R&D and Technology Transfer

From a research point of view, the task of text clustering presents a great challenge, especially in a multilingual context. While a number of document-clustering techniques exist, they all lack the fundamental ability to provide sensible descriptions (labels) of the output document groups. This has been the primary focus of the Carrot² project – to extract sensible groups of documents on related topics, but most of all to provide a short, comprehensive description of these clusters.

At the time of writing, the project includes a number of original text-clustering algorithms and auxiliary components for text processing. In 2004, Carrot² was awarded a special prize for research tools in the finals of the European Academic Software Award competition. The rough-set-based clustering algorithm included in the project received best paper award at the 2005 IEEE Web Intelligence conference. We are also proud to have a number of deployments worldwide, many references in research literature and a few sibling open-source projects using Carrot² components. Constantly growing commercial interest in text-clustering services and algorithms resulted in the establishment in 2005 of a spin-off company, Carrot Search. The company took over the maintenance and further development of the project.

Links:
http://www.carrot2.org
http://www.carrot-search.com

Please contact:
Dawid Weiss
Institute of Computer Science, Poznan University of Technology, Poland
E-mail: dawid.weiss@cs.put.poznan.pl

Figure 1: Information flow inside Carrot2 – a set of search results is clustered into topic groups and then shown back to the user in a variety of ways (hierarchy of topics, graph of relationships, etc).

Quantifying WiMAX Performance

by Kostas Pentikousis, Ilkka Harjula, Esa Piri and Jarno Pinola

According to some estimates, by 2010 WiMAX operators will cover areas inhabited by more than 650 million people. Yet current deployment lags similar previous expectations and predictions. Moreover, non-vendor, third-party empirical evaluations of the technology are far from common. We recently concluded our VoIP and IPTV synthetic traffic measurement and analysis studies over fixed WiMAX at the VTT Converging Networks Laboratory in Oulu, Finland, and briefly report our results here. Moreover, we review our simulation-based studies of extending WiMAX in demanding, mountainous terrains, which are not considered to be the main target deployment environments.

WiMAX, often cited as technology that could serve as the substrate for next-generation mobile broadband networks, stands for Worldwide Interoperability for Microwave Access and is based on IEEE 802.16 standards. WiMAX networks can provide point-to-point and point-to-multipoint broadband IP connectivity to both fixed and mobile hosts, with Quality of Service (QoS) guarantees and robust security.

In theory, and according to vendor field trials and demonstrations, WiMAX can deliver cell bitrates greater than 100 Mb/s, covering large areas (up to 50 km radius from a single base station site using directional antennae), and serving tens of subscribers. These are impressive figures. However, the currently available, commercial off-the-shelf (COTS) equipment delivers significantly less application-layer throughput. At VTT we empirically evaluated fixed WiMAX, aiming to improve our understanding of what is realistically possible using COTS equipment. In particular, we employed the VTT Converging Networks Laboratory (CNL) infrastructure, which includes a fixed WiMAX base station (BS) and two subscriber stations (SS), and studied Voice over IP (VoIP) and live IPTV streaming uplink and downlink performance (see Figure 1). Fixed WiMAX was used both as backhaul for voice and data services as well as a last-mile network access technology. CNL is connected via GEANT2 to the Internet.

We employed multiple competing traffic sources over a point-to-multipoint
WiMAX topology; its capacity was measured in terms of number of synthetic bidirectional VoIP ‘calls’ between subscriber stations while concurrently delivering a variable number of video streams with negligible loss. For VoIP, we considered several scenarios using both Speex and G.723.1 codecs. The video stream was captured from a live TV channel transmission and retransmitted in H.264/AVC format.

We measured throughput, packet loss and one-way delay for both line-of-sight (LOS) and non-line-of-sight (NLOS) conditions. We also calculated mean opinion scores (MOS) based on the ITU Telecommunication Standardization Sector (ITU-T) E-model, for the experiments with G.723.1. We accurately measured one-way delay by employing a software-only implementation of the IEEE 1588 Precision Timing Protocol (PTP). Finally, in order to put our results in perspective, we repeated the measurements after adding IEEE 802.11g access points (APs) in the topology of Figure 1. We found that VoIP flows carrying single sample payloads are clearly underperforming and that application-layer VoIP aggregation can more than triple the number of lossless VoIP flows in the downlink without any network or hardware support. Moreover, our results indicate that a single WiMAX subscriber station can backhaul VoIP traffic from at least two Wi-Fi APs. Due to space constraints we cannot cover all our results in this article, and refer readers to the peer-reviewed publications available from the links below.

