In *Proceedings of the 8th ACM SIGSPATIAL Workshop on Geographic Information Retrieval (GIR'14)*, Dallas, TX, November 2014

# Using Minimaps to Enable Toponym Resolution with an Effective 100% Rate of Recall*

Hanan Samet
Department of Computer Science
Center for Automation Research
Institute for Advanced Computer Studies
University of Maryland at College Park
hjs@cs.umd.edu

## ABSTRACT

A number of systems have been recently constructed that make use of a map query interface to access documents by the locations that they mention. These mentions are often ambiguous in the sense that many interpretations exist for the locations which are not always expressed along with all the necessary qualifiers. In other words, users are assumed to be able to make the appropriate identification based either on knowledge of prior queries or the nature of the document containing the references as well as knowledge of the target audience. The disambiguation process is known as toponym resolution. The map query interface results in the placement of icons and links to the appropriate documents at the corresponding location on the map. Assuming that all toponyms have been recognized (i.e., 100% rate of recall for toponym recognition), it is shown how to achieve an effective 100% rate of recall for toponym resolution for all interpretations of a toponym that the toponym recognition process associates with at least one document. This is done with the aid of a minimap that shows all of these interpretations which means that a user has access to all documents that mention a specific location as long as the textual specification to the location has been recognized as a location rather than as the name of another entity such as a person, company, organization, etc. It also assumes that the user is capable of determining the correct interpretation of each toponym. This is important as it enables the determination of precision and recall.

*This work was supported in part by the National Science Foundation under Grants IIS-10-18475, IIS-12-19023, and IIS-13-20791, as well Microsoft Research, Google Research, and NVIDIA.


GIR'14,November 04 2014, Dallas, TX, USA.

978-1-4503-3135-7/14/11 ...$15.00.
http://dx.doi.org/10.1145/2675354.2675698.

## Categories and Subject Descriptors

H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing

## General Terms

Algorithms, Design, Performance

## Keywords

geotagging, precision, recall, minimap, toponym resolution

## 1. INTRODUCTION

The increasing use of the world wide web and the increasing availability of devices with a GPS capability has led to an explosion in the number of applications that make use of location information. This has also led to a wide variety of ways of expressing it. They range from the conventional explicit methods that are geometric in nature that use latitude-longitude pairs of numbers (e.g., QUILT [43] and the SAND Browser [11, 38]) to implicit ones such as textual (name or description of the location which can even be expressed verbally), the user's IP address, as well as those that rely on gesturing or pointing to a map which is equivalent to using a map query interface (e.g., NewsStand [23,37,40,41,45]). The textual specification has found increasing use due to the ease of entering it when using a device such as a smartphone for querying such spatially-referenced data. For example, see Figure 1 which consists of a subset of a series of advertisements for AT&T pointing out the importance of context in understanding textual specifications of locations. In particular, we see the utility in (a) forming an email address for London, (b) determining the performance of a team vis-a-vis a sports league in Dublin, (c) checking the weather in terms of interpreting the temperature unit in Mexico, and (d) finding a restaurant serving local food in China. Of course, once users use a textual specification to query the data, they also welcome the capability to query data expressed and stored textually.

An advantage of the textual specification of location data (termed a *toponym*) is that it acts like a polymorphic type in the sense that one size fits all. In particular, depending on the application which makes use of this information, a toponym such as "San Francisco" can be interpreted both as a point or as an area, and, most importantly, the user need not be concerned with making this choice. The drawback of the textual specification of location data is that we are


(a)


(b)


(c)


(d)

Figure 1: Example of a series of four advertisements for AT&T from `http://markkuphoto.com/` with photos by Markku Lahdesmaki highlighting the need to understand textual specifications of locations. (a) Forming an email address for London, (b) determining the performance of a team vis-a-vis a sports league in Dublin, (c) checking the weather in terms of interpreting the temperature unit in Mexico, and (d) finding a restaurant in China. Courtesy of AT&T Intellectual Property. Used with permission.

not always sure that the reference corresponds to a location (e.g., "Paris Hilton" which can be the name of a person or a location which in this case is the hotel). Moreover, there are many possible locations with the same name (e.g., "Paris"). and thus the "Paris" toponym is ambiguous. These ambiguities arise when processing documents such as newspaper articles, tweets, blogs, etc.

The study of techniques for resolving these ambiguities is known as *Geographic Information Retrieval* or *GIR* for short. These techniques have been applied in many different textual domains including Web pages [5, 30, 31], blogs [32], encyclopedia articles [17, 44], tweets [13–15, 42], spreadsheets [1, 3, 27], the hidden Web [26], and news articles [6, 12, 21, 22, 24, 25, 33, 40, 45]. These approaches can be further classified as either being rule-based [7–9, 16, 31, 34, 47] or statistical [18–20]. The ambiguities are resolved in a number of ways. One way is by taking advantage of the fact that toponyms often appear together as in lists or tables where we can make use of clues such as prominence, proximity, and sibling (e.g., [2, 4, 24]). Another way is by taking advantage of knowledge of the nature or scope of the document in which the toponym appears. We have three possible scopes [28, 46].

1. The provider scope which corresponds to the identity and location of the content provider.

2. The content scope which corresponds to the type of content that the publication contains (e.g., a financial newspaper).

