
FOUNDATIONS OF DISTRIBUTED COMPUTING

Abstract

This project studies the fundamental mathematical questions arising in the context of distributed computing. On the one hand, we develop novel distributed algorithms, e.g. for computing the median in a network, or for approximating a dominating set using local information only.

On the other hand, we derive impossibility and lower bound results that show the optimality of certain algorithms or models and that answer the question of what cannot be achieved in a distributed system.

Section

Distributed Computing Group

Coordinator

Prof. Dr. Roger Wattenhofer

TIK Participants

Roland Flury
Olga Goussevskaia
Thomas Locher

Yvonne-Anne Oswald

Stefan Schmid

Pascal von Rickenbach

Funding

EU DYNAMO

SNF Decentralized Internetworking

Web

www.dcg.ethz.ch

FOUNDATIONS OF DISTRIBUTED COMPUTING

Distributed Algorithms

What can be computed locally? Though this question is vital for the distributed computing community, results have been rare. The importance of locality stems from the desire of achieving a global goal based on local information only. This is not only one of the key challenges when developing fast distributed algorithms, but it also improves fault-tolerance and may be a stepping stone towards understanding dynamic networks. Recent advances in networking and ever-growing distributed systems drive the need for a thorough understanding of locality issues. We study the impact of locality in the classic message-passing model where the distributed system is modeled as a graph: Nodes represent the processors and two nodes can communicate if and only if they share an edge in the graph.

We present time lower bounds for the distributed approximation of minimum vertex cover and related problems such as minimum dominating set. In particular we examine algorithms with k communication rounds and we determine the number of rounds required in order to achieve a constant or even only a polylogarithmic approximation ratio. The latter lower bounds also hold for the construction of maximal matchings and maximal independent sets. We describe the first distributed approximation algorithm yielding a non-trivial approximation ratio for the minimum dominating set problem in a constant number of rounds based on LP relaxation techniques. For small k -fold dominating sets in general graphs and unit disk graphs we propose fast approximation algorithms requiring only small messages.

In another work, we provide an almost tight classification of the possible tradeoff between the amount of local information and the quality of the global solution for general covering and packing problems. Specifically, we give distributed algorithms which achieve a non-trivial approximation ratio for general covering and packing problems. Moreover, we prove these algorithms to be close to optimal by proving a lower bound on the approximability of packing problems given that each node has to base its decision on information from its k -neighborhood.

Another problem we study is graph coloring in a distributed setting. For algorithms where every node of the network can only send its initial color to all its neighbors, we show a new lower bound on the number of colors of the computed coloring. If such one-round algorithms are iteratively applied, we derive an even smaller lower bound, thus improving previous results considerably.

While the time lower bounds for the approximation of minimum vertex cover and related problems hold even if the message sizes are unbounded, there are interesting problems that become hard to solve once a reasonable maximum on the message size is enforced. For example, it is easy to find the k^{th} smallest element in any graph where each node holds a single element if we allow

unbounded message sizes: A simple flooding-echo procedure on a pre-computed spanning tree where all nodes send a single message containing the numbers of all its children to its respective parent solves the problem. This simple approach is inefficient in the worst case if only a constant number of all n elements can be packed into a single message. We propose a randomized algorithm whose time complexity is much lower with high probability. Additionally, a deterministic algorithm is presented which considerably improves the best known time bound for deterministic algorithms. Moreover, we prove a time lower bound for any randomized or deterministic algorithm, implying that the randomized algorithm is asymptotically optimal.

Another problem that we study is the synchronization of distributed clocks. In particular we study the gradient clock synchronization (GCS) problem, in which the worst-case clock skew between neighboring nodes has to be minimized. For several intuitive clock synchronization algorithms, which attempt to minimize the skew at all times, we show that the clock skew between neighboring nodes can be quite significant: All algorithms incur a worst-case clock skew in the order of D between neighboring nodes, where D denotes the diameter of the network. Additionally, we present an oblivious algorithm with a worst-case skew in the order of the square root of D , which is the best result for gradient clock synchronization to date.

$$\sum_{Q \in \binom{N}{k}} \prod_{j \in Q} \frac{p_j}{1-p_j} = \frac{1}{k} \sum_{i=1}^m \left(\frac{p_i}{1-p_i} \cdot \sum_{Q \in \binom{N \setminus \{i\}}{k-1}} \prod_{j \in Q} \frac{p_j}{1-p_j} \right)$$

$$\frac{1}{k} \sum_{i=1}^m \left(\frac{p_i}{1-p_i} \cdot \sum_{Q \in \binom{N \setminus \{i\}}{k-1}} \prod_{j \in Q} \frac{p_j}{1-p_j} \right)$$

$$\frac{1}{k} \sum_{i=1}^m \frac{p_i}{1-p_i} \cdot \left(\sum_{Q \in \binom{N \setminus \{i\}}{k-1}} \prod_{j \in Q} \frac{p_j}{1-p_j} \right) = \frac{1}{k} \sum_{i=1}^m \frac{p_i}{1-p_i} \cdot A$$

Fig. 1: Formulas

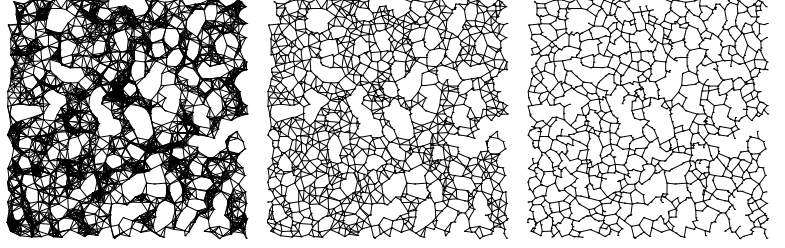


Fig. 2: Controlling the topology

Wireless Communication

A most important lesson in wireless networking is that concurrent transmissions may cause interference. If too many devices transmit simultaneously, the interference caused by these transmissions will prevent a receiver from receiving the signal. On the other hand, if too few nodes transmit at the same time, valuable bandwidth is wasted and the overall throughput may suffer. Hence, the classic problem faced by any MAC layer or scheduling protocol is to select neither too many nor too few devices for concurrent transmission. Thus, the crucial question to be answered is: How much time is required to schedule a given set of directed links between pairs of nodes that indicate communication requests?

