Localness and Urbanness in Geographic Crowd Work

Sukrit Venkatagiri
Department of Computer Science and Center for HCI
Virginia Tech
Arlington, VA, USA
sukrit@vt.edu

Jacob Thebault-Spieker
Department of Computer Science and Center for HCI
Virginia Tech
Blacksburg, VA, USA
jthebaultspieker@vt.edu

Aliza Camacho
Wellesley College
Wellesley, MA, USA
acamacho@wellesley.edu

Kurt Luther
Department of Computer Science and Center for HCI
Virginia Tech
Arlington, VA, USA
kluther@vt.edu

CSCW'18, November 3-7, 2018, Jersey City, NJ, USA.

Abstract
Localness is increasingly viewed as fundamental to crowd work, from volunteer mapping in OpenStreetMap to paid verification of local information on Amazon Mechanical Turk (MTurk). However, it remains unclear what makes localness beneficial. We investigate how the common meaning of localness, *physical proximity*, influences people's success in geographic crowd work and propose a novel second dimension of localness, *urbanness distance*. In a study of volunteer work in Humanitarian OpenStreetMap Team (HOT), we find that success decreases when volunteers live (a) farther away or (b) in differently urban (or rural) areas. In a second study of crowdsourced image geolocation on Amazon Mechanical Turk, proximity and urbanness distance provides no additional performance benefits. Our results suggest that previous conceptions of localness are too simple. Instead, localness may be a multidimensional concept, varying based on the kind of task or incentive.

Author Keywords
OpenStreetMap; Humanitarian OpenStreetMap; geolocation; localness; geographic HCI; population density

Introduction and Motivation
Geographic crowd work is a fundamental aspect of major crowdsourcing platforms, from volunteers contributing to OpenStreetMap, writing about geographic topics on

Author Bios

Sukrit Venkatagiri is a Computer Science Ph.D. student at Virginia Tech. He studies and builds real-time crowdsourcing systems to augment and scale-up the expertise of investigators in journalism and other fields. He is interested in leveraging the local familiarity and skills of individuals in the crowd to combat misinformation.

Jacob Thebault-Spieker is a Postdoctoral Associate at Virginia Tech. Geographic variations in online content or services mean that rural and low-SES regions are disproportionately underserved. Jacob's doctoral work used robust geostatistical methods to measure and understand the mechanisms that underlie these geographic biases.

Aliza Camacho is an undergraduate student majoring in Anthropology and Computer Science at Wellesley College.

Kurt Luther is an Assistant Professor of Computer Science and (by courtesy) History at Virginia Tech, where he directs the Crowd Intelligence Lab.

Wikipedia, to paid workers on Amazon Mechanical Turk helping verify local information [5, 6]. Current geographic crowd work places an emphasis on being *local*.

The First Law of Geography [12] states: "near things are more related than distant things", so leveraging the local expertise of crowds nearby makes sense, especially with such a large pool of potential workers online. Prior work shows that local geographic crowd work produces richer [4], more diverse [13], or more accurate [1] content, and thus some researchers [10, 4, 11] argue that local contributions help counteract some of the systematic geographic biases [8, 2, 7, 9] that disadvantage rural areas in geographic crowdsourcing systems. Low SES and rural places receive lower quality volunteered geographic crowd work, partly due to how geographic crowd volunteers spend their time [11].

In this research, we test the feasibility of a multidimensional conception of localness, focusing on both the common dimension of *proximity* and hypothesizing a second dimension of localness informed by prior work: *urbanness distance*. Urbanness distance is the difference in the urbanness of where a person lives and the urbanness of where they perform geographic crowd work. We study this across two different geographic crowd platforms, with different tasks and incentive structures. Our research questions are:

RQ1: Does proximity affect an individual's success in geographic crowd work?
RQ2: Does urbanness distance affect an individual's success in geographic crowd work?

We also highlight difficulties in recruiting geographically diverse crowd workers on MTurk. We then discuss the implications of multdimensional localness for mitigating rural biases in geographic crowd systems.

Study 1: Humanitarian OpenStreetMap

Our first study focuses on Humanitarian OpenStreetMap Team (HOT), where volunteers perform the generative task of tracing satellite imagery. The HOT community has built a structured contribution workflow designed for disaster and other humanitarian mapping efforts globally, and presents an opportune field site to study our research questions.
In HOT, tasks can be labeled as either (a)'validated', i.e., accepted as done; or (b) 'invalidated', indicating they must be re-mapped or fixed.

