Exact Computation and Bounds for the Coupling Time in Queueing Systems

Extended Abstract

Sebastian Samain and Ana Bušić
Inria Paris DI ENS, École normale supérieure, CNRS, PSL Research University, Paris
Paris, France
sebastien.samain@inria.fr

ABSTRACT

This paper is a work in progress on the exact computation and bounds of the expected coupling time for finite-state Markov chains. We give an exact formula in terms of generating series. We show how this may help to bound the expected coupling time for queueing networks.

CCS CONCEPTS

• **Mathematics of computing → Generating functions**; *Queueing theory*; • **Theory of computation → Regular languages**;

KEYWORDS

Perfect Simulation, Algorithmic Complexity, Queueing Systems, Automata

ACM Reference Format:

Sebastian Samain and Ana Bušić. 2017. Exact Computation and Bounds for the Coupling Time in Queueing Systems: Extended Abstract. In *Proceedings of 11th EAI International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS 2017)*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3150928.3150965

1 INTRODUCTION

Propp and Wilson [8] used a coupling from the past scheme to derive a simulation algorithm, providing unbiased samples of the stationary distribution of a finite-state irreducible Markov chain in finite expected time. The main idea is to start trajectories from all initial states $x \in \mathcal{X}$ at some time in the past until time $t = 0$. If the end state is the same for all trajectories, then the chain has coupled and the end state has the stationary distribution of the Markov chain. Otherwise, the simulations are started from further in the past. The analysis of the complexity of the algorithm requires analysis of the complexity of one iteration and analysis of the number of iterations. Until now, most of the efforts were focused on the the complexity of one iteration (see [5, 8] for the monotone and ani-monotone case and [1–3, 6, 7] for the non-monotone case).

The focus of his paper is on the expected number of iterations, closely related to the coupling time of the Markov chain, i.e. the time needed for the coalescence of trajectories started from all initial states. We provide a systematic framework to compute the expectation of the coupling time. As this may be very challenging in the general case, our second goal is to compute upper and lower bounds on this expectation. The results obtained are applied to queueing networks.

VALUETOOLS 2017, December 5–7, 2017, Venice, Italy

ACM ISBN 978-1-4503-6346-4/17/12.
https://doi.org/10.1145/3150928.3150965

2 EXACT COMPUTATION OF THE EXPECTED COUPLING TIME

We consider an irreducible aperiodic Markov chain over a finite state space $\mathcal{X}$. It is known in this case that there exists a stochastic (or Markov) complete automaton S that simulates the Markov chain started at any state that has the same transition kernel and that couples almost surely in finite time. We can represent the transition rule of this automaton by a semi-group action $\triangleright$ over some alphabet $\mathcal{A}$. Once we have build such simulation, we can build a second stochastic automaton $\mathcal{A}ut_S$, over $\mathcal{P}(\mathcal{X})$, which associates the set $U = \{t \triangleright a | t \in T\}$ with $T \in \mathcal{P}(\mathcal{X})$ and $a \in \mathcal{A}$ the . We thus have:

Proposition 2.1. *The coupling time for the stochastic automaton S that simulates the Markov chain with the transition kernel P is the hitting time of the set $A = \{\{x\} | x \in \mathcal{X}\}$ in $\mathcal{A}ut_S$ from the initial condition $\mathcal{X}$.*

To compute this hitting time, we proceed as follows:

(1) We express the hitting language L of A from $\mathcal{X}$, i.e. the language of words w such that:
- $\mathcal{X} \triangleright w \in A$
- $\forall u$ strict prefix of w, $\mathcal{X} \triangleright u \notin A$

We highlight the fact that L is distinct from the language recognized by $\mathcal{A}ut_S$.

(2) Let $p_1, ..., p_{|\mathcal{A}|}$ be a set of $|\mathcal{A}|$ formal variables and let φ be a bijection between this set of formal variables and $\mathcal{A}$. We then transform the language L into a generating series G in $|\mathcal{A}|$ letters such that the coefficient in front of the formal monomial $\prod_{i \in [1,|\mathcal{A}|]} p_i^{k_i}$ is equal to the number of words in L that have k_1 letters $\varphi(p_1)$, ..., $k_{|\mathcal{A}|}$ letters $\varphi(p_{|\mathcal{A}|})$. There are systematic ways to do that by expressing the language in a way that every word of L is counted only once (by finding a non-ambiguous rational expression for instance).

(3) We compute $D(G)$, where $D = \sum_{i=1}^{|\mathcal{A}|} p_i \frac{\partial}{\partial p_i}$.

