Does Reciprocal Gratefulness in Twitter Predict Neighborhood Safety?: Comparing 911 Calls Where Users Reside or Use Social Media

Ann Marie White,[1] Linxiao Bai,[1] Christopher Homan,[2] Melanie Funchess,[3] Catherine Cerulli,[1] Amen Ptah,[1] Deepak Pandita,[1] Henry Kautz[1]

[1] University of Rochester, [2] Rochester Institute of Technology, [3] Mental Health Association of Rochester
{annmarie_white, catherine_cerulli, amen_ptah}@urmc.rochester.edu, {lbai2, dpandit2}@ur.rochester.edu, cmh@cs.rit.edu, mfunchess@mharochester.org, kautz@cs.rochester.edu

Abstract

Is there a relationship between urban neighborhood safety and helpful or supportive user networks on Twitter? An interdisciplinary, community-partnered team analyzed one year (2013-2014) of geo-tagged tweets from a 16-county region to generate a network of users who expressed gratefulness to one another. Call counts to 911 (2013-2015) around locations in the urban center indicate safety-oriented activity in residential areas. We compared frequencies of 911-related police activity across geographic pockets (200m radius or 200×200 meter^2 areas) where mutually helpful or thanked users lived or frequently tweeted, to pockets randomly selected in proximity. The 13 naturally helpful users with predicted home locations in this city lived where fewer 911 calls for police service were initiated over time, on average. Our results show that close neighborhood locations where 79 naturally helpful users had strong local geographic ties in this urban area, as evidenced by Twitter use patterns, functioned as safe zones needing fewer policing services on average than surrounding areas. We discuss the implications of predicting real-world needs for police services based on supportive qualities of one's social media neighborhoods.

Introduction

Individuals influence the caring present in their surroundings and promote social system features that lead to a variety of prosocial benefits. Social support is often acknowledged among those witnessing positive social benefits when caring actions are performed (Cohen 1982), and is a class of attributed prosocial behavior referred to as natural helping. Natural helpers are defined as, "Individuals, in our neighborhoods, who residents naturally turn to, or seek out in difficult times because of his or her concern, interest and innate understanding" (Patterson and Memmott 1992). Attributions of helpfulness and prosocial behavior, experienced as gratefulness in real or online environments, can promote mental wellness and resiliency (Lindström and Eriksson 2005, Dominijanni et al. 2015).

Empathy among neighbors also produces collective capacities that reduce risks for violence (White and Funchess 2012). Neighbors who feel socially cohesive and able to exert control together, or help each other's wellbeing, can dampen crime or signals of risk such as police activity on a street (Sampson et al. 1997). While natural helping in physical neighborhoods can foster social connectedness and promote safety, wellness, and protection from violence, can natural-helping capacity in virtual, social-media-based settings mirror or predict these qualities? Can protective influences within social media signal urban neighborhood resilience among residents in precise locations?

Members of our academic-community partnership witnessed events where both violence and protection from violence in neighborhood life played out in social media. Here, we ask if locations with greater safety in urban spaces co-locate with social media networks that recognize and acknowledge prosocial or helpful behaviors. The hypotheses, emerging from this community-based participatory data science approach, are: (H1) Precise neighborhood locations with naturally helpful residents, identified via Twitter user networks that expressed mutual gratitude, are more protected from police service needs. (H2) The 911 police services occur less within neighborhood parcels where social-media-based natural helpers frequently tweet.

Related Work

Home Location Inference. Researchers have used social media to infer and predict the health of individuals (Sadilek and Kautz 2013), or to discover risks of violence such as alcohol use in predicted home location, for comparison across geographic areas (Hossain et al. 2016).

Natural helper detection. Social network analysis of social media data reveals tightly-connected communities who interact together in virtual public spaces, independent of geospatial proximity. Social media data like tweets can reveal helping and prosocial subnetworks: users who reciprocal "thank" or are mutually grateful also create greater prosocial and mental wellness-oriented content in Twitter (Dominijanni et al. 2015).

