RESEARCH PROGRESS REPORT

EFFECTIVE IDENTIFICATION AND ENGAGEMENT OF TRANSPORTATION STAKEHOLDERS USING GEOSPATIAL ANALYTICS AND ONLINE ADVERTISING

George Orfanidis
Guided By - Dr. Okan Pala
Computer Science Department | Center of Geospatial Analytics
North Carolina State University

ABSTRACT

How does one effectively identify transportation stakeholders using spatial and demographic analysis in order to reach them using popular advertising platforms? As a part of the project development phase, the public outreach and engagement stage is one of the most crucial parts of transportation project planning. In this stage, the public is informed of the future plans and of the benefits for the broader community. Direct and indirect feedback is gathered. Organizing and implementing public hearings, while effective, falls short on involving all stakeholder groups who would be affected by the implemented project. One of the reasons for that is the difficulty in identifying and engaging certain transportation project stakeholders. It has been observed that local stakeholders, people that are in close proximity to the project, are more likely to get involved in the planning process compared to commuting stakeholders, people that live farther away but regularly travel through or to the project area. This may cause a one-sided representation in public hearings that may not reveal the actual value (or lack thereof) of the proposed transportation project. Therefore, in order to have equality in the representation of opposing points of view, it is important to identify and engage all stakeholders that have an honest interest in the outcome of the project. With the new techniques of identifying and engaging groups of stakeholders, we aim to alleviate some of these concerns of misrepresentation. The objective is to engage all stakeholders (both local and commuting) that will have a meaningful and constructive contribution to ensure that the transportation project team is able to gather feedback from both points of view. For this, we introduce targeting/ advertising using geospatial and demographic data along with geospatial analytics and online advertising. In terms of spatial/location targeting, we have developed a tool by combining existing options from popular advertising platforms in a novel way. This helps with the precise identification of local and commuting stakeholders. In order to digitally reach stakeholders with a high possibility of active participation, we also employed demographic analysis. For this, we performed a thorough literature review to create demographic profiles of possible stakeholders who are more relevant or might have a greater interest in specific transportation projects. We developed a tool that pinpoints areas that have a significant population density of demographic categories of interest which are used as specific targeting locations for online outreach campaigns. Finally, the outputs of these two tools will be used as input to a Robotic Process Automation (RPA) tool that will automatically create the targeted advertising campaigns. The overall purpose and effort of identifying stakeholders and appropriately targeting them is to have an equal, fair, and full representation of stakeholders as a part of the decision making and improvement process for

a given transportation project.

INTRODUCTION

Background

The planning process of a transportation project, composed of multiple stages, is multi-dimensional and complex. Each time the North Carolina Department of Transportation (NCDOT) plans a project, six steps are followed including the acquisition of public input on the proposed project. Organizing and implementing public hearings, while effective, falls short on involving all stakeholder groups who would be affected by the implemented project. Specifically, the vast majority of the involved stakeholders have been those who are local to the proposed project and do not want construction near their home or work location. Outreach methods such as publicly announcing the proposed projects via social media or sending informational notices through the mail (North Carolina Department of Transportation, 2019) have not produced the desired increase in the participation rate of stakeholders.

Problem Statement

Data from previous outreach and engagement campaigns have revealed public representation flaws. Local stakeholders, people that are in close proximity to the project's location, are more likely to attend public hearings compared to those who are commuting stakeholders, people that live farther away but regularly travel through or to the project area. This creates a one-sided perspective that may not reveal both the positive and negative effects of the project. Additionally, participation data has shown that many people become involved in the planning process simply because they are skeptical about long term construction nearby. Finally, some stakeholders do not live in the vicinity of the project location and therefore they are unaware of the existence of the project at all. In order to mitigate the cascading effects of these flaws, an innovative public outreach and engagement strategy should be implemented. It is evident that populations that may have a greater interest in the outcome of the project should be identified in addition to general local and commuting stakeholders in order to have an equal representation of both sides and therefore be able to implement the project in a way where everybody is satisfied by the outcome. This project's aim is to increase public awareness of proposed transportation projects by identifying and targeting affected stakeholders using spatial analysis (to target the local and commuting stakeholders)

and demographic analysis (to target areas with a high concentration of people particularly interested in the given project).

Previous Work

A geospatial analysis tool was developed by the Center of Geospatial Analytics at North Carolina State University. Its functionality can assist in the improvement of the public outreach and engagement stage of the transportation planning process. The project will utilize Network analysis and driving polygons with demographic data that has been previously created [7]. The goal of the driving polygons is to create a clear and comprehensive mapping of the areas around the transportation project based on drive time. This will be used further in conjunction with spatial and demographic analysis for precise targeting.

