Report on requirement analysis and design of PAUSANIAS

Akrivi Vlachou, Postdoctoral Researcher

George Tsatsanifos, PhD Student

Abstract

Nowadays an increasing amount of web-accessible information on spatial objects becomes available to the public every day. Apart from the spatial location of an object (e.g., a point of interest), additional descriptive information typically includes textual description as well as various ratings, often user generated. Modern applications employ spatio-textual queries, which take into account both the spatial location of an object and its textual similarity to retrieve the most relevant objects. However, existing applications provide a limited functionality to the users. For example, several meaningful queries cannot be expressed by existing approaches and motivate our novel prototype system. The goal of our research is to introduce a novel framework, called Pausanias, for supporting ranked spatial-keyword search over web-accessible geotagged data. In this deliverable, we provide an anal-

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union European Social Fund

OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING investing in knowledge society

MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS MANAGING AUTHORITY

NSRF 2007-2013 programme for development EUROPEAN SOCIAL FUND

Co- financed by Greece and the European Union

ysis of requirements that should be fulfilled by the proposed prototype and a concise description of the system design of Pausanias.

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

Contents

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union European Social Fund OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING investing in knowledge society MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS MANAGING AUTHORITY NSRF 2007-2013 programme for development EUROPEAN SOCIAL FUND Co- financed by Greece and the European Union

1 Introduction

Search engines, such as Google and Yahoo!, provide efficient retrieval and ranking of web pages based on queries consisting of a set of given keywords. Recent studies show that 20% of all Web queries also have location constraints, i.e., also refer to the location of a geotagged web page. An increasing number of applications support location-based keyword search, including Google Maps, Bing Maps, Yahoo! Local, and Yelp. Such applications depict points of interest on the map and combine their location with the keywords provided by the associated document(s). The posed queries consist of two conditions: a set of keywords and a spatial location. The goal is to find points of interest with these keywords close to the location. We refer to such a query as spatial-keyword query. Moreover, mobile devices nowadays are enhanced with built-in GPS receivers, which permits applications (such as search engines or yellow page services) to acquire the location of the user implicitly, and provide location-based services. For instance, Google Mobile App provides a simple search service for smartphones where the location of the user is automatically captured and employed to retrieve results relevant to her current location. As an example, a search for pizza results in a list of pizza restaurants nearby the user. Given the popularity of spatial-keyword queries and their wide applicability in practical scenarios, it is critical to (i) establish mechanisms for efficient processing of spatial-keyword queries, and (ii) support more expressive query formulation by means of novel query

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union
European Social Fund
OPERATIONAL PROGRAMME
EDUCATION AND LIFELONG LEARNING
investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
Co- financed by Greece and the European Union
NSRF 2007-2013
programme for development
EUROPEAN SOCIAL FUND

types. Although studies on both keyword search and spatial queries do exist, the problem of combining the search capabilities of both simultaneously has received little attention.

During our research, we have identified the shortcomings and limitations of existing approaches for spatial-keyword search. Such limitations belong to two broad categories: (a) performance-related, and (b) functionality-related. Our prototype will address both categories. Shortcomings in the first category that affect the performance of existing techniques, will be addressed by our novel indexing technique. We will evaluate different metrics to measure the performance, including execution time, space requirements for index structures, maintenance overhead in case of dynamic data. The second category contains additional limitations that relate to the search capabilities and the functionality provided by existing systems. Our novel prototype system handles meaningful queries that cannot be expressed by existing approaches. In the following, we identify the requirements that our prototype should fulfill in order to provide novel and useful spatial-keyword search capabilities, while at the same time advancing the state-of-the-art significantly.

