To appear in the *Journal of Location Based Services*
Vol. 00, No. 00, Month 20XX, 1–13

Special Issue

Spatio-temporal Analysis of Meta-data Semantics of Market Shares Over Large Public Geosocial Media Data

Abdulaziz Almaslukh[a,b], Amr Magdy[a,b], Sergio J. Rey[a,b,c*]
[a] *Center for Geospatial Sciences*;
[b] *Department of Computer Science and Engineering*;
[c] *School of Public Policy*;
University of California
Riverside, CA

(*Received January 2018*)

Monitoring market share changes over space and time is an essential and continuous task for commercial companies and their third-party local agents to adjust their sale campaigns and marketing efforts for profit maximization. This paper uses social media data as a cheap and up-to-date source to reveal the implicit semantics that are embedded in the meta-data of public geosocial datasets. We use Twitter data as a prime example of rich geosocial data. This data is associated with several meta-data attributes. Using this meta-data, we perform a geospatial analysis for the source platform from which a tweet is posted, e.g., from Apple or Android device. Our analysis studies all counties in US connected states over two years 2016-2017. We show that market structure at the national level masks substantial variation at the county scale. Moreover, we find strong spatial autocorrelation in platform distribution and market share in the US. In addition, we show interesting changes over the two years that motivates further analysis at different spatial and temporal levels. Our results are supported with visual maps of location quotients and market dominance, in addition to formal test results of spatial autocorrelation, and spatial Markov analysis.

Keywords: geospatial analysis, Twitter data, meta-data semantics, PySAL

1. Introduction

Geosocial media data, e.g., tweets and Facebook posts, has entered an unprecedented flourishing era with the widespread use of mobile users and devices. Everyday, 328+ million active Twitter users generate 500+ million tweets (Twitter Stats 2017), while 1.45+ billion Facebook users post 3.2+ billion comments and likes (Facebook Statistics 2018). The vast majority of such data comes from mobile users, specifically, 80+% of Twitter users and 85+% of Facebook users are mobile. The mobility of this data is combined with rich *user-generated content* including keywords/hashtags, news items, and social interactions, in addition to rich *meta-data* including exact/estimated location, timestamp, language, user information, and platform information. A plethora of applications have exploited both user-generated content and meta-data information. So, time, location, and user information are combined with user opinions and news to provide a holistic

This work was supported by the U.S. National Science Foundation under Grant SES-1733705.
*Corresponding Author. Email: sergio.rey@ucr.edu

functionality. Examples include news extraction (Boston Explosions 2013; Sankaranarayanan et al. 2009), rescue services (China Floods 2012; Hurricane Irma 2017; Hurricane Harvey 2017), event analysis (Abdelhaq et al. 2013; Ukraine Unrest 2014), scientific research (Twitter Political Sciences 2017; Twitter Sociology 2017), and geo-targeted advertising (Twitter GeoAds 2012). However, semantics extraction in the literature of geosocial media (Bermingham and Smeaton 2010; Meij et al. 2012; Xu et al. 2016) has mainly focused on mining user-generated content, e.g., tweet text, and link them to real-world entities, e.g., persons or places, or concepts like Wikipedia articles. On the contrary, the semantics embedded in meta-data information are still underutilized while it can be used in several applications and use cases on geosocial media data. Market share analysis of software platforms is one of such use cases that can exploit meta-data information in geosocial media data to provide cheap and up-to-date analysis for commercial companies and their local agents.

In this paper, we apply methods of exploratory spatio-temporal data analysis to the meta-data information of geosocial media to reveal the spatial structure of platform market share in USA. Specifically, we use a large dataset of 1 billions tweet messages available from public Twitter APIs that span two years (2016-2017) and include geo-locations within USA. Each tweet message is associated with *source* meta-data attribute that identifies the platform from which this tweet is posted, e.g., Apple, Android, Windows, Blackberry, and so on. We use a combination of time, location, and source attributes to analyze the market share of each platform within different US counties over the past two years. Our analysis draws upon different spatial analysis measures and time slices analyzing the spatio-temporal dynamics of mobile users usage of different platforms.

Our analysis shows that market share of different platforms is spatially autocorrelated with interesting clusters over different US regions. This is shown visually on maps of US counties and verified through formal tests that show spatial dependence. We also show market dominance of different platforms in US counties in both years through visual maps and formal tests of global and local spatial autocorrelation. The results show significant dominance for Apple devices in many of the counties and states with changes over time. We discuss in details the changes over different spatial regions and time slices in the following sections.

The rest of this paper is organized as follows. Section 2 presents the related work. Section 3 introduces the source data and our processing procedures. Section 4 presents our spatial analytical methodology and its application to the indexed data. We close the paper with a summary of the key findings and a discussion of future areas of research.

2. Related Work

Our work is twofold: (a) analyzing platform market shares, and (b) analyzing social media in a spatial context. This is related to three main areas namely, *platform market share analysis*, *spatio-temporal analysis on social media*, and *semantics in social media*, each outlined below.

