Using Matching Algorithms for Improving Locations in Cadastral Maps

Eliyahu SAFRA and Yerach DOYTSHER, Israel

Key words: Integration, Matching, Spatial-Join, Cadastre.

SUMMARY

When integrating two different cadastral maps, points represent the same real-world location should be identified and matched. The commonly used Nearest-Neighbor (NN for short) join may be applied in order to match between objects from the two sources. Yet, this method has several evident drawbacks. First, the NN join is not symmetric, i.e. it depends on the matching direction. Second, in order to make the integration more accurate, one needs way to evaluate the quality of each match. Integration of regular vector datasets was studied in the past. This research investigates the specific features of integration of cadastral maps. When integrating cadastral maps, the main issue after identifying the matching points, is the points locations accuracy. The matching process can be used to improve the locations accuracy by interpolating the locations from the two sources. However, typically part of the points from each source does not appear in the other source, so their location cannot be easily corrected. Moreover, each pair of matched points is in some confidence, and it is not clear which matches should be included in the result. By defining pairs of matched points and points appearing in one source only, improving the locations of not matched points by finding a systematic distortion between the sources in some area is becoming fissile. Correcting the locations of not matched points is based on linear and/or non-linear rubber-sheeting transformations.

Samples containing result of the integration of two large cadastre maps are presented and discussed. The results show that, without correcting the not matched points, the topology relations between neighboring nodes can be damaged. Furthermore, by applying new matching methods instead of the commonly used NN (Nearest-Neighbor) method the quality of the resulting map can be significantly improved.

Using Matching Algorithms for Improving Locations in Cadastral Maps

Eliyahu SAFRA and Yerach DOYTSHER, Israel

1. INTRODUCTION

1.1 Geographic Information Systems

The commonly GI (Geographic Information) System is nowadays the main framework used for storage, retrieval, mapping, and analysis of geographic data. LIS (Land Information System) are the GIS adaptation for handling Cadastre data and land management. In these systems spatial features are stored in a coordinate system, which references a particular location on Earth. Descriptive attributes in tabular form are associated with spatial features. Spatial data and associated attributes in the same coordinate system can then be layered together for analysis as well as mapping in general and cadastral mapping in particular.

1.1.1 Representation Mode

Two common approaches for representing spatial data in a geospatial database are available:
– **Raster** – A cellular data structure composed of rows and columns. Groups of pixels with the same value represent features. This representation is the output of sensors, such as satellite, aerial photographs etc.
– **Vector** – A coordinate-based data structure used to represent geographic features, which represent entities in the world. A *dataset* is a collection of features about a certain subject e.g. junctions, roads, or buildings.

In this work, as dealing with cadastral data, we consider data in Vector mode, which is the most common, as well as the most convenient way to represent data in the database. In that mode, every object has its spatial part, which may embody both geometry (i.e. location and shape) topology (i.e. spatial relationship among objects, such as adjacency), and an alphanumeric (or descriptive) part to describe other common attributes.

1.1.2 Representation of Geographic Objects in the Vector Mode

There are three common ways to represent the spatial part of a geographic feature in the database:
1. **Point** – A point is a zero dimensional object that holds its latitude and longitude (or northing and easting). It is used to represent speck features e.g. springs and junctions.
2. **Line** (polyline) – One dimensional object consists of consequential connected segments (a segment is part of a polyline limited by two points). It is used to represent linear features e.g. rivers and roads.
3. **Polygon** – Two-dimensional object consist of a closed line (a line in which the start point is identical to the end point). It is used to represent features with area e.g. ponds and buildings.

In our case, when dealing with cadastral data, points will represent parcel corners (vertices), line will represent parcel fronts, and polygons will represent cadastral parcels.

1.2 Data Integration

With the evolution of fast networking, data integration has become a major challenge for both industry and research. In many instances, in order to solve analysis problems we must integrate data from different sources. The problem arises from the fact that each organization collects the data to fill its own needs. In many instances, in order to solve real world problems, data from one source alone is insufficient; as a consequence, we may need to gather data from several sources. In our scenario of dealing with cadastral data, using just one source, specifically the cadastral (graphical) maps, will result in having non accurate and non complete cadastral information.

In order to use the collected data we have to *integrate* the data from the different sources in a seamless manner, so that we can query the data from the different sources as easily and fluently as if it came from a single source.