(4) For every monomial $m = \prod_{i=1}^{|\mathcal{A}|} p_i^{k_i}$ in the formal variables $p_1, ..., p_{|\mathcal{A}|}$, $([m]G)m(\mathbb{P}(\varphi(p_1), ..., \mathbb{P}(\varphi(p_{|\mathcal{A}|}))$ is the probability to get a word of L, hence a coupling word for the stochastic automaton S, with k_1 letters $\varphi(p_1)$, ..., $k_{|\mathcal{A}|}$ letters $\varphi(p_{|\mathcal{A}|})$. Hence we get that $D(G)(\mathbb{P}(\varphi(p_1), ..., \mathbb{P}(\varphi(p_{|\mathcal{A}|})) = E[T_c]$, with T_c the coupling time for the stochastic automaton S.

By decomposing the sequence of states starting from $\mathcal{X}$ in $\mathcal{A}ut_S$ implied by a word w according to self-avoiding walks on this automaton, we get the following general result:

THEOREM 2.2. *Let S be a stochastic automaton, $L(x, A)$ the hitting language of $A \in \mathcal{P}(\mathcal{X})$ from $x \in \mathcal{X}$ in $\mathcal{A}ut_S$, $L(x, y)$ the hitting language of $y \in \mathcal{X}$ from $x \in \mathcal{X}$ in $\mathcal{A}ut_S$ and*

- *$A_{x,A}$ and $A_{C,x,y}$ the generating series of $L(x, A)$ and $L(x, y)$ on the automaton obtained from $\mathcal{A}ut_S$ by keeping only the states of C.*
- *$R_{E,x}$ the generating series of the language of first returns to x that go only through the states of E.*
- *$ChCon(C_x, C_{y,A})$ the set of words whose letters are maximal strongly connected components, whose first letter is C_x the strongly connected component containing x and whose last letter is $C_{y,A}$ the strongly connected component of y in the automaton obtained from $\mathcal{A}ut_S$ by deleting A.*
- *for u a word over the maximal strongly connected components of $\mathcal{A}ut_S$, $\mathfrak{P}ass(u)$ the set of words (v, w) such that there exists a sequence of states feasible in $\mathcal{A}ut_S$ that passes successively through each letter of u and such that the last state visited in u_i is v_i and the first state visited in u_{i+1} is w_i.*
- *$\gamma = \varphi^{-1}$.*
- *$\triangleright_t = \cup_{i=0}^{+\infty} \triangleright^i$.*

Then:

$$A_{x,A}(p) = \sum_{y \in A} \sum_{u \in ChCon(C_x, C_{y,A})} L_u(p)$$

$$L_u(p) = \sum_{(v,w) \in \mathfrak{P}ass(u)} \frac{1}{1 - R_{C_x,x}(p)} A_{C_x,x,\{v_0\}}(p)$$

$$\gamma\left(\sum_{a \in (v_0 \triangleright_t .)^{-1}(w_0)} a\right) S_{u,v,w}(p) \frac{1}{1 - R_{C_{y,A}, w_{|u|-1}}(p)} A_{C_{y,A}, w_{|u|-1}, \{y\}}(p)$$

$$S_{u,v,w}(p) = \left(\prod_{i=2}^{|u|-2} \frac{1}{1 - R_{u_i, w_{i-1}}(p)} A_{u_i, w_{i-1}, \{v_i\}}(p) \gamma\left(\sum_{a \in (v_i \triangleright_t .)^{-1}(w_i)} a\right)\right)$$

In the case of a single queue, a similar method has been used in [4], with a slightly simpler way of decomposing coupling words in this particular case.

3 COMPUTING BOUNDS FOR THE EXPECTED COUPLING TIME

This part is still a work in progress, but we hope to extend the technique that we developed here for independent queues for general queueing networks.

Computing the exact expectation of the coupling time is a rather complicated task. So one may prefer to get bounds to simplify the analysis of the expected number of iterations for perfect simulation.

To reach this goal, one needs to establish a criteria of comparison on generating series that allows to compare the associated conditional expectation (i.e. $\frac{D(G)(\mathbb{P}(\phi(p_1)),\ldots,\mathbb{P}(\phi(p_{|\mathcal{A}|})))}{G(\mathbb{P}(\phi(p_1)),\ldots,\mathbb{P}(\phi(p_{|\mathcal{A}|})))}$). This is done by the following binary relation:

Definition 3.1. We define the binary relation $\leq$ over the generating series in $p_1, \ldots, p_n$ by: $G \leq H$ iff $\exists M$ a set of monomials in the variables $p_1, \ldots, p_n$ such that

