Proceedings of the Doctoral School of Computer Science and Information Technologies

Proceedings of Topics "Use of declarative programming"

Zsolt Zombori, János Csorba, Péter Szeredi: Static Type Inference for the Q language using Constraint Logic Programming

Abstract: We describe an application of Prolog: a type inference tool for the Q functional language. Q is a terse vector processing language, a descendant of APL, which is getting more and more popular, especially in financial applications. Q is a dynamically typed language, much like Prolog. Extending Q with static typing improves both the readability of programs and programmer productivity, as type errors are discovered by the tool at compile time, rather than through debugging the program execution. We map the task of type inference onto a constraint satisfaction problem and use constraint logic programming, in particular the Constraint Handling Rules extension of Prolog. We determine the possible type values for each program expression and detect inconsistencies. As most built-in function names of Q are overloaded, i.e. their meaning depends on the argument types, a quite complex system of constraints had to be implemented.

János Csorba, Zsolt Zombori, Péter Szeredi: Pros and Cons of Using CHR for Type Inference.

Abstract. We report on using logic programming and in particular the Constraint Handling Rules extension of Prolog to provide static type analysis for the Q functional language. We discuss some of the merits and diculties of CHR that we came across during implementation of a type inference tool.

Zsolt Zombori, Péter Szeredi: Loop Elimination, a Sound Optimisation Technique for PTTP Related Theorem Proving. Acta Cybernetica 20 (2012) 441–458.

Abstract: In this paper we present loop elimination, an important optimisation technique for first-order theorem proving based on Prolog technology, such as the Prolog Technology Theorem Prover or the DLog Description Logic Reasoner. Although several loop checking techniques exist for logic programs, to the best of our knowledge, we are the first to examine the interaction of loop checking with ancestor resolution. Our main contribution is a rigorous proof of the soundness of loop elimination.

Zsolt Zombori: Expressive Description Logic Reasoning Using First-Order Resolution

Abstract. Description Logic languages are being used more and more frequently for knowledge representation. This creates an increasing demand for efficient automated DL reasoning. Although several DL reasoners are available, they typically do not scale well with the increase of data. In this paper we present a resolution based approach: we show how to transform description logic axioms to a set of function-free clauses of first-order logic. These axioms can afterwards be used for a well scaling,

query oriented data reasoning. The transformation is given for the SHIQ DL language, but we show how to extend the result to a RIQ DL knowledge base. The method described has been implemented in a module of cthe DLog reasoner openly available on SourceForge to download.

Proceedings of Topics "Use of GPGPU applications"

Balázs Csébfalvi: Cosine-Weighted B-Spline Interpolation: A Fast and High-Quality Reconstruction Scheme for the Body-Centered Cubic Lattice. IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS 19:(9) pp. 1455-1466. (2013)

Abstract: In this paper, Cosine-Weighted B-spline (CWB) filters are proposed for interpolation on the optimal Body-Centered Cubic (BCC) lattice. We demonstrate that our CWB filters can well exploit the fast trilinear texture-fetching capability of modern GPUs, and outperform the state-of-the-art box-spline filters not just in terms of efficiency, but in terms of visual quality and numerical accuracy as well. Furthermore, we rigorously show that the CWB filters are better tailored to the BCC lattice than the previously proposed quasi-interpolating BCC B-spline filters, since they form a Riesz basis; exactly reproduce the original signal at the lattice points; but still provide the same approximation order.

Márton József Tóth, Balázs Csébfalvi: Shape Transformation of Multidimensional Density Functions using Distribution Interpolation of the Radon Transforms. In: 9th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (GRAPP). Lisszabon, Portugal, 2014.01.05-2014.01.08. pp. 5-12.

Abstract: In this paper, we extend 1D distribution interpolation to 2D and 3D by using the Radon transform. Our algorithm is fundamentally different from previous shape transformation techniques, since it considers the objects to be interpolated as density distributions rather than level sets of Implicit Functions (IF). First, we perform distribution interpolation on the precalculated Radon transforms of two different density functions, and then an intermediate density function is obtained by an inverse Radon transform. This approach guarantees a smooth transition along all the directions the Radon transform is calculated for. Unlike the IF methods, our technique is able to interpolate between features that do not even overlap and it does not require a one dimension higher object representation. We will demonstrate that these advantageous properties can be well exploited for 3D modeling and metamorphosis.

Balázs Csebfalvi, Balázs Tóth, Stefan Bruckner, Eduard Gröller: Illumination-Driven Opacity Modulation for Expressive Volume Rendering. In: Vision, Modeling & Visualization. Magdeburg, Germany, 2012.11.12-2012.11.14. pp. 103-109.

Abstract: Using classical volume visualization, typically a couple of isosurface layers are rendered semi-transparently to show the internal structures contained in the data. However, the opacity transfer function is often difficult to specify such that all the isosurfaces are of high contrast and sufficiently perceivable. In this paper, we propose a volumerendering technique which ensures that the different layers contribute to fairly different regions of the image space. Since the overlapping between the effected

regions is reduced, an outer translucent isosurface does not decrease significantly the contrast of a partially hidden inner isosurface. Therefore, the layers of the data become visually well separated. Traditional transfer functions assign color and opacity values to the voxels depending on the density and the gradient. In contrast, we assign also different illumination directions to different materials, and modulate the opacities view-dependently based on the surface normals and the directions of the light sources, which are fixed to the viewing angle. We will demonstrate that this model allows an expressive visualization of volumetric data.

Viktor Vad, Balázs Csébfalvi, Moncef Gabbouj: Calibration of the Marschner-Lobb Signal on CC, BCC, and FCC Lattices. In: EuroVis. Bécs, Austria, 2012.06.05-2012.06.08. pp. 19-23.

Abstract: The well-known Marschner-Lobb (ML) signal has been originally proposed for visually comparing 3D resampling filters applied on the traditional Cartesian Cubic (CC) lattice. Recently, this popular benchmark is also used for evaluating reconstruction schemes designed for the optimal Body-Centered Cubic (BCC) lattice and the suboptimal Face-Centered Cubic (FCC) lattice. Nevertheless, to the best of our knowledge, it has not been thoroughly studied whether the ML signal meets the assumptions that the theory of optimal regular volume sampling is based on. In this paper, we try to find equivalent CC, BCC, and FCC representations for unbiased comparisons. For the continuous reconstruction, we use comparable approximations of the ideal low-pass filter, and increase the sampling frequency until the aliasing effects completely vanish. Based on these experiments, we show that the ML signal is appropriate for comparing the CC and BCC lattices, but it is inappropriate for fairly comparing the FCC lattice to the CC and BCC lattices regarding the visual quality of the corresponding reconstructions. In fact, the ML signal very strongly prefers the FCC sampling due to the special shape of its spectrum. However, this property can hardly be expected from a practical signal.

Balázs Hajagos, László Szécsi, Balázs Csébfalvi: Fast silhouette and crease edge synthesis with geometry shaders. In: Spring Conference on Computer Graphics (SCCG). Smolenice, Slovakia, 2012.05.02-2012.05.04. pp. 78-83.

Abstract: We describe a technique to generate the geometry of silhouette and crease strokes in geometry shaders. This allows for single-pass, small-overhead rendering of conventional meshes. We exploit classic triangle adjacency information, but also introduce crease difference vectors associated with vertices in order to identify crease edges.

Gábor Jakab, Tamás Huszár, Balázs Csébfalvi: Iterative CT Reconstruction on the GPU. In: Hatodik Magyar Számítógépes Grafika és Geometria Konferencia. Budapest, Hungary, 2012.02.21-2012.02.22. pp. 124-131.

Abstract: The computing power of modern GPUs makes them very suitable for Computed Tomography (CT) image reconstruction. Apart from accelerating the reconstruction, their extra computing performance compared to conventional CPUs can be used to increase image quality in several ways. In this paper we present our upgraded GPU based iterative reconstruction algorithm, including ML-TR (Maximum Likelihood algorithm for Transmission tomography), MAP estimation using MRF Gibbs priors with Huber function. We experimentally evaluated the impact of wide

range of reconstruction parameters on the image noise level and the resolution, and we compared the performance of the reconstruction components in phantom studies with special focus on image noise, resolution and ring artefacts. Achieving more performance we extended our solution with multi-GPU support.

Márton Tóth, Dávid Dvorszki, Balázs Csébfalvi: GPU-Accelerated Segmentation of Medical Volume Data. In: Hatodik Magyar Számítógépes Grafika és Geometria Konferencia. Budapest, Hungary, 2012.02.21-2012.02.22. pp. 101-106.

Abstract: In this paper, a novel segmentation technique is introduced that is based on an abstract distance transform. The user can select a voxel as a seed point, from which the other voxels can be reached along different paths. For each voxel, we determine a path that is of minimal cost. Taking the density of the seed point as a reference, the cost of a path is calculated as an aggregate deviation of the densities corresponding to voxels visited along the given path. After having the cost of the cheapest path assigned to each voxel, a simple thresholding is used to obtain a segmentation mask. We demonstrate that this approach is competitive to the popular level set segmentation, but can be implemented more efficiently on recent Graphics Processing Units (GPU). Previously, using high-level shader languages for direct programming of the graphics pipeline, the implementation of different segmentation methods was difficult. However, the Open Computing Language (OpenCL) allows a more flexible general-purpose programming of recent GPUs. We show that, utilizing this new Application Programming Interface (API), our algorithm can be easily realized.