1.2.1 Problems in Data Integration

– **Ontology differences –** We have one world. There are, however, many ways to define its *ontology*. These different ways derive from different points of view and they may all be correct.
– **Schema differences –** Typically different sources contain different schemas, the schema of the resultset should be different from those in the sources, and in some way contain the data from all sources.
– **Entities identification** – Every source is a collection of objects. Each world entity is typically represented by (at most) one object in the source. However, since integration is done upon several sources, each entity may be represented by several different objects. Thus, we need to identify, which objects represent the same entity.

1.2.2 Object Fusion

When two objects describe the same entity, we can simply show them both to the user, however, this may be confusing and make the data analysis more time consuming. Another option is to perform *object fusion* i.e. to combine the two.
Object fusion is a simple task when the different sources use the same key for object identification, since objects are fused when they have an identical key.
When there is no such key, we can use other attributes for identification e.g. name, phone-number etc'. In this case, however, the fusion result may contain errors, whose number depends on the quality of the attributes that were used, and on the quality of the fusion algorithm.

1.3 Integration of Spatial Datasets

The additional component in spatial databases, compared with other databases, is that each object has a spatial part; thus, the first stage of any integration is to place the two sources on one map. This is done by matching their coordinate systems (i.e., converting the old Israeli cadastral maps from Cassini-Soldner grid to Israel Transverse Mercator grid and datum), and detecting systematical displacement. After this is done, we want to remove redundancies, i.e. to present each geographic entity by one object on the map, which leads to object fusion. In the fusion process, the location, which is attached to every object, can be used as a key to identify the entity it represents.

There are problems, however, that still demand a solution:

– Spatial databases inevitably contain *errors* of some kind. The very fact that spatial datasets represent real world phenomena in an abstract and generalized form necessarily signifies that they can never be truly accurate. Errors may occur at any stage of the database construction process. Data collection is subject to the accuracy of the particular technique being used. For example, data derived from remotely sensed images will include errors due to the characteristics of the airborne platforms and sensor systems (Ware and Jones, 1998)
Data resulting from the processing of existing datasets may include additional errors due to error propagation. For example, a coverage obtained by overlaying an existing pair of datasets will contain an error that is a function of the error contained in each of the source datasets.

– The data may be in different datum, different coordinate systems, and different representations (i.e. point, line, or polygon).

– It is an also common occurrence that entities (i.e. being in the world), which are represented by objects (i.e. feature in the dataset) in one dataset, are not represented in another. This may happen because of different generalization level or because of differences in the time period in which the data was gathered.

These factors entail that location-related object fusion process be a careful one, so we can benefit from the usage of the many different sources.

1.4 Different Aspects of Cadastral Data

Cadastre is the method of registering land, designed to ensure the rights of individuals and the state of their property. In Israel as in many other countries, the cadastral data has been and collected for the duration of many decades or even centuries by using different measuring methods and techniques, and in varying levels of accuracy.

Until recent years, measurement results were recorded in field books, and used to determine the boundaries of the cadastral blocks and the parcels, as well as other features (buildings, fences, electric poles, etc.). All these measurements were depicted graphically on field plan sheets. Maps of the cadastral blocks were prepared based on the field sheet, consisting of all parcels in the block and all included features. These cadastral maps contain neither the measured data nor any dimensions whatsoever of the parcel boundaries and serve only as a graphical presentation of parcel layout. Measurements for cadastral mapping were performed by using the chain surveying method (until the 1970s), then by using the polar method of theodolite and electro-optical distance measuring (TotalStation), and lately by using the GPS technology.

In addition to these different measuring methods, the accuracy level of the national control network has been changed from a very low accuracy based mainly on local adjustment of separate traverses to a rigorous adjustment as a uniform and accurate network. The previous Israeli Cassini-Soldner grid has been replaced during the 1990s by ITM (Israel Transverse Mercator) grid and new datum.

Accurate, accessible and updated cadastral information constitutes the basis for planning and implementation of a variety of real estate related operations in many areas. In the present form of cadastral maps ("graphic cadastre") in Israel, the existing cadastral information does not fulfill these needs. In order to improve the quality of the current level of the graphic cadastre, it is needed to integrate data from several sources. The existing sources for obtaining such data are field measurements of land boundaries; digitizing existing maps; and processing the existing surveying data.