Often, a graph model is applied to capture the characteristics of wireless networks. In this model a node is assumed to receive a message correctly if no other node in close proximity transmits at the same time. Thus, the problem boils down to solving variants of coloring or independent set problems. Clearly, this is a tremendous simplification of the physical reality. Particularly, interference caused by different transmitters may accumulate and is not binary, i.e., does not stop at any specific border. Moreover, a node may successfully receive a message in spite of being in the transmission range of other simultaneous transmitters. In fact, a message is successfully received by a node if the ratio between the received signal strength and the ambient noise plus interference from other nodes exceeds a certain hardware-specific threshold. Adopting this notion of signal-to-noise-plus-interference ratio (SINR) is also known as the physical model.

Not surprisingly, experiments demonstrate that graph-based models capture reality inadequately. In addition, we show that the theoretical limits of any protocol based on graph-based models can be broken by a protocol for the physical model.

In wireless networks it is typically required that all nodes are connected. We therefore study the scheduling complexity of connecting a given number of nodes located at arbitrary positions by some communication tree. We present an algorithm constructing such a tree, assigning power levels and time slots to each link such that in polylogarithmic time, all transmissions are received. Hence, theoretically, there is no fundamental scalability problem for scheduling in wireless networks.

For arbitrary topologies, we prove that even in worst-case networks, the number of time slots required to successfully schedule all links of an arbitrary topology is proportional to the squared logarithm of the number of network nodes times a static interference measure.

In another work, we propose a new interference measure, disturbance. Most known approaches suffer from extremely bad performance in certain scenarios even if disturbance is low. We present an algorithm which achieves a speed up that can be exponential in the size of the network compared to previously known solutions.

Even if all nodes send with the same power level, scheduling remains difficult. We also study the problem of scheduling in the geometric SINR model, which explicitly uses the fact that nodes are distributed in the Euclidean plane. We prove that the problem is NP-complete and devise approximation algorithms.

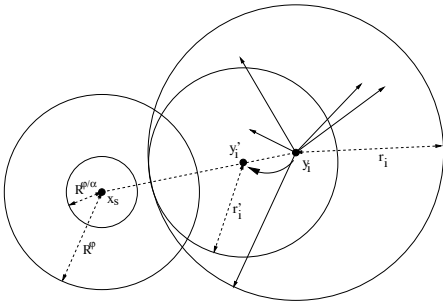


Fig. 5: Scheduling complexity analysis

Selected Publications

Raphael Eidenbenz, Yvonne Anne Oswald,
Stefan Schmid, and Roger Wattenhofer
Manipulation in Games
ISAAC 2007

Fabian Kuhn, Thomas Locher,
and Roger Wattenhofer
Tight Bounds for Distributed Selection
SPAA 2007

Thomas Moscibroda, Yvonne Anne Oswald,
and Roger Wattenhofer
How Optimal are Wireless Scheduling Protocols?
INFOCOM 2007

Olga Goussevskaia, Yvonne Anne Oswald,
and Roger Wattenhofer
Complexity in Geometric SINR
MOBIHOC 2007

Thomas Locher and Roger Wattenhofer
Oblivious Gradient Clock Synchronization
DISC 2006

Thomas Moscibroda and Roger Wattenhofer
**The Complexity of Connectivity in Wireless
Networks**
INFOCOM 2006

Thomas Moscibroda, Roger Wattenhofer,
and Aaron Zollinger
**Topology Control Meets SINR: The Scheduling
Complexity of Arbitrary Topologies**
MOBIHOC 2006

Fabian Kuhn and Roger Wattenhofer
On the Complexity of Distributed Graph Coloring
PODC 2006

Hagit Attiya, Fabian Kuhn, Greg Plaxton,
Mirjam Wattenhofer, and Roger Wattenhofer
Efficient Adaptive Collect Using Randomization
Distributed Computing, 18(3):179–188, 2006

Hagit Attiya, Fabian Kuhn, Greg Plaxton,
Mirjam Wattenhofer, and Roger Wattenhofer
**Fault-Tolerant Clustering in Ad Hoc and Sensor
Networks**
ICDCS 2006

Fabian Kuhn, Thomas Moscibroda,
and Roger Wattenhofer
The Price of Being Near-Sighted
SODA 2006

Fabian Kuhn and Roger Wattenhofer
**Constant-Time Distributed Dominating Set
Approximation**
Springer Journal for Distributed Computing, 17(4):
May 2005

Fabian Kuhn, Thomas Moscibroda,
and Roger Wattenhofer
What Cannot Be Computed Locally!
PODC 2004

Fabian Kuhn and Roger Wattenhofer
**Dynamic Analysis of the Arrow Distributed
Protocol**
SPAA 2004