To answer our research questions, we systematically determined crowd volunteers' home locations by manually coding 1,385 of their 'About' pages. We then calculated the success rates of their contributions (whether they were validated or invalidated). To answer RQ1, we computed their physical distance from the locations they contributed to. To answer RQ2, we computed their urbanness distance from locations they contributed to. Here, urbanness distance is measured in terms of the difference in population density between a volunteer's home location and that of the locations they contribute to. For instance, if a volunteer lived in New York City, but helped map a rural area, their urbanness distance would be quite large.

We then ran a mixed effects binomial logistic regression because each of the values for our validation states (the dependent variable in our analysis) is categorical. Our dependent variable was $validation\_status$, and our fixed-effects variables were $proximity\_km\_log2$ and urb_dist_log2. We also included a random effect for each contributor, to account for individual variation. The results of this regression are shown in Table 1.

Table 1 shows that within HOT, being from a place that is either 1) near to the HOT task or 2) similarly urban (or rural) has a positive impact on geographic crowd work success.

Table 1: Results of Binomial Mixed-Effects Logistic Regression

	validation_status
proximity_km_log2	$-0.068^{\dagger}$
	(0.036)
urb_dist_log2	-0.075^{**}
	(0.014)
Constant	4.115^{**}
	(0.471)

†p≤0.1; *p≤0.05; **p≤0.01

More formally, both proximity and urbanness have an influence on crowd volunteers' odds of being validated.

With regard to our proximity results, we find that being from a place further away from the task location very likely (p = 0.06) decreases a crowd volunteer's odds of having their mapping work validated. Consider two HOT volunteers, one who lives twice as far away from the HOT task as the other. Holding all else constant, we would expect the crowd volunteer who is farther away to have 7% lower odds of being validated than the one who is physically closer.

With regard to our urbanness distance results, we find that being from a place that has a different urbanness than the task location decreases a crowd volunteer's odds of having their mapping work validated. Consider two HOT volunteers contributing to the same task, in a place with a population density of 20 people/km^2. Both crowd volunteers are equidistant from the task, but one crowd volunteer lives in a place with a higher population density of 40 people/km^2, and another lives in a place with an even higher population density of 80 people/km^2; i.e., the latter crowd volunteer is 'half as urbaness local' as the former. We expect the crowd volunteer with the larger urbanness distance to have 7% lower odds of their task being validated than the crowd volunteer with the shorter urbanness distance.

Taking our results together, we find that both proximity and urbanness distance influence the success of geographic crowd work, on a volunteer-based mapping task in HOT.

Study 1: Humanitarian OpenStreetMap

Our first study focuses on Humanitarian OpenStreetMap Team (HOT), where volunteers perform the generative task of tracing satellite imagery. The HOT community has built a structured contribution workflow designed for disaster and other humanitarian mapping efforts globally, and presents an opportune field site to study our research questions.
In HOT, tasks can be labeled as either (a)'validated', i.e., accepted as done; or (b) 'invalidated', indicating they must be re-mapped or fixed.

To answer our research questions, we systematically determined crowd volunteers' home locations by manually coding 1,385 of their 'About' pages. We then calculated the success rates of their contributions (whether they were validated or invalidated). To answer RQ1, we computed their physical distance from the locations they contributed to. To answer RQ2, we computed their urbanness distance from locations they contributed to. Here, urbanness distance is measured in terms of the difference in population density between a volunteer's home location and that of the locations they contribute to. For instance, if a volunteer lived in New York City, but helped map a rural area, their urbanness distance would be quite large.

We then ran a mixed effects binomial logistic regression because each of the values for our validation states (the dependent variable in our analysis) is categorical. Our dependent variable was $validation\_status$, and our fixed-effects variables were $proximity\_km\_log2$ and urb_dist_log2. We also included a random effect for each contributor, to account for individual variation. The results of this regression are shown in Table 1.

Table 1 shows that within HOT, being from a place that is either 1) near to the HOT task or 2) similarly urban (or rural) has a positive impact on geographic crowd work success.
More formally, both proximity and urbanness have an influence on crowd volunteers' odds of being validated.

With regard to our proximity results, we find that being from a place further away from the task location very likely (p = 0.06) decreases a crowd volunteer's odds of having their

mapping work validated. Consider two HOT volunteers, one who lives twice as far away from the HOT task as the other. Holding all else constant, we would expect the crowd volunteer who is farther away to have 7% lower odds of being validated than the one who is physically closer.