1.5 The Contribution of this Work

In this work we deal with location based *spatial object fusion*. The aim of the research is to find objects which correspond to the same real world entity, in the different sources. We assume that the sources are in the same coordinate system, thus, we deal with problems which arise either from different measuring and computing methods, low accuracy or from differences in the generalization level. Two different new algorithms to perform this task are presented (in addition to a third standard and common method). Each algorithm is better than the others in different situations e.g. different densities and different accuracies. We also present the results of extensive tests that illustrate the weaknesses and strengths of these algorithms, under varying assumptions about the density of each spatial database and the degree of overlap between these databases.

The outline of this paper is as follows. In Section 2, we formally define the problem and describe ways to measure the quality of the results. Section 3 describes the different algorithms for object fusion. Section 4 discusses the tests and presents their results. Section 5 contains Related-work. Finally, Conclusions and a discussion in issues demanding further research appear in section 6.

2. DEFINITIONS

2.1 Problem Definition

Geographic Information System is a repository for geographic objects, or just objects for short. Each *object* contains information about a real world entity; a real world entity is described uniquely in the system by one object. In the system the objects are organized in *datasets*, where each dataset contains objects about a certain subject.

The input for fusion process is two or more datasets from different sources. The datasets should overlap on both the subject and the described area. In a preprocessing filtering stage, objects, that dose not belong to the intersection between the datasets, are eliminated. It is still assumed however, that the overlap between the datasets, after the filtering stage, is not complete, i.e. there are world entities which appear only in one dataset.

When geographic datasets are integrated, an important task is to identify when two or more objects, from the different sources, represent the same entity and fuse those objects to a single object. Since each entity is represented by at most one object in a dataset, a fusion set may contain at most one object from each dataset. Thus, the matching between objects is 1:1.

A fusion algorithm operates over several input datasets and generates (non empty) fusion sets. Each fusion set is a collection of objects that are believed to represent the same world entity. A fusion set can contain at most one object from each source. A fusion set is *complete* if it contains all the objects that represent one given entity (but it may contain other objects as well). A fusion set is *sound* if all the objects in it represent the same entity (there may be, however, other objects which represent the same entity as well). A fusion set is *correct* if it is both complete and sound i.e. it contains all the objects which represent an entity and nothing else.

In this work we deal with the problem of creating fusion sets using only the location of the objects. This task could be a simple one if the datasets are accurate and the representation (e.g. polygon) of the objects is identical. Since this is not the case, as mentioned in Section 1, we need to estimate the fusion sets.

The presented algorithms deal with fusing two datasets, denoted as $A = \{ a_1 \ldots a_m \}$ and $B = \{ b_1 \ldots b_n \}$. Thus, each fusion set contains either one object i.e. $\{a_i\}$ or $\{b_j\}$ where $1 \leq i \leq m$, $1 \leq j \leq n$, or two objects i.e. $\{a_i, b_j\}$. A singleton fusion set (i.e. $\{a_i\}$ or $\{b_j\}$) is correct if the entity the object represents does not appear in the other dataset. A pair fusion set (i.e. $\{a_i, b_j\}$) is correct, if the two objects $\{a_i\}$ and $\{b_j\}$ correspond to the same world entity, or, in other words, are corresponding objects.

2.2 Measuring the Quality of the Result

Since we do not know the correct fusion sets but try to estimate them, we need to evaluate the quality of our estimation. This is done in terms of *recall* (i.e. completeness of retrieval) and *precision* (i.e. purity of retrieval) as in information retrieval.
Note that the evaluation of the recall and precision can be done *only* if we know the relationships between objects and the entities they represent, so we can compute the correct results. In real world datasets, frequently, this is not the case.

2.2.1 Recall and Precision

Empirical studies of retrieval performance in IR (Information Retrieval) have shown a tendency for precision to decline as recall increases. The trade-of between precision and recall is inherent, not merely an inconvenient empirical finding[1]. Ascertaining the

[1] Look in http://www.sims.berkeley.edu/research/projects/oasis/trade.html

relationship between them, however, does not fall within the scope of this work. Rather, these parameters will be used to measure the quality of the result. Hereafter we define recall and precision for our context.
he measures counts the number of entities represented by correct fusion sets versus the number of entities represented in the different sources or the result set. The intuition is to measure how much we succeed in describing the reality by integrating the data. The definition of the recall and precision is:

$$Recall = \frac{\text{\# of correctly described entities}}{\text{\# of entities}}$$

$$Precision = \frac{\text{\# of correctly described entities}}{\text{\# of fusion sets in the result}}$$

3. FUSION METHODS

In this section, we will describe three join methods. First, the commonly used Nearest-Neighbor method. Then, two other methods that we have developed. The Mutually-Nearest method and Normalized-Weights method.

