A Tool for Matching Crowd-sourced and Authoritative Geospatial Data

Heshan Du, Natasha Alechina, Glen Hart
Computer Science
University of Nottingham
Nottingham, UK

Michael Jackson
Nottingham Geospatial Institute
University of Nottingham
Nottingham, UK

Abstract— **A software tool MatchMaps was designed to facilitate matching between geospatial features held in different datasets, especially an authoritative dataset and a crowd-sourced dataset. MatchMaps is not fully automatic; it requires human interaction to resolve problematic matching cases. Previous studies have shown that this approach results in higher precision and recall than those achieved by fully automatic tools. In this study, we aim to evaluate MatchMaps with respect to the amount of human effort required for matching, and compare it with a fully manual matching process.**

***Keywords:** geospatial data matching; crowd-sourced geospatial data; validation; human effort; user evaluation*

I. INTRODUCTION

Maps, whether digital or paper-based, are a common feature of our daily life. They typically provide a two-dimensional representation of geographic features, such as roads, rivers, buildings, places, etc., in the real world (i.e. a topographic base) over which other "thematic" information may be displayed such as density of population or crime statistics. The information represented provides both an indication of where on the earth's surface an object of interest is (i.e. its geometry) and lexical information on what that geometry represents (e.g. a road and its name such as "High Street"). There will also typically be additional "meta-data" which may provide details such as when the information was recorded and the quality of the data.

Where the map data has been surveyed and classified using formal quality assurance procedures, for example by a national mapping agency, we refer to the resulting data as "authoritative". Traditionally, because mapping required the use of specialist surveying tools and an associated high-level of expertise, most national level mapping was carried-out by governmental agencies or specialist mapping companies. Maps produced by the general public were mostly of the "sketch map" form which indicated where key features were in relative terms but which could not be relied upon for precise location, completeness or consistency.

This situation has radically changed in the last few years by a number of technological developments and changes by governments in the release of associated data (e.g. precise Global Navigation Satellite System data and satellite imagery). Perhaps the most important of these developments is the mobile phone capable of accurately recording its position combined with the use of simple-to-use applications being able to delimit physical features on the ground and tag the resulting geometry with information describing the nature, purpose and use of that feature. This "crowd-sourced data" may be actively collected as a volunteer activity by citizens (Citizen Science, Volunteered Geographic Information) [1, 2] or passively acquired as a bi-product of an application the main purpose of which is something else. Such data can hugely enrich the information content in the traditional map. However, currently, compared to authoritative data, crowd-sourced data is usually less geometrically accurate, less formally structured and lacks the associated metadata that allows it to be used in situations where commercial, policy or life-critical use is involved. Despite this, crowd-sourced data offers great potential as it often contains richer user-based information, can reflect real world changes (e.g. new constructions of buildings) more quickly, and has a much lower acquisition cost [3]. With the rapid development of crowd-sourced data in recent years, it has become increasingly desirable to use authoritative and crowd-sourced geospatial data synergistically, trying to take the best out of both.

However, achieving the conflation of authoritative and crowd-sourced data is far from straightforward. In different geospatial datasets, different terminologies or vocabularies are often used to describe spatial features. For example, the same restaurant can be classified as *Restaurant* in one dataset, whilst as *Place_to_Eat* in another database and simply as a brand-name in others (e.g. McDonald's). An identically spelt word, even within a single language, can often have many different meanings. Whilst an authoritative dataset will have a defined taxonomy or ontology where a word should have a precise definition, the "crowd" may not follow such rules and may use several descriptions for a common object some of which may be local vernacular terms. For example, the word *College* may mean an institution within a university in one dataset, refer to a government secondary school in another and a private language training establishment in a third. Other terms may be used inconsistently, for instance, one person may include McDonald's within the category restaurant whilst others may not.

