Uniformité d'un échantillonnage épidémique

Yann Busnel, Roberto Beraldi, Roberto Baldoni

Uniformité d'un échantillonnage épidémique

Yann Busnel[1], Roberto Beraldi[2] et Roberto Baldoni[2]

[1] *LINA / Université de Nantes, 2 rue de la Houssinière, BP 92208, F-44322 Nantes Cedex 03, France*
[2] *Dipartimento di Informatica e Sistemistica, Universitá di Roma – La Sapienza, Via Ariosto 25, I-00185 Roma, Italy*

Au sein d'un ensemble de nœuds, un service d'échantillonnage aléatoire idéal se doit de retourner un pointeur vers un nœud, correspondant à un échantillon indépendant sans biais du groupe considéré. Cet article porte sur un service d'échantillonnage issue de la notion de *mélange de vues*. Chaque nœud possède une vision locale du système, constituée d'un ensemble de c pointeurs vers des nœuds. Afin d'implémenter correctement un service d'échantillonnage uniforme, cette vue pouvant évoluer au fil du temps, doit contenir un échantillon uniforme de taille c. Pour cela, des couples de nœuds échangent périodiquement une partie de leur vue locale (*procédure de mélange*). Cet article propose des résultats prouvés dans [BBB09a, BBB09b] que *(i)* à partir de n'importe quelle distribution des vues locales, après une série de mélanges suffisamment longue, chaque vue contiendra irrémédiablement un échantillon uniforme de taille c ; *(ii)* une fois la propriété précédente vérifié, toute série de mélange consécutive ne modifiera pas la propriété d'uniformité et *(iii)* une *borne inférieure* de la vitesse de convergence.

Keywords: Echantillonnage de réseau; Protocole épidémique; Distribution uniforme; Analyse théorique; Processus stochastique

1 Introduction

Uniform peer sampling service has been shown recently to be a basic building block for several applications in large-scale distributed systems [JVG+07] as information dissemination, counting, clock synchronization, *etc*. Working on the top of a biased peer sampling can affect either performance, correctness or both of a given application. A sequence of invocations to a peer sampling service returns a sequence of samples of the peers belonging to the system. If samples are unbiased random samples of the system, the peer sampling service is called *uniform*. There are two main approaches to implement uniform random sampling, *random walk* and *gossip-based* protocols.

A random walk on a given graph is a sequential process that consists in visiting the nodes of the graph according to a random order induced by the way the walker is allowed to move. More precisely, the walker moves from one node to one of its neighbors that is selected uniformly at random. The key property of a random walk is that, after a suitable number of steps, called the mixing-time, the visited node is the same as drawn from a uniform distribution [Bol01]. Thus, random walk-based peer sampling mechanisms aim at implementing a biased random walk. Unfortunately, the mixing-time depends on the topological property of the graph, which is generally unknown. Thus, for the reached node to be uniformly sampled, the length of the walk has to be properly tuned. Moreover, this technique may incur in a long delay to return a sample.

This paper focuses on uniform peer sampling based on gossip protocols. We consider a system formed by n peers (*i.e.*, nodes), each provided with a local view of size $c \leq n$. Each node runs a simple *shuffling protocol* where pair of nodes regularly and continuously swaps part of their local views (*shuffle operation*). This protocol is similar to the ones used in [JVG+07, VGvS05, BTV06]. The shuffling protocol aims that local views eventually represent a uniform random sample of the system. The main results presented in this paper show formally that :

1. starting from any non-uniform distribution of nodes in the local views, after a sufficiently long sequence of pairwise shuffle operations executed by the shuffling protocol, each local view represents a uniform random sample of size c among the whole system (Theorem 3.2) ;
2. once previous property has been established, any sequence of successive shuffle operations does not modify the previous property (Corollary 3.3) ;

3. using this protocol, optimal setting can be identified in term of convergence speed (Theorem 4.1).

