Building Self-Organized Image Retrieval Network

Stanislav Barton
Faculty of Informatics
Masaryk University
Brno, Czech Republic
barton@fi.muni.cz

Vlastislav Dohnal
Faculty of Informatics
Masaryk University
Brno, Czech Republic
dohnal@fi.muni.cz

Jan Sedmidubsky
Faculty of Informatics
Masaryk University
Brno, Czech Republic
xsedmid@fi.muni.cz

Pavel Zezula
Faculty of Informatics
Masaryk University
Brno, Czech Republic
zezula@fi.muni.cz

ABSTRACT

We propose a self-organized content-based Image Retrieval Network (IRN) that is inspired by a Metric Social Network (MSN) search system. The proposed network model is strictly data-owner oriented so no data redistribution among peers is needed in order to efficiently process queries. Thus a shared database where each peer is fully in charge of its data, is created. The self-organization of the network is obtained by exploiting the social-network approach of the MSN – the connections between peers in the network are created as social-network relationships formed on the basis of a query-answer principle. The knowledge of answers to previous queries is used to fast navigate to peers, possibly containing the best answers to new queries. Additionally, the network uses a randomized mechanism to explore new and unvisited parts of the network. In this way, the self-adaptability and robustness of the system are achieved. The proposed concepts are verified using a real network consisting of 2,000 peers containing descriptive features of 10 million images from the Flickr Photo Sharing system.

Categories and Subject Descriptors

H.3 [**Information systems**]: Information storage and retrieval

General Terms

Algorithms, Design, Experimentation

Keywords

content-based image retrieval, peer-to-peer network, self-organized system, social networking

1. INTRODUCTION AND MOTIVATION

The content-based information retrieval recently attracts bigger attention. Digital multimedia follows the current trend in the growth of the digital content available. The multimedia is originated either by digitizing already created content or just publishing and making the content available through specialized databases like digital libraries or galleries. Since the large collections are already available the search in such challenging data is becoming very up-to-date.

Unstructured networks like Kazaa, Gnutella or Freenet are used for file sharing. Yet, the files shared in these networks are searched according to their names. On the other hand, systems built to facilitate the content-based information retrieval work rather in the structured manner (M-Chord [17], MCAN, VPT*, GHT*[2] or distributed hashtable approaches). Such types of systems require the data redistribution among peers, mainly for clustering reasons, making their deployment in the large web-scale shared multimedia database usable in limited extent in terms of scalability.

The unstructured peer-to-peer systems emerged in order to enable the sharing of the content in a data-owner oriented way [24]. Such approaches enabling the content-based search in peer-to-peer networks represent Linari et al. [15] and MON [14].

In this paper, we would like to propose a distributed system that will be able to scale to large extents, representing tens of millions of entities on thousands of peers and enable the content-based multimedia retrieval from a shared database. In this paper we prove the concept of such design using 10 million digital images incorporated in the shared database – the Image Retrieval Network (IRN).

This paper is organized as follows. The related work is presented in Section 1.1 and the contribution of this work is summarized in Section 1.2. In Section 2 the reader is provided with the necessary background. Section 3 introduces the structure of the proposed content-based information retrieval network. Consequently, Section 4 presents the experience with the proposed structure. Finally, the paper is concluded and the directions of future work are discussed in Section 5.

1.1 Related Work

Large-scale searching and indexing methods for digital images have been the center of attention of many academic and commercial research groups. Those efforts can be roughly divided into two major ways of retrieving images. Firstly, it is using the annotations or other attributes associated with each image. Secondly, it is the content of the image that is being compared to retrieve the most similar images to the query one. When the former approach is used, large data collections can be managed efficiently because traditional well-established indexing and searching techniques can be applied. In the latter case, features are extracted from each digi-

LSDS-IR'08, October 30, 2008, Napa Valley, California, USA.

tal image and those usually represented by a very high-dimensional vector stored in an index allowing to compare the similarity of the acquired feature vectors. Finally, some recent social networking and self-organization work is reviewed in the last subsection.

1.1.1 Annotation-Based Retrieval

A number of systems use this approach to index Web-scale image collections, e.g. Google Images, Yahoo!, ExaLead [8], or PicSearch [19]. Because these systems may suffer from a lack of trustworthy image annotations, there are attempts to increase the quality of image labeling via interactive public-domain games, e.g. The ESP Game [7] or Google Image Labeler [10]. Another example of an attribute-based search application is photo searching by geographic locations where the photos were taken. This functionality is provided for instance by the Flickr system [27] that allows users to query about 30 million images using such geotag.

Another direction of research is an automatic image annotating based on the analysis of image content. The ALIPR [13] system uses a training set of images which is separated into a number of concept categories; these categories are annotated. The indexed images are categorized and annotated by means of extracting visual features from them and by applying statistical methods. Searching is then based purely on the assigned annotations. The ALIPR [1] has a public demo, its internal functionality is provided by an older system SIMPLIcity [26].

1.1.2 Content-Based Retrieval

The content-based image retrieval is an orthogonal approach where the images are searched according to their visual content similarity. Such techniques typically extract a characterization (signature, descriptor) of the image content, which is then used for indexing and searching. Recent comprehensive surveys [25, 6] describe a number of approaches and their applications. These systems are usually specialized and tuned for a specific application domain and/or they manage relatively small data collections. The Tiltomo project [23] indexes about 140,000 images from Flickr and searches them by color and texture characteristics; it also allows to combine it with subject searching. The ImBrowse [11] system allows to search a collection of about 750,000 images by color, texture, shapes and combinations of these (employing five different search engines).

