普松分佈的推廣

Generalization of Poisson Distribution

楊效適

摘要. “統計表徵” (statistical token) 是一個由 “統計表徵產生器” (statistical token generator) 所產生的隨機數。所謂的 “統計表徵產生器” 是一個不均勻 (non-uniform) 的隨機數產生器。當使用者以統計表徵作爲通訊系統擷取頻道 (channel access) 之依據時, 此系統具有隨機性, 而且發生 “碰撞” (collision) 的或然率較低, “統計表徵產生器” 具有可調變的特性, 可產生出不同的統計表徵分佈, 當我們以或然率理論推導統計表徵分佈時, 我們可發現一個較普松分佈廣義的新分佈。

Abstract: Statistical token is a random number generated from statistical token generator which is a non-uniform random number generator, when users use the statistical token as the medium access control technology of communication system, a contention protocol with lower collision rate than CSMA/CD protocol can be built.

By using probability theory, a very concise formula was derived to describe the distribution of statistical token model.

1. 緒言

競爭型通訊協定 (contention protocol) 是一種數據通訊 (data communication) 系統常使用的通訊協定。由於數據通訊中的資料流 (data flow) 具有隨機性, 使得適用於隨機性資料流的競爭型通訊協定廣泛被採用。分析數據通訊系統與競爭型通訊協定常使用或然率理論。經過多年的分析與研究, 對於一些較初始 (primary) 競爭型通訊協定可以傳統隨機分佈 (普松分佈) 描述。

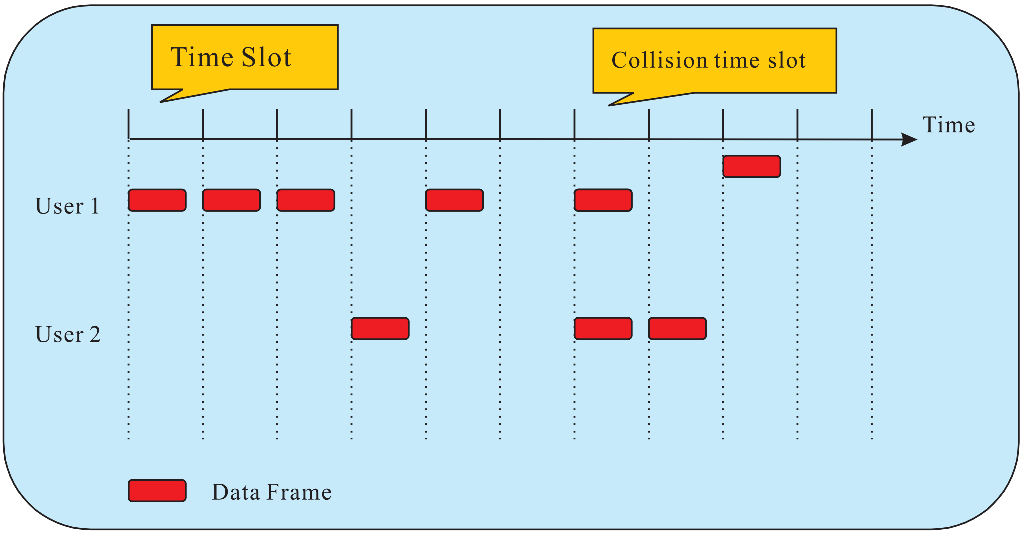
開發較優良的競爭型通訊協定是通訊系統研究項目之一, 相對於一種新通訊協定的提出, 新隨機模型也應建立。如此方能對於新通訊協定有深入的了解。本文提出一種新競爭型通訊協

定與其對應的隨機分佈。新競爭型通訊協定較初始競爭型通訊協定有更多的控制因素, 其對應的隨機分佈也較普松分佈更爲廣義。

2. ALOHA 通訊協定

ALOHA 通訊協定是美國夏威夷大學所發展出來的一種競爭型通訊協定。也是最早的一種競爭型通訊協定 (1970年代)。其架構也成爲後來以太網路 (Ethernet) 的基本架構。

所謂 ALOHA 通訊協定是指: 通訊頻道區分爲固定長度的時槽 (time slot), 同時資料框的長度也固定。每個時槽的長度足以正確的傳輸此一資料框, 同時使用者必需在時槽的起點開始傳送, 不得任意傳送 (參考圖一)。



