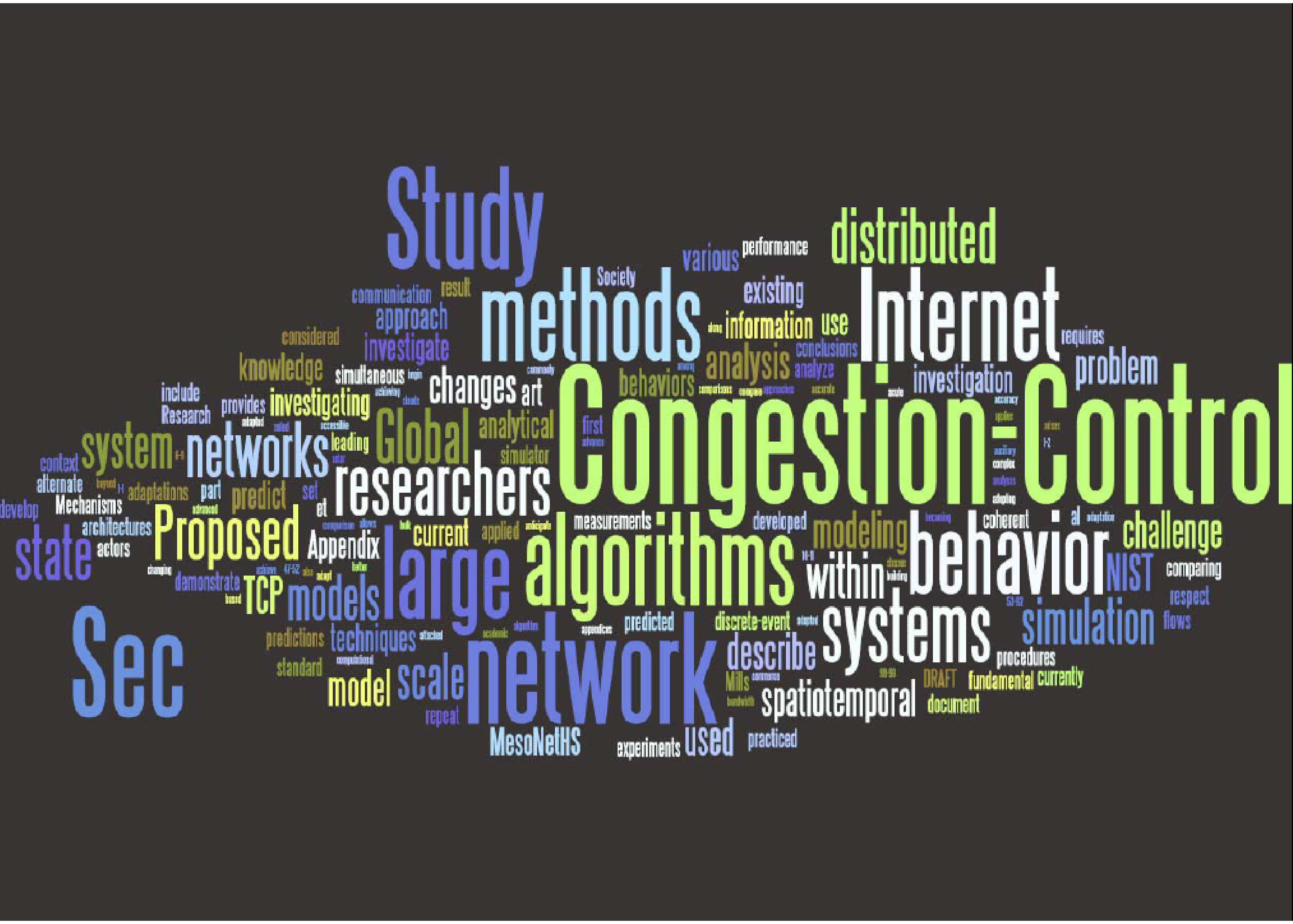
# Chapter 1 – Introduction



# 1 Introduction

Society is becoming increasingly reliant on large networked information systems for commerce, communication, education, entertainment and government. “[Despite] society’s profound dependence on networks, fundamental knowledge about them is primitive. [Global] communication…networks have quite advanced technological implementations but their behavior under stress still cannot be predicted reliably…. *There is no science today that offers the fundamental knowledge necessary to design large complex networks [so] that their behaviors can be predicted prior to building them.*” [104] This lack of knowledge grows more acute as society moves toward service-oriented architectures [102-103] that deploy software, platforms and infrastructure as distributed services accessible through networks.

Why are large distributed systems so difficult to predict? Such systems exhibit global behavior that arises from independent decisions made by many simultaneous actors, which adapt their behavior based on local measurements of system state. As a result of actor adaptations, global system behavior may change, influencing subsequent measurements, and leading to further adaptations. This continuous cycle of measurement and adaptation produces a time-varying global behavior that drives the performance experienced by individual actors within spatiotemporal regions of a large distributed system. Thus, to truly understand and predict behaviors in such systems requires techniques to model and analyze designs at large scale. Such techniques are currently beyond the state of the art, as practiced by network researchers.