2 Requirement analysis

An increasing number of applications support location-based queries, which retrieve the most interesting spatial objects based on their geographic location. Recently, spatio-textual queries have attracted much attention, as such

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union European Social Fund
OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
Co- financed by Greece and the European Union
NSRF 2007-2013 programme for development
EUROPEAN SOCIAL FUND

queries combine location-based retrieval with textual information that describes the spatial objects. Most of the existing queries only focus on retrieving objects that satisfy a spatial constraint ranked by their spatial-textual similarity to the query point. However, users are quite often interested in spatial objects (*data objects*) based on the quality of other facilities (*feature objects*) that are located in their vicinity. Such features objects are typically described by non-spatial numerical attributes such as quality or ratings, in addition to the textual information that describes their characteristics. For example, several approaches [4, 3, 5, 6] focus on ranking the data objects based on their spatio-textual similarity to a query point and some keywords. Our work focuses on ranking the data objects based on the quality and relevance of the facilities in their spatial neighborhood.

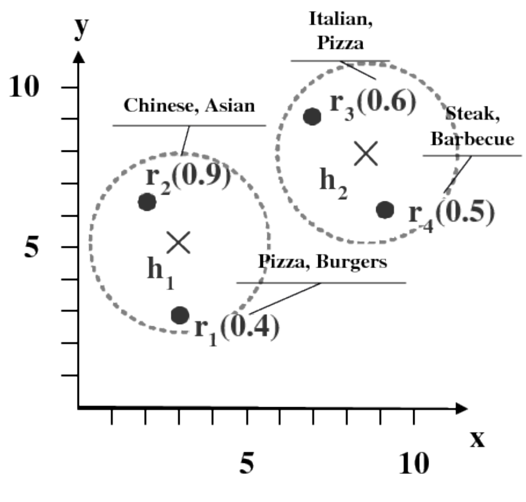

Figure 1: Example.

In our scenario, we have objects of different types such as hotels or restau-

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

rants. Each of them has a location and some keywords that characterize them. In addition, the objects have also numerical attributes such as price or rating. In the simplest case, the user is interested only in one type of feature such as hotel, but want to combine all different criteria, i.e., for example consider a tourist that is looking for a hotel with a swimming pool near the city center and that has a low price. Current approaches, focus only on the spatio-textual aspect or on ranking of the object based on numerical values. In the more generic case, more that one type of objects may be combined. Consider for example, a tourist that looks for *"hotels that have nearby a good* ***Italian*** *restaurant that serves* ***pizza****"*. Fig. 2 depicts a spatial area containing hotels (data objects) and restaurants (feature objects). The quality of the restaurants based on existing reviews is depicted next to the restaurant. Each restaurant also has textual information, such as pizza or steak, which describes additional characteristics of the restaurant. The tourist specifies also a spatial constraint (in the figure depicted as a range around each hotel) to restrict the distance of the restaurant to the hotel. Obviously, the hotel h_2 is the best option for a tourist that poses the aforementioned query. In the general case, more than one type of feature objects may exist in order to support queries such as *"hotels that have nearby a good* ***Italian*** *restaurant that serves* ***pizza*** *and a cheap coffeehouse that serves* ***muffins****"*. Even though spatial preference queries have been studied before [8, 9, 7], their definition ignores the available textual information. In our example, the spatial preference query would correspond to a tourist that searches for *"hotels that*

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union
European Social Fund
OPERATIONAL PROGRAMME
EDUCATION AND LIFELONG LEARNING
investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
Co- financed by Greece and the European Union
NSRF 2007-2013
programme for development
EUROPEAN SOCIAL FUND

are nearby a good restaurant" and the hotel h_1 would always be retrieved, irrespective of the textual information. To this end, we propose a novel and more expressive query type, called *spatio-textual preference query*, for ranked retrieval of data objects based the textual relevance and the non-spatial score of feature objects in their spatial neighborhood.

Furthermore, a main difference of top-k spatio-textual preference queries compared to traditional spatial preference queries [8, 9, 7] is that the rank of a data object changes depending on the query keywords, which renders techniques [7] that rely on materialization inappropriate. Moreover, processing spatial preference queries is costly in terms of both I/Os and execution time [8, 9], because it may require searching the spatial neighborhood of all data objects before reporting the top-k. Thus, extending spatial preference queries for supporting also textual information is challenging, since the new query type is more expensive due to the overhead imposed by the similarity of the query keywords to the facilities' textual descriptions.