The initial phase of creating said driving polygons was to collect relevant data to conduct network analysis. A variety of road data was collected from different sources to construct a full transportation network. Furthermore, demographic data from the American Community Survey (ACS) was incorporated to estimate populations of specific demographic categories within the borders of the driving polygons, which were constructed with center at the transportation project's location. (see Figure 1)

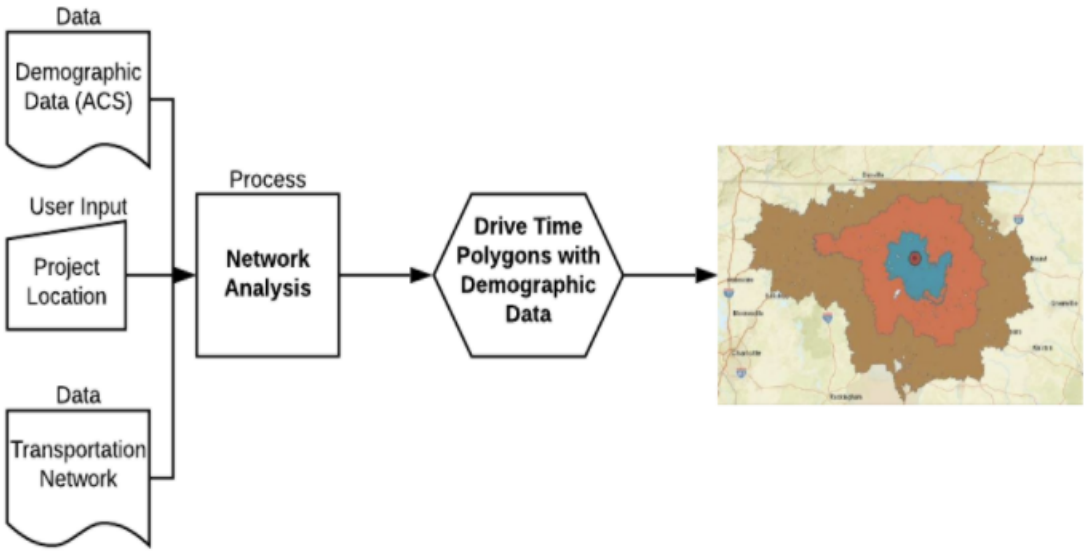

Figure 1: Flow Chart of the already existing tool.

Each drive-time polygon shows how far someone can travel in a specified amount of time from the center of the transportation project. The specified drive times in the above picture are 10 minutes, 30 minutes, and 50 minutes indicated by the colored concentric polygons from smallest to largest respectively. Each polygon is connected with an approximate population of specific demographic categories. We will leverage this spatial analysis tool to better identify and target transportation project stakeholders.

Our Approach

Location Tool

The drive-time polygons will be used to identify areas with the population affected by a specific transportation project. The outer drive-time polygon will determine the last area for targeting. Everything outside of that border will be considered irrelevant for targeting and as a result, will be excluded. People from the excluded areas will not receive any type of advertisement that is related to the transportation project. We assume that people outside that region won't be affected since they are too far away from the project's site (more than an hour away). Consequently, we need to target the stakeholders up to the outer driving polygon border that may be affected by the implemented transportation project.

The most effective way to map as accurately as possible the drive-time polygons in order to determine the locations that need to be excluded is by the use of Zip Codes. In Google Ads' platform, targeting by Zip Codes is the most precise and computationally feasible way to emulate the drive-time polygons. Based on which drive-time polygon needs to be mapped a specific relationship between the Zip Code layer and drive-time polygon layer is selected. If the outer drive-time polygon needs to be mapped, the location script will identify Zip Codes that have their center in the outer drive-time polygon else if the inner drive-time polygon needs to be mapped the location script will select Zip Codes that intersect with the inner drive-time polygon (Figure 2). To create the list of excluded Zip Codes the selection is inverted, meaning that Zip Codes that do not have their center in or do not intersect with the corresponding drive-time polygon are selected. Since all of North Carolina has been dealt with in terms of exclusion, the final step is to exclude all the other states and countries to leave us the final area of interest.

Google Ads' platform restricts the number of excluded locations to a thousand. This forces the smallest drive time polygon to include at least three Zip codes. However, it is very rare that this restriction will ever have any effect in this type of targeting.

- The total number of Zip Codes in North Carolina: 808
- Total number of countries as of 2020 (United States are excluded): 194
- Total number of states (North Carolina is excluded): 49
- Zip Codes included in the drive-time polygon >= 3

There are 808 Zip codes in North Carolina however, 48 of them are not recognized by Google Ads' platform for privacy purposes. Consequently, the available number of Zip Codes for processing is 760. The locations that will always be in the excluded list are given by Countries + States (North Carolina is excluded) = 194 + 49 = 243. So, knowing that Google Ads allows for only 1000 location exclusions, we still have 757 to work with (in terms of Zip Codes), 1000 - 243 = 757 slots for Zip codes exclusions used later in the process. At least three Zip Codes should be included within the drive-time polygon, total number of Zip Codes in North Carolina - available slots for Zip Codes exclusions = 760 - 757 = 3.

