Final Report

Contract BDV24-562-02

EVALUATING COMMUNITY BUILDING EFFECTIVENESS OF TRANSPORTATION INVESTMENTS: USING TRADITIONAL AND BIG DATA ORIENTED ANALYTICAL APPROACHES

Naveen Eluru, Ph.D.

Samiul Hasan, Ph. D.

Essam Radwan, Ph.D., P.E.

Sabreena Anowar, Ph.D.

Mehedi Hasnat, B.Sc.

University of Central Florida

Department of Civil, Environmental & Construction Engineering

Department of Industrial Engineering & Management Systems

Orlando, FL 32816

November 2017

Disclaimer

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the State of Florida Department of Transportation.

Technical Report

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle *EVALUATING COMMUNITY BUILDING EFFECTIVENESS OF TRANSPORTATION INVESTMENTS: USING TRADITIONAL AND BIG DATA ORIENTED ANALYTICAL APPROACHES*		5. Report Date November 2017 6. Performing Organization Code
7. Author(s) Naveen Eluru, Ph.D., Samiul Hasan, Ph.D., Essam Radwan, Ph.D.		8. Performing Organization Report No.
9. Performing Organization Name and Address Department of Civil, Environmental & Construction Engineering Department of Industrial Engineering & Management Systems University of Central Florida, Orlando, FL 32816		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address		13. Type of Report and Period Covered Final 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

Research was conducted to identify current practices in evaluating community impacts of transportation infrastrcuture investments. An extensive survey of contemporary literature was conducted towards that end. Variation in property price is the most commonly evaluated indicator of community impact of new and improved transportation facilities. While the results reported are mixed; a majority of the studies found that improved accessibility provided by the facility improvement usually resulted in property price increase. Based on the review, several measures of effectiveness were proposed including property price/rent variation, pedestrian/bike crash distributions, proportion of severe crashes, crime rate and ridership changes, land use mix change, proportion of transit/bike/walk commuters, and jobs proximity index. These indicators/measures can be developed by collating appropriate data from different sources using the ArcGIS platform. Data for developing the indicators can be collected from different sources including American Fact Finder, Florida Geographic Data Library (FDGL), Florida Department of Transportation (FDOT), Florida Department of Revenue (FDOR), US Census Bureau, Environmental Protection Agency (EPA) and other online data repositories. In addition, social media data from Twitter was collected and analysed. Our analysis of the collected data indicated that social media data do have some readily available indicators for measuring community building impacts of transportation projects.

17. Key Words	18. Distribution Statement
COMMUNITY BUILDING, TRANSPORTATION INVESTMENT IMPACTS, SOCIAL MEDIA DATA	No restrictions.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	49	

Executive Summary

Transportation infrastructure investments not only enhance mobility and connection between regions, but also play a major role in shaping and transforming surrounding communities. Therefore, any evaluation of the impacts of infrastructure investments has to consider impacts on both the system users and the communities affected. The proposed research effort attempts to examine the role of transportation infrastructure investments in community building measures. Toward that end, a two-pronged research strategy was adopted combining both traditional and big data oriented analytical approaches.

In the first part of the research, we conducted an exhaustive literature review to identify state-of-practice and state-of-art of examining community impacts of different system components of transportation system. We found that changes in property prices was the most commonly evaluated indicator of community impact of new and improved transportation facilities (for example, roadway expansion, transit facility improvement, bikeshare facility installation). The results reported in the studies were mixed; however, the majority of them found that improved accessibility provided by the facility improvement usually resulted in property price increase. Informed from the review, several measures of effectiveness were proposed and the potential data sources for developing the measures were also identified including property price/rent variation, pedestrian/bike crash distributions, proportion of severe crashes, crime rate and ridership changes, land use mix change, proportion of transit/bike/walk commuters, and jobs proximity index.

In the second part of the research, we collected and analyzed social media data collected from Twitter to examine community perception of the major transportation projects in the Central Florida region. We described the procedure of collecting data from Twitter using query scripts. Our analysis of the collected data indicated that social media data do have some readily available indicators for measuring community building impacts of transportation projects. Other indicators can also be developed by running sophisticated data mining techniques.

As part of this research, we will continue collecting data for further analysis. In the next phase of this project, we will generate the proposed measures of effectiveness for selected transportation projects. In addition, we will conduct topic and sentiment analysis using the collected data. These analysis techniques applied over the filtered data will enable us to gather valuable insights on how transportation investments help to build communities.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

1.1 BACKGROUND

Transportation infrastructure investments are intended to facilitate and enhance the movement of people and goods. However, in addition to building connections across regions and affecting the mobility of the system users, these investments impact land use, urban residential location decisions and activity patterns, economic growth, overall quality of life and community well-being (Andersson et al., 2010a). Further, emerging transportation infrastructure (such as connected vehicles and infrastructure, driverless cars, electric cars) and analytics (social media and big data approaches, machine learning methods) are likely to play a major role in transforming existing cities into Smart Cities comprised of Smart Communities. Given the critical role of transportation, it is important to examine the influence of transportation projects on overall community building, quality of life and well-being.

Transportation infrastructure investments include investments in building a new roadway, extending or improving capacity of an existing roadway, introducing new transit facilities, installing additional stations or stops to expand transit coverage, installing walk and bike infrastructure. The impacts of these investments can be classified into two broad categories: transportation system effects that result in direct benefits for system users (drivers, passengers, companies) and community (social and economic) effects that affect the community as a whole. There are well-defined performance measures, based on engineering and economic criteria, for assessing the direct system user benefits. For example, how a new facility leads to reduced journey time or reduced travel cost. On the other hand, such indicators are scarce for assessing the impacts of transportation projects on community.

In recent years, there is growing interest toward evaluating community impacts in the research community and policy makers. It has stemmed from the recognition that transportation projects that benefit a subset of users might create negative externalities for the adjacent community members. For instance, a highway expansion might provide better accessibility and faster travel times between an origin (such as suburbs) and a destination (such as central business district). However, it is likely to expose the residents of the communities adjacent to the highway to increased air or noise pollution or even divide the existing community and reduce accessibility to local amenities (social exclusion). Any evaluation of the impact of the highway expansion has to consider impact on system users and communities affected.

So, what is community impact? Simply put, *"these are the effects that any transportation project or investment has on adjacent neighborhoods and communities."* It includes *"the quality of the local environment as experienced by people who live, work or visit there as a consequence of changes in noise, views, walking environment, land use mix and community cohesion (the quality of interactions among neighbors). Related impacts on property values can also be included, and differential impacts on vulnerable population groups may also be covered"* under this definition (http://bca.transportationeconomics.org/benefits/community-impacts). Clearly, the concept is qualitative and subjective. The influence on community members is far from homogenous. Thus, comprehensive community impact assessments are inherently complex than assessing system user impacts and a single cumulative index or measure is not generally sufficient. Both positive and negative impacts need to be assessed – the positive impacts would certainly give indication of the success of the project while the negative impacts would help formulate mitigating measures to improve community well-being. A general overview of the interaction between system effects and community effects is represented in Figure 1.

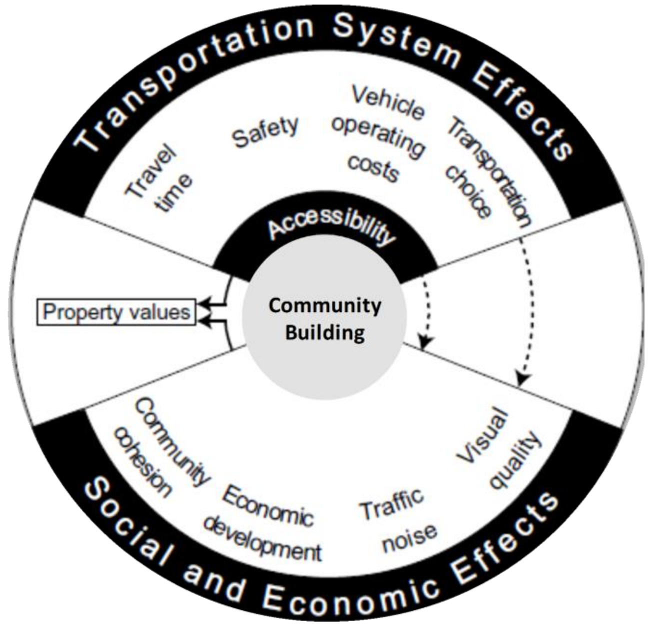


Figure 1: Interaction of System and Community Effect
(Source: Forkenbrok and Weisbrod, 2001; Figure 1.2)

1.2 CURRENT RESEARCH

The proposed research effort is geared towards examining the role of transportation infrastructure investments in community building measures. Towards that end, we adopt a two-pronged research approach. First, we do an exhaustive literature review to identify state-of-practice and state-of-art of examining community impacts of various transportation features. Informed from the review, several measures of effectiveness are proposed and the potential data sources are identified. Second, we collect social media data and analyze it to examine community perception of the major transportation projects in the Central Florida region.

The report is organized as follows. Chapter 2 contains the literature review followed by the measures of effectiveness in Chapter 3. In Chapter 4, we discuss the social media data collection procedure and analysis results while Chapter 5 concludes the report.

CHAPTER 2

2.1 SUMMARY OF REVIEWED LITERATURE

There is scarcity of literature that evaluate the community development impact of transportation projects and investments. Our objectives are:

- review and compile contemporary studies on this issue (since the 2000's)
- identify and document the indicators used by previous research efforts
- summarize the results obtained

Towards that end, more than 50 publications were reviewed including published academic research – within and beyond transportation domain (social science, health, urban planning, urban economics, environment), non-academic articles, and published governmental reports. This report provides a complete compilation of reviewed works (attached matrix of studies), and a summary of key findings. To be sure, different projects are aimed at modifying/improving/developing different components of the transportation system. Figure 2 identifies the components – Infrastructure, transit facility and non-motorized facility - that we focused our review on.

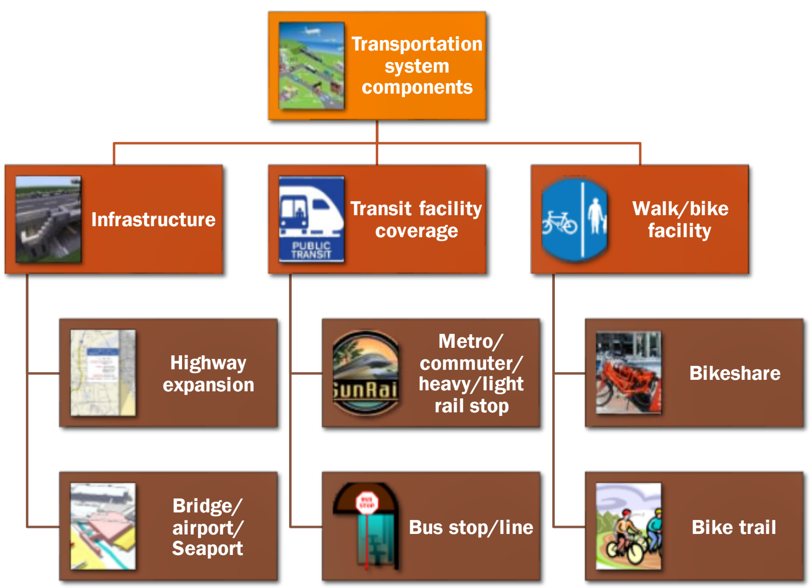


Figure 2: Transportation System Components Chosen for Review

2.2 INVESTMENT IN INFRASTRUCTURE

From our review, we have found that there is vast empirical literature on the effects of improved accessibility brought about by new or improved roadway infrastructure (such as roads, bridges, airports and seaports). Table 1 lists the studies that we reviewed in this category. Several observations can be made from the table.