3.1 Nearest Neighbor (NN) Method

The nearest neighbor Method is the common used tool to fuse different sources. The operation (denoted as *NN*) takes two datasets *A* and *B*, and to each objects of *A* matches the closest object in *B*. This basic method is being implemented by many of the GIS software producers, and for example, ESRI (Minami, 2000) offers a spatial nearest neighbor join as a standard tool which is part of their ArcView and ArcInfo software packages.

There are several drawbacks in using this operation for object fusion. This operation may result both in data loss and data duplication of *B* objects, since one object from *B* may be the closest to more than one object in *A*, and another object to no object in *A*. This operation is also asymmetric, since the result of $NN(A,B) \neq NN(B,A)$.

3.2 Mutually-Nearest (MUTU) Method

The first and simplest novel fusion algorithm presented in this work is Mutually-Nearest Method (MUTU for short). This Method fuses two objects *iff* ("if and only if") each object is the closest to the other. Other objects, that do not have a mutual nearest neighbor object generate singleton fusion sets i.e. fusion sets with only one object.

3.2.1 Match Definition

Let the datasets be *A* and *B*.
Let $a \in A$, $b \in B$ be objects.
If for any *x* such that $x \in B$, *b* is closer than *x* to *a*. **And** for any *y* such that $y \in A$, *a* is closer than *y* to *b*.
Then the set *{a,b}* is a pair fusion set.

Note that since closer mean *strictly* closer, when there are two objects e.g. ($x, y \in A$ and $x \neq y$) with identical minimal distance to an object $b \in B$, there would be no pair fusion set creation.

3.2.2 Algorithm Definition

The algorithm definition is depicted in Figure 1.

3.2.3 "Too Far" – Distance Limit

Theoretically, any two objects (from different datasets) may correspond to the same entity; however, if the distance between two objects exceeds a certain threshold, the probability for a correspondence is low. The **"Too Far"** parameter is a threshold such that if the distance between two objects exceeds it, these objects will not be in the same fusion set.
The definition of this parameter is as the sum of the errors in the integrated datasets, and is identical to the error interval. Thus, we change algorithm - in Figure 1 - line 5 - as follows: when the distance between two mutual nearest objects exceeds the **"Too Far"** value, the algorithm will generate two singleton fusion sets, rather than one pair fusion set.

3.2.4 Properties of the Mutually-Nearest Method

The Mutually-Nearest has two main advantages. First, the fusion operator is symmetric. In addition, each object from A and B appears in the result set precisely once, thus there is neither data loss nor duplication.

The main drawback of the Mutually-Nearest is that each object has only one candidate for a match, the closest one. While in a dense area, and in complicated situations, we may want to check several other candidates as well, and possibly select one that is not the closest.

3.3 Normalized Weights Method

The Normalized-Weights Method (NW for short) is more complicated than the previous methods. The NW method consists of four steps. The steps are:
1. Setting probabilities.
2. Creating table of mutual probabilities.
3. Normalizing the table.
4. Selecting sets.

Because of space limitation, we won't describe the NW method here. A full description appears in (Beeri et al., 2004; Beeri et al., 2005)

Algorithm 1 Mutual Nearest Fusion

Input–Two datasets $A = \{a_1 \ldots a_m\}$ and $B = \{b_1 \ldots b_n\}$.

Output–The fusion results of A and B.

for all $a \in A$ **do** {stage-1}
    Find b–the nearest object from B to a
    Find $\hat{a}$–the nearest object from A to b
    if $a = \hat{a}$ **then**
        Create the fusion set $\{a, b\}$
    else
        Create the fusion set $\{a\}$
    end if
end for
for all $b \in B$ **do** {stage-2}
    if b is not in any fusion set **then**
        Create the fusion set $\{b\}$
    end if
end for

Figure 1-The Mutually-Nearest method

3.4 Comparison between the three Methods

We described here three different novel methods for object fusion; in this section we will briefly compare some of their features.

3.4.1 Methods Results Comparison

There is an important difference between the results of the MUTU method on one side hand and those of the NW and NN methods in the other hand. In the MUTU method, each object appears in precisely one fusion set, since each object either meets the requirements of Mutually-Nearest, or else creates a singleton fusion set. In the other methods each object may not appear at all, i.e. neither as part of a pair nor as a singleton, or appear in more than one fusion set, since in these methods we create fusion sets by threshold on the confidence values.