To summarize our prototype will support queries based on three different type of information: i) spatial information ii) keywords and iii) numerical attributes. The combination of those three types of information enables expressing many meaningful queries. This is a main difference to existing approaches [4, 3, 5, 6] that focus on ranking the data objects based on their spatio-textual similarity to a query point and some keywords. Moreover, the user of our novel system, is able to pose queries that combine different types of data objects, such as hotels and restaurants. Current approaches [1,

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union European Social Fund
OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
Co- financed by Greece and the European Union
NSRF 2007-2013 programme for development
EUROPEAN SOCIAL FUND

10, 2] focus on finding a set of data objects that are close to each other and relevant to a given query, whereas in our approach we rank the data objects based on the facilities in their spatial neighborhood. In addition, our prototype also supports queries based on the current approaches that support only spatial and textual information. Thus, we conclude that our prototype will provide to the user more enhanced search capabilities than provided by existing approaches.

3 The Web-Interface

In Figure 2 we present what the web-interface of the Pausanias web-application looks like. First and foremost, the interface consists of a map object which is accompanied by a form. The user can specify the appropriated keywords separated by either commas or white spaces.

In Figure 3, we present an example of enacting a spatio-textual query for restaurants serving “pizza” and “burger” and hotels with “internet” and “sauna” that are also located around the Grand Central Terminal in NYC. We note that only the restaurants and the hotels that are within 1 mile distance from the train station are qualified and from these the ones that match best the user’s description and also are highly ranked (e.g. a five-star hotel is always a better choice than a four-star hotel when both hotels are equally similar to the user’s keywords) are illustrated in the browser along with relevant information. The user must explicitly express by clicking on

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program ”Education and Lifelong Learning” (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union European Social Fund
OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
NSRF 2007-2013 programme for development
EUROPEAN SOCIAL FUND
Co- financed by Greece and the European Union

Figure 2: The web-interface of an application leveraging spatio-textual search.

the appropriate radio button the type of search he desires, range search in our case that is.

Regarding the influence operator (the last option available in the form), it facilitates search which is performed according to criteria that combine textual similarity and rank score with distance. More specifically, the score achieved by a particular location diminishes with distance, and thus, matching hotels and restaurants that are also close to the specified areas are more preferable. The mechanism that incorporates both criteria is explained thoroughly in the following section.

Furthermore, we have paid special attention on our site being as user-friendly as possible. For example, you can easily notice in Figure 3 that all

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union European Social Fund OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING investing in knowledge society MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS MANAGING AUTHORITY Co- financed by Greece and the European Union NSRF 2007-2013 programme for development EUROPEAN SOCIAL FUND

