

# Фрактальные методы в информационных технологиях обработки, анализа и классификации больших потоков астрономических данных

© А.В. Мышев

© А.В. Дунин

Национальный исследовательский ядерный университет (МИФИ) – Обнинский институт атомной энергетики (ИАТЭ),  
Обнинск, Россия

[mishev@iate.obninsk.ru](mailto:mishev@iate.obninsk.ru)

**Аннотация.** Рассмотрена фрактальная парадигма построения моделей и логических схем алгоритмов и процедур информационных технологий обработки, анализа и классификации больших потоков астрономических данных об орбитах и траекториях малых тел. Методология построения таких моделей и схем основана на построении оценок критериев близости и связанности орбит и траекторий в пространстве возможных состояний с использованием соответствующего математического аппарата фрактальных размерностей. Логическая, алгоритмическая и содержательная сущности фрактальной парадигмы заключаются в следующем. Во-первых, обработка и анализ потока данных орбит и траекторий состоят в том, чтобы определить, образует ли он фрактальную структуру. Если да, то определить центры фрактальной связанности потока и получить оценки индекса информационной связанности орбит или траекторий. Во-вторых, нужно выделить монофрактальные структуры в потоке и классифицировать их по признаку принадлежности к классам перколирующего фрактала или фрактального агрегата.

**Ключевые слова:** связанность орбит, фрактальные меры, фрактальная размерность, перколирующий фрактал, фрактальный агрегат.

## Fractal Methods in Information Technologies for Processing, Analyzing and Classifying Large Flows of Astronomical Data

© A.V. Myshev

© A.V. Dynin

National Research Nuclear University MEPHI (IATE),  
Obninsk, Russia

[mishev@iate.obninsk.ru](mailto:mishev@iate.obninsk.ru)

**Abstract.** The fractal paradigm of constructing models and logical schemes of algorithms and procedures for information processing, analysis and classification of large flows of astronomical data on the orbits and trajectories of small bodies is considered. The methodology for constructing such models and schemes is based on the construction of estimates of proximity and connectivity criteria for orbits and trajectories in the space of possible states using the corresponding mathematical apparatus of fractal dimensions. The logical, algorithmic and substantial essence of the fractal paradigm is as follows. Firstly, the processing and analysis of the data flow of orbits and trajectories is to determine whether it forms a fractal structure. If yes, then it determines the centers of fractal connectivity of the flow and obtains estimates of the index of information connectivity of orbits or trajectories. Secondly, it is necessary to isolate the monofractal structures in the flow and classify them according to the attribute of belonging to the classes of a percolating fractal or a fractal aggregate.

**Keywords:** connectedness orbits, fractal measures, fractal dimension, percolating fractal, fractal aggregate.

### 1 Введение

Рассмотрены новые подходы к построению

---

Труды XIX Международной конференции «Аналитика и управление данными в областях с интенсивным использованием данных» (DAMDID/ RCDL'2017), Москва, Россия, 10–13 октября 2017 года

моделей и логических схем алгоритмов и процедур информационных технологий обработки, анализа и классификации больших потоков астрономических данных об орбитах и траекториях малых тел. Методология построения таких моделей и схем основана на построении оценок критериев близости и связанности орбит и траекторий в пространстве возможных состояний с помощью соответствующего математического аппарата фрактальных

размерностей. Логическая, алгоритмическая и содержательная сущности методов и технологий фрактальной парадигмы заключаются, во-первых, в обработке и анализе потока данных орбит и траекторий с тем, чтобы определить, образует ли он фрактальную структуру (если да, то необходимо определить центры фрактальной связанности потока и получить оценки индекса информационной связанности орбит или траекторий), во-вторых, в выделении монофрактальных структур в потоке и классификации их по признаку принадлежности к классам перколирующего фрактала или фрактального агрегата. Фрактальная парадигма в методологии разработки и реализации информационных технологий обработки, анализа и классификации больших потоков данных, в отличие от традиционных методов и способов [1–4], позволяет учитывать как свойства регулярности и нерегулярности структуры пространства состояний информационной шкалы данных потока, так и их динамическую и информационную связанность.

