Designing a Peer-to-Peer Architecture for Distributed Image Retrieval

Akrivi Vlachou, Christos Doulkeridis,
Dimitrios Mavroeidis, and Michalis Vazirgiannis

Department of Informatics
Univ. of Economics and Business
Athens, Greece
{avlachou, cdoulk, dmavr, mvazirg}@aueb.gr

Abstract. The World Wide Web provides an enormous amount of images easily accessible to everybody. The main challenge is to provide efficient search mechanisms for image content that are truly scalable and can support full coverage of web contents. In this paper, we present an architecture that adopts the peer-to-peer (P2P) paradigm for indexing, searching and ranking of image content. The ultimate goal of our architecture is to provide an adaptive search mechanism for image content, enhanced with learning, relying on image features, user-defined annotations and user feedback. Thus, we present PIRES, a scalable decentralized and distributed infrastructure for building a search engine for image content capitalizing on P2P technology. In the following, we first present the core scientific and technological objectives of PIRES, and then we present some preliminary experimental results of our prototype.

Keywords: Peer-to-peer, distributed search, image retrieval.

1 Introduction

The advent of the World Wide Web in conjunction with efficient centralized search engines like Google and Yahoo has made an enormous amount of information easily accessible to everybody. Search engines provide efficient ranking thus easing identification of important results. Image content is only partially covered by web search engines, although it is evident that there is a tremendous wealth of digital images available on computers or other devices around the world. This is probably due to the fact that image content induces further complicated problems regarding the search: current centralized image search facilities do not sufficiently support metadata management, semantics, advanced media querying etc.

The widespread use of digital image equipment enables end-users to capture and edit their own image content, sometimes of high intellectual or commercial value. The centralized character of Web search raises issues regarding royalties,

N. Boujemaa, M. Detyniecki, and A. Nürnberger (Eds.): AMR 2007, LNCS 4918, pp. 182–195, 2008.

in the case of protected content, censorship, and to some degree information monopoly. Moreover current tools for image retrieval are limited in query expressiveness and lack semantic capabilities.

The P2P paradigm for file sharing, and recently covering WWW content [10,3], is a challenging alternative to centralized search and ranking. In a large scale network of peers, the storage and access load requirements per peer are much lighter than for a centralized Google-like server farm; thus more powerful techniques from information retrieval, statistical learning, and ontological reasoning can be employed on each peer's local search engine towards boosting the quality of search results. In addition, the virtual monopoly of information imposed by centralized search engines calls for more efforts towards information sharing (in our case image content), where the participants enjoy similar capabilities at least at the structural level.

In this context, the objective of this paper is to research, design and prototype a distributed indexing and searching toolkit, for image content, that capitalizes on the P2P paradigm using the Web as dissemination means. The main characteristic of PIRES (P2P Image Retrieval System) is that both content and indexing information is distributed across a large scale network of peers willing to share resources. The PIRES scenario assumes that users are making available part of their image content to the wider community, in exchange for the ability to search and access other users' content. Image content is tagged/annotated by the users in a way that it is indexable and therefore searchable. PIRES provides the means to manage the features extracted from images, to organize and index metadata in a distributed manner, to provide efficient P2P search mechanisms, and to facilitate adaptive ranking based on user feedback. Users participating in PIRES have a rich query mechanism enabling mixed queries based on keywords, ontology-based classifications, image samples assisted by a personalization mechanism based on adaptive user feedback.

In this paper, we present the design of the overall P2P architecture of PIRES. We describe the requirements of our approach, the research challenges that we wish to address, and the specific system components we focus on, namely feature extraction and image annotation, P2P organization and indexing, P2P search and access, and P2P personalized ranking. Currently, we have implemented a prototype simulator to study the feasibility and performance of the proposed P2P architecture. A full implementation of PIRES as well as testing in a real setting is left for future work.