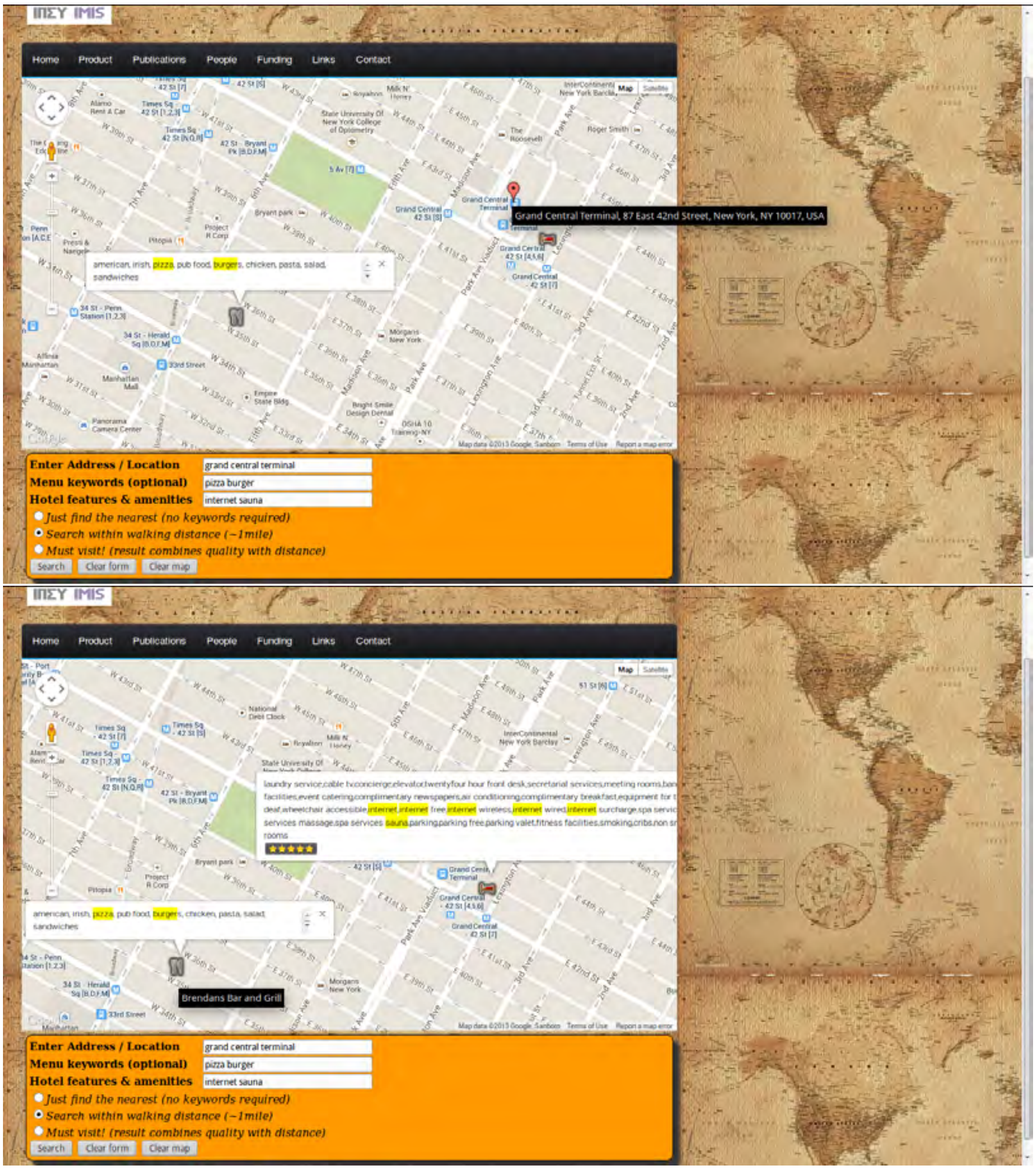

Figure 3: An example of a spatio-textual range search.

matched keywords are highlighted. We have also chosen different markers for each feature class (e.g. restaurants and hotels) denoting their nature. Addi-

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union
European Social Fund
OPERATIONAL PROGRAMME
EDUCATION AND LIFELONG LEARNING
investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
Co-financed by Greece and the European Union
NSRF 2007-2013
programme for development
EUROPEAN SOCIAL FUND

tionally, when the user hovers or clicks on a marker we provide information associated with this particular landmark/location. At any time, if the user makes a mistake, for example he clicks by mistake on the map and placing a marker on a landmark he is not actually interested in, he can reset the form or remove the markers from the map by clicking on the appropriate buttons of the form, namely “Clear form” and “Clear map”.

We have made special considerations towards enhancing the performance of the application. For instance, in Figure 4 the increased number of specified query landmarks by the user has no or little impact on performance. In particular, we processed and geolocated off-line all data in order to derive additional information, such as the address given the latitude and the longitude of a certain landmark. Of course, this could take place ad-hoc as the user issues his queries, but the Google Maps API imposes certain limitations on its resources and services, e.g. no more than four or five geolocation requests can be processed per second, in order to limit the consumed bandwidth per user, prevent DDOS attacks, etc.

