International Journal of Computer and Communication Technology

Volume 6 | Issue 2 Article 15

April 2015

PERFORMANCE OPTIMIZATION IN NETWORKS SERVING HETEROGENEOUS FLOWS

B. VENKATESWARLU
St.Anns College of Engineering & Technology, Chirala, AP, India, bvenkatesh18@gmail.com

AVS SUDHAKARA RAO
St.Anns College of Engineering & Technology, Chirala, AP, India, ande_sudhakar@yahoo.co.in

P. HARINI
St.Anns College of Engineering & Technology, Chirala, AP, India, hpogadadanda@yahoo.co.in

PERFORMANCE OPTIMIZATION IN NETWORKS SERVING HETEROGENEOUS FLOWS

[1]B. VENKATESWARLU, [2]AVS SUDHAKARA RAO & [3]P. HARINI

[1,2,3]St.Anns College of Engineering & Technology, Chirala, AP, India
Email: bvenkatesh18@gmail.com, ande_sudhakar@yahoo.co.in, hpogadadanda@yahoo.co.in

Abstract: Channel-aware scheduling strategies have emerged as an effective mechanism for improving the throughput of wireless data users by exploiting rate variations. The improvement in throughput comes however at the expense of an increase in the variability of the service rate received over time. While the larger variability only has a limited impact on delay-tolerant data transfers, it does severely affect delay-sensitive applications. In order to examine the merits of channel-aware scheduling for the latter users, we consider a wireless system supporting a combination of streaming and elastic traffic. We first examine a scenario with rate-adaptive streaming traffic, and analyze the flow-level performance in terms of transfer delays and user throughputs for various canonical resource sharing schemes. Simulation experiments demonstrate that the analytical results yield remarkably accurate estimates, and indicate that channelaware scheduling achieves significant performance gains. Next we investigate a scenario where the streaming sources have an intrinsic rate profile and stringent delay requirements. In that case, channel-aware scheduling yields only modest performance gains, and may even be harmful.

Keywords: *Cross-layer optimization, dynamic load balancing, flow rate control, heterogeneous traffic, routing, scheduling, utility maximization*

1. INTRODUCTION

We have witnessed the development of increasingly sophisticated optimization and control techniques to address a variety of resource allocation problems for communication networks. Much of this investigation has focused primarily on optimizing functions of long-term performance metrics such as throughput subject to network stability. Two types of traffic can be distinguished: elastic traffic with controllable packet injection rates generated by file transfer or other delay to applications, and inelastic traffic with fixed packet injection rates generated by delay-sensitive applications. Much of the existing work focuses on the existence of either the inelastic traffic alone. The integration of elastic and inelastic flows in single-hop wireless systems has been studied in and has been extended to a multiple-hop network in, however with the restriction of every flow having a single route. In, the coexistence of inelastic and elastic flows has also been considered in a more general setup.

However, previous utility maximization-based solutions do not distinguish inelastic packets and elastic packets at the packet level. Thus, the inelastic packets need to compete with elastic packets for link bandwidths, so these two types of flows have comparable delay performance. Yet, inelastic flows model delay-sensitive traffic and must be served with higher priority as they traverse the network. Our framework differs from earlier utility maximization-based approaches in that we give strictly higher service priority to inelastic packets, i.e., at every link, elastic packets can be transmitted only when there are no inelastic packets waiting for service. This prioritization decouples the inelastic packets and elastic packets at the *link (or, equivalently, packet) level* and will result in small delays for inelastic flows. Note that even though two types of flows are decoupled at the link level, to provide high utilization for elastic traffic, the inelastic flows must smartly distribute their load among their available routes. To that end, we developed our algorithm to maximize the network utility defined by elastic flows under the prioritization, which provides new coupling methods at the *flow level* that are different from previous utility maximization-based solutions.

We believe the main contributions of this work to be the following.

• The mathematical formulation of the utility maximization problem for elastic rate control subject to inelastic traffic requirements of fixed rate and service prioritization.

• The development of a distributed joint load-balancing and rate-control algorithm that gives strict service priority to inelastic packets while guaranteeing optimal resource utilization for elastic traffic. The description and the optimality of this algorithm are provided for both the fluid model and the actual stochastic network.