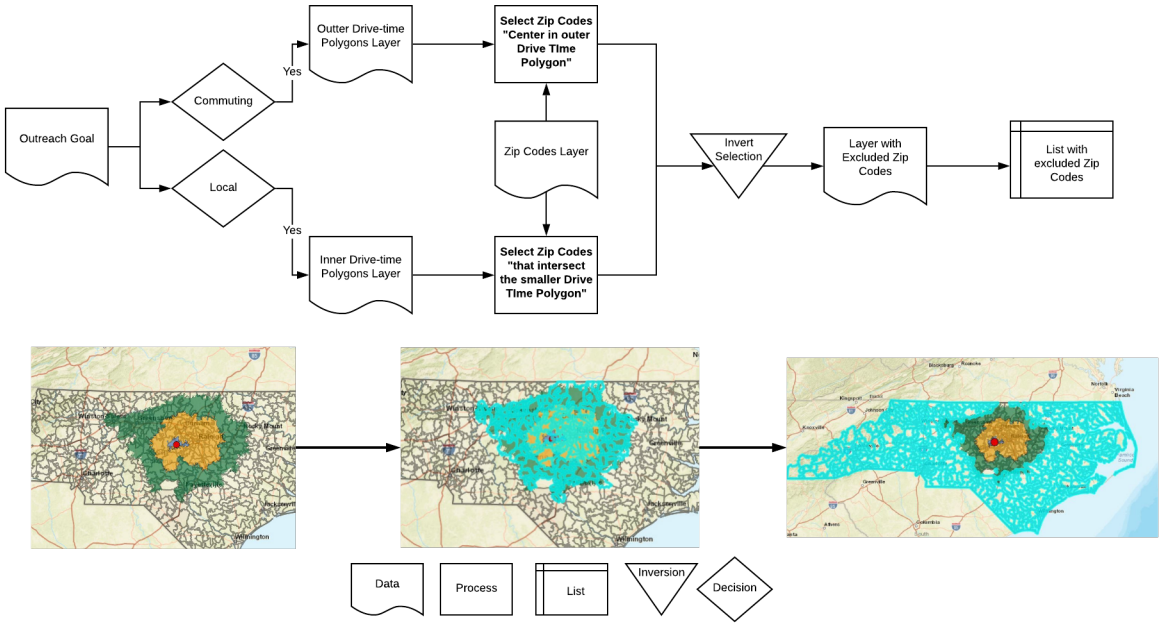

Figure 2: Identification of locations that need to be excluded.

Demographic Tool

We performed a robust literature review to identify the demographics of people who may have a greater interest or be more relevant to a specific transportation project. We aim to target and

engage specific population groups whose feedback and input may be more constructive in the planning and development process. For instance, consider people who regularly use public transportation to go to work, their feedback on a new public transportation project may be more applicable. Demographic profiles were created for stakeholders that use public transit (Demographic Category 1) and stakeholders who commute by car (Demographic Category 2). Stakeholders that may be affected by the implementation of a public transportation project can now be identified more accurately with the use of the corresponding demographic profile.

Demographic Category 1: Public Transit [1]

1. **Age:** Younger people use public transit more often. The majority of trips have been taken by people between 25 and 54 years in age.
2. **Ethnicity:** The single largest group of riders consists of White/Caucasian riders (40%). Combined communities of color make up a majority of riders (60%).
3. **Gender:** Female account for 55.5% while males account for 44.5%.
4. **Household Income:** People below the poverty level are less likely to own, or have access to, a personal vehicle. It has been observed that households with annual incomes less than $25,000 were 10 times more likely, on average, to be zero-vehicle households than households with annual incomes of $75,000 and above.
5. **Household size:** Two people are the most common transit rider household size followed by one-person households.
6. **Education:** Higher education implies more frequent use of public transit.
7. **Population density:** Public transit is more frequent in urban areas

Demographic Category 2: Commuting by Car [2]

1. **Age:** As age increases the car use increases as well.
2. **Gender:** Car use is predicted by being male
3. **Income:** As income increases, car use increases as well

[1] Information has been acquired from [1],[4],[6],[8]
[2] Information has been acquired from [4]

4. **Household size:** The greater the household size the smallest the percentage of households that use public transportation. So, we assume that these households may use an automobile to commute

5. **Population density:** Car use is more frequent in rural areas

Utilizing the tool that was previously created by NCSU (Discussed in section: Previous Work) in conjunction with the relevant demographic profiles that have been identified from the literature review, a population density layer will be constructed. By overlaying the demographic density layer with the drive-time polygons the areas with high concentration of stakeholders with specific demographic profiles will be pinpointed (see Figure 3). Finally, these areas will be used as specific targeting locations for online outreach campaigns (see Figure 4).