- The most commonly investigated indicator of community development is the sales price of properties including residential, commercial, retail, office, food, plaza, industrial, and vacant land as a result of new highway development, expansion of highway, construction/opening of new bridge/tunnel, opening of tolled roads, and expansion of airport facility

- Hedonic regression technique is the most prevalent methodology applied
- The variation in sales price is investigated as a function of proximity (how far the properties are located from the roadway) and the noise level within a certain buffer distance
- The results obtained are mixed. However, the majority of the studies found that increased accessibility brought by the facility increases values of residential properties
- As expected, nuisance from noise negatively impacts property price. The price reduction is of the order of 1-3%. Andersson et al. (2010b) found that road noise has larger negative impact than rail noise
- Hamersma et al. (2017) investigated resident's satisfaction due to new highway construction and found that new residents who moved to the neighborhood after highway construction expressed more satisfaction than the existing residents
- Kang and Cervero (2009) found that conversion of freeway to greenway increases property price

Table 1: Literature on Roadway Infrastructure

Study	Region	Evaluated Measure	Property Type	Dependent Variable and Methodology	Result
Levkovich et al., 2016	Netherlands	Proximity (distance from interchange and highway)	Residential	Housing price, Repeat sales/ difference-in-difference	• Positive effect of increased accessibility outweighs the negative effects
Gingerich et al., 2013	Windsor, Canada	Proximity (properties within 800m buffer of highway ramp)	Commercial, retail, office, food, plaza, industrial, and vacant	Sales price, Spatial regression model	• No significant correlation except for a negative impact on price of vacant land
Iacono and Levinson, 2011	Minnesota, USA	Proximity (dummy for location within ¼ -1mi of upgraded highway)	Residential	Sales price, Hedonic regression	• 100-m increase in distance from the nearest access point on an upgraded highway link reduced property price by 0.3% • Proximity to expanded highway's Right of Way (ROW) reduces housing price upto ¼ mile
Blanco and Flindell, 2011	London and Birmingham, UK	Road traffic noise (sound level)	Residential	Offer price, Hedonic regression	• Residents of different geographic region have different willingness-to-pay for lower noise levels
Brandt and Maennig, 2011	Hamburg, Germany	Proximity (dummy for location of house on a wide road) Air and rail traffic noise (sound level)	Residential (condominiums)	Listing price, Hedonic spatial lag regression	• Property prices reduce by 0.23% following a 1 dB(A) increase in road noise
Andersson et al., 2010b	Lerum, Sweden	Road and rail noise (sound level)	Residential (single-family)	Sales price, Hedonic regression	• Road noise has a larger negative impact on the property price than railway noise
Martinez and Viegas, 2009	Lisbon, Portugal	Proximity (distance from network)	Residential	Asking price, Hedonic spatial lag regression model	• Proximity to urban ring roads and radial networks increase property values • Proximity to motorways and roadways with increased office buildings decrease property values
Kim et al., 2007	Seoul, South Korea	Proximity (distance to highway, arterial road, minor arterial)	Residential	Land price, Hedonic regression	• 1% increase in traffic noise reduces property price by 1.3%

		Road traffic noise (sound level)			
Cervero and Duncan, 2002	Santa Clara, USA	Proximity (within ½ mile distance from grade separated freeways or highway interchange)	Office and commercial land	Transaction price, weighted Hedonic regression	• Property location within ½ mile of thoroughfares was associated with lower land values
Hamersma et al., 2017	Netherlands	Highway development	Residential	Residents' satisfaction, Structural equation model	• Residents living in areas closest to highway development has lower satisfaction • Small proportion of the residents perceived an increase in residential satisfaction due to the highway development
Meijers et al., 2013	Netherlands	Construction of a new bridge/tunnel	Residential	Housing price, Hedonic regression	• Increased accessibility increases housing price
	Seoul, South Korea	Freeway replaced by urban stream and linear park	Residential and commercial	Land value, Multilevel hedonic regression	• The conversion resulted in increased land value within 500 meters of the freeway and greenway
Riebel et al., 2008	Los Angeles, USA	Expansion of highway	Residential	Sales price, Combined hedonic spline regression	• Maximum increase in price is observed at a moderate distance from the expanded highway
Theebe, 2004	Netherlands	Expansion of airport and construction of railways	Residential	Sales price, Hedonic regression	• Noise reduced housing price by 3%-10%
Boarnet and Chalermpong, 2003	California, USA	New tolled roads	Residential (single-family)	Sales price, Hedonic regression	• Accessibility benefits created by the new tolled road increase the housing price
Smersh and Smith, 2000	Jacksonville, USA	Construction of bridge	Residential	Sales price, Repeat sales regression	• Differential effects are found at different ends of the bridge

2.3 INVESTMENT IN TRANSIT

We considered rail and bus transit system in our review. The majority of the studies focus on rail transit. Rail transit system comprised of heavy rail, commuter rail, rapid/high speed rail, metro/subway, and/or light rail. Investment in rail transport system is reported to affect local economy at macro-, meso-, and micro-level (Banister and Thurstain-Goodwin, 2011). Macroeconomic studies use aggregate time-series data and examine the linkage between infrastructure and regional growth measured in terms of GDP or employment growth or population growth (Atack et al., 2010). At the meso-level, agglomeration economies, such as how traffic congestion impact productivity in cities and labor market effects are assessed. In micro-level studies, land and property market effects are examined. The findings from these studies provide guidance for the adoption and implementation of transit finance strategies and thus their importance is widely recognized in the transportation economics and planning literature (Ko and Cao, 2013). For the purpose of this review, we focus our attention on micro-level studies. Table 2, Table 3, and Table 4 list the studies that we reviewed in this regard. Several observations can be made from these tables.

- The impact of accessibility benefits of rail facilities is mostly investigated by examining the values of properties sold before and after the opening of the facility. Some researchers have explored pre-opening anticipatory effects of rail transit lines on property values ("announcement effect") as well (Li, 2016; McMillen and McDonald, 2004; Bae et al., 2003) and found that announcement of new facility opening increases property price
- Property values are represented in terms of sales/transaction price, assessed market value, or rental rates. For residential properties, these data are extracted from the assessor's data, parcel data, or multiple listing service (MLS) data while the rental rates were obtained either from self-administered surveys or rental offices of apartment complex
- Controlling for a wide range of other features such as physical attributes of the housing and neighborhood characteristics, the impact of rail system on the residential and non-residential stock has mainly been examined through proxies of rail accessibility, proximity, and service quality measures (Armstrong and Rodriguez, 2006; Debrezion et al., 2011)
- The studies are mainly cross-sectional. A few studies used repeated sales price data (McMillen and McDonald, 2004; Grimes and Young, 2010) or employed difference-in-difference methodology based on openings of stations (Gibbons and Machin, 2005; Li, 2016)
- Hedonic pricing models and its extensions are the most prevalent methodology applied; the functional forms vary from study to study
- While there are plenty of studies investigating the price changes in residential property types, limited efforts were devoted to non-residential properties – lack of data being the major hindrance
- Although the results are mixed, most studies concluded that investment in rail corridors generally increases property prices. According to urban economics, this is the due to the increase in the accessibility of the corridor relative to the whole transportation network.

However, the accessibility benefits seem to be localized and decline with distance, both for residential and non-residential properties (Ko and Cao, 2012). In addition, we also observed that railways stations impact residential and non-residential property types separately. The extent of the impact area of railway stations is larger for residential properties, whereas the impact of a railway station on commercial properties is limited to immediately adjacent areas (Debrezion et al., 2011)

- Several researchers examined the impact of rail transit on incidence of crimes. Among these, Tay et al. (2013) and Robin et al. (2003) didn't find any significant correlation. However, Bowes and Ihlanfeldt (2001) reported increased crime rate within half-mile radius of rail station
- Among other effects, researchers have investigated how rail transit is associated with vehicle ownership, vehicle miles traveled, transit ridership, and health

In contrast, it was surprising to find that only a handful of studies have investigated the impact of bus transit, although bus transit has a larger network in the region and carries a larger share of transit passengers. Due to their extensive network, effects of bus transit system on property values and community development is more likely to be regional as opposed to the localized (as it is for rail transit). Table 5 lists the studies that we reviewed in this regard. The following observations can be made from these tables.

- Only a few studies attempted to examine the effect of bus transit accessibility. Interestingly, researchers found that proximity to bus stops has no significant association with property price but it negatively impacted apartment rents

Table 2: Literature on Rail Transit Impact on Property Price/Rent

Study	Region	Type of Rail	Effect Evaluated (Measure)	Property Type	Dependent Variable and Methodology	Main Results
Li, 2016	Beijing, China	Metro	Accessibility (distance to the closest station (<1 km))	Residential	List price, Hedonic regression	• 3.8% price increase for properties located within 1 km from the closest station
Ko and Cao, 2013	Minneapolis, USA	Light rail	Accessibility (network distance from station)	Commercial, industrial	Sales price, Hedonic regression	• Price increases non-linearly for properties located within 0.9 miles of stations
Gingerich et al., 2013	Windsor, Canada	Light rail	Proximity (properties within 200/400m buffer of rail line)	Commercial, retail, office, food, plaza, industrial, vacant	Sales price, Hedonic spatial lag regression	• Industrial property price increases with increased proximity • The reverse impact is observed for food and commercial services
Mayor et al., 2012	Dublin, Ireland	Commuter rail, light rail, train	Accessibility, proximity (indicator variables for house location within 250m-2km of stations and Right of Way (ROW))	Residential	Purchase price, Hedonic regression	• Properties within 500m-2km of light rail stations experience 7-17% higher price • Properties within 250m-500m of train stations experience 7-8% higher price
Duncan, 2011	San Diego, USA	Light rail	Accessibility (network distance to the nearest station)	Residential (condominiums)	Sales price, Hedonic regression	• Station proximity with good pedestrian environment increase condo price
Debrezion et al., 2011	Amsterdam, Rotterdam and Enschede, Netherlands	Commuter rail	Accessibility (network distance to the nearest and most frequently used station) Service quality (service quality index)	Residential	Transaction price, Hedonic regression	• Housing price is more affected by the distance from the most frequently used station
Andersson et al., 2010a	Taiwan	High speed rail	Accessibility (network distance to the station)	Residential	Sales price, Hedonic regression	• High ticket price and inaccessible locations results in small or negligible increase in land values
Koster et al., 2010	Netherlands	Passenger rail	Accessibility (network distance to the nearest station)	Residential	Repeated sales price, Hedonic regression	• Property values increase by about 1.5−2% with every km reduction in distance from the nearest railway station

Martinez and Viegas, 2009	Lisbon, Portugal	Metro, light rail	Accessibility (walk time to the station)	Residential	Advertised asking price, Hedonic spatial lag regression	• Proximity to rail facility increases property asking price • Increase amount varies with varying accessibility
Shin et al., 2007	Seoul, South Korea	Subway	Accessibility (distance and walk time to the nearest station)	Residential (apartments)	Actual sales price, Hedonic spatial lag regression	• 1% increase in walking time reduces sales price by 0.017%-0.021% • 1% increase in system wide accessibility reduces sales price by 0.051%-0.076%
Hess and Almeida, 2006	New York, USA	Light rail	Accessibility (straight line and network walk distance)	Residential	Assessed value, Hedonic regression	• Properties within ¼ mile of train stations experience 2-5% higher price • Effects vary in magnitude for different stations in the system – premium is higher in high income area stations
Armstrong and Rodriguez, 2006	Eastern Massachusetts, USA	Commuter rail	Accessibility (network distance from station by foot and by car) Proximity to right-of-way (drive time to the nearest highway interchange and commuter ferry boat)	Residential (single-family)	Sales price, Hedonic spatial lag regression	• Properties within ½ mile buffer of stations experience 9.6%-10.1% higher price • 1-minute increase in drive time, property values decrease by 1.6% • Every 100ft distance from ROW increases property values between $73.21-$289.72
Celik and Yankaya, 2006	Izmir, Turkey	Subway	Accessibility (distance from subway station)	Residential (multi-family)	Asking price, Hedonic regression	• 1-meter additional distance decreases the property values by $4.76
Gibbons and Machin, 2005	London, UK	Subway	Accessibility (distance to the nearest station) Proximity (distance to the ROW)	Residential	Sales price, Hedonic spatial lag regression	• 1-km reduction in distance increase property values by 1.5%
Bae et al., 2003	Seoul, South Korea	Subway	Proximity (distance to the ROW)	Residential (condominiums)	Sales price, Hedonic spatial lag regression	• Distance to ROW impacted sales price prior to the opening of subway line

Clower and Weinstein, 2002	Dallas, USA	Light rail	Accessibility (distance from station)	Office, retail, industrial, residential (single and multi-family)	Assessed value, aggregate change in value	• Price of office properties within ¼ mile of rail station increased by 24.7% • Price of residential properties within ¼ mile of rail station increased by 38.2% • Industrial properties located further away experienced larger gains • Negligible increase for retail was observed
Cervero and Duncan, 2002	Santa Clara, USA	Light rail, commuter rail	Accessibility (distance from station)	Office and commercial land	Transaction price, weighted Hedonic regression	• Commercial parcels within ¼ mile of light rail station experienced 20% higher price • No capitalization premiums for properties in close proximity to commuter rail station
Bowes and Ihlanfeldt, 2001	Atlanta, USA	Heavy rail	Accessibility (distance from station) Proximity (distance from ROW)	Residential (single-family)	Sales price, Hedonic regression	• Properties within ¼ mile of rail stations have their price reduced by 19% • Price increase for houses located within 1-3 miles
Knaap et al., 2001	Portland, USA	Light rail	Accessibility (distance from station)	Vacant residential land	Sales price, Hedonic regression	• Announcement effect on property sale price was observed