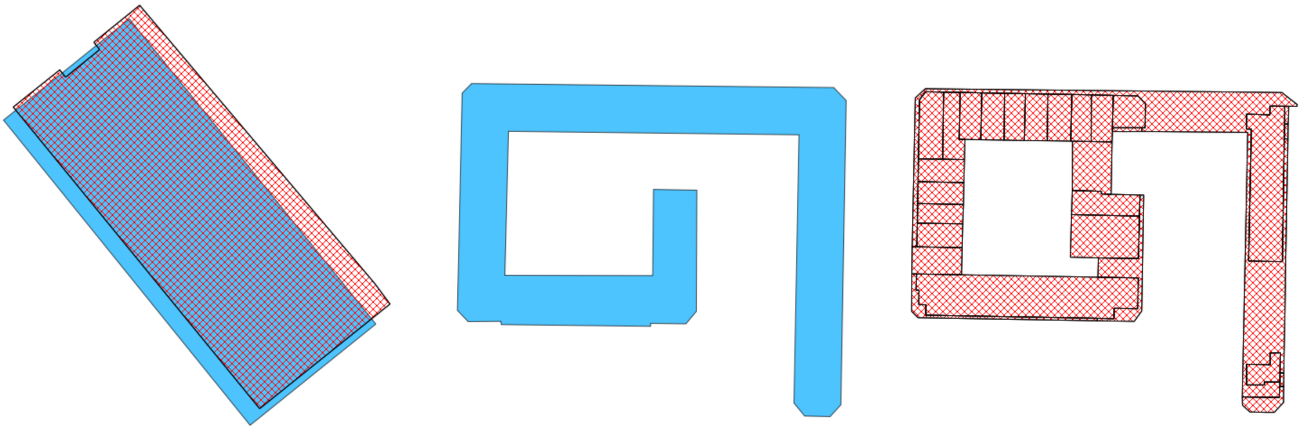

Figure 1. **a.** Boots in OSGB (stippled) data and OSM (solid) data; **b.** Wyndham Court in OSM data; **c.** Wyndham Court in OSGB data

For the same geographic area or the same set of spatial features, different geospatial data sources will have different representative geometries. Features may be represented in one dataset, but not in the other. The scale or accuracy of the geometry capture may vary. Even where the same precision of measurement is adopted, different points may be captured to represent the boundary of a feature so that two independently captured representations of a single object will always differ in some respect. Fig. 1 shows two locations, Boots (the Chemists Ltd) and Wyndham Court in Southampton UK, represented in Ordnance Survey of Great Britain (OSGB) [4], the Great Britain's national mapping agency, and OpenStreetMap (OSM) [5] datasets. The position and shape of Boots are represented differently in OSGB (stippled) data and OSM (solid) data (Fig. 1.a). Wyndham Court is represented as a whole in OSM (Fig. 1.b), and as several flats and shops in OSGB (Fig. 1.c).

The research challenge to be addressed if authoritative and crowd-sourced data are to be used together as a single conflated data source is to resolve the disparate geometries and lexical descriptors of a single feature into a single unambiguous object description. Therefore, in order to maximize the synergistic use of authoritative and crowd-sourced geospatial datasets, it is essential to establish correspondences (matches) between them. The matches between spatial features or geometries are of the following four types.

1. A *sameAs* match states that two spatial features represent the same real world object.
2. A *partOf* match states that one spatial feature represents a real world object which is part of what the other spatial feature refers to.
3. A *BufferedEqual* match states that two geometries probably describe the same real word location.
4. A *BufferedPartOf* match states that one geometry describes a real word location which is probably part of another location which is described by the other geometry.

A *sameAs* or *partOf* match is between two spatial features. It is generated using both location (e.g. geometries) and lexical information (e.g. names). *BufferedEqual* and *BufferedPartOf* matches are generated just for geometries, without considering any lexical information.

When generating matches, it is important to ensure their correctness, since the existence of wrong matches can result in misleading information and misuse of geospatial data [6]. For example, if a clinic is stated *sameAs* a bank, then people who want to go to a clinic may be guided to a bank instead. If this association is made deliberately with the intention of misleading semantic web search engines or applications, it is called "semantic spam" [7].

Our work concentrates on matching geographic features held in different datasets with no shared form of digital identity. To generate matches, we are relying on matching attribute information and location information. Doing so by hand can be extremely time-consuming and automatic matching (such as [8-12]) using a geographic information system or some other system will still require manual checking of output matches finally. There are at least four processes that need to be considered:

1. Semantic and syntactic understanding of the datasets;
2. Translation to a common data model or representation for comparison (this may involve several different forms for different aspects of the attribution);
3. Automatic matching;
4. Manual checking and correction.

Of these (1) and (2) are largely manual. (1) would ideally capture as much of the semantics in machine-readable forms as possible as a means to aid automatic matching (3). (3) and (4) should minimize the amount of manual intervention required in (4). (4) should only present items for manual inspection where there are uncertainties or ambiguities, the majority of matches should not need to be presented. A user interface should present items for checking such that it is clear what the ambiguity is and provide tools to enable either corrections to be

Figure 2. Graphical user interface of MatchMaps with Boxes 1-4

made or at least the match to be rejected. Corrective actions may result in matches being recomputed and could result in further ambiguities or uncertainties. These would in turn need to be presented for checking. Any tool should maintain a list of the actions taken by the operator.