The rest of the paper is organized as follows: We present related work in Section 2, while in Section 3 we introduce the overall architecture and the main roles in the system. Afterwards, in Section 4, we briefly outline the individual modules that realize the proposed functionality. In Section 5 we describe the P2P organization and indexing, while P2P searching is presented in Section 6. Thereafter, in Section 7, we describe the adaptive P2P ranking employed in PIRES. In Section 8, we present some preliminary experimental results. Finally, we conclude in Section 9.

2 Related Work - State of the Art

The advent of image search engines, such as Google Image Search (`http://images.google.com/`), together with the popularity of photo sharing systems, like `flickr.com`, indicate that contemporary users show increasing interests in capturing, storing, sharing and searching digital image content. Numerous commercial systems are available online, making the task of successful retrieval of multimedia content particularly challenging. A non-exhaustive list of such systems includes: search engines, like AltaVista Image Search (`http://www.altavista.com/image/default`) and Riya (`www.riya.com`), content-based retrieval tools, like GIFT, the GNU Image-Finding Tool (`http://www.gnu.org/software/gift/`), and news/media portals, such as the REUTERS image library search and retrieval tool (`http://www.carydesign.co.uk/ril/`).

A number of EU research projects in relevant areas that have recently concluded or are under development imply the significance and increasing interest for multimedia (text, image, audio-visual content) search and retrieval. *MESH* (Multimedia Semantic Syndication for Enhanced News Services) (`http://www.mesh-ip.eu/`) intends to apply advanced multimedia analysis to provide personalized multimedia summaries and novel access mechanisms for news to end users. The *K-Space* network of excellence (`http://www.kspace-noe.net/`) also conducts research relevant to content-based image analysis. Projects related to advanced search technologies for digital audio-visual content are coordinated by the *CHORUS* coordination action. Two P2P-related projects, under this action, are *SAPIR* (`http://sysrun.haifa.il.ibm.com/sapir/`) and *Victory* (`http://www.victory-eu.org/`). *SAPIR* is the most similar research project to our work.

Distributed and peer-to-peer (P2P) image retrieval has attracted some research interest lately. In [6], a system called FuzzyPeer is presented for answering similarity queries in P2P networks, like *"find the top-k images which are similar to a given sketch"*. In [11], the authors argue against the appropriateness of structured P2P systems for queries based on image similarity metrics (such as color histograms). They investigate how unstructured P2P networks can be exploited to provide such enhanced searching facilities. Müller and Henrich also study P2P content-based image retrieval in unstructured P2P networks, based on compact summaries in [12]. They assume a small number (compared to the number of peers) of globally known clusters, so that queries can be directed to only a subset of peers with relevant content. In contrast, our approach does not make any assumptions about the distribution of peer contents, although better performance can be achieved if the data is clustered. In [7], the design and implementation of a system for P2P content-based multimedia retrieval is presented.

An overview of adaptivity features that a P2P image retrieval system should demonstrate is presented in [13]. A relevance feedback mechanism for P2P content-based image retrieval is proposed in [8]. Further, while there exists sufficient work on the issue of distributed ranking (top-k retrieval), there are only a few works that study the issue of P2P ranking, for example [1,9]. All these research papers are related to (parts of) the functionality that PIRES aims to provide.

In comparison to previous work in P2P image retrieval, our approach focuses mainly on scalability. We propose a feasible P2P system that encompasses and combines salient features, such as decentralization, high availability, fault-tolerance, relatively low response times as well as respect for peer autonomy. These features are hardly combined in existing research in the relevant literature.

3 Architecture

The PIRES architecture utilizes a P2P infrastructure for supporting the deployment of a scalable enhanced search engine for multimedia content, addressing (and particularly suitable to) future user needs and search requirements. A high-level abstraction of the PIRES architecture, showing the different participating entities (peers), their interconnections and functionality is shown in Figure 1. Collaborating peers in the PIRES search engine consist of: 1) *content providers* and 2) *information brokers*. These roles are not mutually exclusive. For instance, a media portal may utilize peers with substantial processing and storage capabilities as information brokers, which provide content as well as indexing and retrieval functionality at the same node. Other participants, such as mobile users running mobile client software on their personal computers may only act as content providers. Below we discuss these roles in more detail.

