recommendations. The following subsections will briefly explain how each component works.

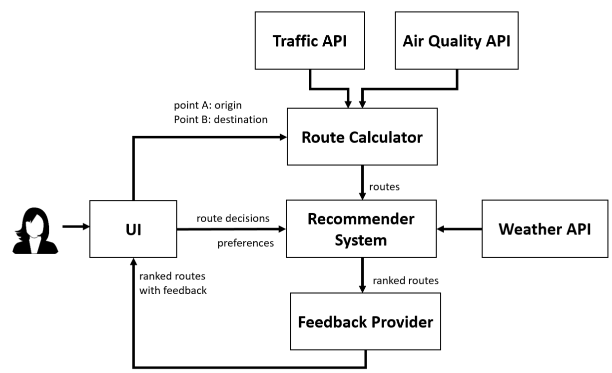

Fig. 1. Smart Resident System Architecture

B. Route Calculator

We used an existing algorithm, adapted from Graph Hopper [18] to generate different routes between point A and point B. The algorithm is based on Dijkstra's shortest route algorithm [19]. We adapted the algorithm so that it does not just rely on distance when calculating the routes. Instead, it generates a variety of routes based on several factors: distance, air quality, and traffic. The implementation can be found at [20]. Currently traffic and air quality are simulated variables. Traffic has a value from 0 to 5, where 0 means there is no traffic at all, and 5 means there is very heavy traffic. Moreover, air quality's value is between 0 to 5, where 0 is very harmful polluted air, and 5 is very clean and healthy air. The route calculator calculates three alternative routes for various means of transportation (car, bicycle, walking):

- The shortest route by distance.
- The shortest route by traffic and distance.
- The shortest route by air quality, traffic, and distance.

C. Recommender System

The recommender system primarily uses content-based filtering to generate recommendations. The content is the history of transportation decisions that the user has made. However, to combat the cold start problem, we also base the recommendations on user's transportation preferences. The implementation can be found at [21].

We identified the following variables that may affect the decision of the user to use a certain means of transportation: tempreature, weather condition, traffic, availability of bike lanes, availability of sidewalks, trip distance, trip time, day of week, and time of day. The availability of bike lanes and sidewalks is a number from 0 to 5, where 0 means there are no lanes or sidewalks at all, and 5 means there are bike lanes or sidewalks everywhere in the route.

We used our intution to identify some variables such as bike lanes, and literature to identify other variables such as weather and tempreature [22].

Initially, the system does not have historical data to use to give recommendations. Hence, the system initially asks the user a series of questions to elicit the user preferences to build an initial model. The questions aim at knowing when the user is likely to use a certain means of transportation based on various factors such as traffic, tempreature, weather condition, etc. The estimated time for answering the questions is no longer than five minutes. Figure 2 gives an example of eliciting bicycle-tempreature data from the user.

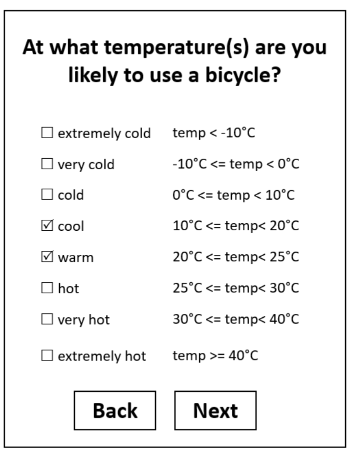

Fig. 2. Eliciting user preferences

Observation Matrix: Whenever the user makes a transportation decision, it is stored in the observations matrix as a new row. For instance, the first row in Figure 3 desribes the situation where a user drove a car, tempreature was -5°C, traffic was moderate, the distance was 10 miles, air quality was good, and there were not many bike lanes or sidewalks in the recent an observation is. It is calculated in terms of the number of days that have elapsed since the recommender system was first used. For instance, the fourth observation in Figure 3 is the most recent since it has been recorded four days after the recommender system has initially been used. The reason why we incorporate freshness is because we want the recommender system to give more weight to newer observations. As the user uses the system more, his or her behavior will change more. Hence, the system will be more influenced by the more recent observations.

For proving the concept, we simulated the data in the observation matrix. We used random normal distribution to generate the categorical and continuous variables in the observation matrix.