Table 3: Literature on Rail Transit Impact on Crime

Study	Region	Type of Rail	Measure Evaluated	Methodology	Main Results
Tay et al., 2013	Calgary, ‘Canada	Light rail	Number of crimes	Observational before and after analysis	• Crime rates varied (increased, decreased or remained unchanged) in the surrounding communities
Robin et al., 2003	Los Angeles, USA	Light rail	Number of crime (neighborhood and municipality wide)	Piecewise regression model (before and after analysis)	• No significant association between transit facility and crime incidence was observed
Bowes and Ihlanfeldt, 2001	Atlanta, USA	Heavy rail	Census tract crime density	Linear regression	• Increased crime rate for tracts within ½ mile distance of railway stations

Table 4: Literature on Other Impacts of Rail Transit

Study	Region	Type of Rail	Measure Evaluated	Dependent Variable	Methodology	Main Results
Shen et al., 2016	Shanghai, China	Metro	Competitiveness as mobility tool	Vehicle ownership	Binary logit/ Nested logit	• High quality rail service can reduce vehicle ownership
Huang and Chao, 2014	Taipei, Taiwan	Metro	Competitiveness as mobility tool	Vehicle ownership	Count regression (difference-in-difference)	• Extending metro coverage with improved level of service can reduce vehicle ownership
Cao and Schoner, 2014	Minnesota, USA	Light rail	Transit use (use of transit for commute and non-commute purpose)	-	Propensity score matching	• Residents who lived in the area prior the line was opened use transit more frequently • 50-80% increase in ridership
Bhattacharjee and Goetz, 2012	Denver, USA	Light rail	Congestion on adjacent highways	Vehicle Miles Traveled (VMT)	Temporal and spatial mapping	• Light rail reduces congestion, but for a short period of time
Senior, 2009	London, UK	Light rail	Transit use (Changes in frequencies of rail and bus use, modal switching)	-	Before and after analysis	• In the rail corridor, in both short and medium term, rail ridership increased while ridership of bus decreased • Higher frequency of rail usage was observed in the rail corridor
Brown and Werner, 2007	Minnesota, USA	Light rail	Health (bouts of activity) Transit use (ridership)	-	Before and after analysis	• Walk to station was associated with moderate activity bouts • After opening of a new stop, the ridership increased by 19%
Lee and Chang, 2006	Seoul, South Korea	High speed rail	Transit use (change in number of passenger trips)	-	Before and after analysis (1 year)	• Ridership increased in the corridor where high speed rail stations are located • Ridership decreased in other conventional rail corridors where high speed rail stations are not directly accessible
Bowes and Ihlanfeldt, 2001	Atlanta, USA	Heavy rail	Commercial development	Retail employment density	Random effects regression	• No significant impact

Table 5: Literature on Bus Transit System

Study	Region	Type of Rail	Effect Evaluated (Measure)	Property Type	Dependent Variable and Methodology	Main Results
Cao and Hough, 2008	Fargo, USA	Bus transit	Proximity (distance from route)	Residential (apartments)	Monthly rent, Hedonic regression	• Apartments located within 1/8 mile of bus routes are $18.41 cheaper than other apartments
Bina et al., 2006	Texas, USA	Bus transit	Accessibility (density of bus stop)	Residential (apartments)	Monthly rent, Hedonic regression	• Bus stop density negatively impacts rent
Celik and Yankaya, 2006	Izmir, Turkey	Bus transit	Accessibility (distance from bus stop)	Residential (multi-family)	Asking price, Hedonic regression	• No significant effect on property values
Combs, 2017	Bogota, Columbia	Bus rapid transit	Changes in travel pattern (tour frequency)	-	Count regression	• No substantial impact on lower income households to meet daily mobility needs
Combs and Rodriguez, 2014	Bogota, Columbia	Bus rapid transit	Competitiveness as mobility tool (vehicle ownership)	-	Difference-in-difference	• Reduces vehicle ownership in high income households • Reverse impact for low income households
Cervero and Kang, 2011	Seoul, Korea	Bus rapid transit	Proximity (distance from bus stop)	Residential, non-residential	Land use type, Multinomial logit Land price, Hedonic regression	• Land price increased by 10%
Munoz-Raskin, 2010	Bogota, Columbia	Bus rapid transit	Accessibility (properties within 10 minutes of walking distance of the system)	Residential	Housing price, Hedonic regression	• Price of middle-income properties increase • Reverse impact for low-income properties

2.4 INVESTMENT IN WALK/BIKE FACILITIES

Given the wide ranging implications of over-reliance of automobiles for personal travel, policy makers are trying to promote non-motorized modes as potential alternatives, at least for short distance utilitarian trips. Recently, governments are investing more in infrastructure facilitating walking and biking to popularize them among the general public. Although the positive impacts of cycling are widely known, there are very few studies that actually studied community impact. Table 6 lists the studies that we reviewed in this regard. Several observations can be made from these tables.

- Of the four studies on bike facilities, two are on bikeshare and two on bike trails. Properties in the vicinity of bikeshare stations experience higher prices (El-Geneidy et al., 2015) while bikeshare stations also induce economic and retail activities (Buehler and Humrey, 2015). Interestingly, bike trails negatively impacted housing price in suburban areas (Krizek, 2006)
- Walkability is an important attribute that has been linked to quality of life in many ways. Health related benefits of physical exercise and walking, mental health benefits of reduced social isolation and increased social interaction are a few of the many positive impacts on quality of life that can result from a walkable neighborhood. While the health and environmental implications of walkable communities are being extensively studied, the social benefits have not been investigated as broadly. The few studies that we found, almost all of them reported that increased walkability increases property price. A negative association of mortgage default probability with walkability of neighborhood was found in Rauterkus et al. (2010)

Table 6: Literature on Walk/Bike Facilities

Study	Region	Type of Facility	Measure	Property Type	Dependent Variable and Methodology	Result
El-Geneidy et al., 2015	Montreal, Canada	Bikeshare (BIXI)	Presence of bikeshare stations	Residential	Repeated sales price, Multilevel longitudinal hedonic regression	• Presence of bikeshare system in a neighborhood increases the property value by 2.7%
Pivo and Fischer, 2011	USA	-	Walkability via Walkscore	Office, retail, apartment, industrial	Market value, income return, capital return, total return, Linear regression	• 10-point increase in walkability increases office, retail and apartment values by 1-9% • No effect on industrial properties
Rogers et al., 2011	New Hampshire, USA	-	Walkability	-	Social capital, Correlation	• Neighborhood walkability is positively linked with community well-being
Rauterkus and Miller, 2011	Alabama, USA	-	Walkability via Walkscore	Residential, commercial	Sales price, Linear regression	• Increased walkability increase land value and the effect is stable over time
Rauterkus et al., 2010	Chicago, Jacksonville and San Francisco, USA	-	Walkability via Walkscore	Residential	Mortgage default, Probit regression	• Walkability is associated with a lower mortgage default probability in high income areas • Mortgage default probability increases with higher walk Scores in low income areas
Krizek, 2006	Minneapolis, USA	Bike trails and lanes	Proximity to bike facilities	Residential	Sales price, Linear regression	• In suburban areas, bike facilities negatively impact home values
Buehler and Humrey, 2015	Washington DC, USA	Bikeshare (Capital Bike)	Economic (Users' willingness to spend, perception of business owner)	-	Intercept survey of users and business	• 23% of the patrons were likely to spend more due to bikeshare facility • 20% of the business thought bikeshare had a positive impact on sales
Merom et al., 2003	Sydney, Australia	Bike trail	Trail usage Walking and cycling activity	-	Before and after analysis (bike count, change in walking and cycling hours)	• Mean daily bike count increased • Trail usage was higher among bike owners living near the trail

CHAPTER 3

3.1 MEASURES OF EFFECTIVENESS (MOE) AND DATA SOURCE

There are several ongoing major transportation projects in the Central Florida region including second phase of SunRail commuter rail extension, I-4 expansion, pedestrian and bicycling facility installation, and bicycle sharing system (Juice) introduction. Although the regional boundary encompasses nine counties (Brevard, Flagler, Lake, Marion, Orange, Osceola, Seminole, Sumter and Volusia) within District 5, Polk county within District 1 and part of Indian River county in District 4 of FDOT, we confine our study to only District 5 counties (see Figure 3).

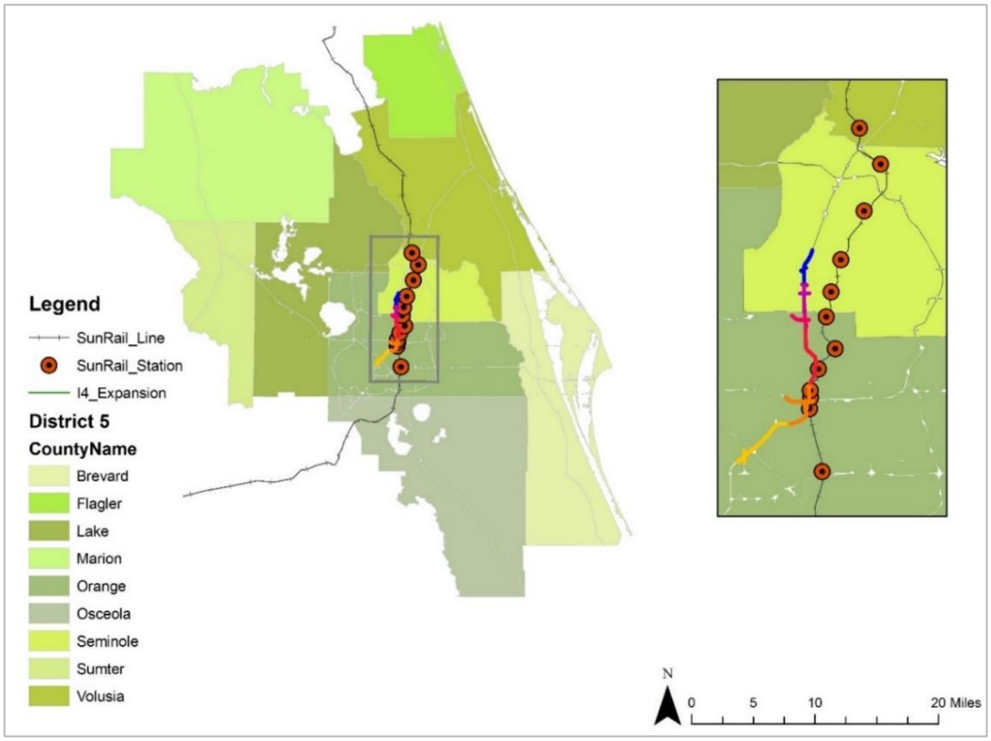


Figure 3: Study Area with Two Major On-going Transportation Investment Projects

To be sure, the development of measures of effectiveness is a data intensive process. These indicators/measures can be developed by collating appropriate data from different sources using the ArcGIS platform. These data for developing the indicators can be collected from different data sources including American Fact Finder, Florida Geographic Data Library (FDGL), Florida Department of Transportation (FDOT), Florida Department of Revenue (FDOR), US Census Bureau, Environmental Protection Agency (EPA) and other online data repositories. Of these, the Smart Location Database (SLD) that is obtainable from the EPA website, summarizes more than 90 different indicators associated with the built environment and location efficiency along with various demographic and employment statistics. Most of the attributes of the database are available for all U.S. block groups and developed for the year 2010. So, it is a good starting point for developing base year community building assessment measures. The potential data sources are presented in Table 7.

