

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

DOI: 10.18137/RNU.V9I187.22.03.P.003

УДК 681.31

А.Г. Басыров, В.В. Кузнецов, М.Н. Лупашко

АЛГОРИТМЫ РЕСУРСОСБЕРЕГАЮЩЕГО АВТОНОМНОГО КОНТРОЛЯ МАССИВА ДАННЫХ НАБЛЮДЕНИЯ В БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТАХ

Аннотация. Представлены и исследованы алгоритмы анализа качества блоков данных, применяемые в рамках подхода к организации распределенных информационных процессов, реализуемых в условиях временных и энергетических ограничений. Предложен показатель качества ресурсосберегающего анализа данных. Проведено имитационное моделирование процесса анализа качества блоков данных предложенными алгоритмами.

Ключевые слова: анализ качества информации, ресурсосберегающий информационный процесс, предварительная обработка данных.

A.G. Basyrov, V.V. Kuznetsov, M.N. Lupashko

ALGORITHMS FOR RESOURCE-SAVING AUTONOMOUS MONITORING OF AN ARRAY OF SURVEILLANCE DATA IN UNMANNED AERIAL VEHICLES

Abstract. The article presents and investigates algorithms for analyzing the quality of data blocks used as part of the approach to organizing distributed information processes implemented under time and energy constraints. An indicator of the quality of resource-saving data analysis is proposed. Simulation modeling of the process of analyzing the quality of data blocks by the proposed algorithms is carried out.

Keywords: information quality analysis, resource-saving information process, preprocessing of data.

Введение

Современные системы распределенного сбора информации, как правило, функционируют в условиях жестких энергетических и временных ограничений. Собранная беспилотными летательными аппаратами информация должна передаваться в центр, где она обрабатывается и затем используется по целевому назначению. Если объем собранной информации незначителен, то ее можно передать без предварительной обработки. Если же передача информации требует значительных энергетических затрат, и велика вероятность отбраковки большого количества данных, то целесообразно предварительно обработать собранную информацию вычислительным устройством на борту летательного аппарата для сокращения ее объема, например, путем удаления незначущих или непригодных для дальнейшего использования данных.

Если собранная информация представляет собой набор блоков данных (далее – БД), то ее предварительная обработка заключается в последовательном анализе и отбраковке

Басыров Александр Геннадьевич

доктор технических наук, профессор, начальник кафедры информационно-вычислительных систем и сетей. Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург. Сфера научных интересов: информационные вычислительные системы, высокопроизводительная обработка информации. Автор более 150 опубликованных научных работ.

Электронный адрес: vka_24kaf_1@mil.ru

Кузнецов Вадим Викторович

кандидат технических наук, доцент, доцент кафедры информационно-вычислительных систем и сетей. Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург. Сфера научных интересов: компьютерные технологии, надежность программного обеспечения. Автор более 30 опубликованных научных работ.

Электронный адрес: vka_24kaf_1@mil.ru

Лунашко Максим Николаевич

начальник 553-го Военного представительства Министерства обороны РФ. Сфера научных интересов: информационные вычислительные системы. Автор более 10 опубликованных научных работ.

Электронный адрес: vka_24kaf_1@mil.ru

непригодных блоков. Кроме того, возможны и другие действия над БД, например, распознавание, сжатие, фильтрация [5; 6].

Предварительная обработка данных на борту беспилотного летательного аппарата требует определенного временного и энергетического ресурса; при этом, однако, можно сократить затраты времени и энергии на передачу больших массивов данных.

Проблема заключается в определении подхода к ограниченному анализу собранных данных в условиях неопределенности их пригодности к дальнейшему использованию.

Используемые методы контроля на основе ускоренных испытаний [3; 4] в рассматриваемой постановке задачи неприемлемы из-за малого размера партии продукции, отсутствия статистики по отказам и существенным ограничениям на время и ресурсы контроля БД.

Таким образом, возникает задача формализации показателя качества ресурсосберегающего анализа данных, разработки алгоритмов анализа качества данных и анализа эффективности их применения.

Показатель качества ресурсосберегающего анализа данных

Показатель качества ресурсосберегающего анализа данных должен учитывать результативность и ресурсоемкость процесса такого анализа. Результативность оценивается долей выявленных дефектных БД, а ресурсоемкость – количеством проверок.