MatchMaps is a tool for establishing correspondences between authoritative and crowd-sourced geospatial datasets, and in particular, the correctness of the correspondences. In our previous work [13], we described the use of qualitative spatial logic reasoning, description logic reasoning and truth maintenance techniques in MatchMaps to locate problematic matches which are responsible for logical contradictions. In [14], the performance of a preliminary version of MatchMaps was compared to two fully automatic matching tools LogMap [15] and KnoFuss [16]. The precision and recall of MatchMaps (on a small dataset with manually computed ground truth) are much higher, mainly because LogMap and KnoFuss do not make explicit use of spatial information. In this paper, we focus on the MatchMaps validation process where human experts are involved to decide the correctness of problematic matches and remove incorrect ones. We describe the use of the system MatchMaps by the developer and users from Ordnance Survey of Great Britain and evaluate the effort required for validating matches.

It is apparent to the authors that the use of this technique is wider than matching authoritative and crowd-sourced data and will have applications wherever two digital maps and associated geospatial (vector) data are coincident.

II. Actions for Removing Incorrect Matches

Since initial matches generated by MatchMaps could contain errors, MatchMaps uses logic reasoning to detect problematic matches (which could lead to logical contradictions). Logical errors are discovered using description logic reasoning and qualitative spatial logic reasoning. An example of logical mistakes discovered using description logic reasoning is: two objects are matched to be the same, but one is a bank, the other is a pub. A logical contradiction arises because bank and pub are disjoint concepts, containing no common elements. An example of logical mistakes discovered using qualitative spatial logic reasoning is: two objects in the same dataset are near, but their “corresponding” objects in the other dataset are far away. A logical contradiction arises because two objects cannot be near and far at the same time.

If any logical contradiction exists, minimal sets of statements for deriving it are generated and visualized to users. Users are asked to decide the correctness of matches involved in such minimal sets of statements and remove the wrong ones. MatchMaps allows users to take four types of actions, as explained below.

Retract: If a match is found to be incorrect, then it is appropriate to retract it. A retracted match will be removed from the output. If *a partOf c* is retracted, then *a sameAs c* will be retracted automatically. Similarly, a *BufferedEqual* match could be retracted automatically as a result of retracting a related *BufferedPartOf* match.

Confirm: If a match is found to be correct, then it is appropriate to confirm it. A confirmed match will be used to validate the correctness of other matches, i.e. any match which contradicts a confirmed match will be removed automatically. If *a sameAs c* is confirmed, then *a partOf c* and *c partOf a* will be confirmed automatically. Similar rules apply when confirming *BufferedEqual* matches.

Strong Retract: If a match is found to be incorrect and all matches "similar" to it are also incorrect, then it is appropriate to use "strong retract" to retract all of these wrong matches at a time. The consequences of "strong retract" different kinds of matches are as follows.

- If *a partOf c* is strongly retracted, then *a partOf x* is retracted for any feature *x* differing from *a* (*a* is not *partOf* any other feature *x*).
- If *a sameAs c* is strongly retracted, then *a sameAs x* is retracted for any feature *x* differing from *a* (*a* is not *sameAs* any other feature *x*), *c sameAs x* is retracted for any feature *x* differing from *c* (*c* is not *sameAs* any other feature *x*).
- If *a BufferedPartOf c* is strongly retracted, then *a BufferedPartOf x* is retraced for any geometry *x* differing from *a* (*a* is not *BufferedPartOf* any other geometry *x*).
- If *a BufferedEqual c* is strongly retracted, then *a BufferedEqual x* is retracted for any geometry *x* differing from *a* (*a* is not *BufferedEqual* any other geometry *x*), *c BufferedEqual x* is retracted for any geometry *x* differing from *c* (*c* is not *BufferedEqual* any other geometry *x*).

For example, in the case shown in Fig. 1, if MatchMaps asks whether Wyndham Court is *partOf* a shop in it, then an effective action is "strong retract". As a consequence, Wyndham Court will not be stated as *partOf* any other feature in output matches. Users need to be careful not to overuse "strong retract". For example, if *a partOf c* is found to be wrong, but it is possible that *a* is *partOf* some other feature in input data, then it is appropriate to use "retract" rather than "strong retract". But if *a* is definitely not *partOf* any other feature in input data, then "strong retract" is appropriate.

Strong Confirm: If an exact correct match (a match which is correct and it entails all other correct matches) is found, then it is appropriate to use "strong confirm". The consequences of "strong confirm" different kinds of matches are as follows.

