MODELING AND ANALYSIS OF INFORMATION SYSTEMS, VOL. 27, NO. 1, 2020
journal homepage: www.mais-journal.ru

THEORY OF COMPUTING

On a Segment Partition for Entropy Estimation

E. A. Timofeev[1]

DOI: 10.18255/1818-1015-2020-1-40-47

[1]P. G. Demidov Yaroslavl State University, 14 Sovetskaya, Yaroslavl 150003, Russia.

MSC2020: 94A17
Research article
Full text in Russian

Received November 23, 2019
After revision February 18, 2020
Accepted February 28, 2020

Let Q_n be a partition of the interval $[0, 1]$ defines as

$$\begin{aligned}
&Q_1 = \{0, q^2, q, 1\}.\\
&Q'_{n+1} = qQ_n \cap q^2 Q_n, \quad Q''_{n+1} = q^2 + qQ_n \cap qQ_n, \quad Q'''_{n+1} = q^2 + qQ_n \cap q + q^2 Q_n,\\
&Q_{n+1} = Q'_{n+1} \cup Q''_{n+1} \cup Q'''_{n+1},
\end{aligned}$$

where $q^2 + q = 1$.
The sequence $d = 1, 2, 1, 0, 1, 2, 1, 0, 1, 0, 1, 2, 1, 0, 1, 2, 1, \dots$ defines as follows.

$$\begin{aligned}
&d_1 = 1,\ d_2 = 2,\ d_4 = 0;\\
&d[2F_{2n} + 1 : 2F_{2n+1} + 1] = d[1 : 2F_{2n-1} + 1];\\
&\quad n = 0, 1, 2, \dots;\\
&d[2F_{2n+1} + 2 : 2F_{2n+1} + 2F_{2n-2}] = d[2F_{2n-1} + 2 : 2F_{2n}];\\
&d[2F_{2n+1} + 2F_{2n-2} + 1 : 2F_{2n+1} + 2F_{2n-1} + 1] = d[1 : 2F_{2n-3} + 1];\\
&d[2F_{2n+1} + 2F_{2n-1} + 2 : 2F_{2n+2}] = d[2F_{2n-1} + 2 : 2F_{2n}];\\
&\quad n = 1, 2, 3, \dots;
\end{aligned}$$

where F_n are Fibonacci numbers ($F_{-1} = 0, F_0 = F_1 = 1$).
The main result of this paper.
Theorem.

$$Q'_n = 1 - Q'''_n = \left\{ \sum_{i=1}^{k} q^{n+d_i},\ k = 0, 1, \dots, m_n \right\},$$

$$Q''_n = 1 - Q''_n = \left\{ q^2 + \sum_{i=m_n}^{k} q^{n+d_i},\ k = m_n - 1, m_n, \dots, m_{n+1} \right\},$$

where $m_{2n} = 2F_{2n-2},\ m_{2n+1} = 2F_{2n-1} + 1$.

Keywords: measure; metric; entropy; estimation; unbiased; self-similarity; Bernoulli measure

INFORMATION ABOUT THE AUTHORS

Evgeniy Alexandrovich Timofeev | orcid.org/0000-0002-3094-4390. E-mail: timofeevEA@gmail.com
Sc.D., professor.

For citation: E. A. Timofeev, “On a Segment Partition for Entropy Estimation”, *Modeling and analysis of information systems*, vol. 27, no. 1, pp. 40-47, 2020.

Об одном разбиении отрезка, применяемом для оценки энтропии

Е. А. Тимофеев[1]

DOI: 10.18255/1818-1015-2020-1-40-47

[1]Ярославский государственный университет им. П. Г. Демидова, ул. Советская, 14, Ярославль, 150003 Россия.

УДК 519.17
Научная статья
Полный текст на русском языке

Получена 23 ноября 2019 г.
После доработки 18 февраля 2020 г.
Принята к публикации 28 февраля 2020 г.