Márton Tóth, Dávid Dvorszki, Balázs Csébfalvi: Robust Volume Segmentation using an Abstract Distance Transform. In: 16th Central European Seminar on Computer Graphics (CESCG). Smolenice, Slovakia, 2012.04.29-2012.05.01. pp. 211-217.

Abstract: In this paper, a novel segmentation technique is introduced that is based on an abstract distance transform. The user can select a voxel as a seed point, from which the other voxels can be reached along different paths. For each voxel, we determine a path that is of minimal cost. Taking the density of the seed point as a reference, the cost of a path is calculated as an aggregate deviation of the densities corresponding to voxels visited along the given path. After having the cost of the cheapest path assigned to each voxel, a simple thresholding is used to obtain a segmentation mask. We demonstrate that this approach is competitive to the popular level set segmentation, but can be implemented more efficiently on recent Graphics Processing Units (GPU). Previously, using high-level shader languages for direct programming of the graphics pipeline, the implementation of different segmentation methods was difficult. However, new languages such as OpenCL or CUDA provide a more flexible environment for GPGPU (General Purpose computing on the GPU) programming. We show that, utilizing this new technology, our algorithm can be easily realized.

László Szirmay-Kalos, Milán Magdics, Balázs Tóth: Multiple Importance Sampling for PET. IEEE TRANSACTIONS ON MEDICAL IMAGING 33:(4) pp. 970-978. Paper TMI2300932. (2014)

Abstract: This paper proposes the application of multiple importance sampling in fully 3-D positron emission tomography to speed up the iterative reconstruction process. The proposed method combines the results of lines of responses (LOR) driven and

voxel driven projections keeping their advantages, like importance sampling, performance and parallel execution on graphics processing units. Voxel driven methods can focus on point like features while LOR driven approaches are efficient in reconstructing homogeneous regions. The theoretical basis of the combination is the application of the mixture of the samples generated by the individual importance sampling methods, emphasizing a particular method where it is better than others. The proposed algorithms are built into the Tera-tomo system.

Márta Zsolt, Szirmay-Kalos László: Partial Volume Effect Correction using Segmented CT Images with Distance Mapping. In: Szirmay-Kalos László, Renner Gábor (szerk.) VI. Magyar Számítógépes Grafika és Geometria Konferencia. Budapest, Hungary, 2012.02.20-2012.02.21. Budapest: Neumann János Számítógép-tudományi Társaság, pp. 106-113. ISBN: 978-963-313-048-3

Abstract: This paper proposes a Partial Volume Effect correction algorithm for PET/CTs, which improves the PET data based on the registered and segmented CT volume. The algorithm is based on mathematical morphology operations and identifies those regions where both the PET and the CT data have boundaries. The correction is restricted to these regions thus we can avoid the migration of artifacts and boundaries present in the CT but irrelevant for the positron density. Thus, the algorithm also maintains activity correctness.

Magdics Milán, Tóth Balázs: Stochastic Iteration in PET Reconstruction. In: Szirmay-Kalos László, Renner Gábor (szerk.) VI. Magyar Számítógépes Grafika és Geometria Konferencia. Budapest, Hungary, 2012.02.20-2012.02.21. Budapest: Neumann János Számítógép-tudományi Társaság, pp. 132-138. Paper 19. ISBN: 978-963-313-048-3

Abstract: This paper presents a modification of the classical Expectation Maximization scheme for fully 3D PET reconstruction. The proposed technique is based on the recognition that system matrix elements are high dimensional integrals that are approximated from discrete samples in each iteration step. If the sample locations are deterministic, then the approximation error will lead to a modified reconstruction result. However, when the sample positions are random, then the iteration may not converge at all. In our scheme, we use random samples, but average the results of forward projections in each iteration step. We show that this method is equivalent to using much higher number of samples at the cost of evaluating just a few of them in a single iteration step. With these advancements, over 128^3 resolution voxel arrays can be reconstructed in a few minutes.

Proceedings of Topics "Modeling and solving complex optimization problems "

OPTIMAL DEDICATED PROTECTION APPROACH TO SHARED RISK LINK GROUP FAILURES USING NETWORK CODING

Peter Babarczi∗, Janos Tapolcai∗, Pin-Han Ho†, Muriel Medard‡ ∗Dept. of Telecommunications and Media Informatics, Budapest University of Technology and Economics, Hungary †Dept. of Electrical and Computer Engineering, University of Waterloo, Canada ‡Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA Email: ∗{babarczi,tapolcai}@tmit.bme.hu,

†p4ho@uwaterloo.ca, ‡medard@mit.edu In Proc. IEEE International Conference on Communications (ICC), Ottawa, ON, Canada, 2012.

Abstract: Abstract—Survivable routing serves as a key role in connection-oriented communication networks for chieving desired service availability for each connection. This is particularly critical for the success of all-optical mesh networks where each lightpath carries a huge amount of data. Currently, 1 + 1 dedicated path protection appears to be the most widely deployed network resilience mechanism because it offers instantaneous recovery from network failures. However, 1 + 1 protection consumes almost twice as much capacity as required, which imposes a stringent constraint on network resource utilization. In addition, finding an SRLG-disjoint path is essential for 1 + 1 protection, which is nonetheless subject to non-trivial computation complexity and may fail in some SRLG scenarios. To address these problems, we introduce a novel framework of 1+1 protection, called Generalized Dedicated Protection (GDP), for achieving instantaneous recovery from any SRLG failure event. It is demonstrated, that finding a non-bifurcated optimal solution for GDP is NP-complete. Thus, the paper presents a novel scheme applying Generalized Dedicated Protection and Network Coding (GDP-NC) to ensure both optimal resource utilization among dedicated protection approaches and instantaneous recovery for single unicast flows, which can be split into multiple parts in all-optical networks. We demonstrate that the proposed GDP-NC survivable routing problem is polynomial-time solvable, owing to the ability to bifurcate flows. This flexibility comes at the expense of additional hardware for linear combination operations for the optical flows.

COST COMPARISON OF 1+1 PATH PROTECTION SCHEMES: A CASE FOR CODING

Harald Øverby, Gergely Biczók Dept. of Telematics Norwegian Univ. of Science and Technology Email:{haraldov,gbiczok}@item.ntnu.no Péter Babarczi and János Tapolcai Dept. of Telecommunications and Media Informatics Budapest Univ. of Technology and Economics Email: {babarczi,tapolcai}@tmit.bme.hu In Proc. IEEE International Conference on Communications (ICC), Ottawa, ON, Canada, 2012.

Abstract: Communication networks have to provide a high level of resilience in order to ensure sufficient Quality of Service for mission-critical services. Currently, dedicated 1+1 path protection is implemented in backbone networks to provide the necessary resilience. On the other hand, there are several possible realization strategies for 1+1 path protection functionality (1PPF), utilizing both diversity- and network coding. In this paper we consider the cost aspects of the different realization strategies. We evaluate the cost of providing 1PPF both analytically and empirically in realistic network topologies. Our results show that both diversity and network coding can provide 1PPF with reduced cost compared to traditional 1+1 path protection, even in case of short paths and strict coding restrictions. Specifically, the network coding scheme could be used as a cost-efficient and potentially all-optical realization of 1PPF.

IP FAST REROUTE WITH REMOTE LOOP-FREE ALTERNATES: THE UNIT LINK COST CASE

Levente Csikor, Gabor Retvari High Speed Networks Laboratory, Department of Telecommunications and Media Informatics Budapest University of Technology and Economics Email: {csikor, retvari}@tmit.bme.hu Springer Optimization Letters, 7, no. 4, 2013 pp 723-730.

Abstract: Up to not so long ago, Loop-Free Alternates (LFA) was the only viable option for providing fast protection in pure IP and MPLS/LDP networks. Unfortunately, LFA cannot provide protection for all possible failure cases in general. Recently, the IETF has initiated the Remote Loop-Free Alternates (rLFA) technique, as a simple extension to LFA, to boost the fraction of failure cases covered by fast protection. Before further standardization and deployment, however, it is crucial to determine to what extent rLFA can improve the level of protection in a general IP network, as well as to find optimization methods to tweak a network for 100% rLFA coverage. In this paper, we take the first steps towards this goal by solving these problems in the special, but practically relevant, case when each network link is of unit cost. We also provide preliminary numerical evaluations conducted on real IP network topologies, which suggest that rLFA significantly improves the level of protection, and most networks need only 2 − 3 new links to be added to attain 100% failure case coverage.*Index Terms*—IP Fast ReRoute, Remote Loop-Free Alternates,link protection, heuristics, unit link costs

CLUSTERING OF TICK DATA TO REDUCE STORAGE SPACE

Gabor I. Nagy* and Krisztian Buza** * Budapest University of Technology/Department of Telecommunications and Media Informatics, Budapest, Hungary ** Budapest University of Technology/Department of Computer Science and Information Theory, Budapest, Hungary nagyg@tmit.bme.hu, buza@cs.bme.hu In: INES 2012 - 16th International Conference on Intelligent Engineering Systems. Lisbon, Portugal, 2012.06.13-2012.06.15. (IEEE)pp. 6-12.(ISBN: 978-1-4673-2695-7) mycite: 135378

Abstract: Tick data is one of the most prominent types of temporal data, as it can be used to represent data in various domains such as geophysics or finance. Storage of tick data is a challenging problem because two criteria have to be fulfilled simultaneously: the storage structure should allow fast execution of queries and the data should not occupy too much space on the hard disk or in the main memory. In this paper, we present a clustering-based solution, and we introduce a new clustering algorithm, SOPAC, that is designed to support the storage of tick data. Our approach is based on the search for a partitional clustering that optimizes storage space. We evaluate our algorithm both on publicly available real-world datasets, as well as real-world tick data from the financial domain. We also investigate on task-specific benchmarks, how well our approach estimates the optimum. Our experiments show that, for the tick data storage problem, our algorithm substantially outperforms - both in terms of statistical significance and practical relevance - state-of-the-art clustering algorithms.