A great research effort is still put into this area, let us mention at least activities integrated in the Chorus [4] initiative. To the best of our knowledge, there is no technology for general image similarity search which would scale to data volumes comparable to the number of images covered by the Web (tens of millions to billions). Also there are efforts that combine both content and non-content based search paradigms like [9].

1.1.3 Social Networking and Self-Organization

The approach presented is an unstructured P2P network that organizes its self in order to get better results more efficiently. Other systems on a similar basis working as Semantic Overlay Networks are [15, 14]. Nonetheless, the presentation of these systems lack the discussion of the scalability issues therefore their usage on a dataset as large as ours is questionable. The self-organization of our system stems from the principles of social networking. This paradigm is recently becoming very popular due to its rational incorporation of both randomness and exploitation of previous knowledge. Query routing algorithms in such systems are presented in [3]. The maintenance of the social networking networks is discussed in [18].

1.2 Contributions

The proposed system incorporates two aspects to be able to scale to large extents: the unstructureness and self-organization that provides a good approximation with low query costs. In [21] the authors present a Semantic Overlay Network (SON) that adopts a metric space to achieve the extensibility of their system to search within multimedia data. We give an outline of this system's principles in Section 3.1. We got inspired in this SON, reworked it and deployed it as a stand-alone system where no additional underlying structure to maintain the data is needed as it is in the case of the SON approaches.

The system proposed in this paper addresses these challenges, so the overall contributions are:

large-scale the properties of the system are evaluated on the dataset that comprises of 10,000,000 images belonging to 2,000 users mapped onto 2,000 peers.

improved self-organization and self-adaptability the ability to self-organize is that autonomous entities (peers) group together into certain structures without any explicit rule. The whole system can evolve either in time or space. Our system organizes itself in a way that local knowledge is sufficient to localize relevant data sources and random factors are used to explore unvisited parts of the network.

data-owner orientation no data redistribution is needed in order to index and access the data in the network, therefore the peer is fully in charge of its own data contribution to the database.

content-based image retrieval simple perceptual quantities drawn from the so-called *low-level* semantic fields represented by five MPEG-7 feature descriptors are used to extract feature vectors from the original digital image.

2. BACKGROUND

A very useful, if not necessary, search paradigm is to quantify the *proximity*, *similarity*, or *dissimilarity* of a query object versus the objects stored in a database to be searched. Roughly speaking, objects that are *near* a given query object form the query response set. A useful abstraction for nearness is provided by the mathematical notion of *metric space* [12]. We consider the problem of organizing and searching large datasets from the perspective of metric spaces, sometimes conveniently labeled distance spaces. This research branch has been surveyed in recent books [20, 29].

2.1 Metric Space

Suppose a metric space $\mathcal{M} = (\mathcal{D}, d)$ defined for a domain of objects (or the objects' *keys* or *indexed features*) $\mathcal{D}$ and a total (distance) function d. In this metric space, the properties of the function $d : \mathcal{D} \times \mathcal{D} \mapsto \mathbb{R}$, sometimes called the metric space postulates, are typically characterized as:

$$\begin{array}{ll}
\forall x, y \in \mathcal{D}, d(x, y) \geq 0 & \text{non-negativity,} \\
\forall x, y \in \mathcal{D}, d(x, y) = d(y, x) & \text{symmetry,} \\
\forall x, y \in \mathcal{D}, x = y \Leftrightarrow d(x, y) = 0 & \text{identity,} \\
\forall x, y, z \in \mathcal{D}, d(x, z) \leq d(x, y) + d(y, z) & \text{triangle inequality.}
\end{array}$$

Examples of distance functions are L_p metrics (City-block (L_1), Euclidean (L_2), or maximum distance (L_∞)), the edit distance, or the quadratic-form distance, to name a few. For image features, weighted versions of L_1 and L_2 metrics are usually used.

2.2 Similarity Queries with Timestamps

Probably the most common type of similarity query is the *range query* $R(q, r)$. The query is specified by a query object $q \in \mathcal{D}$, with some query radius r as the distance constraint. From a database $X \subset \mathcal{D}$, the query retrieves all objects found within the distance r from q, formally:

$$R(q, r) = \{o \in X, d(o, q) \leq r\}.$$

The similarity queries are extended with *timestamps* to know when the query was issued. From now and on, we denote the range query as $R(q, r, t)$ and the range query will be the kind of queries that our system processes and on whose answers builds the information to locate similar data.

2.3 MPEG-7 Feature Descriptors

To transform a digital image into data in metric space, usually an extraction of characteristic features is used. For our image dataset, we use the following five MPEG-7 feature descriptors: the *color structure* (CS), *color layout* (CL), *scalable color* (SC), *edge histogram* (EH), and *homogeneous texture* (HT). Specifically, CS, CL, and SC express the spatial distribution of colors in an image. The EH captures local density of edge elements and their directions (sometimes called the *structure* or *layout*); it acts as a simple and robust representation of shapes. Finally, HT is a texture descriptor. These descriptors are typically represented as vectors and the MPEG-7 [16] contains a specific distance measure for each of the descriptors. These measures express the *dissimilarity* of two particular descriptors and all of them satisfy the metric postulates.