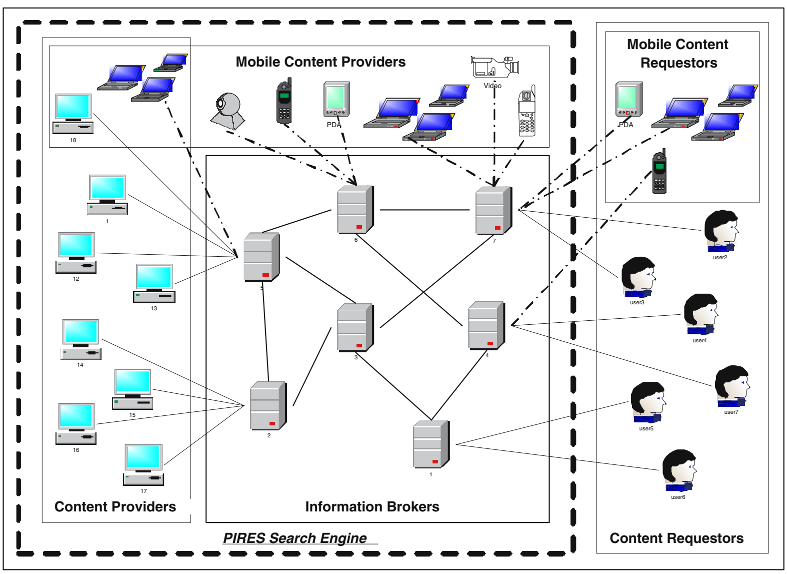


Fig. 1. High level view of the PIRES architecture

Content providers are nodes that produce or store images that they would like to publish. We stress here that only metadata is indexed and managed by PIRES; the actual content remains on the participating peers that own it. Using the PIRES client software, peers can register to a PIRES information broker peer and publish their content. This process involves manual annotation of the material to facilitate the indexing and retrieval potential of interested parties. One of the features of the PIRES client software is that it encapsulates a feature extraction functionality. Using this embedded service, a peer can automatically or semi-automatically extract a set of features for its local images.

Besides stationary content providers, PIRES supports mobile content providers, which are able to dynamically capture content and make it widely available. We envisage reporters or tourists as potential candidates for this role, allowing instant access to "fresh" multimedia content. Mobile content providers may additionally/optionally upload the actual multimedia content to dedicated servers – instead of only metadata – to ensure continuous availability.

Information brokers consist of more powerful and less volatile peers in the network. They realize a decentralized indexing service. In addition to the basic form of metadata that is generated by the content providers, information brokers may employ more sophisticated (and thus demanding) algorithms that could not be executed in lightweight peers, due to lack or processing power or lack of more widespread knowledge on the rest of the multimedia content in the network. By participating in PIRES, information brokers gain access to a wealth of "fresh" user-created multimedia content, which is organized and indexed in a way that allows more elaborate search facilities. We envisage news portals, agencies or third-party multimedia repositories as potential PIRES information brokers. With PIRES these entities increase the visibility of their content, but most importantly they enrich their content collection, by getting access to thousands or millions of independent and undiscovered multimedia sources (content providers).

Content requestors are not necessarily contributing content to PIRES search engine; however their role is crucial to the architecture, as they represent the users of the search engine. Moreover user feedback is exploited to provide improved ranking of results. Content requestors enjoy a rich repertoire of query and searching facilities seamlessly integrated in a simple/intuitive but powerful interface. PIRES supports a variety of different search modes and a mixture of them.

In more technical terms, PIRES relies on a super-peer architecture. Because super-peer infrastructures [15] harness the merits of both centralized and distributed architectures, we make a design choice of a super-peer architecture for the P2P network interconnection.

