The One and Many Maps: Participatory and Temporal Diversities in OpenStreetMap*

Tyng–Ruey Chuang†
Institute of Information Science
Academia Sinica
Taipei, Taiwan

Dong–Po Deng‡
Institute of Information Science
Academia Sinica
Taipei, Taiwan

Chun–Chen Hsu§
Institute of Information Science
Academia Sinica
Taipei, Taiwan

Rob Lemmens
Faculty of Geo–Information Science and Earth Observation (ITC)
University of Twente
Enschede, Netherlands

ABSTRACT

OpenStreetMap is an open and collaborative project with thousands of people contributing GPS traces and other data into the making of a global map of places and networks. It is open in the sense that everyone can contribute to the project, and results from the project are free for everyone to reuse. This is contrary to traditional cartography where often a central authority controls the making of the map and its release. Is OpenStreetMap more democratic, and in what sense? Is OpenStreetMap more relevant to the mass, and how can we judge?

We define and use several metrics to measure temporal properties of defined areas in OpenStreetMap, and to sample modes of participation in these areas. These metrics are used to graph the datasets representing the current OpenStreetMap so as to reveal unevenness in user participation and data temporality. We use the dataset about Taiwan as a test case to observe participatory and temporal diversities among different areas of Taiwan in OpenStreetMap.

Categories and Subject Descriptors

H.2.8 [**Database Applications**]: Spatial Database and GIS; H.3.5 [**Online Information System**]: Web-based services; H.5.3 [**Group and Organization Interfaces**]: Collaborative computing

General Terms

Design, Human Factors, Measurement

Keywords

Cartography, Metrics, OpenStreetMap, Participation, Temporality

*This research is supported in part by the National Science Council of Taiwan (grant no. NSC101-2119-M-001-004 and NSC102-2627-M-001-009).

†Tyng–Ruey Chuang is also with the Research Center for Information Technology Innovation, Academia Sinica.

‡Dong–Po Deng is also a PhD candidate at the Faculty of Geo–Information Science and Earth Observation (ITC), University of Twente.

§Chun–Chen Hsu is also a PhD candidate at the Department of Computer Science and Information Engineering, National Taiwan University.

ACM SIGSPATIAL GEOCROWD'13, Nov. 5, 2013. Orlando, FL, USA.
ACM ISBN 978-1-4503-2528-8 /13/11.

1. BACKGROUND AND MOTIVATION

The Web has changed the way of geospatial information production and sharing. Online mapping services enable people not only to consume but also produce geospatial information [15]. The term *Volunteered Geographic Information* (VGI) was coined to describe collaborative mapping activities as well as voluntary contributions of geographic data from the mappers [5]. OpenStreetMap (OSM) is one of the representative examples of VGI. OpenStreetMap is a wiki–style online mapping platform in which tens of thousands of people voluntarily contribute geospatial data into the making of a global map [7]. Its peer production model demonstrates that more and more mapping activities are done by the citizens. It represents the success of a collective form of geospatial content creation.

Collaborative geospatial content creation is not a new concept in the field of geographic information. The idea can be tracked back to *Public Participation Geographic Information Systems* (PPGIS) in which non–governmental organizations, grassroots groups, and community–based organizations use GIS to broaden public involvement in policymaking [14]. With the use of open source software and the facilitation of an online framework for collecting geospatial data, OpenStreetMap, however, created a new paradigm of collaborative mapping. OpenStreetMap offers a venue for rapid convergence of information technologies, geospatial information, and user communities. As witnessed by the rapid increase of contributors and their contributions, OpenStreetMap has shown the promise of geospatial data collaboration and sharing. We view the characteristics of data collaboration in OpenStreetMap as the subjects of VGI research.

The geographic information science community has started to study OpenStreetMap in particular about user motiva-

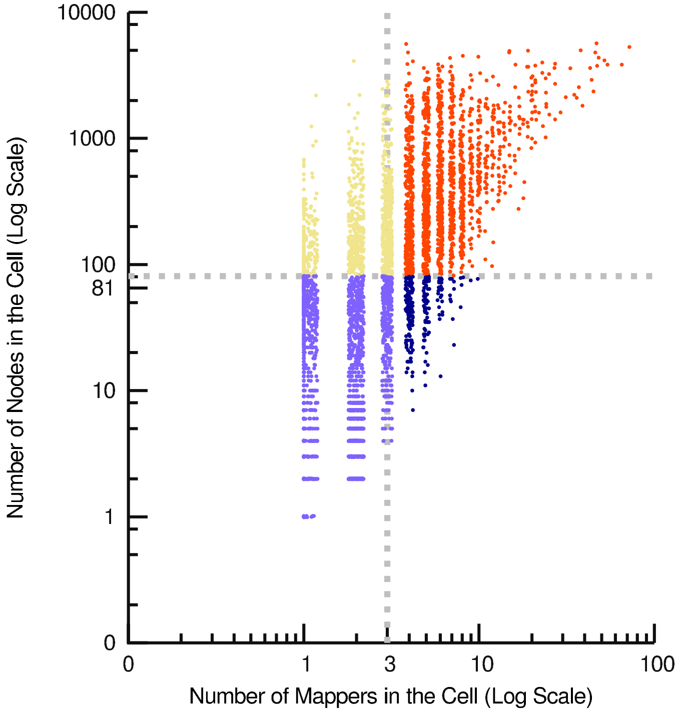

Figure 1: Distribution of the cells by both mapper count and node count.