As part of a team of researchers [105] at NIST, we are investigating methods to model and analyze distributed systems, such as the Internet, computational grids, service-oriented architectures and computing clouds. As part of this investigation, the study reported here develops, applies and evaluates a coherent set of modeling and analysis methods for distributed information systems of large spatiotemporal scale. The methods are adapted from techniques often applied by NIST scientists to study physical systems.

In this study, we develop methods to investigate global system behavior within the context of a challenge problem: comparing some proposed changes to the standard congestion control algorithm [9-10] for the Internet. Congestion control procedures are implemented as part of the transmission control protocol (TCP) that operates within every computer attached to the global Internet. Numerous researchers [46-51] have forecast changes in relationships among bandwidth and propagation delay as the speed of network links increases. These researchers predict that the current version of TCP will prove inadequate, leading to substantial underutilization in network resources and preventing end users from achieving high transfer rates. Such predictions have stimulated researchers to propose alternate congestion control algorithms [52-61] intended to achieve higher network utilization and better user performance. Evaluating the implications of adopting proposed changes to TCP congestion control procedures requires investigating global behaviors that result when such changes are deployed on a large scale throughout an Internet-like network. The current study provides such an investigation.

We begin (in Chapter 2) with a discussion of the challenge problem and the current state of the art with respect to investigating proposed Internet congestion control algorithms. We outline various approaches that we considered for modeling and analysis and we describe the approach we selected. We introduce five hard problems we needed to solve in order to implement our approach and we discuss the solutions we adopted. In Chapter 3, we describe MesoNet, a medium scale, discrete-event simulation model that we created for use in this study. MesoNet allowed us to expose candidate congestion control algorithms to a wide variety of network conditions. We subjected MesoNet to sensitivity analyses, as documented in Chapter 4 and in Appendix C. These sensitivity analyses helped us to gain confidence that MesoNet provides a suitable model for TCP networks, and also enabled us to identify the most important parameters influencing MesoNet behavior. As part of our sensitivity analyses, we employ a NIST-developed, 10-step graphical analysis process, which is described in Appendix D. In Chapter 5, we explain our models for various congestion control algorithms and we document key empirical comparisons used to verify model correctness. The bulk of the study consists of six experiments, which we describe in Chapters 6 through 9. As we discuss in Chapter 2, these experiments were not constructed as an integral campaign, but rather arose through a process of iterative refinement, where findings from previous experiments suggested useful directions for subsequent experiments. We first compare (Chapter 6) congestion control regimes in a large, fast network simulation and then repeat the comparison (Chapter 7) in a simulated network with smaller size and slower speeds. In Chapter 8, we enlarge the traffic classes considered, while comparing the congestion control algorithms in a network where some flows use standard TCP and some use alternate algorithms. In Chapter 9, we repeat an experiment from Chapter 8 but in a larger, faster simulated network, where theorists suggest alternate congestion control algorithms could provide best advantage. Taken together, these experiments compare the behavior of seven congestion control algorithms under a wide range of simulated conditions. We generate sufficient information to draw some conclusions in Chapter 10 about the congestion control algorithms. Chapter 10 also provides an evaluation of the methods that we developed and applied. We include some appendices to document auxiliary investigation of analytical (Appendix A) and hybrid (Appendix B) models of TCP networks.

This study may interest two different audiences: (1) those seeking to understand and evaluate methods to model and analyze behavior in large, distributed information systems and (2) those aiming to compare proposed changes in algorithms for the Internet. Readers in the first audience can expect to learn about various modeling, experiment design and statistical analysis techniques applied to study dynamics in complex systems. In addition, such readers may benefit from our findings with regard to the strengths and weaknesses of the techniques we applied. Readers in the second audience can expect to learn how to model a data communications network with a manageable set of parameters. In addition, such readers may benefit from learning how we let measurement data (rather than preconceived metrics) drive our comparison of alternative congestion control algorithms. Mindful of these two different audiences, we attempt to provide a sufficient level of explanation to engage every reader. We explain our modeling and analysis methods in detail so that networking experts can follow our methods. And we provide sufficient tutorial information to allow those readers who are not networking experts to follow our challenge problem and related findings. Where appropriate, we also provide references to additional sources where readers in each audience can pursue more information.

We can summarize the contributions of this study along several lines. First, we define and demonstrate a coherent set of modeling and analysis methods that can be used to investigate behavior in distributed systems of large spatiotemporal scale. The methods we develop represent an advance in the state of the art, as currently practiced by network researchers. Second, we evaluate our modeling and analysis methods in the context of a challenge problem that investigates behavior of various proposed Internet congestion control algorithms. The challenge problem is of current interest to industrial and academic researchers within the Internet Congestion Control Research Group (ICCRG) of the Internet Research Task Force (IRTF). Third, we provide conclusions and recommendations with respect to the congestion control algorithms that we study. We demonstrate that our methods lead to insights that have not been obtained using existing methods. Fourth, we describe a medium-scale, discrete-event network simulator that we developed for our study. The simulator, called MesoNet, can be efficiently parameterized and allows feasible simulation of high-speed networks transporting hundreds of thousands of simultaneous flows. The most commonly used network simulators are incapable of supporting such large-scale models. Fifth, we suggest an approach that might improve the accuracy of existing analytical models for Internet congestion control algorithms. We anticipate future work to include improved analytical models within existing fluid-flow simulation frameworks in an effort to obtain accurate predictions regarding spatiotemporal behavior in large networks.