

Michal Zabovsky – Matus Chochlik – Karol Matiascko – Eva Valentikova *

DYNAMIC ARCHITECTURE FOR ANALYTICAL ITS SERVICES

The emergence of various technologies allowing the real-time collection, processing and transfer of data between stationary traffic devices, moving vehicles and data centers provides many new possibilities for Intelligent Transportation Systems (ITS). However, the need for mobile access to analytical data leads to problems with mobile computers connectivity. Mobile computers often suffer from limited connectivity or lack of network access. Existing data replication solutions are not well suited for mobile scenarios as well as algorithms used for data replication, in spite of the fact that it is the most desirable environment. In the following paper we introduce the algorithm for analytical data replication forming dynamic architecture including mobile computers connected by wireless network based on IEEE 802.11.

Keywords: Intelligent transportation systems, distributed database systems, data replication, mobile networks, vehicular networks

1. Introduction

The emergence of various technologies allowing the real-time collection, processing and transfer of data between stationary traffic devices, moving vehicles and data centers provides many new possibilities for Intelligent Transportation Systems (ITS). Vehicular networks (VANETs), as a subset of Mobile Ad-Hoc Networks (MANETs) are gaining importance because they allow local data exchange and limited internetwork access for vehicles moving together in a traffic flow. The vehicles which are outside the range of a stationary communication gateway can be still connected through nearby vehicles.

These networks have however very specific characteristics mainly due to the fact that the mobile elements (vehicles) are moving at high speeds, which has considerable impact on the performance of functions like connection initiation, addressing, routing, etc. On the other hand, mobile nodes could provide similar communication services as stationary nodes. The basic problem is to decide when a mobile node could be included in a communication model as a stationary node and when as regular mobile node – a client. Considering a distributed data delivery system or distributed databases system, we need to create a replication model in which mobile nodes could be included in a replication schema.

2. Problem definition

Distributed database systems and distributed database management systems have been developed in response to trend of distributed computation. In distributed models of computations several sites are connected via a communications network. However, the need for mobile access to large databases leads to problems with mobile computer connectivity. Mobile computers often suffer from

limited connectivity or complete lack of network access. Existing replicated databases are not well suited for mobile scenarios as well as algorithms used for data replication, in spite of the fact that it is the most desirable environment.

A basic problem comes with the assumption that the single network can not satisfy communication needs. This is typical for mobile communication networks used for personal communication as well as wide area communication networks. For our architecture we can consider a virtual wide area network characterized by common communication network parameters such as bandwidth, latency and others. Nevertheless, communication parameters must be continuously re-evaluated since the geospatial network is much more sensitive to communication parameters changes than commonly used local area networks.

Another principal problem is typical for mobile networks. The communication network consists of static and mobile nodes. Hence we need to take into consideration two kinds of characteristics for nodes and communication network (vertices and edges in graph representation). In principle, static and mobile nodes need each different management by the architecture itself.

Since a virtual network allows us to serve fundamental needs for huge information based services such as analytical information distribution, we will use the static model for basic data distribution across a network. The replication itself provides mechanism for data distribution to that part of network where the information portion is consumed. Hence, data replication benefits follow our needs for analytical data distribution.

In general, wireless communication networks keep track of a mobile client's location through a profile. This profile contains not only client's current position (or location) but also another

* Michal Zabovsky, Matus Chochlik, Karol Matiascko, Eva Valentikova

Department of Informatics, Faculty of Management Science and Informatics, University of Zilina, Slovakia,
E-mail: michal.zabovsky@fri.uniza.sk

information used for billing and authentication [1]. We recommend using profiles for application and services implemented in ITS. Since the connectivity, bandwidth, data storage and processing capacity of the individual elements in ITS's virtual architecture varies in time, the system must also handle the load balancing and prioritization of the distributed information [2].

For the fragment allocation problem the assumption comes with WAN described as a network. The network consists of sites $S = \{S_1, S_2, \dots, S_m\}$, on which a set of transactions $T = \{T_1, T_2, \dots, T_q\}$ is running and a set of fragments $F = \{F_1, F_2, \dots, F_n\}$ into which global data can be distributed [3]. This is the generalized problem and it could be adopted for more specific purposes focusing the optimal data distribution.

The principal problem of distribution follows two definitions of optimality: minimal cost and maximal performance. The cost function consists of the cost of storing each F_j on site S_k . Existing models and implementations are mostly based on a read/write pattern for the first definition of optimality or response time/throughput at each site for the second one. Both optimizations can be performed either statically or dynamically.

Static optimization solves the problem of partitioning all global relations during the fragmentation phase of distributed database design [4]. Adaptive optimization tries to make suboptimal solution by using some kind of heuristic algorithm taking in consideration runtime behavioral statistics of a distributed database [5] [6]. Alternative methods are discussed e.g. in [7].