- If *a partOf c* is strongly confirmed, then *a partOf c* is confirmed, and all matches involving *a* except for *a partOf x* (*x* is *c* or *a*) will be retracted (*a* is not *partOf* any feature other than *c* and itself).
- If *a sameAs c* is strongly confirmed, then *a sameAs c* is confirmed, and all matches involving *a or c* except for *a sameAs c*, *a partOf c* and *c partOf a* will be retracted (*a* is only *sameAs c* and vice versa).
- If *a BufferedPartOf c* is strongly confirmed, then *a BufferedPartOf c* is confirmed, and all matches involving *a* except for *a BufferedPartOf x* (*x* is *c* or *a*) will be retracted (*a* is not *BufferedPartOf* any geometry other than *c* and itself).
- If *a BufferedEqual c* is strongly confirmed, then *a*

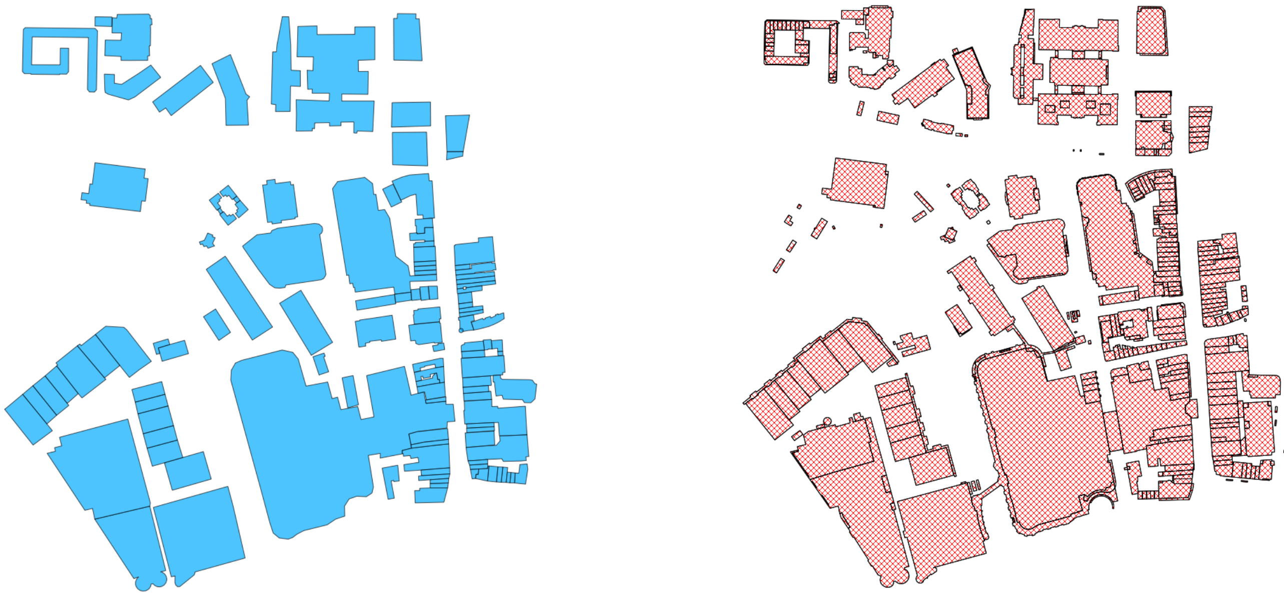

Figure 3. Data used for evaluation: a small area of Southampton, UK.