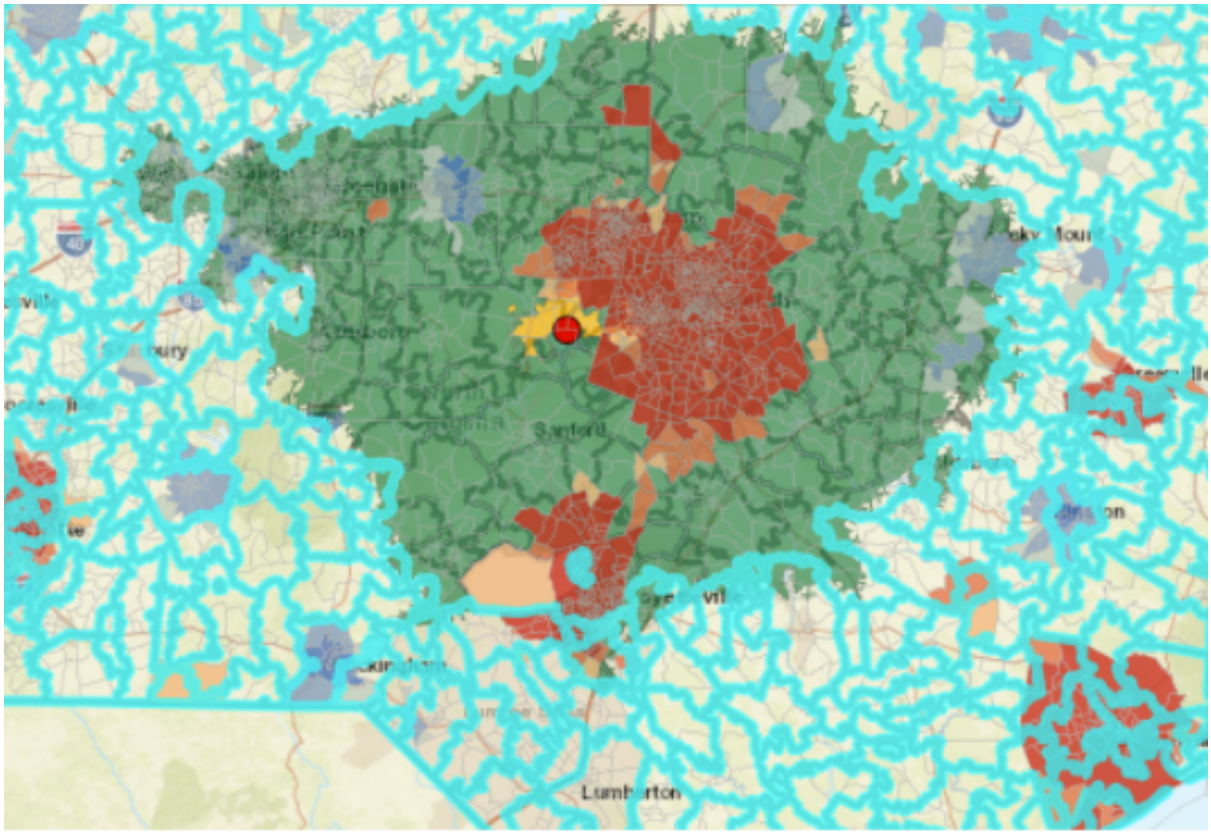

Figure 3: Hotspots of stakeholders with specific demographic profiles

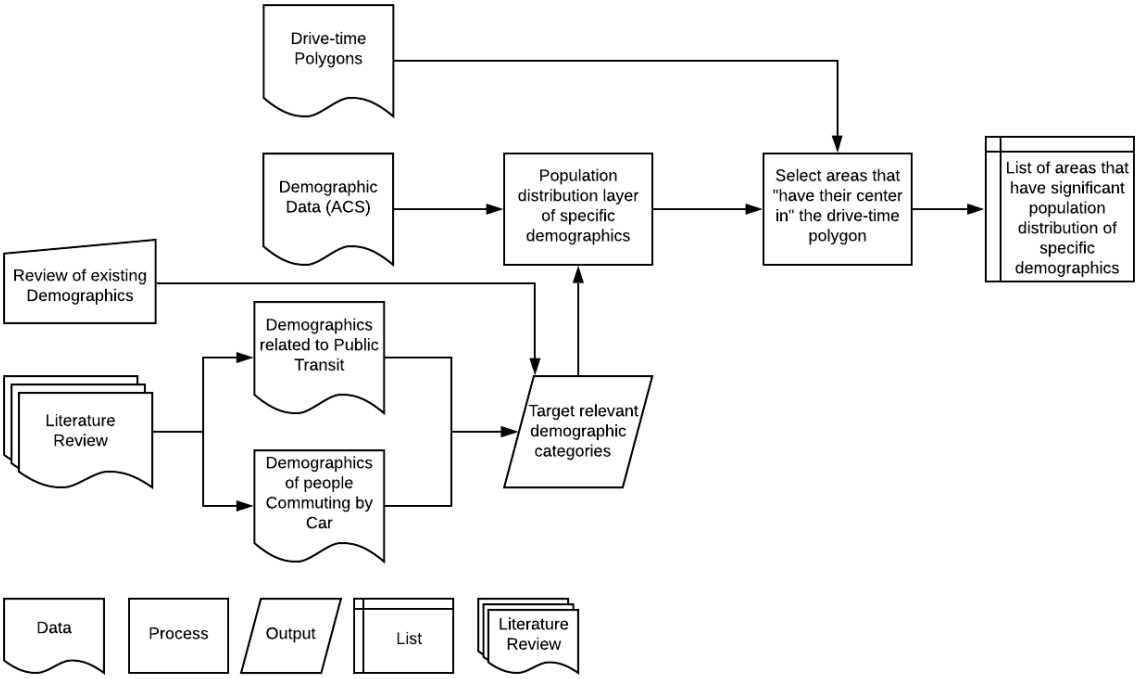

Figure 4: Identification of areas with significant density of specific demographics

Creation of Campaigns

Three outreach goals have been established aiming to capture local and commuting stakeholders as well as stakeholders that may have a greater interest in the implementation of the proposed transportation project (based on their demographic profiles). To successfully target and engage the stakeholders that belong to the above categories we will use our developed spatial and demographic tools in conjunction with the targeting options available in popular advertising platforms. Google's targeting option "People in or regularly in targeted locations" will be mainly used. As explained by Marving in the article "Google is expanding when it shows ads to 'people in targeted locations' ", instead of showing ads to people only when they are physically located in your targeted locations at the time of their search, it will also include people who regularly commute or travel to your targeted locations even when they aren't physically there when they perform a search [2]. Our campaign will also send ads to people who are not in the targeting location but who make searches relevant to the transportation project and regularly commute to the central area. The idea is that the location-targeted campaigns can reach people with ads targeted to their work locations when they are home and vice versa [2].

1. Target Local Stakeholders

We aim to target and engage stakeholders that are local to the project area. This is done using the shortest drive-time polygon and the Google Ads targeting option “People in or regularly in targeted locations”. For targeting local stakeholder in particular, the Google ads exclusion process mentioned above will exclude all areas of the world besides this smallest drive-time polygon (see Figure 5). In this way we only target stakeholders whose main residence is located in close proximity to the project’s site, people who for the most part are constantly in the smallest drive-time polygon. Additionally, the “ad projection time” option will be utilized to further restrict the pool of stakeholders that are eligible to receive the advertisement. For instance, let’s consider a specified campaign aiming to target people that have their main resident inside the small polygon, it will project ads between 7 pm and 12 am, a time period that most companies are closed and consequently employees have returned to their residences.