ROUTER VIRTUALIZATION FOR IMPROVING IP-LEVEL RESILIENCE

Janos Tapolcai, Gabor Retvari MTA-BME Future Internet Research Group, High-Speed Networks Laboratory (HSNLab),Dept. of Telecommunications and Media Informatics, Budapest University of Technology,Email: {tapolcai, retvari}@tmit.bme.hu In Proc. IEEE INFOCOM, Turin, Italy, 2013.

Abstract: IP-level failure protection based on the IP Fast ReRoute/Loop-Free Alternates (LFA) specification has become industrial requirement recently. The success of LFA lies in its inherent simplicity, but this comes at the expense of letting certain failure scenarios go unprotected. Realizing full failure coverage with LFA so

far has only been possible through completely reengineering the network around LFA-compliant design patterns. In this paper, we show that attaining high LFA coverage is possible without any alteration to the installed IP infrastructure, by introducing a carefully designed virtual overlay on top of the physical network that provides LFAs to otherwise unprotected routers. We study the problem of how to provision the overlay to maximize LFA coverage, we find that this problem is NPcomplete, and we give Integer Linear Programs to solve it. We also propose novel methods to work-around the limitations of current LFA implementations concerning Shared Risk Link Groups (SRLGs), which might be of independent interest. Our numerical evaluations suggest that router virtualization is an efficient tool for improving LFA-based resilience in real topologies. Index Terms—IP Fast ReRoute, Loop-Free Alternates, router virtualization

ON ACHIEVING ALL-OPTICAL FAILURE RESTORATION VIA MONITORING TRAILS

Janos Tapolcai_, Pin-Han Hoy, Peter Babarczi_, Lajos Ronyaiz _ MTA-BME Future Internet Research Group, High-Speed Networks Laboratory (HSNLab),Dept. of Telecommunications and Media Informatics, Budapest University of Technology, ftapolcai, babarczig@tmit.bme.hu y Dept. of Electrical and Computer Engineering, University of Waterloo, Canada, p4ho@uwaterloo.ca z Computer and Automation Research Institute Hungarian Academy of Sciences and BME, ronyai@sztaki.hu In Proc. IEEE INFOCOM Mini-Symposium, Turin, Italy, 2013.

Abstract: The paper investigates a novel monitoring trail (mtrail) scenario that can enable any shared protection scheme for achieving all-optical and ultra-fast failure restoration. Given a set of working (W-LPs) and protection (P-LPs) lightpaths, we firstly define the neighborhood of a node, which is a set of links whose failure states should be known to the node in restoration of the corresponding W-LPs. A set of m-trails is routed such that each node can localize any failure in its neighborhood according to the ON-OFF status of the traversing m-trails. Bound analysis is performed on the minimum bandwidth required for the m-trails. Extensive simulation is conducted to verify the proposed scheme

STATELESS MULTI-STAGE DISSEMINATION OF INFORMATION:SOURCE ROUTING REVISITED

Janos Tapolcai_, Andras Gulyasy_, Zalan Heszbergery_, Jozsef Biro_, Peter Babarczi_, Dirk Trossenz _MTA-BME Future Internet Research Group1, Budapest University of Technology and Economics (BME), Hungaryy Hungarian Academy of Science (MTA) Information system research group1, Hungary zComputer Laboratory2, Cambridge University, United Kingdom Email: _tapolcai, gulyas, heszberger, biro, babarczi@tmit.bme.hu, dirk.trossen@cl.cam.ac.uky In Proc. IEEE Globecom - Next Generation Networking and Internet Symposium, 2012.

Abstract: Large-scale information distribution has been increasingly attracting attention, be it through uptake in new services or through recent research efforts in fields like informationcentric networking. The core issue to be addressed is the more efficient distribution of information to a large set of receivers. Avoiding state in the forwarding elements is crucial for any scheme to be successful. This paper addresses this challenge by revisiting the idea of in-packet Bloom filters and source routing. As

opposed to the traditional in-packet Bloom filter concept which represent the trees flatly as sets, we build our filter by enclosing limited information about the structure of the tree, namely its stage decomposition, which helps to get rid of typical Bloom filter illnesses as infinite loops and false positive forwarding. Our analytical and simulation results show that by using this information we obtain more succinct tree representation while still maintaining forwarding efficiency.

FAST FAILURE LOCALIZATION IN ALL-OPTICAL NETWORKS WITH LENGTH-CONSTRAINED MONITORING TRAILS

Eva Hosszu_, J´anos Tapolcai_, Lajos Ronyaiy, Peter Soproni_, Peter Babarczi_, and Pin-Han Hoz _MTA-BME Future Internet Research Group1, Budapest University of Technology and Economics, Hungary y Computer and Automation Research Institute Hungarian Academy of Sciences and BME, Hungary z Dept. of Electrical and Computer Engineering, University of Waterloo, Canada Email: _fhosszu, tapolcai, soproni, babarczig@tmit.bme.hu, ronyai@sztaki.huy, p4ho@uwaterloo.caz In: 4th International Workshop on Reliable Networks Design and Modeling. Saint Petersburg, Russia, 2012.10.03-2012.10.05. pp. 1-7.

Abstract: Monitoring trails (m-trails) have been extensively studied as an alternative to the conventional link-based monitoring approach by using multi-hop supervisory lightpaths in all-optical networks. However, none of the previous studies have investigated the effect of length constraints upon the m-trail formation, which nonetheless correspond to the failure localization time. This paper addresses the above issue and formulates a new m-trail allocation problem, where the relationship between the number of m-trails versus the maximum hop count is explored. First, the paper investigates the theoretical bounds of allocating m-trails with at most k hops via an optimal group testing construction. Secondly, a novel meta-heuristic approach based on bacterial evolutionary algorithm for solving the lengthconstrained m-trail allocation problem is introduced. Through extensive simulations the performance gap of the proposed algorithm to the lower bound is presented on a wide diversity of topologies. Index Terms—monitoring trail; combinatorial group testing; bacterial evolutionary algorithm; length limit

COMPRESSING IP FORWARDING TABLES: TOWARDS ENTROPY BOUNDS AND BEYOND

Gábor Rétvári, János Tapolcai, Attila Korösi, András Majdán, Zalán Heszberger Department of Telecommunications and Media Informatics Budapest University of Technology and Economics {retvari,tapolcai,korosi,majdan,heszi}@tmit.bme.hu ACM Sigcomm 2013

Abstract: Lately, there has been an upsurge of interest in compressed data structures, aiming to pack ever larger quantities of informationinto constrained memory without sacrificing the efficiency of standard operations, like random access, search, or update. The main goal of this paper is to demonstrate how data compression can benefit the networking community, by showing how to squeeze the IP Forwarding Information Base (FIB), the giant table consulted by IP routers to make forwarding decisions, into information-theoretical entropy bounds, with essentially zero cost on longest prefix match and FIB update. First, we adopt the state-of-theart in compressed data structures, yielding a static entropycompressed FIB representation with asymptotically

optimal lookup. Then, we re-design the venerable prefix tree, used commonly for IP lookup for at least 20 years in IP routers, to also admit entropy bounds and support lookup in optimal time and update in nearly optimal time. Evaluations on a Linux kernel prototype indicate that our compressors encode a FIB comprising more than 440K prefixes to just about 100–400 KBytes of memory, with a threefold increase in lookup throughput and no penalty on FIB updates.

ON A PROBLEM OF R´ENYI AND KATONA

Eva Hosszu∗Department of Telecommunications and Media Informatics Budapest University of Technology and Economics hosszu@tmit.bme.hu Janos Tapolcai∗Department of Telecommunications and Media Informatics Budapest University of Technology and Economics tapolcai@tmit.bme.hu Gabor Wiener§ Department of Computer Science and Information Theory Budapest University of Technology and Economics wiener@cs.bme.hu
8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications June 4-7, 2013, Veszprém, Hungary.

Abstract: We are dealing with the classical problem of determining the minimum size of a separating system consisting of sets of size k. The problem was raised by R´enyi, the first and most important results are due to Katona; Wegener, Luzgin and Ahlswede also proved important bounds. We give a simple, short proof of a strengthening of Katona's main theorem determining the minimum size of a separating system of k-sets.