В работе изучается разбиение отрезка, которое строится по следующему правилу:

$$\begin{aligned}
&Q_1 = \{0, q^2, q, 1\}.\\
&Q'_{n+1} = qQ_n \cap q^2Q_n,\quad Q''_{n+1} = q^2 + qQ_n \cap qQ_n,\quad Q'''_{n+1} = q^2 + qQ_n \cap q + q^2Q_n,\\
&Q_{n+1} = Q'_{n+1} \cup Q''_{n+1} \cup Q'''_{n+1},
\end{aligned}$$

где $q^2 + q = 1$.
Введем последовательность чисел $d = 1, 2, 1, 0, 1, 2, 1, 0, 1, 0, 1, 2, 1, 0, 1, 2, 1, \dots$, положив

$$\begin{aligned}
&d_1 = 1,\ d_2 = 2,\ d_4 = 0;\\
&d[2F_{2n} + 1 : 2F_{2n+1} + 1] = d[1 : 2F_{2n-1} + 1];\\
&\quad n = 0, 1, 2, \dots;\\
&d[2F_{2n+1} + 2 : 2F_{2n+1} + 2F_{2n-2}] = d[2F_{2n-1} + 2 : 2F_{2n}];\\
&d[2F_{2n+1} + 2F_{2n-2} + 1 : 2F_{2n+1} + 2F_{2n-1} + 1] = d[1 : 2F_{2n-3} + 1];\\
&d[2F_{2n+1} + 2F_{2n-1} + 2 : 2F_{2n+2}] = d[2F_{2n-1} + 2 : 2F_{2n}];\\
&\quad n = 1, 2, 3, \dots;
\end{aligned}$$

где F_n – числа Фибоначчи ($F_{-1} = 0, F_0 = F_1 = 1$).
Основной результат работы.
Теорема.

$$Q'_n = 1 - Q'''_n = \left\{ \sum_{i=1}^{k} q^{n+d_i},\ k = 0, 1, \dots, m_n \right\},$$

$$Q''_n = 1 - Q''_n = \left\{ q^2 + \sum_{i=m_n}^{k} q^{n+d_i},\ k = m_n - 1, m_n, \dots, m_{n+1} \right\},$$

где $m_{2n} = 2F_{2n-2},\ m_{2n+1} = 2F_{2n-1} + 1$.

Ключевые слова: мера; метрика; энтропия; оценка; несмещенность; самоподобие; мера Бернулли

ИНФОРМАЦИЯ ОБ АВТОРАХ

Евгений Александрович Тимофеев | orcid.org/0000-0002-3094-4390. E-mail: timofeevEA@gmail.com
доктор. физ.-мат. наук., профессор кафедры теоретической информатики.

Для цитирования: E. A. Timofeev, “On a Segment Partition for Entropy Estimation”, *Modeling and analysis of information systems*, vol. 27, no. 1, pp. 40-47, 2020.

В работе [1] для обоснования несмещенности оценки энтропии применялась последовательность разбиений отрезка, которая строилась по следующим рекуррентным правилам:

$$\begin{aligned}
&Q_1 = \{0, q^2, q, 1\}.\\
&Q'_{n+1} = qQ_n \cap q^2 Q_n, \quad Q''_{n+1} = q^2 + qQ_n \cap qQ_n, \quad Q'''_{n+1} = q^2 + qQ_n \cap q + q^2 Q_n,\\
&Q_{n+1} = Q'_{n+1} \cup Q''_{n+1} \cup Q'''_{n+1},
\end{aligned} \tag{1}$$

где $q^2 + q = 1$.

В настоящей работе будет показано, что Q_n – измельчающая последовательность разбиений и отрезки разбиения Q_n имеют длины q^n, q^{n+1}, q^{n+2}. Найдено рекуррентное задание длин отрезков разбиения Q_n.

Введем последовательность чисел $d = 1, 2, 1, 0, 1, 2, 1, 0, 1, 0, 1, 2, 1, 0, 1, 2, 1, \dots$, положив

$$\begin{aligned}
&d_1 = 1,\ d_2 = 2,\ d_4 = 0;\\
&d[2F_{2n} + 1 : 2F_{2n+1} + 1] = d[1 : 2F_{2n-1} + 1];\\
&\quad n = 0, 1, 2, \dots;\\
&d[2F_{2n+1} + 2 : 2F_{2n+1} + 2F_{2n-2}] = d[2F_{2n-1} + 2 : 2F_{2n}];\\
&d[2F_{2n+1} + 2F_{2n-2} + 1 : 2F_{2n+1} + 2F_{2n-1} + 1] = d[1 : 2F_{2n-3} + 1];\\
&d[2F_{2n+1} + 2F_{2n-1} + 2 : 2F_{2n+2}] = d[2F_{2n-1} + 2 : 2F_{2n}];\\
&\quad n = 1, 2, 3, \dots;
\end{aligned} \tag{2}$$

где F_n – числа Фибоначчи ($F_{-1} = 0, F_0 = F_1 = 1$).