Content providers are simple peers (henceforth called peers) that keep multimedia content, usually generated on their own. Each peer joins the collaborative search engine, by connecting to one of the information brokers that act as super-peers, using the basic bootstrapping protocol. The information brokers are entities that maintain summaries/indexes of content/data that reside on content providers.

The choice of the proposed super-peer architecture is motivated and driven by our main requirement for scalability. Currently deployed super-peer networks, e.g. eMule (`www.emule-project.net`) and KaZaA (`www.kazaa.com`), support large number of users, demonstrating the appropriateness of super-peer architectures when scalability is required. In addition, peer autonomy is respected as the actual content/data remain on the owner peer, and only metadata is indexed and handled by the super-peer network.

In order to accommodate the proposed functionality, we identify the following main components of the proposed system:

– Feature extraction and image annotation
– P2P organization and indexing
– P2P search and access
– P2P personalized ranking

In the following, we provide a short overview of the system, and then we describe in more detail the individual components, their functionality, and how their collective combination realizes the proposed system.

4 System Overview

The overall objective of PIRES is to provide novel, dynamic and enhanced image retrieval services, over a P2P architecture of collaborative computers. For this purpose, a set of basic functionalities on each peer are required, varying from low level feature extraction and annotation of image data, to organizing, indexing and searching the extracted metadata in a distributed way.

4.1 Local Feature Extraction and Annotation

Content-based analysis and retrieval of multimedia documents requires methods for feature extraction. Features may include both text-based features (such as key words, annotations) and visual features (such as color, texture, shape, faces). Combing the low level features with the high level concepts of the keyword annotations bridges the gap between actual image content and the user's sense of similarity. There already exists rich literature on existing techniques of visual feature extraction [14]. However, finding an automatic transition from the low-level features to semantic entities or equivalently the automatic extraction of high-level characteristics is an extremely hard task [2].

For images, existing image segmentation and object recognition algorithms cannot yet reach a level of precision which would be adequate for automatic annotation. In PIRES, we aim to combine knowledge-driven image analysis and text-mining technologies, to support the user in the process of semantic annotation of the images and image-parts. This process is assisted by ontological information or some domain specific taxonomy, that enables annotation of images not only with keywords that can provide syntactic matching, but rather with semantic concepts that allow more enhanced semantic matching.

Reuse of already generated semantic metadata can be exploited to support automatic text-based and image-based methods also for other resources. Unfortunately, even with such a combined strategy it cannot be realistically expected that fully-automatic semantic metadata generation will always yield acceptable results. Therefore, we propose a semi-automatic approach, where users give feedback, correct and extend the results of image segmentation, image recognition and text mining algorithms. To achieve that, user-friendly annotation tools are needed, and they will be implemented in our future work.

4.2 Decentralized and Distributed Indexing of Image Content

The purpose of the advanced feature extraction mechanism is to feed a distributed indexing mechanism. Locally extracted metadata and user annotations are mapped to the global ontology, in order to: 1) create common semantics, and 2) enable more sophisticated retrieval mechanisms than exact matching.

Advanced P2P mechanisms will be designed to identify thematically coherent communities of participants and thus organize content in a way (we envisage semantic overlay networks [4] as a potential candidate) that enables efficient searching.

4.3 P2P Search and Access Mechanisms for Image Content

The search facilities in PIRES cover both spontaneous and subscription queries. The search is formulated based on a rich set of advanced features: keywords, metadata tags, semantics, media samples or a mixture of them. Additional parameters that will expand queries in the search process will be contextual attributes (i.e. location, time, device features, personal profiles). Furthermore, we envisage an advanced personalization process that will capitalize on mining query/usage logs and interaction histories. The combination of semantic/ontology methodologies and the statistical ones, will offer users the possibility to have a much more precise and to the point interaction with the distributed image content.

4.4 Peer-to-Peer Ranking

PIRES aims to use novel P2P ranking algorithms that assist users in their search. The issues to be taken into account in the ranking process deal with: 1) the limited knowledge of content any entity of PIRES has, and 2) how to incorporate user feedback in order to provide adaptive ranking of results.