Means of Transport	Temperature	Traffic	Distance	Air Quality	Bike Lanes	Sidewalks	Time of day	Day	Freshness
Car	-5°C	3	10m..	4	2	2	7:00 PM	MON	1
Transit	20°C	4	6m	3	3	3	8:30 AM	TUE	2
Car	5°C	5	4.5m	2	3	3	8:00 AM	TUE	3
Bike	15°C	2	3m	3	4	4	12:00 PM	WE	4
...	...	...	...	...	...	...	...	...	...

Fig. 3. Observation Matrix

User Profile: When a new observation is recorded, it is added accumulatively to the user profile. The user profile is a matrix that models the decisions and preferences the user. Figure 4 shows an example of a user profile matrix. However, due to lack of space, not all the variables are shown.

The columns are the different means of transportation, and the rows represent the variables that may influence the user's decision. We broke down the main variables (e.g. temperature) into sub-variables. For instance, the temperature sub-variables were defined based on various ranges. For instance, the "very cold" sub-variable represents the temperature between -10° C and 0° C, whereas the warm temperature represents the temperature between 20° C and 25° C. Our ranges were inspired by the suggestions made by the national weather service [23].

			Means of transportation					
			Walk	Cycle	Transit	Drive	Cycle + Transit	Total
Temperature	extremely cold	temp < -10°C	0	0	5	12	0	17
	very cold	-10°C <=temp < 0°C	0	0	9	12	0	21
	cold	0°C <=temp < 10°C	1	3	10	30	5	49
	cool	10°C <=temp< 20°C	4	10	15	5	3	37
	warm	20°C <=temp< 25°C	12	20	5	2	3	42
	hot	25°C <=temp< 30°C	2	13	15	10	3	43
	very hot	30°C <=temp< 40°C	0	1	10	20	5	36
	extremely hot	temp > =40°C	0	0	0	1	0	1
Traffic	very heavy traffic	5 <=traffic < 4	5	3	10	30	5	53
	heavy traffic	4 < =traffic < 3	4	10	10	5	3	32
	Moderate traffic	3 <= traffic < 2	20	20	5	2	3	50
	Light traffic	2 <=traffic < 1	2	13	15	10	3	33
	Very light traffic	traffic<=1	0	1	10	20	5	36
Distance	Very short distance	distance<1m	12	20	5	2	1	40
	short distance	1m<=distance<3m	16	15	10	10	2	53
	moderate distance	3m<=distance<5m	1	10	15	15	3	44
	a little long distance	5m<=distance<8m	12	20	14	20	3	69
	long distance	8m<=distance<10m	0	5	30	15	3	53
	Very long distance	distance>=10m	0	0	20	20	5	45

Fig 4. User Profile Matrix

Each observation in the observation matrix is used to update the user profile matrix. For instance, the first observation in Figure 3 updates the user profile in this way. For each value in the observation, we look up the relevant means of transportation vs the appropriate variable. Let's take the first observation in the observations matrix as an example. The means of transportation in the first observation is car. Hence, we will be updating some cells that fall under the "Drive" column in the user profile matrix. Further, the temperature in the first observation is -5° C. This temperature corresponds to the "very cold" temperature sub-variable. Hence, we add 1 to the cell "Drive vs. very cold temperature" (1 represents the freshness value.) The traffic value is 3, which is considered "moderate" traffic. Therefore, we add 1 to the cell "car vs moderate traffic". We continue to update the matrix for each value in the observation. We follow the same process in the second observation. However, this time the freshness is 2. Hence, we add 2 to the corresponding cells. For instance, the temperature being 20 C corresponds to the variable warm. Hence, we add 2 to the cell "transit vs warm".

The algorithm for updating the matrix based on an observation is as follows:

Algorithm 1: update_user_profile

input: an observation (O) in the observation matrix (OM), the user profile matrix (M).

output: the user profile matrix (M) being updated

1. I=O["means_of_transportation"]; //transportation index
2. **foreach** entry in O **do**:
3. variable_index=entry.index();
4. M[variable_index][I]+=O["freshness"];
5. **end**

Recommendations: Making recommendations is simply ranking different routes produced by the route calculator. After the ranking is done, we simply present the routes to the user in the order they were ranked. The ranking is essentially based on the user profile matrix. To rank each route, the recommender system gets every route as a set of attributes. For instance, Table I shows an example of the attributes of a certain route. In addition to the route attributes, the recommender system also receives weather information (see Table II).