3. The reader or serving scope which corresponds to the location of the readers and their demographics (e.g., a publication that focuses on immigrants from a particular country or region of the world). This is captured by making use of local lexicons [25, 33].

The process of identifying and disambiguating textual references to spatial data is frequently referred to collectively as *geotagging*, as well as *geoparsing* (or more precisely as *spatial named entity (SNE) geoparsing* [29]), with little or no distinction being made between them in the sense of the context in which they are invoked. We believe that this context is important since the underlying processes are very different as are the techniques/measures for evaluating them. Hence, we propose to use the term *spatiotextual data handling*, or *geodetexting* for short, to refer to the collective task; the term *geoparsing* to refer to the process of identifying and disambiguating textual references to spatial data in queries; and the term *geotagging* to refer to the process of identifying and disambiguating textual references to spatial data in the underlying data that is being queried. This data will often be placed in specific tables based on the disambiguation to facilitate subsequent access (e.g., via a spatial index). Both geoparsing and geotagging are evaluated by the same measures of precision and recall [35].

As in the conventional implementations of geoparsing and geotagging, in our treatment, geodetexting consists of two tasks: toponym recognition and toponym resolution, although other researchers (e.g., Marrero et al. [29]) add a classification task to the toponym recognition task which provides a type for the spatial entity (e.g., "Michigan" is a state, lake, etc.) and a localization task to the toponym resolution task which geocodes the spatial entity (e.g., the appropriate GPS coordinates). We argue that the evaluation measures for the recognition and resolution tasks are very different and depend on the nature of the underlying data that is being queried. In the case of geoparsing, there exist tools to make use of context such as the location from which one is querying as well as knowledge of prior queries to resolve ambiguously specified toponyms. For example, consider posing the query "Alexandria" when in College Park, MD on a number of different smartphones. Figure 2, shows the results for the Apple iOS5 Maps by Google (2a), iOS Maps by Google (2b), Android by Google (2c), Windows Phone on Windows 8 (2d), and Apple iOS6 and iOS7 (2e). Notice that all but Apple iOS6 and iOS7 place "Alexandria" in Virginia rather than in Egypt as done by Apple iOS6 and iOS7 (see [39] for a comparative study of smartphone mobile app mapping APIs). These decisions are made with the aid of large knowledge bases like Google's Knowledge Graph[1] and Knowledge Vault [10], and Microsoft's Satori Knowledge Base[2]. The availability of choices is also pointed out well in the series of ads by AT&T in Figure 1 about the importance of context in answering queries that involve a textual specification of a location. In contrast, in the case of geotagging, the ambiguity is resolved by using a number of factors such as the source document in which the location is specified or knowledge of the intended reader of the document in which the ambiguously specified location appears (reader or serving scope as described earlier). In addition use is made of knowledge of the remaining contents of the document in which the purported spatial reference is embedded, which generally is quite large in terms of the number of terms. In contrast, in the case of geoparsing, there may be no text in which the location specification is embedded.

[1] http://www.google.com/insidesearch/features/search/knowledge.html

[2] http://www.bing.com/blogs/site_blogs/b/search/archive/2013/03/21/satorii.aspx

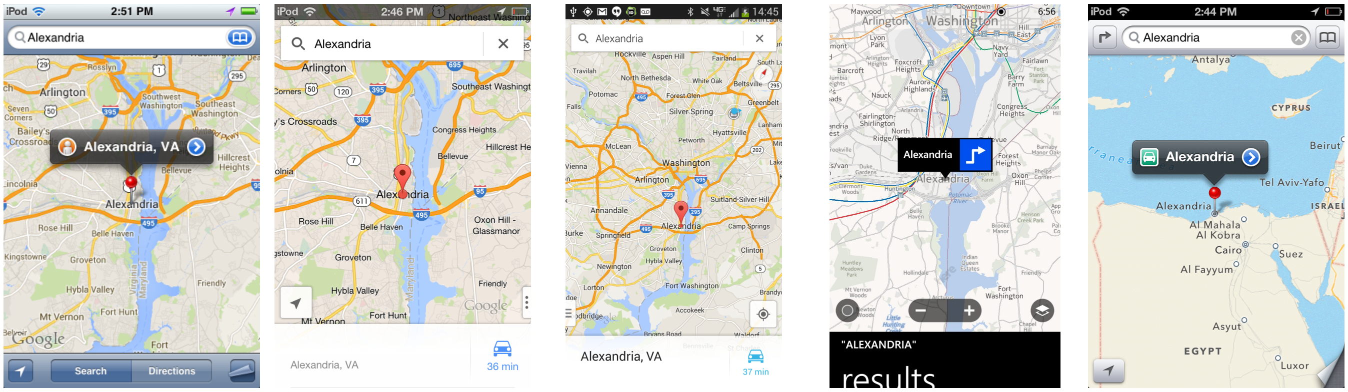
## 2. PRECISION AND RECALL FOR TOPONYM RECOGNITION AND RESOLUTION

Toponym recognition is just a matter of deciding whether or not a term corresponds to a toponym. It is evaluated using both precision and recall. For toponym recognition, the precision measure is the ratio of the number of times the identification of a term as a toponym is correct and the number of such identifications as toponyms. It does not take into account the number of times a term which is a toponym has failed to be identified/classified as a toponym. This is the role of the recall measure which is the ratio of the number of terms that have correctly been identified as toponyms and the number of terms that have been processed that are indeed toponyms. Determining which terms are indeed toponyms can be done by manually annotating the documents containing the toponyms, which may be a tedious process especially if there are many documents. Also it assumes that the user is capable of determining whether the term is used as a toponym.