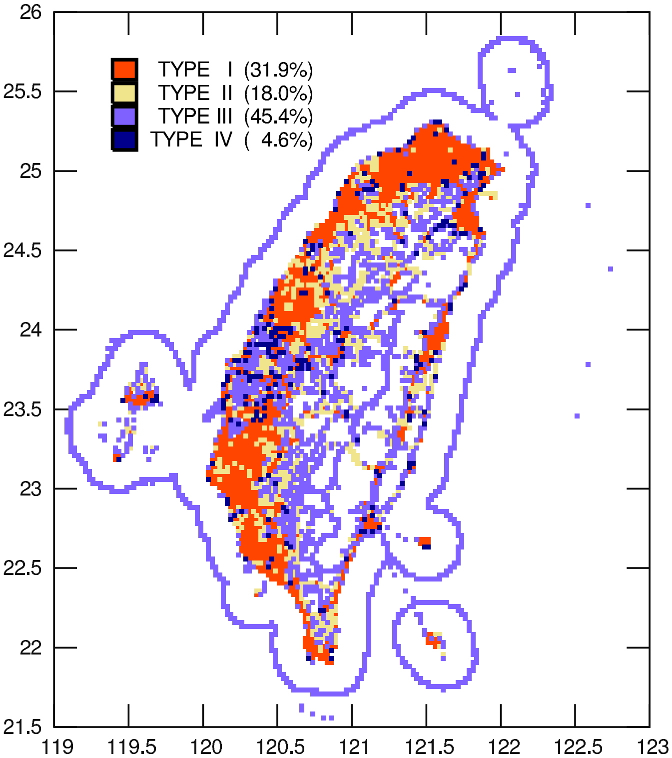

Figure 2: Mapping the cells in Taiwan by their types. (*c.f.* Figure 1)

tions for participating in the OpenStreetMap mapping activities [1, 2, 10, 11, 12]. It has been observed that the instrumentality of local knowledge (about the areas where the contributors have lived or traveled to) is a key to invoke them to map, especially when they see the areas they care about are missing or erroneously mapped [1]. The current state of OpenStreetMap actually is an assembly of many edits and updates over a period of time. Every edit or update should be a meaningful unit in the understanding of data collaboration activities in OpenStreetMap.

In this paper we look for ways to systematically and efficiently discover data collaboration patterns and diversities in OpenStreetMap. It is an initial study of the OpenStreetMap dataset (at least about the part of Taiwan) by developing a set of metrics to summarize user participation and spatiotemporal variations of updates in defined areas of OpenStreetMap. By exploring the different manners in which data are added to the OpenStreetMap dataset, and to reveal variations by visualization, we hope to see the OpenStreetMap not as one collective map but as many overlapping maps concurrently in the making where each has its own characteristics.

2. DATA, METRICS, AND GRAPHING

'Node', 'Way', and 'Relation' are the three fundamental object types in the OpenStreetMap data model. Both 'Node' and 'Way' are used to define geometry objects, while 'Relation' is used to define geographic or logical relations between the geometry objects. A 'Node' is a geospatial point in the form of latitude and longitude coordinates. A 'Way' consists of an ordered list of nodes. If the first node and the last node are the same in the ordered list, the way can be an 'Area' or 'Closed Way'. Each geometry object consists of a version number, a changeset number, an ID, an account name of the contributor, the date when it was updated or created, and a 'Tag'. A 'Tag' is used to describe an attribute of the geometry object. For example, a shop can be a 'Tag' of a 'Node', a road is a 'Tag' of a 'Way', and a building can be a 'Tag' of 'Area'. Moreover, multiple tags can be attached to a single geometry object. A contributor in OpenStreetMap is called a 'mapper'. A changeset is 'a group of edits' made within a certain time frame by one mapper. An edit can be a creation or an update of a node or a way. The same mapper can have multiple active changesets at the same time. Once a geometry object (node or way) is updated, it should be annotated with the new version number. The version numbers are actually used to control the updates and creations of edits.

2.1 Metrics

Let c be a geographic region — a *cell* — and let D_c be the set of all the *surviving nodes* in the current OpenStreetMap dataset which is used to render a map of c. For now, we analyze only `node` elements in the OpenStreetMap dataset; `way` and `relation` elements are not considered. As for *surviving*, we mean nodes that are actually in the snapshot of dataset that is used to produce the current map. It is noted there are *historical* data items that are no longer in the snapshot, and are no longer used to render maps. Historical nodes are not considered in this study.

A node element has, among others, the following attributes: `id` (node id), `timestamp`, `uid` (user id), and `lat` and `lon` (latitude and longitude of the node). A data item has version information and is associated to a changeset. Neither do we

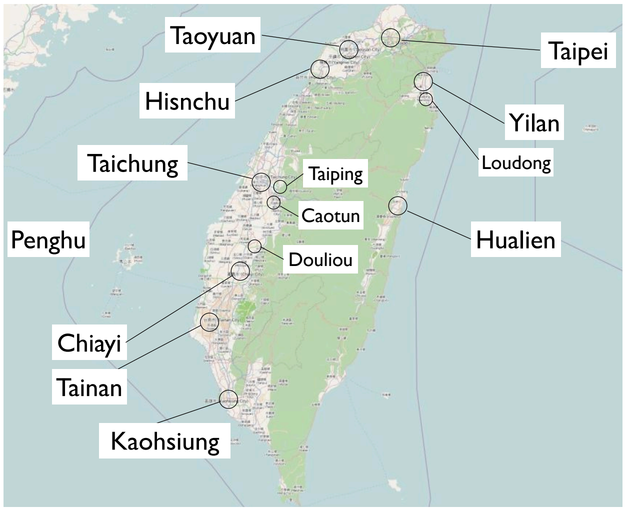

Figure 3: Cities in Taiwan.