Another issue that should be taken into account is the personalization of the rankings. This is due to the fact that different users could expect different results even when they provide the same query. In order to address this issue we employ a novel query expansion mechanism, based on relevance feedback that ensures the adaptiveness of the query ranking framework to the specific user needs.

In the next sections, we focus on the P2P aspects of the image content retrieval system, which justify its innovation and how PIRES enhances the state of the art, both from the scientific and technological point of view.

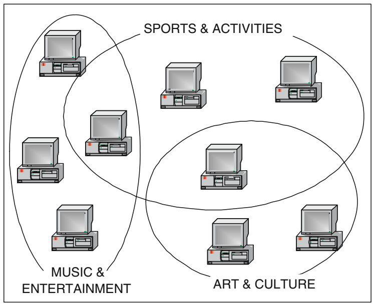


Fig. 2. The notion of semantic overlays in a P2P network

5 P2P Organization and Indexing

Users decide which features (metadata) to publish to the search engine, in order to make their content searchable. This is a deviation from the centralized search engine paradigm that relies on crawling: the user decides what meta-data information to publish and initiates the indexing phase. Thus the adopted model is push-based, similar to [3,4], compared to the traditional pull-based model for collecting web data. Furthermore, before making content and metadata publicly available, users may add annotations that enrich the expressiveness of their metadata, facilitating access by other users and increasing visibility. The client-software transmits to the information broker the instances and the mappings of the extracted features that will be appended on the appropriate representation file at the information broker. This reduces network traffic, as only changes to the file are communicated.

We stress here that the actual image content resides on the peers - only the metadata extracted are published (for instance for images only, a low-quality preview of the picture can be published as well). It should also be stressed, that digital rights management issues can be determined at this point, by having each user explicitly specify the access rights for its content. However, as such issues are out of the technical scope of this paper, they will not be consider further here.

5.1 Distributed Metadata Organization

A topic related to global indexing is the dynamic formation of interest-based communities in a transparent way to the end-user. In this manner, content is organized into thematic categories, with several advantages like more efficient query processing and creation of thematically focused groups. Towards this end, PIRES relies on the decentralized generation of Semantic Overlay Networks (SONs) as a scalable and efficient solution to P2P content organization and subsequent searching [3]. SONs (see also Fig. 2) have been proposed as a mechanism that reduces the cost of searching in large-scale P2P networks, while increasing the quality of the results, since queries are directed only to specific subsets of peers that hold relevant contents to the query [4]. Regarding the actual construction of these overlays, PIRES adopts P2P clustering and classification techniques either for a completely unsupervised scenario or for a semi-supervised setting respectively.

5.2 Distributed Indexing

Within the generated overlay networks, distributed indexing is required in order to serve requests for content. Information brokers accumulate metadata information that is published by the content providers. These metadata are used to populate local indices that describe the content of the connected peers. The precision/accuracy of these indices strongly depends on the quality of the feature extraction process.

In the distributed indexing phase, super-peers need to consolidate their local indices, in order to provide a global distributed index. This index needs to be distributed for several reasons: 1) scalability, 2) load-balancing, 3) fault-tolerance, 4) in order to avoid censoring phenomena. This exchange of local indices can be performed by broadcasting, in the absence of more sophisticated dissemination mechanisms.

Ontologies influence the global index too, since they provide information of the keywords and the topics of the material that is indexed at a super-peer and can be retrieved from its peers. For instance, we can use them to quickly discard super-peers that can not contribute to the final result, thus decreasing communication costs.

Our prototype implementation uses distributed indexing based on clustering over the features extracted by the images to support the retrieval of similar based on content images.

