образуют последовательность $\{12, 17, 32, 36\}$, однако не существует такой подстановочной матрицы P, что $A_2 > P$.

ЛИТЕРАТУРА

1. *Бар-Гнар Р. И., Фомичев В. М.* О минимальных примитивных матрицах // Прикладная дискретная математика. Приложение. 2014. №7. С. 7–9.
2. *Фомичев В. М.* Свойства минимальных примитивных орграфов // Прикладная дискретная математика. 2015. №2(28). С. 86–96.

УДК 519.226, 519.244.3, 519.244.8 DOI 10.17223/2226308X/11/3

ПРОВЕРКА ГИПОТЕЗЫ О ВЛОЖЕНИИ С ДОПУСКОМ ДЛЯ ДИСКРЕТНЫХ СЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Н. М. Меженная

Последовательность X является подпоследовательностью с допуском d последовательности Y, если X получается из Y удалением несмежных отрезков не более чем из d знаков. В этом случае говорят, что X может быть вложена в Y с допуском d. Предложен последовательный критерий проверки гипотезы о вложении с допуском d для дискретных случайных последовательностей над конечным алфавитом и изучены его свойства. Вероятность ошибки первого рода (вероятность отклонения верной гипотезы о вложении с допуском) построенного критерия равна нулю. Трудоёмкость предложенной процедуры пропорциональна длине вкладываемой последовательности, что по порядку намного меньше трудоёмкости тотального опробования. Получено выражение для вероятности ошибки второго рода при альтернативной гипотезе о том, что рассматриваемые дискретные последовательности образованы независимыми в совокупности случайными величинами с равномерными распределениями на конечном алфавите.

Ключевые слова: *плотное вложение, вложение с допуском, последовательный критерий, гипотеза о независимости, вероятности ошибок первого и второго рода, дискретная случайная последовательность.*

Введение

Пусть $X_n = (x_1, \ldots, x_n)$ и $Y_m = (y_1, \ldots, y_m)$ — последовательности элементов множества $A_N = \{0, \ldots, N-1\}$, $N \geqslant 2$, длин n и $m \geqslant 1 + (d+1)(n-1)$ соответственно. Последовательность X_n может быть *вложена с допуском* $d \geqslant 1$ в начало последовательности Y_m, если существуют такие натуральные числа

$$1 = j_1 < j_2 < \ldots < j_n \leqslant m, \quad j_{k+1} - j_k \in \{1, 2, \ldots, d+1\}, \quad k = 1, \ldots, n-1, \qquad (1)$$

что $x_k = y_{j_k}$, $k = 1, \ldots, n$. В этом случае X_n является *подпоследовательностью* Y_m с допуском d. Вложение с допуском $d = 1$ в [1] названо *плотным*.

В [1] найдена верхняя оценка для вероятности того, что заданная двоичная последовательность может быть плотно вложена в последовательность независимых двоичных случайных величин с равномерными распределениями. В [2] этот результат обобщен на последовательности со значениями в любом конечном алфавите. Получены неулучшаемые нижняя и верхняя оценки для вероятности плотного вложения и указаны классы последовательностей, на которых они достигаются. Обобщение понятия плотного вложения на вложение с допуском проведено в [3]. Задача о вложениях двоичных последовательностей и её практическое применение рассмотрены в [4, 5].

Построение критерия и его свойства

Рассмотрим задачу о проверке гипотезы H_{0n} о том, что X_n является подпоследовательностью с допуском d последовательности Y_m независимых равномерно распределённых на множестве A_N случайных величин.

Самый простой способ проверки гипотезы H_{0n} состоит в том, чтобы опробовать все $(d+1)^{n-1}$ удовлетворяющих (1) наборов $(j_1, j_2, \ldots, j_n)$ мест вложения с допуском d последовательности X_n в начало последовательности Y_m. При этом вероятность отклонить гипотезу H_{0n}, если она верна, равна нулю. В [2] показано, что при $d = 1$ вероятность ошибочного принятия гипотезы H_{0n} убывает экспоненциально быстро. К очевидным недостаткам такой процедуры естественно отнести её большую вычислительную сложность.

Критерий согласия с гипотезой H_{0n}, не требующий проверки всех вариантов вложения, при $d = 1$ предложен в [6, 7]. В настоящей работе проведём обобщение этих результатов на случай произвольного $d \geqslant 2$.

Пусть $j_1 = 1$; $j_k = \min\{t > j_{k-1} : x_k = y_t\}$, $k = 2, \ldots, n$; $V_1 = 1$; $V_k = V_k(X_n) = j_k - j_{k-1}$, $k = 2, \ldots, n$; $T_k = V_2 + \ldots + V_k$.

Построим критерий $\mathcal{T}$ по следующему правилу. Последовательно по $k = 2, \ldots, n$ вычисляем значение T_k. Если $x_1 = y_1$ и на k-м шаге неравенство

$$T_k \leqslant (d+1)(k-1) \tag{2}$$

не выполнено, то гипотеза H_{0n} отклоняется. В противном случае продолжаем проверку. Если $x_1 = y_1$ и при всех $k = 2, \ldots, n$ выполнено (2), то гипотеза H_{0n} принимается.