Теорема 1. *Пусть последовательность d задана в (2), тогда*

$$Q'_n = 1 - Q'''_n = \left\{ \sum_{i=1}^{k} q^{n+d_i},\ k = 0, 1, \dots, m_n \right\}, \tag{3}$$

$$Q''_n = 1 - Q''_n = \left\{ q^2 + \sum_{i=m_n}^{k} q^{n+d_i},\ k = m_n - 1, m_n, \dots, m_{n+1} \right\}, \tag{4}$$

где $m_{2n} = 2F_{2n-2},\ m_{2n+1} = 2F_{2n-1} + 1$.

Замечание 1. *Последовательность d приведена и в «Энциклопедии целочисленных последовательностей»[1] под номером A191329, где она задана формулой*

$$d_n = [nq + n] \mod 2 + [nq] \mod 2.$$

Замечание 2. *Из (3) получаем формулу*

$$q^{2-n} = \sum_{i=1}^{m_n} q^{d_i}.$$

Доказательство. Для упрощения задания множеств Q_n докажем несколько вспомогательных лемм

[1] https://oeis.org/

Лемма 1. $Q'_n = 1 - Q'''_n$, $Q''_n = 1 - Q''_n$.

Доказательство. Проведем индукцию по n.

При $n = 1$ условия леммы выполнены.

Предположим, что они выполнены для n и покажем, что они выполнены при $n + 1$.

Из предположений следует, что $Q_n = 1 - Q_n$.

Из определения разбиения (1) следует, что

$$1 - Q'''_{n+1} = q - qQ_n \cap q^2 - q^2 Q_n = qQ_n \cap q^2 Q_n = Q'_{n+1}.$$

$$1 - Q''_{n+1} = q - qQ_n \cap 1 - qQ_n = q(1 - Q_n) \cap q^2 + q(1 - Q_n) = qQ_n \cap q^2 + qQ_n = Q''_{n+1}.$$

□

Лемма 2. *Справедливы соотношения*

$$Q'_n = qQ'_{n-1} \cup qQ''_{n-1}. \tag{5}$$

$$Q'_n \cap qQ'_n = qQ'_{n-1}, \tag{6}$$

$$Q'_n \cap qQ''_n = qQ''_{n-1}, \tag{7}$$

$$Q''_n \subset q^2 + qQ'_n, \tag{8}$$

$$Q''_n \subset qQ'''_n, \tag{9}$$

$$Q'_{n-1} \subset Q'_n, \tag{10}$$

$$Q''_{n-1} \subset Q''_n, \tag{11}$$

Доказательство. Проведем индукцию по n.

При $n = 2$ из (1) имеем

$$Q_2 = \{0, q^3, q^2, q, q + q^4, 1\},$$

поэтому условия леммы выполнены.

Предположим, что они выполнены для n и покажем, что они выполнены при $n + 1$.

Докажем (5). Из определения разбиения (1) следует, что

$$Q'_{n+1} = \left(qQ'_n \cup qQ''_n\right) \cap \left(q^2 Q'_n \cup q^2 Q''_n \cup q^2 Q'''_n\right) = q\left[\left(Q'_n \cap qQ'_n\right) \cup \left(Q'_n \cap qQ''_n\right) \cup \left(Q''_n \cap qQ'''_n\right)\right].$$

Применяя (6), (7), (9), получим

$$Q'_{n+1} = q\left(qQ'_{n-1} \cup qQ''_{n-1} \cup Q''_n\right).$$

Применяя (5), получим

$$Q'_{n+1} = q\left(Q'_n \cup Q''_n\right). \tag{12}$$

Докажем (6). Из последнего равенства (12) имеем

$$Q'_{n+1} \cap qQ'_{n+1} = \left(qQ'_n \cup qQ''_n\right) \cap \left(q^2 Q'_n \cup q^2 Q''_n\right) = q\left[\left(Q'_n \cap qQ'_n\right) \cup \left(Q'_n \cap qQ''_n\right)\right].$$

Применяя (6), (7), (5), получим

$$Q'_{n+1} \cap qQ'_{n+1} = q\left(qQ'_{n-1} \cup qQ''_{n-1}\right) = qQ'_n.$$

Докажем (7). Из определения разбиения (1) следует, что

$$Q''_{n+1} = q^2 + qQ'_n \cap qQ'''_n. \tag{13}$$

Отсюда и из (12) следует, что

$$Q'_{n+1} \cap qQ''_{n+1} = \left(qQ'_n \cup qQ''_n\right) \cap q^3 + q^2Q'_n \cap q^2Q'''_n = q\left[\left(Q'_n \cup Q''_n\right) \cap q^2 + qQ'_n \cap qQ'''_n\right] = \\ = q\left[Q''_n \cap q^2 + qQ'_n \cap qQ'''_n\right].$$