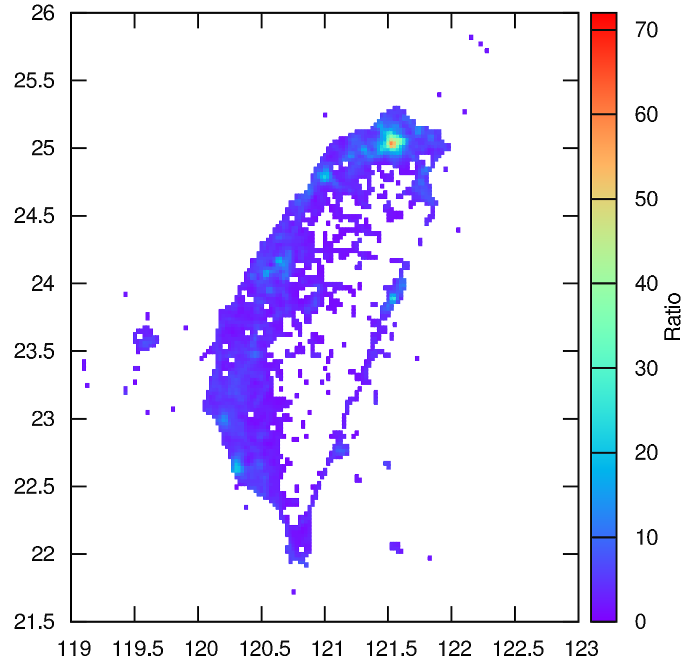

Figure 4: Spatial distribution of mappers over area.

consider such information. For a cell c with a definite boundary, one can compute the set D_c from the OpenStreetMap dataset (as one can calculate whether a node is positioned within c).

To measure participatory and temporal differences among cells in OpenStreetMap, we define the following functions on D_c. When the context is clear, we omit the subscript c and simply write $D = \{d_0, d_1, \ldots, \ldots, d_{n-1}\}$ where d_i, $0 \leq i \leq n-1$, is a node in the cell, and n is the total number of nodes. We use m to denote the number of mappers who contribute to the nodes in D. For convenience, we fix a date (say, today) as the day of reference when comparing the ages (in day) among all the nodes. For nodes whose `timestamp` values fall into the range of the fixed date, they have age 0. For nodes that are time–stamped the day before it, they have age 1, and so on.

For a node d_i, we write

$$d_i = (k_i, t_i, u_i, p_i)$$

where k_i is the node id of d_i, t_i the age, u_i the user id of its contributor, and p_i the position (*i.e.*, the pair of its `lat` and `lon` values). Note that, by definition, geographically p_i is within the boundary of c for all $0 \leq i \leq n-1$. Without loss of generality, we require

$$t_0 \geq t_1 \geq t_2 \geq \ldots \geq t_{n-1}$$

That is, the nodes are sorted by their ages with the newest one being d_{n-1} and the oldest one being d_0. Nodes of the same age appear consecutively but their ordering does not matter to us. We use $d = (d_0, d_1, d_2, \ldots, d_{n-1})$ to denote a sequence when is clear in the context. Likewise, we write

$$\begin{aligned} k &= (k_0, k_1, k_2, \ldots, k_{n-1}), \\ t &= (t_0, t_1, t_2, \ldots, t_{n-1}), \\ u &= (u_0, u_1, u_2, \ldots, u_{n-1}), \\ p &= (p_0, p_1, p_2, \ldots, p_{n-1}). \end{aligned}$$

Note that a sequence can be a bag in which repeating elements may occur (for example in the case of t and u).

Starting from the above, we then compute, compare, and visualize various characteristics of OpenStreetMap cells. A cell can be any geographical region such as an administrative district or a bounding box. For simplicity and for easy comparison, however, we divide the globe surface into "rectangular" tiles by their geographic coordinates. As an example, in our study of Taiwan in OpenStreetMap in Section 3, each cell is a tile that measures 0.025 degree apart in parallel (latitude and longitude) which result in approximately a rectangle of 2.5 *km* by 2.5 *km*.

In general, we use $area_c$ to denote the area covered by a cell c, and we use pop_c for the people population in region c. When it is clear in context, we omit the subscript and simply write *area* and *pop*.

We define several metrics to measure the various characteristics of a cell. These metrics present some aspects of the mapping activities in a cell as summarized from attributes of the nodes in the cell. As these OpenStreetMap cells can be colored based on their metrics, this results in a cartography of mapping activities based on the OpenStreetMap dataset. For any two metrics, we too can graph all cells on a plane by the pairs of values respectively from the two metrics. This may give insights into possible patterns in the mapping activities.

In the following, we define metrics for measuring node and mapper density, and for measuring node age and temporality. We also illustrate the way we will use to graph cells on a plane by two metrics.

2.1.1 Node and Mapper Density

The following measures the densities of nodes, as well as those of their contributors, *i.e.*, the mappers.

node over area ratio — $n/area$

mapper over area ratio — $m/area$

node over population ratio — n/pop

mapper over population ratio — m/pop

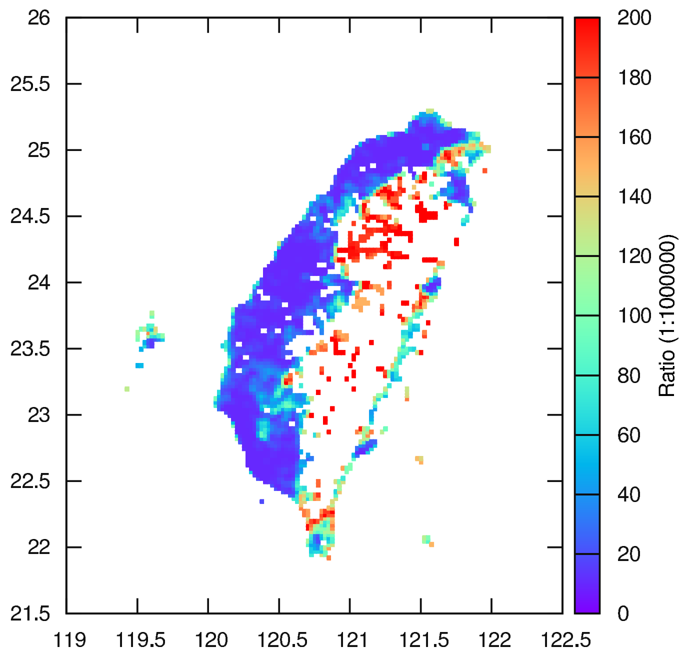

Figure 5: Spatial distribution of mappers over population.