Validity of 911 calls data in studies of violence. Studies have employed 911 call data to examine instances of violence and effects of preventative interventions, demonstrating validity for assessing safety-related features of settings and event histories (Kothari et al. 2012).

Methods

Study Design

A rectangular area roughly 4 miles in width and 7 miles in length that aligns with the city boundaries defined the subset of tweets and 911 calls of interest. To permit hypothesis testing of mean differences in neighborhood police presence linked to safety-related activity of 911 calls, we compared Twitter-user network patterns across a similar time period. Social network analysis identified users based on linguistic features of tweets intentionally sent ("mentioned") to other users that acknowledge helpfulness. Users in the helping network were snapped to precise geographic home locations or grid locations where they tweeted. We visualized the 911 call data and tweet locations for users on maps. An example of tweets and 911 calls pinned near a home is shown at pagliaccl.github.io/rpdPresentation.html.

We applied t-tests (Whitley and Ball 2002) to compare means of the 911 call-related police activity distributions in small geographies of immediate proximity to helpers' key locations (home residence or where they frequently tweet). We conducted two sets of hypothesis testing with either the user home or the tweet activity location as the unit of analysis. We used Python 2.7 and scipy.stat package to assess the relationship between helpers' local activity life space and neighborhood police presence.

Data

Twitter. Using DataSift, we collected 7 million geo-tagged tweets for a period of 1 year (July 2013 and June 2014) from a 16-county region of a northern state. This region includes one main urban area and surrounding suburban and rural communities. The urban center is recognized nationally as having high rates of poverty and violence.

Safety-related activity and police presence. A public-access property information dataset provided detailed assessment and zoning information about each address parcel in the city, including number of housing units. A data use agreement provided all the approximately 800,000 911 calls for service involving the urban addresses during 2013-2015. A "situation found" field substantiates the police service need based on what police find and act on at the address referenced in the call, including emergencies (e.g., mental health, 2%; violence, 5%; quality of life threats or accidents, 34%), performing a routine police activity that maintains quality of life (e.g., school crossing or traffic violation, 11%), and other situations (48%).

Outcome: Police Service Calls

Calls in street address formats were geo-coded to the WGS-84 Web Mercator Auxiliary Sphere (WMAS), planar projection coordinates, using ArcGIS and a NY State geocoding service (gis.ny.gov). The volume of 911 service calls (V) measured the degree of insecurity or safety-related needs. Higher counts of police services pinned to a unit location of analysis indicate greater safety-related needs. A call density map via ArcGIS is presented in Figure 1.

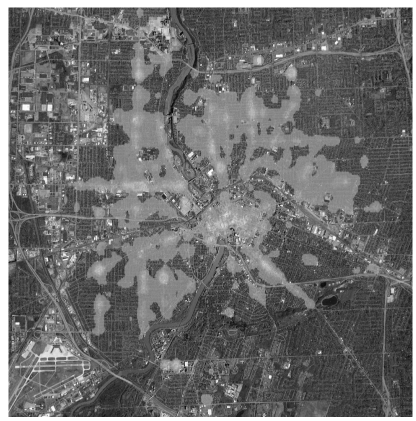

Figure 1. Heat map of all police activity reported in 911 from 2013-2015 within the urban bounding box

We also binned 911 calls after discretizing the urban area into 200m×200m (meter) cells. This Eulerian approach creates a grid of squares to form units of analysis. To account for population density, we computed a normalized call volume (NCV=V/F) for each unit cell using the count of family units zoned within (F) (Haas et al. 2011).