Применяя (8), (9), получим

$$Q'_{n+1} \cap qQ''_{n+1} = qQ''_n.$$

Докажем (8). Подставив (12) и (13), получим, что нужно доказать вложение

$$Q''_{n+1} = q^2 + qQ'_n \cap qQ'''_n \subset q^2 + q^2Q'_n \cup q^2 + q^2Q''_n.$$

Применяя (5), получим

$$q^2 + q\left(qQ'_{n-1} \cup qQ''_{n-1}\right) \cap qQ'''_n \subset q^2 + q^2Q'_n \cup q^2 + q^2Q''_n.$$

Последнее вложение выполняется по (10), (11).

Докажем (9). Из леммы 1 и (12) следует, что

$$Q'''_{n+1} = q^2 + qQ''_n \cap q^2 + qQ'''_n. \tag{14}$$

Подставив (13), получим, что нужно доказать вложение

$$Q''_{n+1} = q^2 + qQ'_n \cap qQ'''_n \subset q^3 + q^2Q''_n \cup q^3 + q^2Q'''_n = Q'''_{n+1}.$$

Для этого достаточно доказать, что

$$Q'''_n \subset q^2 + qQ''_n \cup q^2 + qQ'''_n.$$

Из леммы 1 и (5) следует, что

$$Q'''_n = q^2 + qQ''_{n-1} \cap q^2 + qQ'''_{n-1}.$$

Подставляя, получим вложение

$$q^2 + qQ''_{n-1} \cap q^2 + qQ'''_{n-1} \subset q^2 + qQ''_n \cup q^2 + qQ'''_n,$$

которое выполняется по (10), (11).

Докажем (10). Подставив (5), (12), получим, что нужно доказать вложение

$$Q'_n = qQ'_{n-1} \cup qQ''_{n-1} \subset qQ'_n \cup qQ''_n = Q'_{n+1},$$

которое выполняется по (10), (11).

Докажем (11). Подставив (13), получим, что нужно доказать вложение

$$Q''_n \subset Q''_{n+1} = q^2 + qQ'_n \cap qQ'''_n.$$

Применяя (10), (11), получим

$$Q''_{n+1} = q^2 + qQ'_n \cap qQ''_n \supset q^2 + qQ'_{n-1} \cap qQ'''_{n-1} = Q''_n.$$

□

Введем обозначения.

Пусть $d = d_1, d_2, \dots, d_m$ – последовательность целых чисел, тогда положим

$$\overline{d} = d_m, d_{m-1}, \dots, d_1, \tag{15}$$

$$d^* = 0, d_3, d_4, \dots, d_m, \tag{16}$$

$$Q(d) = \left\{ \sum_{i=1}^{k} q^{d_i},\ k = 0, 1, \dots, m \right\}. \tag{17}$$

Лемма 3.

$$Q'_n = q^n Q(d'_n), \tag{18}$$

$$Q''_n = q^2 + q^n Q(d''_n), \tag{19}$$

где последовательности d'_n, d''_n удовлетворяют рекуррентным уравнениям

$$d'_{2n} = d'_{2n-1} d''_{2n-1}, \quad d''_{2n} = d'_{2n-1}, \tag{20}$$

$$\begin{aligned} d'_{2n-1} &= d'_{2n-3} d''_{2n-3} d'_{2n-3}, \\ d'^*_{2n-1} &= d''_{2n-1} d'_{2n-3}, \\ d''_{2n-1} &= d''_{2n-3} d'_{2n-5} d''_{2n-3}. \end{aligned} \tag{21}$$

Доказательство. Проведем индукцию по n.

$$\begin{aligned} Q'_3 &= q^3 Q(d'_3) = \{0, q^4, q^3, q^2\},\ d'_3 = 1, 2, 1; \\ Q''_3 &= q^2 + q^3 Q(d''_3) = \{q^2, q\},\ d''_3 = 0; \end{aligned}$$

$$\begin{aligned} Q'_4 &= q^4 Q(d'_4) = \{0, q^5, q^4, q^3, q^2\},\ d'_4 = 1, 2, 1, 0; \\ Q''_4 &= q^2 + q^4 Q(d''_4) = \{q^2, q^2 + q^5, q^2 + q^4, q\},\ d''_4 = 1, 2, 1; \end{aligned}$$

$$\begin{aligned} Q'_5 &= q^5 Q(d'_5) = \{0, q^6, q^5, q^4, q^3, q^3 + q^6, q^3 + q^5, q^2\},\ d'_5 = 1, 2, 1, 0, 1, 2, 1; \\ Q''_5 &= q^2 + q^5 Q(d''_5) = \{q^2, q^2 + q^5, q^2 + q^4, q\},\ d''_5 = 0, 1, 0; \end{aligned}$$

Поэтому лемма справедлива при $n \le 5$. Предположим, что утверждения леммы выполняются при некотором n, и покажем, что они выполнены для $n + 1$.

