Coupon Collector Problem on Graphs

Mohammed Barmaki[#1]
[#]*EcoleNationaledesSciencesAppliquées& Department of Informatics, Logistic and Mathematics&IbnTofail Kénitra Morocco*

Abstract

In this article, we start from the combinatorial version of the coupon collector problem, in order to generalize it to the infinitely generated groups. We introduce it analogously to the waiting time in order to complete an n-collection and then after, we establish the graph invariance associated with a finitely generated group. We compute the average of this waiting time for monogenic and free groups.

Keywords — *Random walk, Group, Markov Chain, Isoperimetric profile.*

I. INTRODUCTION

The coupon collector problem is a standard combinatorial problem which consists of estimating the expected number of coupons purchases needed to complete the collection of all possible n different types.

By generalizing this problem, let the coupons be obtained by using an arbitrary probability distribution, and then consider other related process, the problem has been found to be modelled by many practical situations. The usefulness of this model has been seriously hampered by the computational difficulties in obtaining any numerical results concerning moments or distributions (see [1,2]).

In this article we give an interpretation of the problem of the collector in terms of random walks.

For more details on random walks one can see [4,5,6,9,10;11]. We will establish, as in [3,7], an invariant of locally compact groups.

Therefore, we consider a finitely generated group G with neutral element e. Let S be a finite minimal generating subset of G. We define for all x of $G-\{e\}$ the length of x by :
$l_s(x) = Min\{\, k \in \mathbb{N}^* \,;\, \exists\, s_1, s_2, \ldots, s_k :\ x = s_1, s_2 \ldots . s_k\},\ l_s(e) = 0$.

Let $S^k = \{\, x \in G \,|\, l_s(x) = k\}, and\ B_n(S) = \{\, x \in G^+ | l_s(x) \le n\}$, where $G^+ = \cup_{n \in \mathbb{N}} S^n$.

For a positive integer n we consider $\Omega = G^{\mathbb{N}}$.

Let $X_i : \Omega \to G$ be the i-th canonical projection.

We denote by $A|B$ the event A such that B is realized.

We define a probability P on Ω, such that (X_n) is a Marrkov chain, with transition rule given by:

$\forall k \in \mathbb{N}, \forall s \in S, \forall g \in B_{n-1}(S) \cap X_k(\Omega)$, $P\,(X_{k+1} = g \mid X_k = g\,) = \frac{l_S(g)}{Card\,(\,B_{n-1}(S)\,)}$, and

$$P\,(X_{k+1} = gs \mid X_k = g\,) = (1 - \frac{l_S(g)}{Card\,(\,B_{n-1}(S)\,)})\,\frac{1}{card\,(S)},$$

$$\forall k \in \mathbb{N}, \forall g \in (\,B_n(S) - B_{n-1}(S)\,) \cap X_k(\Omega), P\,(X_{k+1} = g \mid X_k = g\,) = 1$$

and $P\,(X_{k+1} = l\,|X_k = g\,) = 0$ in the other cases. And X_0 has a uniform $(B_{n-1}(S))$ distribution.

We define the random variable $U_n = card\{k\,; l_S(X_k) \le n-1\}$, which is the waiting time of arrival at the first element of length n in G.

We remark that for finite groups $U_n = +\infty$from a certain rank, in the next we suppose that G is infinite.We are interested in asymptotic behavior of$\psi_S(n) = E(U_n\,|X_0 = e)$in the folllowing sense:

For two real valued functions f, g defined on a discrete subset of $]0, +\infty[$, we define the relation $f \preccurlyeq g$ by:

$$\exists\, \alpha, \beta \in\ ]0, +\infty[,\ \forall\, x \in\ ]0, +\infty[\ ;\ \bar{f}(x) \le \alpha \bar{g}(\beta x) + \alpha$$

where $\bar{f}$and $\bar{g}$are the linear interpolations of f and g. When $f \preccurlyeq g$ and $g \preccurlyeq f$, we write $f \asymp g$ .The asymptotic behavior of f is the equivalence class of f for the relation $\asymp$.