3.4.2 Complexity

Suppose that the two datasets *A* and *B* have *m* and *n* objects. The running time of the MUTU method is $O(m \cdot n)$. However, this can be improved by using a spatial index that can find the nearest and second-nearest neighbors in logarithmic time to $O(m \cdot log(n) + n \cdot log(m))$.

The runtime of the NW method is $O(i \cdot (m \cdot n))$, where i is the number of iterations in the normalization process (note that *i* is known only after the algorithm has finished). In the tests, the number of iterations was always an order of magnitude less than the number of objects.

4. TESTS

This section contains tests that we perform to check the performance of the different methods on data that imitate cadastral data. The tests aim at answering the following questions: First, what are the performances of the three methods on cadastral data? And secondly, which of the three methods should be used for cadastral data?

4.1 Dataset Generator

We found that using real-world data as is, for evaluating the three methods, is an insufficient approach. This is due to the fact that in many cases it was hard to check whether a fusion set is correct or incorrect. We therefore implemented an experiment framework that allowed us to test our algorithms on partially synthetically data. There are two main advantages in testing the methods using this framework. First we were able to vary the datasets in size, error, density and overlap (i.e. the number of world entities that appear in both sets). Second, it was very simple to check whether a fusion set is correct or incorrect.

In this section we first describe the features of the datasets generation framework and then the results of the tests with the synthetic data are presented.

4.1.1 The Generation Process

To best imitate the features of cadastral data, we used as our "world" a collection of points defining all parcel corners of a cadastral block in a rural zone near Haifa (in the northern part of Israel). These points were measured by a private licensed surveyor for the Survey of Israel (the governmental agency responsible for geodesy, mapping, cadastre and GIS) during a re-parceliziation process of the area. Based on these points, in each test two different datasets were created. Each dataset contained a subset of the points according to some size factor.

The object creation for establishing a dataset is an iterative process. The generator randomly selects an entity *e*, then it checks if *e* is already represented in the dataset. If the entity *e* is not represented in the dataset, the generator creates an object *o* with location in a circle with a center at the location of *e* and which has a radius that is equal to the error-interval[2] parameter. The location of *o* is calculated using two random numbers. One is the direction *a*

[2] Error-interval – the maximal distance between an object and the entity it represents.

(360 ≥ a ≥ 0) in a uniform distribution, and the second is the distance *d (Error-interval ≥ d ≥ 0)* in a Gaussian distribution. The location of *o* is in distance *d* from *e*, and in direction *a* from *e* with respect to the north. The objects generation process ends when the number of objects is reached.

4.1.2 Generator Parameters

The tests were designed to check the algorithms for various degrees of overlap. The world in all the tests is identical; the tests differ only in the number of objects in each dataset, and thus the degree of the overlap differs as well. The sizes of the two sources in each test are identical. We also checked the algorithms using datasets with different sizes, but the results we present here are sufficient to demonstrate the main conclusions.