With regard to our urbanness distance results, we find that being from a place that has a different urbanness than the task location decreases a crowd volunteer's odds of having their mapping work validated. Consider two HOT volunteers contributing to the same task, in a place with a population density of 20 people/km^2. Both crowd volunteers are equidistant from the task, but one crowd volunteer lives in a place with a higher population density of 40 people/km^2, and another lives in a place with an even higher population density of 80 people/km^2; i.e., the latter crowd volunteer is 'half as urbaness local' as the former. We expect the crowd volunteer with the larger urbanness distance to have 7% lower odds of their task being validated than the crowd volunteer with the shorter urbanness distance.

Taking our results together, we find that both proximity and urbanness distance influence the success of geographic crowd work, on a volunteer-based mapping task in HOT.

Table 2: Number of Crowd Workers in Each Urbanness Level

Urbaness Level	CA	IL	Total
1	201	33	234
2	39	27	66
3	44	1	45
4	9	10	19
5	2	12	14
6	1	31	4

Study 2: Crowdsourced Image Geolocation

In our second study, we focus on a different geographic crowd work task and in a different context: the analytical task of image geolocation performed by paid crowd workers on Amazon Mechanical Turk (MTurk). In prior work, we designed an interface to facilitate crowd-supported image geolocation [5], where crowd workers compare 16 cells of satellite imagery with a provided aerial diagram, and mark each cell as Yes/Maybe if it matches, or No if not. In our study, all tasks had exactly one correct cell and 15 decoys.

Using this system, we conducted an experiment to study how an individual crowd worker's localness influences their performance in image geolocation. First, we systematically selected two images based on on geographic and visual diversity, among other criteria [5]. The images were located in Chicago and Waverly, IL. We then generated aerial diagrams of the two images with the help of a professional from UpWork.

Next, to answer RQ1, we hired two groups of crowd workers, using MTurk's location qualifications, to perform the image geolocation task. We hired local crowd workers from IL (same location as the image), and non-local workers from California (CA), another populous US state (to ensure a large pool of workers), yet far away from IL.

To answer RQ2, we classified our two image locations based on their urbanness, using the the NCHS Urban–Rural Classification Scheme for Counties [3], which describes six county-level urbanness categories ranging from most urban (level 1) to most rural (level 6). We use this to calculate workers' urbanness distance.

We paid workers $1.21 for our task with an estimated completion time of 10 minutes ($7.25/hr). Workers completed an online consent form, a survey with demographic questions, and the image geolocation task. We aimed to hire an equal number of workers from CA and IL. Despite our best efforts, including a number different recruitment strategies over a two-week period, we were able to hire 323 workers from CA, but only 101 from IL. Table 2 shows the number of workers in each urbanness level, after filtering out low-quality workers' data. Given our difficulties in hiring geographically-targeted workers, we report results only for Waverly, IL.

Results. Overall, workers from both states performed well in identifying the single correct cell, with a true positive rate of 72.97% for CA workers and 80.23% for IL. The false positive rate was low at 10.81% for CA workers and 13.11% for IL. There was no statistically significant difference between CA and IL workers' true positive rate (Mann–Whitney U-test, W = 11424, p = 0.117) or false positive rate (Mann–Whitney U-test, W = 11804, p = 0.174). We found similar results when workers were grouped based on their urbanness distance: they all performed equally well and there was no statistically significant difference between groups.

Discussion and Future Work

Putting Our Results in Context. Comparing Studies 1 and 2, we find mixed results in response to our research questions. With regard to the question posed for the workshop, our results point to two important conclusions. First, there are contexts where a multi-dimensional conception of localness (being physically nearby, or from a place that is similarly urban/rural to the task) matters in geographic crowd work. However, there are other contexts in which rural areas may benefit substantially from the wisdom of the crowd. In Study 2, workers performed paid, analytic, geographic crowd work equally well, regardless of how far away or how differently urban their home was from the task location.

Implications for Mitigating Geographic Biases. One difference between our two studies is the type of task in each. Study 1 used a synthesis task, while Study 2 used an analytical task. Our results show that for a synthesis task, 'proximity' or 'urbanness' localness affects worker success. Thus, for synthesis tasks, incorporating rural perspectives may be critical. Conversely for analytical tasks, geographic crowd work systems may be able to benefit from the crowd's ability to achieve high-quality work, regardless of 'proximity' or 'urbanness' localness between the crowd worker and the task location. The second difference between our two studies is the incentive structures. This suggests that in volunteer crowd work settings, rural places may be disadvantaged, because localness is important. However, paying for geographic crowd work may allow rural regions to take advantage of the crowd. For example, paid geographic crowd work and/or analytical task design could be used to mitigate geographic biases like lower quality content in rural areas (e.g., [2, 8, 7]).

