

УДК 519.6

Использование интерпретируемого классификатора для метода логического анализа данных для идентификации электрорадиоизделий

Е.Е. Машинец^{1,*}, Р.И. Кузьмич², А.А. Ступина¹, И.А. Гайдук¹

¹ Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнёва, просп. им. газ. «Красноярский рабочий», 31, Красноярск, 660037, Россия

² Сибирский федеральный университет, пр. Свободный, 79, Красноярск, 660041, Россия

*E-mail: ris2005@mail.ru

Аннотация. Электронная компонентная база (ЭКБ), применяемая при комплектации бортовой аппаратуры космических аппаратов (КА), должна соответствовать следующим специфическим характеристикам. Космические аппараты выполняют разнообразные функциональные задачи, что приводит к необходимости использования широкой номенклатуры типоминиатур ЭКБ при крайне небольших размерах партий ЭКБ. При производстве широкой функциональной номенклатуры ЭКБ требуется применение разнообразных технологий, многие из которых уникальны и отсутствуют в Российской Федерации. К сверхжестким требованиям по надежности и стойкости (к дестабилизирующим факторам космического пространства) приводят и жесткие требования к сроку активного существования при полном отсутствии возможностей ремонта. Поскольку в Российской Федерации отсутствует специализированное производство ЭКБ для использования в ракетно-космической технике, комплектация бортовой аппаратуры производится посредством применения ЭКБ иностранного производства и комплектации бортовой аппаратуры ЭКБ отечественного производства через специализированные испытательные технические центры с проведением операций сплошного входного контроля ЭКБ, дополнительных отбраковочных испытаний, диагностического неразрушающего контроля и с применением выборочного разрушающего физического анализа. Для задачи отбора электронной компонентной базы при комплектации бортовой аппаратуры КА показано применение интерпретируемого классификатора для метода логического анализа данных.

Ключевые слова: классификация, метод логического анализа, признак, закономерность

Using an interpreted classifier for a logical data analysis method for identifying semiconductor devices

Е.Е. Mashinec^{1,*}, R.I. Kuzmich², A.A. Stupina¹, I.A. Haiduk¹

¹ Reshetnev Siberian State University of Science and Technology, 31, Krasnoyarsky Rabochy Av., Krasnoyarsk, 660037, Russia

² Siberian Federal University, 79, Svobodny pr., Krasnoyarsk, 660041, Russia

*E-mail: ris2005@mail.ru

Abstract. The electronic component base (ECB) used to complete the onboard equipment of spacecraft (SC) must comply with the following specific characteristics. Spacecraft perform a variety of functional tasks, which leads to the need to use a wide range of ECB standard types with extremely small sizes of ECB batches. The production of a wide functional range ECB requires the use of a variety of technologies, many of which are unique and absent in the Russian Federation. Strict requirements for the period of active existence in the complete absence of repair possibilities lead to super-stringent requirements for reliability and resistance (to destabilizing factors of outer space). Since in the Russian Federation there is no specialized production of electronic components for use in rocket and space technology, the onboard equipment is completed by using foreign-made electronic components and completing on-board equipment for electronic components of domestic production through specialized test technical centers with continuous incoming control of electronic components, additional screening tests, diagnostic non-destructive testing and using selective destructive physical analysis. For the problem of selecting an electronic component base when completing the spacecraft onboard equipment, the use of an interpreted classifier for the method of logical data analysis is shown.

Keywords: classification, method of logical analysis, sign, pattern

Задача комплектации электронной компонентной базы для использования в ракетно-космической технике решается посредством применения электрорадиоизделий иностранного и отечественного производства через специализированные испытательные технические центры с обязательным проведением ряда специфических операций по отбраковочным испытаниям [1]. Для повышения точности отбора потенциально ненадежных изделий электронной компонентной базы космического применения, был предложен подход к построению интерпретируемого классификатора для метода логического анализа данных [3].

Дано два непересекающихся множества Ω^+ и Ω^- n -мерных векторов некоторой выборки данных, принадлежащих положительному или отрицательному классу. Число бинарных признаков, после этапа бинаризации [4], может быть велико, поэтому для исключения избыточных переменных в исходной выборке данных во множестве переменных определяется некоторое подмножество S . Используя данное подмножество можно с высокой точностью различить положительные наблюдения от отрицательных, после чего используются проекции Ω_s^+ и Ω_s^- множеств Ω^+ и Ω^- на S [5].

Рассматриваемый метод основывается на понятии закономерности. Положительной закономерностью называется подкуб пространства булевых переменных B_2^t , который пересекается с множеством Ω_s^+ и может иметь некоторое ограниченное число общих элементов с множеством Ω_s^- . Аналогично задается и отрицательная закономерность.