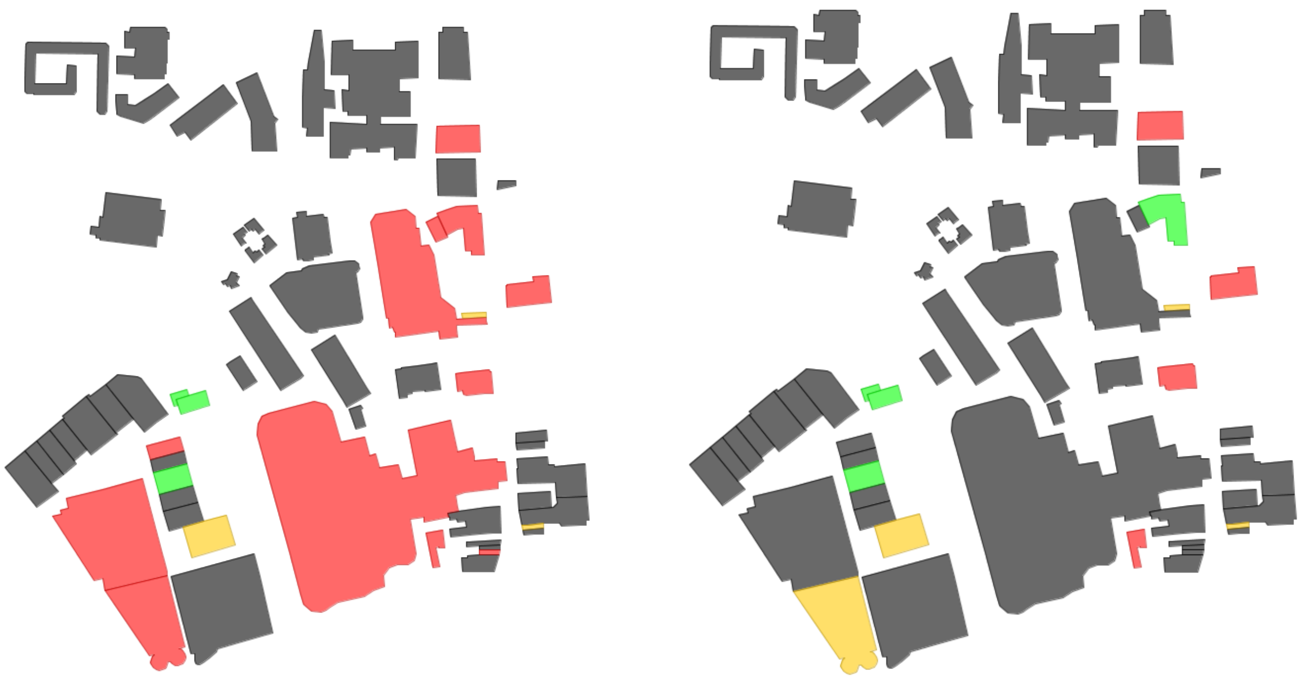

Figure 4. Object Matchng Results without Validation (left) vs. Results with Validation (right): TP: Black; FP: Red; TN: Yellow; FN: Green.

BufferedEqual c is confirmed, and all matches involving *a or c* except for *a BufferedEqual c*, *a BufferedPartOf c* and *c BufferedPartOf a* will be retracted (*a* is only *BufferedEqual c* and vice versa).

For example, in the case shown in Fig. 1, if MatchMaps asks whether Boots in OSGB *sameAs* Boots in OSM, then an effective action is "strong confirm".

III. Graphical User Interface of MatchMaps

A user interaction window of MatchMaps consists of four main parts circled by the four black boxes, as shown in Fig. 2.

Box 1 shows a match whose correctness needs to be decided by users. Users are allowed to switch between *sameAs* and *partOf*, and between *BufferedEqual* and *BufferedPartOf*. As show in Fig. 2, the match displayed in Box 1 is:

A102308429986 *sameAs* Co116824709.

Features or geometries starting with an A (stands for Authoritative) are from OSGB data; those starting with a C (stands for Crowd-sourced) are from OSM data. Users could change the displayed match to

A102308429986 *partOf* Co116824709.

If a *sameAs* or *partOf* match is displayed, then an information button is available for checking lexical information of the involved spatial features. For example, the spatial feature A102308429986 is labelled as "Antalya, Upper Parliament Street, General Commercial", Co116824709 is labelled as "Antalya Takeaway, Fast Food".

Box 2 and **Box 3** both display corresponding geometries of the match in Box 1, to help users decide whether the match is correct. The focuses of Box 2 and Box 3 are different. In Fig. 2, Box 2 focuses on the geometry of the authoritative feature A102308429986 which is displayed as a solid green polygon. Box 2 shows the boundary line of the crowd-sourced feature Co116824709 for comparison. Box 3 focuses on Co116824709 and displays A102308429986 for comparison. Three map navigation tools "Zoom in/out", "Pan tool" and "Zoom to full extent" are provided to help users examine the geometries.

The match displayed in Fig. 2 requires users to decide its correctness, because A102308429986 and Co116824709 are both labelled as "Antalya", their geometries overlap but have different sizes, shapes and positions.

In **Box 4**, six buttons are provided to users:

- **Retract:** retract the match displayed in Box 1.
- **Confirm:** confirm the match displayed in Box 1.
- **Strong Retract:** "strong retract" the match displayed in Box 1.
- **Strong Confirm:** "strong confirm" the match displayed in Box 1.
- **Next:** look at the next match. Users are advised to click this button, if they are not sure about the correctness of the current match, or they expect dealing with the next match could be more effective (i.e.

"strong retract" or "strong confirm" is applicable to the next match).

- **Quit:** quit the current round of checking (there are several rounds of checking, and each consists of checking several matches). Users are advised to click this button, if they really get stuck.

The graphical user interface of MatchMaps is implemented using libraries of OpenJUMP [17], and the JCS Conflation Suite [18] is used to process two dimensional geometries.

IV. DEVELOPER EVALUATION

The performance of MatchMaps is evaluated by its developer (the first author), for matching spatial objects in a small area of Southampton, UK. The data used for evaluation is shown in Fig. 3 and its statistics are summarized in Table I.