In layman's terms, high precision means a low number of false positives, and high recall means a low number of false negatives. Intuitively, the false positives are the terms that are wrongly classified as toponyms, and the false negatives are the toponyms that were not classified as toponyms and hence missed, It is clear that we want to keep the numbers of both false positives and false negatives low.

Toponym resolution is much more complex than toponym recognition as it is a matter of determining the correct interpretation of a term that has been identified as a toponym. Thus the two processes are related in the sense that they are executed in sequence. Again, toponym resolution is evaluated in terms of both precision and recall. For toponym resolution, precision is the ratio of the number of times the identification of a term as a toponym by the toponym recognition process has been correctly resolved and the number of times that a term has been identified as a toponym by the toponym recognition process. The denominator includes terms that have been misidentified as toponyms by the toponym recognition process, but does not take into account the number of times that a term which is a toponym has failed to be identified/classified as a toponym by the toponym recognition process. Again, this is the role of the recall measure which is the ratio of the number of toponyms that have been correctly resolved and the number of terms that have been processed that are indeed toponyms, regardless of whether or not the toponym recognition process has classified them as toponyms. This means that if a toponym has not been recognized, then it is deemed as not being resolved correctly even though the toponym resolution process could have possibly resolved it where it given an opportunity to do so, and thus the toponym resolution recall rate could be lower than it need be. This definition of recall has been used in the comparative study in [22].

Thus it would seem that it may be appropriate to conduct a study of the efficacy of the toponym resolution process which is independent of the toponym recognition process in which case the only measure of interest is accuracy which corresponds to the number of toponyms that have been correctly resolved. Such a study is possible using the method of Lieberman and Samet [22] as the toponym recognition and resolution processes are independent. However, it is




(a) (b) (c) (d) (e)

Figure 2: Results for query string "Alexandria" when in "College Park, MD" (a) Apple iOS5 Maps by Google, (b) iOS Maps by Google, (c) Android Maps by Google, (d) Here Maps on Windows Phone, and (e) Apple iOS6 and iOS7 Maps,

not possible for many of the existing geodetexting systems such as Thomson Reuters's OpenCalais[3] and Yahoo's Placemaker[4] since they cannot resolve a toponym if they could not recognize it.

Toponym recognition is relatively simple in comparison with toponym resolution in the sense that in some implementations, the precision of toponym recognition is relatively high due to being overly cautious in recognizing toponyms thereby leading to more variation in the recall which is much lower, while in other implementations, like NewsStand, the precision is lower but the recall is much higher (e.g., [21] which compared NewsStand with Thomson Reuters's OpenCalais and Yahoo's Placemaker). On the other hand, in many implementations, while the precision of toponym resolution can be high, the recall of toponym resolution varies more and is usually lower than that of toponym recognition (e.g., [22] which compared NewsStand with Thomson Reuters's OpenCalais and Yahoo's Placemaker).

Our experience with the NewsStand system reveals that one of the problems with evaluating toponym resolution is that most of the time the toponyms are recognized and are not ambiguous. This is not surprising when one examines Gazetteers and notes how few of the entries have multiple interpretations, and, when they do, the number is not very high (e.g., [22]). We have first-hand experience here in that when we sampled the incoming news articles at random (either by topic or by location), we found the performance of the geodetexter to be quite good in the sense that the toponyms have most often been both recognized and resolved correctly. This is because we have evaluated the performance of the geodetexter on all occurrences of toponyms in a document or a document collection. However, in the case of ambiguous toponyms, many toponym resolution errors still occur. Thus ideally, what we would really like to do is evaluate the performance of the toponym resolution component of the geodetexter when restricted to ambiguous toponyms. This is because we have observed that there is extreme variability with individual toponyms due to the number of different interpretations of particular toponyms and the number of prominent interpretations of them that exist.

Once an evaluation of the performance of a particular geodetexter has been conducted, it is natural to ask how it fares in comparison to other geodetexters. This is best achieved by running the comparison on the same data known as a *ground truth*. It is usually accomplished with the aid of corpuses which are collections of representative documents for the context being evaluated. In the case of news, this could be a collection of articles which vary to account for different content and content providers in terms of factors including, but not limited to, geography, content focus, circulation, time, language, recentness, reader education level, and article length. As we see, the number of factors is large thereby requiring that the size of the corpus be large as well as making the process of compiling it nontrivial in terms of the needed labor expenditure. There is also the possibility that permission to use will not be granted by the content providers, which can be circumvented by the reference to a url, although this may work only for a limited time due to the risk that the url may go away with the passage of time, thereby hampering the attainment of our goal.

A number of studies of geodetexters have been conducted with different corpuses (e.g., [18,21,22]). The difficulty with all corpuses is that regardless of their absolute size in terms of the number of articles and the total amount of text that they contain, they are nevertheless relatively small for evaluating toponym resolution. In particular, the ambiguities that we wish to be able to handle do not occur with great frequency or on a regular basis. Therefore, they are not likely to occur in the text corpus unless the corpus compiler makes a deliberate effort to populate the corpus with articles with multiple occurrences of ambiguous toponyms which would be unlikely to occur in a corpus. On the other hand, most corpuses are fine for toponym recognition as virtually all documents invariably mention some locations and the identity of the locations is not a factor in the toponym recognition process assuming a sufficiently large and complete Gazetteer.

