Resource Allocation and Cross-Layer Control in Wireless Networks
Resource Allocation and Cross-Layer Control in Wireless Networks

Leonidas Georgiadis
Dept. of Electrical and Computer Engineering
Aristotle University of Thessaloniki
Thessaloniki 54124, Greece
leonid@auth.gr

Michael J. Neely
Dept. of Electrical Engineering
University of Southern California
Los Angeles, CA 90089, USA
mjneely@usc.edu

Leandros Tassiulas
Computer Engineering and Telecommunications Dept.
University of Thessaly
Volos, Greece
leandros@uth.gr

Full text available at: http://dx.doi.org/10.1561/1300000001
Editorial Scope

Foundations and Trends® in Networking will publish survey and tutorial articles in the following topics:

- Ad Hoc Wireless Networks
- Sensor Networks
- Optical Networks
- Local Area Networks
- Satellite and Hybrid Networks
- Cellular Networks
- Internet and Web Services
- Protocols and Cross-Layer Design
- Network Coding
- Energy-Efficiency
- Incentives/Pricing/Utility-based
- Games (co-operative or not)
- Security
- Scalability
- Topology
- Control/Graph-theoretic models
- Dynamics and Asymptotic Behavior of Networks

Information for Librarians

Foundations and Trends® in Networking, 2006, Volume 1, 4 issues. ISSN paper version 1554-057X, ISSN online version 1554-0588. Also available as a combined paper and online subscription.
Resource Allocation and Cross-Layer Control in Wireless Networks

Leonidas Georgiadis\(^1\), Michael J. Neely\(^2\) and Leandros Tassiulas\(^3\)

\(^1\) Aristotle University of Thessaloniki, Thessaloniki 54124, Greece, leonid@auth.gr
\(^2\) University of Southern California, Los Angeles, CA 90089, USA, mjneely@usc.edu
\(^3\) University of Thessaly, Volos, Greece, leandros@uth.gr

Abstract

Information flow in a telecommunication network is accomplished through the interaction of mechanisms at various design layers with the end goal of supporting the information exchange needs of the applications. In wireless networks in particular, the different layers interact in a nontrivial manner in order to support information transfer. In this text we will present abstract models that capture the cross-layer interaction from the physical to transport layer in wireless network architectures including cellular, ad-hoc and sensor networks as well as hybrid wireless-wireline. The model allows for arbitrary network topologies as well as traffic forwarding modes, including datagrams and virtual circuits. Furthermore the time varying nature of a wireless network, due either to fading channels or to changing connectivity due to mobility, is adequately captured in our model to allow for state dependent network control policies. Quantitative performance measures that capture the quality of service requirements in these systems depending on the supported applications are discussed, including throughput maximization,
energy consumption minimization, rate utility function maximization as well as general performance functionals. Cross-layer control algorithms with optimal or suboptimal performance with respect to the above measures are presented and analyzed. A detailed exposition of the related analysis and design techniques is provided.
Contents

1 Introduction 1

2 The Network Model and Operational Assumptions 7

2.1 Link rate function examples for different networks 9
2.2 Routing and network layer queueing 17
2.3 Flow control and the transport layer 20
2.4 Discussion of the assumptions 21

3 Stability and Network Capacity 25

3.1 Queue stability 25
3.2 The network layer capacity region 29
3.3 The capacity of one hop networks 36

4 Dynamic Control for Network Stability 41

4.1 Scheduling in an ON/OFF downlink 41
4.2 Network model 45
4.3 The stabilizing dynamic backpressure algorithm 48
4.4 Lyapunov stability 51
4.5 Lyapunov drift for networks 56
4.6 Time varying arrival rates 59
In cross-layer designs of wireless networks, a number of physical and access layer parameters are jointly controlled and in synergy with higher layer functions like transport and routing. Furthermore, state information associated with a specific layer becomes available across layers as certain functions might benefit from that information. Typical physical and access layer functions include power control and channel allocation, where the latter corresponds to carrier and frequency selection in OFDM, spreading code and rate adjustment in spread spectrum, as well as time slot allocation in TDMA systems. Additional choices in certain wireless network designs may include the selection of the modulation constellation or the coding rate, both based on the channel quality and the desired rates [55, 156]. Due to the interference properties of wireless communication, the communication links between pairs of nodes in a multinode wireless environment cannot be viewed independently but rather as interacting entities where the bit rate of one is a function of choices for the physical and access layer parameters of the others. Our cross-layer model in this text captures the interaction of these mechanisms, where all the physical and access layer parameters are collectively represented through a control vector $I(t)$. 