As a metric function for this data, we use an aggregation of the metric functions for each extracted feature as a weighted sum of the individual feature distances. The particular distances are normalized before weighted and summed. Table 1 summarizes the features we use, their respective distance measures, and their weights in the aggregate metric function.

MPEG-7 Feature	Metric	Weight
Scalable Color	L_1 metric	2
Color Structure	L_1 metric	3
Color Layout	sum of L_2	2
Edge Histogram	special	4
Homogeneous Texture	special	0.5

Table 1: Extracted features summary.

3. SELF-ORGANIZED IMAGE RETRIEVAL NETWORK

The proposed self-organized image retrieval network is based on a system called Metric Social Network (MSN) presented in [21]. The MSN is designed to work as a Semantic Overlay Network, i.e., to work as an overlay to an existing P2P network to lower the query costs by approximating answers to queries posed. As the underlying network authors used M-Chord [17] that is a structured P2P network in which the data are redistributed to achieve better data clustering.

The aim of this paper is to propose a self-organized image retrieval network that would represent a shared database for content-based image retrieval. We would like the network to operate independently on any other network or system to meet the principles of unstructured P2P system, i.e., that the system can be joined or left by any peer without a larger impact on the network itself. Therefore, a great emphasis in the design is put on the requirement of no data redistribution in the network.

In this section, we give a brief outline of the MSN structure followed by a review of enhancements made towards the proposed system.

3.1 Metric Social Network Overview

Peers of the MSN are capable of querying other network peers. Interconnection of peers is based on the query-answer paradigm, i.e., the relationships among peers are established according to metadata created upon results of processed queries. Each peer maintains metadata about queries it has asked or answered, called a *query history*. This represents the peer's knowledge about the network and is exploited by a routing algorithm. Every entry E_i (the metadata about a processed query) stored in the query history identifies peers that participated in the query answering. This way, relationships are formed among peers with respect to the particular query. Therefore, each entry in the query history has associated the processed query (the query object, radius and timestamp), an *acquaintance* and a list of *friends*. The acquaintance (weak relationship) is the peer leading to the community of peers that return the best answer to the query. Whereas the friends (strong relationships) are the peers that contributed to the query answer significantly. The detailed architecture of the MSN can be found in [21].

3.1.1 Adaptive Query Routing Algorithm

The major contribution over other systems is the adaptive query routing based on the notion of *confusability* of processed queries. In [22], Shepard proposed a *Universal Law of Generalization*. The law states that the probability of confusing two items o_1 and o_2 is a negative exponential function of the distance $d(o_1, o_2)$. The authors of MSN extended this notion to also take into account the intersection of ball regions defined by the queries and the time when the queries were processed.

This extended confusability is then used to choose the most promising peers to forward the processed query to. Firstly, the similar queries (template queries) – according to the extended confusability – are retrieved from the query history of the current peer. Afterwards, the processed query is forwarded to the acquaintances associated with the template queries. The routing stops when the maximum hop count is reached.

When the query is processed, the pieces of the answer are ordered by the amount of images they contributed to the answer. Upon this ordering, the metadata about the processed query – the new relationships among the peers – are created and sent to the particular peers for storage in their query history. Thus the system sees good answers in those that comprise from large amounts of objects.

3.1.2 Properties

In the MSN, there is no automatic exploration implemented, no background actions are done by peers automatically, and peers do not exchange any summary profiles about their data. Peer's knowledge is gradually built and improved using just the information about results of processed queries. The system follows the properties characteristic for good self-organized and robust systems.

Self-organization – MSN organizes itself in a way that local knowledge is sufficient to locate relevant data sources.

Self-adaptability – a self-adaptable network should subsequently evolve and tune its parameters according to changing user requirements or data in terms of adding or deleting objects but also in terms of joining and leaving peers. As MSN is

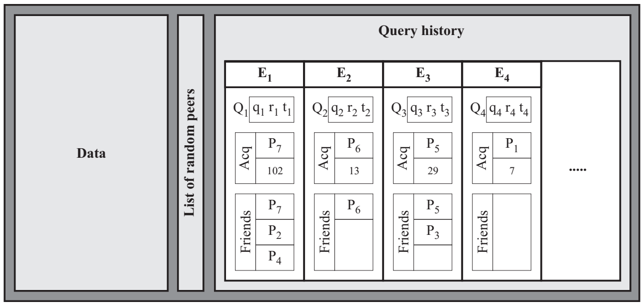

Figure 1: Architecture of peer of the proposed IRN.

based on the query-answer paradigm, peers' knowledge is still evolving and improving when new queries come. Obsolete queries stored in query histories are gradually removed and information about peers' joinings and failures is quickly propagated through the network.

Robustness – a robust system has no single point of failure, is resistant against sundry errors, and can repair or correct damage without any external help. Entities of the system can automatically detect, diagnose and recover from failures as a result of internal errors or external inconsistencies. The MSN has no single point of failure because all peers are equal. The routing algorithm is resistant against critical failures of the peers – when a contacted peer does not respond, a query is forwarded to another one.