圖一. ALOHA 通訊協定

現假設有 M 個使用者, 且每個使用者平均在 n 個時槽的時間內送出一個資料框。

爲了模擬 ALOHA 通訊協定, 我們可假設每個使用者有一範圍爲 $(0, n-1)$ 且平均分佈的隨機數產生器, 其產生之隨機數 i (出現機率 $= p$), 即代表這個使用者在時槽 i 送出資料框。

我們現在想知道; 在某一時槽內有 k 個使用者送出資料框或然率 $P_0(k)$。從上面的假設我們可看出 $P_0(k)$ 就是一個二項式分佈:

$$P_0(k) = C(M,k)\cdot p^k \cdot (1-p)^{M-k} \tag{2.1}$$

$$P(k) = \frac{\mu^k}{k!}\cdot e^{-\mu} \tag{2.2}$$

當 $M \to \infty$, $M_p = M/n = \mu$ 時, $P_0(k)$ 近似於 $P(k)$。

其中 $K \geq 2$是表示: 同一時槽內有二個或二個以上的使用者送出資料框而發生碰撞。

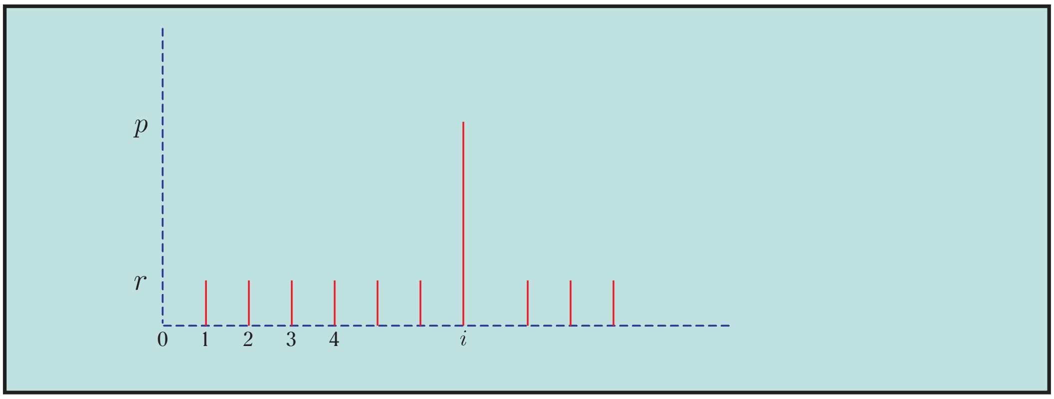
而 $K = 0$ 是表示: 此一時槽沒有任何使用者送出資料框。我們最感興趣的是 $K = 1$ 的情況, 此一情況表示: 資料框成功的送出, 而 $K = 1$ 的或然率 (或百分比), 就是此一通道的效能 (throughput)。

3. 統計表徵 (statistical token)

上一節中對於 ALOHA 通訊協定的分析所使用的模型: 每個使用者有一個均勻的隨機數產生器。每個使用者依照所得到的隨機數做爲送出資料框的時槽指標 (time slot index)。這個隨機數產生器決定了這個模型的隨機性, 同時也決定了此一模型 (或通訊協定) 效能的極限: 普松分佈。

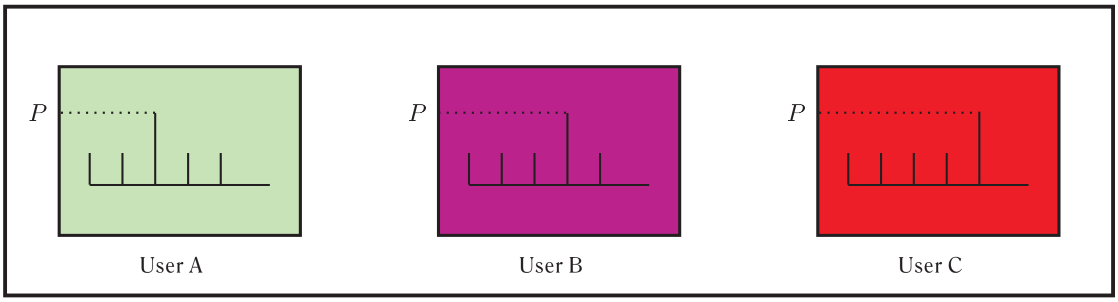
當我們設想此一模型內可以改變的因素時, 我們很直覺的就可發現: 均勻的隨機數產生器是一個可改變的因素。而此一因素有可能導致此一模型分佈的改變。