For example, in the case of news, the NewsStand system consistently ingests approximately 50,000 articles per day

[3] http://opencalais.com
[4] http://developer.yahoo.com/geo/placemaker

from about 10,000 sources. It has been running for six years and thus has ingested over 100 million news articles during this time. The corpuses and data sets used in most evaluations do not have such volumes of data. Nevertheless, even using one day of news is not enough to contain a sufficiently large number of ambiguous toponyms. Moreover, most of the corpuses contain static data while it is preferable to use dynamic data such as the news articles given that language and usage do change over time.

We suggest to evaluate toponym resolution only for ambiguous toponyms. Furthermore, we propose to do so for each interpretation of the toponym that exists in our data set, and only for these interpretations. These are the interpretations found by the application of a manual document annotation process regardless of whether the toponym recognition process has actually found them, The annotation involves a considerable amount of work as it requires human action and assumes that the user is capable of determining the correct interpretation. However, its advantage is that it can also be used to evaluate the toponym recognition process although resolving ambiguity is not part of toponym recognition. At times, doing such a complete job is impossible. In particular, this is the case when (1) sampling elements of the data set at random in search for toponyms, (2) annotating the entire data set is not feasible due to size and time, (3) not having access permission restrictions as being behind a pay wall, or (4) there exist quotas in terms of data volume as is the case when using such systems as Google Translator or Microsoft Translator to translate foreign language data.

In this case, we propose to resort to the next best alternative which assumes that the toponym recognition process is perfect in the sense of not missing any toponyms(i.e., 100% rate of recall). In particular, we gather all pairs of toponyms and interpretations that have been generated as a result of being recognized by the toponym recognition process and resolved by the toponym resolution process. For each such pair, we compute the precision and recall defined as follows.

Precision is the ratio of the number of times that the toponym-interpretation pair $(t, i)$ has been correctly resolved and the number of such pairs that have been generated as a result of being recognized by the toponym recognition process and resolved by the toponym resolution process.

Recall is defined as the ratio of the number of times that the toponym-interpretation pair $(t, i)$ has been correctly resolved and the number of such pairs that actually exist in the subset of the data set with which we are dealing (i.e., that we have recognized), The number of such pairs that exist is determined by a manual process such as the annotation described above. Thus we see that the denominators of the precision and recall are different whereas the numerators are the same. However, unlike when using the conventional definitions of precision and recall for toponym resolution, the above definitions of precision and recall may result in the recall rate being both greater than the precision rate (as when there are fewer actual instances of the toponym-interpretation pair) and also smaller than the precision rate (as when there are more actual instances of the toponym-interpretation pair).

It is important to note that these formulations of precision and recall are for a particular toponym-interpretation pair. Nevertheless, they can be used for the entire toponym resolution process by summing them and computing an average for both precision and recall. The advantage of this formulation is that it is not biased towards the high side by the myriad of correct toponym resolutions which occur often when one considers how many toponyms have just one interpretation in most Gazetteers, even large ones.

# 3. EXAMPLE OF HOW TO ACHIEVE AN EFFECTIVE 100% RATE OF RECALL FOR TOPONYM RESOLUTION

The formulations of precision and recall for toponym resolution presented in Section 2 that focus only on ambiguous toponyms can be incorporated in a practical setting to yield a retrieval process, based on textual specifications of locations, that, assuming 100% rate of recall for toponym recognition, has an effective 100% rate of recall for toponym resolution. We demonstrate how it can be achieved by explaining its implementation in NewsStand [40, 45]. This is a system built by us that enables retrieving news articles by determining the actual locations that are mentioned in them and then accessing them using a map query interface both on a desktop and on a smartphone [37, 41].

NewsStand crawls the web looking for RSS news feeds and collects the articles that they transmit. It determines the geographic locations mentioned in each article by applying an appropriate geotagging process and also tries to determine the geographic focus or foci (i.e., the key location(s) in the article). In addition, NewsStand aggregates news articles by topic based on content similarity (via use of a clustering method based on TF-IDF [36]) so that articles about the same news event are grouped into the same cluster, also, at times, referred to equivalently as a *topic* (for more details, see [23]). It can be sampled by going to `http://newsstand.umiacs.umd.edu`.

Figure 3 displays a screen shot of NewsStand's output to the example query "What is happening at location X on September 4, 2014", where, in this example, X corresponds to a location somewhere in Africa, Europe, and part of the Americas. The output is a map where each icon (i.e., symbol on the map), termed a *marker*, represents a set of articles on the same and/or different topics where the main property that is shared by all of the articles is that they are associated with (i.e., mention) the corresponding map location. The type of the icon conveys information about the news category into which the majority of the article topics associated with the location fall (e.g., general news, business, science and technology, entertainment, health, sports, etc.). It can be set/reset by toggling the appropriate buttons on the top of the screen shot).