3.2 Proposed Enhancements

In order to make the MSN work as a stand-alone network, we firstly enhance the design with an ability to locally maintain data. Additionally, we introduce random factors to achieve better search results. As we mentioned above, MSN relies on a data clustering provided by the underlying indexing structure. Finally, we introduce a protocol to allow peers to join the network. The peer architecture of the proposed IRN is depicted in In Figure 1.

3.2.1 Local Data Organization

Each peer of the proposed network is capable of maintaining and organizing its data. For the data organization, any indexing structure suitable for metric data or even sequential file can be used. In our implementation, we used M-tree [5].

3.2.2 Randomness

As was mentioned above, the MSN was designed as a Semantic Overlay Network over an existing structure. It relied on a fact that the underlying structure redistributed the data among peers in a way, that the data was as much clustered as possible. That is not possible if we want to maintain the data-owner orientation of IRN. Therefore we had to introduce random factors to the design to make the localization of proper peers more robust in order to process a posed query. Our system exploits random factors in a way that the routing algorithm contacts not only the most promising peers but also random peers.

Initialization.

Initially, when a new peer joins the network, its list of random peers is empty and it has to obtain some. Albeit the list is empty, the peer must know at least one another peer (because of its successful joining the network). In the same situation, any peer can end up

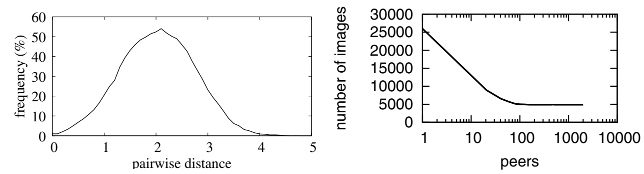

Figure 2: (a) Distance density of the image dataset. (b) Number of photos per peer (user).

since its list of random peers can get underfilled or invalidated due to dynamics of peer participation (churn). So, these peers obtain new random peers by sending a request to all their known peers. A peer receiving such a request decides with a random function whether to respond to this request or not, and then forwards the request to all its random peers. If a peer responds to this request, it is inserted to the list of random peers of the peer which originated the request. This is repeated until TTL (time-to-live) is zero. We set TTL to four. We empirically verified that 30-50 random peers is sufficient for networks consisting of hundreds or thousands of peers.

A $Peer$ deletes another $Peer_x$ from its list of random peers if $Peer_x$ is considered as invalid (e.g. after it leaves the network) or if $Peer_x$ has been identified as a friend or an acquaintance of $Peer$ for three times. This limit can also be interpreted as $Peer$ has three entries in its query history which point to $Peer_x$.

Randomness in Routing Algorithm.

At each step of query forwarding, the query Q is also forwarded to a random peer with a certain probability. This probability is inversely proportional to the value of maximum confusability of the template queries and the processed one. For instance, if the maximum confusability is equal to 0.73, the query is also forwarded with the probability of $27\,\%$ to a peer chosen from the list of random peers. The random peer is always used in the case when the peer has less than five entries stored in its query history. To avoid the infinite contacting of random peers, the query forwarding is stopped when a predefined number of hops is reached. After analyzing the peers' answers by the peer that initiated the query processing, the created entry E corresponding to the processed query is stored in the query history of: the peer that initiated the query, all friends and also peers that were contacted in the random manner during routing.

3.2.3 Bootstrap Protocol

We assume that a new peer which is willing to join the network is given few addresses of peers already participating in the network. In order to allow the peer to ask queries, it has to obtain a list of random peers. This procedure has been described in Section 3.2.2. After successful filling the list of random peers, it has to issue a few queries to the network asking for its own data, in order to share it. In this way, the network is informed about the data newly available. This procedure of data announcement can be optimized by selecting the queries carefully – using centers of data clusters is an effective option, for example.

4. EXPERIMENTAL EVALUATION

In this section, we present our experience with the proposed self-organized network used for content-based image retrieval. We have used real-life data represented by extracted features from 10 million images acquired from the Flickr [28] Photo Sharing system.

4.1 Contrasting IRN and Related Systems

Even though Section 1.1 contains a lot of relevant systems that also do the content-based information retrieval on images, it is very difficult to compare our system with them. Firstly, it is due to the fact that the foundations of those systems are not publicly available and secondly, because they use different data and the scalability issues are unclear so it is not obvious that those systems would be able to handle such amount of data. Secondly, our proposed network is intrinsically unstructured what makes it among those systems unique. This also makes it incomparable with M-Chord structure, what is also the reason why we cannot compare the IRN with the MSN semantic overlay network. Therefore, we try to measure the network's properties using self-contained measures that would give the reader the insight into the network's performance.

4.2 Data

The object in the dataset is the extracted characteristic features of the original image and is represented as one 280-dimensional vector. The distribution of the dataset is quite uniform and such a high-dimensional data space is extremely sparse. With respect to the network size, the number of vectors in the dataset is up to 10 million. Figure 2 (a) depicts distance density of the largest dataset using the aggregated distance function.

In this dataset, the original user that uploaded the image to Flickr is known, thus each peer in the network represents one user. Each peer then contains images of one user. Figure 2 (b) demonstrates the photo count per user in the image dataset, please notice that the x-axis is in log scale. By the hardware infrastructure used, the maximum number of peers was 2,000.