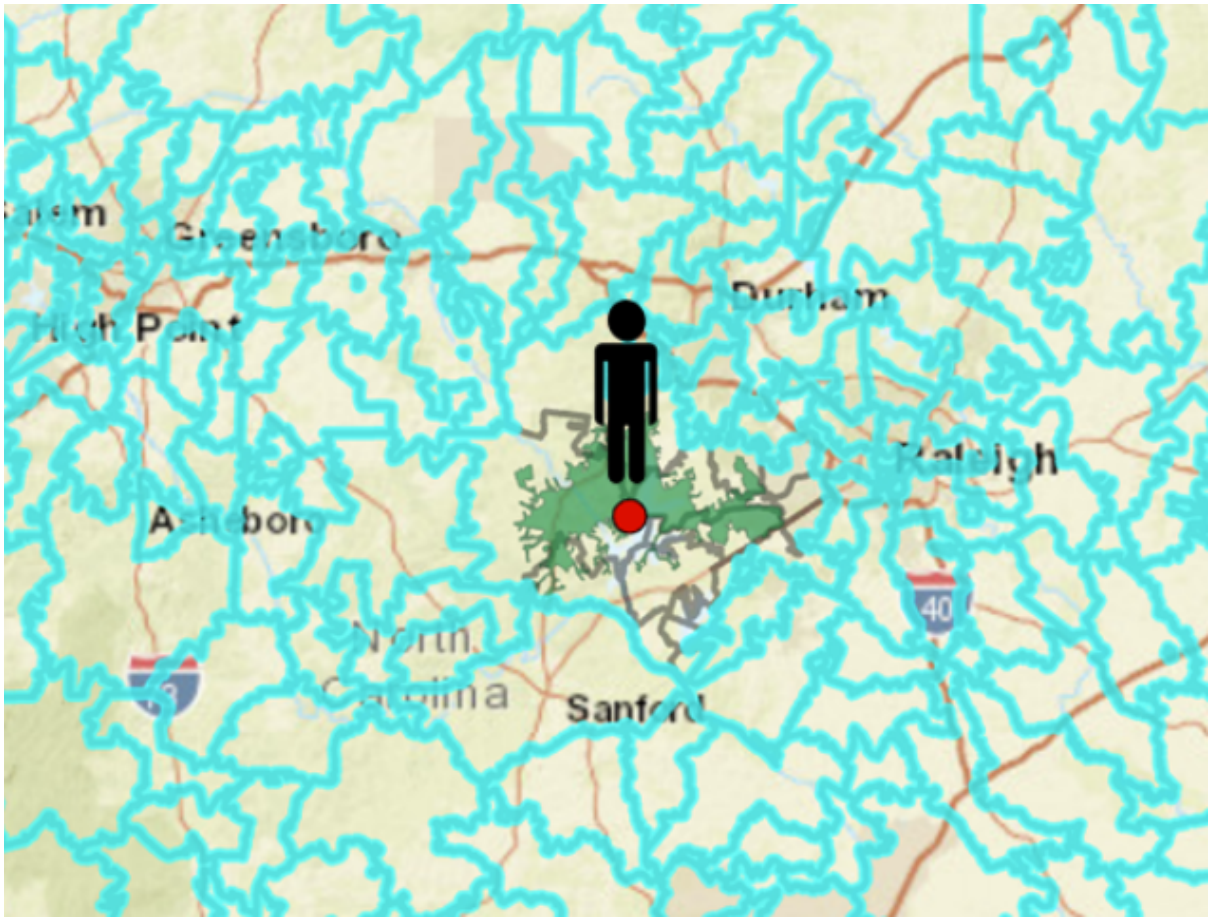

Figure 5: Identification of areas that need to be exclude based on the shortest drive-time polygon

2. Target commuting stakeholders that regularly commute to or through the project's location

We aim to target commuting stakeholders that regularly commute to or through the project's site. The Google Ads targeting option "People in or regularly in targeted locations", as its name indicates, will send advertisements to people that are often in the targeted location but also to people who regularly commute or travel to the targeted location even when they are not physically there [2]. Using our location tool, we aim to further restrict the areas that regularly commuting stakeholders reside in. Stakeholders who are within the bounds of the outer drive-time polygon (green region) will only be eligible to receive our advertisements. Stakeholders who are located outside the outer drive-time polygon (green region) will not be targeted no matter how often they travel to the targeted location (yellow region) since they are in regions that have been excluded. Again, the Google Ads exclusion process using the targeting option "People in or regularly in targeted locations" is used to eliminate all areas of the world except locations bounded by the outermost drive-time polygon (green region) (see Figure 6).

3. Target stakeholders with specific demographic profiles

We aim to target stakeholders who are more relevant or might have a greater interest in specific transportation projects based on their demographic profiles. The location tool in conjunction with the demographic tool will pinpoint areas that have a significant population of stakeholders with a specific demographic profile (see Figure 3).

Robotic Process Automation

It would be very time consuming to individually run the spatial and geographical tools and use their results to structure advertising campaigns for several transportation projects. Thus, Robotic Process Automation will be used to enable the creation of targeted advertising in the most efficient way. Robotic Process Automation, RPA, stands at the process-driven side of the continuum known as Artificial Intelligence. Along with Robotics and Artificial Intelligence technology, scientists and software developers have also constructed programs whose aim is to automate repetitive tasks. Gartner, the well-known news website, defines Robotic Process Automation as follows: "RPA tools perform [if, then, else] statements on structured data, typically using a combination of user interface interactions, or by connecting to APIs to drive client servers, mainframes or HTML code." [9] RPA will call the corresponding spatial and