3. Model for client-based ITS services

Considering advanced services provided by ITS, we need to take into account significant data portions transported by a communication network. Since for low data consuming services common GSM networks can be used, this kind of network is inappropriate for large data transmissions. Of course, advanced communication protocols could be used e.g. HSDPA, but the question of low cost independent solution is still relevant.

A basic model for adaptive data replication including mobile nodes connected via IEEE 802.11 standards was introduced in [8] [9]. By the static fragment allocation the fragments are located at the sites from which they are most frequently accessed. Since the distributed database system is rather dynamic, the main problem with previous, static allocation of fragments comes with changing workload. This occurs when the access frequencies to various portions of database from a particular site vary with time. Even very simple methods for dynamic data allocations are able to improve the system throughput by 30 percent. Experimental evaluation of dynamic data allocation strategies can be found in [10]. To determine when a re-allocation is needed, algorithms proposed in [10] maintain weighted counters of the number of access from each site to each block. For effective estimation the aging factor is used to update counters. The main problem can be divided into the two problems: how to detect changes in workload and how to dynam-

ically re-allocate fragments of database result in improved throughput.

Considering mobile nodes, the communication network characteristics must be evaluated carefully since they form a transmission cost matrix. Basic characteristics for static nodes, such a throughput and latency (or round-trip time), can be estimated reliably by tools such a *tstat* or *pathchar*. For wireless clients is estimation more difficult. Technology for wireless communication based on the IEEE 802.11 standards provides connection with variable transmission characteristics. Rather than throughput and latency a *signal-to-noise ratio* (SNR) is a characteristic that has to be involved in a replication model. SNR affects both characteristics throughput and latency in a significant way as shown in [11]. SNR directly impacts the performance of a wireless connection. A higher SNR value (in dB) means that the signal strength is stronger in relation to the noise levels, which allows higher data rates and fewer retransmissions.

The linear mathematical model for throughput prediction based on previous observations can be defined as [11]:

$$T = T_{max} \quad SNR > SNR_C \quad (1)$$

$$T = A \times (SNR - T_0) \quad SNR \leq SNR_C \quad (2)$$

Where T_{max} is a saturation throughput, A defines slope, T_0 is a breaking point where T_{max} is changing to curve described by A , SNR_0 is a cutoff SNR specified by a hardware vendor and SNR_C defines a critical threshold. A respective exponential model is also described in [11], for the proposed solution linear algorithm is sufficient enough to describe communication network characteristics. Now we can define a set SNR_C of m elements, where each value snr_{ci} represents the critical threshold for site i as $SNR_C = \{snr_{c1}, snr_{c2}, \dots, snr_{cm}\}$. When the site i is a static node, the critical threshold is zero. Finally, we need to implement a function which returns the current SNR value for mobile nodes. Such a function is necessary to implement on each site from the replication schema since it depends on a particular configuration.

4. Adaptive algorithm

Based on previous observations, we defined the adaptive replication algorithm [8] which manages mobile node access into the replication schema. For ITS architecture the main modification is based on an assumption that mobile nodes are using analytical data only for reading. Since we can spread the replication schema by this algorithm, it is increasing availability of large data set to be replicated on mobile nodes. By experimental evaluation we simulated transfer of fragments inside the replication schema by using the algorithm without and with mobility management.

The modified algorithm for analytical data distribution in ITS architecture has in comparison with the algorithm proposed in [8] only one phase. The expansion test is used for spreading data over mobile nodes. The second test, test of contraction, is not used

since we are managing to write operations only on static nodes. The replication schema in this architecture is not necessarily associated with the database but the data distribution itself.

Test of expansion:

- Step 1: The control process examines read counters for each fragment containing analytical data.
- Step 2: The site with the highest counter value is marked as a candidate for fragment re-allocation.
- Step 3: If the candidate is the site on which the fragment is currently located, go to step 6.
- Step 4: For the mobile nodes get SNR and SNR_c values. For static nodes return $SNR = 100$ and $SNR_c = 0$.
- Step 5: If $SNR > SNR_c$ then copy the fragment from the original site to the candidate site. Otherwise choose the site with the highest counter value from the set of unmarked sites and mark it as a candidate for fragment re-allocation – then go to step 3.
- Step 6: Wait for a specified number of transactions to be completed and then go to step 1.

The experimental results show that for regular connection when a mobile is not suffering from a limited access the number of transactions per second (TPS) stays on the same level. TPS is one of the most common benchmark criteria, thus we present the replication performance in this value. For mobile nodes with variable

SNR the situation is much more interesting. The adaptive algorithm shows a significantly better TPS number and this gain from implementation can be used for a better mobility management. By SNR checking we can easily improve the overall performance and select mobile nodes for analytical (and thus large) data transmission.

