

**МОДЕЛЬ РАСПРЕДЕЛЕННОГО  
ХРАНЕНИЯ ДАННЫХ В КЛАСТЕРАХ  
ПЕРСПЕКТИВНЫХ КОСМИЧЕСКИХ  
АППАРАТОВ ДИСТАНЦИОННОГО  
ЗОНДИРОВАНИЯ ЗЕМЛИ**

**MODEL OF DISTRIBUTED  
DATA STORAGE  
IN PERSPECTIVE REMOTE  
SENSING SATELLITE  
CLUSTERS**

*Рассмотрена проблема организации распределенного хранения данных в кластерах космических аппаратов дистанционного зондирования Земли (ДЗЗ) в условиях ограниченности возможностей бортовой аппаратуры и доступности аппаратов.*

*Предложен подход к построению распределенной системы хранения данных в кластере космических аппаратов на основе восстанавливающего кодирования.*

*Представлена модель распределённого хранения, позволяющая оценивать вероятность успешного восстановления результатов наблюдения кластером ДЗЗ с учетом доступности аппаратов ДЗЗ в конкретный момент. Также в модели учтена вероятность начала сбора данных космическим аппаратом (сервером или ретранслятором) до момента окончания распределения данных внутри кластера.*

**Ключевые слова:** *распределённые системы хранения, вероятность успешного восстановления, кластеры космических аппаратов.*

*In the article is considered a problem of maintaining distributed data storage in perspective remote sensing satellite clusters. It takes in to account the limited onboard resources and availability of such satellites.*

*An approach to the distributed data storage based on the erasure codes is proposed.*

*The model of distributed storage which allows to estimate the probability of successful remote sensing data reconstruction at cluster is presented. It takes into account the probability of satellites availability and probability of beginning data collection with satellite-server or satellite-repeater before end of data distribution.*

**Keywords:** *distributed data storage systems, probability of successful reconstruction, clusters of satellites.*

**Введение**

С каждым годом растет спрос на данные, полученные от космических средств дистанционного зондирования Земли (ДЗЗ). Например, в 2014 году на орбиты вокруг Земли запущено рекордное число космических аппаратов (КА) с аппаратурой съемки Земли: 26 спутников по действующим гражданским программам дистанционного зондирования Земли (ДЗЗ). В основном

<sup>1</sup> Адъюнкт кафедры информационно-вычислительных систем и сетей Военно-космической академии им. А.Ф. Можайского.

© Максимов В.А., 2017.

<sup>2</sup> Кандидат технических наук, доцент кафедры информационно-вычислительных систем и сетей Военно-космической академии им. А.Ф. Можайского.

© Молчанов О.Е., 2017.

данные ДЗЗ представлены изображениями подстилающей поверхности, зарегистрированными в оптическом диапазоне спектра электромагнитного излучения [1].

Тенденциями современных космических технологий ДЗЗ являются создание и применение многоспутниковых орбитальных группировок (ОГ) ДЗЗ на базе микроспутников (МС), способных превзойти по эффективности существующие орбитальные группировки из традиционных КА при сравнимой стоимости. Перспективным направлением развития МС помимо дальнейшей микроминиатюризации является построение из отдельных КА орбитальных структур (кластеров): «созвездий» и «роёв». Объединение МС в кластеры позволяет существенно расширить

функциональные возможности МС, повысить их производительность и живучесть [2].

В США начиная с 2006 года ведется целый ряд НИОКР, посвященных военному использованию МС, в том числе [3]:

- программа COSMIC (Concepts and Operations for Space and Missile Defense Integration Capabilities) создания космических систем военного назначения на базе МС (реализуется с 2006 года);

- концепция оперативного доступа в космос ORS (Operationally Responsive Space) с целью непосредственной поддержки боевых действий, (реализуется с 2007 года);

- программа непосредственного использования космических систем при ведении боевых действий SeeMe (Space Enabled Effects for Military Engagements) (начата в конце 2012 года);

- государственная программа Tacsat (Tactical Satellite) США по созданию быстроразворачиваемых группировок малых КА (начата в конце 2004 года);

- программа System F6 (Future, Fast, Flexible, Fractionated, Free-Flying Spacecraft united by Information eXchange) для проектирования, создания и демонстрации возможностей фрагментированного КА [4];

- программа ALASA (Airborne Launch Assist Space Access), в рамках которой создается платформа воздушного базирования для быстрого и экономичного выведения на орбиту космических аппаратов массой до 50–100 кг (начата в конце 2012 года).

Под кластером следует понимать совокупность однотипных МС или МС различного целевого назначения, совместно решающих общую задачу и воспринимаемых потребителем как единое целое («виртуальный» космический аппарат). Кластеры МС имеют ряд отличий от других орбитальных группировок. Во-первых, в существующих ОГ полноразмерных КА и МС отсутствует непосредственная связь между КА, или же эта связь опосредована наземным комплексом управления или специальным комплексом. Во-вторых, баллистическое построение кластера предполагает нахождение всех его МС в относительной близости друг от друга, допускающей информационный обмен между ними.