6 P2P Search and Access

Search requests (henceforth also termed queries) for image content are initiated by content requestors that wish to retrieve and exploit fresh content. Notice that the role of content requestors can correspond to end users, content providers and information brokers. By utilizing its P2P indexing scheme, PIRES supports not only most of the conventional query models, but also provides a flexible infrastructure, where future user requirements or more human-centric query interfaces (that resemble more the way human requests are posed in the real world) can be accommodated. PIRES supports the following non-conventional query/search models:

Semantic Keyword-Based Queries. The user provides a set of keywords that are transparently associated with parts of the ontology to exploit semantic inferences. This query type differs from the traditional keyword-based search, since it goes beyond exact matching techniques supporting semantic relationships. We also consider queries expressing semantic or knowledge based relationships.

Query by Thematic Category. In addition to providing keywords, users belonging to specific communities and interested in particular topics may issue queries within specific thematic areas (i.e. in terms of taxonomy paths), in order to increase the search precision. In this case, a generic query that would normally have low precision can be targeted to a specific thematic category, thus increasing the probability of high quality results and at the same time decreasing the incurred processing cost. The choice of thematic category implies some data organization (manual or semi-automatic), e.g. into a globally known taxonomy (such as DMOZ), so that the user is supplied with a list of categories at query time.

Similarity-Based Queries. Instead of traditional text based requests an object is given by the user and similar objects based on a given image are retrieved. These queries belong to the Query by Example (QBE) category and are evaluated based on the feature space in which the images are represented. Each image is represented by a high-dimensional vector and the closest images to the given image are retrieved.

Mixed Queries. A combination of keyword-based, thematic category and query by example. Mixed queries aim to enhance the expressiveness of the query interface by combining the aforementioned query types. These queries are particularly useful when users cannot explicitly specify what they seek in terms of (for example) keywords only. Nevertheless, it is often the case that users know other parts of information that can help shaping a more specific query, for example the thematic category corresponding to the query. This facility is usually missing from current search engines.

Subscription-Based Queries. PIRES also supports subscription-based queries. Rather than having clients explicitly requesting the information they need, data is sent to them as soon as it is available. Clients submit a subscription to receive updates on new image content that is published based on a variety of criteria (similar to pull-based queries) and then disconnect.

From the aspect of the Information Brokers, the design of the indexing and organization of the metadata takes into account the efficient query processing of all mentioned types of queries. Based on the distributed organization, when a query is posed to a super-peer (initiator super-peer) the initiator must propagate the query to those peers (based on the global index) that influence the result set. After all queried super-peers respond, the initiator super-peer has to merge the result into a global result set, which are returned to the content requestor that had submitted the query.

7 Ranking the Retrieved Image Metadata

Searching in a large-scale environment is practically useless in the absence of efficient ranking. According to the experience from traditional search engines, users seldom browse through the entire list of results, they rather focus on the first few results only. It is therefore crucial to provide ranking of the retrieved images and this brings up several challenging research questions, i.e., how to rank image content? A trivial ranking is based on media-specific information, e.g. image resolution, size, date. On the other hand, it would be advantageous to contemplate a quality metric for ranking the results using the annotations. This metric can be based on (for example) how often a keyword occurs.

The user interface also provides certain extensions to make ranking more flexible and thus useful to the user. For example, the user can rank the retrieved results according to location, time, etc., which can be particularly useful when searching for nearby or ”fresh” content. Thus, we identify at this stage, the following ranking criteria: 1) Semantic-based, 2) Context-based (Location-based, Time-aware, Media-specific).

Another interesting extension is to use **query feedback** to better rank matching data for future queries. This functionality is provided at the screens that depict the query results in the client software and can be used to fine-tune the local indexes and also enable the provisions of personalized rankings.

We observe that the practical appropriateness of the ranking of results depends heavily on the utilization of a “good” similarity measure. A “good” similarity measure should be able to identify conceptually similar objects, thus helping users to retrieve relevant images based on their queries. Many researchers have argued that instead of trying to define a “global” similarity measure systems should be adaptive, exploiting the feedback provided by the users.