II. STABILITY OF THE ASYMPTOTIC BEHAVIOR OF THE AVERAGE WAITING TIME $\boldsymbol{\psi_S(n)}$

In this section we show that the asymptotic behavior of $\psi_S(n)$is independent of the generating set S, which allow us to construct an invariant of the Cayley graph (G^+, S).

Proposition 1.
If S and S' are two minimum generating sets of G such that $G^+ = \cup_{n\in\mathbb{N}} S^n = \cup_{n\in\mathbb{N}} S'^n$,then $\psi_S(n) \asymp \psi_{S'}(n)$.

Proof.
There exists a positive integer p such that $S \subset \cup_{K=0}^{p} S'^k$. By an induction on i, one gets

$$\{ i, l_S(X_i) \leq n-1 \} \subset \{ i, l_{S'}(X_i) \leq np-1 \},$$

hence

$$E(U_n) \leq E(U'_{np}).$$

It follows that $\psi(n) \leq \psi(np)$hence $\psi_S(n) \preccurlyeq \psi_{S'}(n)$, and exchanging the roles of S and S' we obtain the result.■
In the sequal, we denote $\psi(n)$the asymptotic behavior of $\psi_S(n)$.

III. FOR INFINITE MONOGENIC GROUP

In this section we prove the following result.

Theorem 3.1.
If G is an infinite monogenic group, then the average waiting time on G satisfies $\psi(n) \asymp nln(n)$.

Proof.
The group G is monogenic, then there exists $a \in G$ such that $S = \{a\}$ is a minimal generating subset of G .We can write $U_n = card\{k\,; l_S(X_k) < n\} = \sum_{i\geq 1} 1_{\{X_i \in B_{n-1}\}}$ so $U_n = 1_{\{X_0 \in B_{n-1}\}} + V_n$ where$V_n = \sum_{i\geq 1} 1_{\{X_i \in B_{n-1}\}}$
For $k \in \{0, \ldots\ldots, n-1\}$, $let\, u(n,k) = E(U_n \,|l_S(X_0) = k)$, so$\psi(n) = u(n,0)$.
Then for all $k \in \{0, \ldots\ldots, n-1\}$,

$$\begin{aligned} u(n,k) &= E(U_n\,|l_S(X_0) = k) \\ &= E(V_n\,|l_S(X_0) = k) + 1 \\ &= E(V_n\,|l_S(X_0) = k\,,\, l_S(X_1) = k\,)\,\mathrm{P}(l_S(X_1) = k\,|\,l_S(X_0) = k\,) + E(V_n\,|l_S(X_0) = k\,,\, l_S(X_1) = k+1)\mathrm{P}(l_S(X_1) = k+1\,|\,l_S(X_0) = k\,) + 1 \\ &= E(V_n\,|l_S(X_1) = k\,)\,\mathrm{P}(l_S(X_1) = k\,|\,l_S(X_0) = k\,) + E(V_n\,|l_S(X_1) = k+1)\mathrm{P}(l_S(X_1) = k+1\,|\,l_S(X_0) = k\,) + 1 \\ &= \frac{k}{n}u(n,k) + \frac{n-1}{n}u(n,k+1) + 1, \end{aligned}$$

Hence$u(n,k) = u(n,k+1) + \frac{n}{n-k}$ so, $u(n,0) = \sum_0^{n-1} \frac{n}{n-k} \asymp nln(n)$and it follows that $\psi(n) \asymp nln(n)$. ■

IV. FOR A FREE GROUP

Free groups form a fairly representative class of non-amenable groups (see [8]).For a free group, we have the following result.

Theorem 4.1.
Let G be a free group with $p$generators,$p > 1$. Then the average waiting time of the visit of the n −th ring is $\psi(n) \asymp n$.

Proof.