Platform market share analysis. Several studies have been conducted to analyze mobile platform market share, e.g., Atlas (2017); Statista (2018); Movoto (2015); Kantar (2017); Statista B (2017); Jumptap (2011), due to its importance in continuously monitoring the market status and study opportunities for local growth or significance of shrinking resources on a localized scope for profit maximization. These studies use different data sources, e.g., user interviews and surveys Kantar (2017); Statista B (2017), cell phone data Statista (2018), and social media data Movoto (2015), and conducted at different spatial resolutions, e.g., national level Atlas (2017); Statista (2018); Statista B (2017); Kantar (2017) or state level Movoto (2015). However, the results of these studies are not consistent and differ based on different factors, including the temporal variability of market shares from time to another Statista B (2017); Jumptap (2011).

Unlike all existing work, our work is the first to provide such analysis on the county level, which is the finest granularity compared to all existing studies. This is enabled by the granularity of user location updates on social media. In addition, our analysis enables month-to-month updated results due to using a cheap data source that is publicly available and continuously updated by a sheer amount of users. This allows high adaptivity with market shares temporal variability.

Spatial analysis on social media. Spatial information in social media data are used in different applications to link user activities to their natural spatial extent. This included events detection and analysis (Abdelhaq et al. 2013; Sakaki et al. 2010) where events naturally have space and time extents, analyzing population demographics in different countries (Magdy et al. 2014), discovering localized news stories (Magdy et al. 2016), visualizing social media messages (Marcus et al. 2011; Weber and Garimella 2014), and targeting users by location in geo-ads (Twitter GeoAds 2012). However, none of existing work has exploited the information about the underlying device or platform to analyze market-related data using social media messages. Such data could be a super cheap and accurate source for market analysis compared to the expensive user studies or customized crowd-sourcing. It also provides a more frequently updated and finer granular source of data compared to the other means of performing such analysis.

Semantics in social media. Semantic analysis on social media has been active among different analysis tasks on popular social media data. This literature includes different types of semantic analysis. The first type is adapting the traditional definition of semantic analysis on text documents to social media data with its short text and new characteristics (Bermingham and Smeaton 2010; Meij et al. 2012). This work links each social media message to either real-world entities, e.g., persons or places, or concepts, e.g., Wikipedia articles. The main objective is to identify a set of meaningful entities that reveal the topic of this message. The second type is to extract context-specific semantic information out of social media messages. For example, identifying and analyzing social events through user-generated web data have become very popular with the rise of social media (Abdelhaq et al. 2013; Sakaki et al. 2010). Other examples include analyzing health-related posts (Public Health Emergency 2015; Twitter Chicago Foodborne 2014), extracting news stories (Boston Explosions 2013; Sankaranarayanan et al. 2009), and improving responding to emergencies through analyzing real-time social media (China Floods 2012; Hurricane Irma 2017; Hurricane Harvey 2017).

Our work is of the second type where we extract context-specific information from the social media messages. However, our analysis is distinguished from previous work in two aspects. First, to the best of our knowledge, we are the first to use the platform *source* attribute to reveal the structure of market dynamics derived from social media usage at a sub-national spatial scale. Second, our analysis focus solely on meta-data information emphasizing the importance of such information in revealing new semantics from social media data. On the contrary, all previous work focus on the textual content and consider it as the only source of semantics.

3. Datasets and Processing

3.1 *Datasets*

We use a public Twitter dataset that is collected through Twitter public Streaming APIs over two years (2016-2017). The dataset collects only tweets with geo-location latitude/longitude information, either exact or uncertain locations represented as Minimum Bounding Rectangles (MBRs). Although a small percentage of tweets are geotagged, using them still gives correlating results with studies that uses actual sales data, e.g., Atlas (2017), which shows potential representativeness in this small sample to the underlying market sales. In addition, social media provides a cheap and continuously updated data source compared to existing studies as detailed in Section 2.

Our analysis uses tweets that only lie within USA connected states, with total of 1 billion tweets[1]. Each tweet includes several meta-data attributes including timestamp, location, and source platform attributes that are used in our study.

The main attribute used in our analysis, besides time and location, is the source attribute. This attribute indicates from which source the tweet is posted. This is mainly a source platform, e.g., iPhone, iPad, Windows device, Android, Blackberry, and so on. In fact, this attribute has a lot of distinct values. We aggregate these values based on the manufacturing company. For example, the values that contain iPhone, iPad, iOS, or Mac all becomes "Apple"[2]. Having this, we are able to analyze the market share of each manufacturer in different spatial regions and temporal slices. This attribute also contains other values, such as application names. For example, geo-tagged tweets come from applications such as Instagram and Foursquare in 2017 is approximately 18.5% in most states. Also, there are 8.2% of geo-tagged tweets belong to other applications; which could be web browsers for instance. Therefore, we are actually using 73% of the geo-tagged data to analyze the market share. Figure 1 shows the percentages breakdown of different sources in our Twitter dataset.