TABLE I. DATA USED FOR EVALUATION

Geometry		**Spatial Object**	
OSM	**OSGB**	**OSM**	**OSGB**
119	417	62	933

One of the main objectives of this evaluation is to show the effectiveness (regarding precision, recall and time) of using validation in MatchMaps. For the same set of input data, we calculate and compare the precision and recall of output matches generated by MatchMaps *without* validation (just using geometry matching and object matching) and *with* validation (using logical reasoning to detect problematic matches and using human decisions to remove incorrect matches). Another objective of this evaluation is to compare the time spent by using MatchMaps with validation against an estimated minimal amount of time required by experts to match all the objects manually.

The experiments are performed on an Intel(R) Core(TM) 2 Duo CPU E8400 @ 3.00 GHz, 4.00 GB RAM desktop computer. Times for automated processes are in seconds, averaged over 5 runs. When matching geometries, a level of tolerance 20 meters is used to tolerate slight differences. The ground truth is a subset of the ground truth for a larger dataset established in [13]. For the study area, the ground truth contains 632 *sameAs* and *partOf* matches (*a sameAs b*, *a partOf b* and *b partOf a* are counted as one). Based on the ground truth, each OSM object is classified into one of the four groups: "correctly matched" (True Positive or TP), "incorrectly matched" (False Positive or FP), "correctly not matched" (True Negative or TN) and "incorrectly not matched" (False Negative or FN). The size of each group is the number of OSM spatial objects in it. For example, for the Wyndham Court in OSM data, though there are hundreds of *partOf* matches involving it, it is only counted as one element in "Correctly Matched".

The minimal amount of effort required for matching the objects in the study area manually is estimated as follows. Assuming that an expert generates every match in the ground truth one by one by clicking two spatial objects on the maps (Fig. 3), and each match is generated like this using 3 seconds (1 second for each click, one for deciding the type of match), then the total time for generating all the matches in the ground truth is 31.6 minutes, about half an hour. This estimate is very optimistic, without taking into account the time spent in checking and comparing lexical information. The real time for matching the objects manually can be much longer, depending on the experience and knowledge of experts.

TABLE II. MATCHING OSM SPATIAL OBJECTS TO OSGB WITHOUT VALIDATION

TP	**FP**	**TN**	**FN**	**Precision**	**Recall**	**Time**
44	12	3	3	0.78	0.75	11s

TABLE III. MATCHING OSM SPATIAL OBJECTS TO OSGB WITH VALIDATION BY THE DEVELOPER

TP	**FP**	**TN**	**FN**	**Precision**	**Recall**	**Time (automated)**
50	4	4	4	0.92	0.86	47s

TABLE IV. HUMAN EFFORT OF THE DEVELOPER

Action	**Decision Time**
12	71s

The matching results generated by MatchMaps without and with validation are visualized in Fig. 4. Their statistics are summarized in Table II and Table III. The time in Table II is the total time for loading input data, generating geometry matches, generating object matches using geometry matches, and saving output matches. The time in Table III not only covers the time counted in Table II, but also the time for the *automated* part (except the human decision time which is displayed in Table IV) of validating matches using qualitative spatial logic and description logic reasoning. The human effort required in validation is summarized in Table IV. Over five runs, the average decision time by the developer is 71s. The total time for generating and validating object matches by the developer is around 118s on average.

Comparing Table II and Table III, the precision and recall of output matches are both improved by more than 10%, at the expenses of around 2 minutes validation process. Compared to the estimated minimal amount of time for matching manually, the total time for generating and validating object matches by the developer using MatchMaps is much less (118 seconds < 31.6 minutes).

V. USER EVALUATION

The user evaluation of MatchMaps aims to determine how much human effort and time is required to produce a mapping between two small (about 100 buildings) datasets using the tool. The data used for evaluation is shown in Fig. 3 and its statistics is summarized in Table I. Participants are recruited from the University of Nottingham and Ordnance Survey of Great Britain in Southampton. They are asked to use the tool, during which to decide whether some matches generated by MatchMaps are correct or not, and take actions to remove incorrect ones. The time required to make such decisions and take actions is automatically logged. The only information kept from the study is an automatically produced log of times and users' decisions.