Рассматриваемые логические схемы алгоритмов и процедур технологий обработки, анализа и классификации больших потоков данных построены на основе теории фрактальных размерностей пространственных и временных структур, алгоритмическая и содержательная сущность которых заключается в следующем. Во-первых, обработка потока данных состоит в том, чтобы определить, образует ли он фрактальную структуру. Если да, то необходимо определить центры фрактальной связанности потока данных и получить оценки индекса информационной связанности. Во-вторых, алгоритмы и процедуры технологий анализа и классификации обработанного потока позволяют выделить монофрактальные структуры, если поток образует мультифрактал, и классифицировать их по признаку принадлежности к классам перколирующего фрактала или фрактального агрегата, а также оценить меру расхождения между геометрическими и информационными фрактальными размерностями, как индикатора единства количественных и качественных характеристик потока.

## 2 Астрономические данные (обработка, анализ и классификация)

### 2.1 Постановка задачи

Для обработки потоков астрономических данных об орбитах и траекториях объектов космического пространства получены новые критерии их близости. В качестве критерия близости двух орбит введена количественная оценка фрактальной меры на множестве по-парных расстояний между соответствующими парами точек орбит. Такая оценка фрактальной меры является критерием связанности двух орбит, посредством которого отражается степень их геометрической и информационной близости. Основная посылка и смысл введенной сущности объясняются и поясняются следующей логической схемой и

алгоритмом.

Во-первых, на орбитах определяются реперные (или опорные) точки, в качестве которых могут выступать перигелий или афелий орбиты. Во-вторых, относительно реперных точек выбирается  $N_0$  дискретных точек с шагом дискретизации по истинной аномалии  $\Delta v = 360^\circ / N_0$ , где значение  $N_0$  определяется из следующих условий: 1) статистической значимости и репрезентативности выборки; 2) уровня надежности оценки фрактальной меры. В-третьих, на каждом  $i$ -ом шаге дискретизации по  $\Delta v$  вычисляется расстояние  $r_i$  между соответствующими точками на орбитах. В-четвертых, после завершения предыдущего шага по всем дискретным точкам орбит определяются  $r_{min}$  и  $r_{max}$ , вычисляется  $\Delta = |r_{max} - r_{min}|$ , который разбивается на  $K$  подинтервалов. В-пятых, для заданного уровня геометрической близости двух орбит  $r_{дог}$  (радиус сферы доверительности) вычисляется оценка уровня доверительности  $p_{дог}$ , который отражает долю точек  $r_i \in \Delta$ , для которых справедливо условие  $r_i \leq r_{дог}$ . Оценка  $p_{дог}$  является критерием и количественной мерой фрактальной природы близости двух орбит, т. е. для заданного  $r_{дог}$  и с каким значением  $p_{дог}$  можно считать связанными две орбиты. Для получения количественной оценки фрактальной меры на  $K$  подинтервалах множества точек  $r_i \in \Delta$ , учитывающей одновременно емкостные и информационные размерности фрактала, используется формула для оценки универсальной фрактальной размерности  $d_b$ , которая является синергией и обобщением обозначенных выше фрактальных размерностей. Она была получена одним из авторов, а содержательно-смысловое значение величины  $d_b$  более полно и конструктивно описано в [5]. Оценка значения величины  $d_b$  определяется следующим выражением

$$d_b = \lim_{\varepsilon \rightarrow 0} \frac{B(\varepsilon)}{\log(1/\varepsilon)} = \lim_{\varepsilon \rightarrow 0} \frac{\sum_{i=1}^K p_i \log \sum_{j=1}^K (1 - \rho_{ij}) p_j}{\log \varepsilon}, \quad (1)$$

где  $p_i$  – вероятность попадания значения  $r_i$  в  $i$ -й подинтервал  $\Delta = |r_{min} - r_{max}|$ ;  $\varepsilon$  – длина подинтервала для заданного разбиения интервала  $\Delta$ ;  $\rho_{ij}$  – рандомизированная метрика между центрами  $j$ -го и  $i$ -го подинтервалов;  $B(\varepsilon)$  –  $B$ -энтропия.

Рандомизированная метрика  $\rho_{ij}$  определяется по следующей формуле

$$\rho_{ij} = |r_i - r_j| / |r|, \quad (2)$$

где  $|r_i - r_j|$  – это расстояние (геометрическое или информационное) между  $i$ -м и  $j$ -м подинтервалами;  $|r|$  – длина интервала  $\Delta$ . Для вычисления оценки емкостной фрактальной размерности множества  $\{r_i\}$ , которая отражает свойства фрактальной геометрии «дырявого» множества точек  $r_i \in \Delta$ , использовалась известная формула [6]

$$d_f = - \lim_{\varepsilon \rightarrow 0} \frac{\log N(\varepsilon)}{\log \varepsilon}, \quad (3)$$

где  $N(\varepsilon)$  – число покрытий множества точек  $r_i \in \Delta$ ;  $\varepsilon$  – радиус сферы покрытия. Емкостная фрактальная

размерность  $d_f$  вводится и определяется как оценка меры нерегулярности топологической и геометрической структуры множества точек  $r_i$ .