3.2 *Data Access and Querying*

As a preprocessing step to facilitate efficient data access and querying, we index the whole dataset using a spatial grid index. The grid index covers the latitude/longitude boundaries of USA connected states. The gird is divided into 100x100 equal-space cells. Each grid cell contains a list of data objects where the objects' locations lie inside the cell boundaries. If the object location is uncertain and represented as a Minimum Bounding Rectangle (MBR), then the rectangle centroid is considered as a representative point location. Data of each month, from January 2016 to December 2017, is loaded in a separate index, totaling 24 index structures over the two years, each index with 30GB storage on the average.

The spatio-temporal analysis is then performed through querying each grid index

[1] Tweets can be also filtered to eliminate non-human messages, e.g., chatbot messages, as in Castellini et al. (2017). As this is not the main contribution of this article, we eliminated the effect of automated messages by limiting each user to have only one tweet in the analyzed datasets. Thus, users who post plenty of tweets, such as bots, do not skew the analysis results.

[2] In fact, the source metadata might include semi-structured text that describe the utility used to post the tweet, such as "twitter for iphone" and "twitter for android". We first parse this semi-structured text based on pre-defined patterns mined from all distinct textual values that present in the dataset. Then, we combine all Apple related sources (e.g., iPhone, iPad, and iOS) into a single category called "Apple". The same exact pre-processing is done for Android tweets.

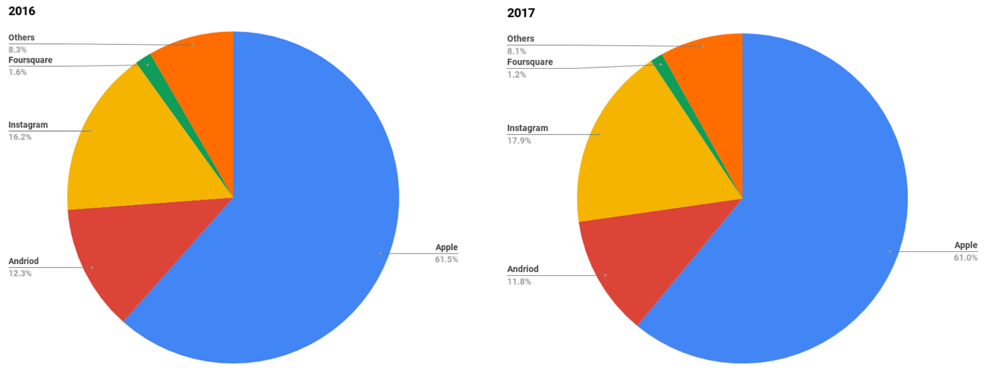

Figure 1. Percentages breakdown of different sources in the analyzed Twitter dataset

with the polygons of US states and counties. Each query retrieves the count of each social media activity or platform for a certain state/county in a certain month, repeating over all states/counties and all months. Then, the counts are aggregated over higher temporal levels, e.g., year. The aggregated counts are then fed to the PySAL library (Rey et al. 2015) to perform the geospatial analysis and generate maps with the different analysis measures and spatial test results over different time slices as detailed in the analysis section.

4. Spatio-temporal Analysis of Platform Market Structure

We apply three sets of analytical methods to examine the market structure for platform use. Our processing and querying of the data show insignificance of all platforms except Apple and Android platforms. As such our analysis will focus on Apple and Android platforms over space and time. Section 4.1 presents a market concentration analysis, Section 4.2 presents a static spatial autocorrelation analysis, and Section 4.3 presents spatial Markov analysis.

4.1 *Platform Concentration over Space and Time*

We draw on recent developments in exploratory space-time data analysis to investigate our dataset. We first consider a measure of the market penetration of a particular platform (Apple, Android, Windows, ...etc) based on a *location quotient* defined as:

$$LQ_{i,r,t} = \frac{a_{i,r,t}/\sum_i a_{i,r,t}}{\sum_r a_{i,r,t}/\sum_r \sum_i a_{i,r,t}} \tag{1}$$

Where $a_{i,r,t}$ is the number of unique social media users of type or platform i in location r occurring during time period t. The location quotient helps identify areas displaying heightened activity relative to that observed in some reference geography. For the latter, we use the national scale as the relevant baseline. More specifically, the numerator in the LQ measures the market share for activity i in location r at time period t, while the denominator serves to compare this local share to the share observed in the broader geography. Areas with $LQ_{i,r,t} >>> 1$ are highlighted as *hot spots* for that particular activity.