Показатель качества ресурсосберегающего анализа данных определим выражением

$$k = \frac{n_{\text{дп}}}{n_{\text{д}}} \left(1 - \frac{n_{\text{п}}}{N} \right), \quad (1)$$

где $n_{\text{п}}$ – количество проверенных БД ($n_{\text{п}} \leq N$); $n_{\text{д}}$ – количество дефектных БД ($n_{\text{д}} \geq 1$); $n_{\text{дп}}$ – количество выявленных в результате проверки дефектных БД ($n_{\text{дп}} \leq n_{\text{д}}$, $n_{\text{дп}} \leq n_{\text{п}}$); N – общее количество БД.

При этом свойствами показателя качества являются:

– область значений $[0,1]$;

– все параметры $(n_{дп}, n_{д}, n_{п}, N)$ – целые положительные числа.

Взаимосвязь параметров $(n_{дп}, n_{д}, n_{п}, N)$ проиллюстрирована на Рисунке 1.

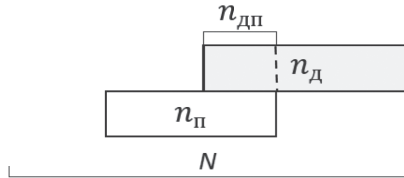


Рисунок 1. Взаимосвязь параметров предварительной обработки данных

Физический смысл показателя (1) состоит в произведении результативности $\frac{n_{дп}}{n_{д}}$ поиска дефектных БД на дополнительную величину затрат $\frac{n_{п}}{N}$, то есть экономию, связанную с этим поиском.

Рассмотрим следующие типовые варианты размещения дефектных блоков в памяти:

- дефектные БД распределены равномерно в памяти;
- дефектные БД распределены компактно в некоторой области памяти;
- дефектные БД распределены компактно в нескольких областях памяти.

Пример указанных вариантов размещения в памяти четырех дефектных блоков из 20 представлен на Рисунке 2. Эти варианты имеют место на практике и могут быть связаны как со сбоями аппаратно-программного обеспечения, так и особенностями сбора целевой информации.

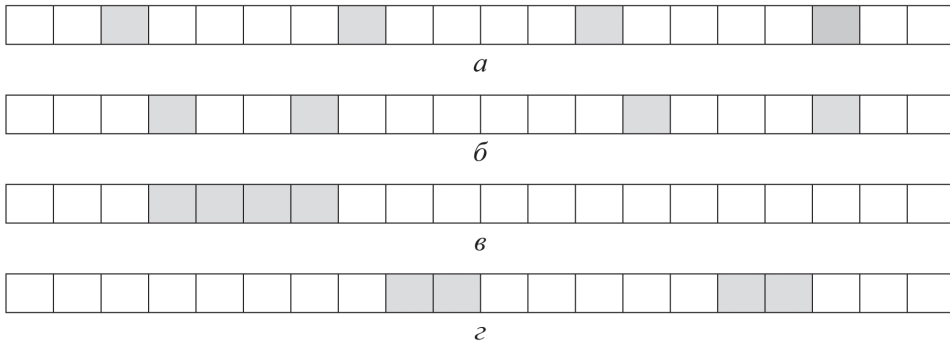


Рисунок 2. Варианты размещения дефектных блоков данных в памяти: *a* – равномерное распределение дефектных блоков в памяти; *б* – случайное распределение дефектных блоков в памяти; *в* – компактное распределение дефектных блоков в одной области памяти; *г* – компактное распределение дефектных блоков в нескольких областях памяти

Предложенный показатель качества может использоваться для анализа эффективности применения алгоритмов ресурсосберегающего анализа качества данных.

Алгоритмы ресурсосберегающего анализа качества данных

Рассмотрим следующие варианты порядка проверки, то есть предварительной обработки БД в памяти:

– последовательно – блоки проверяются поочередно в порядке их размещения в памяти;
 – в случайном порядке – номер очередного проверяемого блока формируется по равномерному закону распределения чисел;

– по правилу половинного деления номеров блоков: номер первого БД определяется как целая часть от $\left\lceil \frac{N}{2} \right\rceil$; номер второго – как целая часть от $\left\lceil \frac{N}{4} \right\rceil$; третьего – как целая часть от $\left\lceil \frac{3N}{24} \right\rceil$; четвертого – как целая часть от $\left\lceil \frac{N}{8} \right\rceil$; пятого – как целая часть от $\left\lceil \frac{3N}{8} \right\rceil$; шестого – как целая часть от $\left\lceil \frac{5N}{8} \right\rceil$; седьмого – как целая часть от $\left\lceil \frac{7N}{8} \right\rceil$ и т. д.