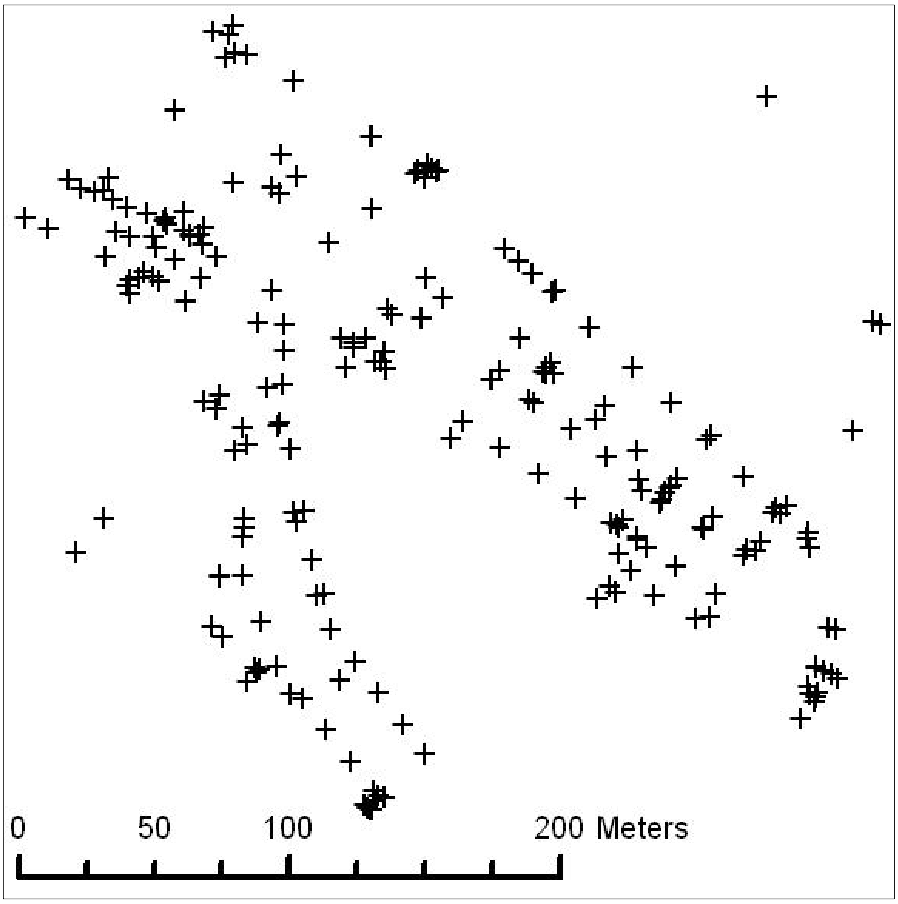

Figure 2 - A small fragment of the world

The test parameters are as follows:

1. Standard deviation of the error:

a. First dataset 1.25m (imitating a digitized cadastral block at the scale of 1:2,500 and summarizing the inaccuracy within the measuring and plotting process and then a digitizing process)

b. Second dataset 0.3m (imitating a restoring process of the cadastral boundaries and then an accuarate field surveying of these boundaries)
2. part of the entities that are represented in each dataset, 40%, 50%, …, 90%

A partial presentation of the world is depicted in Figure 3. To get a sense of the difficulty in finding the correct fusion sets, an enlarged fragment of the test (where the size factor is 0.9) with identifiers are attached to the objects.

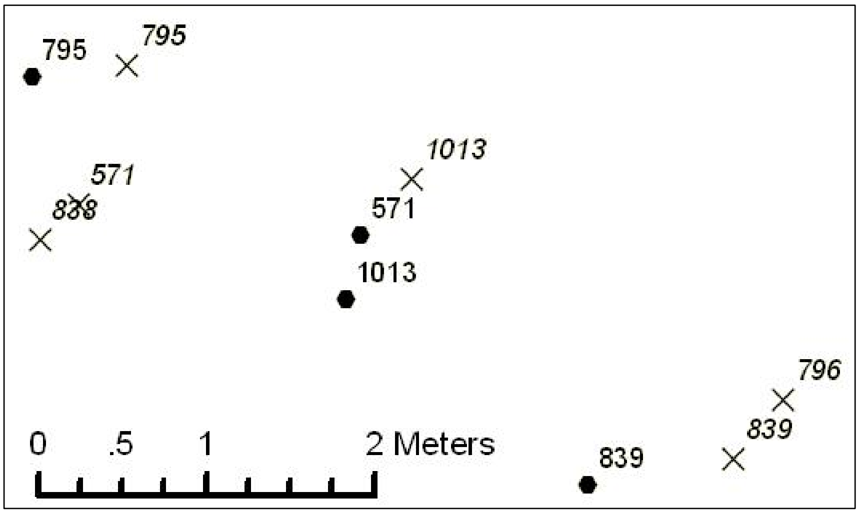

Figure 3 - Enlarged Fragment of the test area

The figure illustrates that the closest selection of an object is not always correct (as for entity 571). Note that when the distance between entities is larger, or when the Error-interval is smaller, such a mix up situation can not happen, and finding fusion sets is simpler.

4.1.3 Tests Results

Figure 4 and Figure 5 describe the final results of our tests. Figure 4 describe the Harmonic mean of Recall and Precision. Figure 5 describe the Recall and the Precision in separated graphs (its legend is identical to the legend of Figure 4).