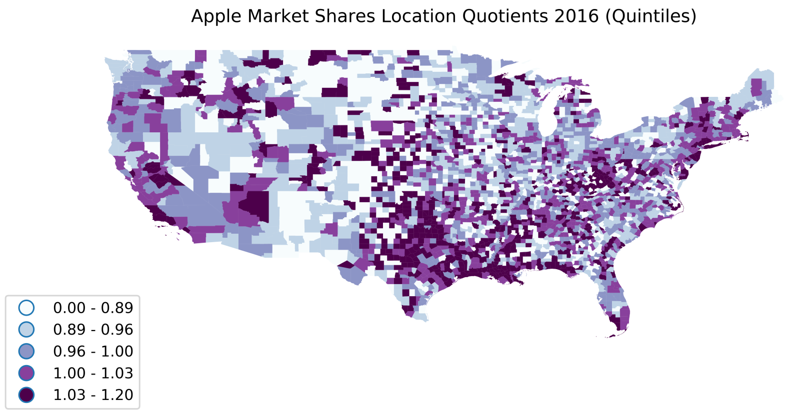

Figure 2. Apple Market Share Location Quotients 2016

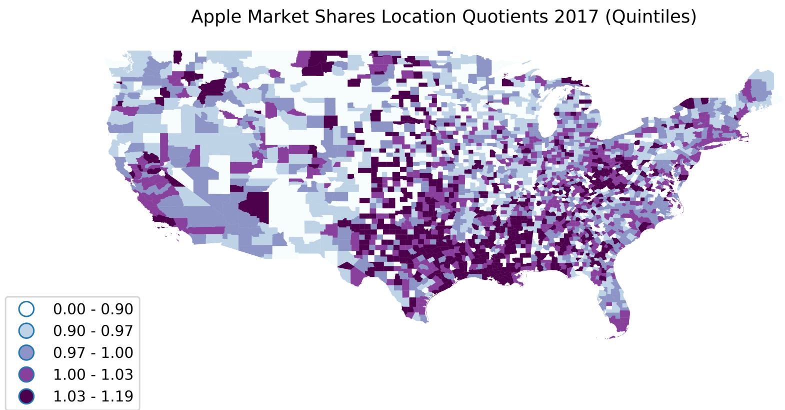

Figure 3. Apple Market Share Location Quotients 2017

Figures 2 and 3 show the location quotients of the Apple platform in 2016 and 2017.[1] Apple platform location quotients (LQ), Figures 2 and 3, show a spatially-correlated distribution where the concentration regions are clustered so that low values of LQ are near each other and so high values. Taking Figure 2 as an example, we can notice low values clustered at the east end of the southwestern states. These low values become higher gradually when moving to the west and to the north. Similarly, high values are clustered near the southern borders and become lower gradually when moving towards the north. This splits the northern area into a gradient that have high values clustered in the west end and become lower towards the east end. These clusters clearly show a spatial clustering for LQ for Apple platform in 2016 over USA.

Comparing the two different years for each platform gives actually different insights over the temporal dimension. For Apple platform, the two years are represented in Figure 2 (2016) and Figure 3 (2017). Comparing the two figures shows an obvious change in Apple concentration in the western region with less concentration in 2017 compared to 2016. Several counties in Washington, Oregon, California, and Nevada are showing less values for location quotients (LQ) in 2017 compared to 2016. On the contrary, 2017 shows slightly higher concentration for Apple near

[1]Because Apple and Android comprise almost the entirety of the US mobile market, their market shares in a given county will be each other's complement. Hence we present figures for Apple only.

Table 1. Tests for Spatial Autocorrelation in Market Shares

Year	Platform	I	$E[I]$	p-value
2016	Apple LQ	0.1027	-0.0003	0.001
	Android LQ	0.0564	-0.0003	0.001
2017	Apple LQ	0.1042	-0.0003	0.001
	Android LQ	0.0804	-0.0003	0.001

the east coast region, yet, less obvious visually than the concentration drop in the west. This might indicate generally a better growth of Android usage in the west compared to the total of Apple platforms, e.g., iPhone, iPad, Mac, and so on.

Finally, comparing the two platforms in terms of concentration is defined by the spatial region and time period. In 2016, Android shows higher concentration in the west and the middle compared to the east, while Apple looks to have a more homogeneous distribution all over the place with less concentration in the east end of the southwestern area. In 2017, the distinction is more obvious. Android shows a high concentration in the west and lower concentration in the east while Apple shows clearly the opposite. These are interesting changes that raise questions about underlying causal mechanisms that should be pursued in future work.

4.2 *Spatial Autocorrelation in Market Structure*

We formally examine these spatial patterns through tests for spatial autocorrelation in the Location Quotients using Moran's I:

$$I_{LQ} = \frac{z_{LQ}^{'} C z_{LQ}}{z_{LQ}^{'} z_{LQ}} \tag{2}$$

where z_{LQ} is the $n \times 1$ vector of county Location Quotients for a given platform in a given year, expressed in deviation from the mean, and C is a spatial weights matrix based on contiguity, where $c_{r,s} = 1$ if counties r and s are contiguous, otherwise $c_{r,s} = 0$.

We evaluate the statistical significance of these statistics based on 999 random spatial permutations of the observed values for the location quotients to develop a reference distribution for Moran's I under the null of spatial randomness.

The results of these tests are reported in Table 1. In both years, and for both platforms, the tests confirm that counties in which each platform is performing differently from its national benchmark are not randomly distributed but display positive spatial autocorrelation.

The tests in Table 1 provide an indication that overall the map patterns of market shares are not random over the two periods. More specifically, this points to the clustering of market shares in space. To complement this global perspective, we turn to a local analysis.