Sample of Users Thanked/Attributed as Helpful

We replicate methods in Dominijanni et al. 2015, based on Garlaschelli and Loffredo 2004, to construct a thanking network. We identified user pairs (u,v) who mentioned each other and acknowledged helpfulness or thanked one another (tweets with at least one keyword of "thanks", "thx", "thnx", "ty", or "thank" created a directed tie, u → v). e.g. "*@username thanks jack, means a lot*" and "*@username thanks! Good seeing you today!*" These pairs form a gratefulness subgraph, the @username mentions network, where users who were attributed to be helpful by others and who also expressed gratefulness themselves are nodes. Edges between nodes indicate a reciprocated acknowledgment of helpfulness between two users who directly tweet one another. We infer helpfulness by measuring where people share gratitude with each other. In this network subgraph, a terminal helper (TH) is a user who is a terminal node (n=250), and a central helper (CH) is a non-terminal node (n=165) (Newman 2010). While 79 of the 415 mutually grateful users identified ever tweeted inside the city, the home location algorithm described in Hossain et al. 2016 labelled 10 CH and 3 TH users as city residents.

Results

H1: Analysis of user home location. For each of the 13 residents, we examined a neighborhood influence area equivalent to a 200m radius circle centered at the predicted home location. To test if the presence of a natural helper results in a different distribution of call volume V in a neighborhood, 500 randomly selected points inside a circular ring area (300m inner radius to 400m outer radius distance from helper's home) comprised a comparison group V'. This matched sample for pairwise comparison is as if a helper's peer neighbor lived at each selected point.

For each "dummy" peer neighbor, we recorded call volume in a 200m radius as V'_i , and an average call volume ($\overline{V}'$) for the 500 random points. For each helper cross-referenced to live in the urban area we derived a call volume V, paired with each user an average $\overline{V}'$ as its case control, and calculated a pairwise difference (V-$\overline{V}'$) (see table).

Subject	1	2	3	4	5	6	7	8	9	10	11	12	13
V	344	0	780	0	0	2	402	267	878	349	0	0	313
$\overline{V}'$	551.9	0.5	758.9	0.4	0	0.8	563.5	550.4	873.6	371.7	0	0.1	591.6
V-$\overline{V}'$	-207.9	-0.5	21.1	-0.4	0	1.2	-161.5	-283.4	4.4	-22.7	0	-0.1	-278.6

Table 1. Total 911 call volumes by helpers' homes (V) and average call volumes by random neighboring points ($\overline{V}'$)

A one-sided pairwise t-test examined if the average V' ($\overline{V}'$) was significantly greater than V across the pairs of home locations. We rejected the null hypothesis $V \geq \overline{V}'$ (mean pair difference=-71.4, sd=32.2, t-stat=2.22, p<.03). We failed to reject the null hypothesis that either variance of V and $\overline{V}'$ were different (F-stat=.81, p<.72; Snedecor and Cochran, 1989), or that the distributions of V-$\overline{V}'$ were non-normal (Pearson's chi-square test, p<.31). A natural helper's home location area had less emergency calls or police presence than randomly selected peer neighbors, on average, when we tested different random seed generators.

H2: Analysis of social media activity locations. We compared cells that were helpers' tweeting locations that exceeded a *k* frequency of tweets sent (varied from 0 to 100+, in increasing increments of 5) to locations with no helper tweet activity. As *k* criterion grows, confidence increases that a cell is a meaningful social media use life space (as compared to signal noise). A total of 1,328 cells recorded tweets from any helper at least once: terminal helpers (THs) ever tweeted in 499 cells, and central helpers (CHs) tweeted in 987 cells. Units with no helper tweet activity were a control comparison (n=2,992).

The difference in mean call volumes (NCV) between selected helper and control tweeting activity locations were compared via a two-sample t-test for equal means (unequal variance, one-sided). Cells with more tweet activity by any helper have less 911-reports of police activity. NCVs ranged from 43.1 to 3.5, respectively across all *k* in increasing order, as compared to the control group mean average of 22.4 calls. Cells with helpers of any centrality (both TH and CH) had lower police service needs, on average, than areas without any social media-based helper activity (t-stats>=6.04, p<.001). As the figure below illustrates, the t-statistic that corresponds to the NCV mean difference between helper and comparison groups increases as *k* tweet count increases. As the *k* tweet frequency increases beyond 65 tweets (n=35 cells), the null hypothesis of equal means is consistently rejected (higher t-statistics, lower p-values). Overall, as the strength of helpers' social media activity increases, less 911-related active police presence is observed on average, as compared to locations where no helpers tweeted.