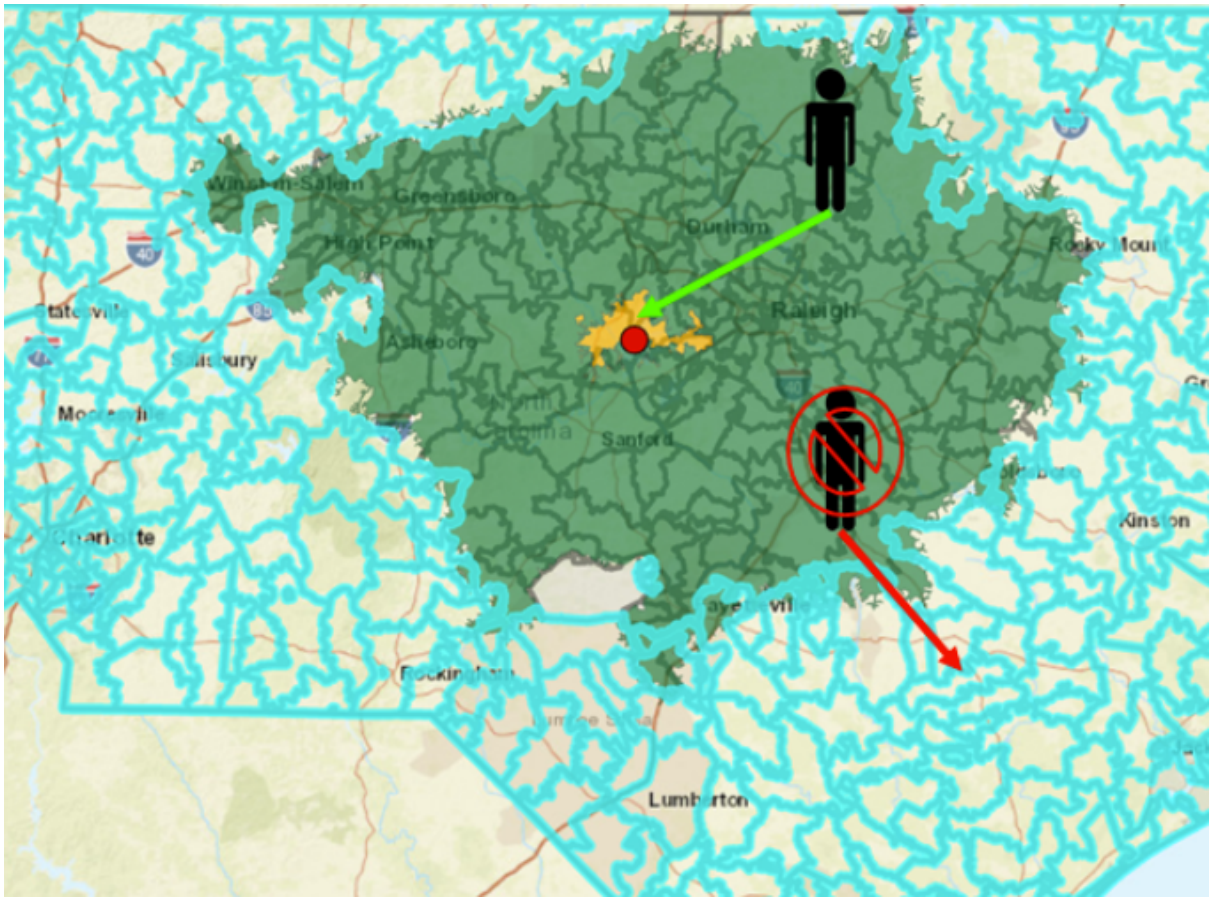

Figure 6: Identification of areas that need to be exclude based on the shortest drive-time polygon

demographic analysis scripts that will run on ArcGIS Pro. The results of the scripts will be saved locally for future use by the RPA. Then, the automated script will locate the CSV files with the outcome of the analysis and use them as input to set up the location targeting in a platform such as Google Ads.

Spatial Engagement Portal (SPEP)

Once the affected and relevant stakeholders have been attracted, they will be directed to the personalized Spatial Engagement Portal (SPEP). The SPEP portal will work with ArcGIS Story Maps and Surveys123. Its aim is to provide detailed information and engage the stakeholders with the perceived project outcomes through a user friendly, highly engaging interface. Additionally, the attracted stakeholders will be able to provide their feedback on the project and any suggestions that they may have. Finally, they will be asked to complete

a quick survey that will help us evaluate our innovative public outreach and engagement approach.

Evaluation

It is vital to assess the success of our proposed public outreach and engagement approach. Some formal methods will be followed which will allow us to gauge the success of our approach compared to other already existing approaches. The guidelines that were developed by the NCDOT in conjunction with the guidelines from the Florida DOT and the Texas DOT will be adjusted and used to measure the effectiveness of our approach. The study will run on two different NCDOT transportation projects. For both projects, two public outreach/engagement methods will be used, the one that NCDOT is currently using and our innovative approach. Finally, the results will be studied to reveal the areas that require improvement and also examine the scientific importance of our findings

Methodology

After a robust literature review, measures for evaluating the effectiveness of public involvement activities have been identified and adjusted to meet our needs. Survey123 will be integrated into our SPEP portal that will help with the evaluation of our targeting and engagement approach. Additionally, Survey123 will allow us to collect location information that is relevant to the affected population. "A logical first step in developing a performance methodology for public involvement is to identify appropriate goals (or what the agency wants to achieve)"[3].