TABLE V. USER EVALUATION OF MATCHMAPS

User	TP	FP	TN	FN	Precision	Recall	Action	Decision Time
1	49	6	4	3	0.89	0.84	13	311s
2	48	8	3	3	0.86	0.83	47	409s
3	46	10	3	3	0.82	0.79	18	390s
4	48	8	3	3	0.86	0.83	13	250s
5	50	6	3	3	0.89	0.86	34	925s
6	50	6	3	3	0.89	0.86	12	104s
7	47	8	3	4	0.85	0.81	32	574s
8	48	8	3	3	0.86	0.83	42	203s
9	48	8	3	3	0.86	0.83	32	363s
10	49	7	3	3	0.88	0.84	19	388s
11	48	8	3	3	0.86	0.83	29	301s
12	49	7	3	3	0.88	0.84	13	230s
Average					0.87	0.83	25	371s

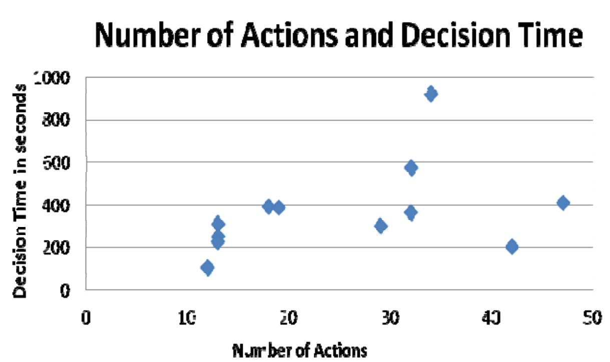

Figure 5. Number of Actions and Decision time

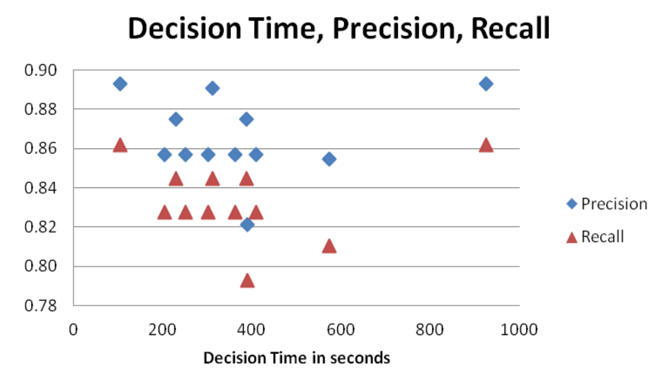

Figure 6. Decision time, Precision, Recall

The statistics of matching results generated by different users are summarized in Table V. The precision, recall, number of actions taken and decision time spent vary from user to user. On average, users generate matches with a precision of 0.87 and a recall of 0.83 by taking 25 actions using about 6 minutes. Compared to the results without any validation in Table II, the precision and recall are improved by 9% and 8% respectively. The time used is much less than the estimated minimal amount of time required by a fully manual matching process (31.6 minutes). The best matching results generated by users are almost as good as those generated by the developer (see User 6), where a precision of about 0.9 and a recall of 0.86 are obtained, taking only one or two minutes to decide the correctness of matches.

We also looked at whether the decision time is influenced by the number of actions, and whether the precision and recall are influenced by the decision time or the number of actions. As shown in Fig. 5, there is no clear correlation between decision time and the number of actions. Within almost the same amount of time, users could take very different numbers of actions. For example, using about 400s, User 2 took 47 actions and User 3 took 18 actions. Users could take almost the same number of actions using quite different time, for example, taking 32-34 actions, User 9 spent 363s but User 5 used 925s. As shown in Fig. 6 and Fig. 7, neither the decision time nor the number of actions influences the precision or recall. User 5 and User 6 obtained the same precision and recall, but took very different numbers of actions and amounts of decision time.

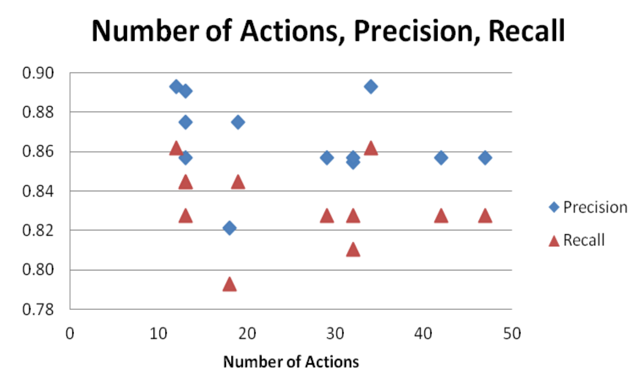

Figure 7. Number of Actions, Precision, Recall

Interestingly, there is strong correlation between precision and recall, as shown in Fig. 8. Users who obtain higher precision often obtain higher recall of outputs. The main deterministic factors for the precision and recall probably include users' different levels of understanding of the data matching problem and using MatchMaps.

VI. CONCLUSION

MatchMaps is a tool which matches spatial features in different geospatial datasets and involves users in the process of validating matches. For matching objects in a small study area, the performance of MatchMaps is evaluated by the developer and users from Ordnance Survey of Great Britain. On average, the human validation improves the precision and recall of output matches by typically 10%. The decision time required is much less than that of a fully manual matching process.