Table 7: Data Sources

Name	Description	Web Link for Accessing Data
American Fact Finder	Population and Economic census, American Community Survey (ACS), American Housing Survey (AHS)	https://factfinder.census.gov/
Florida Department of Revenue	Parcel level sales data, Land use data	http://floridarevenue.com/
Florida Geographic Data Library	Spatial layers of transportation data in Florida	http://www.fgdl.org/
Bureau of Labor Statistics	Employment data	https://www.bls.gov/
US Census Bureau	Population and Economic census, American Community Survey (ACS)	https://www.census.gov/
Bureau of Transportation Statistics	Spatial layers of data by mode	https://www.transtats.bts.gov/
GISInventory	Spatial layers	https://www.gisinventory.net/
Housing and Transportation Affordability Index	Housing and transportation cost	http://htaindex.cnt.org/map/
Environmental Protection Agency	Spatial layer of smart location indicators	https://edg.epa.gov/data/
US Government Open Data	Employment and education data	https://catalog.data.gov/dataset/
US Department of Housing and Urban Development	Spatial layers of jobs and labor market	https://egis-hud.opendata.arcgis.com/
Bureau of Transportation Statistics	Transportation facilities, networks and infrastructures	https://www.rita.dot.gov/

Employing the above identified data sources and informed from the literature review, we propose several measures of effectiveness to evaluate the community building effects of the major projects currently underway in Central Florida. A discussion of these measures of effectiveness and some preliminary analysis results are provided in the ensuing discussion.

3.1.1 Property Price/Rent Variation

The changes in price/rent, before and after the investment, could be examined by creating different sized circular/polygon buffers (0.25/0.5/1/2 mile) around the transportation facility under consideration. Network distance between the parcel centroid and the nearest rail station, bikeshare station, and rail Right of Way (ROW), highway ROW could be used as accessibility and proximity indicators, respectively. The sales/rent data obtained from Florida Department of Revenue (FDOR) will be employed for the analysis. This MOE can be developed for SunRail expansion, I-4 expansion and bikeshare/bike trail projects.

3.1.2 Pedestrian/Bike Crashes

For this, disaggregate level geocoded crash data before and after the opening of the stations/highway expansion is needed. Then, circular/polygon buffers (0.25/0.5/1 mile) centered around the rail stations/highway could be created to identify the surrounding communities. Later on, the pedestrian/bike crashes in the communities within the buffer could be counted and analyzed.

3.1.3 Proportion of Severe Crashes

Enhancement in highway facilities allows faster travel. However, higher vehicle speed is an important indicator for increased crash casualties. Therefore, safety and sustainability of neighborhoods adjacent to roadway facilities can be compromised if they are exposed to vehicle speed above acceptable level. So, proportion of severe crashes before and after a highway infrastructural improvement could be a useful MOE to evaluate the community impact of such projects. Disaggregate level crash data from FDOT is needed for this analysis. After geocoding the crash data, it could be intersected with the highway buffer and count of crashes per severity before and after the expansion could be computed.

3.1.4 Bus Transit Ridership

This MOE can be evaluated for both SunRail extension and I-4 expansion projects. For instance, to check, if the opening of the commuter rail station impacted the bus ridership, we can conduct a before and after analysis of the bus ridership within the rail station buffer (0.25/0.5/1/2 mile). The data on bus ridership can be obtained from LYNX. The bus stops within the buffer need to be identified first and then the quarterly ridership data could be combined to get the boarding and alighting before and after the station opening. Figure 4 shows the bus stops within 2-mile buffer of SunRail stops and Table 8 shows the variation in boarding and alighting before and after the opening of the SunRail stations in the bus stops within the buffer.

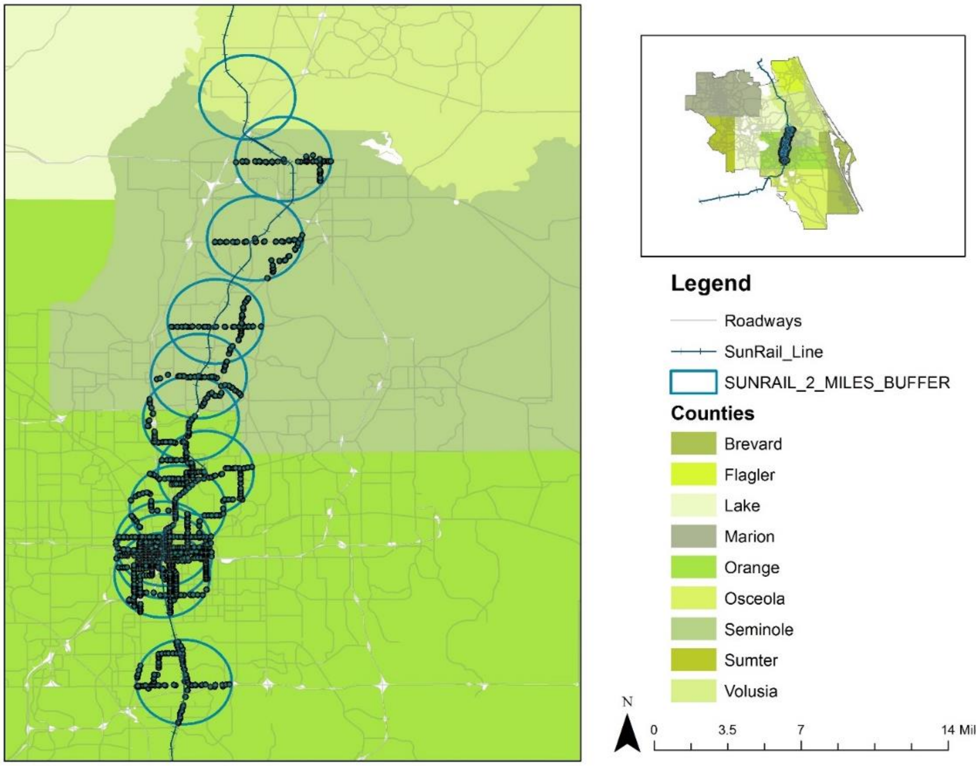


Figure 4 : Bus Stops within Sun Rail Station Buffer (2mi)

Table 8: Variation in Boarding and Alighting in Bus Stations within Station Buffer (2 mi)

Sl. No	SunRail Station Name	No of Stops	Boarding		Alighting	
			Before	After	Before	After
1	Sand Lake Station	108	3,261.53	4,376.83	3,104.32	4,472.54
2	AMTRAK Station	416	1,708.08	2,679.28	1,673.61	2,732.93
3	Church Street Station	441	1,621.83	2,678.01	1,671.18	2,643.64
4	Lynx Central Station	412	8,831.26	11,460.34	9,053.42	12,445.96
5	FL Hospital	206	17,515.10	45,411.11	17,454.28	44,356.22
6	Winter Park	142	2,532.30	4,041.22	2,370.28	4,075.20
7	Maitland	73	1,165.73	1,285.67	1,147.92	1,312.84
8	Altamonte Springs	29	2,291.61	5,510.90	2,447.56	5,237.13
9	Longwood	54	1,512.64	1,675.80	1,582.77	1,763.40
10	Lake Mary	43	857.43	866.65	800.05	837.10
11	Sandford Station	2	28.03	13.47	25.71	15.61
	Total		**41,325.54**	**79,999.29**	**41,331.10**	**79,892.57**

3.1.5 Crime Rate

For this, disaggregate level crime data before and after the opening of the commuter rail stations from the Florida Department of Law Enforcement (FDLE) is needed. A spatial layer of crime density can also be obtained from ArcGIS online resource. Then, circular buffers (0.25/0.5/1 mile) could be created centered around the stations to identify the surrounding communities. Afterwards, the incidences of crime in the communities within the buffer could be counted and analyzed. The data availability could be a restriction for this measure.

3.1.6 Noise and Air Pollution Level

The noise and air pollution level data from fixed monitoring stations could be collected from the Environmental Protection Agency (EPA). Then land use regression models can be developed (dependent variable would be air pollution and predictors would be land-use and built environment data collected at various buffers in ArcGIS) and using the regression, we can rasterize the area and predict noise and air pollution in places where we didn't conduct measurements. The rasterized data can then be intersected with the station or highway buffer and noise and air pollution level in the communities within the buffer can be measured.

3.1.7 Average Commuting Time

Circular/polygon buffers (0.25/0.5/1 mile) centered around the rail stations/highway facility could be created to identify the adjacent communities. From the origin and destination data from National Household Travel Survey (NHTS), we can calculate the average commuting distance on the network. Using the average speed limit on the network, we can obtain the average commuting time of households within the rail station/highway buffer. In addition, ACS provides estimates of the number of households in different travel time to work category, ranging from less than 5 minutes to more than 90 minutes, at 5 minute intervals until 45 minutes, afterwards at 15 and 30 minute intervals. From this, we can calculate proportion of households in each travel time category within the buffer.

3.1.8 Proportion of Transit/Bike/Walk Commuters

Different sized buffers (0.25/0.5/1 mile) centered around the rail stations/highway facility could be created to identify the adjacent communities. From the ACS data, mode share for work at the census block level can be obtained and then the proportion of transit/walk/bike commuters within the buffer can be computed.

3.1.9 Land Use Development Type

This MOE can be evaluated for all three investment projects. For example, it can be evaluated for expansion of SunRail facility in the following way. Circular buffers (0.25/0.5/1 mile) centered around the rail stations could be created to identify the adjacent communities. Afterwards, the generalized land use data from FGDL repository could be intersected with the station buffer and then the changes in land use types could be evaluated before and after the station opening.

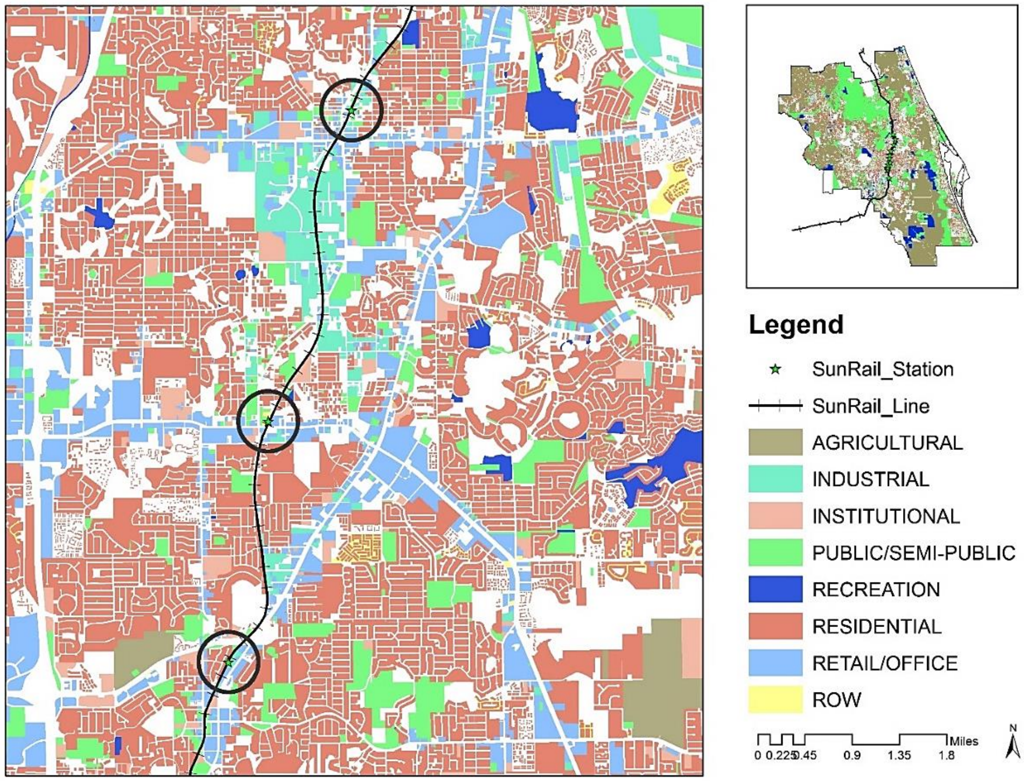


Figure 5: Land Use within Rail Buffer in 2010 (Before Station Opening)