The overall replication cost is reduced mostly for an unstable communication network as shown in Figure 1. Local maximum is presented about the critical SNR (SNR_c). It means that at this point even our algorithm is not able to improve the performance significantly due to the variability of communication network characteristics. For bad network parameters, on the other hand, the introduced algorithm is significantly better than a non-managed replication. The observed results show about 60 percent better overall response time. This is given by the fact that the proposed algorithm avoids the replica transmission to the node with communication problems. Hence the unwilling transmission is not causing problem of Nash equilibrium in the communication network [12].

5. Conclusion and future work

Performance in distributed database systems is heavily dependent on allocation of data among the sites of a replication schema. The static allocation provides only a limited response to workload changes. The situation is even worse when mobile nodes are included in the replication schema. We presented the algorithm for dynamic re-allocation of data with mobile computers included in the replication schema. The proposed algorithm offers a significantly increased performance for nomadic nodes with a limited connection. Our experiments make a practical case for future development of algorithms for changing environment such as intelligent transportation systems, location aware application and information systems for mobile users.

Acknowledgement:

This contribution is the result of the project implementation: Centre of excellence for systems and services of intelligent transport, ITMS 26220120028 supported by the Research & Development Operational Programme funded by the ERDF.

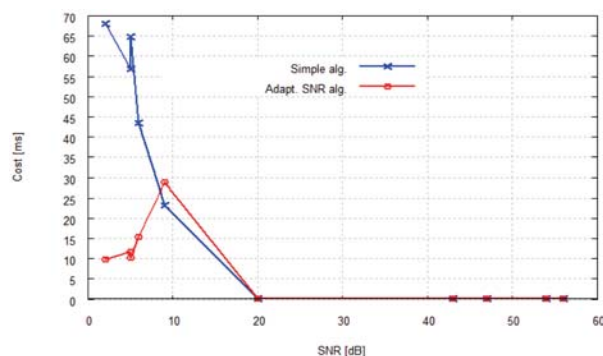


Fig. 1 Adaptive replication algorithm - replication cost



Agentúra
Ministerstva školstva, vedy, výskumu a športu SR
pre štrukturálne fondy EÚ

"Podporujeme výskumne aktivity na Slovensku/Projekt je spolufinancovaný zo zdrojov EÚ."

References:

- [1] TIAN, K. Q., COX, D. C.: *Mobility Management in Wireless Networks : Data Replication Strategies and Applications*. Kluwer Academic Publishers, 2004.
- [2] MARTINCOVA, P., GRONDZAK, K.: Effective distributed query processing on the grid. In: *2nd International Workshop on Grid Computing for Complex Problems - GCCP 2006*, Bratislava, 2006, pp. 129-135.
- [3] HUANG, Y-F., CHEN, J-H.: Fragment Allocation in Distributed Database Design. *Journal of Information Science and Engineering*, No. 17, pp. 491-509, 2001.
- [4] JANACEK J.: *Mathematical programming [Matematicke programovani]*. Zilina : EDIS-ZU, 1999 (in Czech).
- [5] WOLFSON, O., JAJODIA, S., HUANG, Y.: An Adaptive Data Replication Algorithm. *ACM Transactions on Database Systems (TODS)*, Vol. 22, No. 2, pp. 255-314, 1997.
- [6] GRONDZAK, K. et al.: Distributed Solving of combinatorial problems. In *2nd International Workshop on Grid Computing for Complex Problems - GCCP 2006*, Bratislava, 2006, pp. 82-87.
- [7] ARTHUR, L., CORCORAN, HALE, J.: A Genetic Algorithm for Fragment Allocation in a Distributed Database System. In: *Proc. of the 1994 ACM symposium on Applied computing*, Phoenix, Arizona, United States, 1994, pp. 247-250.
- [8] MATIASKO, K., ZABOVSKY, M.: Adaptive Replication Algorithm for Mobile Computers. *Acta Electronica et Informatica*, Vol. 8, No. 2, pp. 18-21, 2008.
- [9] MATIASKO, K., ZABOVSKY, M.: A Data Replication for Mobile Environment. *J. of Electrical Engineering*, Vol. 59, No. 5, pp. 277-280, 2008.
- [10] BRUNSTROM, A., LEUTENEGGER, S. T., SIMHA, R.: *Experimental Evaluation of Dynamic Data Allocation Strategies in a Distributed Database with Changing Workload*. Institute for Computer Applications in Science and Engineering (ICASE), Technical Report TR-95-2, 1995.
- [11] CHEN N., CHEN, J. K., RAPPAPORT, TH. S.: Measured Traffic Statistics and Throughput of IEEE 802.11b Public WLAN Hotspots with Three Different Applications, *IEEE Transactions on Wireless Communications*, Vol. 5, No. 11, pp. 3296-3305, 2006.
- [12] ALMENDRAL, J. A., LUIZ, L., VINCENT, CH., SANJUAN, M. A. F.: Congestion Schemes and Nash Equilibrium in Complex Networks. *Physica A*, No. 355, pp. 602-618, 2005.