ROUTER VIRTUALIZATION FOR IMPROVING IP-LEVEL RESILIENCE

Janos Tapolcai, Gabor Retvari MTA-BME Future Internet Research Group, High-Speed Networks Laboratory (HSNLab),Dept. of Telecommunications and Media Informatics, Budapest University of Technology,Email: ftapolcai, retvarig@tmit.bme.hu In Proc. IEEE INFOCOM, Turin, Italy, 2013.

Abstract: IP-level failure protection based on the IP Fast ReRoute/Loop-Free Alternates (LFA) specification has become industrial requirement recently. The success of LFA lies in its inherent simplicity, but this comes at the expense of letting certain failure scenarios go unprotected. Realizing full failure coverage with LFA so far has only been possible through completely reengineering the network around LFA-compliant design patterns. In this paper, we show that attaining high LFA coverage is possible without any alteration to the installed IP infrastructure, by introducing a carefully designed virtual overlay on top of the physical network that provides LFAs to otherwise unprotected routers. We study the problem of how to provision the overlay to maximize LFA coverage, we find that this problem is NPcomplete, and we give Integer Linear Programs to solve it. We also propose novel methods to work-around the limitations of current LFA implementations concerning Shared Risk Link Groups (SRLGs), which might be of independent interest. Our numerical evaluations suggest that router virtualization is an efficient tool for improving LFA-based resilience in real topologies. Index Terms—IP Fast ReRoute, Loop-Free Alternates, router virtualization

SCALABLE FORWARDING FOR INFORMATION-CENTRIC NETWORKS

Weizhen Yang1 Imperial College London, UK Dirk Trossen Cambridge University Cambridge, UK János Tapolcai* Budapest University of Technology and Economics, Hungary In Proc. IEEE International Conference on Communications (ICC) - NGN, Budapest, Hungary, 2013

Abstract: Information-centric networking (ICN)1 is a new communication paradigm, which has been increasingly attracting attention in the wider research community. Its focus on information provides an alternative to the endpoint-centric model of today's Internet. While architectural foundations for ICN have been laid out in many ongoing efforts, solutions to forwarding information in such new networking environment still remain a challenge. Our criteria for a solution to this challenge are efficiency in terms of achievable link speed and scalability in terms of supported sizes of the network, while supporting multicast as a native operation. Our solution in this paper provides scalability through an extensible addressing format while keeping the forwarding operation efficient in terms of required state as well as forwarding performance.

SUFFICIENT CONDITIONS FOR PROTECTION ROUTING IN IP NETWORKS

János Tapolcai Springer Optimization Letters, 7, no. 4, 2013 pp 723-730.

Abstract: Providing fully distributed, fault tolerant, hop-by-hop routing is one of the key challenges for intra-domain IP networks. This can be achieved by storing two nexthops for each destination node in the forwarding table of the routers, and the packets are forwarded to primary next-hop (PNH), unless PNH is unreachable and secondary next-hop (SNH) is used instead. We follow the architecture by Kwong et al. in On the feasibility and efficacy of protection routing in IP networks, University of Pennsylvania (2010), where the routing tables are configured in a centralized way, while the forwarding and failure recovery is in a fully distributed way without relying on any encapsulation and signaling mechanisms for failure notification, to meet the standard IP forwarding paradigm. A network is protected if no single link of node failure results in forwarding loops. Kwong et al. (On the feasibility and efficacy of protection routing in IP networks, University of Pennsylvania 2010) conjectured that network node connectivity is not sufficient for a network to be protectable. In this paper we showthat this conjecture is in contradictionwith a conjuncture byHasunuma (Discrete Math 234(1–3):149–157, 2001; in Graph-Theoretic Concepts in Computer Science, Springer, Berlin, pp. 235–245, 2002), and show that every four connected maximal planar graph and every underlying graph of a 2-connected line digraph has feasible protection routing.

OPTIMIZING IGP LINK COSTS FOR IMPROVING IP-LEVEL RESILIENCE WITH LOOP-FREE ALTERNATES

Levente Csikor∗, Janos Tapolcai, Gabor Retvari HSNLab, Dept. of Telecommunications and Media Informatics, Budapest University of Technology and Economics COMPUTER COMMUNICATIONS 10.1016: pp. 1-11. (2012)

Abstract: The IP Fast ReRoute–Loop-Free Alternates (LFA) standard is a simple and easily deployable technique to provide fast failure protection right in the IP layer. To our days, most major IP device vendors have products on the market that support LFA out of the box. Unfortunately, LFA usually cannot protect all possible failure scenarios in a general network topology. Therefore, it is crucial to develop LFA-based network optimization tools in order to assist operators in deciding whether deploying LFA in

their network will supply sufficient resiliency. In this paper, we give a new graph theoretical framework for analyzing LFA failure case coverage, and then we investigate how to optimize the Interior Gateway Protocol (IGP) link costs in order to maximize the number of protected failure scenarios. We show that this problem is NP-complete even in a very restricted formulation, and we give an exact algorithm as well as a complete family of heuristics to solve it. Our simulation studies indicate that a deliberate tuning of the approximation strategy can significantly improve the quality of the IGP link costs, and we conclude that LFA cost optimization has the potential for boosting LFAbased resilience in most operational networks significantly.

COMPRESSING IP FORWARDING TABLES: TOWARDS ENTROPY BOUNDS AND BEYOND

Gábor Rétvári, János Tapolcai, Attila K˝orösi, András Majdán, Zalán Heszberger, Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics, {retvari,tapolcai,korosi,majdan,heszi}@tmit.bme.hu ACM Sigcomm 2013

Abstract: Lately, there has been an upsurge of interest in compressed data structures, aiming to pack ever larger quantities of information into constrained memory without sacrificing the efficiency of standard operations, like random access, search, or update. The main goal of this paper is to demonstrate how data compression can benefit the networking community, by showing how to squeeze the IP Forwarding Information Base (FIB), the giant table consulted by IP routers to make forwarding decisions, into information-theoretical entropy bounds, with essentially zero cost on longest prefix match and FIB update. First, we adopt the state-of-theart in compressed data structures, yielding a static entropycompressed FIB representation with asymptotically optimal lookup. Then, we re-design the venerable prefix tree, used commonly for IP lookup for at least 20 years in IP routers, to also admit entropy bounds and support lookup in optimal time and update in nearly optimal time. Evaluations on a Linux kernel prototype indicate that our compressors encode a FIB comprising more than 440K prefixes to just about 100–400 KBytes of memory, with a threefold increase in lookup throughput and no penalty on FIB updates.

COMMENTS ON ‘AVAILABILITY FORMULATIONS FOR SEGMENT PROTECTION’

P´eter Babarczi, Member, IEEE, J´anos Tapolcai, Member, IEEE, and Massimo Tornatore, Member, IEEE, IEEE Transactions on Communications 2013

Abstract: In this comment, we present some remarks on the availability evaluation of overlap dedicated segment protection (o-DSP) method in Tornatore et al., 2010, ,,Availability Formulations for Segment Protection” [1]. We show how to correctly apply the pivotal decomposition availability-evaluation method in directed graphs (which was claimed to inapplicable in [1]), such that it gives the same (exact) connection availability value as the generating function method proposed in the original paper.

Index Terms—Dedicated segment protection, availability evaluation, conditional decomposition, pivotal decomposition, factoring.

PHYSICAL IMPAIRMENTS OF MONITORING TRAILS IN ALL OPTICAL TRANSPARENT NETWORKS

Éva Hosszu, Ehsan Sarvghad Moghaddam, János Tapolcai, Dániel Mazroa MTA-BME Lendület Future Internet Research Group, Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics, Budapest, Hungary, E-mail: {hosszu,sarvghad,tapolcai,mazroa}@tmit.bme.hu IET Networks, 2013.

Abstract: In recent years, following the deployment of wavelength division multiplexing networks, fault detection and localisation has become a challenging issue in networks with high reliability. Optical layer monitoring schemes based on monitoring trails (m-trail) are considered an efficient way to localise a single fault unambiguously in all-optical networks. In spite of the extensive work on the m-trail concept, the issue has not been validated from the feasibility point of view. Previous works on the m-trail monitoring scheme have focused mainly on algorithm design for minimising the number of monitors, however, none of them have observed that length limitations should be considered as well. The authors investigate the physical constraints of launching m-trails, mainly focusing on the maximum length that each m-trail may have and describe an algorithm that solves the length constrained m-trail formulation problem. Numerous simulations were implemented in a physical layer simulator for observing qualitative parameters in different m-trail lengths. The authors propose 15 000 km as a length limit for each m-trail for out-of-band monitoring. Finally, an algorithmic solution is provided for monitoring trail design problem satisfying these physical constraints.

Proceedings of Topics "Distributed evaluation and topological analysis"

Zoltán Czirkos, Márta Rencz, & Gábor Hosszú: Improving Attack Aggregation Methods Using Distributed Hash Tables, International Conference on Internet Monitoring and Protection - ICIMP 2012. Stuttgart, Germany, 2012.05.27-2012.06.01, Paper 30046

Abstract: Collaborative intrusion detection has several difficult subtasks to handle. Large amount of data generated by intrusion detection probes has to be handled to spot intrusions. Also, when correlating the pieces of evidence, the connection between them has to be revealed as well, as it may be the case that they are part of a complex, large-scale attack. In this article, we present a peer-to-peer network based intrusion detection system, which is able to handle the intrusion detection data efficiently while maintaining the accuracy of centralized approaches of correlation. The system is built on a distributed hash table, for which keys are assigned to each piece of intrusion data in a preprocessing step. This method allows one to make well-known correlation mechanisms work in a distributed environment.