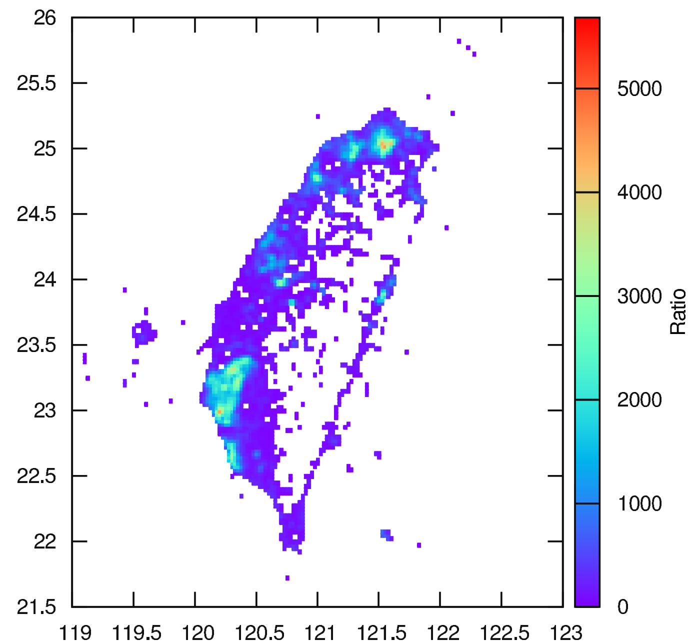

Figure 6: Spatial distribution of nodes over area.

These ratios, when computed for all cells, shall show how well the OpenStreetMap nodes and mappers are covering a region. We are especially interested in the mapper–population ratios to learn about places where OpenStreetMap mappers are under/over–represented. Likewise, the node–population ratio shall be a good index on the coverage of OpenStreetMap data when compared to the volume of man–made geospatial features in a region (assuming population count relates linearly to feature volume). Note that when all cells are of the same area size, node–area ratio and mapper–area ratio can be replaced by node count and mapper count.

2.1.2 Node Age and Temporality

We are interested in the ages of the nodes in a cell, as well as the temporality about the nodes as they are added to the cell. These help answer these questions: How old is the map, as judged by the ages of the nodes in the cells? How about temporal consistence (or the lack of it) in the map?

Recall the following auxiliary functions for a sequence s

$$
\begin{aligned}
\min s &= s_k, \text{where } s_k \leq s_i \text{ for all } s_i \in s \\
\max s &= s_k, \text{where } s_k \geq s_i \text{ for all } s_i \in s \\
\bar{s} &= \frac{\sum_i s_i}{n} \\
cv(s) &= \frac{\sqrt{\sum_i (s_i - \bar{s})^2}}{\bar{s}}
\end{aligned}
$$

where $\min s$ is the minimum of s, $\max s$ the maximum of s, $\bar{s}$ the average of s, and $cv(s)$ the coefficient of variation for elements in s.

The 4-tuple $\langle \min t, \max t, \bar{t}, cv(t) \rangle$ measures the age characteristics of the nodes in a cell. That is,

age of the newest node — $\min s = t_{n-1}$

age of the oldest node — $\max s = t_0$

average age of the nodes — $\bar{t}$

variance of the node ages — $cv(t)$

We also study the modes of mapping. Are the nodes in a cell added in a "burst mode" (because of a mapping party, perhaps), or are they added in an evenly "spread out" manner over time? To answer these questions, age characteristics alone are not sufficient. We define a sequence $g = (g_0, g_1, \ldots, g_{n-2})$ to measure the gaps between any two consecutive elements in t. That is, $g_i = t_{i+1} - t_i$ which is the gap in days between the dates when the two nodes d_{i+1} and d_i were added into the cell.

The 4-tuple $\langle \min g, \max g, \bar{g}, cv(g) \rangle$ measures the day–gap characteristics of node–addition in a cell. That is,

minimal no. of days between two additions — $\min g$

maximal no. of days between two additions — $\max g$

average no. of days between two additions — $\bar{g}$

variance of the gaps between additions — $cv(g)$

A cell with a large value of $cv(g)$ indicates the nodes were added in uneven intervals. A small $\max g$ value indicates the nodes were added into the cell periodically as the interval between any two additions is short. In such a case, one may say the cell is well taken care of as it is continuously updated. On the other hand, a large $\max g$ value indicates there exists a large time gap between the dates of two additions. This may be a sign of temporal unevenness among the mapping activities: some mapping was done long before/after the others.

2.2 Graphing Cells by Two Metrics

As multiple metrics are in use, a cell can be measured in two metrics and the two results compared. Often we will compare the two sets of measurement over all cells to see if