Hovering over the name of a location on the map ("Moscow" in Figure 3) causes the generation of an info bubble containing the headline from a representative article on the dominant topic associated with Moscow, which on this date is the Ukraine/Russia crisis. These topics are obtained by the clustering process. The hovering action also causes the markers at all other locations on the map that are associated with this representative article to be replaced by orange balls. In this example, these locations correspond, in part, to the countries involved in, or affected by, the Ukraine/Russia crisis. Unfortunately, some of these locations may lie outside the geographic span of the map that is currently visible in the screen shot (e.g., in part of North America and the Far East). This is overcome by having the hovering action also cause the generation of a minimap which shows orange

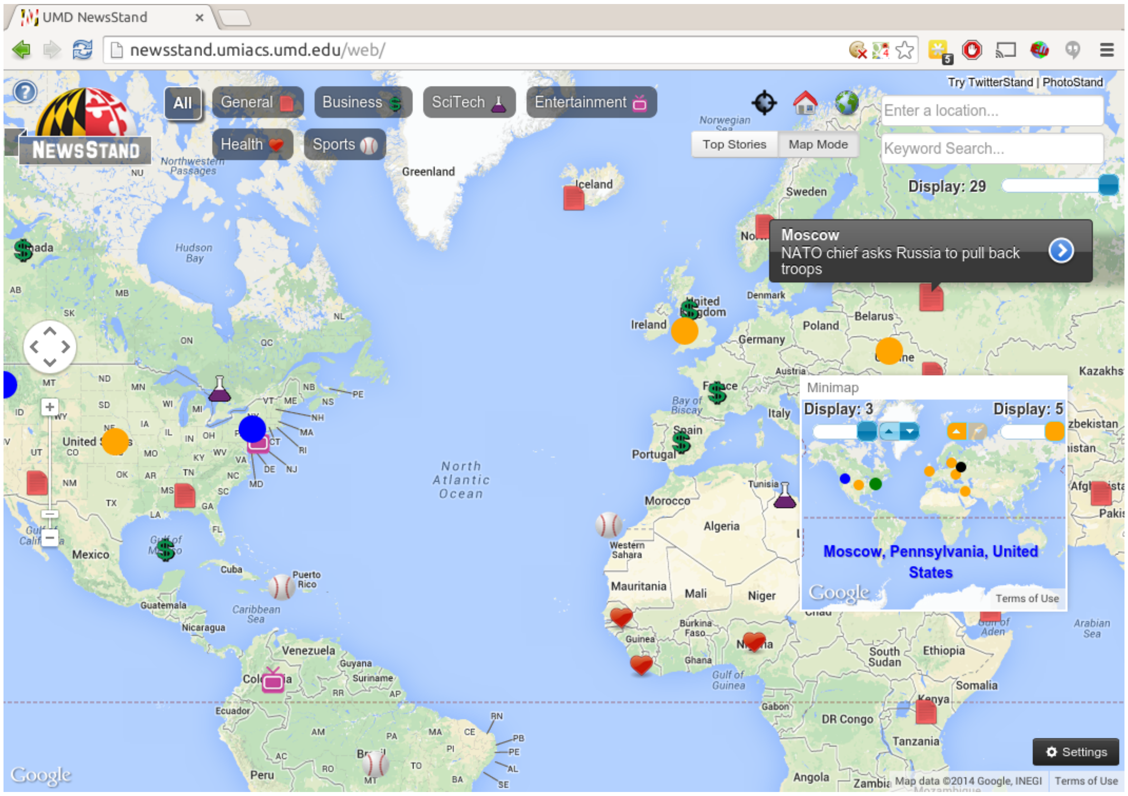



Figure 3: Example screen shot for the output of the query "What is happening at location X on September 4, 2014", where X corresponds to some location in Africa, Europe, and part of the Americas. The output is a representative headline for the Ukraine/Russia crisis at X=Moscow, Russia. Blue balls show alternative interpretations for Moscow in the states of Idaho and Pennsylvania in the USA. Orange balls show other locations associated with the Ukraine/Russia crisis. The minimap summarizes the orange and blue balls for the entire world while highlighting the "Pennsylvania" interpretation of "Moscow".

balls at all of the locations in the world that are associated with the representative article. This lets users see easily the selected geographic focus, without having to leave their area of browsing interest on the main map, and is independent of the current level of zoom.

The hovering action over the name $n$ of a location $l$ also causes the generation of blue balls at all other locations $k$ with the same name $n$ on both the map and minimap, whenever at least one article cluster is associated with interpretation $k$ of $n$. This enables us to detect quickly toponym resolution errors by providing access to all articles determined to mention a particular location name $n$ for any interpretation $k$ of $n$ where this interpretation is made by the toponym resolution process. This is the case as long as at least one article is associated with interpretation $k$ even though $k$ may not be the correct interpretation of the occurrence of $n$ for the article cluster in question, thereby letting the user make the final decision. Effectively, by enabling the user to examine all mentions of $n$ for the correct interpretation (assuming 100% rate of recall for toponym recognition with lower precision as there could be many terms that were wrongly classified as toponyms), we have 100% rate of recall for toponym resolution for the interpretations of a location that are in our Gazetteer. Of course, the precision of toponym resolution will be lower on account of taking all of the interpretations into account in forming the denominator of the precision, but at least we do not miss any. Notice that in some sense we are ranking our article clusters where the highest ranked one is associated with the queried location on the main map and the lower ranked ones are associated with the locations corresponding to the blue balls on the minimap.

