A Refined Mean Field Approximation

Nicolas Gast, Benny Houdt

A Refined Mean Field Approximation

Nicolas Gast
Inria
Grenoble, France
nicolas.gast@inria.fr

Benny Van Houdt
University of Antwerp
Antwerp, Belgium
benny.vanhoudt@uantwerpen.be

EXTENDED ABSTRACT

Stochastic models have been used to assess the performance of computer (and other) systems for many decades. As a direct analysis of large and complex stochastic models is often prohibitive, approximations methods to study their behavior have been devised. One very popular approximation method relies on mean field theory. Its widespread use can be explained by the relative ease involved to define and solve a mean field model in combination with its high accuracy for large systems.

Although mean field theory provides no guarantees with respect to the accuracy of a finite system of size N, there are many examples where the accuracy was demonstrated (using simulation) to be quite high even for systems of moderate size, e.g., for $N \approx 50$. Nevertheless this accuracy very much depends on the exact parameter settings, for instance mean field models for load balancing systems are known to be quite inaccurate under high loads even for moderate sized systems.

The main contribution of this paper is to establish a result that applies to a broad class of mean field models, including density dependent population processes [3] and the class of discrete-time models of [1], and that provides a significantly more accurate approximation for finite N than the classical mean field approximation. When considering a mean field model $X^{(N)}$ described by the density-dependent population process of Kurtz, where $X_i^{(N)}(t)$ is the fraction of objects in state i at time t, the classical mean field approximation shows that if the corresponding ODE model has a unique attractor π, then the stationary measure of $X^{(N)}$ concentrates on π. We establish conditions that show that for a function h, there exists a constant V_h such that

$$\mathbf{E}^{(N)}\left[h(X^{(N)})\right] = h(\pi) + \frac{V_h}{N} + o\left(\frac{1}{N}\right).$$

The full version of the paper is available in [2].

The constant V_h is expressed as a function of the Jacobian and the Hessian matrices of the drift (in π) and the solution of a single Lyapunov equation. As such it can be computed efficiently even when the number of possible states i of an object is large (less than one second for a 500-dimensional model).

The refined mean field model that we propose consists in approximating $\mathbf{E}^{(N)}[h(X^{(N)})]$ by $h(\pi) + V_h/N$ and maintains many of the attractive features of the classic mean field approximation that resulted in its widespread use, but significantly improves upon its accuracy for finite N. We demonstrate that this refined approximation significantly improves the accuracy of the mean field approximation by using a variety of mean field models that include the coupon replication model [4] and a number of load balancing schemes [5–7]. In some special cases the constant V_h can be expressed in closed form, while for remaining cases the computation time for a numerical evaluation of V_h is negligible.

To give a flavor of the accuracy of the refined approximation, we provide some examples in Table 1 picked from the full version of the paper [2]. This table illustrates that the refined approximation is very accurate, even for $N = 10$. Moreover, it is much more accurate than the classic mean field approximation.

	Coupon	Supermarket	Pull/push
Simulation ($N = 10$)	1.530	2.804	2.304
Refined mean field ($N = 10$)	1.517	2.751	2.295
Mean field ($N = \infty$)	1.250	2.353	1.636

Table 1: Illustration of the refined approximation accuracy. The columns correspond to the mean number of users with coupon i or the average queue length for the supermarket model or the pull/push model ($\rho = 0.9$).

REFERENCES

[1] Michel Benaim and Jean-Yves Le Boudec. 2008. A class of mean field interaction models for computer and communication systems. *Performance Evaluation* 65, 11 (2008), 823–838.

[2] Nicolas Gast and Benny Van Houdt. 2017. A Refined Mean Field Approximation. *Proc. ACM Meas. Anal. Comput. Syst.* 1, 2, Article 33 (Dec. 2017), 28 pages. https://doi.org/10.1145/3154491

[3] Thomas G Kurtz. 1970. Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes. *Journal of Applied Probability* 7 (1970), 49–58.

[4] L. Massoulié and M. Vojnović. 2005. Coupon Replication Systems. *SIGMETRICS Perform. Eval. Rev.* 33, 1 (June 2005), 2–13. https://doi.org/10.1145/1071690.1064215

[5] W. Minnebo and B. Van Houdt. 2014. A Fair Comparison of Pull and Push Strategies in Large Distributed Networks. *IEEE/ACM Transactions on Networking* 22 (2014), 996–1006. Issue 3.

[6] Michael Mitzenmacher. 2001. The power of two choices in randomized load balancing. *IEEE Transactions on Parallel and Distributed Systems* 12, 10 (2001), 1094–1104.

[7] Nikita Dmitrievna Vvedenskaya, Roland L'vovich Dobrushin, and Fridrikh Izrailevich Karpelevich. 1996. Queueing system with selection of the shortest of two queues: An asymptotic approach. *Problemy Peredachi Informatsii* 32, 1 (1996), 20–34.