A local indicator of spatial association (LISA) (Anselin 1995) is defined as:

$$I_r = \frac{z_r}{m_2}(z_r \sum_s c_{r,s} z_s) \tag{3}$$

where z_r is the market share in county r and $m_2 = \sum_s z_s^2/n$, and $c_{r,s}$ is the element in row r column s from the spatial weights matrix defined in Equation (2). all other terms as previously defined. The LISA provides a location specific measure

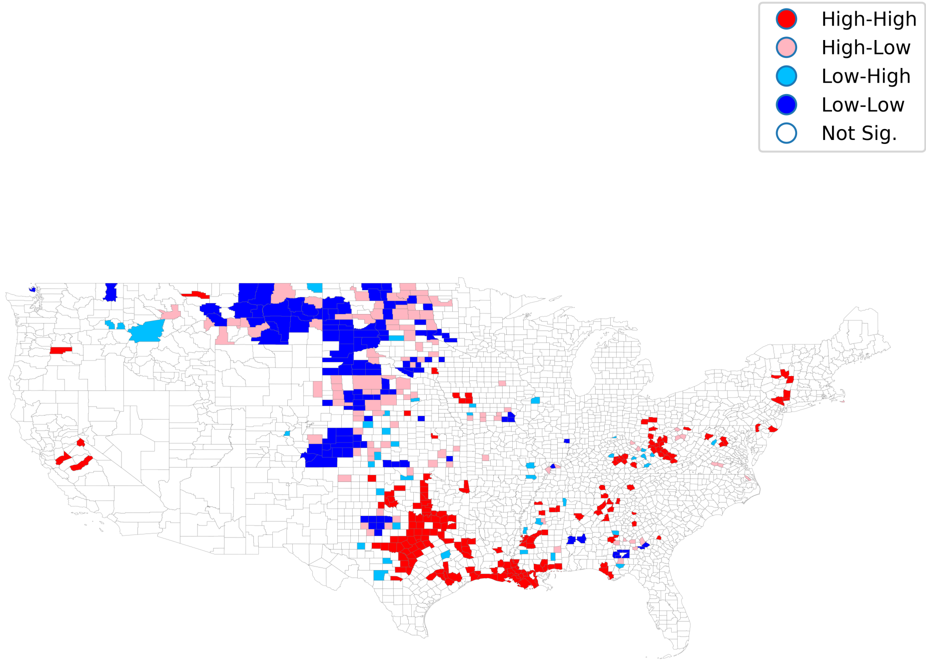

Figure 4. Apple Market Share Local Indicators of Spatial Association 2016

of spatial association. The LISAs can be categorized as representing four different types of spatial association. Counties with high market shares that are neighbored by other counties with high market shares for the same platform represent represent so called “hot-spots” (high-high). A second type of positive spatial association occurs when a county with low market shares is surrounded by other low market share counties, otherwise known as a “cold-spot” (low-low). The other two forms of spatial association can be viewed as negative association in the sense that the market share in the county is inversely related to that found in neighbor hing counties, either the county has a high market share in a neighborhood with low shares (high-low), or the county market is share is low but the platform market shares are higher in the neighboring counties (low-high).

Figures 4 and 5 show the LISA results for Apple market shares. These cluster maps report the LISA values that were found to be statistically significant (p=0.05), where significance is based on conditional local permutations. The dominant type of local spatial association for Apple shares is clustering of like values in space with 179 hot-spots and 119 cold-spots in 2016. In 2017 the number of hot-spots grows to 237, and the number of cold-spots drops to 89. Less common are the negative spatial association counties: in 2016 there were 40 low-high locations, and 103 cases of high-low, while in 2017 there were 35 low-high and 91 high-low locations.

These local lenses provide useful complements to the global pictures that emerged in Figures 2 and 3. For example, visual inspection of the global patterns suggests Apple dominance is particularly pronounced in the south-central part of the country for both years, with a discernible weaker presence in the north-central regions. The local statistics provide insights as to the particular counties within the broader patterns that may be driving the findings of global spatial clustering, particularly the hot-spot counties in the south-central and cold-spot counties in the north-central.