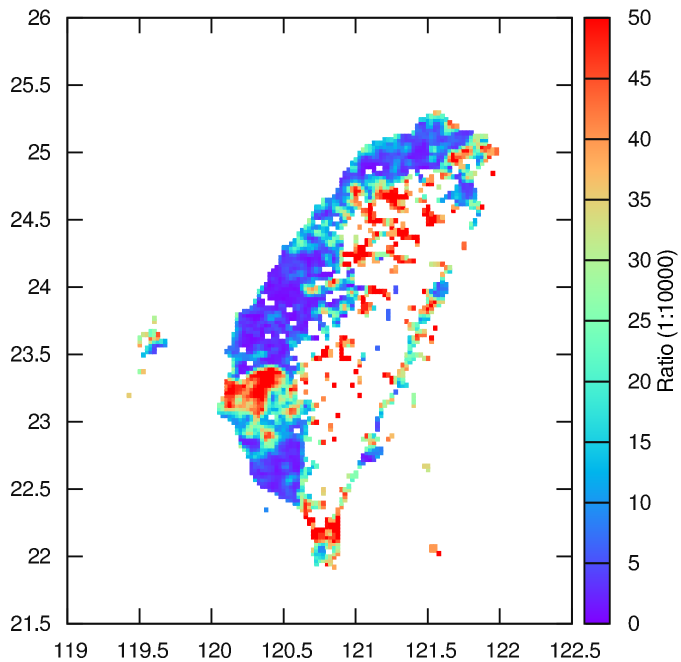

Figure 7: Spatial distribution of nodes over population

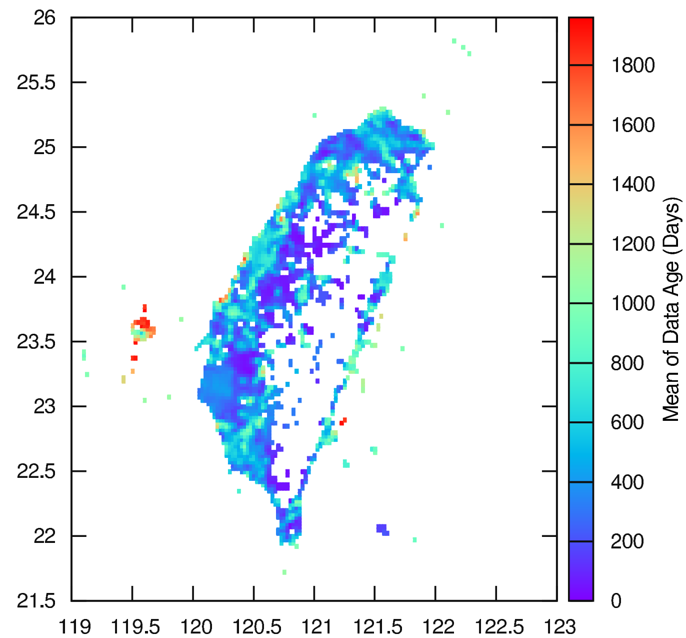

Figure 8: Spatial distribution of average node age.

there are patterns. We use the following example to show a simple way to graph all cells by two metrics, and to map the cells accordingly.

Recall n is the number of nodes in a cell, and m the number of mappers. For a cell, we view the pair (n, m) as a point in the plane, and we do this for all the cells. This produces a visualization of the distribution of the cells by both mapper count and node count. One caveat is that more than one cells may have the same measurement. For example, there are many cells in which each has exactly one mapper contributing only one node. They all fall into $(1, 1)$. To get a proper visualization, we perturb the value (n, m) by two random variables $-1 < a, b < 1$, and put it at point $(n+a, m+b)$ on the plan. Figure 1 is such a visualization of all cells in Taiwan by both mapper count and node count.

Figure 1 uses log scale on both axes. The two gray lines show the two medians: The mapper counts have a median of 3 while the median for node counts is 81. The two lines separate the cells on the plan into four quadrants. The upper right quadrant contains cells in which each has more mappers than the median mapper count, as well as more nodes than the median node count. If we view the four quadrants as four types, we can color the cells by their types and produce a color map of Taiwan, see Figure 2.

From Figures 1 and 2, several observations can be made. We observe that there are some cells (18% of the total) with 3 mappers or less, but each has more nodes than the median node count. Many of these cells have node counts in the thousands which are comparable to, or more than, the counts in cells with much more mappers. A general observation is that, however, the cells with more mappers most of the case do produce more nodes. The result is similar to the findings reported in the literature [3, 8].

3. TAIWAN IN OpenStreetMap

We observe the part of Taiwan in the current OpenStreetMap dataset using the metrics and graphing techniques described in Section 2. The OpenStreetMap mapping activities in Taiwan started in 2007. Up to now, there are about 400 mappers taking part in mapping Taiwan on and for OpenStreetMap. Most mappers seem to concentrate on the urban areas. Apparently Taiwan urban areas carry more geospatial information than what the rural and natural areas do in the current OpenStreetMap. This section sets to take a closer look at the actual dataset. We hope to reveal certain patterns and answer some questions about mapping activities in Taiwan. In the following discussion, several cities in Taiwan will be mentioned. Please consult Figure 3 for the locations of these cities.

3.1 Distributions of Nodes and Mappers

Figure 4 illustrates Taiwan by the ratios of mapper number over area size. It is obvious that in the urban areas the mapper density is higher. Taipei has the highest density of mappers. Other urban areas such as Taoyuan, Hsinchu, Taichung, Tainan, Kaohsiung, and Hualien have mapper densities higher than those of other areas. Figure 5 illustrates the ratios of mapper number over population size. When compared to Figure 4, ones may be surprised to find out some of the high ratio areas occur at the natural and mountain areas. Although these areas have smaller populations, but when compared to the urban areas, they have relatively more mappers. Many of these areas are scenic areas (forest parks or national parks). Most of the mappers probably are tourists. Mapping these places is part of their trips.