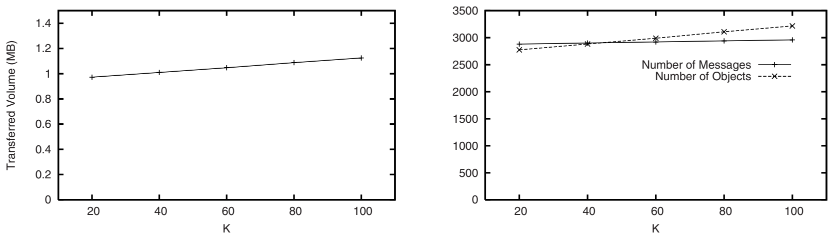
As we have mentioned in Section 6, we aim to use semantic relations/concepts that are provided by an ontology/hierarchical taxonomy to improve the effectiveness of input queries. In the interest of creating personalized rankings, the framework includes a personalization mechanism that learns locally (at peer level) the weights of ontology concepts based on relevance feedback. Since peers utilize the ontology relations for enhancing queries with semantics, the local weight learning process results into a personalized ranking for each peer, thus adapting the query results based on the personalized needs of each peer.

Accessing image content. The query result set that is displayed to the user that posed a query contains only the metadata of the actual object, such as description and a thumbnail. The original (high quality) image file is hosted by the content provider in general. In PIRES, mechanisms will be implemented to support mechanisms that supervise the transactions of files. The content requestors access the original files hosted by the content providers, through the client software and the information broker infrastructure.

Navigating. The query result set containing the metadata of the actual objects is displayed to the user that posed a query as a list, where each row of the list contains a thumbnail and media-specific features, such as size and date. The user may navigate through the list and choose (by clicking) any object he is interested in. Then a more detailed description of the object is displayed where the thematic and semantic categories based on the ontology are represented. The user may navigate through all associated categories or request objects similar to the object based on the displayed semantics.

8 Experimental Results

We have implemented a prototype simulator of the proposed system, in order to study experimentally its properties and test whether PIRES satisfies its design requirements. In our prototype, a super-peer architecture is implemented, and



(a) Transferred volume (b) Number of messages and objects

Fig. 3. Transferred volume, number of messages and objects transmitted during search

peers store image content described by feature extracted by the images. In particular, we use the VEC dataset, which consists of 1M 45-dimensional vectors of color image features. The algorithms used in the experiments are described in [5].

The P2P network topology used in the experiments consists of 200 interconnected super-peers in a random graph topology, built using the GT-ITM topology generator[1], with average connectivity degree equal to 4. Each super-peer is assigned 10 peers, making a total of 2000 peers in the system. At this stage, we study only queries by example based on the image feature space, i.e. given an n-dimensional query object, retrieve the top-k most similar objects to the query. This can also be expressed as retrieval of those objects that are similar to the query with a certain similarity threshold. Similarity is based on a distance function and in this set of experiments we use the Euclidean distance. All simulations run on 3.8GHz Dual Core AMD processors with 2GB RAM, and in all cases we show the average values over all queries.

At first, we measure the volume transferred through the network, for answering a query. Note that this volume corresponds only to the volume of the features values. In Fig. 3(a), we see that the volume is in the order of 1MB, for different number of requested results. Given the size of the simulated network, this is considered quite tolerable for our network. The value k on the x-axis denotes the number of results requested by the user, i.e. the top-20 to top-100 most similar images to the query.

Next, the number of required messages and the number of objects transferred in the network is measured. In Fig. 3(b), we show these numbers again for varying values of k. The increasing tension in both curves is due to the increasing k values. Also, it is important to notice that the number of messages increases very slowly with the number of results, which is a sign in favor of the scalability of the approach with number of results. The number of objects retrieved increases faster, but this is expected and due to its explicit relationship with the k value.