Каждому варианту соответствует свой алгоритм предварительного анализа качества данных, представленный ниже.

Алгоритм 1. Последовательный анализ блоков данных.

Шаг 1. Начало.

Шаг 2. $i := 1$.

Шаг 3. Проверка пригодности i -го БД.

Шаг 4. $i := i + 1$.

Шаг 5. Если $i > N$, то переход на шаг 6, иначе – на шаг 3.

Шаг 6. Конец.

Алгоритм 2. Анализ блоков данных в случайном порядке.

Шаг 1. Начало.

Шаг 2. Сформировать массив G номеров БД $g_j = j, g_j \in G, j = \overline{1, N}$.

Шаг 3. $n := N$.

Шаг 4. На основе датчика случайных чисел сгенерировать целое число b в интервале от 1 до n .

Шаг 5. Проверка пригодности b -го БД.

Шаг 6. Удалить из массива G элемент g_b .

Шаг 7. $n := n - 1$.

Шаг 8. Если $n < 1$, то переход на шаг 9, иначе – на шаг 4.

Шаг 9. Конец.

Алгоритм 3. Анализ блоков данных по правилу половинного деления их номеров.

Шаг 1. Начало.

Шаг 2. $c := 1$.

Шаг 3. $i := 2$.

Шаг 4. $j := 1$.

Шаг 5. $b := \lceil j \cdot N / i \rceil$.

Шаг 6. Проверка пригодности b -го БД.

Шаг 7. $c := c + 1$.

Шаг 8. Если $c > N$, то переход на шаг 12, иначе – на шаг 9.

Шаг 9. $j := j + 2$.

Шаг 10. Если $j > i$, то переход на шаг 11, иначе – на шаг 5.

Шаг 11. $i := i \cdot 2$. Переход на шаг 4.

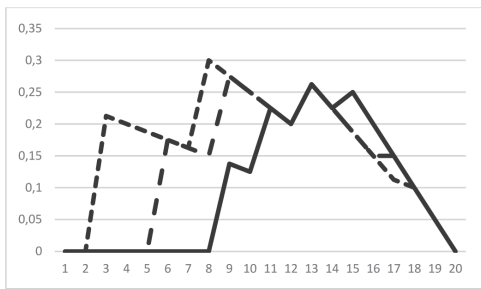
Шаг 12. Конец.

Достоинством представленных алгоритмов является их простота, а с точки зрения вычислительной сложности они являются линейными.

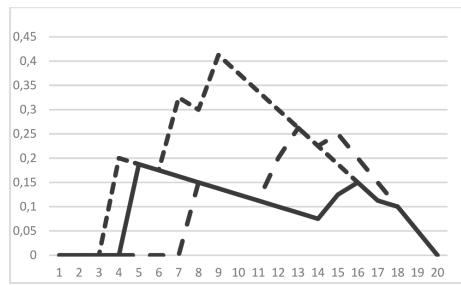
На Рисунке 3 представлены зависимости значения показателя (1) от количества проверок пригодности БД для трех описанных алгоритмов при расположении четырех дефектных БД из 20, изображенных на Рисунке 2.

Изображение вариантов размещения БД на графиках (см. Рисунок 3) имеет вид

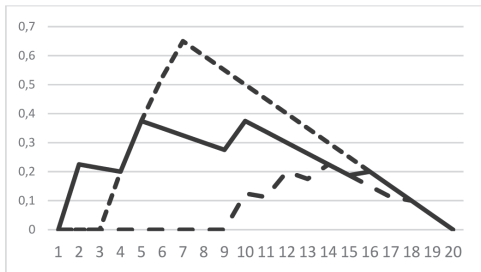
- — последовательный анализ (алгоритм 1);
- — анализ в случайном порядке (алгоритм 2);
- — анализ половинным делением номеров БД (алгоритм 3).



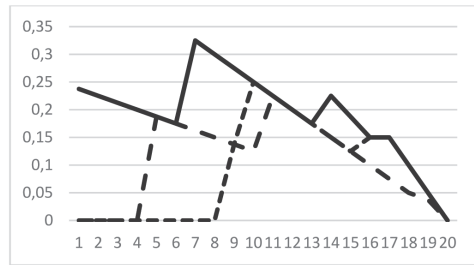
а



б



в



г

Рисунок 3. Пример зависимости значения показателя (1) от количества проверок для различных вариантов размещения дефектных блоков данных: а – равномерное (см. Рисунок 2, а); б – случайное (см. Рисунок 2, б); в – компактное распределение дефектных блоков в одной области памяти (см. Рисунок 2, в); г – компактное распределение дефектных блоков в нескольких областях памяти (см. Рисунок 2, г)

Из представленных графиков видно, что для большинства вариантов размещения дефектных БД преимуществом по показателю (1) обладает алгоритм анализа БД по правилу половинного деления их номеров.