Zoltán Czirkos, György Bognár, & Gábor Hosszú: Pseudo Reliable Broadcast in the Kademlia P2P System. Third International Conference on Advances in Computer Engineering – ACE 2012. Amsterdam, Holland, 2012.06.07-2012.06.08, Paper 6

Abstract: Structured peer-to-peer (P2P) networks are capable of fast and efficient look-up operations as a distributed hash table (DHT); however, they generally do not support complex queries of any kind. To perform such, controlled broadcast messages are used. By using the structure inherent to these networks, one can implement broadcast messages efficiently. In this article a broadcast algorithm for the Kademlia topology is presented. The algorithm uses a replication mechanism similar to that of the storage and retrieval mechanism of Kademlia. This allows for increased reliability and speed of the broadcast as well. A model is presented, which allows one to calculate the required level of replication for any reliability desired at runtime, and is validated with simulation as well.

Raymond Eliza Ivan Pardede, Loránd Lehel Tóth, Gábor Hosszú, Ferenc Kovács, "Glyph Identification Based on Topological Analysis", Scientific Workshop organized by the PhD school on Computer Science in the framework of the project TÁMOP-4.2.2/B-10/1-2010-0009. March 9, 2012, pp. 99-103.

Abstract: This paper proposes a novel mathematical model that describes the logical relationship among glyphs belong to the same script. The proposed model is presented as three logical layers namely Topology, Visual Identity, and Phonetic layers. In the Topology Layer, a unique glyph that represents a grapheme is defined by a set of geometrical properties. Furthermore in the Topology Layer, the relation between two different glyphs is defined by the number of topological transformation steps required to transform the shape of one glyph into another. In the Visual Identity Layer, the glyphs of a single grapheme share some topological attributes in common. Some graphemes of a script from different age may have similar Common Identity. For that particular case, to be able to distinguish them, the evaluation must be extended by means of phonetic attribute of the graphemes. The article presents a potential implementation of the proposed three-layer hierarchical grapheme model.

Gábor Hosszú: Heritage of Scribes. Second Edition. Budapest, 2012. ISBN 978-963-88437-4-6, 310 pages

Abstract: The Heritage of Scribes introduces the history and development of three members of the Rovas (pronounced as r-o-w-a-sh) script family: the Carpathian Basin Rovas, the Khazarian Rovas, and the Szekely-Hungarian Rovas. The historical and linguistic statements in the book are based on the published theories and statements of acknowledged scholars, historians, archaeologists, and linguists. The author provides detailed descriptions of the three Rovas scripts, presents their relationships, connections to other scripts, and explains the most significant Rovas relics. Based on the discovered relations, the author introduces the systematic description of the Rovas glyphs in the Rovas Atlas together with a comprehensive genealogy of each character as well.

Gergely Endre Tóth, Gábor Fodor, Raymond Pardede, Loránd Lehel Tóth, György András Jeney, and Gábor Hosszú: Survey of the Clustering Methods. Informatika. Vol. XIV. No. 2. Jun. 2012. pp. 45-48.

Abstract: This article presents the fundamental parts and classes of the cluster analysis, and shows their advantages and disadvantages along with the applied methods. The article also elaborates each cluster analysis method and explains how to choose between the available methods. Moreover, the technique to analyse the result is also presented. The last part of this article covers how the cluster analysis methods can be applied for data mining and knowledge discovery.

Zoltán Czirkos, Márta Rencz & Gábor Hosszú: A distributed hash table assisted intrusion detection system. IARIA International Journal on Advances in Security, Vol. 5., N. 3&4, 2012, pp. 134-143.

Abstract: Using collaborative intrusion detection to sense network intrusions comes at a price of handling an enormous amount of data generated by detection probes, and the problem of properly correlating the evidence collected at different parts of the network. The correlation between the recorded events has to be revealed, as it may be the case that they are part of a complex, large-scale attack, even if they manifested at different parts of the network. In this paper we describe the inner workings a peer-to-peer network based intrusion detection system, which is able to handle the intrusion detection data efficiently while maintaining the accuracy of centralized approaches of correlation. The system is built on a distributed hash table, for which keys are assigned to each piece of intrusion data in a preprocessing step. The network traffic requirements of such a system, and the load balancing that can be achieved by using the Kademlia peer-to-peer overlay network are discussed as well.

Zoltán Czirkos & Gábor Hosszú: Solution for the Broadcasting in the Kademlia Peer-to-Peer Overlay. Computer Networks – The International Journal of Computer and Telecommunications Networking. 57 (2013) pp. 1853-1862, Elsevier

Abstract: Structured peer-to-peer networks are capable of fast and efficient lookup operations as a distributed hash table. The topology of these networks makes it possible to send broadcast messages among nodes, either for the purpose of providing a complex query service for participants, or to disseminate information valuable for all nodes. In this article a broadcast algorithm for the Kademlia XOR topology is presented. The algorithm, which was developed specifically for Kademlia, uses replication mechanisms similar to that of the storage and retrieval service of this overlay topology. This allows for increased reliability and speed of the broadcast, and also efficient operation, as the routing table for lookups are already available and can be used. An analytical model is presented, which can be used to calculate the required level of replication for any desired reliability at runtime, and is validated with simulation as well.

Gábor Fodor, Ádám T. Balogh, Gábor Hosszú, Ferenc Kovács: Screening of Congenital Heart Diseases by Murmurs Using Telemedical Phonocardiography. In: EMBC12, 34th Annual International IEEE EMBS Conference, San Diego, USA, August 28 – September 1, 2012, pp. 6100-6103.

Abstract: An improved fetal first heart sound model is presented in this paper. Preliminary studies showed that an extended model can describe the waveform of the first heart sound significantly better than the former one, thus it is more suitable for cluster analysis. Using the new model around 40% decrease in the normalized root mean square error (NRMSE) was achieved during the fitting procedure. The NRMSE between the original and the reconstructed first heart sounds varied around 9%. The details of the comparison are presented. 100 fetal heart sound records were randomly selected out of 2400 measurements. Each one is a standard 20-minute CTG test carried out after the 28^{th} week of gestation and recorded on individual patients. Cluster analysis was performed based on the model parameters. The cluster analysis partitioned the datasets into more than one cluster at 24% of the records. One record is selected for presentation purposes where two significantly different clusters were found, which may refer to fetal movements or changes in the cardiovascular system. The reason is still unknown and requires further exploration. The results show that the presented way of investigation can be a promising method for long term measurements and information about the fetal well-being and balance of the autonomous nervous system might be revealed.

Zoltán Czirkos, Gábor Hosszú: Enhancing Collaborative Intrusion Detection Methods Using a Kademlia Overlay Network, 18th EUNICE Conference on Information and Communications Technologies, 29-31 Augustus 2012, Budapest, Hungary, co-sponsored by IFIP WG 6.2 and IFIP WG 6.6 with Springer IFIP-LNCS Proceedings, pp. 52-63

Abstract: The two important problems of collaborative intrusion detection are aggregation and correlation of intrusion events. The enormous amount of data generated by detection probes requires significant network and computational capacity to be processed. In this article we show that a distributed hash table based approach can reduce both network and computational load of intrusion detection, while providing almost the same accuracy of detection as centralized solutions. The efficiency of data storage can be improved by selecting Kademlia as the underlying overlay network topology, as its routing can easily adapt to the dynamic properties of such an application.

Hosszú, Gábor: “Mathematical Statistical Examinations on Script Relics” chapter in book, Data Mining and Analysis in Engineering Field, Editor: Vishal Bhatnagar, Information Science Reference, Hershey, New York, USA

Abstract: The chapter presents statistical evaluations of script relics. Its concept is exploiting mathematical statistical methods to extract hidden correlations among different script relics. Examining the genealogy of the graphemes of scripts is necessary for exploring the evolution of the writing systems, reading undeciphered inscriptions, and deciphering undeciphered scripts. The chapter focuses on the cluster analysis as one of the most popular mathematical statistical method. The chapter presents the application of the clustering in the classification of Rovash (pronounced “rove-ash”, an alternative spelling: Rovas) relics. The various Rovash scripts were used by nations in the Eurasian Steppe and in the Carpathian Basin. The specialty of the rovash paleography that the Rovash script family shows a vital evolution during the last centuries; therefore, it is ideal subject to test the models of the evolution of the glyphs. The most important Rovash script is the Szekely-Hungarian Rovash. Cluster analysis algorithms were applied for determining the common sets among the significant Szekely-Hungarian Rovash alphabets. The determined rovash relic ties prove the usefulness of the clustering methods in the rovash paleography.