The minimap in Figure 3 shows the how the multiple interpretations of the toponym Moscow are handled here. A black ball on the minimap marks the location over which we are currently hovering (i.e., Moscow in Russia). Up and down arrows on the minimap enable scrolling through the orange and blue balls and outputting the corresponding location names. Scrolling through the blue balls enables seeing all of the interpretations of the location name determined by NewsStand. The primary interpretation is the one anchoring the minimap, namely in Russia for this example. The order in which the remaining interpretations are presented by the scrolling process reflects the confidence (from high to low) in the interpretation. Green and red balls on the minimap correspond to the current blue and orange balls, respectively, in the scrolling process. Hovering in the minimap over an orange ball yields the name of the location while hovering over a blue ball yields both the name of the location and its containing location on the minimap (e.g., "Moscow, Pennsylvania, United States" in the figure) which is needed to differentiate among them as all blue balls have the same name. Figure 4 is a screenshot for the same query as in Figure 3 except that it results from hovering over the location "Moscow, Pennsylvania". The figure shows an info bubble with headlines from a number of clusters associated with it. Interestingly, the text of the constituent articles only contains the term "Moscow" with no qualifying information, yet the toponym resolution process has correctly identified them as being associated with "Moscow, Pennsylvania" by use of of context—that is, the articles refer to other locations in "Pennsylvania" such as "Scranton".

The advantage of the minimap is that it acts as a very compact summary of all possible interpretations made by NewsStand for a particular toponym. Thus, users have access to all possible data that is associated with an ambiguously specified toponym. Moreover, if the toponym resolution process makes an error in choosing the correct interpretation, then the user can still find it by going through all of the interpretations that are presented with the blue balls. The only problem is that the order in which the user processes the various interpretations may not be the most optimal as the user will have to process a number of irrelevant interpretations. In other words, the toponym resolution process has not ranked the interpretations in an optimal manner for this particular toponym. Notice the similarity to the way in which the Google search engine, or any other search engine, returns results for a search query. In particular, the search engine returns many (actually too many) results and ranks them according to some formulas that it thinks yield the most useful results first although they may not be ranked in the optimal order for the particular user. Nevertheless, all of the results are present, its just that the user will have to go through all of them, which takes time.

## 4. CONCLUDING REMARKS

Assuming a 100% rate of recall for toponym recognition, we have shown how to achieve an effective 100% rate of recall for toponym resolution by the addition of a minimap that shows all interpretations for a textual specification of a location which are associated with at least one document. This means that a user has access to all documents that mention a specific location as long as the textual specification to the location has been recognized as a location rather than as the name of another entity such as a person, company, organization, etc. It is analogous to the long list of relevant documents that are presented to a user in response to queries to a conventional search engine through which the user must wade. The interesting aspect here is that the map is a very concise representation of the choices whereas in the response to a query to a conventional search engine, there is no way to get an overview of the possible responses other than to page through them screen by screen. In the case of the minimap, users could also eventually make use of a zoom operation to get more interpretations if such a high volume exists but this is rare as although there are many interpretations for a location, it is rare for more than just a few to have been actually mentioned in some documents. This is especially the case for specific news sources which do not span all locations on the globe.

## 5. REFERENCES


[1] M. D. Adelfio, M. D. Lieberman, H. Samet, and K. A. Firozvi. Ontuition: Intuitive data exploration via ontology navigation. In *GIS*,pages 540–541, San Jose, CA, 2010.

[2] M. D. Adelfio and H. Samet. GeoWhiz: Using common categories for toponym resolution. *SIGSPATIAL*, pages 542–545, Orlando, FL, 2013.

[3] M. D. Adelfio and H. Samet. Schema extraction for tabular data on the web. *PVLDB*, 6(6):421–432, 2013.

[4] M. D. Adelfio and H. Samet. Structured toponym resolution using combined hierarchical place categories. In *GIR*, pages 49–56, Orlando, FL, 2013.

[5] E. Amitay, N. Har'El, R. Sivan, and A. Soffer. Web-a-Where: Geotagging web content. In *SIGIR*, pages 273–280, Sheffield, United Kingdom, 2004.

[6] D. Buscaldi and B. Magnini. Grounding toponyms in an Italian local news corpus. In *GIR*, article 15, Zurich, Switzerland, 2010.

[7] D. Buscaldi and P. Rosso. A conceptual density-based approach for the disambiguation of toponyms. *IJGIS*, 22(3):301–313, 2008.

[8] P. Clough. Extracting metadata for spatially-aware information retrieval on the internet. In *GIR*, pages 25–30, Bremen, Germany, 2005.

[9] J. Ding, L. Gravano, and N. Shivakumar. Computing geographical scopes of web resources. In *VLDB*, pages 545–556, Cairo, Egypt, 2000.

[10] X. L. Dong, E. Gabrilovich, G. Heitz, W. Horn, N. Lao, K. Murphy, T. Strohmann, S. Sun, and W. Zhang. Knowledge Vault: A web-scale approach to probabilistic knowledge fusion. In *SIGKDD*, New York, 2014.

[11] C. Esperança and H. Samet. Experience with SAND/Tcl: a scripting tool for spatial databases. *JVLC*, 13(2):229–255, 2002.

[12] E. Garbin and I. Mani. Disambiguating toponyms in news. In *HLT*, pages 363–370, Morristown, NJ, 2005.

[13] J. Gelernter and S. Balaji. An algorithm for local geoparsing of microtext. *GeoInformatica*, 17(4):635–667, 2013.