Результаты имитационного моделирования процессов анализа качества данных

Для оценивания эффективности применения алгоритмов предварительной обработки БД была построена имитационная модель процесса анализа пригодности БД.

Имитационное моделирование заключается в следующем.

Генерируются все возможные варианты массива пригодных и дефектных БД заданного размера N .

Каждым из рассматриваемых алгоритмов предварительной обработки БД проводится анализ пригодности сгенерированных БД. На каждом j -м прогоне модели рассчитывается значение функционала

$$\varphi_i = \frac{1}{N} \sum_{i=1}^N i \frac{n_{\text{дп}}}{n_{\text{д}}} \left(1 - \frac{i}{N}\right). \quad (2)$$

После выполнения всех η итераций моделирования оцениваются два статистических параметра функционала (2):

– среднее значение

$$\bar{\varphi} = \frac{1}{\eta} \sum_{j=1}^{\eta} \varphi_j;$$

– частота выборки для каждого k -го алгоритма

$$\bar{h}_k = \frac{1}{\eta} \sum_{j=1}^{\eta} h_{kj},$$

где $h_{kj} = 1$, если $h_{kj} = \max(h_{1j}, h_{2j}, \dots, h_{\alpha j})$; $h_{kj} = 0$ – в противном случае; $j = \overline{1, \eta}$, $k = \overline{1, \alpha}$, α – количество применяемых алгоритмов предварительной обработки БД.

Исходные данные модели: доля μ дефектных БД в общем множестве БД и количество η итераций моделирования.

Результаты имитационного моделирования представлены на Рисунке 4.

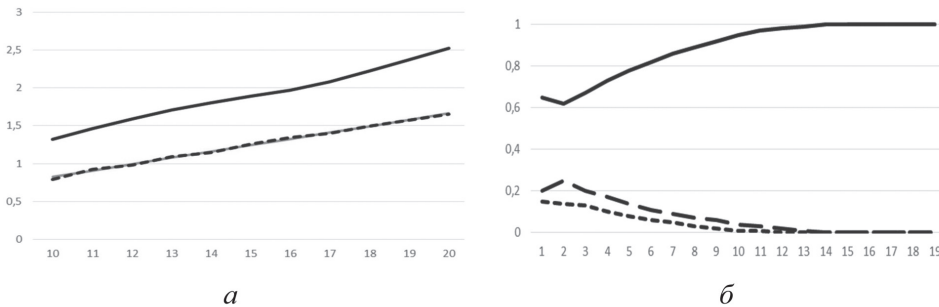


Рисунок 4. Зависимости характеристик показателя качества от количества дефектных БД для $N = 20$: а – зависимость $\bar{\varphi}$ от количества дефектных БД; б – зависимость \bar{h}_k от количества дефектных БД

Представленные зависимости наглядно демонстрируют преимущество алгоритма анализа БД по правилу половинного деления их номеров.

Заключение

Анализ проведенных исследований показывает, что рассмотренные алгоритмы не покрывают все возможные подходы к выбору последовательности анализируемых БД. При этом предложенный показатель качества ресурсосберегающего анализа данных и имитационная модель анализа качества данных могут быть использованы для оценивания эффективности применения любого алгоритма анализа БД.

Кроме того, выбор алгоритма анализа значительно упрощается, если известен характер размещения дефектных БД в памяти устройства.

Из представленных алгоритмов в условиях полной неопределенности размещения дефектных БД в памяти наиболее результативным является алгоритм, определяющий последовательность анализа БД на основе половинного деления их номеров.

Следует отметить, что для обеспечения ресурсосберегающего анализа данных необходимо определить не только порядок (алгоритм) проверки БД, но и найти оптимальное количество проверок в условиях неопределенности качества БД. Этот выбор связан с обоснованием одной из двух стратегий: первая подразумевает, что после проверки некоторого подмножества БД остальные БД считаются дефектными, вторая заключается в проверке некоторого подмножества БД и принятии гипотезы, что все остальные БД являются пригодными. Первая стратегия связана с риском забраковать пригодные БД (ошибка первого рода), вторая – допускает пропуск брака (ошибка второго рода) [4].