Zoltán Czirkos, György Bognár, Gábor Hosszú: Pseudo Reliable Broadcast in the Kademlia P2P System. In: ACEEE International Journal on Communication IJCom, ISSN 2158-7558 (Online) ISSN 2158-754X (Print), published by ACEEE USA

Abstract: Kademlia is a structured peer-to-peer (P2P) application level network, which implements a distributed hash table (DHT). Its key-value storage and lookup service is made efficient and reliable by its well-designed binary tree topology and dense mesh of connections between participant nodes. While it can carry out data storage and retrieval in logarithmic time if the key assigned to the value in question is precisely known, no complex queries of any kind are supported. In this article a broadcast algorithm for the Kademlia network is presented, which can be used to implement such queries. The replication scheme utilized is compatible with the lookup algorithm of Kademlia, and it uses the same routing tables. The reliability (coverage) of the algorithm is increased by assigning the responsibility of disseminating the broadcast message to many nodes at the same time. The article presents a model validated with simulation as well. The model can be used by nodes at runtime to calculate the required level of replication for any desired level of coverage. This calculation can take node churn, packet loss ratio and the size of the overlay into account.

Gábor Fodor, Ferenc Kovács, Gábor Hosszú: Screening of congenital heart diseases by fetal phonocardiography, Journal of Medical Engineering & Technology. Informa healthcare, London.

Abstract: A large proportion of congenital heart diseases (CHD) remain undetected during pregnancy or even after birth. Many of them generate turbulent blood flow, resulting heart murmur. Doppler ultrasound cardiotocography (CTG) is suitable for the assessment of the fetal heart rate and some derived parameters, but it is inadequate for detecting heart murmurs. Although comprehensive examination can be carried out with echocardiography, it is expensive and requires expertise; therefore, it is not applicable for widespread screening. This paper presents a new possibility for screening for some CHDs using phonocardiography, which can be combined with Doppler ultrasound CTG as an extension of it. Furthermore it can be carried out at home allowing repeated measurements, which increases also the reliability of filtering out innocent murmurs. The diagnostic capability of this screening method is supported by a large number of evaluated fetal heart sound records. Moreover, according to experiences pregnant women prefer this reliable, easy to use method, which facilitates their examination.

Zoltán Czirkos & Gábor Hosszú: Komondor: Dynamic Properties of Network Traffic of the Komondor Intrusion Detection System. ENIGMA Brazilian Journal of Information Security & Cryptography

Abstract: In this article we present a distributed network intrusion detection method and consider its network traffic. Network hosts using this method create a distributed hash table, which is an overlay network of nodes with equal responsibility to share data of detected intrusion attempts. Information collected by the nodes does not remain scattered at the probes; rather it is collected and used to create a collective knowledge base. Being able to process data in a distributed manner, nodes are able to enhance security of each other. The network traffic costs of such application of an overlay network is considered, and a broadcast algorithm is presented, which uses replication to enhance its reliability even in the presence of network errors or malicious hosts.

Gábor Hosszú: “A novel computerized paleographical method for determining the evolution of graphemes” chapter in book, Encyclopedia of Information Science and Technology, Third edition, Editor: Mehdi Khosrow-Pour, Information Science Reference, Hershey, New York, USA

Abstract: Examining the genealogy of the graphemes is useful for exploring the evolution of the writing systems, reading undeciphered inscriptions, and deciphering undeciphered scripts. Researching the evolution of various scripts needs sophisticated and automatized procedures for exploring the fine links among the different writing systems. The traditional method compares the whole grapheme repertoire of the scripts. However, it is inefficient if the graphemes in the writing system were borrowed from different scripts. The proposed, novel procedure is based on the topological comparison of each glyph of the graphemes. Its algorithm is based on the global minimization of the topological differences in each descent branch of the graphemes.

Péter Horváth, Gábor Hosszú, Ferenc Kovács: A Proposed Method for Optimized ASIP Synthesis. *Microelectronics Journal*. Elsevier.*benyújtás előtt.*

Abstract: Due to the rapid technology advancement in integrated circuit era, the need for the high computation performance together with the increasing complexity and manufacturing costs raised the demand for high-performance reconfigurable designs; therefore, the application-specific instruction set processors (ASIPs) are widely used in SoC design. The automated generation of software tools for ASIPs is a commonly used technique, but the automated hardware model generation is less frequently applied. Oppositely, the final register-transfer level implementations are usually created, wholly or partly, manually. This paper presents a novel approach for automated hardware model generation for ASIPs. The new solution is based on a novel abstract ASIP model and a modeling language (Algorithmic Microarchitecture Description Language, AMDL) optimized for this architecture model. The proposed AMDL-based pre-synthesis method is based on a set of pre-defined VHDL implementation schemes, which ensure the qualities of the automatically generated register-transfer level models in terms of resource requirement and operation frequency. The design framework implementing the algorithms required by the synthesis method is also presented.

Proceedings of Topics “Graph theory and geometrical problems”

Áron Lászka, Levente Buttyán, Dávid Szeszlér: Designing robust network topologies for wireless sensor networks in adversarial environments. PERVASIVE AND MOBILE COMPUTING 9:(4) pp. 546-563. (2013)

Abstract: In this paper, we address the problem of deploying sink nodes in a wireless sensor network such that the resulting network topology be robust. In order to measure network robustness, we propose a new metric, called persistence, which better captures the notion of robustness than the widely known connectivity based metrics. We study two variants of the sink deployment problem: sink selection and sink

placement. We prove that both problems are NP-hard, and show how the problem of sink placement can be traced back to the problem of sink selection using an optimal search space reduction technique, which may be of independent interest. To solve the problem of sink selection, we propose efficient heuristic algorithms. Finally, we provide experimental results on the performance of our proposed algorithms.

Gábor Wiener: Rounds in combinatorial search, Algorithmica 67, pp. 315-323. (2013)

Abstract A set system H ⊆ 2[m] is said to be separating if for every pair of distinct elements x, y ∈ [m] there exists a set H ∈ H such that H contains exactly one of them. The search complexity of a separating system H ⊆ 2[m] is the minimum number of questions of type "x ∈ H? " (where H ∈ H) needed in the worst case to determine a hidden element x ∈ [m]. If we receive the answer before asking a new question then we speak of the adaptive complexity, denoted by c(H); if the questions are all fixed beforehand then we speak of the nonadaptive complexity, denoted by cna(H). If we are allowed to ask the questions in at most k rounds then we speak of the k-round complexity of H, denoted by ck(H). It is clear that |H| ≥ cna(H) = c1(H) ≥ c2(H) ≥ . . . ≥ cm(H) = c(H). A group of problems raised by G. O. H. Katona is to characterize those separating systems for which some of these inequalities are tight. In this paper we are discussing set systems H with the property |H| = ck(H) for any k ≥ 3. We give a necessary condition for this property by proving a theorem about traces of hypergraphs which also has its own interest.

Klaus Jansen, Stefan Kratsch, Dániel Marx, Ildikó Schlotter: Bin packing with fixed number of bins revisited, Journal of Computer and System Sciences, Volume 79, Issue 1, February 2013, pp. 39-49

Abstract: As Bin Packing is NP-hard already for k = 2 bins, it is unlikely to be solvable in polynomial time even if the number of bins is a fixed constant. However, if the sizes of the items are polynomially bounded integers, then the problem can be solved in time nO(k) for an input of length n by dynamic programming. We show, by proving the W[1]-hardness of Unary Bin Packing (where the sizes are given in unary encoding), that this running time cannot be improved to f (k) · nO(1) for any function f (k) (under standard complexity assumptions). On the other hand, we provide an algorithm for Bin Packing that obtains in time 2O(k log2 k) + O(n) a solution with additive error at most 1, i.e., either finds a packing into k +1 bins or decides that k bins do not suffice.

Zoltán Ádám Mann, Pál András Papp. Predicting algorithmic complexity through structure analysis and compression. APPLIED SOFT COMPUTING 13:(8) pp. 3582-3596. (2013)

Abstract: The complexity of an algorithm is usually specied by the maximum number of steps made by the algorithm, as a function of the size of the input. However, as dierent inputs of equal size can yield dramatically dierent algorithm runtime, the size of the input is not always an appropriate basis for predicting algorithm runtime. In this paper, we argue that the compressed size of the input is more appropriate for this purpose. In particular, we devise a genetic algorithm for compressing a graph by nding the most compact description of its structure, and we demonstrate how the compressed size of the problem instance correlates with the runtime of an exact

algorithm for two hard combinatorial problems (graph coloring and Boolean satisability).

Zoltán Ádám Mann, Anikó Szajkó. Average-case complexity of backtrack search for coloring sparse random graphs. JOURNAL OF COMPUTER AND SYSTEM SCIENCES 79:(8) pp. 1287-1301. (2013)

Abstract: We investigate asymptotically the expected number of steps taken by backtrack search for k-coloring random graphs Gn;p(n) or proving non-k-colorability, where p(n) is an arbitrary sequence tending to 0, and k is constant. Contrary to the case of constant p, where the expected runtime is known to be O(1), we prove that here the expected runtime tends to innity. We establish how the asymptotic behaviour of the expected number of steps depends on the sequence p(n). In particular, for p(n) = d=n, where d is a constant, the runtime is always exponential, but it can be also polynomial if p(n) decreases suciently slowly, e.g. for p(n) = 1= ln n.

Zoltán Ádám Mann, Tamás Szép: Accelerating backtrack search through a best-first-search strategy. Submitted to International Journal of Applied Mathematics and Computer Science.

