connecting other nodes to the core network and to each other. The connections are established using WDM (Wavelength Division Multiplexing) technology.

Furthermore, a network-monitoring system for the OAN network was designed. In the design the usability of Simple Network Management Protocol (SNMP), standardised Management Information Bases (MIBs), and ready-to-use management software was considered.

Next-Generation SONET/SDH in Use

The next-generation SONET/SDH enables new types of services with more efficient network usage to be easily implemented by utilising existing infrastructure.

Corporations require diverse services (eg voice, VPN, data storage, and Internet connection services) from operators. Traditionally the different services are provided through technology-specific transport pipes. However, the next-generation SDH enables the simultaneous transport of heterogeneous services over one wavelength, thereby saving network-building and maintenance costs.

Usually a virtual private connection (VPN) is used to bridge operators' access points. In some applications however, it is desirable to transport the native network signal without extracting packets or frames. Normally the datacom protocols rely on 8B/10B coding, which causes a 25 percent increase in bandwidth. Using the next-generation SDH, which maps 8B/10B-coded data into 64B/65B-coded sequences, the required bandwidth is substantially decreased.

The ability to dynamically reallocate bandwidth allows Bandwidth on Demand (BoD) services. This will revolutionise the network service industry, since the users are able to specify their bandwidth requirements according to the time of day.

Link:
http://www.vtt.fi/tte/tte21/oan/

Contact:
Kari Seppänen
VTT Information Technology, Finland
Tel: +358 9 456 5610
E-mail: Kari.Seppanen@vtt.fi

KOPI – An Online Plagiarism Search and Information Portal

by Máté Pataki

Due to the growing number of digitally available documents, the problem of plagiarism is becoming increasingly serious. The aim of the KOPI project, carried out at the Department of Distributed Systems (DSD) of SZTAKI, was to develop an online plagiarism search portal, which would help digital libraries to protect their documents, and teachers to identify copied texts or publications.

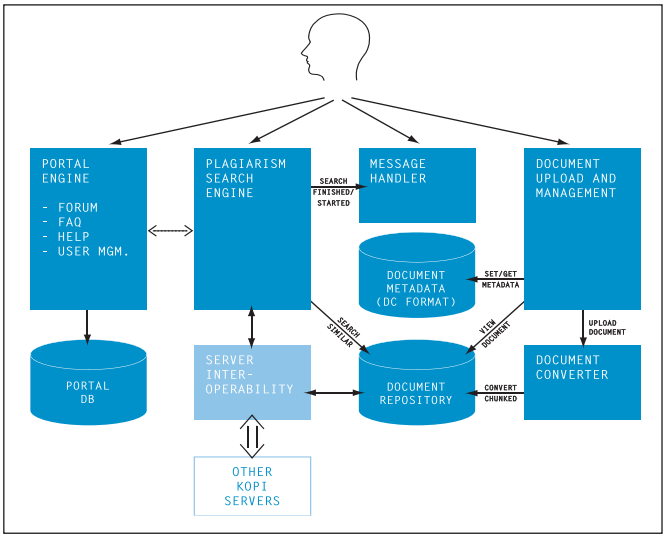
The work was performed as part of a joint project between SZTAKI and Monash University in Melbourne, with a series of investigations being made on document chunking and overlap detection, two techniques on which the detection of plagiarism is based. Continuing this work, development of the portal KOPI commenced in 2003, funded by the Hungarian Government. The portal will become available to users by the end of June 2004.

There are two different approaches to fighting plagiarism. The first is the protection of the document by preventing it from being copied or misused, and the second is the recognition of plagiarism. Protection is an important issue but it can cause difficulties for legal users. Moreover, all kinds of protection will be cracked in time. According to our view, the most effective technique in fighting plagiarism is the fast detection of document overlapping: in other words, there is no sense in copying a digital document if the copy can be detected within minutes. This method is used to protect documents that are part of the KOPI system from illegal use.