• The extension of the base algorithm to a virtual queue based operation that enables further delay reduction for both traffic types with a nominal and controllable sacrifice in the network utilization.

• The relaxation of the static route assumption for the elastic flows to achieve higher utilization of the network resources through dynamic, multipath routing while maintaining the prioritization requirements. This leads to a novel two-stage queuing architecture that complies with the prioritization requirements of the design.

2. RELATED WORK

We study a model of controlled queueing network, which operates and makes control decisions in discrete time t = 0, 1, 2, ... The key feature of the model is that each network control action has two effects. First, network has a finite set N p of “traffic” processing nodes, with queues, and each control action has associated “queueing control” which affects traffic (customer) arrival rates to processing nodes, their processing (service) rates, routing between processing nodes, etc. Second, network generates a finite number of “commodity (utility) flows,” forming steno; namely, each control action k generates amounts ban(k), n ∈ Nu, of the commodities. In addition, the available set of control options depends on some underlying random network “mode,” modeled by an erotic Markov chain.

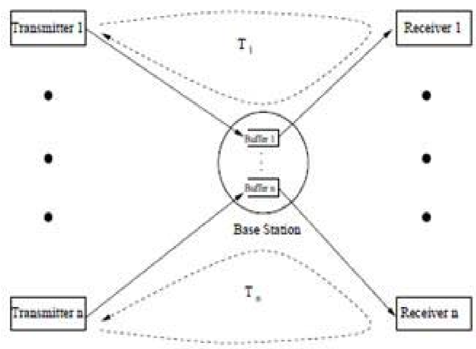


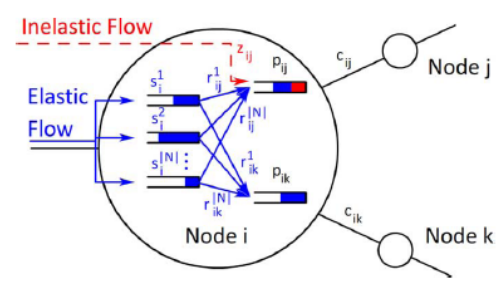
Figure 1: Topology of the Network with Feedback Delays

Let X = (Xn, n ∈ Nu), be the vector of long-term average rates at which commodities are generated under a given control strategy. We seek to find a dynamic control strategy which maximizes a concave utility function H(X) of average commodity rates, subject to the constraint that the network remains stable, that is, roughly speaking, queues at the processing nodes remain bounded. (The utility function H need not be strictly concave.) In this paper we introduce a dynamic control algorithm, called Greedy Primal- Dual (GPD) algorithm, which solves the above problem (asymptotically, as described shortly), under the natural assumption that, roughly speaking, it is feasible at all to keep the network stable (even if we ignore utility optimization). As we will see, the algorithm is very parsimonious, and naturally decomposes to become a decentralized algorithm in special cases when different network elements can make "their own" control decisions independently. Since both commodity generation and queuing control actions depend on a chosen network control action, our model accommodates, in general, scenarios in which "currently available" choices of commodity generation rates, traffic arrival and service rates, and routing are mutually interdependent in arbitrary way. As we will see, this feature is very useful in modeling many systems arising in applications. In applications, different commodity types may have different meanings, and some of the commodities may be "physical" and some "virtual." In telecommunication systems a commodity may be a traffic flow, which may (or may not) need to go through and be processed by the processing network (this is modeled by "coupling" the generated commodity amounts and amounts of arrived traffic to some nodes). A commodity may also correspond to a monetary award (or penalty), associated with a control action. Or, going back to telecommunications and in particular wireless systems, a commodity may be energy or power consumed by a control action. Thus, a commodity may be virtual in the sense that it serves simply to keep track of and optimize certain performance measures. For example, the GPD algorithm can be used to control a queuing network at the lowest average cost (or power consumption), while keeping queues stable. We will demonstrate there that the GPD algorithm can create virtual processing nodes as well; for example, to solve the stated optimization problem, subject to additional desired constraints on the average commodity generation rates. To summarize, our abstract view that network control actions have double effect of controlling queues on one hand, and generating some commodities on the other hand, allows the model to accommodate a large variety of applications and scenarios.