Для эффективного применения кластеров МС целесообразна реализация функционально распределенной обработки информации [5], предполагающей разделение задач съема информации микроспутниками и передачи ее на некоторый выделенный КА для обработки и последующей выдачи на наземные пункты управления.

При решении подобных задач распределенной обработки информации необходимо также решать задачи надежного хранения результатов наблюдения в пределах кластера КА МС до момента их обработки и выдачи на наземный пункт управления.

Проблема надежного хранения результатов наблюдения связана с особенностями конструкции МС и организации работы кластера:

- в силу малых размеров МС невозможно организовать систему хранения данных больших объемов для хранения результатов наблюдения;

- в силу ограничения на энергетические характеристики МС не может находиться в постоянно активном режиме (обмен данными с другими МС) и, как следствие, может оказаться недоступным в момент запроса данных от КА-сервера;

- в силу ограничения на массогабаритные характеристики на МС невозможна установка мощных передатчиков для организации высокоскоростного обмена данными как внутри кластера, так и с НАКУ.

В качестве одного из решений такой задачи предлагается распределенное хранение результатов наблюдения в пределах кластера МС ДЗЗ, аналогичное распределенному хранению данных в сигнальных сетях (WSN), основанное на кодах Рида – Соломона.

#### **Постановка задачи**

Для успешного решения задач дистанционного зондирования Земли каждый МС после проведения сеанса наблюдения должен передать полученные данные на КА-сервер для последующей обработки, либо КА-ретранслятор для передачи на Землю. Однако немедленная передача данных не всегда возможна ввиду различных факторов: помехи, не позволяющие маломощным передатчикам МС осуществить передачу данных; возможные особенности орбитального построения кластера МС; особенности программы управления работой кластера. В таких случаях необходима организация надежного хранения данных на МС до момента востребования их со стороны КА-сервера либо КА-ретранслятора. В случае если каждый МС будет хранить только результаты собственных наблюдений, существует вероятность, что в момент затребования данных он окажется неактивным (например, из-за разряженных батарей либо возникшего сбоя или отказа).

Для увеличения вероятности получить результаты наблюдений необходимо ввести информационную избыточность в хранимые данные и организовать ее распределение по системе. Для

решения данной задачи предлагается следующая схема применения кодов Рида – Соломона.

1. Данные наблюдения отдельного МС делятся на  $l = n - m$  блоков.

2.  $m$  избыточных блоков генерируются с помощью кодов Рида – Соломона на основе  $l$  блоков исходных данных,  $n$  – полное число блоков данных.

3.  $n$  блоков распределяется по всем  $n$  МС, включая МС-источник данных.

Очевидно, что степень избыточности данных  $m \in [0, n - 1]$ . В случае  $m = 0$  каждый МС хранит только собственные результаты наблюдения. В случае  $m = n - 1$  каждый МС хранит полную копию результатов наблюдения всех МС.

Процесс перераспределения внутри кластера закодированных данных занимает некоторое время, которое зависит от скорости обмена данными между МС и их состояния (активен или нет). Будем предполагать, что МС остается активным, пока не передаст все  $(n - 1)$  блоков данных соседям, после чего он удаляет результаты наблюдения из памяти, оставляя там только свою  $1/n$  часть данных. В случае если в момент перераспределения данных они запрашиваются сервером, то МС отправляет все данные на КА-сервер, прекращая перераспределение по кластеру.

Когда КА-сервер запрашивает данные от МС, происходит следующая последовательность действий.

1. КА-сервер передает широкоэвещательный сигнал на сбор данных о наблюдении.

2. Все МС, находящиеся в активном состоянии, начинают передачу данных.

3. В случае отсутствия ответа от МС КА-сервер помечает его как неактивный и инициирует процедуру восстановления данных с данного МС.

4. Когда данные собраны, КА-сервер прекращает сессию сбора данных и начинает их обработку.

5. Активные МС приступают к очередному сеансу наблюдения и перераспределению данных.

6. Когда пассивные МС становятся активными, они получают сигнал от соседей о начале очередного сеанса работы. Происходит удаление старых данных и инициируется очередной сеанс наблюдения.

Задача работы – найти зависимость вероятности успешного восстановления результатов наблюдения от вероятности нахождения МС в активном состоянии и степени избыточности данных.

### Модель распределённого хранения данных

Под надежностью хранения данных в данной работе будем понимать вероятность успешного восстановления КА-сервера результатов наблюдения МС –  $P_{Rc}$ .

Благодаря использованию кода Рида – Соломона данные могут быть восстановлены при потере блоков, числом меньше  $m$ . Тогда вероятность успешного восстановления КА-сервером результатов наблюдения МС будет определяться как вероятность того, что число активных МС больше  $(m - n)$ :

$$P_{nk} = \sum_{k=n-m}^n \binom{n}{k} p_a^k (1 - p_a)^{n-k} . \quad (1)$$

Выражение (1) может быть использовано для определения нижней границы вероятности нахождения МС в активном состоянии  $p_a$ , исходя из требований к вероятности успешного восстановления  $P_{nk}$ .