Figure 6 is a map of Taiwan showing the ratios of node number over area size, and Figure 7 is a map on the ratios of node number over population size. Similar to Figure 4, the urban areas have higher node density (over area size) as shown in Figure 6. A large area in the south, roughly covering the Tainan City and its surrounding area, have very high

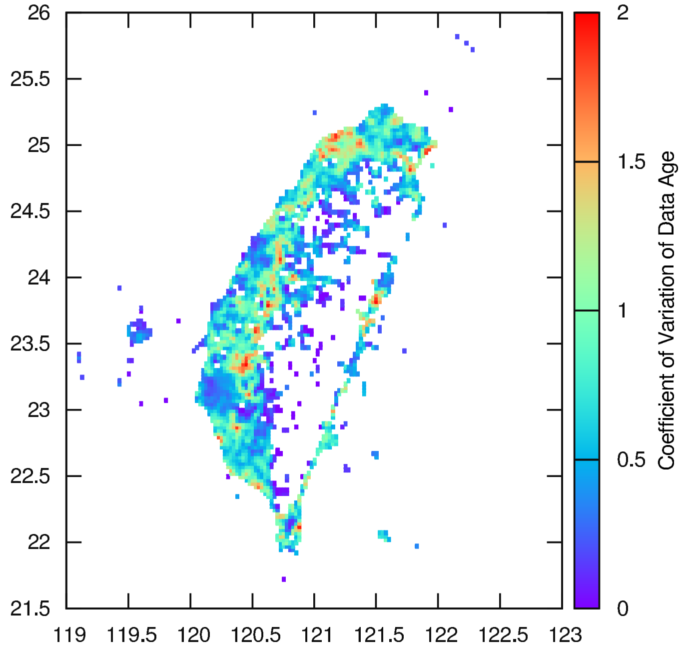

Figure 9: Spatial distribution of the variance of node age.

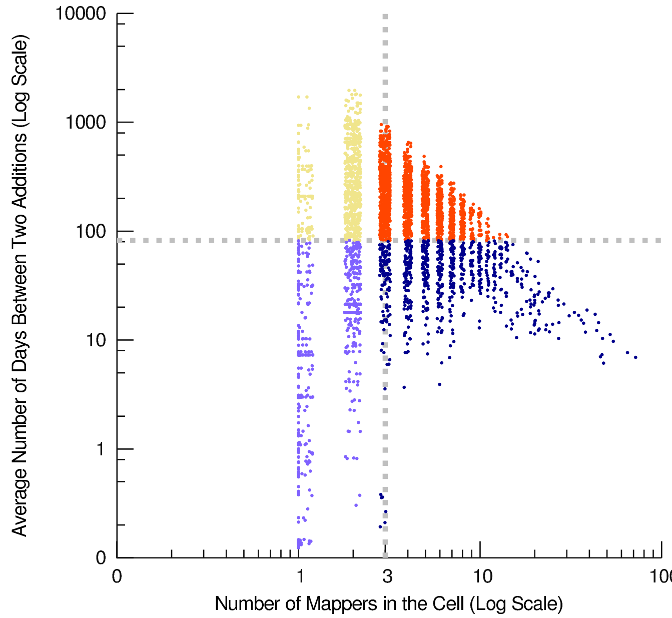

Figure 10: Distribution of the cells by both mapper count and average time gap between two additions.

node density. This is not an entirely urban area. This area also stands out in Figure 7, showing the area to have the highest node count over population size ratio. The nodes there are even denser than those in Taipei when they are normalized by population sizes. This can be re–stated as the people at Tainan has more detailed map (in the Tainan area) than the people in Taipei have (in the Taipei area). It turns out there is a persistent and productive mapper who has been adding nodes to the Tainan City area. The OpenStreetMap in this area has been mapped in detail almost by him/her alone.

3.2 Map Age and Update Interval

OpenStreetMap is a community–based mapping project involving people of different backgrounds. One participates in the mapping at a time that is convenient to him/her. Every mapper can modify and delete other mappers' contributions. The map as presented by the current OpenStreetMap dataset is the result of many revisions and modifications. Some parts of the map may contain nodes that were put there long ago, some are newly added, and all are subject to modifications in the future. The age of a node is the number of days it has survived in the dataset. A young node means it is newly added or modified. An old node probably means that it has been informative or complete for a long time so nobody has the need to modify or delete it. But old nodes could also survive just because the areas they are in are less accessible to mappers. These old nodes can even be incorrect or incomplete. Naturally there is a need to know which areas in the map are "older" than the others hence are in need of attention. Also, it is helpful to know areas where the nodes are updated in an unevenly fashion (which implies geospatial information in these areas may not be temporally consistent).

Figure 8 shows the spatial distribution of average node age. In average, the cells with the oldest nodes occur at Pinghu Islands west of the main Taiwan island. Most of the youngest nodes are located in rural, natural, or mountain areas. The nodes in the urban areas are relative older than the nodes in suburban areas. The spatial patterns of average node age reveal that mapping activities are moving from urban areas to suburban areas, and even to natural and mountain areas. Figure 9 is the spatial distribution of the variance of node age. The large area of the highest age variation appears in Taoyuan which is in Northern Taiwan. Another area having high variance of node age is a narrow belt from Taiping, Caotun, Douliou, to Chiayi, which are rural areas. Some places in the east coast of Taiwan, for example Loudong and Hualien, recently have new mapping activities so the variance of node age is high in these places.

Two cells with the same average node age, say x days, can be very different in how the nodes were added. One cell may have all its nodes added at once x days ago. The other may have nodes being added continuously, and the average of their ages happens to be x. The sequence of gaps, in days, between every two consecutive node additions, in our view, say more about the ways a cell has been maintained. In Figure 10, we graph the distribution of all the cells both by the number of mappers in the cell, and by the average length, in days, of the gaps between any two consecutive node additions. The cells in the first quadrant (Type I) have a large number of mappers, and on average a long period between two node updates. The cells with a small number of mappers and long update period in the second quadrant (Type II). The cells with a small number of mappers and a short update period is in the third quadrant (Type III), and the cells with a large number of mappers and a short update period in the forth quadrant (Type IV). The spatial distribution of the cells by the four types is illustrated in Figure 11 which we find to be informative. Taipei, Taichung, and Kaohsiung — the three major cities in Taiwan — on average have older