Based on route attributes and weather information, the recommender system identifies the relevant variables that it needs to look up in the user profile matrix. For instance, since the distance is 4.5 miles, it corresponds to the variable "moderate distance". The traffic variable corresponds to the variable "light traffic". Furthermore, since the temperature is 30°C, it corresponds to the variable "very hot".

Next the recommender system makes correlations between the variables and the means of transportation. The algorithm treats all variables as binary. Hence, the algorithm calculates the phi correlation between every variable in the matrix as well as the relevant means of transportation [24]. For instance, Table III shows an example of calculating the phi correlation between bike and warm temperature. In this case, it is 0.146 (weak correlation). Similarly, the algorithm calculates the correlation between other variables and the relevant means of transportation.

TABLE I. ROUTE ATTRIBUTES

Variable	Value
Means of transportation	Bike
Distance	4.5 miles
Bike Lanes	Full
Traffic	2
Side walks	Full
Route Info	<<coordinates info>>

TABLE II. WEATHER CONDITIONS

Variable	Value
Temperature	30°C
Sky	Clear
Wind	No
Rain	No
Storm	No

TABLE III. EXAMPLE OF PHI CORRELATION

	Warm Temperature (W)	Not Warm Temperature (W')	Total
Bike (B)	50	20	70
Not Bike (B')	120	100	220
Total	170	120	

After the correlations have been calculated, they are converted into the fisher transformation. This transformation produces a function that is normally distributed rather than skewed. Next the mean of all fisher values is calculated to calculate the score of the route. The algorithm can be expressed as follow:

Algorithm 2: calculate_route_score

input: the route attributes (RA) of a certain route, the index that represents the means of transportation in the user profile matrix (trans_index), the user profile matrix (M)

output: a score that will be used to rank the route

1. **foreach** attribute in RA **do**:
2. correlation=corr(attribute,trans_index);
3. fisher=fisher(correlation);
4. fisher_list.add(fisher);
5. }
6. score=mean(fisher_list);
7. **end**

D. Feedback Provider

The feedback provider component aims at motivating residents to make a transportation decision that is environmentally and budget friendly as well as healthy. The feedback provider uses gamification in the form of visual feedback to motivate residents. Gamification engages users and make them more productive [25].

In addition to distance and time for each route, the system shows the route impact that is related to the resident's health and budget, as well as the environment. In particular, the system shows calories burnt, cents saved, and the pollution caused by the trip. The system makes it clear to the resident as to whether the route impact is positive (such as calories burnt or cents saved), or negative (such as pollution emitted by a vehicle). Furthermore, the system highlights the most or least eco-friendly route (see Figure 5).

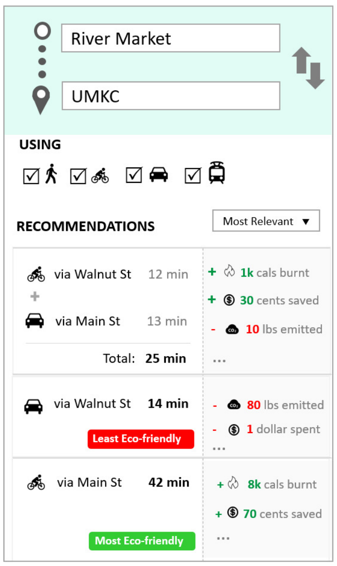

Fig 5. Route Feedback

IV. CONCLUSION AND FUTURE WORK

In this paper, we introduced Smart Resident, a system that helps users make transportation decisions that are environmentally and budget friendly as well as healthy. Unlike traditional transportation guidance systems such as Google Maps, Smart Resident uses a personalized recommender systems that make transportation recommendations based on the user's history of transportation, and other factors such as weather condition, traffic, time, etc. Further, the system provides the user with feedback on the recommended routes in an attempt to persuade the user to make healthier and more eco-friendly choices. In the future, we are planning to evaluate the accuracy of the recommendations by testing it with a random sample of users. Furthermore, we are planning to test the effect of the feedback provider on user's choices.