In this work, we consider the optimal control of networks that serve heterogeneous traffic types with diverse demands, namely inelastic and elastic traffic. We formulated a new network optimization problem, proposed novel queuing architecture, and developed a distributed load-balancing and congestion control algorithm with provably optimal performance. We also provided an important improvement to our joint algorithm to achieve better delay performance by introducing new design parameters together with a set of virtual queues. We have also extended our algorithm to the case of allowing elastic flows to choose their routes dynamically, which will further utilize the resource available in the network. Future research of this topic includes the following.

1. One future direction is to extend our results to multihop wireless with fading channels and interference and develop joint load-balancing/congestion control/routing/scheduling algorithms.
2. We considered a time-slotted system and assumed that the network is perfectly synchronized. The impact of possible synchronism on the algorithm performance needs to be studied.
3. We adopted a link-centric formulation, which assumes instantaneous arrivals of the packets at all the links on their routes. An alternative is to consider a node-centric formulation, where packets are sequentially transferred, and a source only requires the information of the queues at the source.
4. So far, we have focused on the stability and long-term guarantees for the traffic types. We aim to investigate oscillatory behavior and delay characteristics in our future work.
5. In this work, we assume the routes and the supportability of the inelastic flow are given.

The system architecture is as shown in Fig.2.



Queueing architecture for dynamic routing.

Figure 2: System Architecture

3. ALGORITHM

In our discussion, we will abbreviate the aggregate elastic and inelastic rates $y_l(x_e)$, and $z_l(x_i)$, with y_l, and z_l for brevity. We note that condition (4) aims to capture the network stability condition in the fluid model by guaranteeing that the total load on a link is below the link capacity, and condition (5) guarantees that inelastic flows receive enough bandwidth to satisfy its rate demands. Thus, the optimization problem is to maximize the sum of utilities of elastic flows when guaranteeing that inelastic flows are supported.

$$\max_{\mathbf{x}\geq 0} \sum_{f_e \in \mathcal{F}_e} U_e(x_e)$$

$$s.t. \quad y_l(\mathbf{x}_e) + z_l(\mathbf{x}_i) \leq c_l \qquad \forall\, l \in \mathcal{L}$$

$$\sum_{r=1}^{|\mathcal{R}_i|} x_i^{(r)} = a_i \qquad \forall\, f_i \in \mathcal{F}_i.$$

Joint Congestion Control and Load-Balancing Algorithm for the FNO Problem:

- Queue evolution for link l

$$\dot{p}_l(t) := \frac{dp(t)}{dt} = (z_l(t) + y_l(t) - c_l)^+_{p_l(t)}$$

where $(v(t))^+_{p(t)}$ is zero if $v(t) < 0$ and $p(t) = 0$; and $v(t)$ otherwise.

- Congestion controller for elastic flow f_e

$$x_e(t) = U_e'^{-1}(q_{\mathbf{R}_e}(t)).$$

- Load balancing implemented for inelastic flow f_i

$$\dot{x}_i^{(r)}(t) = \left(\bar{q}_i(t) - q_{\mathbf{R}_i^{(r)}}(t)\right)^+_{x_i^{(r)}(t)}$$

where $\bar{q}_i(t)$ satisfies

$$\sum_{r=1}^{|\mathcal{R}_i|} \left(\bar{q}_i(t) - q_{\mathbf{R}_i^{(r)}}(t)\right)^+_{x_i^{(r)}(t)} = 0$$

and $\sum_{r=1}^{|\mathcal{R}_i|} x_i^{(r)}(0) = a_i$.

The intuition behind the load-balancing algorithm described above is to shift the inelastic flows to less heavily loaded routes to allow for the maximum network utilization for elastic flows. In the algorithm, a source needs all the queue information along its route. However, as we mentioned in Section II, we can send queue information hop by hop and still achieve stability even if this information is delayed. Thus, this algorithm can be implemented fully distributed.

Next, we will show the stability and optimality of our joint congestion control and load-balancing algorithm.