- $H = \sum_{m \in M} m H_{[m]}$
- $\forall m \in M, \forall (j, k)$ monomials in $p_1, \ldots, p_n$ such that $deg_{tot}(j) > deg_{tot}(k)$, $([j]H_{[m]})([k]G) \geq ([k]H_{[m]})([j]G)$

THEOREM 3.2. *Let G and H two formal series in $p_1, \ldots, p_n$ and $\mathbb{P}$ a probability distribution over $p_1, \ldots, p_n$. Then*

$$G \leq H \implies \frac{D(G)(\mathbb{P}(p_1), \ldots, \mathbb{P}(p_n))}{G(\mathbb{P}(p_1), \ldots, \mathbb{P}(p_n))} \leq \frac{D(H)(\mathbb{P}(p_1), \ldots, \mathbb{P}(p_n))}{H(\mathbb{P}(p_1), \ldots, \mathbb{P}(p_n))}$$

with $D = \sum_{i=1}^{n} p_i \frac{\partial}{\partial p_i}$ the derivation operator defined in the previous section.

THEOREM 3.3. *Let S a stochastic automaton that simulates a queueing network of k independent queues, $k \in \mathbb{N}^*$, T its coupling time, $\mathbb{P}$ a probability distribution on $\mathcal{A}$ the alphabet chosen for the S, γ a bijection from $\mathcal{A}$ to a set $\{p_i | i \in [1, |\mathcal{A}|]\}$ of formal variables and G the generating series of the coupling words for S restricted to states such that $k-1$ fixed queues are empty (i.e. it is the generating series of the coupling words for one queue such that the underlying stochastic simulation coincides with S). Then*

$$\min_{j \in [1,k]} \sum_{i=1}^{k} \frac{D(G)(p_{2i-1}, p_{2i})(\mathbb{P}_\gamma(p_1), \ldots, \mathbb{P}_\gamma(p_{|\mathcal{A}|}))}{G(p_{2i-1}, p_{2i})(\mathbb{P}_\gamma(p_1), \ldots, \mathbb{P}_\gamma(p_{|\mathcal{A}|}))} + \sum_{\substack{i=1 \\ i \neq j}}^{k} \frac{D(L_i)(\mathbb{P}_\gamma(p_1), \ldots, \mathbb{P}_\gamma(p_{|\mathcal{A}|}))}{L_i(\mathbb{P}_\gamma(p_1), \ldots, \mathbb{P}_\gamma(p_{|\mathcal{A}|}))} \leq E_{\mathbb{P}}[T]$$

with $L_i = \sum_{n=0}^{+\infty} (p_{2i-1} + p_{2i})^n = \frac{1}{1-(p_{2i-1}+p_{2i})}$

As there is an explicit formula for computing the generating series of the coupling time of a single queue, this generating series, and hence this lower bound, can be easily numerically computed.

4 CONCLUSIONS

In this paper we have shown a systematic way to compute exactly the expected coupling time. We illustrated how this can be used to get a lower bound in the case of a network of independent queues. The extension of the techniques in the last section to prove bounds on general queueing networks is still a work in progress.

Acknowledgments: Research supported by ANR grant ANR-16-CE05-0008.

REFERENCES

[1] A. Bouillard, A. Busic, and C. Rovetta. Perfect sampling for closed queueing networks. *Perform. Eval.*, 79:146–159, 2014.

[2] A. Bušić, B. Gaujal, and F. Pin. Perfect sampling of Markov chains with piecewise homogeneous events. *Performance Evaluation*, 69(6):247 – 266, 2012.

[3] A. Bušić, B. Gaujal, and J.-M. Vincent. Perfect simulation and non-monotone markovian systems. In *Proceedings of the 3rd International Conference on Performance Evaluation Methodologies and Tools*, ValueTools '08, pages 27:1–27:10. ICST, 2008.

[4] P. Flajolet and F. Guillemin. The formal theory of birth-and-death processes, lattice path combinatorics and continued fractions. *Adv. in Appl. Probab.*, 32(3):750–778, 2000.

[5] O. Häggström and K. Nelander. Exact sampling from anti-monotone systems. *Statist. Neerlandica*, 52(3):360–380, 1998.

[6] M. Huber. Perfect sampling using bounding chains. *The Annals of Applied Probability*, 14(2):734–753, 2004.

[7] W. S. Kendall and J. Møller. Perfect simulation using dominating processes on ordered spaces, with application to locally stable point processes. *Adv. in Appl. Probab.*, 32(3):844–865, 2000.

[8] J. G. Propp and D. B. Wilson. Exact sampling with coupled markov chains and applications to statistical mechanics. *Random Structures & Algorithms*, 9(1-2):223–252, 1996.