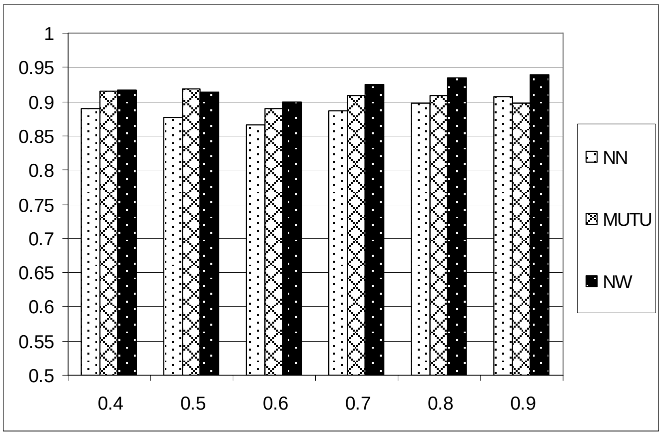

Figure 4 - Harmonic mean of Recall and Precision

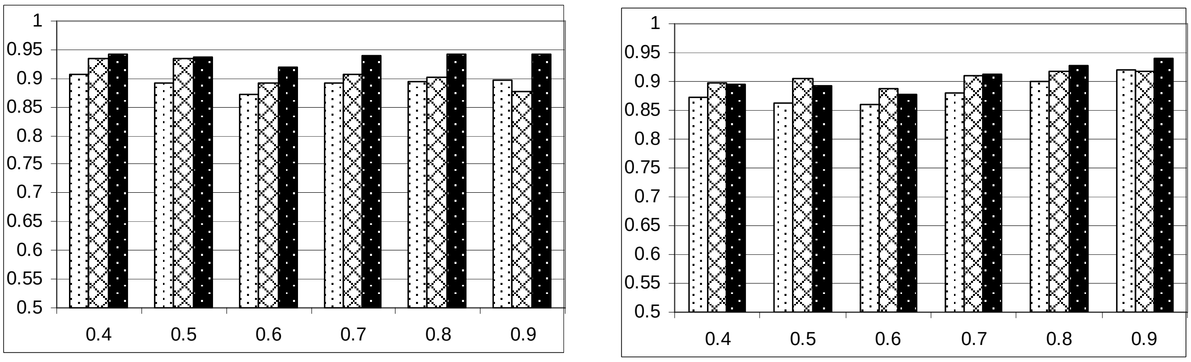

Figure 5 - Recall (in the left), Precision (on the right) of the different methods

The tests show that:

1. The NW has the best or close to the best result under all circumstances.
2. The precision of the MUTU method is higher than the NN method in most cases.
3. For small overlap the precision of the MUTU is the highest. For large overlap the precision of the NW is the highest.

5. RELATED WORK

As accurate cadastral information is the basis for planning real estate related operations (Effenberg et al., 1999), integration of cadastral data from multiple sources is required (Durdin, 1993). An alternative approach of resurveying by modern and accurate techniques all land boundaries is practiaclly a non realistic solution (Arvanitis and Koukopoulou, 1999). Based on these facts, much effort has been invested over the years in developing methods to convert analogue (graphical) cadastral maps into digital information (Hvidegaard, 1987; Vonderohe et al., 1990; and, Williamson, 1996). The option of integrating external information with digital data derived from maps by applying geometric or cadastral constraints was considered by the cadastral community (Morgenstern et al., 1989; Hesse et al., 1990; Tamim and Schaffrin, 1995).

6. CONCLUSIONS AND FUTURE WORK

In this work we discussed how we can fuse cadastral data on the sole basis of location; in addition, we present research about parameters (e.g. *TooFar*) to adapt fusion operation to different situations (e.g. different error intervals). Three methods, with different degrees of complexity, are discussed and their results are compared. The first method, match to each object, its nearest neighbor. The second method, checks whether the nearest object is a good candidate for a match or whether it is not. The third method takes all the objects within the distance bound as candidates and selects the fusion sets with the highest degree of confidence using threshold value.

The tests we have made have shown that the greater the numbers of things taken into account, the better the results. However, since checking things abounds in computation time, there are instances when such a process may be spared altogether.

Much work, however, remains to be done. First, in the proposed algorithms we assume that each entity is presented by one object from each data set at the most, i.e. a 1:1 matching. There are cases, however, when the matching should be one to many, as in the generalization problem. Second, we only[3] use the location attribute of each object to identify it, in many instances there are other attributes, both spatial (e.g. area, perimeter or shape of polygon) and alphanumeric (e.g. name), which may help us in the matching process, as well as in increasing the recall and precision.

More important work involves the fusion of more than two datasets. It may be argued that multiple sources fusion may be done sequentially, i.e. the fusions of two sources, thereafter the fusion of the result set to the third source and so on. But since any fusion inevitably contains errors, it is not clear if such an approach leads to good results. A possible solution is to do the fusion of all the sources simultaneously. This, however, is a complicated process, since the number of possible fusion sets is *exponential* in the sources number; thus, such algorithm should be designed thriftily.