4 Architecture and Implementation

Our implementation leverages conventional principles for modular construction of enterprise applications. A three-tier architecture is a client-server architecture in which the functional process logic, data access, computer data storage and user interface are developed and maintained as independent

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union European Social Fund
OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
Co- financed by Greece and the European Union
NSRF 2007-2013 programme for development
EUROPEAN SOCIAL FUND

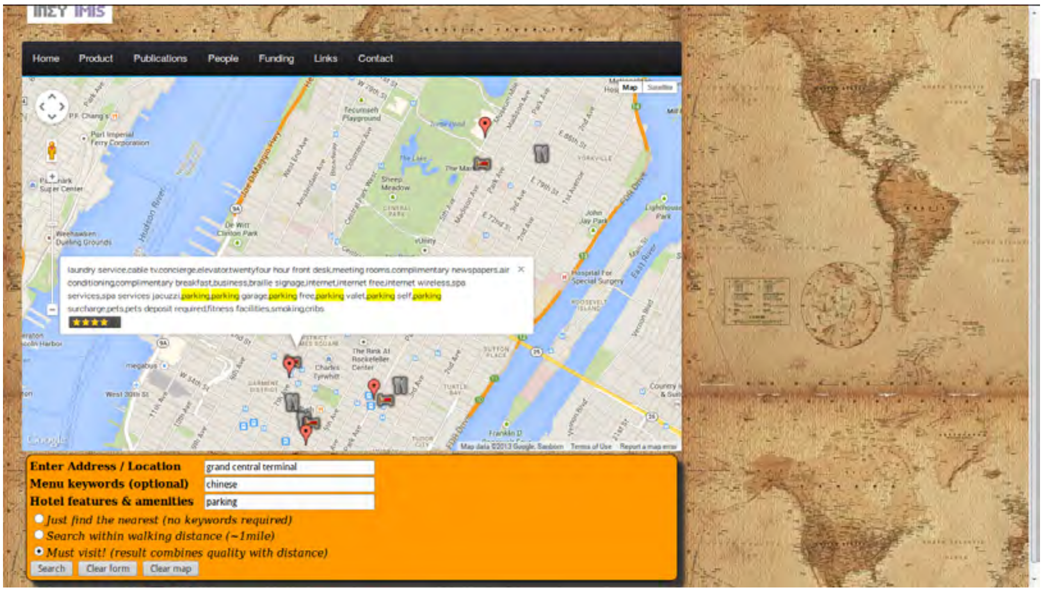

Figure 4: Processing simultaneously numerous markers.

modules on separate platforms. Three-tier architecture is a software design pattern and a well-established software architecture. Its three tiers are the presentation tier, application tier and data tier. Three-tier architecture allows any one of the three tiers to be upgraded or replaced independently. The user interface is implemented on a desktop PC and uses a standard graphical user interface with different modules running on the application server. The relational database management system on the database server contains the computer data storage logic. The middle tiers are usually multitiered. As shown in Figure 5, the three tiers in a three-tier architecture are:

Presentation Tier Occupies the top level and displays information related to services available on a website. This tier communicates with other tiers by sending results to the browser and other tiers in the network.

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union
European Social Fund
OPERATIONAL PROGRAMME
EDUCATION AND LIFELONG LEARNING
investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
Co- financed by Greece and the European Union
NSRF
2007-2013
programme for development
EUROPEAN SOCIAL FUND

Application Tier Also called the middle tier, logic tier, business logic, or service tier, this tier is pulled from the presentation tier. It controls application functionality by performing detailed processing.

Data Tier Houses database servers where information is stored and retrieved. Data in this tier is kept independent of application servers or business logic.

