Ускорение вычисления тестовой функции при оценке качества решающей функции NAGTA

Н. П. Кочетова[1] А. Б. Фролов[2]

Для неадаптивного группового тестирования предложен способ ускорения вычисления тестовой функции при оценке качества порогового решающего правила принятия решения о позитивности множества образцов. Ускорение достигается использованием наряду с дуальными блоками комбинаторной блок схемы ее блоков, соответствующих элементам анализируемого множества.

1. Введение

Идея группового тестирования инициализирована еше в 1943 г. Р. Дорфманом [1] в связи задачей массового тестирования характерного заболевания по анализу крови. Целью группового тестирования является выявление т.н. позитивного множества редких позитивных образцов из заданного множества образцов по результатам тестирования определенным образом составляемых групп образцов на предмет наличия или отсутсвия позитивного образца в группе. Она получила развитие в приложениях комбинаторики [2], в частности, с использованием комбинаторных блок-схем для представления и анализа используемых групп образцов [3]. В [4] изучаются алгоритмы неадаптивного группового тестирования в условиях шумов, сопровождающих измерения при групповом тестировании. В настоящей работе рассматривается идеальная схема группового тестирования (без шумов) и решается задача ускорения вычисления тестовой функции при оценке качества пороговой функции принятия или отклонении решения о позитивности анализируемого множества образцов. Ускорение достигается использованием блоков комбинаторной блоксхемы вместо ее дуальных блоков. Рассматриваются комбинаторные линейные и квадратичные трансверсальные блок-схемы [5].

[1] *Кочетова Наталья Петровна* — студентка каф. МКМ института ИВТ МЭИ, e-mail: iNatashka99@yandex.ru.

Kochetova Natalya Petrovna — student, MPEI, Institute of Information and Computing Technologies, Department of Mathematical and Computer Modeling.

[2] *Фролов Александр Борисович* — профессор каф. МКМ института ИВТ МЭИ, e-mail: frolovab@mail.ru.

Frolov Alexander Borisovich — professor, MPEI, Institute of Information and Computing Technologies, Department of Mathematical and Computer Modeling.

2. Основные понятия и формулировка результата

Пусть X есть множество m образцов и $\mathcal{A}$ есть множество его подмножеств -тестов. Пара (X, A) называется (m, n)-NAGTA (non adaptive group algoritm) [3]. Обозначим $U,\ U \subset X$ множество позитивных образцов. Результатом группового тестирования является т.н. *тестовая функция* $R_U(A) : \mathcal{A} \to \{0, 1\}$ с областью единичных значений (*тестовым множеством*)

$$R_U^{(1)} = \{A : A \cap U \neq \emptyset\}\,. \tag{1}$$

По (1) нужно выделить подмножество U позитивных образцов. По (1) и X можно вычислить множество

$$U^* = X \setminus \cup_{A \notin R_U^{(1)}} A. \tag{2}$$

Если $|U^*| = |U|$, то по результату тестирования позитивное множество определяется однозначно: $U = U^*$, и если $U \neq U^*$, то $R_U^{(1)} \neq R_{U^*}^{(1)}$. Но в реальных условиях мощность $|U|$ позитивного множества заранее неизвестна, но можно предположить, что $|U| \leq s$, где $s \leq m$ - заданное целое число. Тогда $(X, \mathcal{A})$ называется $(X, \mathcal{A})$ - NAGTA с порогом s, если $R_U \neq R_V$ всякий раз, когда $U, V \subseteq X, |U| \leq s, |V| \leq s$, и $U \neq V$ [3]. Таким образом, практически используется пороговое правило принятия решения о позитивности полученного по результату тестирования множества U^*, определяемое предикатом $|U^*| \leq s$. Для оценки качества этого правила применительно к конкретной $(X, \mathcal{A})$ - NAGTA с порогом s предлагается использовать экспертные решения для каждого множества $U,\ |U| \leq s$, определяемые предикатом $|U^*| = |U|$. Тогда можно характеризовать решения порогового правила по каждому такому множеству: tp, если $|U^*| \leq s$ и $|U^*| = |U|$; fp, если $|U^*| \leq s$ и $|U| < |U^*|$; fn, если $|U^*| > s$ и $|U^*| = |U|$; tn, если $|U^*| > s$ и $|U| < |U^*|$;

При этом качество пороговой функции определится по количествам n_{tp}, n_{fp}, n_{fn}, n_{tn} таких характеристик, полученных по всем множествам $U, |U| \leq s$ и для каждого из них придется вычислять тестовое множество $R_U^{(1)}$.

Замена индивидуального тестирования групповым тестированием целесообразна только при выполнении неравенства: $|\mathcal{A}| < |X|$. Такие комбинаторные схемы можно встретить в классе систем троек Штейнера или комбинаторных трансверсальных дизайнов.