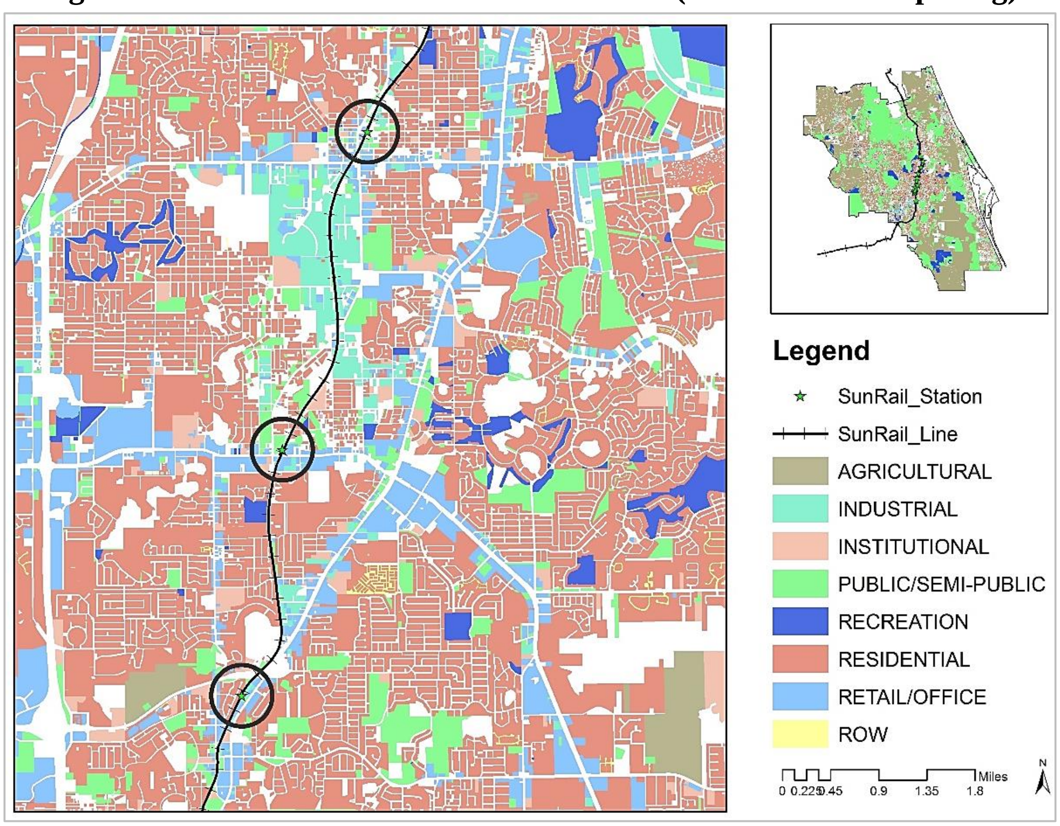


Figure 6: Land Use within Rail Buffer in 2015 (After Station Opening)

3.1.10 Land Use Mix Change

Buffers of different sizes (0.25/0.5/1 mile) centered around rail stations/highway extension could be created to identify the adjacent communities. After that the generalized land use data from FGDL repository could be intersected with the station buffer and then the land use mix could be calculated before and after the station opening using the following equation: Land use mix = -$\sum_k \frac{[p_k ln\, p_k]}{\ln(K)}$, where: p_k is the proportion of the developed land in the kth land use type.

3.1.11 Accessibility to Amenities

Accessibility to amenities (hospitals, schools/colleges, fire stations, restaurants, coffee shops, bars, grocery stores, book stores, shopping malls) is an important component of community desirability and attractiveness. This MOE can be evaluated for all three of the projects mentioned before. Buffers of different sizes (0.25/0.5/1 mile) around the transportation facility under consideration and layer of points of interests could be intersected to count the number of these points of interests in the communities within the buffer. In addition, network distance from the centroid of the parcels within the buffers to various amenities can be calculated. For this, layers of different points of interests are needed. Figure 7 shows restaurants and parks around the 1-4 expansion sites.

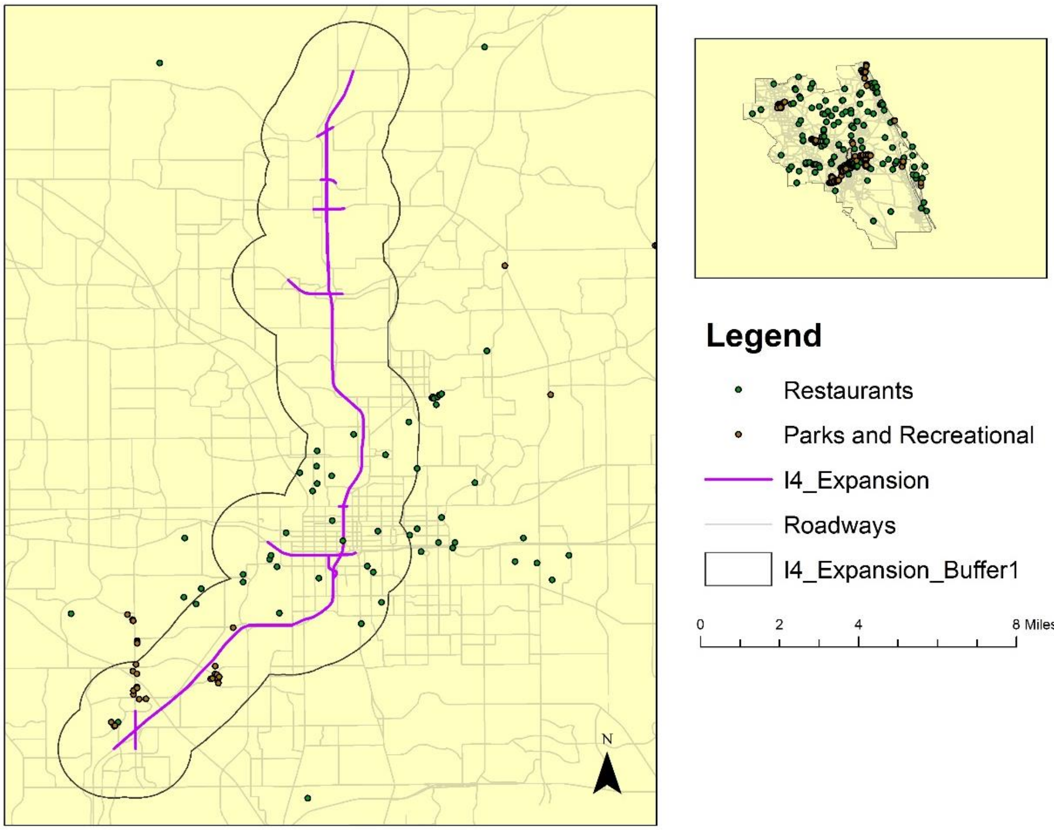


Figure 7: Restaurants and Parks around the I-4 Expansion

3.1.12 Jobs Proximity Index

The jobs proximity index quantifies the accessibility of a given residential neighborhood as a function of its distance to all job locations within a CBSA. This layer can be obtained from US Department of Housing and Development and can be intersected with the station/highway buffers (0.25/0.5/1-mile) to see how the index is varying across the communities within the buffer before and after the extension projects. Figure 8 shows jobs proximity around the 1-4 expansion sites.

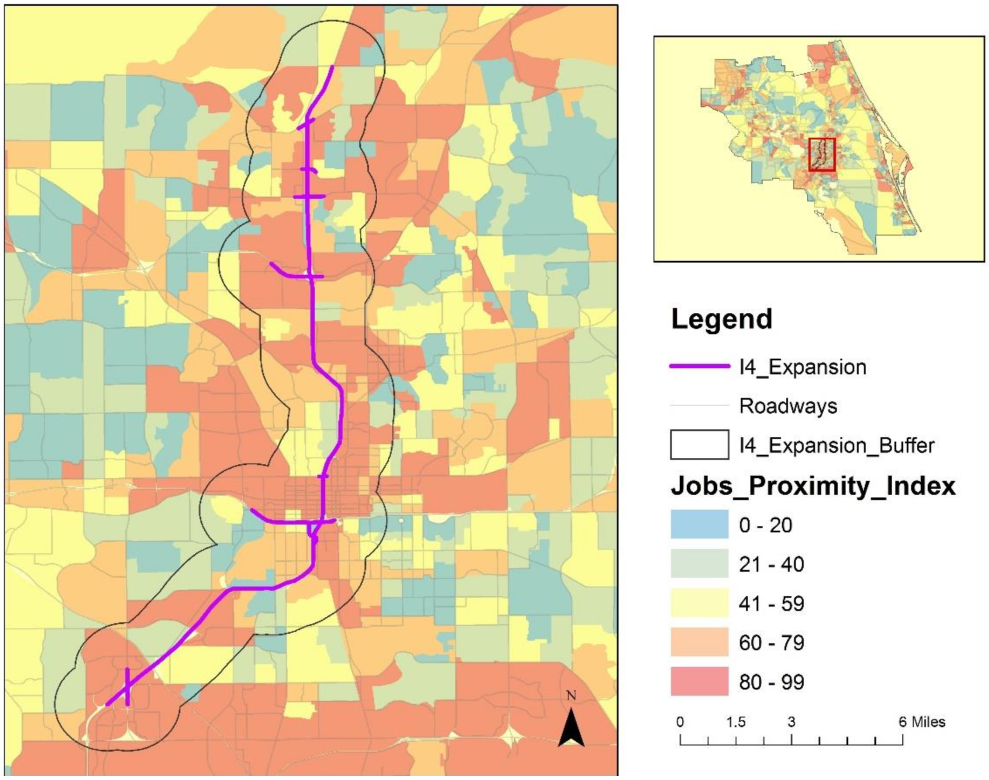


Figure 8: Job Proximity around Highway Buffer (1km)

3.1.13 Connectivity Index

It is calculated as the ratio of the street segments to intersections or the number of roadway links divided by the number of roadway nodes (cul-de-sacs included) or ratio of intersections to dead ends (including cul-de-sacs). The higher the values, the more is the connectivity. As connectivity increases, travel distances decrease and route options increase allowing more direct travel between destinations. This indicator can be calculated by loading the road network data in ArcGIS and intersecting the layer with the roadway polygon buffer (0.25/0.5/1 mile) and counting the number of links, intersections and dead-ends.

3.1.14 Area of Parks

Circular buffer ranging from 0.25-1 mile could be created around the bikeshare stations and intersected with the park layer in ArcGIS. Then the total park area within the bikeshare station buffer could be calculated to see the change in accessibility to these facilities with installation of new bikshare stations.

Table 9: Summary of MOEs and Data Sources

MOE	Relevant Project	Data Source
Property Price/Rent Variation	SunRail, I-4 Ultimate, Bikeshare/SunTrail	FDOT, FDOR
Pedestrian/Bike Crashes	SunRail, I-4 Ultimate	FDOT,S4A
Proportion of Severe crashes	I-4 Ultimate	FDOT,S4A
Bus Transit Ridership	SunRail, I-4 Ultimate	FDOT, LYNX
Crime Rate	SunRail	FDLE
Noise and Air Pollution Level	SunRail, I-4 Ultimate	EPA
Average Commuting Time	SunRail, I-4 Ultimate	NHTS, ACS
Proportion of Transit/Bike/Walk Commuters	SunRail, I-4 Ultimate	NHTS, ACS
Land Use Development Type	SunRail, I-4 Ultimate, Bikeshare/SunTrail	FGDL
Land Use Mix Change	SunRail, I-4 Ultimate, Bikeshare/SunTrail	FGDL
Accessibility to Amenities	SunRail, I-4 Ultimate, Bikeshare/SunTrail	FGDL
Jobs Proximity Index	SunRail, I-4 Ultimate	EPA
Connectivity Index	I-4 Ultimate	FDOT
Area of Park	Bikeshare/SunTrail	FGDL

CHAPTER 4

4.1 INTRODUCTION

Toward understanding public feedback on several established and ongoing transportation projects in the Central Florida region, we have extensively collected social media data. For the project, we have selected Twitter as a reliable data source as it is the most widely used social media platform in the USA with 67 million active users (Omnicore, 2017). Twitter is a micro blogging service used to share views, activities, and thoughts through a 140 character long message called ‘tweet’. Apart from the text portion of a tweet, there are a number of features which carry important clues to latent attributes of social media users. With twitter, one can extract spatial (geo-tagged) and temporal (time-stamped) information for a longer period of time and for large samples without accessing personal details or the content of the tweets (Frias-Martinez et al., 2012; Hasan and Ukkusuri, 2015).

4.2 DATA COLLECTION PROCESS FROM TWITTER

Among various social media platforms (like Facebook, Flickr, Instagram etc.) Twitter is a potential data source as it is collectable through simple web scraping and has a wide range of information within each post (tweets) (Hasan and Ukkusuri, 2015).