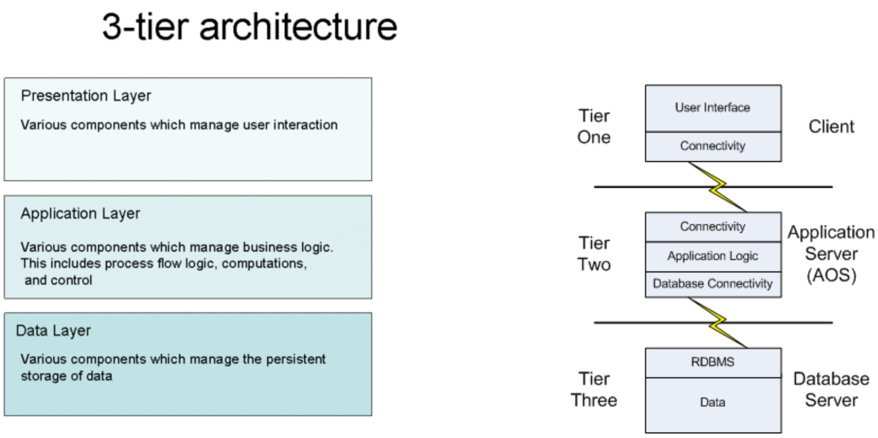

Figure 5: A three-tier architecture example.

4.1 The Presentation Layer

In terms of launching the website, we used exclusively open-source software. In particular, we used Java Server Pages (JSP) for the top-level of our web-application. To elaborate, Java Server Pages (JSP) is a technology for developing web pages that support dynamic content which helps developers insert java code in HTML pages by making use of special JSP tags. A Java Server

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union European Social Fund
OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
NSRF 2007-2013 programme for development
EUROPEAN SOCIAL FUND
Co- financed by Greece and the European Union

Pages component is a type of Java servlet that is designed to fulfill the role of a user interface for a Java web application. JSPs code combines HTML, XHTML code, XML elements, and embedded JSP actions and commands. Using JSP, you can collect input from users through web page forms, present records from a database or any other source, and create web pages dynamically. JSP tags can be used for a variety of purposes, such as retrieving information from a database or registering user preferences, accessing Java Beans components, passing control between pages and sharing information between requests, pages etc.

JSP development involves writing code of simple syntax to handle the tools for manipulating the elements of the service layer, namely the Enterprise Java Beans (EJB) launched in the application server. A scriptlet can contain any number of Java language statements, variable or method declarations, or expressions that are valid in the page scripting language. A scriptlet can be either of the form `<% code fragment%>`, or the XML equivalent `<jsp:scriptlet> code fragment </jsp:scriptlet>`. Any text, HTML tags, or JSP elements must be outside the scriptlet. A JSP declaration declares one or more variables or methods that you can use in Java code later in the JSP file. It can be either of the form `<%! declaration; [ declaration; ]+ ... %>`, or `<jsp:declaration> code fragment </jsp:declaration>`. A JSP expression element contains a scripting language expression that is evaluated, converted to a String, and inserted where the expression appears in the JSP file. More importantly, because the value

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union European Social Fund
OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS MANAGING AUTHORITY
NSRF 2007-2013 programme for development EUROPEAN SOCIAL FUND
Co- financed by Greece and the European Union

of an expression is converted to a String, you can use an expression within a line of text, whether or not it is tagged with HTML, in a JSP file. In our code we make extensive use of JSP expressions in order to produce dynamically JavaScript and HTML code mostly, given the user's queries and preferences. JSP expressions are of the form `<%= expression%>` , or `<jsp:expression> expression </jsp:expression>`, and can include any type of a Java statement or a function call that returns valid code. Next, JSP actions use constructs in XML syntax to control the behavior of the servlet engine. You can dynamically insert a file, reuse JavaBeans components, forward the user to another page, or generate HTML for the Java plug-in. They are of the form `<jsp:action_name attribute=''value'' />`, and they essentially constitute predefined functions. Last but not least, we make use at many occasions of a control flow statement, either for decision making and branching, or repeating actions and dynamically generated HTML code.

Furthermore, our web-application also uses the Google Maps API. Google has developed since 2005 a service for retrieving map images, and web services for performing geocoding, generating driving directions, and obtaining elevation profiles. In fact, well over 350,000 web sites use the Google Maps API, making it the most heavily used web application development API today, today into its third major version. It is a free service, and currently does not contain advertisements. By using the Google Maps API, it is possible to embed Google Maps site into an external website like ours, on to which site specific data can be overlaid. In addition, our JavaScript code which is

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union European Social Fund
OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
Co- financed by Greece and the European Union
NSRF 2007-2013 programme for development
EUROPEAN SOCIAL FUND

executed in the user's browser is used to manage the different overlays of the map, place markers and set their properties, (e.g. icons, animations, attached messages), register event handlers and listeners to handle events like clicking on the map, initialize or reset the map and the form, and others. In some cases JavaScript is disabled by some users in an attempt to protect themselves from its security vulnerabilities and malicious code. Unfortunately, we have no way of treating this impediment currently, and most features of our application are unavailable to those users.