[14] N. Gramsky and H. Samet. Seeder finder - identifying additional needles in the Twitter haystack. In *LBSN*, pages 44–53, Orlando, FL, 2013.

[15] A. Jackoway, H. Samet, and J. Sankaranarayanan. Identification of live news events using Twitter. In *LBSN*, pages 25–32, Chicago, 2011.

[16] C. B. Jones, R. S. Purves, P. D. Clough, and H. Joho. Modelling vague places with knowledge from the Web. *IJGIS*, 22(10):1045–1065, 2008.

[17] W. Kienreich, M. Granitzer, and M. Lux. Geospatial anchoring of encyclopedia articles. In *IV*, pages 211–215, London, 2006.

[18] J. L. Leidner. *Toponym Resolution in Text: Annotation, Evaluation and Applications of Spatial Grounding of Place Names*. PhD thesis, University of Edinburgh, Edinburgh, Scotland, UK, 2006.

[19] J. L. Leidner, G. Sinclair, and B. Webber. Grounding spatial named entities for information extraction and question answering. In *HLT-NAACL 2003 Workshop on Analysis of Geographic References*, pages 31–38, Edmonton, Canada, 2003.

[20] H. Li, R. K. Srihari, C. Niu, and W. Li. InfoXtract location normalization: a hybrid approach to geographic references in information extraction. In *HLT-NAACL 2003 Workshop on Analysis of Geographic References*, pages 39–44, Edmonton, Canada, 2003.

[21] M. D. Lieberman and H. Samet. Multifaceted toponym recognition for streaming news. In *SIGIR*, pages 843–852, Beijing, China, 2011.

[22] M. D. Lieberman and H. Samet. Adaptive context features for toponym resolution in streaming news. In *SIGIR*, pages 731–740, Portland, OR, 2012.

[23] M. D. Lieberman and H. Samet. Supporting rapid processing and interactive map-based exploration of streaming news. In *GIS*, pages 179–188, Redondo Beach, CA, 2012.

[24] M. D. Lieberman, H. Samet, and J. Sankaranarayanan. Geotagging: Using proximity, sibling, and prominence clues to understand comma groups. In *GIR*, article 6, Zurich, Switzerland, 2010.

[25] M. D. Lieberman, H. Samet, and J. Sankaranarayanan. Geotagging with local lexicons to build indexes for textually-specified spatial data. In *ICDE*, pages 201–212, Long Beach, CA, 2010.

[26] M. D. Lieberman, H. Samet, J. Sankaranarayanan, and J. Sperling. STEWARD: architecture of a spatio-textual search engine. In *GIS*, pages 186–193, Seattle, WA, 2007.

[27] M. D. Lieberman, H. Samet, J. Sankaranarayanan, and J. Sperling. Spatio-textual spreadsheets: Geotagging via spatial coherence. In *GIS*, pages 524–527, Seattle, WA, 2009.