To the best of our knowledge, these results have never been formally proved before, despite the fact that there is empirical evidence shown in many papers [JVG+07, VGvS05], that protocols based on view shuffling can provide continuously a uniform sampling.

Let us remark that this result complements the one presented in [BGK+08]. Indeed, authors of [BGK+08] propose a protocol based on view shuffling and formally prove that this protocol converges to a uniform peer sampling also in the presence of byzantine peers. Each run of their protocol leads, after a sufficiently long sequence of shuffle operations, to verify the property : "each local view is a uniform random sample of the system". However, each time a user requires to get a new uniform sample, another instance of this protocol has to be started and it has to converge to a new uniform random sample. Conversely, the shuffling protocol presented in this paper shows that once the local view converges to represent a uniform sample of the system, then successive shuffle operations do not modify the property (Corollary 3.3). Therefore, there is a continuous availability of a uniform random sample without the need to start other instances of the base protocol.

2 Protocol analysis

First of all, after defining a system model and the shuffling protocol (*aka* gossip-based protocol), we derive an analytical model of this protocol, which captures the variation of the system configuration over time. We consider a finite set of n nodes (with $n \geq 2$), which are uniquely identified through a system-wide identifier (ID). Each node i manages a local partial view of the system, denoted V_i of size $c \leq n$ about all the other nodes in the system, including itself.

The view of node i is modeled as a fixed-size set of binary random variables indicating whenever the identifier k appears in V_i or not

$$X_i = (X_{1i}, X_{2i}, \ldots, X_{ni})$$

where

$$X_{ki} = \begin{cases} 1 & \text{if } k \in V_i\,; \\ 0 & \text{otherwise} \end{cases}$$

2.1 Evolution of the system

Let now consider how the system evolves. As explained in [BBB09a, BBB09b], we assume that concurrent operations cannot occur. Thus, we can serialize parallel shuffles in an arbitrary order and assume that only one shuffling operation may take place at a time. Let $P_{ex}(i,j)$ be the probability that i and j make the shuffle, *i.e.*, $P_{ex}(i,j)$ is the probability that the operation $i \diamond j$ takes place.

As all the nodes ideally initiates an exchange at the same rate, we can consider that the initiator node is selected at random among all the n nodes. The target node j is taken at random from ℓ_i (the set of view's items sent by i). Hence

$$P_{ex}(i,j) = \frac{1}{n} \cdot \mathbb{P}[X_{ji} = 1] \cdot \frac{1}{c}.$$

We can describe the global evolution of the system with the following expression :

$\mathbb{P}[X'_{ki} = 1] =$

$$\sum_j P_{ex}(i,j) \cdot \mathbb{P}[X'_{ki} = 1 | i \diamond j] \tag{1a}$$

$$+ \sum_j P_{ex}(j,i) \cdot \mathbb{P}[X'_{ki} = 1 | j \diamond i] \tag{1b}$$

$$+ \left(1 - \sum_j (P_{ex}(i,j) + P_{ex}(j,i))\right) \cdot \mathbb{P}[X_{ki} = 1] \tag{1c}$$

This last equation means that the probability vector of a node follows the view evolution developed in [BBB09a, BBB09b] if it is involved in a view shuffle (Equation 1a and 1b) and remains the same if it is not involved in the last shuffle (Equation 1c).

3 Convergence property of the protocol

Let now consider the converge property of shuffling protocols. In particular, we show that if the shuffling protocol is executed by a system with arbitrary view distribution, then eventually the system converges towards a uniform configuration, *i.e.*, a system in which all the local views represent uniform random samples of the system. Due to space constraint, we only present the main theorem. Proofs are available [BBB09a, BBB09b]. Roughly speaking, a shuffling operation moves the system towards a "more" uniform system, or, in other words, makes the system closer to the uniform configuration.