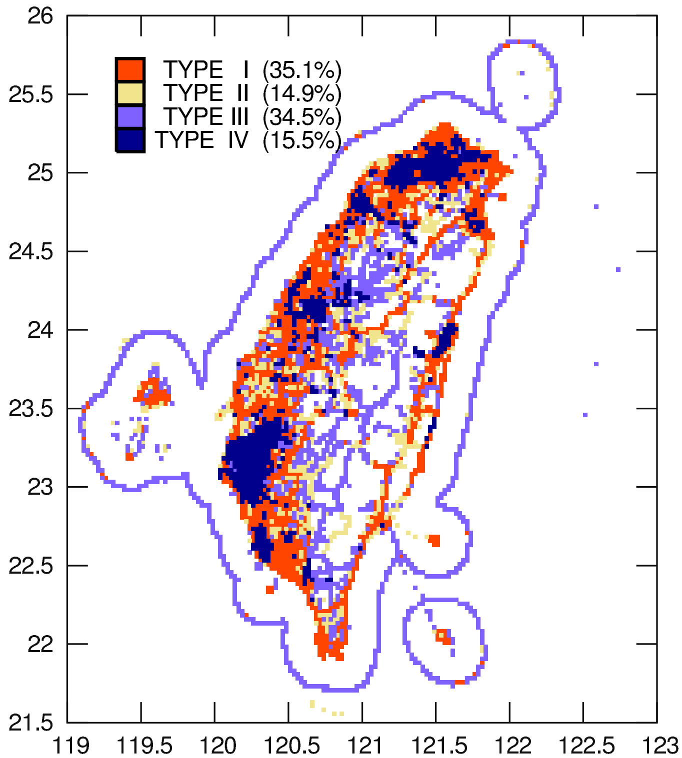

Figure 11: Spatial distribution of the cells by their types. (*c.f.* Figure 10)

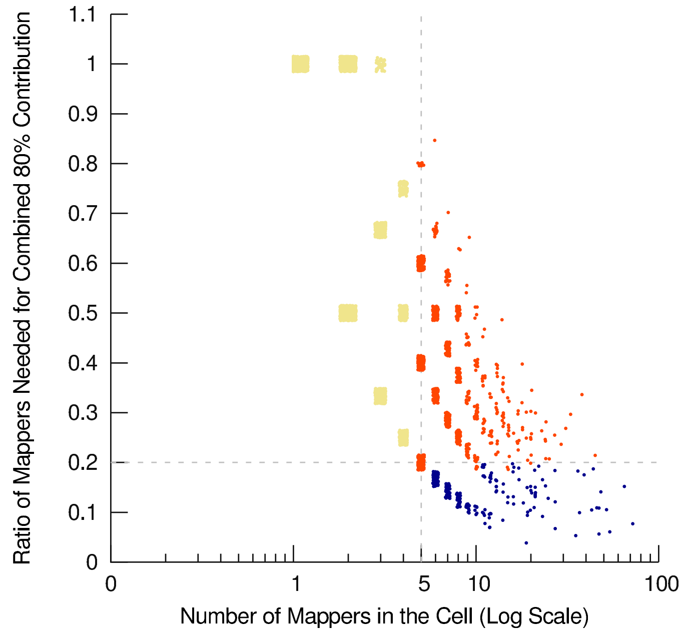

Figure 12: Distribution of the cells by both mapper count and ratio of mappers needed for combined 80% node contribution.

nodes (*c.f.* Figure 8) but they also have shorter gaps between updates. It appears many mappers are still updating the areas. The three cities are mainly consisted of Type IV cells. Interestingly, major road networks consist of Type I cells. That is, many people contribute to the mapping of highways, but once they are done they are infrequently updated (hence long update intervals). On the other hand, administrative boundaries as well as sea links (and the artificial country lines) are TYPE III nodes: Only few people care for them, and once the lines were drew they remain unchanged (hence almost no day between updates).

3.3 The 80/20 Hypothesis

It is often said 80% of the work is done by 20% of the people. Is this true for the contributions to the OpenStreetMap dataset? We put this hypothesis to test by looking into all the cells that constitutes the Taiwan portion of the OpenStreetMap. We do the following. We first sort the mappers in a cell by their contributions (in node count) to the cell. We then add up individual node contributions, from the top mapper to the bottom mapper, until the accumulation just reach 80% of the total node count of the cell. We now know the minimal number of mappers in order to achieve a combined 80% node contribution. We then calculate the ratio of this number to the mapper count of the cell.

Figure 12 graphs the distribution of the cells both by the number of mappers in the cell, and by the above minimal ratio of mappers in the cell in order to achieve a combined 80% contribution. Instead of using the median ratio, we now use the "20%" line to partition the plane. Note that for "less than or equal to 20% of the mappers" to be meaningful, *i. e.* to avoid counting a fraction of one mapper, a cell must have 5 mappers or more. Therefore, the "5 persons" line is used in combination with the "20%" line to partition the plane into 4 quadrants. Note that there is no cell in the 3rd quadrant (which is by design).

The first quadrant (Type I) contains cells where the 80/20 hypothesis is not valid. The second quadrant (Type II) contains cells where each has less than 5 mappers hence the 80/20 rule is not applicable. The fourth quadrant (Type III) contains cells where the 80/20 hypothesis is valid. We may say the mapping in Type I cells is more democratic than that in Type III cells in terms of contribution inclusion. Figure 13 is the spatial distribution of the cells by their types. The 80/20 hypothesis is more likely to be valid in the urban areas, perhaps because there are more mappers per cell and top mappers contribute a lot more. Also note that Type III cells only constitute 5.2% in the entire collection of cells.

