

Diss. ETH No. 18599

**Algorithmic Solutions for Transient Faults
in Communication Networks**
On Swap Edges and Local Algorithms

A dissertation submitted to

ETH Zurich
Swiss Federal Institute of Technology

for the degree of Doctor of Sciences

presented by

Beat Gfeller
Master of Science ETH
born September 15, 1980,
citizen of Röthenbach im Emmenthal, Switzerland

accepted on the recommendation of

Prof. Dr. Peter Widmayer, ETH Zurich, Switzerland
examiner

Prof. Dr. Guido Proietti, University of L'Aquila, Italy
co-examiner

2009

Abstract

Communication networks such as the Internet are ubiquitous today, and play a key role in modern society. Given their widespread use, measures must be taken to ensure the continued operation of these networks even in the presence of failures of individual parts. It is particularly important to efficiently deal with *transient faults*, since these are predominant in many network environments. This tenet is the main motivation for this dissertation, which provides algorithmic solutions for transient faults in communication networks. In this context, we consider network environments with varying characteristics, from rather static to highly dynamic networks.

For handling transient faults in static networks, we use the *swap* approach, which has been designed specifically for this environment: In the absence of any faults, a spanning tree of the underlying network is used for communication. When a link fails, the tree is split into two components, which are then reconnected in a best possible way by adding a single backup link, called a *best swap*. This small and local modification is well-suited for handling transient faults, because it quickly re-establishes communication without a large restructuring effort. Moreover, since the network topology is static, a best swap can be precomputed for every link in the spanning tree before using the network. We present efficient algorithms for precomputing best swaps, for a variety of spanning tree types. In addition, we show how to route messages efficiently through a network using the swap approach.

If too many faults occur in a network, the services it provides will break down. We therefore study the threshold of tolerable faults in networks where transient failures are very frequent. To that end, we consider one of the most fundamental problems in distributed computing, the *consensus* problem, in synchronous networks. In this problem, a number of agents, who can only communicate by exchanging messages through a network, must negotiate a globally unanimous decision. We provide a precise characterization of the number of tolerable failures for reaching such a consensus using a given network.

Finally, we investigate highly dynamic networks, such as wireless ad hoc networks, where in addition to frequent transient faults, the network topology keeps constantly changing as well. In such a

setting, it may be impossible to gather information about the entire network topology. Therefore, a completely different approach is necessary: The task of computing a global solution is distributed among the nodes of the network, and each node only communicates with nearby nodes. Surprisingly, it is possible to find such so-called *local algorithms*, for some fundamental problems in wireless networks. We give efficient local distributed algorithms for two such problems: Computing a maximal independent set, which is a basic building block for distributed coordination, and computing a small connected dominating set, which is useful for energy-efficient message broadcast.

Zusammenfassung

Kommunikationsnetze wie das Internet sind heutzutage allgegenwärtig, und spielen eine wichtige Rolle in der modernen Gesellschaft. In Anbetracht ihrer verbreiteten Nutzung ist es wichtig, Massnahmen zu treffen, um das Funktionieren dieser Netze zu gewährleisten, selbst wenn einzelne Komponenten davon ausfallen. Es ist besonders wichtig, effiziente Mittel gegen *kurzzeitige Fehler* zu finden, da solche Fehler in vielen Netzwerkumgebungen die Mehrheit ausmachen. Dieser Grundgedanke ist die Motivation für diese Dissertation, welche algorithmische Lösungen für kurzzeitige Fehler in Kommunikationsnetzen anbietet. In diesem Zusammenhang werden Netzwerkumgebungen mit verschiedenen Charakteristika betrachtet, von eher statischen bis zu hochgradig dynamischen Netzen.

Um kurzzeitige Fehler in statischen Netzen zu handhaben, benutzen wir das Konzept von *Austauschkanten*, welches für genau diesen Zweck ausgerichtet ist: Solange kein Fehler auftritt, wird zur Kommunikation ein Spannbaum des zugrundeliegenden Netzes benutzt. Wenn eine Verbindung ausfällt, zerfällt das Netz in zwei Komponenten, welche dann auf bestmögliche Weise wieder verbunden werden, indem eine Ersatzkante, die sogenannte *beste Austauschkannte*, zum Netz hinzugefügt wird. Diese kleine und lokale Modifikation eignet sich für das Überbrücken von kurzzeitigen Fehlern, weil sie den Nachrichtenaustausch schnell wieder herstellt, ohne eine aufwändige Umstrukturierung des Netzes zu erfordern. Überdies kann, da das Netz statisch ist, für jede Netzkante eine beste Austauschkannte vor dem Einsatz des Netzes vorberechnet werden. Wir präsentieren effiziente Algorithmen für die Vorbereitung von besten Austauschkannten, für eine Auswahl von Spannbaum-Typen. Zudem zeigen wir, wie Nachrichten auch während des Einsatzes von Austauschkannten effizient durch das Netz gelenkt werden können.

Wenn in einem Netz zu viele Fehler auftreten, kann es seine Dienste nicht mehr verrichten. Daher interessieren wir uns für das maximal erträgliche Mass an kurzzeitigen Fehlern, das ein gegebenes Netz tolerieren kann. Zu diesem Zweck betrachten wir eines der grundlegendsten Probleme im Gebiet des verteilten Rechnens, das *Konsens*-Problem, in taktgesteuerten Netzen. In diesem Problem muss eine Anzahl von Akteuren, welche nur durch den Austausch von Nachrichten über ein Netz kommunizieren können, eine global einstimmige Entscheidung aushandeln. In diesem Teil der Arbeit geben wir

eine exakte Charakterisierung für die maximale Anzahl von Fehlern in einem gegebenen Netz, welche für das Konsens-Problem tolerierbar sind.

Schliesslich untersuchen wir hochgradig dynamische Netze, wie zum Beispiel kabellose Ad Hoc Netze, in denen nicht nur viele kurzzeitige Fehler auftreten, sondern sich auch die Topologie ständig ändert. In solchen Netzen kann es unmöglich sein, eine Gesamtübersicht über die Netzwerktopologie zu gewinnen. Daher ist ein völlig anderer Ansatz notwendig: Die Aufgabe, eine globale Lösung zu berechnen, wird unter den Knoten im Netz aufgeteilt, und jeder Knoten kommuniziert nur mit Knoten in seiner Nähe. Erstaunlicherweise ist es möglich, solche sogenannten *Lokalen Algorithmen* für verschiedene grundlegende Probleme in kabellosen Netzen zu finden. Wir entwerfen effiziente lokale verteilte Algorithmen für zwei solche Probleme: Die Berechnung einer maximalen unabhängigen Menge, welche ein grundlegendes Hilfsmittel für verteilte Koordination ist, und die Berechnung einer kleinen zusammenhängenden Menge, welche nützlich für energie-effizienten Nachrichtenversand ist.