Применяя (5) и индукционное предположение, получим

$$Q'_{n+1} = qQ'_n \cup qQ''_n = q^{n+1} Q(d'_n) \cup q^3 + q^{n+1} Q(d''_n) = q^{2n+1} Q(d'_n d''_n).$$

Следовательно,

$$d'_{n+1} = d'_n d''_n. \tag{22}$$

При нечетном n получаем первое равенство в (20). При четном n, подставляя (20), получим первое равенство в (21).

Из (13) имеем

$$Q''_{2n+1} = q^2 + q^{2n+2} Q(d'_{2n}) \cap q^2 + q^{2n+2} Q(\overline{d'}_{2n}) = q^2 + q^{2n+2} \left[Q(d'_{2n}) \cap Q(\overline{d'}_{2n}) \right].$$

Найдем общую часть последовательностей $d'_{2n} = d'_{2n-1} d''_{2n-1}$ и $\overline{d'}_{2n} = d''_{2n-1} \overline{d'}_{2n-1}$.

Из второго равенства в (21) имеем

$$d'^*_{2n-1} d''_{2n-1} = d''_{2n-1} d'_{2n-3} d''_{2n-1}.$$

Поэтому,

$$d'_{2n} = 1, 2, d''_{2n-1} d'_{2n-3} d''_{2n-1}, \ \overline{d'}_{2n} = d''_{2n-1} d'_{2n-3} d''_{2n-1}, 2, 1.$$

Поскольку $q + q^2 = 1$,

$$Q(d'_{2n}) \cap Q(\overline{d'}_{2n}) = Q(d'^{*}_{2n-1} d''_{2n-1}).$$

Следовательно,

$$d''_{2n+1} = d''_{2n-1} d'_{2n-3} d''_{2n-1}$$

и третье равенство в (21) выполнено.

Подставляя полученное равенство в

$$d'^{*}_{2n+1} = d'^{*}_{2n} d''_{2n} = d'^{*}_{2n-1} d''_{2n-} d'_{2n-1} = d''_{2n-1} d''_{2n-3} d''_{2n-1} d'_{2n-1},$$

получим второе равенство в (21).

Докажем второе равенство в (20).

Из (13) имеем

$$Q''_{2n+2} = q^2 + q^{2n+3} Q(d'_{2n+1}) \cap q^2 + q^{2n+3} Q(\overline{d'}_{2n+1}) = q^2 + q^{2n+3} \left[Q(d'_{2n+1}) \cap Q(\overline{d'}_{2n+1}) \right].$$

Поскольку

$$d'_{2n+1} = d'_{2n-1} d''_{2n-} d'_{2n-1} = \overline{d'}_{2n+1},$$

получаем

$$d''_{2n+2} = d'_{2n+1}.$$

$\square$

Найдем величины $m_n = |d'_n|$.

Из (22) имеем

$$|d''_n| = m_{n+1} - m_n. \tag{23}$$

Подставляя во второе уравнение (21), получим

$$m_{2n} = m_{2n+1} - m_{2n-1}. \tag{24}$$

Подставляя (23) в первое уравнение (20), получим уравнение, равносильное (24).

Подставляя (23), (24) в третье уравнение (20), получим рекуррентное уравнение

$$m_{2n+1} - 4m_{2n-1} + 4m_{2n-3} - m_{2n-5} = 0 \tag{25}$$

с начальным условием $m_1 = 1$, $m_3 = 3$, $m_5 = 7$.

Корни характеристического уравнения равны 1, q^2, q^{-2}, их степени выражаются через числа Фибоначчи, поэтому общее решение уравнения (25) имеет вид

$$m_{2n+1} = C_0 + C_1 F_{2n-1} + C_2 F_{2n-2}.$$

Найдя константы из начальных условий, получим

$$m_{2n+1} = 2F_{2n-1} + 1.$$

Подставляя в(24), получим

$$m_{2n} = 2F_{2n-2}.$$

$\square$

References

[1] E. Timofeev, "Existence of an unbiased consistent entropy estimator for the special Bernoulli measure", *Modeling and Analysis of Information Systems*, vol. 26, no. 2, pp. 267–278, 2019.