4. RELATED AND FUTURE WORK

While the number of contributors as well as their combined contribution to OpenStreetMap is impressive, the quality of user–contributed data is often considered an issue. Previous investigations into the data quality issues of OpenStreetMap have shown that the OpenStreetMap dataset can be fairly accurate, and is mostly comparable to commercial datasets at least in urban areas [3, 6, 8, 9]. Researchers had also developed visual analytics to gain insights into the spatial diversity of OpenStreetMap datasets, *e. g.* to see whether users in different countries would exhibit distinct mapping activities and habits [13]. These visualization tools can provide valuable information when improving the data quality of OpenStreetMap. Neis and Zipf identified active mappers and casual mappers by examining the quantities of their contributions in OpenStreetMap [12]. Their results showed that the contribution patterns in OpenStreetMap

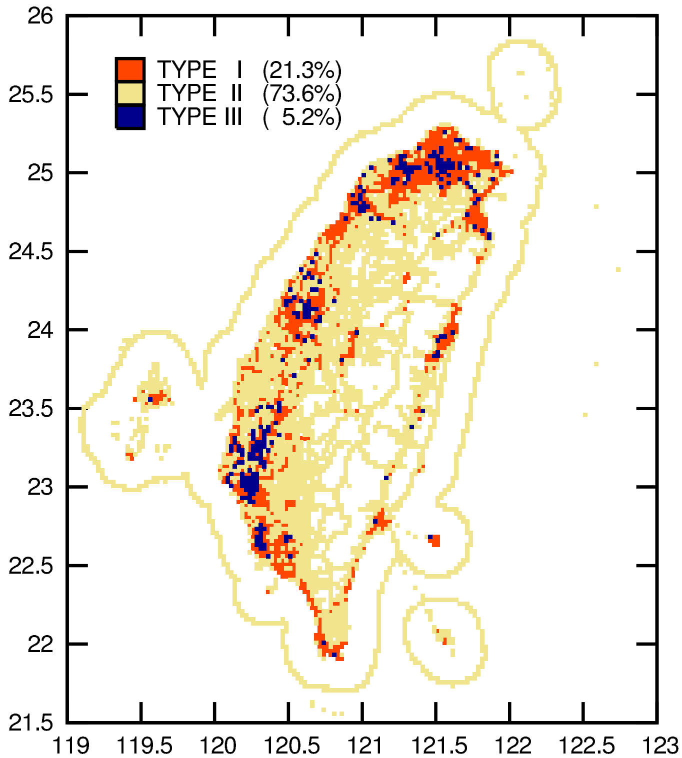

Figure 13: Spatial distribution of the cells by their types. (*c.f.* Figure 12)

corresponded closely to the participation patterns observed in other community–based projects. Mooney and Corcoran examined directly the characteristics of "heavily edited" objects in OpenStreetMap of UK, and they considered these characteristics might be developed as data quality indicators for OpenStreetMap in the future [10]. In general, geospatial data productions can be combinations of crowdsourced data creation models and traditional data production techniques [4]. Data quality issues in hybrid geospatial production models remain an interesting research subject.

This paper is a preliminary study in two sense: We only analyze the Taiwan part of OpenStreetMap, and we only analyze the cells independently (though spatial distribution is visualized and discussed). Because of the time constraint, we have not looked into other geographical areas in OpenStreetMap. Also, as one mapper may contribute to multiple cells, we ought to look into mapping activities across multiple cells. We intend to pursue these directions in the future.

The programs we use to analyze the data are in their early stage of development, and the way we prepare the data for analysis is rather ad hoc. Our tools cannot be easily reused. We are currently considering how better to structure the programs so that they can be easily ported and reused. Metrics–based analysis tools like these can be very useful in improving the data quality in OpenStreetMap as they help discover areas where there is participatory or temporal unevenness in the map making process itself.

5. REFERENCES

[1] N. R. Budhathoki. *Participants' Motivations to Contribute Geographic Information in an Online Community*. PhD thesis, University of Illinois at Urbana-Champaign, 2010.

[2] N. R. Budhathoki and C. Haythornthwaite. Motivation for open collaboration: Crowd and community models and the case of OpenStreetMap. *American Behavioral Scientist*, 57(5):548–575, 2012.

[3] J.-F. Girres and G. Touya. Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS*, 14(4):435–459, 2010.

[4] M. Goodchild. Neogeography and the nature of geographic expertise. *Journal of Location Based Services*, 3(2):82–96, 2009.

[5] M. F. Goodchild. Citizens as sensors: The world of volunteered geography. *GeoJournal*, 69(4):211–221, 2007.

[6] M. Haklay. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design*, 37(4):682–703, 2010.

[7] M. Haklay and P. Weber. OpenStreetMap: User–generated street maps. *IEEE Pervasive Computing*, 7(4):12–18, 2008.

[8] M. M. Haklay, S. Basiouka, V. Antoniou, and A. Ather. How many volunteers does it take to map an area well? The validity of Linus' law to volunteered geographic information. *The Cartographic Journal*, 47(4):491–507, 2010.

[9] I. Ludwig, A. Voss, and M. Krause-Traudes. A comparison of the street networks of Navteq and OpenStreetMap in Germany. In S. Geertman, W. Reinhardt, and F. Toppen, editors, *Advancing Geoinformation Science for a Changing World*. Springer Berlin Heidelberg, 2011.

[10] P. Mooney and P. Corcoran. Characteristics of heavily edited objects in OpenStreetMap. *Future Internet*, 4:285–305, 2012.

[11] P. Neis, D. Zielstra, and A. Zipf. Comparison of volunteered geographic information data contributions and community development for selected world regions. *Future Internet*, 5:282–300, 2013.

[12] P. Neis and A. Zipf. Analyzing the contributor activity of a volunteered geographic information project — the case of OpenStreetMap. *ISPRS International Journal of Geo–Information*, 1(2):146–165, 2012.

[13] O. Roick, J. Hagenauer, and A. Zipf. Osmatrix — gridbased analysis and visualization of OpenStreetMap. In *Proceedings of State of the Map EU 2011*, Vienna, Austria, 2011.

[14] R. Sieber. Public participation geographic information systems: A literature review and framework. *Annals of the Association of American Geographers*, 96(3):491–507, 2006.

[15] A. Turner. *Introduction to Neogeography*. O'Reilly Media, Inc., 2006.