4.3 Measures

The properties of the answer to a processed query have been quantified using following measures:

Recall a ratio of the size of the approximate answer gained from the IRN to the size of the total answer in percent;

Costs a number of peers contacted by the routing algorithm of the IRN in order to process the particular query;

EOPP is normalized error on peers' positions gaining values from the interval $[0, 1]$ which expresses the accuracy of approximate search. The approximate and the total answer are ordered lists of peers (that participated in the answering) by the distance of peer's nearest object to the query object. The formula is defined as:

$$EOPP = \frac{\sum_{i=1}^{|Approx|} Total[Peer_i] - Approx[Peer_i]}{|Approx| \cdot |Total|}$$

where $Approx$ stands for the approximate answer and $Total$ for the total answer. The $Approx[Peer]$ denotes the position of the peer $Peer$ in the ordered list $Approx$. For illustration, if $EOPP = 0$, the approximate answer is a prefix of the total answer. The higher value $EOPP$ has, the more inaccurate the answers are. If the approximate answer was found on 10 peers, and the total is located on 20 peers and $EOPP = 0.1$, the first two peers from the total answer are missed. In the case of $EOPP = 0.5$, the first half of peers from the ordered total answer is missed. Concerning the legibility purposes, notice that the EOPP values are in figures multiplied by 100;

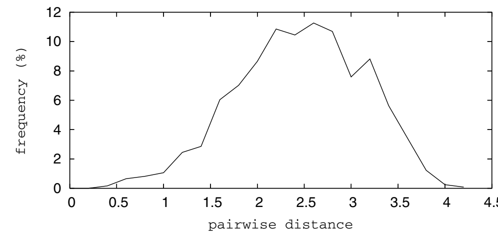

Figure 3: Distance density of the set of 50 training queries.

IRN-NER a number of peers that returned a Non-Empty Result (non empty peer's answer) to the processed query by IRN;

Total-NER a number of peers of the network that have non-empty answer to the processed query;

NER-ratio a ratio of *IRN-NER* and *Total-NER*.

The NER measures are used to justify the costs of the proposed IRN. The EOPP measure is used to measure the goodness of the approximation and to give another point of view on the quality of the retrieved answer other than the recall.

4.4 Self-Organization and Self-Adaptability

The self-adaptability of the proposed network represents the ability to process unseen queries well and to improve response to queries processed repeatedly. Ideally, the recall of the newly processed query should not be lower than the random access and should be acquired with considerable costs. As for repeating queries, their recall should rise together with lowering costs.

To analyze the self-adaptability of the network to newly coming and repeating queries, the following experiment has been conducted. The methodology is to iteratively execute a series of queries whose metadata is stored to the particular peers' query history – *training queries*. This part can be envisioned as learning the network – similar to the learning process in neural networks. After each executed training series, the subset of these queries is used to measure the adaptability of the network – *measuring queries*. The measuring process does not affect the state of the network since the metadata is not stored to the peers' history. Moreover, for each particular query regardless to the state of the network, the peer issuing the query was chosen randomly. The distance density of the set of training queries is in Figure 3.

To give the reader the insight into the dependence of the network size onto the self-adaptability, two networks were used. First one had the size of 2,000 peers and indexed 10 million digital images. The second network comprised of 500 peers and indexed 2.5 million images and was acquired by taking the first 500 peers from the larger network.

Figures 4 and 5 depict the averages measured on these two networks. The set of *training queries* had a size of 50 and the *measuring queries* was a subset of 20 queries from the training set. All the queries were range queries with the radius $r = 1$. On average, the total answer comprised of 1,708 objects on the smaller network and 6,880 on the large one.

Progress of the recall curve in both cases has the same ascending trend. The initial recall in the smaller network is larger than in the larger one – smaller network randomly accessed 52% of its peers and retrieved about 47% of the total answer whereas the larger network accessed about 28% of its peers and achieved recall of 32%. After 20 iterations, the achieved maxima of recall are 80% in the

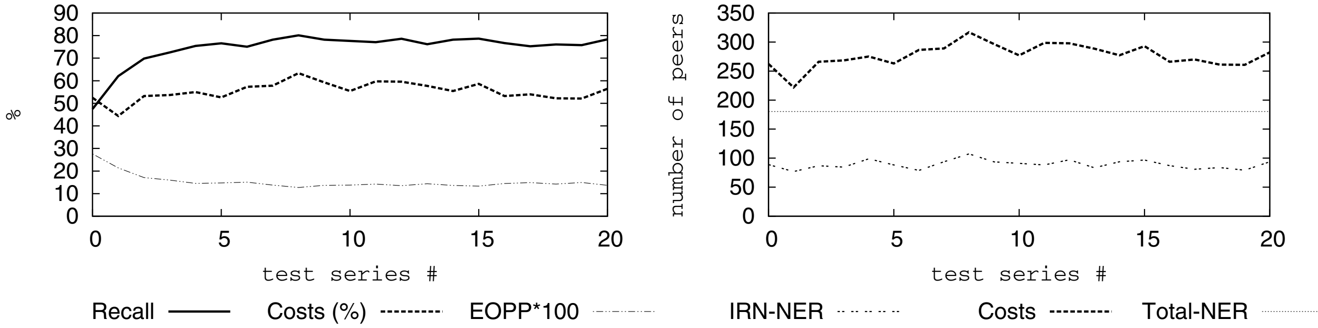

Figure 4: Self-adaptability of the smaller network.