Future Work. In Study 2 we faced difficulties in recruiting geographically-targeted crowd workers, especially those from more rural areas. While prior work has found that workers on MTurk are broadly representative of the general population, our data suggests that only about 5% of MTurk workers are from rural areas (levels 5 and 6 on the NCHS scale), while about 79% are from urban areas (levels 1 and 2). One avenue for future work is studying whether workers on MTurk are geographically representative. Second, more work is needed to understand what specific mechanisms allowed workers in our second study to perform equally well, regardless of where they were from. Third, additional research is needed to understand other salient dimensions of localness, and in what contexts they are advantageous.

Acknowledgements

We thank Rachel Kohler, Anne Hoang, Rifat Sabbir Mansur, Isaac Johnson, and the members of the Crowd Intelligence Lab for their time and invaluable feedback. We also thank the Humanitarian OpenStreetMap contributors and crowd workers on Amazon Mechanical Turk who enabled this research. This research was supported by NSF IIS-1527453 and IIS-1651969.

REFERENCES

1. Melanie Eckle and JoÃčo Porto de Albuquerque. 2015. Quality Assessment of Remote Mapping in OpenStreetMap for Disaster Management Purposes..

In *ISCRAM*. bibtex: eckle2015quality.

2. Mordechai Muki Haklay. 2013. Geography and HCI. (2013), 5–7. http://www.mendeley.com/research/geography-hci/

3. Debora Ingram and Sheila Franco. 2012. NCHS urban-rural classification scheme for counties. *Vital and health statistics. Series 2, Data evaluation and methods research* 154 (2012), 1–65. http://europepmc.org/abstract/med/22783637

4. Isaac L. Johnson, Yilun Lin, Toby Jia-Jun Li, Andrew Hall, Aaron Halfaker, Johannes SchÃűning, and Brent Hecht. 2016. Not at Home on the Range: Peer Production and the Urban/Rural Divide. ACM Press, 13–25. DOI: http://dx.doi.org/10.1145/2858036.2858123

5. Rachel Kohler, John Purviance, and Kurt Luther. 2017. Supporting Image Geolocation with Diagramming and Crowdsourcing. In *Fifth AAAI Conference on Human Computation and Crowdsourcing*. https://aaai.org/ocs/index.php/HCOMP/HCOMP17/paper/view/15812

6. Linna Li, Michael Goodchild, and Bo Xu. 2013. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science* 40, 2 (March 2013), 61–77. DOI: http://dx.doi.org/10.1080/15230406.2013.777139

7. Giovanni Quattrone, Licia Capra, and Pasquale De Meo. 2015. There's No Such Thing As the Perfect Map: Quantifying Bias in Spatial Crowd-sourcing Datasets. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. ACM, New York, NY, USA, 1021–1032. DOI: http://dx.doi.org/10.1145/2675133.2675235

8. Giovanni Quattrone, Afra Mashhadi, and Licia Capra. 2014. Mind the map: the impact of culture and economic affluence on crowd-mapping behaviours. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 934–944.

9. Shilad W. Sen, Heather Ford, David R. Musicant, Mark Graham, Oliver S.B. Keyes, and Brent Hecht. 2015. Barriers to the Localness of Volunteered Geographic Information. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 197–206. DOI: http://dx.doi.org/10.1145/2702123.2702170

10. Jacob Thebault-Spieker, Aaron Halfaker, Loren G Terveen, and Brent Hecht. 2018a. Distance and Attraction: Gravity Models for Geographic Content Production. In *Proceedings of the 36th Annual ACM Conference on Human Factors in Computing Systems*. 13.

11. Jacob Thebault-Spieker, Brent Hecht, and Loren Terveen. 2018b. Geographic Biases Are 'Born, Not Made': Exploring Contributors' Spatiotemporal Behavior in OpenStreetMap. In *Proceedings of the 2018 ACM Conference on Supporting Groupwork (GROUP '18)*. ACM, New York, NY, USA, 71–82. DOI: http://dx.doi.org/10.1145/3148330.3148350

12. W. R. Tobler. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography* 46 (June 1970), 234–240. DOI: http://dx.doi.org/10.2307/143141

13. Dennis Zielstra, Hartwig H. Hochmair, Pascal Neis, and Francesco Tonini. 2014. Areal Delineation of Home Regions from Contribution and Editing Patterns in OpenStreetMap. *ISPRS International Journal of Geo-Information* 3, 4 (Nov. 2014), 1211–1233. DOI: http://dx.doi.org/10.3390/ijgi3041211