ACKNOWLEDGMENT

We express thanks to Ordnance Survey of Great Britain for providing the test data and helping with the user evaluation of MatchMaps.

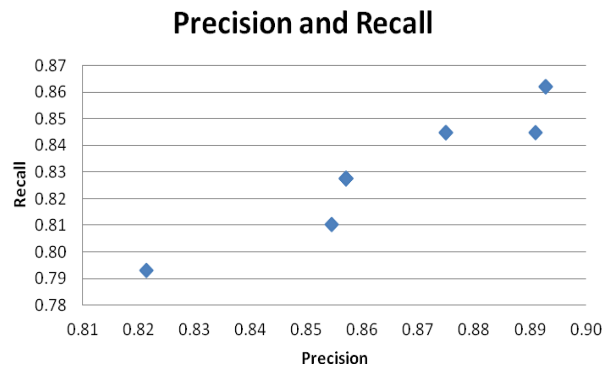

Figure 8. Precision and Recall

REFERENCES

[1] M. F. Goodchild, "Citizens as sensors: the world of volunteered geography," GeoJournal, vol. 69, no. 4, pp. 211–221, 2007.

[2] C. Heipke, "Crowdsourcing geospatial data," Journal of Photogrammetry and Remote Sensing, vol. 65, no. 6, pp. 550 – 557, 2010.

[3] M. J. Jackson, H. Rahemtulla, and J. Morley. "The Synergistic Use of Authenticated and Crowd-Sourced Data for Emergency Response," in Proceedings of the 2nd International Workshop on Validation of Geo-Information Products for Crisis Management, pages 91–99, 2010.

[4] Ordnance Survey. http://www.ordnancesurvey.co.uk, 2012.

[5] OpenStreetMap. http://www.openstreetmap.org, 2014.

[6] G. Hart and C. Dolbear, Linked Data: A Geographic Perspective. CRC Press, 2013.

[7] I. Davis, "Linked Data Spam Vectors." http://blog.iandavis.com/2009/09/linked-data-spam-vectors, 2009.

[8] S. Anand, J. Morley, W. Jiang, H. Du, M. J. Jackson, and G. Hart, "When Worlds Collide: Combining Ordnance Survey and OpenStreetMap data," in Association for Geographic Information (AGI) GeoCommunity '10 Conference, 2010.

[9] H. Du, W. Jiang, S. Anand, J. Morley, G. Hart, and M. Jackson, "An Ontology-based Approach for Geospatial Data Integration," in Proceedings of the 25th International Cartography Conference, no. CO-118, 2011.

[10] T. Koukoletsos, M. Haklay, and C. Ellul, "Assessing Data Completeness of VGI through an Automated Matching Procedure for Linear Data," Transactions in GIS, vol. 16, no. 4, pp. 477–498, 2012.

[11] B. Yang, Y. Zhang, and X. Luan, "A probabilistic relaxation approach for matching road networks," International Journal of Geographical Information Science, vol. 27, no. 2, pp. 319–338, 2013.

[12] H. Fan, A. Zipf, Q. Fu, and P. Neis, "Quality assessment for building footprints data on OpenStreetMap," International Journal of Geographical Information Science, vol. 28, no. 4, pp. 700–719, 2014.

[13] H. Du, H. Nguyen, N. Alechina, B. Logan, M. Jackson, and J. Goodwin, "Using Qualitative Spatial Logic for Validating Crowd-Sourced Geospatial Data," in Proceedings of the 27th Innovative Applications of Artificial Intelligence Conference (IAAI), 2015 (in press).

[14] H. Du, N. Alechina, M. Jackson, and G. Hart, "Matching Geospatial Instances," in Proceedings of the 8th International Workshop on Ontology Matching (P. Shvaiko, J. Euzenat, K. Srinivas, M. Mao, and E. Jim´enez-Ruiz, eds.), vol. 1111, pp. 239–240, CEUR-WS.org, 2013.

[15] E. Jiménez-Ruiz and B. C. Grau, "LogMap: Logic-Based and Scalable Ontology Matching," in Proceedings of the 10th International SemanticWeb Conference, pp. 273–288, 2011.

[16] A. Nikolov, V. Uren, and E. Motta, "KnoFuss: a Comprehensive Architecture for Knowledge Fusion," in Proceedings of the 4th International Conference on Knowledge Capture, pp. 185–186, 2007.

[17] JPP, "openjump." http://www.openjump.org, 2014.

[18] Vivid Solutions, Inc., "JTS Topology Suite." http://www.vividsolutions.com/jts, 2014.