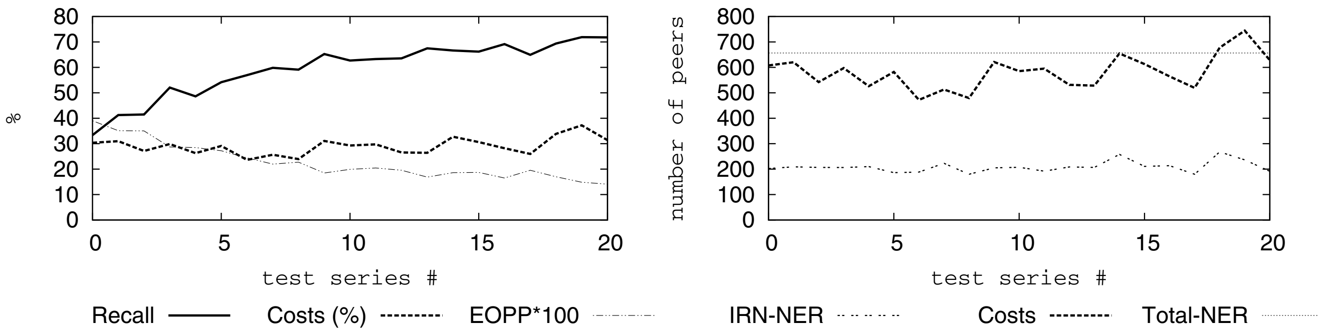

Figure 5: Self-adaptability of the larger network.

smaller network and 71% in the larger one. Notice that in Figure 3 it is illustrated that most of the queries from the training set have their mutual distance greater than one. From the routing algorithm point of view, the presence of metadata about such queries in peers' history does not affect the routing itself – it seems that the queries are processed independently on each other.

To compare the costs, the total answer to each of the training queries was obtained by flooding the whole network. The average value measured is depicted in Figure 4 as the Total-NER. As can be observed in the case of the smaller network, the total answer is spread on average over 36% of the total amount of peers. The image retrieval network has to contact on average 55% of the network after 10 iterations to locate 77% of the total answer. In the case of the larger network, see Figure 5, the total answer is spread on 33% of the peers in the network and the search algorithm contacted on average after 10 iterations 30% of the network to reach 60% of the recall. After 20 iterations, the recall on the larger network is on average 72%. The recall achieved grows slower on the larger network which is due to slower learning ability due to lower proportions of peers contacted in order to locate the answer.

As we see, the numbers measured are proving the scalability of the proposed system, since the spread of the answer among peers grows linearly, the recall does not degrade significantly and still, we observe a decrease in the costs.

The quality of approximation is studied by the EOPP coefficient. Ideally, the EOPP should remain in low numbers what would mean that the routing algorithm contacted peers that contain the most similar parts of the total answer. As can be observed in Figures 4 and 5, the progress is auspicious. The trends should be strictly decreasing, what means that the algorithm retrieves from the network in each step more relevant parts of the answer. As we see the EOPP of the IRN follows such progress.

4.5 Browsing Scenario

Besides the laboratory conditions, we present the experiments that involve a real-life scenario. We have conducted experiments with real users that performed *browsing* in the IRN and we recorded the results gained. The users were asked to start their browsing session using one of 50 images pre-picked from the dataset and after getting the answer to their query, they randomly chose another query image from that answer. This browsing step was repeated for at least 14 times making each of the browsing session containing at least 15 query images. Figure 6 visualizes the browsing interface and one browsing step. One image from the answer to a processed query is chosen as the next query object by clicking on the selected photo. Each user queried the IRN from its own peer. For brevity, we present the results of the browsing experiment only on the larger network.

Figure 7 demonstrates the results of 11 browsing sessions where the results of the corresponding browsing steps are averaged. The zeroth step represents the results gained by processing the query with the first seed photo. Figure 8 gives an insight what was the average distance between the seed photo and a browsed photo in each of the 14 following steps. As we see, the distance usually started from 0.9 in the first step and was on average 1.9 in the last one. The recall presented in Figure 7 came close to 50% only. If we assume that only a pair of query points in two consecutive steps is relevant for future routing, it is obvious that we cannot get better results than roughly 50% of recall. A further training of the network would solve this problem. On the other hand, we can observe