We consider a minimal generating subset $S = \{x_1, \ldots\ldots.., x_p\}$ of G.Keeping the notations introduced in the preceding section, we have

$$u(n,k) = 1 + \frac{k}{p^n}u(n,k) + (1 - \frac{k}{p^n})u(n,k+1),$$

therefore;

$$u(n,k) - u(n,k+1) = \frac{1}{1-\frac{k}{p^n}} \quad ,$$

hence $(n,0) = \sum_{k=0}^{n-1} \frac{1}{1-\frac{k}{p^n}}$, and we obtain $n \leq \psi(n) \leq n \frac{p^n}{p^n - n + 1}$, and the result follows.■

V. LOWER AND UPPER BOUNDS OF $\psi(n)$

We have the following property about the lower bound $\psi(n)$.

Proposition 2.

For any finitely generated group G, $\psi(n) \geq n$.

Proof.

When $X_0 = e$ is realized, then $X_1 = X_0$ or $X_1 \in S$, hence $l_S(X_1) \in \{0,1\}$, and by induction

$l_S(X_0) \leq n-1$, $l_S(X_1) \leq n-1$, … … .. , $l_S(X_{n-1}) \leq n-1$are realized, we get$\{0, \dots \dots, n-1\} \subset \{ i , l_S(X_i) < n \}$, hence$E(U_n | X_0 = e) \geq n$, so we obtain the lower bound.■

Proposition 3.
For any finitely generated group , for non slow random walk , $\psi(n) \preccurlyeq n ln(n)$.

Proof.
We have

$$u(n,k) \leq \frac{k}{card\,(S^n)} u(n,k) + (1 - \frac{k}{card\,(S^n)}) u(n,k+1) + 1$$

and since the random walk on G is not slow then for any n, we have $S^n \subsetneq S^{n+1}$, so for all positive integer n, $card(S^n) \geq n$, then

$$u(n,0) \leq \sum_{k=0}^{n-1} \frac{card(S^n)}{card(S^n) - k} \leq \sum_{k=0}^{n-1} \frac{n}{n-k} \leq n(1 + \ln(n)).$$

REFERENCES

[1] Boneh, Arnon and Hofri, Micha, "The Coupon-Collector Problem Revisited". Computer Science Technical Reports. Paper 807, e-Pubs Perdue University 1989.

[2] Amy N. Myers And Herbert S. WILF, Some new aspects of the coupon collector problem, SIAM J. DISCRETE MATH. Vol. 17, No. 1, pp. 1?17 (2003).

[3] Bendikov, A. Pittet,Ch. Roman Sauer, Spectral distribution and $L2$-isoperimetric profile of Laplace operators on groups, - Arxiv preprint arXiv:0901.0271, 2009 - arxiv.org.

[4] Coulhon,Th.Grigor'yan, A. On diagonal lower bounds for heat kernels on non-compact manifolds and Markov chains. Duke Math. J. 89 (1997), 133-199. MR1458975.

[5] Coulhon,Th.Grigor'yan,A.Pittet, Ch. A geometric approach to on-diagonal heat kernel lower bounds on groups. Ann. Inst. Fourier 51 (2001), 1763-1827. MR1871289.

[6] Gretete, D. Stabilit\'e du comportement des marches aléatoires sur un groupe localement compact, Ann. Inst. Henri PoincaréeProbab. Stat. 44(2008), no. 1, 129-142 (French, with English and French summaries).

[7] Hebisch, W. On heat kernels on Lie groups. Math. Zeit. 210 (1992), 593-605. MR1175724.

[8] *Kesten, H. Full Banach mean values on countable groups, Math. Scand. 7 (1959) 146-156.*

[9] Mustapha, S., C.R. Acad.Sci. Paris, Ser.I 340(2005).

[10] Mustapha, S.,Lower estimates for random walks on a class of amenable. p-adicgroups,Journal URL, vol. 14 , no 51, pp. 1513-1531(2009).

[11] Pittet, C., and Saloff-Coste, L. Random walks on finite rank Solvable Groups}, (May 2002).