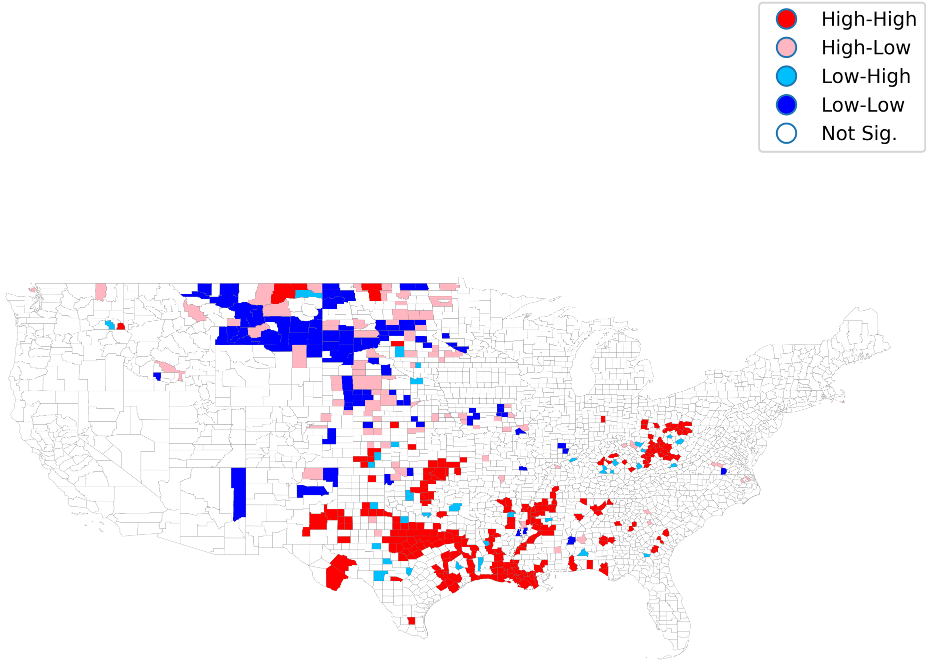

Figure 5. Apple Market Share Local Indicators of Spatial Association 2017

4.3 *Platform Concentration Spatial Dynamics: Markov and Spatial Markov Analysis*

The results of the tests for spatial clustering in market dominance in each time period point to non-random structure in the market, confirming the visual inspection of Figures 2 and 3, and of the local clustering in Figures 4 and 5. An important question related to this clustering is whether it is consistent between the two periods or if there are underlying dynamics?

To examine this question we adopt a discrete Markov chains framework. Here the states of the chain are taken as the five quintiles of the market shares for a given platform. We estimate a first-order probability transition matrix with elements:

$$\hat{p}_{l,k} = \frac{t_{l,k}}{\sum_k t_{l,k}} \tag{4}$$

where $t_{l,k}$ is the number of counties with Apple market shares in the lth quintile in 2016 that transitioned into the kth quintile of market shares in 2017.

Table 2 presents the results of spatial Markov analysis based on the five quantiles C0 (the lowest) to C4 (the highest). The first matrix in Table 2 reports the estimated transition probability matrix across quintiles of the Apple distribution over the 2016-2017 period. On the whole, there is a high degree of movement across the quintiles, as the three central staying probabilities (diagonal elements) are below 0.50, indicating that the Apple market share for a county is more likely than not to move out of its 2016 quintile to a new position in the 2017 distribution. For the second and third quintiles (C1 and C2), the probability is greater for a move upwards into a higher market share quintile rather than downwards, while for the fourth quintile (C3), there is a higher probability of a downward movement.

The global transition probabilities are estimated for the entire set of counties, treating each county's transition in the share distribution as independent of the market shares in the neighboring counties. The strong evidence of spatial autocorrelation we uncovered in the global and local analyses from the previous section suggests that such an assumption may be overly restrictive.

To examine this assumption, we estimate a spatial Markov model (Rey 2001) which extends the classic discrete Markov chain to condition the transition probabilities on the spatial context surrounding a county. More specifically, we first obtain the spatial lag of Android market shares for county r as:

$$LAG_r = \sum_s \frac{c_{r,s}}{\sum_s c_{r,s}} s_s \tag{5}$$

where $c_{r,s}$ is an element from the spatial weights matrix (defined in equation (3)), and then condition the transition probabilities for the market share in county r on the quintile of its spatial lag. The spatial lag is the average of the market shares s_s in the neighboring counties.

The five transition probability tables below the global table are estimates of the transition probabilities for observations whose neighboring counties had shares in a different quintile in the 2016 period. The two formal tests of whether the transition dynamics are different depending on spatial context of a county's market share are both significant. In other words, movement of Apple market shares in a county are not independent of Apple market penetration in neighboring counties at the beginning of the period.

The implications of these differences can be seen by focusing on particular cells from the conditional and unconditional transition matrices. For example, on average a county with a market share that falls in the bottom quintile of the share distribution has a 0.624 probability of remaining there over the year interval. In contrast, if the a county was in that same quintile, but had neighbors who on average were also in the bottom quintile, the probability of the focal county remaining in the bottom quintile rises to 0.636. At the other end of the share distribution, counties in the fifth quintile (C4) had a 0.556 probability of remaining in the top quintile over the transition period. However, counties in the top quintile with neighbors also in the top quintile experienced a higher staying probability (0.668). In other words, the global Markov transition probabilities are masking local spatial context which works to modify the transition dynamics that individual counties experience.

5. Discussion and Conclusions

In this paper, we have analyzed a large public dataset of geosocial Twitter data to extract semantic information that are implicitly embedded in meta-data information. Specifically, we analyzed 1 billion geotagged tweet messages that are located in USA connected states during 2016 and 2017. Our main goal is to explore how meta-data in tweet source, i.e., the platform from which the tweet is posted, may provide information on market segmentation and dynamics. This is useful for commercial companies and their local agents to continuously monitor up-to-date information about market changes using a cheap user-generated source of data. We have shown dominance of Apple and Android platforms throughout US in all time slices, with absence of other platforms, e.g., Windows and Blackberry. In addition, we have shown spatial autocorrelation of both location quotients and market dominance for both platforms, verified through visual maps and formal tests. Our analysis results show interesting clusters that motivate further analysis for different meta-data attributes over different levels of spatial and temporal granularity.