我們將每個使用者的隨機數產生器調整成一種不均勻的分佈, 而且每個使用者的隨機數產生器在某一特定的出現或然率較高。如在圖二中出現 i 的或然率爲 p, 其餘各點的或然率爲 r, 且 p 大於 r。



圖二. 不均勻隨機數產生器

現在假設有 n 個使用者, 其代號 id 介於 0 與 $n-1$ 之間。而其隨機數產生器出現 p 點的位置與其代號 id 相同。當此一模型成立時, 我們可發現: 每個使用者所得到的隨機數 i 並不是平均分佈在 0 與 $n-1$ 的範圍中, 而是較集中於自己的代號 id 的一種分佈 (如圖三)。



圖三. 統計表徵分佈圖

3.1. 統計表徵模型的分佈

現在我們要推導 "統計表徵" 模型的分佈, 如果推導出來的分佈在 $k=1$ 處是有較普松分佈爲高的或然率。那麼, 我們可以確定; 以此種模型爲基礎的通訊協定具有隨機性, 而且有較高的效能。

爲了使數學推導的過程更明瞭, 圖四是用來說明假設。假設如下:

(1) 有 M 個使用者, n 個時槽, 且 $M\le n$。

(2) 每個使用者平均在 n 個時槽的時間內送出一個資料框。

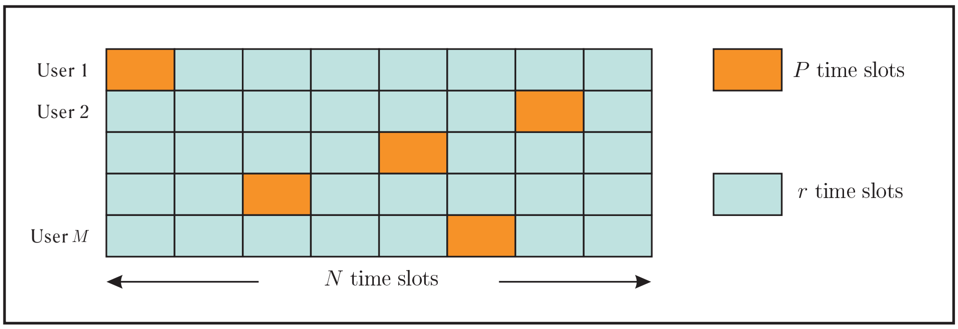
(3) M 個使用者的隨機數產生器分佈如圖三, 其在某一 "特定點" 出現或然率爲 p, 其餘爲 r。

(4) 這些 "特定點" 並不重複, 且 $p>r$。

現今令 $P_{st}(k)$ 爲此一模型的或然率分佈, 而 $P_{st}(k)$ 可表爲:

$$P_{st}(k)=u\cdot P_p(k)+(1-u)\cdot P_r(k)$$

其中 $\mu=M/n$, $P_p(k)$ 爲某一時槽內有對應之高或然率 (p) 時, 出現 k 個資料框的或然率。$P_r(k)$ 爲某一時槽內沒有對應之高或然率時, 出現 k 個資料框的或然率。



圖四. 統計表徵時槽

$$
\begin{aligned}
&P_{st}(k)\\
&= \mu \cdot P_p(k) + (1-\mu) \cdot P_r(k)\\
&= \mu \cdot ((1-p) \cdot C(M-1,k) \cdot r^k \cdot (1-r)^{M-K-1} + p \cdot C(M-1,k-1) \cdot r^{K-1} \cdot (1-r)^{M-k}\\
&\quad +(1-\mu) \cdot C(M,k) \cdot r^k \cdot (1-r)^{M-k}\\
&= \mu \cdot \left((1-p) \cdot \frac{(M-1)!}{(M-k-1)! \cdot k!} \cdot r^k \cdot (1-r)^{M-k-1} + p \cdot \frac{(M-1)!}{(M-k)! \cdot (k-1)!} \cdot r^{k-1} \cdot (1-r)^{M-k}\right)\\
&\quad +(1-\mu) \cdot \frac{M!}{(M-k)! \cdot k!} \cdot r^k \cdot (1-r)^{M-k}
\end{aligned}
$$