Goal 1:Ensure that all interested parties have an opportunity to participate [3] [3]

The aim of this goal is to evaluate our outreach approach. We need to quantitatively measure if we have successfully identified and targeted relevant stakeholders to the transportation project. Qualitative input also plays a large role in distinguishing people that are local or commuting and are positively or negatively affected. The following questions and questions related to the following categories can assist in the evaluation process and measure the degree of satisfaction of the first goal.

1. Demographic Information of stakeholders relevant to transportation projects
    - Race/ethnicity

[3] Information has been acquired from [3], [5]

- Income
- Age
- Education
- Employment
- Parental Status
- Gender
- Household size
- Household vehicles
- Geographic identifier (Address, Zip Code)
- Minority Groups (underrepresented groups, people with mobility impairment)

2. Do you identify yourself as a local or commuting stakeholder with respect to the project's location?

Additionally, to measure the success of the outreach approach we can calculate the number of clicks to the survey, the number of organizations attracted, and the diversity in the organizations attracted. To measure the actual participation we can calculate the rate of engagement by dividing the total number of people that submitted the form by the total number of people that viewed our advertisement.

Goal 2:Ensure that our engagement method is efficient and effective [4]

The aim of this goal is to evaluate the SPEP portal and its role in the engagement process. These measures require dialogue or interaction with members of the public and are both quantitative and qualitative. They measure whether and how members of the public were engaged and what kind of experience those participants had as a result. Taken together, these objective and subjective measures can be combined to describe how meaningful the public engagement efforts were [5]. The following questions can assist in the evaluation process and measure the degree of satisfaction of the second goal.

1. Convenience - Was our approach the most convenient for initial feedback?
    - Were you provided with contact information for individuals that would address any questions you had prior to the meeting?
    - If you had questions prior to the meeting, were they adequately addressed?

[4] Information has been acquired from [3], [5]

- Did you experience any issues accessing the SPEP portal?
- Availability of information in languages other than English [1]
- Was this information dissemination technique more effective than other antiquated techniques(flyers, post, etc.)?
- Rate how convenient was this method for providing your input with regard to the proposed project

2. Clarity of information (These questions will assist with the Story maps evaluation)

- Was the information you heard beneficial in understanding the project?
- Was the information presented clearly?
- Do you believe the information could have been more easily understood if it was presented differently?
- Do you feel that you have been informed early enough about the project?
- Affected parties feel that ample notice was provided of construction projects
- Stakeholders feel that this engagement method adequately informed them
- Stakeholders feel that this engagement method gave the opportunity to actively participate and express their opinion
- Was the information presented in standard English? Technical terms were not present

All the above indicators align with the Public Involvement Performance measurements that the NCDOT published in 2011. Additionally, to formulate the survey's questions, responses from NCDOT's surveys were taken into account. Qualitative responses that are related to our goals were used to create survey questions to further connect previous stakeholders' concerns with our current approach.

Future Work

Employ Data mining techniques to identify trends. Identify demographics of stakeholders that may be directly or indirectly affected. Additionally, to distinguish stakeholders that may be positively or negatively affected by the proposed transportation project.

REFERENCES

[1] "APTA Passenger Characteristics.pdf." Accessed: Jun. 15, 2020. [Online]. Available: http://filecenter.santa-clarita.com/transit/APTA %20Passenger %20Characteristics.pdf.

[2] "Google is expanding when it shows ads to 'people in targeted locations,' " Search Engine Land, May 31, 2019. https://searchengineland.com/google-is-expanding-when-it-shows-ads-to-people-in-targeted-locations -317601 (accessed Jul. 28, 2020).

[3] "pipmfinalreport06-26.pdf." Accessed: Jun. 22, 2020. [Online]. Available: https://fdotwww.blob.core.windows.net/

[4] J. M. Roos, F. Sprei, and U. Holmberg, "Sociodemography, Geography, and Personality as Determinants of Car Driving and Use of Public Transportation," Behavioral Sciences, vol. 10, no. 6, Art. no. 6, Jun. 2020, doi: 10.3390/bs10060093.

[5] G. Griffin, G. Stoeltje, T. Geiselbrecht, C. Simek, B. Ettelman, and M. Metsker-Galarza, "Texas A&M Transportation Institute PRC 17-89 F January 2018," p. 31.

[6] United States. Department Of Transportation. Bureau Of Transportation Statistics, "Transportation Statistics Annual Report 2018," 2018, doi: 10.21949/1502596.

[7] "Transportation_Network_Analysis_Report_xhong2.pdf.".

[8] "Who Rides Public Transportation," p. 86.

[9] Van der Aalst, W., Bichler, M., and Heinzl, A. 2018. Robotic Process Automation. Business & Information Systems Engineering. 60, 4 (2018), 269-272.