Finally, Fig. 4 depicts the response time of the system for varying network transfer times. As response time, we mean the time required to retrieve the

[1] Available at: http://www.cc.gatech.edu/projects/gtitm/

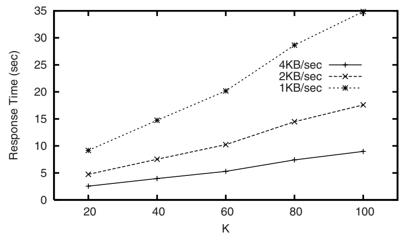


Fig. 4. Response time

number of results specified by the user. Notice that we use very modest transfer times, 1-4KB/sec, that resemble the speed of dial-up connections. In reality, connections between super-peers or peers will be much faster, so we study here a worst case scenario of poor communication, in terms of bandwidth. Even for this setup, for example in the case of 4KB/sec and 20 results, the response time is approx. 2.5 seconds. In general, the response time increases with k, as more results require more time to reach the user.

9 Conclusions

In this paper, we presented an P2P architecture for efficient search mechanism for the image content that is scalable and supports coverage of the web contents. We outlined the requirements for indexing, searching and ranking of image content in a decentralized and distributed environment. Relying on a super-peer architecture, we presented the overall architecture of the PIRES framework and presented the initial results of our prototype implementation. In our future work, we plan to implement and deploy our system, in order to test it in a real network platform like PlanetLab[2].

References

1. Balke, W.-T., Nejdl, W., Siberski, W., Thaden, U.: Progressive distributed top-k retrieval in peer-to-peer networks. In: Proceedings of ICDE, pp. 174–185 (2005)
2. Chang, S.: The holy grail of content-based media analysis. IEEE Multimedia 9(2), 6–10 (2002)
3. Doulkeridis, C., Nørvåg, K., Vazirgiannis, M.: The SOWES approach to P2P web search using semantic overlays. In: Proceedings of WWW, pp. 1027–1028 (2006)
4. Doulkeridis, C., Nørvåg, K., Vazirgiannis, M.: DESENT: Decentralized and distributed semantic overlay generation in P2P networks. IEEE Journal on Selected Areas in Communications (J-SAC) 25(1), 25–34 (2007)
5. Doulkeridis, C., Vlachou, A., Kotidis, Y., Vazirgiannis, M.: Peer-to-peer similarity search in metric spaces. In: Proceedings of VLDB 2007, pp. 986–997 (2007)

[2] `https://www.planet-lab.org/`

6. Kalnis, P., Ng, W.S., Ooi, B.C., Tan, K.-L.: Answering similarity queries in peer-to-peer networks. Inf. Syst. 31(1), 57–72 (2006)
7. King, I., Ng, C.H., Sia, K.C.: Distributed content-based visual information retrieval system on peer-to-peer networks. ACM Trans. Inf. Syst. 22(3), 477–501 (2004)
8. Lee, I., Guan, L.: Content-based image retrieval with automated relevance feedback over distributed peer-to-peer network. In: Proceedings of ISCAS, pp. 5–8 (2004)
9. Michel, S., Triantafillou, P., Weikum, G.: Klee: A framework for distributed top-k query algorithms. In: Proceedings of VLDB, pp. 637–648 (2005)
10. Michel, S., Triantafillou, P., Weikum, G.: Minerva∞: A scalable efficient peer-to-peer search engine. In: Middleware, pp. 60–81 (2005)
11. Müller, W., Boykin, P.O., Sarshar, N., Roychowdhury, V.P.: Comparison of image similarity queries in P2P systems. In: Proceedings of P2P, pp. 98–105 (2006)
12. Müller, W., Henrich, A.: Fast retrieval of high-dimensional feature vectors in P2P networks using compact peer data summaries. In: Proceedings of MIR, pp. 79–86 (2003)
13. Müller, W., Henrich, A., Eisenhardt, M.: Aspects of adaptivity in P2P information retrieval. In: Proceedings of AMR, pp. 1–15 (2005)
14. Rui, Y., Huang, T., Chang, S.: Image retrieval: current techniques, promising directions and open issues. Journal of Visual Communication and Image Representation 10(4), 39–62 (1999)
15. Yang, B., Garcia-Molina, H.: Designing a super-peer network. In: Proceedings of ICDE 2003, p. 49 (2003)