The figure shows that the CH t-statistic value is consistently at or above the TH and CH+TH values for the same *k* tweet count, indicating a greater mean difference in police service calls when CH (or greater centrality in the helping network) determines test locations. When specifying social media use areas that only include users more central to the helping network (CH), we often observe greater NCV differences than when we compare activity areas of less central helpers (TH). Neighborhood cells where users more central to the helping network tweeted a *k* of 50+ times (n=28) have less police presence than cells where they never tweeted (t-stats>=-5.05, p<.001), as do cells with 15+ but less than 26 tweets (p<.05). Neighborhood pockets where users are on the helping network periphery (TH) were highly active on Twitter (n=10 cells with greater than 55 tweets per year, with mean NCVs of 6.12 or less) consistently recorded lower average NCV as compared to cells without any helpers (t-stats >=-6.52, p<.001).

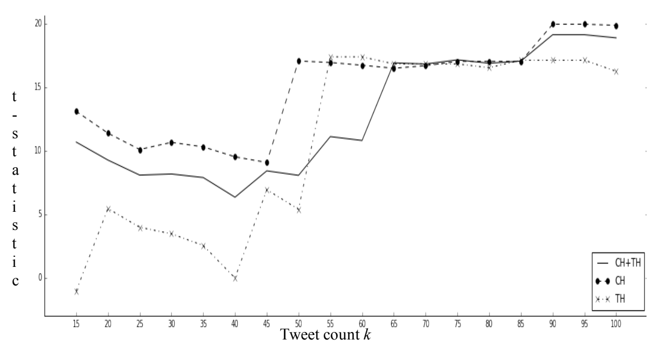

Figure 2. T-statistics for NCV differences by helper groups: Trends in reductions of average NCV across k tweets

Discussion and Conclusion

Identifying natural helper reciprocity within neighborhood locations, measured via mutually grateful social media user network subgraphs, corresponded to precise "colder" police activity spots in urban areas. Patterns of fewer police service needs existed in locations where mutually supportive users reside or are more active in social media. General reciprocity has been found to have implications for community helping behavior (Althoff et al. 2014), but our data cannot infer causality due to gratefulness found in tweets. Nor did we yet test if all users live and travel in safer locations. But, greater centrality (in gratefulness graphs) and activity of helpers in precise locations across time reduces noise in the data to reveal areas with reduced police activity. Greater centrality, a feature that conveys stronger confidence in the helping role performance, corresponded to less observed police activity in neighborhoods. Additional research is needed to interpret the cold spots – are these due to undetected threats (e.g., illegal economic activity), perceived acceptability of threats, or distrust in 911?

Neighborhoods exert a strong and comprehensive influence on the lives of their residents – positive and negative - in a variety of social matters of great import (Phillips and Shonkoff 2000). Prior research demonstrates that peers living within a 255 meter radius can influence other neighbors' problem behaviors, but not peers residing larger distances away (400 or 800m) (Caughy et al. 2013); could other variables or spatial features generate the findings? We did limit spatial autocorrelation threats in adopting this precisely located peer comparison, while controlling for nonrandom location features associated with neighborhood police presence such as residence density. While the precise localized patterns in 911 police responses meets expectations if supportive clusters of neighbor-to-neighbor interactions are indeed present within 200m, the degree of spatial autocorrelation will be explored in future work.