To collect data from Twitter, it requires a set of authentication keys providing an OAuth (Open Authorization) which is a standard for token-based authentication for accessing web data. Through a set of unique OAuth keys, we have used Twitter’s REST Application Program Interface (API) and Stream API to web scrap from twitter web pages. The REST API provides programmatic access to read and write Twitter data, i.e. create a new Tweet, read user profile and follower data etc. and Streaming API continuously delivers new responses to API queries over a long-lived HTTP connection receiving updates on the latest Tweets matching a search query, stay in sync with user profile updates etc. (Twitter Developer Documentation (a)). These developer keys are freely available within a certain query limits for specific types of search requests (Twitter Developer Documentation (b)). In brief, with valid OAuth keys one can search for tweets containing certain keywords and/or a group of keywords, tweets from certain user accounts, specific tweets within a selected geographical boundary box etc. For this project, a set of keyword and some specific Twitter accounts have been selected to collect data. The Appendix sections contain the python scripts used to collect the data.

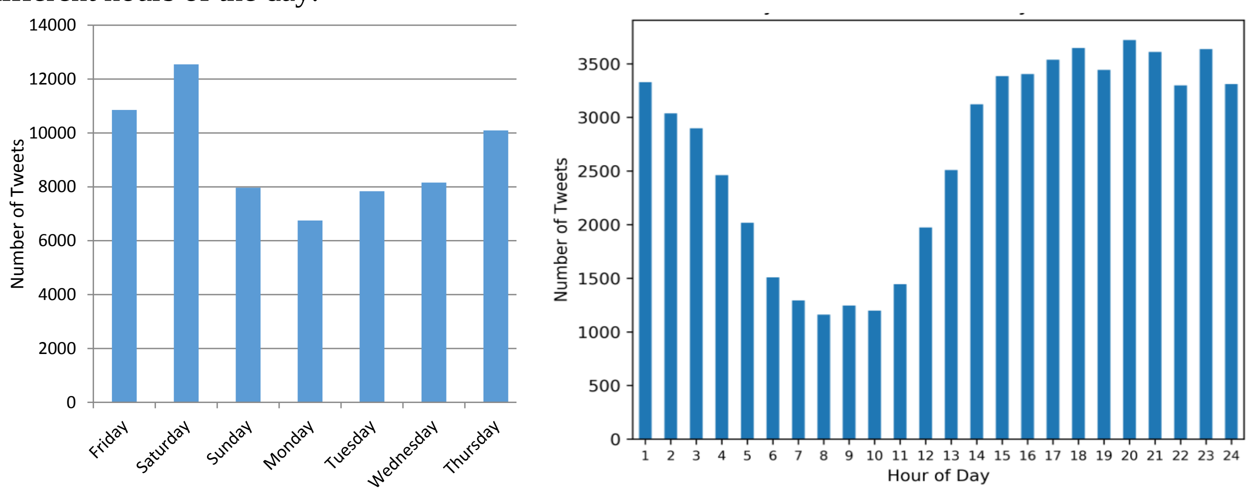
4.2.1 Tweet Search using Specific Keywords

The research team has selected some specific keywords, based on input from FDOT program manager that represent the key components of the transportation infrastructure in the Central Florida region. We mainly focused on several ongoing major transportation projects in the Central Florida region including second phase of SunRail commuter rail extension, I-4 expansion, pedestrian and bicycling facility installation, and bicycle sharing system (Juice) introduction. Within the limitations of twitter search API, data from the last 8 to 9 days can be collected for any specific keyword or a group of keywords. Keeping this condition in mind, data are being collected once in every 7 to 8 days starting from 24 February, 2017. Table 10 shows the collected number of tweets using different keyword and different group of keywords up to August 1, 2017.

Table 10: Tweets Collected using Specific Keyword Search

Sl. No.	Keywords	Total Unique Tweets	Geo-tagged Tweets
1	Florida Bus	2376	36
2	Florida Crime	13172	38
3	Florida Sidewalk	221	41
4	Florida Spring	27891	40
5	Florida Walking	11905	37
6	I4 Construction	1190	40
7	I4 Crash	3830	33
8	I4 Ultimate	144	36
9	Juicebike, juice bike	982	34
10	lynx bus, lynsbusorlando	578	47
11	Sunrail	2302	32
12	Suntrail	8	40
13	Suntran, Suntran Ocala	31	33
14	votran	147	31
	Total Tweets	**64793**	**518**

Figure 9 shows the frequency of the collected tweets in different days of the week and different hours of the day.



(a) **(b)**

Figure 9: Daily and Hourly Posted Tweets from Keyword Search. (a) Daily Number of posted Tweets and (b) Hourly Number of posted Tweets

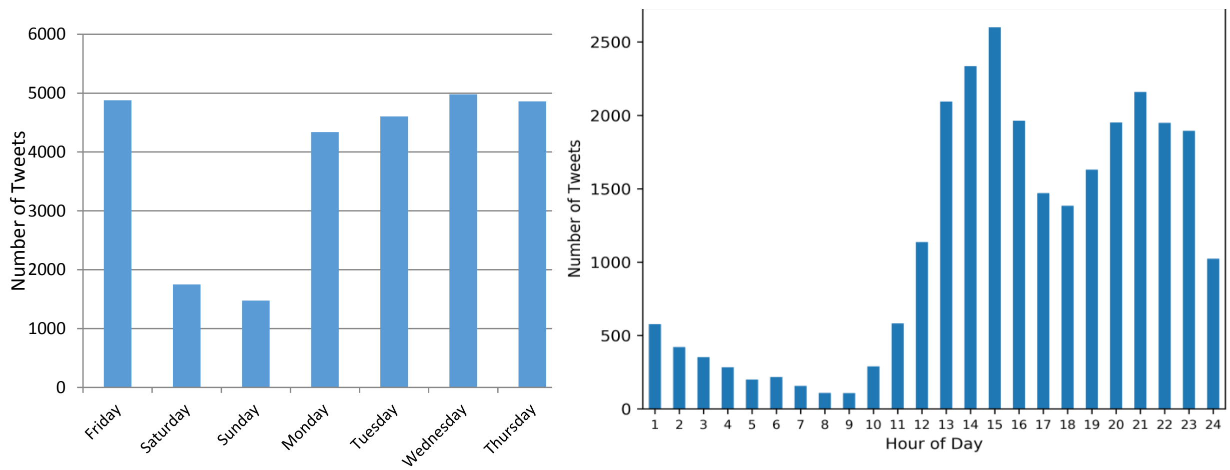
4.2.2 Tweet Search from Specific User Accounts

We have identified Twitter accounts which disseminate important information about the existing and on-going transportation infrastructures in the Central Florida region. In addition, we have collected data from 14 FDOT 511 service Twitter accounts that share incidents and real-time traffic information throughout the state. Each account provides traffic information for specific regions and/or facilities maintained by FDOT. Among these accounts, tweets have been collected from 13 accounts which use English language (Table 2) except the account named 'FL511_Estatal' which uses the Spanish language. For a particular user, Twitter search API restricts the maximum retrievable tweets up to the latest 3240 tweets at a time. Table 11 shows the tweets collected from the 26 user accounts until August 1, 2017.

Table 11: Tweets Collected from Specific User Accounts

User Name	Total Tweets	Created at	Earliest Tweet	Latest Tweet	Duration in Days	Daily Tweets
fl_511_i4	8716	1/12/2012 14:48	2/1/2017 21:00	8/1/2017 12:39	181	48.2
FL511_95Express	5947	1/24/2017 19:51	2/24/2017 21:02	8/1/2017 13:44	158	37.7
fl511_central	14881	10/6/2010 16:54	2/12/2017 0:25	8/1/2017 12:39	171	87.3
fl511_i10	5803	10/6/2010 17:30	1/30/2017 12:52	8/1/2017 11:45	183	31.7
fl511_i75	5772	10/6/2010 17:33	2/17/2017 17:31	8/1/2017 9:39	165	35.1
fl511_i95	6946	10/6/2010 17:37	2/24/2017 13:14	8/1/2017 11:23	158	44.0
fl511_northeast	7477	10/7/2010 12:38	3/11/2017 7:24	8/1/2017 13:34	143	52.2
fl511_panhandl	5438	1/12/2012 14:20	1/28/2017 11:56	8/1/2017 12:44	185	29.4
FL511_SOUTHEAST	13160	5/10/2017 1:42	4/13/2017 12:58	8/1/2017 13:44	110	119.6
fl511_southwest	4060	10/6/2010 17:15	1/20/2017 10:41	7/31/2017 16:00	192	21.1
fl511_state	15712	10/7/2010 12:57	4/29/2017 19:18	8/1/2017 13:44	94	167.6
fl511_tampabay	5731	10/6/2010 17:01	2/11/2017 17:56	8/1/2017 9:30	171	33.6
fl511_turnpike	4448	10/6/2010 17:23	2/4/2017 19:34	8/1/2017 13:05	178	25.0
FL511_Estatal	3215	3/7/2017 20:31	7/21/2017 9:31	8/1/2017 13:44	11	287.7
321Transit	996	8/25/2010 15:58	8/25/2010 16:04	7/31/2017 19:27	2532	0.4
965traffic	6434	4/7/2011 13:54	9/29/2016 13:38	8/1/2017 13:00	306	21.0
BikeWalkCFL	2926	8/29/2013 19:02	8/30/2013 17:58	7/31/2017 15:07	1431	2.0
I4Ultimate	3791	11/25/2014 17:19	1/17/2017 14:20	7/28/2017 22:00	192	19.7
juicebikes	281	3/23/2009 22:59	3/19/2011 21:53	7/31/2017 18:32	2326	0.1
lakexpress	100	8/13/2009 20:37	9/29/2010 18:45	4/28/2017 18:32	2403	0.0
lynxbusorlando	6504	6/4/2009 19:39	4/10/2013 13:46	8/1/2017 13:01	1574	4.1
RideSunRail	3287	5/7/2012 20:50	5/10/2014 15:59	7/31/2017 14:05	1178	2.8
SunRailRider	515	3/24/2011 13:54	4/4/2011 22:41	8/29/2014 11:12	1243	0.4
SunTranTDP2017	29	11/9/2016 15:20	11/9/2016 17:44	6/13/2017 15:08	216	0.1
WazeTrafficOrl	3240	11/3/2014 19:32	9/8/2016 6:07	4/3/2017 16:07	207	15.6
Total Tweets	**135409**	-	-	**Average**	**628**	**43**

The accounts have average activity of 43 tweets per day with ‘FL511_Estatal’ being the most active account posting more than 287 tweets per day. Figure 10 shows the daily and hourly activity of the user accounts.



(a) **(b)**

Figure 10: Daily and Hourly Posted Tweets from User Accounts Search: (a) Daily Number of posted Tweets and (b) Hourly Number of posted Tweets

The accounts have been found to post more tweets from Friday to Friday of the weeks. Also, in hourly basis the most of tweets were found to be posted between the window of 1 PM to 4 PM and 7 PM to 11 PM.

4.3 PRELIMINARY ANALYSIS FOR COMMUNITY BUILDING INDICATORS

We have conducted a preliminary analysis over the collected Twitter datasets to find if there are enough indicators related to community building available in the data. Some of these indicators are readily available from the data. For instance, the number of followers of a Twitter account is a measure of influence of that account and reflects its connectivity with the community. The more followers an account has, the wider is the reach of its posted information/tweets indicating the importance of the project to the community. Another indicator is the number of times a message has been reposted (retweeted) by others; it reflects the importance of specific information to the community. Table 12 shows the number of followers, number of friends and the total number of tweets that have been retweeted at least once.

Table 12: Follower, Friends and Re-Tweet counts of the Twitter Accounts

User Name	Follower Count	Friends Count	Number of Tweets Retweeted
321Transit	527	175	230
965traffic	1594	277	358
BikeWalkCFL	1427	922	1568

fl_511_i4	3144	625	123
FL511_95Express	128	52	61
fl511_central	3053	656	574
fl511_i10	1289	247	215
fl511_i75	530	262	251
fl511_i95	6928	1032	290
fl511_northeast	1648	198	439
fl511_panhandl	1778	207	340
FL511_SOUTHEAST	6080	371	537
fl511_southwest	2366	72	127
fl511_state	2174	173	115
fl511_tampabay	3953	107	229
fl511_turnpike	10262	384	438
flcrimewatch	11	2	0
I4Ultimate	1696	130	261
juicebikes	1131	239	123
lakexpress	118	110	12
lynxbusorlando	4742	273	3698
RideSunRail	13851	592	2283
SunRailRider	847	27	249
SunTranTDP2017	15	99	4
WazeTrafficOrl	134	0	15

From Table 12, it is found that 'RideSunRail' (the account responsible for giving information about Sunrail project) has the highest number of followers (account created at 5/7/2012) and 'SunTranTDP2017' has the lowest number of followers (account created at 11/9/2016). These are the first order followers or the number of users those are directly following the accounts under consideration. It is possible to build a network of each account by collecting the followers of the followers (second order connections of the accounts under consideration). This will help us to find the broader community connected with these accounts and the influence of these accounts. We have also analyzed the trend in the number of followers for a selected number of accounts. Figure 11 shows the trend line of follower gain (or loss) of these four accounts in different time. From figure 11 it is seen that there are little activities in terms of total number of followers in the user accounts. 'BikeWalkCFL' has constant number of followers from the beginning till the latest tweet collected. 'RideSunRail' experienced a small growth in the number of followers during August, 2013 and July, 2017. An increase or decrease in the number of followers is an indication of how the account is attracting or losing users over time.