Our work is an initial attempt at using spatial analytics together with geosocial media data to examine market dynamics below the national scale. The goal has been to introduce the application of exploratory spatial data analysis methods to uncover any patterns in market dynamics. There is abundant evidence that the patterns of relative market dominance between Apple and Android are not spatially random. This raises some intriguing questions for future research regarding the mechanisms that may be responsible for these spatial distributions.

From a practical perspective, the finding of spatial dependence in both the market shares at each point and time, and in the transitional dynamics of those shares may be used to inform targeted marketing interventions designed to garner increased market share. Generally speaking, analysts must not focus on a given county (market) in isolation as the spatial dependence we uncover suggests that different forms of spatial interactions and spillovers may be at work. In particular, interventions in counties with the same market penetration may result in different outcomes due to the spillovers latent in the market dynamics. Analysts should explore the possibilities of positive spillovers where the intervention results in diffusion of market gains beyond the target county and into neighboring counties. At the same time, analysts should be cognizant of negative spatial externalities where poor market penetration in surrounding counties dampens the impact of advertising interventions in a target county.

As with all applications of geotagged social media data, care must be taken in the interpretation of the tweets given the highly mobile nature of individuals who post tweets. This raises issues related to different types of spatial uncertainty associated with social media data. The first source of uncertainty surrounds the well known modifiable areal unit problem that arises when multiple spatial aggregations of the social media data may lead to different quantitative and qualitative conclusions (Arbia 1988). We focus on the county scale aggregation here as we feel that is closer to the notion of a market than the state level scale. The question of whether the findings about market dynamics would change with aggregation to the state scale remains a question for future research. The second area of spatial uncertainty pertains to the actual location of where the tweet was made. In some instances the twitter API provides only a bounding box for a tweet and the precise coordinates need to be inferred. Also, the locational semantics of the tweet (i.e., what the content of the tweet tells us about certain locations.) remains an open issue. We see the disambguation of these different types of spatial uncertainty as an important area for future research.

References

Abdelhaq, H., Sengstock, C., and Gertz, M. (2013). EvenTweet: Online Localized Event Detection from Twitter. In *VLDB*.

Anselin, L. (1995). Local indicators of spatial association-LISA. *Geographical Analysis*, 27(2):93–115.

Arbia, G. (1988). *Spatial data configuration in then statistical analysis of regional economic and related problems*. Kluwer Academic Publishers, Dordrecht.

Atlas (2017). Android v iOS Market Share - 2017 Review. `https://deviceatlas.com/blog/android-v-ios-market-share-2017-review`.

Bermingham, A. and Smeaton, A. F. (2010). Classifying Sentiment in Microblogs: Is Brevity an Advantage? In *CIKM*.

Boston Explosions (2013). After Boston Explosions, People Rush to Twitter for Breaking

News. http://www.latimes.com/business/technology/la-fi-tn-after-boston-explosions-people-rush-to-twitter-for-breaking-news-20130415,0,3729783.story.

Castellini, J., Poggioni, V., and Sorbi, G. (2017). Fake twitter followers detection by denoising autoencoder. In *Proceedings of the International Conference on Web Intelligence (WI)*.

China Floods (2012). Sina Weibo, China Twitter, comes to rescue amid flooding in Beijing. http://thenextweb.com/asia/2012/07/23/sina-weibo-chinas-twitter-comes-to-rescue-amid-flooding-in-beijing/.

Facebook Statistics (2018). Facebook Statistics. http://newsroom.fb.com/company-info/.

Hurricane Harvey (2017). Hurricane Harvey Victims Turn to Twitter and Facebook. http://time.com/4921961/hurricane-harvey-twitter-facebook-social-media/.

Hurricane Irma (2017). In Irma, Emergency Responders New Tools: Twitter and Facebook. https://www.wsj.com/articles/for-hurricane-irma-information-officials-post-on-social-media-1505149661.

Jumptap (2011). IPhone Vs. Android: Which Does Your State Prefer? (STUDY). https://www.huffingtonpost.com/2011/08/05/iphone-android-state_n_919488.html.

Kantar (2017). Apple iOS Surges in US. https://us.kantar.com/tech/mobile/2017/apple-ios-surges-in-us/.

Magdy, A., Aly, A. M., Mokbel, M. F., Elnikety, S., He, Y., Nath, S., and Aref, W. G. (2016). GeoTrend: Spatial Trending Queries on Real-time Microblogs. In *SIGSPATIAL*.

Magdy, A., Ghanem, T., Musleh, M., and Mokbel, M. (2014). Exploiting Geo-tagged Tweets to Understand Localized Language Diversity. In *Proceedings of Workshop on Managing and Mining Enriched Geo-Spatial Data, GeoRich*.

Marcus, A., Bernstein, M. S., Badar, O., Karger, D. R., Madden, S., and Miller, R. C. (2011). Twitinfo: Aggregating and Visualizing Microblogs for Event Exploration. In *CHI*.

Meij, E., Weerkamp, W., and de Rijke, M. (2012). Adding Semantics to Microblog Posts. In *WSDM*.

Movoto (2015). These Awesome Maps Show Which States Like iOS And Android The Most. https://www.movoto.com/blog/opinions/ios-vs-android-map/.

