Chapter 4

Conclusion

4.1 Summary

During the course of the last three years, we have been applying a general methodology to different problems encountered in distinct types of applications. This general methodology consists first of proposing mathematical models for the system at hand. Then, metrics of interest or estimators are derived under the assumptions of the proposed models. Finally, the analytical results are compared with simulated and/or experimental results.

In the first chapter of the thesis, we have proposed two single server finite buffer queueing models for a path between two hosts. The single queue is assumed to model the bottleneck link along the path. Under Poisson arrivals and either exponential or deterministic service times, we express the loss probability, the server utilization and the expected response time in terms of the intensity of the background traffic entering the queue and the buffer size at the bottleneck link. Note that it is the first time that the analysis of the $M/D/1/K$ queue has been pushed so far. Conditional probabilities are further derived for the $M{+}M/M/1/K$ queue model. Taking any pair of these equations and solving for the unknown cross traffic intensity and the unknown buffer size, returns estimates of these unknown parameters, provided that some measurements of the metrics at hand are available. The pairs of estimators – obtained using this procedure over all possible pairs of equations – are then tested through simulations. Pairs of estimators are discarded while others have proved to perform well, which is the case of the pair of estimators relying on measurements of the loss probability and the response time.

In the second chapter of the thesis, we have proposed successively three models for multicast groups. The first model, an $M/M/\infty$ queue under heavy traffic, enables the use of Kalman filtering theory to derive an efficient membership estimator, which is optimal for the $M/M/\infty$ queue under heavy traffic. The second model, an $M/M/\infty$ queue under a general traffic regime, lead to the same estimator, but this time the Wiener filter theory was used. Note that the later theory can be applied to the $M/G/\infty$ queue as long as the distribution of receivers lifetime is not heavy-tailed. However, the canonical factorization of the power spectrum of the measurements process is possible only for the $M/M/\infty$ queue. The third and last model studied, an $M/H_L/\infty$ queue where H_L denotes an L-stage hyperexponential distribution, allows the derivation of an efficient estimator which is optimal among all estimators satisfying a first-order linear auto-regressive equation. To summarize what precedes, we have proposed three models, have followed three distinct approaches and obtained *two* distinct membership estimators. These estimators have been tested on simulations driven by real traces. Their overall performance are almost identical, however the estimator deriving from the $M/M/\infty$ queue model shows a good ability to track high variations in the membership, the other estimator being more on target in flat periods.

In the third chapter, we have proposed Markovian models for agent location mechanisms in a mobile code environment. To the best of our knowledge, it is the first time that such mechanisms are modeled and their performance are formally evaluated. We derive expressions for the expected response time of each mechanism in each model and further express the expected number of forwarders in the decentralized mechanism. The theoretical results are compared to both simulated and experimental results (experiments conducted over both a LAN and a MAN). It appeared in our experiments that forwarders achieve better performance over a MAN and that the centralized technique is preferable in a LAN. The relatively close match between the latter results validates both Markovian models and allows the use of the theoretical formulas as predictors of the performance of the mechanisms. We have investigated the performance of the mechanisms under a wide variety of conditions using the difference between the expected response times of each mechanism. This theoretical comparison has illustrated several unexpected effects when some parameter is changed and revealed that no mechanism is uniformly better than the other.

4.2 Perspectives

Concerning the first part of the thesis (Chapter 1), we have seen that the scheme $P_L\_U\_R$ does not perform especially well in simultaneously estimating the bottleneck bandwidth μ, the cross traffic intensity at the bottleneck link λ and the buffer size at the bottleneck link K. We should therefore use an estimate of μ (provided by *pathrate*, PBM or ROPP for instance) and then use scheme P_L_R to estimate λ and K. The impact of using $\hat{\mu}$ instead of using μ still needs to be investigated. Beside that, the $M{+}M/M/1/K$ queue model may be extended to account for the propagation delay. The latter could be modeled by a normally distributed random variable or an exponentially distributed random variable.

Concerning the estimation of the membership of multicast groups, we plan to work on more realistic models, such as the $M/W/\infty$ queue or the $M/L/\infty$ queue. The first step will be to compute (2.39) for these queues. After that, one may follow the approach of Section 2.6.2 to find the optimal first-order filter, using either (2.6) or (2.54). We actually do not know which auto-regressive equation is preferable. Future research will certainly address this issue. Another point which will also need to be addressed concerns the choice of the ACK probability and the ACK interval for the considered model. Instead of limiting the amount of feedback generated each round, we might limit the amount of feedback generated each I seconds, as already suggested in Remark 2.7.1.

Concerning the research on agent location mechanisms, we plan to revisit the model for the centralized mechanism. Instead of the single queue server investigated in Section 3.4.2, we will consider a multiqueue single server with vacation. Another possibility consists of studying a multiqueue server with feedback, in which a query that has already been served can potentially be replaced in the queue. Besides, we are interested in modeling other agent location mechanisms, especially a mixed mechanism recently implemented in the *ProActive* library. In this mixed approach, both forwarders and a location server are used to ensure communications. Forwarders have a fixed lifetime and mobile objects periodically update their locations at the server. When a message finds no object (forwarder or mobile agent) on a given site, the source queries the location server to have the latest known location of the mobile object. The engineering questions which arise in such mechanisms are "what is the ideal lifetime for forwarders?" and "how often should the mobile object inform the location server of its location?" Future research aim at answering these questions.