

DISCUSSION ON BENCHMARKING OF MEDIA SEARCH BASED ON P2P OVERLAYS

Michal Grega

Mikolaj Leszczuk

AGH University of Technology
Krakow, Poland
grega@kt.agh.edu.pl
leszczuk@kt.agh.edu.pl

Ruben Cuevas Rumin

Isaias Martinez Yelmo

Universidad Carlos III de Madrid
Madrid, Spain
rcuevas@it.uc3m.es
imyelmo@it.uc3m.es

Adriano Fiorese

Universidade de Coimbra
Coimbra, Portugal
University of the State of Santa Catarina
Joinville, Brazil
fiorese@dei.uc.pt

Nicolas Liebau

Technische Universitaet
Darmstadt
Darmstadt, Germany
Nicolas.Liebau@KOM.tudarmstadt.de

Siyu Tang

Delft University of Technology
Delft, The Netherlands
S.Tang@ewi.tudelft.nl

Abstract: This paper, being a summary of [1] presents discussion points in the area of measurement of the performance of the multimedia search service in the P2P overlay networks. As the P2P technology is becoming a popular mean for delivery of multimedia content, the problem of effective search for this content is becoming an object of interest of numerous research groups. This research effort needs a common framework for tuning of search algorithms and evaluation of effectiveness. An overview of the proposed benchmarking system and extended version of this paper is presented in [1].

1. MOTIVATION AND CHALLENGES

The P2P overlays are gaining on popularity as a mean of access to large amounts of multimedia data. A typical P2P system consists of two parts — the search tool and the delivery mechanism. One of the difficulties in designing good mechanisms for distributed autonomous systems is a lack of a unified process for evaluating the efficiency of mechanisms, both in the research community and in the industry. As content distribution systems have widely recognised standard overlays to compare with, in search there is no such base. In this paper we will focus on benchmarking search mechanisms for P2P systems.

The existing benchmarks for evaluation of search performance (mainly accuracy) were designed and are utilised in the local repositories of media. Such metrics cannot be directly applied in the P2P environment because of its nature.

P2P environments are, first of all, characterised by a significant delay in communications. This makes the aspect of search time, which was regarded not important in local repositories of media, worth measuring. The presented approach to the measurement of search time is based on the detailed analysis of single components of this attribute of the search system. P2P systems are distributed by nature — so the media can be separated from its meta-data and both can exist in the network in multiple copies. The high dynamic of the P2P environments may cause the media or the meta-data inaccessible during the search process due to the effect of churn. These properties may influence both the accuracy and cost of search and need to be taken into account while developing the metrics.

A benchmark can be defined as “a standardized problem or test that serves as a basis for evaluation or comparison (as of computer system performance)” (according to the Merriam-Webster English dictionary). The goal of benchmarking is to assess the quality of the benchmarked system and to allow comparison to other, similar systems. Researchers,

especially in the computer science, are used to perform benchmarking according to either official or unofficial standards.

2. DISCUSSION POINTS

- **What to benchmark?**
There are several aspects which can be benchmarked: Search Accuracy, Time, Cost and the interdependences of those
- **How to benchmark?**
How to create a benchmarking environment for the P2P overlays? How to allow for benchmarking in different types of overlays?
- **To use a simulator or to benchmark a real network?**
On one hand – benchmarking using a simulation only does not reflect the real life scenario in 100%. On the other – benchmarking of some parameters in the real network can be very difficult if not impossible
- **Which search services should be benchmarked?**
P2P overlays allow introduction of advanced search services, such as Query by Example, due to their high, potentially, computing power and distribution of media. Should the benchmarks take into account more advanced search systems than simple text queries?
- **How to detect all tradeoffs while creating Benchmarking Framework?**
Some tradeoffs are obvious – such as the Search Accuracy vs. Search Time tradeoff. Detection of other – inter-metric and intra-metric tradeoff may be difficult without a test deployment of the Benchmarking Framework

REFERENCES

- [1] M. Grega, M. Leszczuk, R. C. Rumin, I. Martinez-Yelmo, A. Fiorese, N. Liebau, S. Tang „Benchmarking of Media Search Based on P2P Overlays” *Submitted for the 1P2P4mm workshop, Infoscale 2008, Vico Enquense, Italy.*