REFERENCES

Arvanitis, A., Koukopoulou, S. (1999). Managing Data During the Update of the Greek Cadastre. *CD Proceedings of FIG Commission 3 Annual Meeting and Seminar*, Budapest, Hungary, 7.2, 1-9.

Beeri, C., Kanza, Y., Safra, E., and Sagiv, Y. (2004) Object fusion in geographic information systems. *In Proc. of the 13th International Conference on Very Large Data Bases*, Toronto, Canada.

Beeri, C., Doytsher, Y., Kanza, Y., Safra, E., and Sagiv, Y. (2005) Finding corresponding objects when integrating several geo-spatial datasets. *In Proc. of the 13th ACM*

[3] Generalization – transforming a map from big scale to a smaller scale

International Workshop on Geographic Information Systems, ACM-GIS, Bremen (Germany).
Durdin, P. (1993). Measurement-Based Databases: One Approach to the Integration of Survey and GIS Cadastral Data. *Surveying and Land Information Systems*, 53(1), 41-47.
Effenberg, W.W., Enemark S. Williamson I.P. (1999). Framework for Discussion of Digital Spatial Data Flow within Cadastral Systems. *The Australian Surveyor*, 44(1), 35-43.
Hesse, W.J., Benwell, G.L., Williamson, I.P. (1990). Optimizing, Maintaining and Updating the Spatial Accuracy of Digital Cadastral Databases. *The Australian Surveyor*, 35(2), 109-119.
Hvidegaard, J. (1987). A Concept for the Renovation of the Danish Cadastral Maps. *Proceedings of the Workshop on Cadastral Renovation*, Lausanne, Switzerland, 81-91.
Minami, M. (2000). Using ArcMap. Environmental Systems Research Institute, Inc.
Morgenstern, D., Prell, K.M., Riemer, H.G. (1989). Digitization and Geometrical Improvement of Inhomogeneous Cadastral Maps. *Survey Review*, 30(234), 149-159.
Tamim, N., Schaffrin, B. (1995). A Methodology to Create a Digital Cadastral Overlay through Upgrading Digitized Cadastral Data. *Surveying and Land Information Systems*, 55(1), 3-12.
Vonderohe, A.P., Fisher, S.S., Krohn, D.K. (1990). A Positional Comparison of the Results of Two Digital Cadastral Mapping Methods. *Surveying and Land Informational Systems*, 50(1), 11-24.
Williamson, I.P. (1996). Coordinated Cadastre – A Key to Building Future GIS, *Proceedings of the Regional Conference on Managing Geographic Information Systems for Success*, 60-69.
Ware, J. M., Jones, C. B. (1998) Matching and aligning features in overlayed coverages. In *Proceedings of the 6th ACM International Symposium on Advances in Geographic Information Systems*, pages 28-33.

BIOGRAPHICAL NOTES

Mr. Eliyahu Safra is a PhD candidate at the Technion (Israel Institute of Technology) in Mapping and Geo-Information Engineering. His master degree (2004) is in computer science where he implemented an integrator for geographic datasets. His PhD research includes a complete approach for integrating heterogeneous geospatial datasets.

Prof. Yerach Doytsher graduated from the Technion - Israel Institute of Technology in Civil Engineering in 1967. He received an M.Sc. (1972) and D.Sc. (1979) in Geodetic Engineering also from the Technion. Until 1995 he was involved in geodetic and mapping projects and consultation within the private and public sectors in Israel. Since 1996 he is a faculty staff member in Civil and Environmental Engineering at the Technion, and is currently the head of the Department of Transportation and Geo-Information Engineering. He is also heading the Geodesy and Mapping Research Center at the Technion.

CONTACTS:

Mr. Eliyahu Safra
Technion – Israel Institute of Technology
Faculty of Civil and Environmental Engineering
Department of Transportation and Geo-Information Engineering
Technion City
Haifa 32000
ISRAEL
Tel. + 972 4 8293065
Fax + 972 4 8234757
Email: safra@techunix.technion.ac.il

Prof. Yerach Doytsher
Technion – Israel Institute of Technology
Faculty of Civil and Environmental Engineering
Department of Transportation and Geo-Information Engineering
Technion City
Haifa 32000
ISRAEL
Tel. + 972 4 8293183
Fax + 972 4 8234757
Email: doytsher@technion.ac.il