當 $M \to \infty$, 由 Stirling Formula: $n! \approx \sqrt{2 \cdot \pi \cdot n} \cdot n^n \cdot e^{-n}$

$$
\begin{aligned}
&P_{st}(k)\\
&= \mu \cdot \left((1-p) \cdot \frac{\sqrt{2 \cdot \pi \cdot (M-1)}}{\sqrt{2 \cdot \pi \cdot (M-k-1)}} \cdot \frac{(M-1)^{M-1}}{(M-k-1)^{M-k-1}} \cdot \frac{e^{-(M-1)}}{e^{-(M-k-1)}} \cdot \frac{1}{k!} \cdot r^k \cdot (1-r)^{M-k-1}\right.\\
&\quad \left. + p \cdot \frac{\sqrt{2 \cdot \pi \cdot (M-1)}}{\sqrt{2 \cdot \pi \cdot (M-k)}} \cdot \frac{(M-1)^{M-1}}{(M-k)^{M-k}} \cdot \frac{e^{-(M-1)}}{e^{-(M-k)}} \cdot \frac{1}{(k-1)!} \cdot r^{(k-1)} \cdot (1-r)^{M-k}\right)\\
&\quad +(1-\mu) \cdot \frac{\sqrt{2 \cdot \pi \cdot M}}{\sqrt{2 \cdot \pi \cdot (M-k)}} \cdot \frac{M^M}{(M-k)^{M-k}} \cdot \frac{e^{-M}}{e^{-(M-k)}} \cdot \frac{1}{k!} \cdot r^k \cdot (1-r)^{M-k}\\
&= \mu \cdot \left((1-p) \cdot \left(\frac{M-1}{M-k-1}\right)^{M-k-0.5} \cdot (M-1)^k \cdot r^k \cdot \frac{e^{-k}}{k!} \cdot (1-r)^{M-k-1}\right.\\
&\quad \left. + p \cdot \left(\frac{M-1}{M-k}\right)^{M-k-0.5} \cdot (M-1)^{k-1} \cdot r^{k-1} \cdot \frac{e^{-(k-1)}}{(k-1)!} \cdot (1-r)^{M-k}\right)\\
&\quad +(1-\mu) \cdot \left(\frac{M}{M-k}\right)^{M-k-0.5} \cdot M^k \cdot r^k \cdot \frac{e^{-k}}{k!} \cdot (1-r)^{M-k}.
\end{aligned} \tag{3.1}
$$

3.1 式中 $r = \dfrac{1-p}{n-1}$, $\mu = \dfrac{M}{n}$, 則 3.1 式各項可化簡如下:

(1) $(M-1)^k \cdot r^k = \left(\frac{M-1}{n-1} \cdot (1-p)\right)^k \approx \left(\frac{M}{n} \cdot (1-p)\right)^k = (\mu \cdot (1-p))^k$.

(2) $(M-1)^{k-1} \cdot r^{k-1} = (\mu \cdot (1-p))^{k-1}$.

(3) $(1-r)^{M-k-1} = \left(1 - \frac{1-p}{n-1}\right)^{M-k-1} = \left(1 - \frac{M}{n-1} \cdot \frac{1-p}{M}\right)^{M-k-1}$.

令 $M >> k \Rightarrow M \approx M-k-1$, 則 (3) $\approx \left(1 - \mu \cdot \frac{1-p}{M}\right)^M \approx e^{-\mu \cdot (1-p)}$.

(4) $(1-r)^{M-k} \approx e^{-\mu\cdot(1-p)}$.

(5) $\left(\frac{M-1}{M-k-1}\right)^{M-k-0.5} = \left(1+\frac{k}{M-k-1}\right)^{M-k-0.5} \approx e^k$.

(6) $\left(\frac{M-1}{M-k}\right)^{M-k+0.5} = \left(1+\frac{k-1}{M-k}\right)^{M-k+0.5} \approx e^{(k-1)}$.