4.2 The Service Layer

Furthermore, the application server hosting our website is Apache Tomcat 7 installed on a Debian-based virtual machine of the GRNET cloud (`https://okeanos.grnet.gr/`). This is where the service tier is located. Towards this end, we created the appropriate Enterprise Java Session Beans (EJB) that encapsulate the business logic of our application. Here each class is intended to handle common concerns such as persistence, transactional integrity, and security, and the properties, events, and methods of the beans are exposed to the top-level (web-interface) to be controlled accordingly. More specifically, the Stateful Session Beans are business objects having state: that is, they keep track of which calling client they are dealing with throughout a session, and thus, access to the bean instance is strictly limited to only one client at a time. If concurrent access to a single bean is attempted anyway the container serializes those requests. In some cases, stateful session beans' state may be

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union
European Social Fund
OPERATIONAL PROGRAMME
EDUCATION AND LIFELONG LEARNING
investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
Co- financed by Greece and the European Union
NSRF 2007-2013
programme for development
EUROPEAN SOCIAL FUND

persisted by the container automatically to free up memory after the client hasn't accessed the bean for some time.

To elaborate, the state of a client consists of all parts that constitute a query, namely the address/location the user specified in the textbox of the form, the latitudes and the longitudes of the points of interest that the user specified by clicking on the overlay map, all associated keywords, e.g. menu keywords and hotel amenities, the type of the issued query, e.g. range search, nearest neighbors query, etc., and other details which are associated with the specific user. Moreover, we provide additional functionality on the server side to log each user's queries in an effort to predict his future requests, or recommend additional features that are relevant to the user's query history, apart from the results of the current query. The connectivity of this layer with the data tier is made using the Java Database Connectivity API (JDBC), which is a Java-based data access technology that defines how a client may access a database. Essentially, it provides methods for querying and updating data in a database. Note that JDBC is oriented towards relational databases.

Of course the middle tier can be omitted *but* in such a client-server solution the client would be handling the business logic and that makes the client "thick". A thick client means that it requires heavy traffic with the server, thus making it difficult to use over slower network connections. By introducing the middle layer, the client is only handling presentation logic. This means that only little communication is needed between the client and the middle tier. For instance, in our case omitting the middle layer would

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

involve additional bandwidth consumption and tasks from the client side. First, all qualified tuples that correspond to landmarks, restaurants and hotels (say within walking distance from the specified area) would have to be retrieved, then serialized and transmitted over the network to reach the client machine. Finally, the browser would be burdened with processing the server's data (e.g. with JavaScript) and ranking all qualified tuples to illustrate just the top features. On the other hand, the middle tier determines what data is needed (and where it is located) and acts as a client in relation to a third data tier. Consequently, there are very significant benefits arising from introducing a three-tier architecture such as enhanced scalability, very little delay from the client point of view, and others. On the downside, an increased programming cost and system complexity is involved.

4.3 The Storage Layer

Last but not least, the third tier includes the database and a program to manage read and write access to it. The DBMS of our choice that suited our needs was MySQL 5.5. The schema of the database is quite simple and stores information about hotels, restaurants. More importantly, we also provide an option to perform faster search operations that rely on customized indexes that are capable of handling both the available spatial and the textual information simultaneously. Specifically, significant efforts were made towards designing and developing novel algorithms that would take that most out of customized multi-dimensional data-structures for the secondary memory,

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union European Social Fund
OPERATIONAL PROGRAMME EDUCATION AND LIFELONG LEARNING investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS MANAGING AUTHORITY
NSRF 2007-2013 programme for development EUROPEAN SOCIAL FUND
Co- financed by Greece and the European Union

which aggregate different types of information. In addition, special variants of these indices were also built to exploit the fact that queries focus on other aspects of the data than the ones that are usually considered when building such an index. In order to improve the performance even more, we have designed specialized index structures that will be integrated into our prototype.