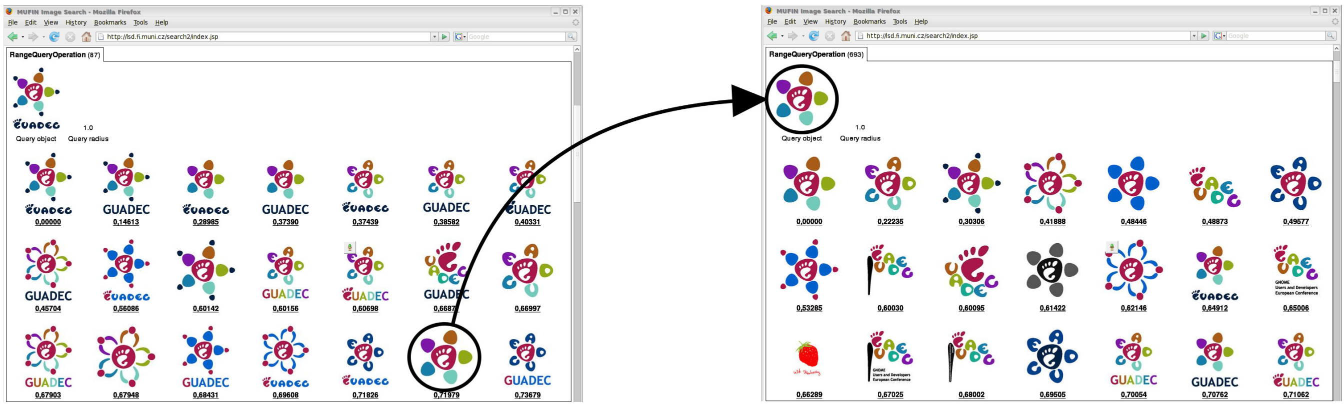

Figure 6: Browsing in the photo database using the results of the previous query in the session.

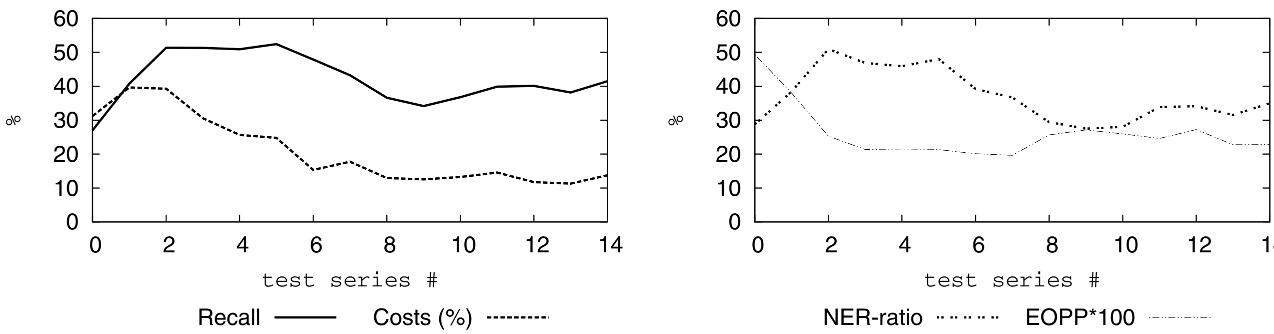

Figure 7: Browsing experiment results.

a dramatic decrease in costs which is attributed to the fact that all queries in each session were issued from the same peer and thus the query history gets filled faster – the peer became more knowledgeable. Before processing the queries in the browsing session, the network was in the initial state – it did not process any training queries.

This experiment was carried out to see how the network reacts when one particular peer processes similar queries. The results showed that the system can react to processing similar queries by a considerable decrease of costs. Because the query objects were still distant, the recall did not improve that dramatically. We attribute this to the quantitative measuring of query answer quality – it favors those pieces of answers (of individual peers) that contain more objects, rather those that whose objects are closer to the query object.

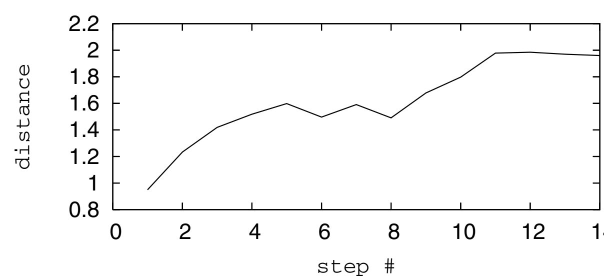

Figure 8: Average distance from the first query image in the browsing session.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a distributed system for a content-based search in multimedia data. The design of the system takes its inspiration in social networks and thus the proposed approach exhibits the ability to self-organize itself towards better processing of repeating queries, and to self-adapt in order to change in the users' tastes or the intrinsic change of data in time. With randomized access we enhanced the previously published system to work as a stand-alone network. This concept has been proven using a digital image dataset that comprised of 10 million images provided by Flickr.

The design of the network is also strictly data-owner oriented, each participant – a peer in the network – maintains its data contribution to the shared database – the Image Retrieval Network – so no data redistribution in order to search the shared database is required. This search approach implies an intrinsic extent of approximation of the query processing that is analyzed by a series of experiments concluding in promising results.

Due to the adopted metric space paradigm, the proposed network is extensible for the search in other multimedia, like music or video clips, or in other non-trivial data types. It is just a matter of a suitable descriptor accompanied with a distance function.

As for the ongoing and future work, the first thing that seems to improve the overall performance of the proposed system, is changing the way of measuring the quality of the answer. This is now done by counting the quantity of the retrieved objects. Instead, we intend to use an approach based on the proximity of retrieved objects to the query object. Hand in hand with this effort goes an implementation of the processing of nearest neighbor type of queries.

Another issue that attracts our attention is the deployment of the network in larger setting than the tested one consisting of 2,000 peers. Due to the infrastructure limitations we have been able to make the experiments only on a network of such a limited size, yet it still formed a shared database of a considerable size. We also plan to thoroughly test the anticipated robustness of the system in terms of its ability to react onto changes in the data in the database. Meaning also the change of peers – joining and leaving the network – and also the change of the peer's data.