(7) $\left(\frac{M}{M-k}\right)^{M-k+0.5} = \left(1+\frac{k}{M-k}\right)^{M-k+0.5} \approx e^k$

故 3.1 式可化簡如下:

$$\begin{aligned} P_{st}(k) = \mu \cdot \Bigg(&(1-p)\cdot e^k \cdot (\mu\cdot(1-p))^k \cdot e^{-k} \cdot e^{-\mu\cdot(1-p)} \cdot \frac{1}{k!} \\ &+p \cdot e^{(k-1)} \cdot (\mu\cdot(1-p))^{k-1} \cdot e^{-(k-1)} \cdot e^{-\mu\cdot(1-p)} \cdot \frac{1}{(k-1)!}\Bigg) \\ &+(1-\mu)\cdot\left(e^k \cdot (\mu\cdot(1-p))^k \cdot e^{-k} \cdot e^{-\mu\cdot(1-p)} \cdot \frac{1}{k!}\right) \\ = \frac{1}{k!} \cdot (\mu\cdot(1-p))^k \cdot e^{-\mu\cdot(1-p)} \cdot &\left(1-\mu\cdot p + \frac{p}{1-p}\cdot k\right) \end{aligned} \tag{3.2}$$

上面 (3.2) 式就是推導出當 $M \leq n$ 時 "統計表徵" 的分佈公式。當我們令 $p = 1/n$ 且 $n >> 1$, 使得 p 趨近於 0 時。(3.2) 式成爲:

$$P_{st}(k) = \frac{\mu^k}{k!} \cdot e^{-\mu} \tag{3.3}$$

我們發現: (3.3) 式就是我們在2節中所推導出的普松分佈。至此, 我們可確定: (3.2) 式是一個較普松分佈更爲廣義的隨機分佈。

表一

$\mu = 1$				
K	$p = 0.1$	0.2	0.3	0.4
0	0.365913	0.359463	0.347610	0.329287
1	0.369978	0.377436	0.392302	0.417097
2	0.184786	0.186921	0.189447	0.190986
3	0.060924	0.059431	0.056371	0.051369
4	0.014943	0.013803	0.011994	0.009681
5	0.002912	0.062515	0.001977	0.001399
6	0.000470	0.000376	0.000265	0.000164
7	0.000065	0.000048	0.000030	0.000016
8	0.000008	0.000005	0.000003	0.000001
9	0.000001	0.000001	0.000000	0.000000

表二

$\mu = 0.8$				
K	$p = 0.1$	0.2	0.3	0.4
0	0.447812	0.442926	0.434119	0.420773
1	0.361365	0.367839	0.380197	0.399982
2	0.144110	0.144706	0.144840	0.143518
3	0.037951	0.036630	0.034202	0.030567
4	0.007437	0.006782	0.005791	0.004580
5	0.001158	0.000986	0.000761	0.000527
6	0.000149	0.000118	0.000082	0.000049
7	0.000016	0.000012	0.000007	0.000004
8	0.000002	0.000001	0.000001	0.000000
9	0.000000	0.000000	0.000000	0.000000

3.2. 統計表徵分佈之數值計算

表一、表二分別爲從 (3.2) 式所計算出的數值。$\mu = 1$ 時, $P_{st}(k=1) > e^{-1} = P(k=1)$。$\mu = 0.8$ 時, $P_{st}(k=1) > 0.8 \cdot e^{-0.8} = P(k=1)$。由此可見: 以統計表徵 (statistical token) 爲通訊協定的系統有較低的碰撞機率。

4. 結論

本文由一種通訊協定的設計而導引出一個新型隨機分佈, 經由數學證明; 此新型隨機分佈可視爲一廣義的普松分佈, 此一新型隨機分佈具有二個控制因素: μ 與 p。較普松分佈多出一個 p 控制因素, 使此一新型分佈有較好的適應性。

本文並未討論 μ 大於*1*與更複雜統計表徵產生器 (不均勻隨機數產生器) 的影響, 筆者相信此一 (類) 新型分佈將對通訊協定 (communication protocol)、編碼技術 (coding technology)、訊務工程 (traffic engineering) 等領域具有適用性。

参考文獻

1. S. S. Lam, L. Kleinrock, Packet Switching in a Multiaccess Broadcast Channel :Performance Evaluation, IEEE Transactions on Communication, APRIL 1975.
2. M. D. Vose, A Linear Algorithm for Generating Random Number With a Given Distribution, IEEE Transactions on Software Engineering, Sept. 1991
3. W. Stalling, Local Networks: An Introduction.

—本文作者曾爲中華電信研究所副研究員—