REFERENCES

[1] Washington State Department of Ecology, "Health Effects from Automobile Emissions", 2018.

[2] US Environmental Protection Agency, Overview of Greenhouse Gases, https://www.epa.gov/ghgemissions/overview-greenhouse-gases, 2018.

[3] F. Kuster, "A European Roadmap for cycling", ECF. 1–14. 2015.

[4] T. Koska, F. Rudolph, "The Role of Walking and Cycling in Reducing Congestion", A Portfolio of Measures., FLOW Project, 2016.

[5] Google Maps, Found at: https://www.google.com/maps , 2018.

[6] Austin, Live From Austin, Texas: The Smart City Challenge, found at: https://www.austintexas.gov/sites/default/files/files/Transportation/Austin_SCCFinal_Volume1_5.25.pdf , 2018.

[7] Smart City App, Improve My City ~ Citizen Requests, Complaints, and Suggestions, Found at: http://smartcityapps.urenio.org/improve-my-city_en.html, 2018.

[8] R. Singh, "The impact of intrinsic and extrinsic motivators on employee engagement in information organizations", Journal of Education for Library and Information, 2016

[9] F. Ricci, L. Rokach, B. Shapira, "Introduction to Recommender Systems Handbook", Recommender Systems Handbook, Springer, 2011.

[10] J. S. Breese, D. Heckerman, C. Kadie, Empirical analysis of predictive algorithms for collaborative filtering. In Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence, 1998.

[11] Allen, R.B., "User Models: Theory, Method, Practice". International J. Man-Machine Studies, 1990.

[12] P. Adamopoulos, A. Tuzhilin, "On Unexpectedness in Recommender Systems: Or How to Better Expect the Unexpected". ACM Transactions on Intelligent Systems and Technology. doi:10.1145/2559952, 2015.

[13] A. I. Schein; A. Popescul; L. H. Ungar; D. M. Pennock. Methods and Metrics for Cold-Start Recommendations. Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2002). New York City, New York: ACM. pp. 253–260. ISBN 1-58113-561-0, 2002.

[14] C. Aggarwal, Recommender Systems: The Textbook. Springer. ISBN 9783319296579, 2016.

[15] G. Adomavicius, A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions". IEEE Transactions on Knowledge and Data Engineering. 17, 2005.

[16] R. Burke, Hybrid Web Recommender Systems, The Adaptive Web, Peter Brusilovsky, Alfred Kobsa, Wolfgang Nejdl (Ed.), Lecture Notes in Computer Science, Springer-Verlag, Berlin, Germany, Lecture Notes in Computer Science, Vol. 4321, pp. 377-408, 2007.

[17] D. Bouneffouf, "Following the User's Interests in Mobile Context-Aware Recommender Systems: The Hybrid-e-greedy Algorithm", Proceedings of the 2012 26th International Conference on Advanced Information Networking and Applications Workshops (PDF), Lecture Notes in Computer Science, IEEE Computer Society, pp. 657–662, 2012.

[18] GraphHopper, a Directions API, Found at: https://www.graphhopper.com/ , 2018.

[19] E. W. Dijkstra, E. W., "A note on two problems in connexion with graphs" (PDF). Numerische Mathematik. 1: 269–271. 1959.

[20] C. Rottinghaus, Smart Resident Route Calculator Implementation, https://bitbucket.org/CollinQuiason/smarthopper, March 2018.

[21] B. Ahmed, M. A. Kuhail, Smart Resident Implementation, https://github.com/bilaleme/Smart-Resident. March 2018.

[22] C. Brandenburg,A. Matzarakis, A. Arnberger1, Weather and cycling – a first approach to the effects of weather conditions on cycling, Meteorol. Appl. 14: 61–67, 2007.

[23] NWS, Found at: https://www.weather.gov/ajk/ForecastTerms , 2018.

[24] B. S. Everitt, The Cambridge Dictionary of Statistics, CUP, 2002.

[25] S. Dale, Gamification: Making work fun, or making fun of work?, Business Information Review, Vol. 31(2) 82–90, 2018