The Computer Journal Vol. 48 No. 1 © The British Computer Society 2005; all rights reserved
DOI:10.1093/comjnl/bxh069

Capsule Reviews

The Capsule Reviews are intended to provide a short succinct review of each paper in the issue in order to bring the content to a wider readership. The Capsule Reviews were compiled by Fairouz Kamarreddine. Professor Kamarreddine is an Associate Editor of *The Computer Journal* and is based in the Department of Mathematical and Computer Sciences at Heriot-Watt University, Edinburgh, UK.

Hashing of databases with the use of metric properties of the hamming space. V. Balakirsky
Hashing of databases is a particular approach to the storage of a collection of items and the retrieval of those items of the collection whose key values match given key values. The key value of an item determines the address for the storage of that item. Collisions occur when different keys have the same addresses. Since there is a trade-off between cost of storage and fast access time and since main memories usually have fast access time and a size limited by increasing cost, databases are stored in a secondary memory with slow access. The number of required access can be reduced if the values of a hash function are stored in the main memory, the records of a database are stored in an external memory and a working memory is used for storing pre-computed values of the hash function. This paper aims at solving the following task: 'given a pattern and a fixed size of working memory, form the set of addresses of records that can disagree with the pattern in the number of positions smaller than the given threshold value'. The paper uses the metric properties of the Hamming space for searching procedures. The author shows that the triangle inequality for Hamming distances generates a rejection rule for the records to be included in the subject of records that can be close to the given pattern. This rejection rule is exploited in the given hashing algorithm. An estimation of the performance of the hashing algorithm is given.

Information flow analysis for fail-secure devices. A. Rae and C. Fidge
It is important that security properties are preserved even in the event of failure of an information security device. However, it is difficult to evaluate the fault behaviour of devices, especially since dependencies between different component faults may vary. Furthermore, composition of local fault modes results in a rapid escalation of the number of cases to be evaluated. This paper takes a graph-theoretic approach to evaluate failure modes. The approach of the paper is to represent the connectivity of components within a device in fault-mode dependency tables used to detect undesirable information flow. The use of information flow here focuses on identifying violations of confidentiality rather than focussing on integrity. Fault-mode dependency tables are defined and shown to apply for local, independent component failures as well as for system-wide fault-modes. A substantial case study is given in the form of a 'Keyboard Switch' which allows sharing of equipment between computers residing in different security domains. Failure in the Keyboard Switch could allow sensitive information to flow from one computer to another. Fault-mode dependency tables are defined for this case study and 12 fault modes are identified in which undesirable information flow can occur. These 12 modes are partitioned into three groups and the interpretation of results is left to the security evaluator. Current and future work of the authors involves the development of a tool which will generate the tables automatically easing the job of the security evaluator.

Note on robust and simple authentication protocol. H.-Y. Chien, R.-C. Wang and C.-C. Yang
In an earlier issue of *The Computer Journal*, one of the authors of this paper proposed ROSI, a robust and simple authentication protocol that is based on low-cost smart cards, supports simple hashing operations and is a highly efficient password-based authentication protocol. Although password-based authentication protocols are the conventional authentication protocols to authenticate remote users, they suffer from various kinds of attacks. ROSI on the other hand, was designed to resist the known attacks. However, ROSI is vulnerable to the denial of service (DOS) attack. This paper illustrates the vulnerability of ROSI to DOS and proposes an improved-ROSI scheme which improves the server's performance. The DOS attack happens because the attacker can intercept the messages sent from the client to the server and vice versa. In other words, the DOS attack results from the requirement of state synchronization between the client and the server. In order to conquer the DOS attack the paper eliminates the state synchronization requirement. In addition, the paper argues that any secure authentication protocol should establish the session key simultaneously. The improved-ROSI scheme consists of two phases: the registration phase where the user submits his identity and password to the server through a secure channel, and the authentication and establishment phase where the user uses the server issued smart card to logon to the sever and establish fresh session keys with the sever. The improved-ROSI protocol is a key authentication protocol with key confirmation which requires five hashing operations on the client and the server. To achieve mutual authentication and key confirmation improved-ROSI requires one more additional message run than ROSI. Improved-ROSI improves the security, the functionality and the server's performance.