Abstract: Backtrack-style exhaustive search algorithms for NP-hard problems tend to have large variance in their runtime. This is because fortunate branching decisions can lead to nding a solution quickly, whereas unfortunate decisions in another run can lead the algorithm to a region of the search space with no solutions. In the literature, frequent restarting has been suggested as a means to overcome this problem. In this paper, we propose a more sophisticated approach: a best-rstsearch heuristic to quickly move between parts of the search space, always concentrating on the most promising region. We describe how this idea can be e-ciently incorporated into a state-of-the-art backtrack search algorithm, without sacricing optimality. Moreover, we demonstrate with substantial empirical results that, for hard solvable problem instances, the new approach provides signicantly higher speed-up than frequent restarting.

Áron Lászka, Dávid Szeszlér, Levente Buttyán: Game-theoretic Robustness of Many-to-one Networks. In: Game Theory for Networks. Vancouver, Canada, 2012.05.24-2012.05.26. Dordrecht: Springer, pp. 88-98. (Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering; 105.)

Abstract. In this paper, we study the robustness of networks that are characterized by many-to-one communications (e.g., access networks and sensor networks) in a game-theoretic model. More specically, we model the interactions between a network operator and an adversary as a two player zero-sum game, where the network operator chooses a spanning tree in the network, the adversary chooses an edge to be removed from the network, and the adversary's payo is proportional to the number of nodes that can no longer reach a designated node through the spanning tree. We show that the payo in every Nash equilibrium of the game is equal to the reciprocal of the persistence of the network. We describe optimal adversarial and operator strategies and give ecient, polynomial-time algorithms to compute optimal strategies. We also generalize our game model to include varying node weights, as well as attacks against nodes.

Áron Lászka, Dávid Szeszlér, Levente Buttyán: Linear Loss Function for the Network Blocking Game: An Efficient Model for Measuring Network Robustness and Link Criticality. In: Decision and Game Theory for Security. Budapest, Hungary, 2012.11.05-2012.11.06. Berlin: Springer, pp. 152-170.

Abstract. In order to design robust networks, rst, one has to be able to measure robustness of network topologies. In [1], a game-theoretic model, the network blocking game, was proposed for this purpose, where a network operator and an attacker interact in a zero-sum game played on a network topology, and the value of the equilibrium payo in this game is interpreted as a measure of robustness of that topology. The payo for a given pair of pure strategies is based on a loss-in-value function. Besides measuring the robustness of network topologies, the model can be also used to identify critical edges that are likely to be attacked. Unfortunately, previously proposed loss-in-value functions are either too simplistic or lead to a game whose equilibrium is not known to be computable in polynomial time. In this paper, we propose a new, linear loss-in-value function, which is meaningful and leads to a game whose equilibrium is eciently computable. Furthermore, we show that the resulting game-theoretic robustness metric is related to the Cheeger constant of the topology graph, which is a well-known metric in graph theory.

Gábor Wiener: On a problem of Rényi and Katona. In: Search Methodologies III., Bielefeld, Germany, 2012.09.03-2012.09.07. p. 73.

Abstract: We are dealing with the classical problem of determining the minimum size of a separating system consisting of sets of size at most k . The problem was raised by R´enyi, the first and most important results are due to Katona; Wegener and Ahlswede also proved important bounds. We give a simple, short proof of a strengthening of Katona’s main theorem determining the minimum size of a separating system of sets of size at most k .

Attila Sali: Minimum representations of closure operations. Workshop 'Algebra Across the Borders' II. 2012.06.17-2012.06.20, Bolyai Institute, Szeged, Hungary

Abstract: In the present paper a distance concept of closures originated in databases is investigated. Two database instances are of distance 0, if they have the same number of attributes and satisfy exactly the same set of functional dependencies. This naturally leads to the poset of closures as a model of changing database. The distance of two databases (closures) is defined to be the distance of the two closures in the Hasse diagram of that poset. We determine the diameter of the poset and show that the distance of two closures is equal to the natural lower bound, that is to the size of the symmetric difference of the collections of closed sets. We also investigate the diameter of the set of databases/closures with a given system of keys, where keys are minimal sets whose closure is the whole underlying set. Sharp upper bounds are given in the case when the minimal keys are 2 (or r)-element sets.

Attila Sali: Note on the size of binary Armstrong codes. 2012 Shanghai Conference on Algebraic Combinatorics, 2012.08.17-2012.08.22, Shanghai Jiao Tong University, p. 27.

Abstract: An Armstrong code Arm(q; k; n) is a code of length n over an alphabet of size q with minimum Hamming distance d = n .. k + 1 and the additional property that for every subset of size k .. 1 = n .. d of the coordinate positions there are two codewords that agree there (so the minimum distance occurs `in all directions'). For example, the code consisting of the rows of an n by n identity matrix is an Arm(q; n.. 1; n) and the code of the n +1 vectors ci = (1; : : : ; 1; 0; : : : ; 0) with i ones followed by n .. i zeroes is an Arm(q; n; n) for all q . In this note we take q = 2, and give necessary and sucient conditions for the existence of an Arm(2; k; n). We show for binary Armstrong codes Arm(2; k; n) that asymptotically n=k 1: 224, while such a code is shown to exist whenever n=k 1: 12. We also construct an Arm(2; n .. 2; n) and Arm(2; n .. 3; n) for all admissible n .

Kristóf Marussy, Krisztián Buza: SUCCESS: A New Approach for Semi-supervised Classification of Time-Series. In: Artificial Intelligence and Soft Computing - 12th International Conference, ICAISC 2013, Zakopane, Lengyelország, 2013.06.09-2013.06.13. Proceedings, Part I (LNAI 7894), pp. 437-447.

Abstract. The growing interest in time-series classication can be attributed to the intensively increasing amount of temporal data collected by widespread sensors. Often, human experts may only review a small portion of all the available data. Therefore, the available labeled data may not be representative enough and semi-supervised techniques may be necessary. In order to construct accurate classiers, semi-supervised techniques learn both from labeled and unlabeled data. In this paper, we introduce a novel semi-supervised time-series classier based on constrained hierarchical clustering and dynamic time warping. We discuss our approach in the framework of graph theory and evaluate it on 44 publicly available real-world time-series datasets from various domains. Our results show that our approach substantially outperforms the state-ofthe-art semi-supervised time-series classier. The results are also justied by statistical signicance tests.

P. Biro, T. Fleiner, R. Irving: Matching Couples with Scarf's Algorithm, 8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications, 2013.06.04.-2013.06.07, Veszprém, Hungary, pp. 55-64.

Abstract: Scarf’s algorithm [18] provides fractional core elements for NTU-games, which are equivalent to fractional stable matchings for stable matching problems on hypergraphs. Bir´o and Fleiner [3] showed that the Scarf algorithm can be extended for capacitated NTU-games. In this setting agents can be involved in more than one coalition at a time, cooperations may be performed with different intensities up to some limits, and the contribution of the agents can also differ in a coalition. The fractional stable solutions for the above model, produced by the extended Scarf algorithm, are called stable allocations. In this paper we apply this solution concept for the Hospitals Residents problem with Couples (HRC). This is one of the most important general stable matching problems due to its relevant applications, also wellknown to be NP-hard. We show that if a stable allocation yielded by the Scarf algorithm turns out to be integral then it provides a stable matching for an instance of HRC, so this method can be used as a heuristic. In an experimental study, we compare this method with other heuristics constructed for HRC that are applied in practice in the American and Scottish resident allocation programs, respectively. Our main

finding is that the Scarf algorithm outperforms all the other known heuristics when the proportion of couples is high.

L. Buttyán, Á. Lászka, D. Szeszlér: A Minimum Cost Source Location Problem for Wireless Sensor Networks, 8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications, 2013.06.04.-2013.06.07, Veszprém, Hungary, pp. 79-88.

Abstract: Wireless sensor networks consist of physically unprotected devices, therefore it is essential to design the network configuration in such a way that it is the most resistant against attacks. However, it is not at all obvious how to measure the robustness of a network to best serve the purposes of wireless sensor networks. In this paper we propose a metric that seems to be appropriate and then we address the problem of assigning the sink role to a subset of vertices such that the arising network is as robust with respect to the proposed metric as possible.

K. Buza, I. Galambos: An Application of Link Prediction in Bipartite Graphs: Personalized Blog Feedback Prediction,8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications, 2013.06.04.-2013.06.07, Veszprém, Hungary, pp. 89-96.

Abstract: The last decade lead to an unbelievable growth of the importance of social media. One of the most interesting challenges associated with social media is predicting which user is expected to comment which blog. In this paper, we formulate the above task as a link prediction problem in bipartite graphs. Recently, sparse matrix factorization became popular for link prediction. We show that the conventional algorithm for the factorization of matrices has great potential to predict new links between blogs and users. However, it fails to capture the true distribution of comments and therefore a straight-forward application of conventional matrix factorization leads to suboptimal predictions that do not even outperform the simple baseline of random predictions in terms of overall precision and recall. In order to alleviate this deficiency, we devise a simple technique which improves the efficiency of the algorithm substantially.

K. Buza, K. Marussy: Hubness-based indicators for semi-supervised time-series classification, 8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications, 2013.06.04.-2013.06.07, Veszprém, Hungary, pp. 97-108.