Public Health Emergency (2015). Public Health Emergency, Department of Health and Human Services. http://nowtrending.hhs.gov/.

Rey, S. J. (2001). Spatial empirics for economic growth and convergence. *Geogr. Anal.*, 33(3):195–214.

Rey, S. J., Anselin, L., Li, X., Pahle, R., Laura, J., Li, W., and Koschinsky, J. (2015). Open geospatial analytics with PySAL. *ISPRS International Journal of Geo-Information*, 4(2):815–836.

Sakaki, T., Okazaki, M., and Matsuo, Y. (2010). Earthquake Shakes Twitter Users: Real-Time Event Detection by Social Sensors. In *WWW*.

Sankaranarayanan, J., Samet, H., Teitler, B. E., Lieberman, M. D., and Sperling, J. (2009). TwitterStand: News in Tweets. In *GIS*.

Statista (2018). Subscriber share held by smartphone operating systems in the United States from 2012 to 2018. https://www.statista.com/statistics/266572/market-share-held-by-smartphone-platforms-in-the-united-states/.

Statista B (2017). Mobile OS in the U.S. https://www.statista.com/study/11519/mobile-operating-system-us-market-statista-dossier/.

Twitter Chicago Foodborne (2014). Health Department Use of Social Media to Identify Foodborne Illness Chicago, Illinois, 20132014. https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6332a1.htm.

Twitter GeoAds (2012). New Enhanced Geo-targeting for Marketers. https://blog.twitter.com/2012/new-enhanced-geo-targeting-for-marketers.

Twitter Political Sciences (2017). The Power of Images: A Computational Investigation of Political Mobilization via Social Media. https://www.nsf.gov/awardsearch/showAward?AWD_ID=1727459.

Twitter Sociology (2017). Twitter Data Changing Future of Population Research. http://news.psu.edu/story/474782/2017/07/17/research/twitter-data-changing-future-population-research.

Twitter Stats (2017). Twitter Statistics. https://www.statista.com/topics/737/twitter/.

Ukraine Unrest (2014). The Twitter War: Social Media's Role in Ukraine Unrest. news.nationalgeographic.com/news/2014/05/140510-ukraine-odessa-russia-kiev-twitter-world/.

Weber, I. and Garimella, V. R. K. (2014). Visualizing User-Defined, Discriminative Geo-Temporal Twitter Activity. In *ICWSM*.

Xu, Z., Zhang, H., Sugumaran, V., Choo, K.-K. R., Mei, L., and Zhu, Y. (2016). Participatory Sensing-based Semantic and Spatial Analysis of Urban Emergency Events Using Mobile Social Media. *EURASIP Journal on Wireless Communications and Networking*, 2016(1).

Table 2. Spatial Markov Analysis of Apple Market Shares 2016-2017

Spatial Markov Test					
Number of classes: 5					
Number of transitions: 3108					
Number of regimes: 5					
Regime names: LAG0, LAG1, LAG2, LAG3, LAG4					
Test			LR		Q
Stat.			196.303		196.012
DOF			80		80
p-value			0.000		0.000
P(H0)	C0	C1	C2	C3	C4
C0	0.624	0.175	0.045	0.034	0.122
C1	0.180	0.451	0.230	0.087	0.052
C2	0.051	0.225	0.400	0.233	0.090
C3	0.024	0.095	0.272	0.428	0.180
C4	0.121	0.053	0.053	0.217	0.556
P(LAG0)	C0	C1	C2	C3	C4
C0	0.636	0.133	0.036	0.040	0.156
C1	0.192	0.392	0.238	0.092	0.085
C2	0.067	0.233	0.311	0.233	0.156
C3	0.093	0.148	0.278	0.370	0.111
C4	0.276	0.114	0.033	0.106	0.472
P(LAG1)	C0	C1	C2	C3	C4
C0	0.616	0.185	0.048	0.041	0.110
C1	0.216	0.488	0.179	0.068	0.049
C2	0.057	0.270	0.328	0.254	0.090
C3	0.019	0.143	0.276	0.333	0.229
C4	0.163	0.023	0.047	0.233	0.535
P(LAG2)	C0	C1	C2	C3	C4
C0	0.562	0.271	0.073	0.042	0.052
C1	0.153	0.497	0.252	0.080	0.018
C2	0.063	0.245	0.415	0.220	0.057
C3	0.041	0.115	0.328	0.344	0.172
C4	0.098	0.073	0.098	0.159	0.573
P(LAG3)	C0	C1	C2	C3	C4
C0	0.714	0.190	0.024	0.000	0.071
C1	0.134	0.402	0.268	0.134	0.062
C2	0.032	0.212	0.487	0.205	0.064
C3	0.007	0.079	0.270	0.461	0.184
C4	0.103	0.068	0.060	0.325	0.444
P(LAG4)	C0	C1	C2	C3	C4
C0	0.577	0.141	0.056	0.028	0.197
C1	0.222	0.444	0.222	0.056	0.056
C2	0.042	0.147	0.411	0.274	0.126
C3	0.011	0.053	0.234	0.527	0.176
C4	0.033	0.014	0.047	0.238	0.668