In the following, we note $h(\mathcal{C})$ the distance between the configuration $\mathcal{C}$ and the uniform one. This distance represents the difference between $\max_{i,j} \mathbb{P}[X_{ji} = 1]$ and the uniform value $\overline{p} = \frac{c}{n}$. First of all, we prove that one shuffling operation only reduce the distance to this uniform configuration :

Lemma 3.1 (Operator $\diamond$ reduces the potential) *Let $\mathcal{C}$ and $\mathcal{C}'$ respectivelly the configuration of the system before and after a* shuffling operation. *Then $h(\mathcal{C}') < h(\mathcal{C})$.*

We are now in the position to state this theorem, independently of the shuffling operation considered :

Theorem 3.2 (Convergence to uniformity)
Let i be the number of shuffling operations executed on a system of n nodes, $\mathcal{C}_0$ be any initial unpartitioned distribution of local views and $\mathcal{C}_i$ be the configuration of the system after those i shuffling operations. Local views built by the shuffling protocol will converge to uniform random samples of the system, i.e.*,*

$$\forall \mathcal{C}_0, \lim_{i \to \infty} h(\mathcal{C}_i) = 0.$$

Proof The claim follows from result comes from Lemma 3.1, as a shuffling operation *strictly* reduce the global potential, independently of the pair involved in the shuffle. Thus, the distance of the current distribution of sample with the uniformity could only monotonically reduce, due to Equation 1. Then, the distribution of the samples converges to the uniform one. □

Let us now show a corollary stating that once local views represent uniform samples of the system, the shuffling protocol keeps this property true forever.

Corollary 3.3 (Operator $\diamond$ preserves uniformity) *Let $\mathcal{C}$ be a uniform unpartitioned distribution of local views. A shuffling operation executed by the shuffling protocol between any pair of two local views X_i and X_j belonging to $\mathcal{C}$ produces a distribution $\mathcal{C}'$ that is uniform.*

Proof Lemma 3.1 gives us that the potential of two views involved in a shuffling operation can only decrease. Given the fact that $\mathcal{C}$ corresponds to the uniform distribution, X_i and X_j are uniform and $P_i = P_j$ are vectors with all elements equal to $\overline{p} = \frac{c}{n}$. Thus, the potential of X_i and X_j are $h(P_i) = h(P_j) = 0$. From Lemma 3.1, after the shuffle, $h(P_i)$ and $h(P_j)$ cannot increase and thus, remain to 0. Then, $\mathcal{C}'$ is the uniform distribution. □

4 Lower bound in convergence speed

Experimental approaches [JVG+07, VGvS05, ...] point out that, in the design of a gossip-based protocol, the size of the exchange set l has to be set to the half of the complete view c, in order to obtain the highest efficiency in term of convergence speed. This conjecture can be intuitively shown as sketched below.

A shuffle operation with a sent vector ℓ between two nodes is equivalent to a shuffle with the complementary of ℓ (*i.e.* $V - \ell$), followed by swapping the ID of these two nodes (*cf.* Figure 1). Indeed, in this figure, the content of V_i after the shuffle on the left side is equivalent to the content of V_i on the right side after *(1)* a shuffle with the sent vector $\ell'_i = V_i - \ell_i$ and *(2)* swapping the node's ID (i becomes k and *vice versa*).

Now, consider $l \leq \frac{c}{2}$. It is obvious that the higher the size of the sent vector, the greater the effectiveness[†] of a shuffle. Moreover, according to the above equivalence, a shuffle with l is equivalent to a shuffle with $c - l$. Thus, for $l \geq \frac{c}{2}$, the lesser the size of the sent vector, the greater the effectiveness of a shuffle. So, the greatest effectiveness is reached for $l = \left\lfloor \frac{c}{2} \right\rfloor$, as confirmed numerically in [BBB09a, BBB09b].

[†] Roughly speaking, *effectiveness* represents how different the shuffled views are from the ones before the shuffle. The higher the difference, the greater the effectiveness.