Joint Congestion Control and Load-Balancing Algorith for the SNO-K Problem:

- Scheduling with Strict Prioritization:
  For each link $l \in \mathcal{L}$, we serve c_l packets from p_l, witl strict priority to inelastic packets, which leads to the following queue evolution:

$$p_l[t+1] = (p_l[t] + z_l[t] + y_l[t] - c_l)^+.$$

- Congestion controller for elastic flow f_e

$$x_e[t] = \min\left\{M, U_e'^{-1}\left(\frac{1}{K} q_{\mathbf{R}_e}[t]\right)\right\}$$

  where M is a positive constant satisfies $M > 2\max_{l\in\mathcal{L}}\{c_l\}$.

- Load balancing implemented for inelastic flow f_i

$$\triangle x_i^{(r)}[t] = \left(\bar{q}_i[t] - q_{\mathbf{R}_i^{(r)}}[t]\right)^+_{x_i^{(r)}[t+1]}$$

  or equivalently

$$x_i^{(r)}[t+1] = \left(x_i^{(r)}[t] + \bar{q}_i[t] - q_{\mathbf{R}_i^{(r)}}[t]\right)^+$$

  where $\bar{q}_i[t]$ satisfies

$$\sum_{r=1}^{|\mathcal{R}_i|} \left(\bar{q}_i[t] - q_{\mathbf{R}_i^{(r)}}[t]\right)^+_{x_i^{(r)}[t+1]} = A_i[t+1] - A_i[t]$$

  and $\sum_{r=1}^{|\mathcal{R}_i|} x_i^{(r)}[0] = A_i[0]$.

4. SIMULATION RESULTS

Individual responsible for testing may prefer to select their own technique and tool based on the test situation. For selecting the appropriate testing process the project should be analyzed with the following three testing concepts:

1. Structural versus functional testing
2. Dynamic versus static testing
3. Manual versus automatic testing

After analyzing through the above testing concepts we divided to test our project in Waterfall model testing methodology. Structural analysis based test sets are tend to uncover errors that occur during coding of the program. The properties of the test set are to reflect the internal structure of the program. Structural testing is designed to verify that the developed system and programs work as specified in the requirement. The objective is to ensure that the product is designed structurally sound and will function correctly. Functional testing ensures that the requirements are properly satisfied by the application system. The functions are those tasks that the system is designed to accomplish. This is not concerned with how processing occurs but rather with the results of the processing. The functional analysis based test sets tend to uncover errors that occurred in implementing requirements or design specifications. After selecting the appropriate testing methodology we have to select the necessary testing technique such as stress testing, execution testing, recovery testing, operation testing, compliance testing and security testing. We are performing operation testing