Another intricacy of a wireless mobile communication network is the fact that the channel and the network topology might be changing in time due to environmental factors and user mobility respectively. That variation might be happening at various time scales from milliseconds in the case of fast fading to several seconds for connectivity variations when two nodes get in and out of coverage of each other as they move. Actions at different layers need to be taken depending on the nature of the variability in order for the network to compensate in an optimal manner. All the relevant parameters of the environment that affect the communication are represented in our model by the topology state variable $S(t)$. The topology state might not be fully available to the access controller, which may observe only a sufficient statistic of that. The collection of bit rates of all communicating pairs of nodes at each time, i.e. the communication topology, is represented by a function $C(t) = C(I(t), S(t))$. Note that the function $C(\cdot, \cdot)$ incorporates among others the dependence of the link rate on the Signal-to-Interference plus Noise Ratio (SINR) through the capacity function of the link. Over the virtual communication topology defined by $C(t)$, the traffic flows from the origin to the destination according to the network and transport layer protocols. Packets may be generated at any network node having as final destination any other network node, potentially several hops away. Furthermore, the traffic forwarding might be either datagram or based on virtual circuits, while multicast traffic may be incorporated as well. The above model captures characteristics and slightly generalizes systems that have been proposed and studied in several papers including [108, 111, 113, 135, 136, 143, 144, 147, 149]. That model is developed in detail in Section 2 while representative examples of typical wireless models and architectures that fit within its scope are discussed there.

The network control mechanism determines the access control vector and the traffic forwarding decisions in order to accomplish certain objectives. The quantitative performance objectives should reflect the requirements posed by the applications. Various objectives have been considered and studied in various papers including the overall throughput, power optimization, utility optimization of the allocated rates as well as optimization of general objective functions of throughput and/or
power. In the current text we present control strategies for achieving these objectives.

The first performance attribute considered is the capacity region of the network defined as the set of all end-to-end traffic load matrices that can be supported under the appropriate selection of the network control policy. That region is characterized in two stages. First the ensemble of all feasible long-term average communication topologies is characterized. The capacity region includes all traffic load matrices such that there is a communication topology from the ensemble for which there is a flow that can carry the traffic load and be feasible for the particular communication topology. Section 3 is devoted to the characterization of the capacity region outlined above.

The capacity region of the network should be distinguished from the capacity region of a specific policy. The latter being the collection of all traffic load matrices that are sustainable by the specific policy. Clearly the capacity region of the network is the union of the individual policy capacity regions, taken over all possible control policies. One way to characterize the performance of a policy is by its capacity region itself. The larger the capacity region the better the performance will be since the network will be stable for a wider range of traffic loads and therefore more robust to traffic fluctuations. Such a performance criterion makes even more sense in the context of wireless ad-hoc networks where both the traffic load as well as the network capacity may vary unpredictably. A policy A is termed “better” than B with respect to their capacity regions, if the capacity region of A is a superset of the capacity region of B. A control policy that is optimal in the sense of having a capacity region that coincides with the network capacity region and is therefore a superset of the capacity region of any other policy was introduced in [143, 147]. That policy, the max weight adaptive back-pressure policy, was generalized later in several ways [111, 115, 135, 149] and it is an essential component of policies that optimize other performance objectives. It is presented in Section 4. The selection of the various control parameters, from the physical to transport layer, is done in two stages in the max weight adaptive back pressure policy. In the first stage all the parameters that affect the transmission rates of the wireless links are selected, i.e. the function $C(I(t), S(t))$ is determined. In the
4 Introduction

Second stage routing and flow control decisions to control multihop traffic forwarding are made. The back pressure policy consists in giving priority in forwarding through a link to traffic classes that have higher backlog differentials. Furthermore, the transmission rate of a link that leads to highly congested regions of the network is throttled down. In that manner, the congestion notification travels backwards all the way to the source and flow control is performed. Proofs of the results based on Lyapunov stability analysis are presented also in Section 4.

The stochastic optimal control problem where the objective is the optimization of a performance functional of the system is considered in Sections 5 and 6. The development of optimal policies for these cases relies on a number of advances including extensions of Lyapunov techniques to enable simultaneous treatment of stability and performance optimization, introduction of virtual cost queues to transform performance constraints into queueing stability problems and introduction of performance state queues to facilitate optimization of time averages. These techniques have been developed in [46, 108, 115, 116, 136, 137] for various performance objectives. More specifically in Section 5, the problem of optimizing a sum of utility functions of the rates allocated to the different traffic flows is considered. That formulation includes the case of the traffic load in the system being out of the capacity region, which case some kind of flow control at the edges of the network needs to be employed. That is done implicitly through the use of performance state queues, allowing adjustment of the optimization accuracy through a parameter. The approach combines techniques similar to those used for optimization of rate utility functions in window flow controlled sessions in wireline networks, with max weight scheduling for dealing with the wireless scheduling. In Section 6, generalization of these techniques for optimization functionals that combine utilities with other objectives like energy expenditure are given and approaches relying on virtual cost queues are developed.

Most of the results presented in the text are robust on the statistics of the temporal model both of the arrivals as well as the topology variation process. The traffic generation processes might be Markov modulated or belong to a sample path ensemble that complies with certain burstiness constraints [35, 148]. Similarly, the variability of the
topology might be modeled by a hidden Markov process. These models are adequate to cover most of the interesting cases that might arise in real networks. The proofs in the text are provided for a traffic generation model that covers all the above cases and it was considered in [115]. The definition of stability that was used implies bounded average backlogs. The emphasis in the presentation is on describing the models and the algorithms with application examples that illustrate the range of possible applications. Representative cases are analyzed in full detail to illustrate the applicability of the analysis techniques, while in other cases the results are described without proofs and references to the literature are provided.
References


References

References
References


References


References


References


References


References


References