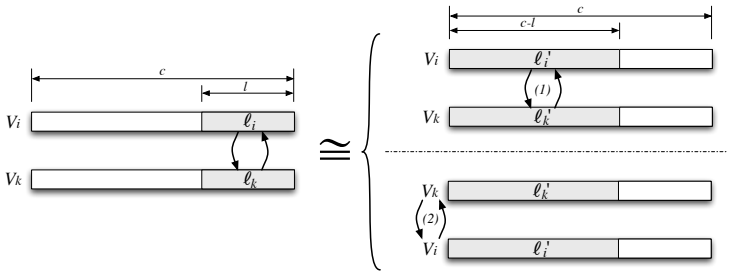

FIG. 1: Intuitive equivalence between a small ℓ value and the opposite $c-\ell$ ones.

Let us proved it formally, in Theorem 4.1 below. First, we have to express in a measure of the effectiveness :

Shuffling Effectiveness The effectiveness of a shuffle operation correspond to the magnitude of difference between an update view and both of the views before the shuffle, *i.e.* : $\mathcal{E}(P') = \min\{P \cdot P', Q \cdot P'\}$ where $P \cdot P'$ represent the scalar product between the vectors P and P'.

Indeed, according to the reasoning above, the core idea of a shuffling operation is to mix at most as possible both views involved in the shuffle. Thus, more different P' is from its initial state P is a good measure, but it has also to be balance with its similarity to the partner ones Q.

Starting from this definition, we should *maximize the effectiveness* of each operation. We prove in [BBB09a] that this maximization is achieve for $l = \frac{1}{2} \cdot c$.

Theorem 4.1 (Greatest Effectiveness of a Shuffle) *Given two probability vector P and Q. The maximum value of the expected effectiveness is reach for* $l = \frac{1}{2} \cdot c$.

5 Concluding Remarks

The paper has provided a theoretical ground to the fact that a shuffling protocol provides eventually nodes with uniform random samples of a system. Before this was only an empirical evidence. Differently from [BGK+08], our analysis shows that the same instance of the shuffling protocol can provide permanently a node with uniform sample of the system. Corollary 1 formally grasps this difference.

In [BBB09a, BBB09b], we also presented a numerical evaluation of the shuffling algorithm on its convergence speed of the local views to uniform random samples. We also formally proved what is the best fraction of the local views to swap in a shuffling operation to get best convergence speed.

Références

[BBB09a] Y. Busnel, R. Beraldi, and R. Baldoni. A Formal Characterization of Uniform Peer Sampling based on View Shuffling. Technical Report 4/09, MIDLAB, June 2009.

[BBB09b] Y. Busnel, Y. Beraldi, and R. Baldoni. A formal characterization of uniform peer sampling based on view shuffling. In *the 2nd IEEE Workshop on Reliability, Availability and Security (WRAS '09)*, Dec. 2009.

[BGK+08] Edward Bortnikov, Maxim Gurevich, Idit Keidar, Gabriel Kliot, and Alexander Shraer. Brahms : byzantine resilient random membership sampling. In *the 27th ACM symposium on Principles of Distributed Computing (PODC '08)*, pages 145–154, Toronto, Canada, 2008. ACM.

[Bol01] Béla Bollobás. *Random Graphs – 2nd Edition*. Cambridge University Press, Cambridge, UK, 2001.

[BTV06] François Bonnet, Frederic Tronel, and Spyros Voulgaris. Brief announcement : Performance analysis of cyclon, an inexpensive membership management for unstructured p2p overlays. In *DISC*, pages 560–562, 2006.

[JVG+07] Márk Jelasity, Spyros Voulgaris, Rachid Guerraoui, Anne-Marie Kermarrec, and Maarten van Steen. Gossip-based peer sampling. *ACM Transaction on Computer System*, 25(3) :8, august 2007.

[VGvS05] Spyros Voulgaris, Daniela Gavidia, and Maarten van Steen. CYCLON : Inexpensive Membership Management for Unstructured P2P Overlays. *Journal of Network System Management*, 13(2) :197–217, june 2005.