Однако  $P_{nk}$  представляет собой условную вероятность восстановления, при условии, что КА-сервер начнет сбор данных после окончания распределения блоков данных по кластеру. Безусловная вероятность будет несколько выше ввиду того, что КА-сервер может начать сбор данных раньше, и ему не понадобится восстанавливать данные (он получит их от МС напрямую). Таким образом, вероятность успешного восстановления КА-сервером результатов наблюдения МС:

$$P_{Rc} = p_d P_{nk} + (1 - p_d) , \quad (2)$$

где  $p_d$  – вероятность того, что распределение завершено до того, как КА-сервер начал сбор данных.

Предположим, что КА-сервер начинает сбор данных в  $N$ -й промежуток времени (между промежутком  $(N - 1)$  и  $N$ ). Процедура распределения данных будет считаться завершённой до момента времени  $N$ , если каждый из МС был активен хотя бы один промежуток времени до момента  $N$ . Тогда вероятность завершения распределения вычисляется следующим образом:

$$p_d(N) = 1 - P_R \text{ (один или несколько МС не были ни разу на промежутке } [0, N]) = 1 - \sum_{i=1}^{n-1} \binom{n-1}{i} (p_p)^i (1 - p_p)^{n-i-1} , \quad (3)$$

где  $p_p = 1 - p_a$ .

Вероятность того, что КА-сервер начнет сбор данных во время  $N$ -го промежутка времени:

$$p_s = e^{-\lambda N} (e^{\lambda N} - 1) , \quad (4)$$

где  $\lambda$  – параметр экспоненциального распределения.

Так как  $p_d$  зависит от  $N$ , то необходимо взять среднее по всему диапазону  $N$ .

Таким образом, итоговая формула для вероятности восстановления данных будет выглядеть следующим образом:

$$P_R = \sum_{N=1}^{\infty} [p_d(N) \cdot p_{nk} + (1 - p_d(N))] \cdot p_s \cdot (5)$$

### Заключение

Приведенная вероятностная модель распределенного хранения данных в кластере МС ДЗЗ описывает основные зависимости параметров хранения данных в распределенной системе, использующей коды Рида – Соломона, а также позволяет рассчитывать вероятность успешного восстановления данных наблюдения в зависимости от параметров системы. Она может применяться для обоснования состава и определения основных параметров хранения данных в группировках МС.

Применение предложенной схемы хранения данных позволяет организовать обработку информации в кластере МС, при этом позволяя экономить энергоресурс и ресурс надежности МС, что приводит к увеличению сроков активного существования кластера МС в целом.

### Литература

1. Савиных В.П. Оптико-электронные системы дистанционного зондирования / В.П. Савиных, В.А. Соломатин. – М. : Недра, 1996. – 315 с.
2. Басыров А.Г., Швецов А.С., Шушаков А.О., Ширококов В.В. Модель распределенной обработки информации в условиях воздействия дестабилизирующих факторов на информационно-телекоммуникационную систему // Современные проблемы науки и образования. – 2015. – № 2. – URL: <http://www.science-education.ru/122-20968> (дата обращения: 30.07.2015).
3. Modular Space: DARPA's F6 Program – URL: [http://www.defenseindustrydaily.com/](http://www.defenseindustrydaily.com/Mod-)

ular-Space-DARPA-Awards-Phase-2-Systems-F6-Contract-06044. (дата обращения: 22.10.2014).

4. Ширококов В.В., Шинкаренко А.Ф. Подход к организации межспутникового взаимодействия в распределенной вычислительной структуре орбитальной группировки микроспутников // Труды Военно-космической академии имени А.Ф. Можайского. – СПб. : ВКА имени А.Ф. Можайского, 2015. – Вып. 646. – С. 77–82.

5. Захаров И.В., Кремез Г.В., Максимов В.А. Построение распределенных запоминающих устройств бортовых вычислительных систем космических аппаратов дистанционного зондирования Земли // Труды Военно-космической академии имени А.Ф. Можайского. – СПб. : ВКА имени А.Ф. Можайского, 2016. – Вып. 652. – С. 160–166.

6. Лоскутов А.И., Дуников А.С., Артюшкин А.Б., Нечай А.А. Математическая модель системы символьной синхронизации наземной приемно-регистрирующей станции телеметрической информации в условиях флуктуаций амплитуды сигнала // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». – 2017. – Выпуск 1. – С. 11–19.

7. Новиков А.Н., Нечай А.А., Малахов А.В. О подходе к обоснованию рациональной номенклатуры эталонной базы измерительных комплексов на основе нечетких моделей // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». – 2017. – Выпуск 1. – С. 72–79.

8. Ширококов В.В., Нечай А.А. Алгоритм планирования энергосберегающей параллельной обработки информации с учетом информационной важности и времени поступления задач // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». – 2017. – Выпуск 1. – С. 88–93.