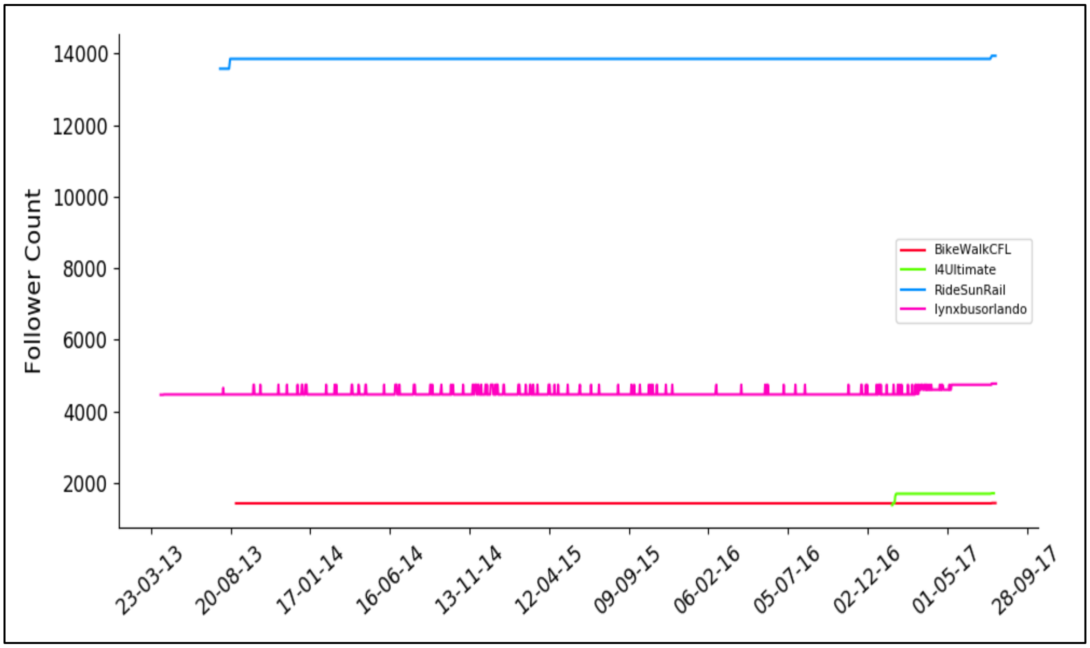


Figure 11: Trend in Total Number of Followers of Twitter Accounts

Tweets found from the keyword search are also good indicators of public opinions about the transportation infrastructures and services provided. A small sample of tweets posted by different individual users is given below:

- patboyle60: "RT @PierreTristam: Loving our very first ride on #sunrail. We need more of these French-like car-busting commuter luxuries"
- brentschmidt: "RT @derekburgan: A sidewalk where you don't burn your feet in Florida. Between this and Pandora's bioluminescence we now have E-Ticket walk\xe2\x80\xa6"
- CrankySlytherin: "Not surprised by Sunrail's numbers. It's still new to the area. Give it time it will grow. Also needs more hours and more destinations."
- Swomack: "@SMurphyCongress What are you doing to help us get additional funding for a better Sunrail system in Central FL? This should be a priority!"
- Indy_Austin: "Without SunRail into the Orlando, passengers will need to transfer at least twice to reach the terminal."
- AlexLaporte12: "RT @M_O_Ski: There is nothing better than walking outside in the morning to the smell of florida. You're a real one if you know what i'm t\xe2\x80\xa6"
- wellman_shana: "RT @JoshuaWMGant: I hear a lot of folks complaining about the construction on i4, but it is coming along nicely. I can't wait to see the fi\xe2\x80\xa6"

A detailed text analysis of the collected tweets will help to reveal topics of issues and public sentiments (i.e. positive or negative) towards a specific transportation service and infrastructure projects. Figure 12 shows 50 most frequent words found in the keyword search process.

(a) **(b)**

Figure 12: Most Frequent Words found from Keyword Search: (a) ‘I4’ and ‘Construction’, (b) ‘florida’ and ‘sidewalk’

The top most words in the tweets found from the combined keyword search of ‘I4’ and ‘construction;’ includes ‘construction’, ‘planned’, ‘blocked’, ‘orange’ etc. Figure 13 shows 50 most frequent words found from user account search.

(a) **(b)**

Figure 13: Most Frequent Words found from User Account Search: (a) ‘I4Ultimate’ and (b) ‘SunRailRider’

The top most words in the tweets of the user account ‘I4Ultimate’ includes ‘blocked’, ‘orange’, ‘I4’ etc. and from user account ‘SunRailRider’ includes ‘SunRail’, ‘train’, ‘minutes’ etc. To understand the whole context, a detail topic analyses will need to be done in the next phase of the research.

CHAPTER 5

5.1 SUMMARY AND CONCLUSIONS

The draft final report summarizes the findings obtained from the two major research tasks of the project. In the first task, we reviewed existing literature from a broad variety of fields to understand how community impact of different transportation facilities or infrastructure investments is evaluated. Our review suggests that addition or improvement or enhancement of transportation facilities do impact contiguous communities and these impacts can be measured through several indicators such as land/housing price change, crime rate change, change in land use development type, transit usage shares. Informed from the review, we proposed several measures of effectiveness for assessing community development impacts of several ongoing transportation projects in the Central Florida region. Potential publicly available data sources that can be useful in developing the measures of effectiveness are also identified and documented. Downloaded data from publicly available data sources are submitted with the report.

In the second task, we collected and analyzed social media data collected from Twitter. Moreover, we have described the procedure of collecting data from Twitter using query scripts. The developed social media query scripts and collected data are accompanied with this report. We have found that social media data have some readily available indicators for measuring community building impacts of transportation projects. Other indicators can also be developed by running sophisticated data mining techniques.

As part of this research project, we will continue collecting data for further analysis. In the next phase of this project, we will filter the data to gather most relevant tweets. We will analyze the influence of transportation infrastructure into people’s live using their activities in social media. The reach and effectiveness of a number of FDOT Twitter accounts in transmitting valuable information will also be measured. In addition, topic and sentiment analysis will be conducted over the collected data. These analysis techniques applied over the filtered data will enable us to gather valuable insights on how transportation investments help to build communities.

REFERENCES

Andersson, D. E., Shyr, O. F., & Fu, J. (2010). Does high-speed rail accessibility influence residential property prices? Hedonic estimates from southern Taiwan. *Journal of Transport Geography, 18*(1), 166-174.

Andersson, H., Jonsson, L., & Ögren, M. (2010). Property prices and exposure to multiple noise sources: Hedonic regression with road and railway noise. *Environmental and Resource Economics, 45*(1), 73-89.

Armstrong, R. J., & Rodriguez, D. A. (2006). An evaluation of the accessibility benefits of commuter rail in eastern Massachusetts using spatial hedonic price functions. *Transportation, 33*(1), 21-43.

Atack, J., Bateman, F., Haines, M., & Margo, R. A. (2010). Did railroads induce or follow economic growth? *Social Science History, 34*(02), 171-197.

Bae, C.-H. C., Jun, M.-J., & Park, H. (2003). The impact of Seoul's subway line 5 on residential property values. *Transport Policy, 10*(2), 85-94.

Banister, D., & Thurstain-Goodwin, M. (2011). Quantification of the non-transport benefits resulting from rail investment. *Journal of Transport Geography, 19*(2), 212-223.

Bhattacharjee, S., & Goetz, A. R. (2012). Impact of light rail on traffic congestion in Denver. *Journal of Transport Geography, 22*, 262-270.

Bina, M., Warburg, V., & Kockelman, K. (2006). Location choice vis-à-vis transportation: Apartment dwellers. *Transportation Research Record: Journal of the Transportation Research Board, 1977*, 93-102.

Blanco, J. C., & Flindell, I. (2011). Property prices in urban areas affected by road traffic noise. *Applied Acoustics, 72*(4), 133-141.

Boarnet, M. G., & Chalermpong, S. (2001). New highways, house prices, and urban development: A case study of toll roads in orange county, ca. *Housing Policy Debate, 12*(3), 575-605.

Bowes, D. R., & Ihlanfeldt, K. R. (2001). Identifying the impacts of rail transit stations on residential property values. *Journal of urban Economics, 50*(1), 1-25.

Brandt, S., & Maennig, W. (2011). Road noise exposure and residential property prices: Evidence from Hamburg. *Transportation Research Part D: Transport and Environment, 16*(1), 23-30.

Buehler, R., & Hamre, A. (2015). Business and bikeshare user perceptions of the economic benefits of capital bikeshare. *Transportation Research Record: Journal of the Transportation Research Board, 2520*, 100-111.

Cao, X., & Hough, J. A. (2012). Hedonic value of transit accessibility: An empirical analysis in a small urban area. *Journal of the Transportation Research Forum, 47*(3), 171-183.

Cao, X. J., & Schoner, J. (2014). The influence of light rail transit on transit use: An exploration of station area residents along the Hiawatha line in Minneapolis. *Transportation Research Part A: Policy and Practice, 59*, 134-143.

Celik, H., & Yankaya, U. (2006). The impact of rail transit investment on the residential property values in developing countries: The case of Izmir subway, turkey. *Property Management, 24*(4), 369-382.

Cervero, R., & Duncan, M. (2002). Transit's value-added effects: Light and commuter rail services and commercial land values. *Transportation Research Record: Journal of the Transportation Research Board*, *1805*, 8-15.

Clower, T. L., & Weinstein, B. L. (2002). The impact of Dallas (Texas) area rapid transit light rail stations on taxable property valuations. *Australasian Journal of Regional Studies, The, 8*(3), 389.

Combs, T. S. (2017). Examining changes in travel patterns among lower wealth households after BRT investment in Bogotá, Colombia. *Journal of Transport Geography*, *60*, 11-20.

Combs, T. S., & Rodríguez, D. A. (2014). Joint impacts of bus rapid transit and urban form on vehicle ownership: New evidence from a quasi-longitudinal analysis in Bogotá, Colombia. *Transportation Research Part A: Policy and Practice*, *69*, 272-285.

Debrezion, G., Pels, E., & Rietveld, P. (2011). The impact of rail transport on real estate prices: An empirical analysis of the Dutch housing market. *Urban Studies*, *48*(5), 997-1015.

Duncan, M. (2011). The impact of transit-oriented development on housing prices in San Diego, CA. *Urban Studies*, *48*(1), 101-127.

El-Geneidy, A., van Lierop, D., & Wasfi, R. (2016). Do people value bicycle sharing? A multilevel longitudinal analysis capturing the impact of bicycle sharing on residential sales in Montréal, Canada. *Transport Policy*, *51*, 174-181.

Forkenbrock, D., & Weisbrod, G. (2001). NCHRP Report 456: Guidebook for Assessing the Social and Economic Effects of Transportation Projects. TRB, National Research Council, Washington, D. C., 2001.

Frias-Martinez, V., Soto, V., Hohwald, H., & Frias-Martinez, E.. Characterizing urban landscapes using geolocated tweets. In Privacy, Security, Risk and Trust (PASSAT), International Conference on and 2012 International Conference on Social Computing (SocialCom), 2012, pp. 239-248, IEEE.

Gibbons, S., & Machin, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics*, *57*(1), 148-169.

Gingerich, K., Maoh, H., & Anderson, W. (2013). Location and transportation effects on nonresidential real estate price regressions in Windsor, Ontario, Canada. *Transportation Research Record: Journal of the Transportation Research Board*, *2397*, 99-107.

Hamersma, M., Heinen, E., Tillema, T., & Arts, J. (2017). New highway development in the Netherlands: A residents' perspective. *Transportation Research Part D: Transport and Environment*, *51*, 326-339.