Temporal functional dependencies and temporal nodes bayesian networks. W. Y. LIU, N. SONG AND H. YAO

Bayesian networks (BNs) act as a framework for managing uncertainty using probability. Dynamic Bayesian networks (DBNs) are used to represent temporal probabilistic information in real world applications. However, the structure of DBNs can become very complex and temporal nodes Bayesian networks (TNBNs) are used to simplify the representation of DBNs. Since temporal functional dependencies are an important data dependency in temporal relational databases, and since temporal data dependencies can be obtained directly from real world semantics instead of using a large amount of data, this paper proposes an approach to construct a TNBN from a set of temporal functional dependencies without extensive computational costs incurred by learning TNBN directly from data. First, the temporal relational model is extended to a probabilistic temporal relational model (PTRM) by adding a probabilistic attribute. Then, three kinds of temporal data dependencies are introduced in PTRMs and the relationships amongst these three kinds are discussed. Since one of the three kinds is none other than probabilistic temporal functional dependency (PTFD) and since the paper shows that PTFD implies the second kind, which in turn implies the third kind which is equivalent to a BN, the paper sets out to construct a TNBN from temporal functional dependencies. A method is provided to carry out this construction and is illustrated by a detailed example.

A resource exchange architecture for peer-to-peer file sharing applications. C.-M. HUANG AND T.-H. HSU

In peer-to-peer (P2P) overlay networks computers can act both as clients and servers. P2P overlay networks have many advantages but also suffer from serious problems. In particular, it is difficult to find and share resources in P2P overlay networks that are associated with a limited number of users. Since file-sharing applications are increasing rapidly, it is important to collaborate and communicate in different P2P file-sharing applications. This paper aims to enable connectivity and universal access among heterogeneous P2P file-sharing networks. To achieve this the authors propose the so-called 'Shoran' architecture which provides a uniform message format (UMF) to ease communication, and enables resources to be easily found. The UMF is based on XML and uses the Resource Description Framework (RDF, developed by the W3C for Web-based metadata) as the standardized query format for resource description. A resource retrieval scheme is proposed which enables resources to be easily found with the help of the so-called ultra-peers. Shoran uses various approaches to integrate ultra-peers with heterogeneous P2P file sharing networks. Moreover, both direct and indirect retrieval mechanisms are adopted. Experimental results and analysis of the proposed architecture are provided.

Does multi-hop communication reduce electromagnetic exposure? J.-P. EBERT, D. HOLLOS, H. KARL AND M. LOBBERS

There is a growing concern about the danger and health hazards caused by electromagnetic radiation from wireless and mobile base stations. Many studies have taken place on assessing the risks of electromagnetic radiation while at the same time, various attempts have been made to turn wireless communication as energy-efficient as possible. This paper studies the emission reduction potential of the multi-hopping approach for two fundamental but different wireless network types: IEEE 802.11b as a representative example for networks with distributed channel access control and HIPERLAN/2 as a representative example for centrally controlled networks. Intuitively, multi-hopping improves energy efficiency and reduces electromagnetic radiation in the wireless case because of the more-than-linear path loss coefficient. However, there is a trade-off between high transmission power and longer transmission time. Such trade-off can be further generalized when adaptive modulation schemes are taken into account. But then, this brings about the question whether multi-hopping still reduces radiation when adaptive modulations are used to maintain system capacity. This paper answers this and other related questions for IEEE 802.11b and HIPERLAN/2.

On using handoff statistics and velocity for location management in cellular wireless networks. K.-Y. LAM, B. Y. LIANG AND C. L. ZHANG

There are two operations in the location management of wireless networks: (i) location update which records the current location of a mobile terminal (MT) into a location database and (ii) paging which searches an MT in a cell using radio signals. An MT may report its location to the location database when it enters a new location area (LA). It is important in location management to keep the cost of location update and paging to a minimum. This paper proposes a probabilistic approach called the Handoff-Velocity Prediction (HVP) scheme, to minimize the paging cost in searching an MT. HVP is based on the assumption that the movement behaviour of MTs has temporal and spatial properties. The paper also provides an algorithm to calculate the optimal distance threshold for location update generation to minimize the cost of location management. Furthermore, the paper studies the cost analysis of HVP and shows how to define optimal values for location update in order to minimize the location management cost. The performance of HVP is compared to a number of other approaches at location management.