As a portal site, KOPI includes common services such as a forum, context-sensitive help, FAQ, and static documents including information on plagiarism and university laws. In addition, two system-specific services are offered to users of the portal: a document upload and management service, and the plagiarism search engine. The first can be used to upload documents (html, rtf, doc, pdf, txt) or a batch of documents (zip), and to attach meta-information to them. The meta-data are stored in Dublin Core meta-data format to make possible future interoperability with other systems.

The uploaded documents can then be compared with each other, with previously uploaded documents, with all documents uploaded by users, or with collections of documents gathered from the Web or documents in digital libraries. The comparison is made offline, reducing waiting time and costs for the user. When the job is finished, the message handler unit sends the results to the user via e-mail.

The heart of the similarity search engine is the chunking method, which is used to chunk the given text into smaller pieces. This task and the conversion of the document to plain text are performed by the document converter subsystem. When comparing documents, only these



The structure of the KOPI system.

chunks or their so-called compressed fingerprints are examined to determine how many common parts the documents have. The KOPI system uses a combination of word chunking and overlapping word chunking to chunk the documents. This new algorithm provides a fast and accurate search, while keeping the size of the database small. (For more information on the chunking methods see the links below.)

In order to perform an efficient plagiarism search, the KOPI system needs to collect as many documents as possible. Four possible sources exist:

- documents on Internet
- digital library collections
- publications and theses from schools, universities, or conference organisers
- material uploaded by the users of the KOPI system.

Documents from the Internet are collected using a Web crawler. Digital libraries with an open interface to the Internet (eg OAI) can also be easily harvested. In the future, it is likely that university students will be requested to submit their theses in digital form, and so within a couple of years a large set of documents will have been collected.

KOPI is currently a stand-alone portal application. Future developments in the frame of a PhD project will target the creation of a distributed KOPI architecture. In such a system, institutes would use their own local copy of the KOPI engine, but could initiate a plagiarism search involving documents over the whole distributed KOPI system.

Links:
KOPI portal: http://kopi.sztaki.hu

Department of Distributed Systems of SZTAKI: http://dsd.sztaki.hu

Plagiarism Detection and Document Chunking Methods, The Twelfth International World Wide Web Conference: http://www2003.org/cdrom/papers/poster/p186/p186-Pataki.html

Match Detect Reveal Project at Monash University Melbourne: http://www.csse.monash.edu.au/projects/MDR/

The Dublin Core Metadata Initiative: http://dublincore.org/

Please contact:
László Kovács, SZTAKI, Hungary
Tel: +36 1 279 6212

Assessing Quality of Service on the Internet

by Peder J. Emstad

Everyone who uses the Internet and mobile phones has an interest in their quality of service. We may notice delays when we browse the Web and transfer files, and we may wonder if our transactions are secure or if we can connect at all. New services, such as voice and video, are constantly being offered based on packet-switching technology, but provide varying quality for the user. A new centre at the Norwegian University of Science and Technology has been established to determine how to assess this quality of service.

Traditionally, telecommunication networks have been established to transport voice sig-nals between geographically disparate users, and for that purpose a worldwide network, albeit of widely varying technical quality, has emerged. However, the days of networks dominated by analogue voice signals are long gone. Today's network of networks, the Internet, has become the general vehicle for the transport of digital data representing all kinds of information and information services. These services are realised through the exchange of digital data between end-users, and between end-users and service providers. They cover most aspects of human activities, both private and professional: general information exchange, e-mail, voice and multimedia services, transfer and storage of medical information, economic transactions, location information – the list is already seemingly endless. In the industrialised world, everyday life as we know it would cease to exist if the Internet were to become permanently inoperable.

In order to address the problem of Quality of Service (QoS), the Centre for Quantifiable Quality of Service in Communication Systems has been established at the Norwegian University of Science and Technology. Research will be based around the continued evolution of packet-switching techniques for mobile networks and the Internet . Society's use of and dependence on such networks increases steadily, even though the QoS in a broad sense is unsatisfactory and at best variable. Services to be