Hasan, S., and Ukkusuri, S.V. (2015). Location contexts of user check-ins to model urban geo life-style patterns, *PloS one*, *10* (5), e0124819.

Hess, D. B., & Almeida, T. M. (2007). Impact of proximity to light rail rapid transit on station-area property values in buffalo, New York. *Urban Studies*, *44*(5-6), 1041-1068.

Huang, W.-H., & Chao, M.-C. (2014). The impacts of the mass rapid transit system on household car ownership in Taipei. *Journal of Sustainable Development of Energy, Water and Environment Systems*, *2*(2), 191-207.

Iacono, M., & Levinson, D. (2011). Location, regional accessibility, and price effects: Evidence from home sales in Hennepin county, Minnesota. *Transportation Research Record: Journal of the Transportation Research Board, 2245*, 87-94.

Kang, C. D., & Cervero, R. (2009). From elevated freeway to urban greenway: Land value impacts of the CGC project in Seoul, Korea. *Urban Studies, 46*(13), 2771-2794.

Kim, K. S., Park, S. J., & Kweon, Y.-J. (2007). Highway traffic noise effects on land price in an urban area. *Transportation Research Part D: Transport and Environment, 12*(4), 275-280.

Knaap, G. J., Ding, C., & Hopkins, L. D. (2001). Do plans matter? The effects of light rail plans on land values in station areas. *Journal of Planning Education and Research, 21*(1), 32-39.

Ko, K., & Cao, X. J. (2013). The impact of Hiawatha light rail on commercial and industrial property values in Minneapolis. *Journal of Public Transportation*, *16*(1), 3.

Koster, H. R., van Ommeren, J. N., & Rietveld, P. (2010). *Estimating the benefits of improved rail access; geographical range and anticipation effects*. Retrieved from http://EconPapers.repec.org/RePEc:tin:wpaper:20100094

Krizek, K. J. (2006). Two approaches to valuing some of bicycle facilities' presumed benefits. *Journal of the American Planning Association, 72*(3), 309-320.

Lee, J.-H., & Chang, J. (2006). Effects of high-speed rail service on shares of intercity passenger ridership in South Korea. *Transportation Research Record: Journal of the Transportation Research Board, 1943*, 31-42.

Levkovich, O., Rouwendal, J., & Marwijk, R. (2016). The effects of highway development on housing prices. *Transportation, 43*(2), 379-405.

Li, T. (2016). The value of access to rail transit in a congested city: Evidence from housing prices in Beijing. Available at SSRN: https://ssrn.com/abstract=2831478 or http://dx.doi.org/10.2139/ssrn.2831478

Martínez, L., & Viegas, J. (2009). Effects of transportation accessibility on residential property values: Hedonic price model in the Lisbon, Portugal, metropolitan area. *Transportation Research Record: Journal of the Transportation Research Board, 2115*, 127-137.

Mayor, K., Lyons, S., Duffy, D., & Tol, R. S. (2012). A hedonic analysis of the value of rail transport in the greater Dublin area. *Journal of Transport Economics and Policy (JTEP), 46*(2), 239-261.

McMillen, D. P., & McDonald, J. (2004). Reaction of house prices to a new rapid transit line: Chicago's midway line, 1983–1999. *Real Estate Economics, 32*(3), 463-486.

Meijers, E., Hoekstra, J., & Spaans, M. (2013). Fixed link, fixed effects? Housing market outcomes of new infrastructure development in the Dutch delta area. *Geografisk Tidsskrift-Danish, 113*(1), 11-24.

Merom, D., Bauman, A., Vita, P., & Close, G. (2003). An environmental intervention to promote walking and cycling—the impact of a newly constructed rail trail in western Sydney. *Preventive Medicine, 36*(2), 235-242.

Munoz-Raskin, R. (2010). Walking accessibility to bus rapid transit: Does it affect property values? The case of Bogotá, Colombia. *Transport Policy, 17*(2), 72-84.

Nijland, H. A., Hartemink, S., van Kamp, I., & van Wee, B. (2007). The influence of sensitivity for road traffic noise on residential location: Does it trigger a process of spatial selection? *The Journal of the Acoustical Society of America, 122*(3), 1595-1601.

Omnicore. Twitter by the Numbers: Stats, Demographics & Fun Facts. https://www.omnicoreagency.com/twitter-statistics/. Accessed June 4, 2017.

Pivo, G., & Fisher, J. D. (2011). The walkability premium in commercial real estate investments. *Real Estate Economics, 39*(2), 185-219.

Rauterkus, S., Thrall, G., & Hangen, E. (2010). Location efficiency and mortgage default. *Journal of Sustainable Real Estate, 2*(1), 117-141.

Rauterkus, S. Y., & Miller, N. (2011). Residential land values and walkability. *Journal of Sustainable Real Estate, 3*(1), 23-43.

Reibel, M., Chernobai, E., & Carney, M. (2008). *House price change and highway construction: Spatial and temporal heterogeneity.* Paper presented at the American Real Estate Society conference.

Rogers, S. H., Halstead, J. M., Gardner, K. H., & Carlson, C. H. (2011). Examining walkability and social capital as indicators of quality of life at the municipal and neighborhood scales. *Applied Research in Quality of Life, 6*(2), 201-213.

Senior, M. L. (2009). Impacts on travel behavior of greater Manchester's light rail investment (Metrolink phase 1): Evidence from household surveys and census data. *Journal of Transport Geography, 17*(3), 187-197.

Shen, Q., Chen, P., & Pan, H. (2016). Factors affecting car ownership and mode choice in rail transit-supported suburbs of a large Chinese city. *Transportation Research Part A: Policy and Practice, 94*, 31-44.

Shin, K., Washington, S., & Choi, K. (2007). Effects of transportation accessibility on residential property values: Application of spatial hedonic price model in Seoul, south Korea, metropolitan area. *Transportation Research Record: Journal of the Transportation Research Board, 1994*, 66-73.

Smersh, G. T., & Smith, M. T. (2000). Accessibility changes and urban house price appreciation: A constrained optimization approach to determining distance effects. *Journal of Housing Economics, 9*(3), 187-196.

Tay, R., Azad, A., Wirasinghe, S., & Hansen, S. (2013). Analysis of the influence of urban rail stations on neighborhood crime. *International Journal of Urban Sciences, 17*(2), 281-289.

Theebe, M. A. (2004). Planes, trains, and automobiles: The impact of traffic noise on house prices. *The Journal of Real Estate Finance and Economics, 28*(2-3), 209-234.

Twitter Developer Documentation (a). https://dev.twitter.com/docs. Last Accessed: June 13, 2017.

Twitter Developer Documentation (b). https://dev.twitter.com/rest/public/rate-limits. Last Accessed: June 13, 2017.

Appendix 1

Python Script for Keyword Search

```
#Compatible with Python 3 versions
#Importing Necessary Modules of Python
import tweepy
import csv
from tweepy import Stream
from tweepy import OAuthHandler
from tweepy.streaming import StreamListener

#Inserting Authentication Keys provided by Twitter Developer
consumer_key = 'A'
consumer_secret = 'B'
access_token = 'C'
access_token_secret = 'D'

#Authenticating in the Python Script
auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)
api = tweepy.API(auth)

#Open the csv file in which we have stored the keywords in a column of the first row
for line in open('0. List_KW.csv','r', encoding='utf-8'):
    keyword = line.strip()

   #Open csv file in which Tweets for each keywords will be saved
    csvFile = open('%s_08_August_tweets.csv' % keyword, 'w', encoding='utf-8')
    csvWriter = csv.writer(csvFile)
    try:
        for tweet in tweepy.Cursor(api.search,q=[keyword],count=100,\
                    lang="en", \
                    since="2017-07-31",
                    until="2017-08-08").items(): #can be used for upto 10 Day
            print ("Running...")
            csvWriter.writerow([tweet.created_at,tweet.text.encode('utf-8'),tweet.user.screen_name,
            tweet.favorite_count, tweet.retweet_count, tweet.id, tweet.coordinates,
            tweet.in_reply_to_screen_name, tweet.in_reply_to_status_id,
            tweet.in_reply_to_status_id_str, tweet.in_reply_to_user_id,tweet.is_quote_status,
            tweet.retweeted, tweet.source, tweet.user,
            tweet.truncated,tweet.user.contributors_enabled,tweet.user.created_at,tweet.user.default_p
```

```
            rofile, tweet.user.time_zone, tweet.user.profile_text_color, tweet.entities,
            tweet.user.description.encode('utf-8'), tweet.user.id])
        print(keyword,' completed')
        csvFile.close()
    except:
        pass
        print(keyword, ' failed')
```

Appendix 2

Python Script for User Account Search

```
#Import necessary modules
import tweepy
import csv, pdb
import time

# Twitter API credentials
twitter_app_auth = {
   'consumer_key': "A",
   'consumer_secret': "B",
   'access_token': "C",
   'access_token_secret': "D",
}


def get_all_tweets(screen_name):
   # Twitter only allows access to a users most recent 3240 tweets with this method

   # authorize twitter, initialize tweepy
   auth =
tweepy.OAuthHandler(twitter_app_auth['consumer_key'],twitter_app_auth['consumer_secret'])

auth.set_access_token(twitter_app_auth['access_token'],twitter_app_auth['access_token_secret'])
   api = tweepy.API(auth)

   # initialize a list to hold all the tweepy Tweets
   alltweets = []

   # make initial request for most recent tweets (200 is the maximum allowed count)
   new_tweets = api.user_timeline(screen_name=screen_name, count=200)

   # save most recent tweets
   alltweets.extend(new_tweets)

   # save the id of the oldest tweet less one
   oldest = alltweets[-1].id - 1

   # keep grabbing tweets until there are no tweets left to grab
   while len(new_tweets) > 0:
```

```
        # print "getting tweets before %s" % (oldest)

        # all subsiquent requests use the max_id param to prevent duplicates
        new_tweets = api.user_timeline(screen_name=screen_name, count=200, max_id=oldest)

        # save most recent tweets
        alltweets.extend(new_tweets)

        # update the id of the oldest tweet less one
        oldest = alltweets[-1].id - 1

    # print "...%s tweets downloaded so far" % (len(alltweets))



    # transform the tweepy tweets into a 2D array that will populate the csv
    outtweets = [[tweet.id_str, tweet.created_at, tweet.text.encode('utf-8'),
tweet.user.screen_name,tweet.user.favourites_count,tweet.retweet_count,
              tweet.coordinates,tweet.in_reply_to_screen_name,tweet.in_reply_to_status_id,
tweet.in_reply_to_status_id_str,tweet.in_reply_to_user_id,
              tweet.is_quote_status,tweet.retweeted, tweet.source,tweet.user,tweet.truncated,
tweet.user.contributors_enabled,tweet.user.created_at,
              tweet.user.default_profile,tweet.user.time_zone,tweet.user.profile_text_color,
tweet.entities,
              tweet.user.description.encode('utf-8')] for tweet in alltweets]



    # write the csv
    with open('%s_August_08_tweets.csv' % screen_name, 'w', encoding='utf-8') as f:
        writer = csv.writer(f)
        writer.writerow(["id", "created_at", "text", "user_Screen_name",
"User_favourites_Count","RT_Count","Tweet_Coordinates",

"In_Reply_to_ScreenName","in_reply_to_status_id","in_reply_to_status_id_str","in_reply_to_user_
id","is_quote_status",
                   "retweeted","tweet_source", "tweet.user", "tweet.truncated",
"user.contributors_enabled","user.created_at",
                   "user.default_profile",
"user.time_zone","user.profile_text_color","tweet.entities","user.description"])
        writer.writerows(outtweets)
    pass
```

```
if __name__ == '__main__':
    # pass in the username of the account you want to download
    start=time.time()
    #out2 = open('Users_Tweets_found.csv', 'w')
    success=0
    failure=0
    failure_user=[]
    count=0

    for line in open('0. List_User.csv','r', encoding='utf-8'):
        #Opened csv file only contains the user names in first column
        account_name = line.strip()
        time_elasped=time.time()-start
        if count >60 and time_elasped<900:
            time.sleep(900-time_elasped)
            start=time.time()
            count=0
        try:
            count+=1
            get_all_tweets(account_name)
            success+=1
            print(account_name,' completed ','total success=', success )
        except:
            failure_user.append(account_name)
            failure+=1
            count+=1
            print(account_name, ' failed', 'total failure=', failure)
```