Abstract: Due to the decreasing cost and wide availability of sensors that measure the change of a quantity over time, machine learning methods focusing on time-series have gained increasing attention during the last few decades. In domains where time-series databases are only emerging, adequate amounts of data annotated by human experts might not be available due to the associated high expenses and required effort. Semi-supervised learning techniques are able to alleviate some of these problems. We exploit the relation between hubness— a phenomenon emergent in time-series data sets due to their high dimensionality—and the effectiveness of semi-supervised learning. We evaluate various hubness-based indicator values on 44 publicly available data sets whether they can predict the effectiveness of semi-supervised learning. We also investigate the use of hubness-based indicators in choosing between two semi-supervised learners for time series. Our results show that selection of the appropriate

learning method is often possible based on indicators available a priori, without labelling the whole data set.

B. Csizmadia, K. Friedl: All Pairs Small Stretch Paths in Weighted Graphs, 8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications, 2013.06.04.-2013.06.07, Veszprém, Hungary, pp. 115-120.

Abstract: In an undirected graph with nonnegative weights we consider the problem of approximating the shortest paths with a multiplicative error 2. For dense graphs we improve the algorithm of Cohen and Zwick and obtain running time O(n52). In order to achieve this, based on the idea of Pandey, Kumar, and Singh, we replace the small dominating set the algorithm of Cohen and Zwick uses by a small set obtained by a greedy process, making the algorithm conceptually simpler.

T. Fleiner, Zs. Jankó, A. Tamura: Stability of generalized network flows, 8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications, 2013.06.04.-2013.06.07, Veszprém, Hungary, pp. 133-142.

Abstract: In this work, we extend the stable marriage theorem of Gale and Shapley by showing a common generalization of Ostrovsky's result on supply chain networks and Fleiner's on stable network flows. Our mathematical model, based on a directed graph and certain choice functions on the stars is much more flexible then previous ones and for this reason we believe that it is also interesting for Economists studying stability in supply chains or two-sided markets. Our main result is proved by applying Tarski's fixed point theorem.

É. Hosszu, J. Tapolcai, G. Wiener: On a problem of Rényi and Katona, 8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications, 2013.06.04.-2013.06.07, Veszprém, Hungary, pp. 229-232.

Abstract: We are dealing with the classical problem of determining the minimum size of a separating system consisting of sets of size k. The problem was raised by R´enyi, the first and most important results are due to Katona; Wegener, Luzgin and Ahlswede also proved important bounds. We give a simple, short proof of a strengthening of Katona's main theorem determining the minimum size of a separating system of k-sets.

A. Recski: Matroid Duality and Voltage-Current Symmetries, 8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications, 2013.06.04.-2013.06.07, Veszprém, Hungary, pp. 419-426-

Abstract: If the underlying graph of a linear electric network is planar, one can dualize the network by replacing the graph with its dual and the devices by their "duals" but what does this latter concept mean? Masao Iri and the present author observed some three decades ago that the two usual answers ("interchange the role of voltage and current" or "take the matroid dual") are not equivalent, leading to two distinct voltage-current symmetries in electric network theory. Using these results two sets of duality

principles are presented, referring to two basic qualitative problems, to the unique solvability of the linear networks and to the interconnectability of linear multiports.

Á. Tóth: The asymptotic value of the independence ratio for the categorical graph power, 8th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications, 2013.06.04.-2013.06.07, Veszprém, Hungary, pp. 489-494.

Abstract: The independence ratio, i(G) of a graph G is the ratio of the independence number and the number of vertices. Its asymptotic value for the categorical graph power is defined as $A(G) = \lim_{k\to\infty} i(G\times k)$, where $G\times k$ denotes the kth categorical power of G. In this paper we give a simple formula for this graph parameter, thereby we answer two questions of Alon and Lubetzky. We also discuss some other open problems related to A(G) which are immediately settled by this result.

Proceedings of Topics "Modelling of distributed software systems (Grid, SOA, Cloud)"

Richárd Kápolnai, Imre Szeberényi, Balázs Goldschmidt: Approximation of repeated scheduling chains of independent jobs of unknown length based on historical data. In: Recent Advances in Computer Science (17th WSEAS International Conference on Computers) Rodosz, Grece, 2013.07.16--2013.07.19, pp. 41--46.

Abstract: We consider the problem of minimizing the makespan when scheduling a Parameter Sweep Application (PSA, set of independent jobs) on identical machines of a parallel computational infrastructure. However, there is no a priori information on the job lengths, and the user intends to execute the whole PSA multiple times and is able to measure individual machine completion times. We also require that any set of jobs assigned to a machine has to be briefly described so an arbitrary schedule may not be suitable. This paper proposes an iterative framework which repeats computing an approximation schedule, executing the PSA and updating the historical database according to the machine completion times. After each iteration, the approximation algorithm further improves some upper bound of the makespan until a 2-approximation is reached. The scheduling algorithm always assigns consecutive jobs called chains to machines keeping the historical database and the machine assignment descriptions brief.

http://devil.iit.bme.hu/~balage/tamop_cloud/WSEAS2013_kapolnai.pdf

Balázs Goldschmidt, István Hartung: Performance analysis of Windows Azure data storage options, 9th International Conference on "Large-Scale Scientific Computations" - Sozopol, Bulgaria

Abstract. Windows Azure provides an IaaS cloud service with virtual machines, web and worker roles and practically unlimited, pay-as-you-go storage options which can be used for applications requiring big data or parallel computing which is important in many _elds including biology, astronomy, nuclear physics and economics. When moving an application or computation task to the cloud it is very important to perform proof of concept performance testing and to carefully choose the proper building

blocks for the given tasks. Windows Azure provides multiple data management options with a relational SQL database for transactional data access, Azure Tables for auto scalable storage of unstructured data, and a blob storage for storing large amounts of binary data which is easily mountable to a given virtual machine. In this paper we present a general performance analysis of the Windows Azure cloud with focus on cloud storage options. We present an environment to perform automated testing of the major features of Azure storage and we also present the preliminary results and suggestions regarding the usage of the di_erent services.

http://devil.iit.bme.hu/~balage/tamop_cloud/LSSC2013_hartung.pdf

Péter Budai, Balázs Goldschmidt: Performance Analysis of Cloud-Based Applications, 9th International Conference on "Large-Scale Scientific Computations" - Sozopol, Bulgaria

Abstract. Cloud computing, and IaaS cloud services in particular suit well to resource-intensive applications by offering on-demand allocation of computing power, storage space and network bandwidth together with pay-as-you-go billing system. Typical cloud applications consist of several interdependent components, all residing on one or more dedicated virtual computers. In order to be able to accurately estimate the resource requirements of a specific component, one must carry out detailed performance analysis. In this paper, we present the general concepts and pitfalls of performance analysis in the cloud environment. Then we present a lightweight distibuted framework that is capable of generating load to and collecting performance metrics from the component instances. The capabilities of our framework will be demonstrated on a case study of the scalability analysis of a distributed MySQL relational database management system.

http://devil.iit.bme.hu/~balage/tamop_cloud/LSSC2013_budai.pdf

Balazs Simon, Balazs Goldschmidt, Imre Szeberenyi: Engineering Instance-Migratable BPEL Business Processes, In: Marko Bohanec (szerk.) Proceedings of the 15th International Multiconference INFORMATION SOCIETY -- IS 2012, Volume A. Ljubljana, Slovenia, 2012.10.08-2012.10.12. Ljubljana: Narodna in Univerzitetna Knjiznica, (2012). pp. 233-236. ISBN: 978-961-264-046-0

Abstract: Business processes are the backbone of business logic in a Service Oriented Architecture. Most software vendors provide BPEL engines to host and execute business processes. A BPEL engine usually stores the states of business process instances in a database; however, the schema of this database may change between engine versions and is most certainly varying between different vendors' engines. Usually, vendors do not provide means to migrate the states of running business process instances to other engine versions or to other vendors' products. This paper proposes a non-intrusive framework and a design pattern for BPEL processes to make their instances migratable by replaying previously recorded SOAP messages.

Balazs Simon, Balazs Goldschmidt, Karoly Kondorosi: A Metamodel for the Web Services Standards, JOURNAL OF GRID COMPUTING (ISSN: 1570-7873) 11:(4) pp. 735-752. Paper 630. (2013) DOI: 10.1007/s10723-013-9273-4

Abstract: Web services provide distributed communication in a platform independent way. The WS-* standards define how middleware aspects (security, reliability, transactions, etc.) can be realized through web services. Although the WS-Policy standard family can be used to configure the various WS-* protocols, they are very hard to construct and to maintain manually. In addition, most SOA products and grid systems implementing these standards provide their own methods for configuring these protocols, making it very difficult to match the various configuration options of different products. This fact inspired us to propose a platform independent metamodel for describing distributed systems of web services including the most important WS-* standards. The present article defines the full metamodel, it specifies the corresponding programming language formally, and it shows the productivity of the framework built around the metamodel through real-life examples. The framework is capable of generating product specific configuration files and source codes, resulting in directly interoperable applications even between different SOA products. The framework could also promote interoperability with grid systems built on WS-* protocols.