By addressing all these efforts we would like to make the network able to work in real world setting and enable its users to share their multimedia data for a content-based search in the web-scale.

6. ACKNOWLEDGEMENTS

Partially supported by the EU IST FP6 project 045128 (SAPIR), the national research project 1ET100300419, the Czech Grant Agency projects 201/07/P240 and 102/05/H050.

7. REFERENCES

[1] ALIPR. http://www.alipr.com/, June 2008.

[2] M. Batko, D. Novak, F. Falchi, and P. Zezula. On scalability of the similarity search in the world of peers. In *Proceedings of INFOSCALE 2006, Hong Kong*, pages 1–12, New York, 2006. ACM Press.

[3] M. Bender, T. Crecelius, M. Kacimi, S. Michel, J. X. Parreira, and G. Weikum. Peer-to-peer information search: Semantic, social, or spiritual? *IEEE Data Eng. Bull.*, 30(2):51–60, 2007.

[4] Chorus. http://www.ist-chorus.org/, June 2008.

[5] P. Ciaccia, M. Patella, and P. Zezula. M-tree: An efficient access method for similarity search in metric spaces. In M. Jarke, M. J. Carey, K. R. Dittrich, F. H. Lochovsky, P. Loucopoulos, and M. A. Jeusfeld, editors, *Proceedings of the 23rd International Conference on Very Large Data Bases (VLDB 1997), Athens, Greece, August 25-29, 1997*, pages 426–435. Morgan Kaufmann, 1997.

[6] R. Datta, D. Joshi, J. Li, and J. Z. Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys*, 2008. To appear.

[7] ESP Game. http://www.espgame.org/, June 2008.

[8] ExaLead. http://www.exalead.com/, June 2008.

[9] C. Gennaro, M. Mordacchini, S. Orlando, and F. Rabitti. Processing complex similarity queries in peer-to-peer networks. In *SAC '08: Proceedings of the 2008 ACM symposium on Applied computing*, pages 473–478, New York, NY, USA, 2008. ACM.

[10] Google Image Labeler. http://images.google.com/imagelabeler/, June 2008.

[11] ImBrowse. http://media-vibrance.itn.liu.se/vinnova/cse.php, June 2008.

[12] J. L. Kelly. *General Topology*. D. Van Nostrand, New York, 1955.

[13] J. Li and J. Z. Wang. Real-time computerized annotation of pictures. In *Proceedings of MULTIMEDIA '06*, pages 911–920, New York, 2006. ACM Press.

[14] A. Linari and M. Patella. Metric overlay networks: Processing similarity queries in p2p databases. In *Proceedings of the 5th International Workshop on Databases, Information Systems and Peer-to-Peer Computing (DBISP2P 2007)*, 2007.

[15] A. Linari and G. Weikum. Efficient peer-to-peer semantic overlay networks based on statistical language models. In *P2PIR '06: Proceedings of the international workshop on Information retrieval in peer-to-peer networks*, pages 9–16, New York, NY, USA, 2006. ACM Press.

[16] B. Manjunath, P. Salembier, and T. Sikora, editors. *Introduction to MPEG-7: Multimedia Content Description Interface*. John Wiley & Sons, Inc., New York, NY, USA, 2002.

[17] D. Novak and P. Zezula. M-Chord: A scalable distributed similarity search structure. In *Proceedings of INFOSCALE '06, Hong Kong*. ACM Press, 2006.

[18] W. Penzo, S. Lodi, F. Mandreoli, R. Martoglia, and S. Sassatelli. Semantic peer, here are the neighbors you want! In *EDBT '08: Proceedings of the 11th international conference on Extending database technology*, pages 26–37, New York, NY, USA, 2008. ACM.

[19] PicSearch. http://www.picsearch.com/, June 2008.

[20] H. Samet. *Foundations of Multidimensional And Metric Data Structures*. The Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann, 2006.

[21] J. Sedmidubský, S. Bartoň, V. Dohnal, and P. Zezula. Adaptive approximate similarity searching through metric social networks. Technical Report FIMU-RS-2007-06, Faculty of Informatics, Masaryk University Brno, http://www.fi.muni.cz/reports/files/2007/FIMU-RS-2007-06.pdf, November 2007.

[22] R. N. Shepard. Toward a universal law of generalization for psychological science. *Science*, 237(4820):1317–1323, 1987.

[23] Tiltomo project. http://www.tiltomo.com/, June 2008.

[24] P. Triantafillou, C. Xiruhaki, M. Koubarakis, and N. Ntarmos. Towards high performance peer-to-peer content and resource sharing systems. In *CIDR*, 2003.

[25] R. C. Veltkamp and M. Tanase. Content-based image retrieval systems: A survey. Technical Report UU-CS-2000-34, Department of CS, Utrecht University, http://give-lab.cs.uu.nl/cbirsurvey/, 2002.

[26] J. Z. Wang, J. Li, and G. Wiederhold. SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(9):947–963, 2001.

[27] Yahoo! http://www.flickr.com/map/, June 2008.

[28] Yahoo! http://www.flickr.com/, June 2008.

[29] P. Zezula, G. Amato, V. Dohnal, and M. Batko. *Similarity Search: The Metric Space Approach*, volume 32 of *Advances in Database Systems*. Springer, 2005.