Besides empirically quantifying COTS WiMAX performance, we also used Monte Carlo simulations to study the performance of the WiMAX physical layer in more challenging radio environments. Radio channel properties in isolated mountainous areas differ significantly from those considered in the main target environments of the WiMAX system specifications. We therefore studied promising ways of compensating for performance loss due to environmental factors by steering the direction of the receiver antenna beams. First, the radio channel properties in the mountainous environment around the Vesuvius volcano in southern Italy were analytically derived based on the bistatic radar equation and the geological properties of the area. Careful analysis revealed that the presence of the mountain increases the length of the channel delay spread significantly as well as spreading the received signal in the spatial domain. This model was merged with the WINNER I channel model in Matlab in order to generate a model that could be used in computer simulations.

Consequently we studied the ability of several MIMO algorithms, especially the so-called pre- and post-FFT EVD (Fast Fourier Transform Eigenvalue Decomposition) beamformers, to compensate for performance loss caused by
DEPLOY: Industrial Deployment of Advanced System Engineering Methods for High Productivity and Dependability

by Alexander Romanovsky

The work of the FP7 ICT DEPLOY Integrated Project is driven by the tasks of achieving and evaluating industrial take-up, initially by DEPLOY industrial partners, of DEPLOY methods and tools, together with the necessary further research on methods and tools.

Formal engineering methods enable greater mastery of complexity than do traditional software engineering processes. It is the central role played by mechanically-analysed formal models throughout system development that enables mastery of complexity. As well as leading to big improvements in system dependability, greater mastery of complexity leads to greater productivity by reducing the expensive test-debug-rework cycle and by facilitating increased reuse of software.

The successful three-year FP6 STREP RODIN project on Rigorous Open Development Environment for Complex Systems (2004-2007, http://rodin.cs.ncl.ac.uk/) researched and developed industrial strength methods and tools paving the way for the technology to be deployed. In particular, RODIN delivered an extensible open source platform, based on Eclipse, for refinement-based formal methods along with a body of work on formal methods for dependable systems. DEPLOY exploits and builds on these results.

In DEPLOY five leading European companies, representing five major sectors: transportation (Siemens), automotive (Bosch), space (Space Systems), telecommunication (Nokia) and business information (SAP), will deploy advanced engineering approaches to further strengthen their development processes in order to improve competitiveness.

Objectives
The overall aim of DEPLOY is to make major advances in engineering methods for dependable systems through the deployment of formal engineering methods. The work is driven by the tasks of achieving and evaluating the industrial take-up of the DEPLOY methods and tools, initially in the five sectors which are key to European industry and society.

The aim will be achieved with a coherent integration of scientific research, technology development and industrial deployment of the technology. The complementary expertise and technological base of the industrial deployment partners and the technology provider partners will be combined to achieve a set of challenging scientific and technological objectives.

Consortium
DEPLOY offers a balanced interplay between industrial deployment, scientific research and tool development, where companies in five sectors join their forces with eight technology providers to meet the goal.

The industrial sectors, transportation (Siemens), automotive (Bosch), space (Space Systems), telecommunication (Nokia) and business information (SAP), comprise a palette of important European base industries of today. The companies possess different maturity levels when it comes to deploying formal approaches.

The five academic partners are world leaders in formal methods research, that have considerable experience in developing and applying resilience methods as well as a wide range of formal approaches.

The tool vendors, Systerel and ClearSy, have long-standing experience in developing tool support for formal engineering methods. CETIC has considerable experience in industrial quality measurement and will be in charge of the assessment activities.

The project is coordinated by Newcastle University with a dedicated Project Office set at the School of Computing Science. Project partners are Newcastle University (UK), Aabo Akademi University (Finland), Bosch (Germany), CETIC (Belgium), ClearSy (France), ETH Zurich (Switzerland), Heinrich-Heine Universität Düsseldorf (Ger-