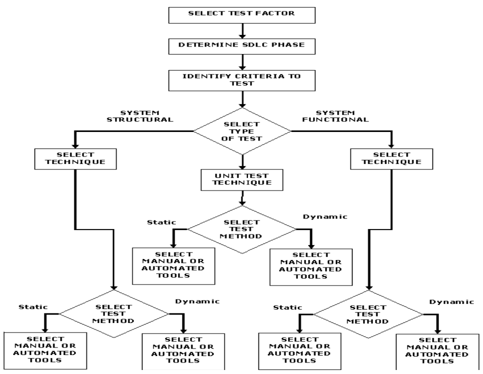


Figure 3: Testing technique and tool selection process

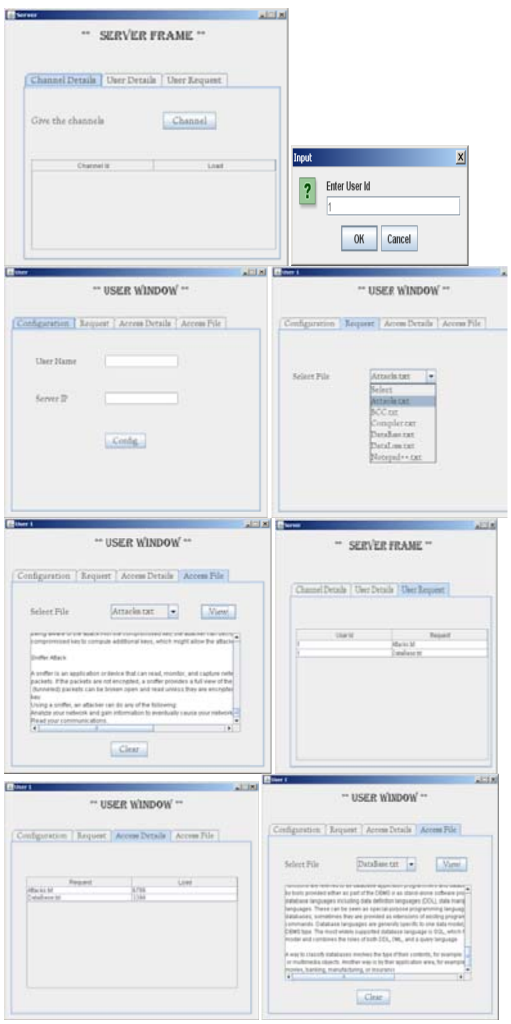


Figure 4:Screen Shots

5. CONCLUSION

We consider the optimal control of networks that serve heterogeneous traffic types with diverse demands, namely inelastic and elastic traffic. We formulated a new network optimization problem, proposed a novel queuing architecture, and developed a distributed load-balancing and congestion control algorithm with provably optimal performance. We also provided an important improvement to our joint algorithm to achieve better delay performance by introducing new design parameters together with a set of virtual queues. We have also extended our algorithm to the case of allowing elastic flows to choose their routes dynamically, which will further utilize the resource available in the network.

One future direction is to extend our results to multi-hop wireless networks with fading channels and interference and develop joint load-balancing/congestion control/routing/scheduling algorithms. Here, we considered a time-slotted system and assumed that the network is perfectly synchronized. The impact of possible synchronism on the algorithm performance needs to be studied. We adopted a link-centric formulation, which assumes instantaneous arrivals of the packets at all the links on their routes. An alternative is to consider a node-centric formulation, where packets are sequentially transferred, and a source only requires the information of the queues at the source. So far, we have focused on the stability and long-term guarantees for the traffic types. We aim to investigate oscillatory behavior and delay characteristics in our futurework. In this work, we assume the routes and the supportability of the inelastic flow are given. Developing corresponding routing and admission control mechanism will make it complete.

REFERENCES:

1. Eryilmaz and R. Srikant, “Resource allocation of multi-hop wireless networks,” in Proc. Int. Zurich Seminar Commun., Feb. 2006, pp. 90–93.
2. E. Altman, T. Bas¸ar, and R. Srikant, “Congestion control as a stochastic control problem with actiondelays,” Automatica, vol. 35, pp. 1937–1950, 1999.
3. A. Eryilmaz and R. Srikant, “Fair resource allocation in wireless networks using queue-length based scheduling and congestion control,” in Proc. IEEE INFOCOM, Miami, FL, Mar. 2005, vol. 3, pp. 1794–1803.
4. L. Bui, A. Eryilmaz, R. Srikant, and X. Wu, “Joint asynchronous congestion control and distributed scheduling for wireless networks,” in Proc. IEEE INFOCOM, 2006.
5. S. Borst and N. Hegde, “Integration of streaming and elastic traffic in wireless networks,” in Proc. IEEE INFOCOM, May 2007, pp. 1884–1892.
6. A. Eryilmaz, R. Srikant, and J. R. Perkins, “Stable scheduling policies for fading wireless channels,” IEEE/ACM Trans. Networking, vol. 13, no. 2, pp. 411–425, Apr. 2005.
7. S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge, U.K.: Cambridge Univ. Press, 2004.
8. L. Georgiadis, M. J. Neely, and L. Tassiulas, Resource Allocation and Cross-Layer Control in Wireless Networks, ser. Foundations and Trends in Networking. Hanover, MA: now, 2006.
9. J. Bolot and A. Shankar, “Dynamic behavior of rate-based flow control mechanisms,” ACM Comput. Commun. Rev., vol. 20, no. 2, pp. 35–49, 1992.
10. P. Hande, S. Zhang, and M. Chiang, “Distributed rate allocation for inelastic flows,” IEEE/ACM Trans. Netw., vol. 15, no. 6, pp. 1240–1253, Dec. 2007.