Our work provides further evidence that social media, when georeferenced, can be a resource among those seeking to identify neighbors who can grow collective action in preventing violence in precise locations. The 911 data tells one only about police activity, but Twitter can help identify where mechanisms for neighborhood improvement are present across large geographic areas. While the limitations of Twitter in advance predicting unique 911 fire emergencies is noted in research (Zaman et al. 2017), useful patterns and networks exist that, when detected together, map onto precise pockets with less need for police services.

Acknowledgements

The National Institute of General Medical Sciences under award number R01GM108337, and the NYSTAR NYS Center of Excellence in Data Science supported this research in part. We thank the Big Data Docents, our community collaborative research board, for their guidance, and Nick Pettiti and Yutong He for their assistance.

References

Althoff, T., Danescu-Niculescu-Mixil, C., and Jurafsky, D. 2014, May. How to ask for a favor: A case study on the success of altruistic requests. In *Proceedings of the Eighth International Conference on Web and Social Media (ICWSM)-2014*. Palo Alto, CA: AAAI Press.

Caughy, M.O.B., Leonard, T., Beron, K., and Murdoch, J. 2013. Defining neighborhood boundaries in studies of spatial dependence in child behavior problems. *International Journal of Health Geographics* 12(1):24.

Cowen, E.L. 1982. Help is where you find it: Four informal helping groups. *American Psychologist* 37(4):385.

Dominijanni, J., White, A. M., Homan, C., Funchess, M., He, H., Tu, X., and Kautz, H. 2015. "Thanking goodness:" Mutual gratitude as signals of wellness in Twitter networks. In *Sunbelt XXXV: Conference of the International Network for Social Network Analysis (INSNA)*, 308. Huntington, WV: INSNA.

Garlaschelli, D., and Loffredo, M.I. 2004. Patterns of link reciprocity in directed networks. *Phys. Rev. Lett.*, 93 (26):8701.

Haas, A.J., Gibbons, D., Dangel, C., and Allgeier, S. 2011. Automated surveillance of 911 call data for detection of possible water contamination incidents. *International Journal of Health Geographics*, 10(1):22.

Hossain, N., Hu, T., Feizi, R., White, A.M., Luo, J. and Kautz, H.A. 2016, May. Precise Localization of Homes and Activities: Detecting Drinking-While-Tweeting Patterns in Communities. In *ICWSM-2016*, 587-590. Palo Alto, CA: AAAI Press.

Kothari, C.L., Rhodes, K.V., Wiley, J.A., Fink, J., Overholt, S., Dichter, M.E., Marcus, S.C. and Cerulli, C. 2012. Protection orders protect against assault and injury. *Journal of Interpersonal Violence*, 27(14):2845-2868.

Lindström, B., and Eriksson, M. 2005. Salutogenesis. *Journal of Epidemiology & Community Health*, 59(6):440-442.

Newman, M. 2010. *Networks*. New York: Oxford Univ. Press.

Patterson, S.L., Memmott, J.L., Brennen, E.M., and Germain, C.B. 1992. Patterns of natural helping in rural areas. *Social Work Research and Abstracts*, 28(3):22-28.

Phillips, D.A., and Shonkoff, J.P. eds. 2000. *From neurons to neighborhoods*. Wash., DC: National Academies Press.

Sadilek, A., and Kautz, H. 2013, Feb. Modeling the impact of lifestyle on health at scale. In *Proceedings of the 6th ACM international conference on web search and data mining*, 637-646.

Sampson, R.J., Raudenbush, S.W., and Earls, F. 1997. Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, 277(5328):918-924.

Snedecor, G. W., and Cochran, W. G. 1989. *Statistical Methods*, (8th Ed.). Ames, IA: Iowa State University Press.

White, A.M., and Funchess, M. 2012. Neighborhood natural helpers as community-based participatory research partners. In *25th Children's Mental Health Research & Policy Conference*. Tampa, FL: Department of Child and Family Studies, USF.

Zaman, A., Hossain, N., and Kautz, H.A., 2017. Twitter911: A Cautionary Tale. *ICWSM-2017,* 712-715. Palo Alto: AAAI Press.