5 Conclusions

Recently, the database research community has lavished attention on spatio-textual queries that retrieve the objects with the highest spatio-textual similarity to a given query. Differently, our research focus on the problem of ranking data objects based on the quality of facilities in their spatial neighborhood and their textual similarity to user-specified keywords. Towards this end, we have designed our prototype to support a novel query type called *top-k spatio-textual preference queries*. We have defined the interface as well as the system architecture of our prototype. In order to keep the response time of our prototype low, we will develop efficient algorithms for processing our novel query type that rely on our novel index structure that alleviates the limitations of state-of-the-art index structures. Currently, we have integrated the basic functionality in our prototype. We believe that our novel query type overcomes the shortcomings of approaches that focus on finding a set of data objects that are close to each other and relevant to a given query, as they ignore other attributes of the object, such as price or rating as well

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

as the relation to other facilities in their spatial neighborhood.

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union
European Social Fund
OPERATIONAL PROGRAMME
EDUCATION AND LIFELONG LEARNING
investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
Co- financed by Greece and the European Union
NSRF 2007-2013
programme for development
EUROPEAN SOCIAL FUND

References

[1] X. Cao, G. Cong, and C. S. Jensen. Retrieving top-k prestige-based relevant spatial web objects. *PVLDB*, 3(1):373–384, 2010.

[2] X. Cao, G. Cong, C. S. Jensen, and B. C. Ooi. Collective spatial keyword querying. In *Proc. of the Int. Conf. on Management of Data (SIGMOD)*, pages 373–384, 2011.

[3] G. Cong, C. S. Jensen, and D. Wu. Efficient retrieval of the top-k most relevant spatial web objects. *PVLDB*, 2(1):337–348, 2009.

[4] I. D. Felipe, V. Hristidis, and N. Rishe. Keyword search on spatial databases. In *Proc. of Int. Conf. on Data Engineering (ICDE)*, pages 656–665, 2008.

[5] Z. Li, K. C. K. Lee, B. Zheng, W.-C. Lee, D. L. Lee, and X. Wang. Ir-tree: An efficient index for geographic document search. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 23(4):585–599, 2011.

[6] J. B. Rocha-Junior, O. Gkorgkas, S. Jonassen, and K. Nørvåg. Efficient processing of top-k spatial keyword queries. In *Proc. of the Int. Symposium on Spatial and Temporal Databases (SSTD)*, pages 205–222, 2011.

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

[7] J. B. Rocha-Junior, A. Vlachou, C. Doulkeridis, and K. Nørvåg. Efficient processing of top-k spatial preference queries. *PVLDB*, 4(2):93–104, 2010.

[8] M. L. Yiu, X. Dai, N. Mamoulis, and M. Vaitis. Top-k spatial preference queries. In *Proc. of Int. Conf. on Data Engineering (ICDE)*, pages 1076–1085, 2007.

[9] M. L. Yiu, H. Lu, N. Mamoulis, and M. Vaitis. Ranking spatial data by quality preferences. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 23(3):433–446, 2011.

[10] D. Zhang, Y. M. Chee, A. Mondal, A. K. H. Tung, and M. Kitsuregawa. Keyword search in spatial databases: Towards searching by document. In *Proc. of Int. Conf. on Data Engineering (ICDE)*, pages 688–699, 2009.

The research project is implemented within the framework of the Action Supporting Postdoctoral Researchers of the Operational Program "Education and Lifelong Learning" (Actions Beneficiary: General Secretariat for Research and Technology), and is co-financed by the European Social Fund (ESF) and the Greek State.

European Union
European Social Fund
OPERATIONAL PROGRAMME
EDUCATION AND LIFELONG LEARNING
investing in knowledge society
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
MANAGING AUTHORITY
Co- financed by Greece and the European Union
NSRF 2007-2013
programme for development
EUROPEAN SOCIAL FUND