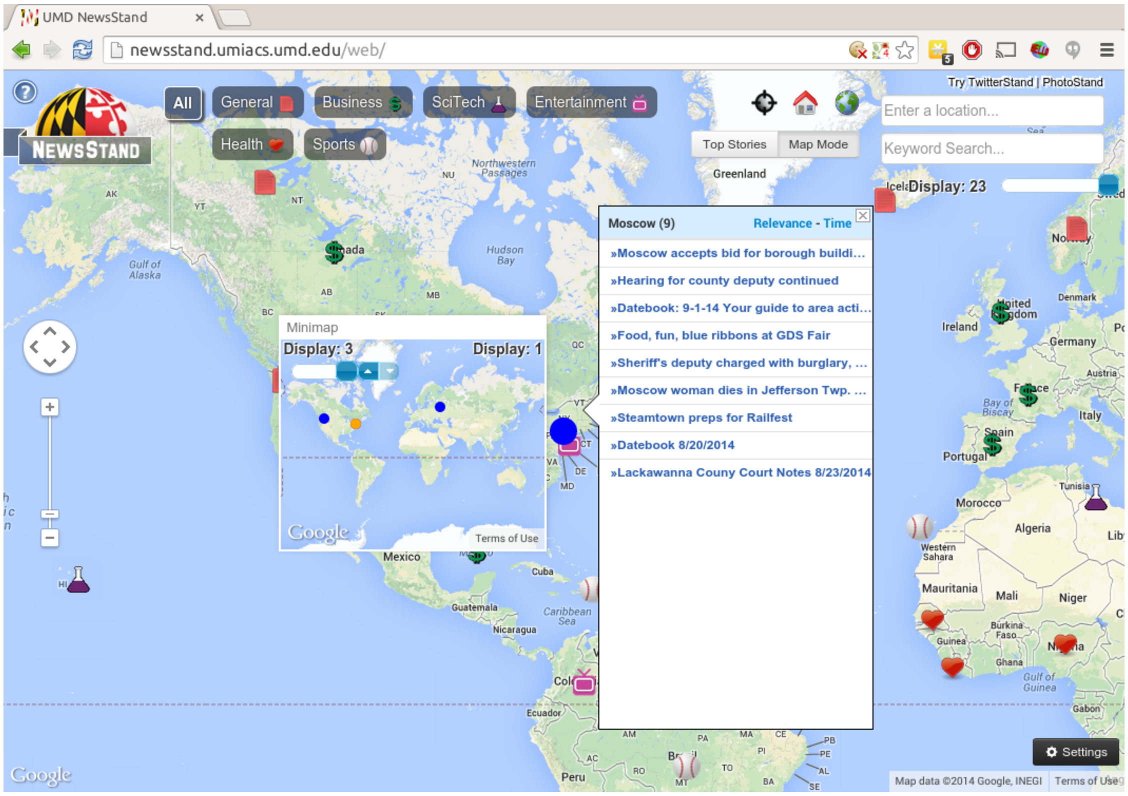



Figure 4: Example screen shot for the output of the query "What is happening at location X on September 4, 2014" where X corresponds to some location in Africa, Europe, and part of the Americas. The output is a set of headlines for article clusters for X=Moscow, Pennsylvania which results from hovering over it and clicking on the rightward pointing arrow in the info bubble at the location. Blue balls show alternative interpretations for "Moscow" in "Russia" and in the state of "Idaho". There are no orange balls as the articles associated with this location are very local.


[28] A. Markowetz, T. Brinkhoff, and B. Seeger. Exploiting the internet as a geospatial database. In *NG2I*, Cambridge, MA, 2003.

[29] M. Marrero, S. Sanchez-Cuadrado, J. M. Lara, and G. Andreadakis. Evaluation of named entity extraction systems. *Adv. in Comp. Ling., Special Issue in Res. in Comp. Sci.*, 41:47–58, 2009.

[30] B. Martins, H. Manguinhas, and J. Borbinha. Extracting and exploring the geo-temporal semantics of textual resources. In *ICSC*, pages 1–9, Santa Clara, CA, 2008.

[31] R. S. Purves, P. Clough, C. B. Jones, A. Arampatzis, B. Bucher, D. Finch, G. Fu, H. Joho, A. K. Syed, S. Vaid, and B. Yang. The design and implementation of SPIRIT: a spatially aware search engine for information retrieval on the internet. *IJGIS*, 21(7):717–745, 2007.

[32] T. Qin, R. Xiao, L. Fang, X. Xie, and L. Zhang. An efficient location extraction algorithm by leveraging web contextual information. In *GIS*, pages 53–60, San Jose, CA, 2010.

[33] G. Quercini, H. Samet, J. Sankaranarayanan, and M. D. Lieberman. Determining the spatial reader scopes of news sources using local lexicons. In *GIS*, pages 43–52, San Jose, CA, 2010.

[34] C. Sallaberry, M. Gaio, J. Lesbegueries, and P. Loustau. A semantic approach for geospatial information extraction from unstructured documents. In *The Geospatial Web*, pages 93–104. Springer, London, 2007.

[35] G. Salton. *Automatic Text Processing: The Transformation Analysis and Retrieval of Information by Computer*. Addison-Wesley, Reading, MA, 1989.

[36] G. Salton and C. Buckley. Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5):513–523, 1988.

[37] H. Samet, M. D. Adelfio, B. C. Fruin, M. D. Lieberman, and B. E. Teitler. Porting a web-based mapping application to a smartphone app. In *GIS*, pages 525–528, Chicago, 2011.

[38] H. Samet, H. Alborzi, F. Brabec, C. Esperança, G. R. Hjaltason, F. Morgan, and E. Tanin. Use of the SAND spatial browser for digital government applications. *CACM*, 46(1):63–66, 2003.

[39] H. Samet, B. C. Fruin, and S. Nutanong. Duking it out at the smartphone mobile app mapping API corral: Apple, Google, and the competition. In *MOBIGIS*, Redondo Beach, CA, 2012.

[40] H. Samet, J. Sankaranarayanan, M. D. Lieberman, M. D. Adelfio, B. C. Fruin, J. M. Lotkowski, D. Panozzo, J. Sperling, and B. E. Teitler. Reading news with maps by exploiting spatial synonyms. *CACM*, 57(10):64–77, 2014.

[41] H. Samet, B. E. Teitler, M. D. Adelfio, and M. D. Lieberman. Adapting a map query interface for a gesturing touch screen interface. In *WWW (Companion Volume)*, pages 257–260, Hyderabad, India, 2011.

[42] J. Sankaranarayanan, H. Samet, B. Teitler, M. D. Lieberman, and J. Sperling. TwitterStand: News in tweets. *GIS*, pages 42–51, Seattle, WA, 2009.

[43] C. A. Shaffer, H. Samet, and R. C. Nelson. QUILT: a geographic information system based on quadtrees. *IJGIS*, 4(2):103–131, 1990.

[44] J. Strötgen, M. Gertz, and P. Popov. Extraction and exploration of spatio-temporal information in documents. In *GIR*, article 16, Zurich, Switzerland, 2010.

[45] B. Teitler, M. D. Lieberman, D. Panozzo, J. Sankaranarayanan, H. Samet, and J. Sperling. NewsStand: A new view on news. In *GIS*, pages 144–153, Irvine, CA, 2008.

[46] C. Wang, X. Xie, L. Wang, Y. Lu, and W.-Y. Ma. Detecting geographic locations from Web resources. In *GIR*, pages 17–24, Bremen, Germany, 2005.

[47] W. Zong, D. Wu, A. Sun, E.-P. Lim, and D. H.-L. Goh. On assigning place names to geography related web pages. In *JCDL*, pages 354–362, New York, 2005.