Линейная трансверсальная комбинаторная блок-схема $TD(2, k, n)$ строится на множестве Y из kn элементов, имеет n^2 блоков $B(a), a \in X$, (множество $\mathcal{B}, |\mathcal{B}| = |X|$) по k элементов и kn дуальных блоков по n элементов (множество $\mathcal{A}, |\mathcal{A}| = |Y|$) [5].

Квадратичная трансверсальная комбинаторная блок-схема $TD(3,k,n)$ строится на множестве Y из kn элементов, имеет n^3 блоков $B(a), a \in X$, (множество $\mathcal{B}, |\mathcal{B}| = |X|$) по k элементов и kn дуальных блоков по n^2 элементов (множество $\mathcal{A}, |\mathcal{A}| = |Y|$) [5].

Если в качестве множества $\mathcal{A}$ принять множество дуальных блоков, то для вычисления тестового множества (1) в случае линейной трансверсальной блок-схемы потребуется выполнить $|U|kn \log_2 n$ операций в поле $GF(n)$ и вдвое больше в случае квадратичной. Это соответствует способу вычисления тестового множества непосредственно по его определению с использованием множества дуальных блоков.

Альтернативным является вычисление тестового множества с построением блоков. Множество (1) получается объединением $|U|$ блоков $B(a)$:

$$R_U^{(1)} = \cup_{a \in U} B(a). \tag{3}$$

Всего при вычислении по (3) затрачивается не более $|U|(t(k-1) + klog_2 kn) = |U|(k(log_2 kn + t) - t)$ операций.

Вычислениями по (3) и далее по (2) с использованием блоков и дуальных блоков трансверсальных дизайнов TD(t,k,n) экспериментально подтверждены данные об их применении, приведенные в Таблице 1.

Таблица 1.

Дизайн	kn (Число групповых тестов)	b (число образцов)	r (число элементов в группе)	Порог s	n_{tp}%	n_{tn}%	n_{fn}%
TD(3,3,4)	12	64	16	1	1	0	0
TD(2,2,8)	16	64	8	1	1	0	0
TD(2,3,8)	24	64	8	2	1	0	0
TD(3,4,8)	32	512	64	2	73	27	0

Если s увеличить на 1, то все числа будут $n_{tp} = 0$, $n_{tn} = 1$, при $s = 1$ для TD(3,4,8) значения $n_{tp} = 1$, $n_{tn} = 0$.

Применительно к TD(3,4,8) вычислительным экспериментом подтверждено, что построенное по тестовому множеству множество из $s = 2$ элементов является множеством положительных образцов. Такими оказываются 73 процентов множеств: $n_{tp} = 93777, n_{tn} = 33649$. По тестовым множествам, построенным по любому из остальных позитивных множеств мощности s, получаются множества, содержащие кроме элементов исходного позитивного множества еще и другие элементы. Добавление таких элементов к исходному множеству не изменяет тестовое множество.

Список литературы

[1] R. Dorfman, “The detection of defective members of large populations”, *Annals of Mathematical Statistics*, **14** (1943), 436–411.

[2] D.-Z. Du and F. K. Hwang, “Combinatorial Group Testing and Its Applications”, *2nd ed. World Scientific Publishing Company*, 2000.

[3] D. Stinson, *Combinatorial designs: constructions and analysis*, Springer, Berlin, Germany, 2003.

[4] DChun Lam Chan, Sidharth Jaggi, Venkatesh Saligrama, Samar Agnihotri, “Non-adaptive Group Testing: Explicit bounds and novel algorithms”, *2012 IEEE International Symposium on Information Theory Proceedings*, 2012.

[5] Paterson M. B., Stinson D. R., “Unied Approach to Combinatorial Key Predistribution Schemes for Sensor Networks”, *Designs, Codes and Cryptography*, **71**:3 (June 2014), 433–457.

Speeding up the computing of the test function when assessing the quality of the NAGTA threshold function

Kochetova N.P., Frolov A.B.

For non-adaptive group testing, a method is proposed to accelerate the computing of the test function when assessing the quality of the threshold decision rule for making a decision on the positivity of a set of samples. Acceleration is achieved by using, along with dual blocks, a combinatorial block diagram of its blocks corresponding to the elements of the analyzed set.

References

[1] R. Dorfman, “The detection of defective members of large populations”, *Annals of Mathematical Statistics*, **14** (1943), 436–411.

[2] D.-Z. Du and F. K. Hwang, “Combinatorial Group Testing and Its Applications”, *2nd ed. World Scientific Publishing Company*, 2000.

[3] D. Stinson, *Combinatorial designs: constructions and analysis*, Springer, Berlin, Germany, 2003.

[4] DC. L. Chan, S. Jaggi, V. Saligrama, S. Agnihotri, “Non-adaptive Group Testing: Explicit bounds and novel algorithms”, *2012 IEEE IS on IT Proc.*, 2012.

[5] Paterson M. B., Stinson D. R., “Unied Approach to Combinatorial Key Predistribution Schemes for Sensor Networks”, *Designs, Codes and Cryptography*, **71**:3 (June 2014), 433–457.