Положительная α -закономерность для $\alpha \in \{0,1\}^t$ – это закономерность, содержащая в себе точку α . Для каждой точки $\alpha \in \Omega_s^+$ найдем максимальную α -закономерность, то есть покрывающую наибольшее число точек Ω_s^+ . Соответствующую α -закономерность зададим с помощью бинарных переменных $Y=(y_1, y_2, \dots, y_t)$ (уравнение 1):

$$y_j = \begin{cases} 1, & \text{если } i\text{-ый признак фиксирован в правиле} \\ 0, & \text{в противном случае.} \end{cases} \quad (1)$$

Существующее условие, говорящее о том, что положительная закономерность не должна покрывать ни одного отрицательного наблюдения, требует, чтобы для каждого наблюдения $\beta \in \Omega_s^-$ переменная y_j принимала значение 1 по меньшей мере для одного j , для которых $\beta_j \neq \alpha_j$ (уравнение 2):

$$\sum_{\substack{j=1 \\ \beta_j \neq \alpha_j}}^t y_j \geq 1 \text{ для любого } \beta \in \Omega_s^- \quad (2)$$

Для повышения устойчивости к ошибкам, возможно, усилить ограничение путем замены числа 1 в правой части неравенства на целое положительное число d .

С другой стороны, положительное наблюдение $\sigma \in \Omega_s^+$ будет тогда входить в рассматриваемую закономерность, когда переменная y_j принимает значение 0 для всех индексов j , для которых $\sigma_j \neq \alpha_j$. Таким образом, суммарное количество положительных наблюдений, покрываемых α -закономерностью, может быть вычислено как (уравнение 3):

$$\sum_{\sigma \in \Omega_s^+} \prod_{\substack{j=1 \\ \sigma_j \neq \alpha_j}}^t (1 - y_j) \quad (3)$$

В итоге, имеем задачу условной псевдодобулевой оптимизации с алгоритмически заданными функциями [2, 3] (уравнения 4 и 5):

$$\sum_{\sigma \in \Omega_s^+} \prod_{\substack{j=1 \\ \sigma_j \neq \alpha_j}}^t (1 - y_j) \rightarrow \max, \quad (4)$$

$$\sum_{\substack{j=1 \\ \beta_j \neq \alpha_j}}^t y_j \geq d \text{ для любого } \beta \in \Omega_s^-, y \in \{0,1\}^t \quad (5)$$

Задача нахождения максимальных отрицательных закономерностей формулируется аналогично. Для повышения устойчивости метода к выбросам следует перейти к построению частичных закономерностей [2, 3], то есть закономерностей, которые могут покрывать некоторое число наблюдений другого класса. Степень частичных закономерностей, как правило, меньше, чем у чистых закономерностей, а покрытие больше.

Отметим, что правила, покрывающие мало наблюдений, статистически не надежны – среди них слишком много правил, допускающих на экзаменующей выборке больше ошибок, чем на обучающей выборке. Поэтому предлагается подход, заключающийся в редукции исходного классификатора. Основная идея состоит в построении классификатора только из самых информативных закономерностей [3]. В результате, применение такого подхода приведет к редукции классификатора без потери его качества классификации.

Для проведения испытаний использовалась выборка данных, состоящая из 70 электрорадиоизделий 1526TL1 без отклонения от заданных характеристик (положительные наблюдения) и 70 электрорадиоизделий 1526TL1 с отклонением (отрицательные наблюдения). Результаты испытаний приведены в таблице 1.

Таблица 1. Сравнение результатов электрорадиоизделий 1526TL1.

Задача оптимизации	Множество правил	Количество правил	Средняя информативность	Покрытие отр. правил	Покрытие пол. правил	Точность классификации, %
Целевая функция (ограничение)	Отр.	61	0,34	36	11	91
	Пол.	53	0,38	10	29	84
Целевая функция (ограничение с процедурой наращивания правил)	Отр.	60	0,43	36	7	91
	Пол.	52	0,45	7	29	84
Целевая функция (ограничение с процедурами наращивания правил и редукции классификатора)	Отр.	22	0,47	39	7	91
	Пол.	21	0,54	6	34	90

Согласно полученным результатам, можно отметить, что подход, основанный на совместном использовании алгоритмических процедур наращивания закономерностей и редукции исходного классификатора, позволяет получить интерпретируемый классификатор с высокой точностью распознавания новых наблюдений, что соответственно повышает точность отбора потенциально ненадежных изделий электронной компонентной базы космического применения.

Список литературы

1. Рожнов, И.П. Алгоритмы автоматической группировки с повышенными требованиями к точности и стабильности результата: монография / И.П. Рожнов, Л.А. Казаковцев, В.И. Орлов, Д.Л. Михнёв. – Москва : ИНФРА-М, 2020. – 192 с. ISBN 978-5-16-016641-4.
2. Кузьмич, Р.И. Модели формирования закономерностей в методе логического анализа данных / Р.И. Кузьмич, И.С. Масич, А.А. Ступина // Системы управления и информационные технологии. – 2017. – №1(67). – С. 33-37.
3. Об одном подходе к построению интерпретируемого классификатора для метода логического анализа данных / Р.И. Кузьмич, А.А. Ступина, В.А. Соколов, Е.Е.

- Машинец // Системы управления и информационные технологии. – 2020. – № 2(80). – С. 34-38.
4. Растрингин, Л.А. Решение задач разношкальной оптимизации методами случайного поиска / Л.А. Растрингин, Э.Э. Фрейманис // Проблемы случайного поиска. – 1988. – № 11. – С. 9-25.
 5. Hammer, P.L. Logical Analysis of Data: From Combinatorial Optimization to Medical Applications / P.L. Hammer, T. Bonates // RUTCOR Research Report 10-2005. –2005.