Замечание 1. Если H_{0n} верна, то существует набор чисел $j_1, \ldots, j_n$, удовлетворяющих (1), и $x_k = y_{j_k}, k = 1, \ldots, n$. Значит, $V_k = j_k - j_{k-1} \leqslant d + 1$, $k = 2, \ldots, n$, и $T_k = V_2 + \ldots + V_k \leqslant (d+1)(k-1)$. Таким образом, вероятность ошибки первого рода критерия $\mathcal{T}$ равна нулю.

Нас интересует вероятность ошибки второго рода при альтернативной гипотезе H_{1n} о том, что последовательность X_n не зависит от последовательности Y_m и тоже состоит из независимых равномерно распределённых на множестве A_N случайных величин, а также среднее число знаков, используемых критерием до принятия решения.

Теорема 1. Вероятность ошибки второго рода критерия $\mathcal{T}$ при $n \geqslant 2$ равна

$$\mathsf{P}\{H_{0n}|H_{1n}\} = \frac{1}{N}\left(1 - \sum_{k=1}^{n-1}\sigma_k\right),$$

где последовательность чисел σ_k имеет производящую функцию

$$\sigma(s) = \sum_{k=0}^{\infty}\sigma_k s^k = 1 - \frac{1-s}{1-s/N}\exp\left\{\sum_{n=1}^{\infty}\frac{s^n}{nN^n}\sum_{m=1}^{dn}\mathrm{C}_{n+m-1}^{n-1}\left(1-\frac{1}{N}\right)^m\right\}. \tag{3}$$

Среднее число шагов до принятия решения при гипотезе H_{1n} равно

$$\frac{1}{N}\left(N - 1 + \sum_{k=1}^{n-2}(k+1)\sigma_k + n\left(1 - \sum_{k=1}^{n-2}\sigma_k\right)\right).$$

Замечание 2. Можно показать, что

$$\mathsf{P}\{V_k = l|H_{1n}\} = N^{-1}(1 - N^{-1})^{l-1},\ l \geqslant 1,\ k = 2, \ldots, n.$$

Согласно [8, теорема 2 § 2 гл. XII, с. 448–449], случайная величина с законом распределения, соответствующим производящей функции (3), является собственной, если $\mathbf{E}V_2 \geqslant d+1$, и имеет конечное математическое ожидание, если $\mathbf{E}V_2 > d+1$. Очевидно, $\mathbf{E}V_2 = N$. Значит, $\sigma(1) = 1$ при $N \geqslant d+1$ и $\sigma'(1) = 1$ при $N \geqslant d+2$.

ЛИТЕРАТУРА

1. *Golic J. Dj.* Constrained embedding probability for two binary strings // SIAM J. Discrete Math. 1996. V. 9. No. 3. P. 360–364.
2. *Михайлов В. Г., Меженная Н. М.* Оценки для вероятности плотного вложения одной дискретной последовательности в другую // Дискретная математика. 2005. Т. 17. № 3. С. 19–27.
3. *Михайлов В. Г., Меженная Н. М.* Нижние оценки для вероятности вложения с произвольным допуском // Вестник Московского государственного технического университета им. Н. Э. Баумана. Сер. Естественные науки. 2012. № 2. С. 3–11.
4. *Donovan D. M., Lefevre J., and Simpson L.* A discussion of constrained binary embeddings with applications to cryptanalysis of irregularly clocked stream ciphers // R. Balakrishnan and C. V. Madhavan (eds.) Discrete Mathematics. Proc. Intern. Conf. on Discr. Math., Indian Institute of Science, Bangalore, December 2006. P. 73–86.
5. *Kholosha A.* Clock-controlled shift registers for key-stream generation // IACR Cryptology ePrint Archive 2001: 61 (2001). http://eprint.iacr.org/2001/061.pdf
6. *Меженная Н. М.* О проверке гипотезы о плотном вложении для дискретных случайных последовательностей // Вестник БГУ. Математика, Информатика. 2017. № 4. С. 9–20.
7. *Меженная Н. М.* Предельные теоремы в задачах о плотном вложении и плотных сериях в дискретных случайных последовательностях: дис. ... канд. физ.-мат. наук. Московский государственный институт электроники и математики. М., 2009.
8. *Феллер В.* Введение в теорию вероятностей и ее приложения: в 2 т. М.: Мир, 1984. Т. 2. 751 с.

УДК 519.7 DOI 10.17223/2226308X/11/4

НИЖНЯЯ ОЦЕНКА МОЩНОСТИ НАИБОЛЬШЕГО МЕТРИЧЕСКИ РЕГУЛЯРНОГО ПОДМНОЖЕСТВА БУЛЕВА КУБА[1]

А. К. Облаухов

Исследуются строго метрические регулярные подмножества булева куба. Представлены итеративные конструкции таких множеств. Получена формула для вычисления количества строго метрически регулярных множеств, получаемых с помощью данных конструкций. Построены специальные семейства метрически регулярных множеств и вычислены мощности множеств из этих семейств. Полученные значения дают нижнюю оценку мощности наибольших метрически регулярных множеств при фиксированном радиусе покрытия.

Ключевые слова: *метрически регулярное множество, метрическое дополнение.*

Рассмотрим $\mathbb{F}_2^n$ — пространство двоичных векторов длины n. *Расстояние Хэмминга* $\mathrm{d}(x, y)$ между двумя векторами $x, y \in \mathbb{F}_2^n$ равно количеству координат, в которых эти векторы различаются.

[1]Работа поддержана грантами РФФИ, проекты № 18-31-00479 и 17-41-543364.