Выбор той или иной стратегии представляет собой отдельную задачу [1; 2].

Литература

1. Басыров А.Г., Кузнецов В.В., Лупашко М.В. Методика выбора структуры специализированной вычислительной системы автономных подвижных объектов, реализующих предварительную обработку информации // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. 2021. Вып. 4. С. 97–103.
2. Басыров А.Г., Лупашко М.Н., Степанов И.В. Стратегия организации информационных процессов в автономных подвижных объектах на основе минимизации энерговременных затрат // Интеллектуальные технологии на транспорте. 2021. № 2. С. 45–48.
3. ГОСТ Р 27.607–2013. Надежность в технике. Управление надежностью. Условия проведения испытаний на безотказность и статистические критерии и методы оценки их результатов. М.: Стандартинформ, 2015. 46 с.
4. ГОСТ Р 50779.10–2000. Статистические методы. Вероятность и основы статистики. Термины и определения. М.: Стандартинформ, 2005. 46 с.
5. Коваль Ю.А., Филиппов М.В. Метод предварительной фильтрации изображений для повышения точности распознавания образов // Инженерный журнал: наука и инновации. 2014. Вып. 12. URL: <http://engjournal.ru/catalog/it/hidden/1307.html>
6. Левенец А.В., Нильга В.В. Алгоритм предварительной обработки измерительных данных для задач их сжатия // Ученые заметки ТОГУ. 2013. Т. 4, № 4. С. 912–918. URL: http://ejournal.khstu.ru/media/2013/TGU_4_180.pdf
7. РД 50-424–83. Методические указания. Надежность в технике. Ускоренные испытания. Основные положения. М.: Издательство стандартов, 1984. 12 с.

References

1. Basyrov A.G., Kuznecov V.V., Lupashko M.V. (2021) *Metodika vybora struktury specializirovannoj vychislitel'noj sistemy avtonomnyh podvizhnyh ob'ektov, realizujushhih predvaritel'nuju obrabotku informacii* [Technique for choosing the structure of a specialized computer system for autonomous mobile objects that implement information preprocessing]. *Vestnik Rossijskogo novogo universiteta. Serija: Slozhnye sistemy: modeli, analiz i upravlenie*, vol. 4, pp. 97– 103 (in Russian).

2. Basyrov A.G., Lupashko M.N., Stepanov I.V. (2021) *Strategija organizacii informacionnyh processov v avtonomnyh podvizhnyh ob#ektah na osnove minimizacii jenergovremennyh zatrat* [Strategy for organizing information processes in autonomous moving objects based on minimizing energy and time costs]. *Intellektual'nye tehnologii na transporte*, No. 2, pp. 45–48 (in Russian).
3. GOST R 27.607–2013. *Nadezhnost' v tehnike. Upravlenie nadezhnost'ju. Uslovija provedenija ispytanij na bezotkaznost' i statisticheskie kriterii i metody ocenki ih rezul'tatov* [GOST R 27.607–2013. Reliability in technology. Reliability management. Conditions for conducting tests for reliability and statistical criteria and methods for evaluating their results]. Moscow, Standartinform Publishing, 46 p. (in Russian).
4. GOST R 50779.10–2000. [GOST R 50779.10–2000. Statistical methods. Probability and bases of statistics. Terms and Definition]. Moscow, Standartinform Publishing, 46 p. (in Russian).
5. Koval' Ju.A., Filippov M.V. (2014) *Metod predvaritel'noj fil'tracii izobrazhenij dlja povyshenija tochnosti raspoznavanija obrazov* [An image pre-filtering method to improve the accuracy of image recognition]. *Inzhenernyj zhurnal: nauka i innovacii*, vol. 12. Available at: <http://engjournal.ru/catalog/it/hidden/1307.html>
6. Levenec A.V., Nil'ga V.V. (2013) *Algoritm predvaritel'noj obrabotki izmeritel'nyh dannyh dlja zadach ih szhatija* [Algorithm for preliminary processing of measurement data for the problems of their compression]. *Uchenye zametki TOGU*, vol. 4, No. 4, pp. 912–918 (in Russian).
7. RD 50-424–83. *Metodicheskie ukazanija. Nadezhnost' v tehnike. Uskorennye ispytanija. Osnovnye polozenija* [Methodical instructions. Reliability in technology. Accelerated testing. Basic provisions]. Moscow, Izdatel'stvo standartov, 12 p. (in Russian).