The segment-page indexing method for large multidimensional queries. G.-H. CHA

Multidimensional indexing methods (MIMs) take into account the page structure of a disk. Since the index pages are widely scattered on a disk due to dynamic page allocation, traditional MIMs have to randomly access many index pages. In particular, MIMs have to randomly access many small index pages and so, they cannot satisfy the performance requirements for large volume of retrievals. This is especially a problem when the application in question requires a large volume of retrievals entailing reads of hundreds of pages in sequence. Performance degradation can be avoided by clustering the related index nodes, however, since existing MIMs take into account the clustering of data and not the clustering of indices and since dynamic index clustering has an overhead, the performance of MIMs cannot be improved by clustering the indices. Furthermore, in most MIMs search, performance is measured by the number of disk page access ignoring any information on the physical adjacency of individual pages. This paper aims to achieve performance improvement by reading a large number of related pages sequentially instead of reading many simple pages randomly. The authors set out to overcome the drawbacks of existing MIMs through a segment-page indexing (SP-indexing) technique. As the name suggests, this technique is based on the concept of 'segments' and considers the disk as partitioned into a collection of segments each of which consists of a set of contiguous pages on disk. This means that less random accesses are needed and hence the cost of I/O is reduced significantly. The paper presents the SP-indexing technique and gives the SP-tree concept designed to demonstrate the effectiveness of SP-indexing. The SP-tree is a multidimensional index structure used to index d-dimensional point data. The SP-tree has a hierarchical structure and divides the data space into pairwise disjoint sets. The paper presents a cost model used to estimate performance of SP-indexing. Such a cost model makes a discrimination between costs of page accesses depending on whether the pages are accessed sequentially or randomly. Since SP-indexing reduces the number of random disk access its performance is superior to that of MIMs. There are cases however where the performance of SP-indexing may be degraded. The paper discusses those cases and shows that performance degradation for updates, inserts and deletes is almost negligible. Different experimental results are given to assess the performance of SP-indexing. The paper concludes that SP-indexing can be used as an alternative index clustering scheme.

A parameterized linear array with a reconfigurable pipelined bus system: LARPBS(p). B. J. D'AURIOL AND R. MOLAKASEEMA

Many optical bus models have been proposed and studied in the literature including the linear array with a reconfigurable pipelined bus system (LARPBS). In these models synchronized concurrent access can be enabled in a pipelined fashion. In particular, LARPBS is an optical bus-based theoretical parallel computation model which has been mostly used to support algorithm development. Since the LARPBS model does not provide for performance prediction of these algorithms, this paper proposes to overcome this limitation by parameterizing the LARPBS model. In doing so, the paper is highly influenced by the bulk synchronous parallel computer (BSP). One of the main purposes of the BSP model is to define an analogous model to the von Neumann machine for sequential computers that can be a model for parallel computation. BSP is parameterized with four parameters: the number of processors/memory components, the processor speed in steps per second, the minimal time between successive synchronizations and the cost in steps per word of delivering data. The BSP model separates computation and communication and uses supersteps to more clearly structure parallel programs, enabling better cost analysis. In their aim to overcome the limitation of performance prediction the authors of this paper both capture in their proposed parameterized LARPBS (LARPBS(p)), the cost of a single operation in an algorithm as well as combine the cost of all such operations. Armed by the fact that in order to estimate the cost of executing an algorithm both computation and communication must be modeled by these operations, the authors divide computation and communication into sequences of supersteps. The LARPBS(p) model uses six parameters which capture the necessary aspects of computation, communication and barrier synchronization. These parameters include the number of processors, the communication latency, the synchronization time, the overhead time, the network capacity and the computation time. The cost analysis of the LARPBS(p) model is given and the cost model is applied to several algorithms. The paper also comments on the application of LARPBS(p) to recursive algorithms.