Формулы (1) и (3) использовались для получения оценок критериев связанности двух орбит, которые описывают и отражают степень пространственной близости двух орбит в определенной пространственной окрестности  $r_{дог}$  с заданным уровнем доверительности  $p_{дог}$ . Логическая схема получения оценок фрактальной меры связанности потока орбит относительно оптимальной опорной орбиты и нахождения такой орбиты в потоке выражается в виде следующего алгоритма.

Первый шаг – произвольно выбирается первая фиксированная орбита потока  $O_j$  ( $j=1 \div L$ , где  $L$  – количество орбит потока) и вычисляются по-парные индексы связанностей  $I_k$  между ней и всеми другими орбитами  $O_k$  ( $k=1 \div L$ ,  $j \neq k$ ,  $k$  – переменный индекс) потока по формулам (1), (2). Результат этого шага – множество индексов связанностей  $K_j = \{I_k\}$  для всех орбит потока относительно выбранной опорной орбиты  $O_j$ . На этом шаге также определяются индексы связанностей  $I_m$ , которые вычисляются по формуле (3) и отражают свойства фрактальной геометрии близости двух орбит, образуя множество  $G_j = \{I_m\}$ .

На втором шаге вычисляется коэффициент связанности потока орбит  $\{O_k\}$  относительно опорной орбиты  $O_j$  по формуле (1), т. е. вычисляется фрактальная размерность  $R_j$  множества  $K_j$ . Процедуры первого и второго шагов выполняются для всего множества орбит  $\{O_j\}$ , где  $j$  – индекс опорной орбиты. На последнем шаге алгоритма получаем множество  $S_{орб} = \{R_j\}$ , элементы которого показывают, насколько та или иная орбита может быть «центром» потока орбит, относительно которого наиболее плотно и компактно группируются орбиты. Условие выбора такого «центра» потока орбит определяется выражением

$$R_j \rightarrow \min_j. \quad (4)$$

Основная посылка и смысл обозначенной логической схемы в информационных технологиях обработки и анализа состоят в том, чтобы дать математическое и логическое описание пространства возможных состояний наблюдаемых малых тел с учетом геометрических, динамических и информационных аспектов их эволюции. Для представления расположения тел в моменты их наблюдений и дальнейшей эволюции разработана математико-логическая схема перехода от орбитального описания положений тел к их геометрическому расположению в реальном пространстве, которая реализована в виде следующей процедуры.

Во-первых, локализуется область пространства (куб или прямоугольный параллелепипед), в котором находятся наблюдаемые объекты-тела. Во-вторых, выделяется элементарный объём разбиения этой области, т. е. область разбиения представляется в виде трёхмерной решётки, узел которой является идентификатором элементарного объёма. В-третьих,

решается задача таксономии и классификации пространственного распределения малых тел на узлах трёхмерной решётки. Задача заключается в следующем: образует ли распределение тел в этом объёме регулярную либо нерегулярную пространственную структуру? Для этого использовался аппарат теории фрактальных размерностей и фрактальной геометрии [5–7]. Решением задачи является выделение объема, определение пространственной геометрии и распределения тел на узлах решётки, т. е. определяется, являются ли соответствующие подмножества узлов решётки фрактальными объектами или регулярными. В-четвертых, решалась задача пространственной кластеризации наблюдаемых объектов на узлах решетки: выделения фрактальных кластеров – перколяционный фрактал или фрактальный агрегат.

## 2.2 Фрагменты результатов обработки и анализа данных наблюдений

Для обработки и анализа данные астрономических наблюдений для малых тел солнечной системы были сгруппированы, исходя из опыта предыдущих исследований [8–11]. Классификация малых тел по группам, полученная, исходя из ограничений на значения элементов их кеплеровских орбит, показана в Таблице 1.

Иллюстрация и отражение результатов обработки и анализа потока астрономических данных на примере объектов-тел группы 7 в виде фрагмента приведены в Таблице 2. В столбцах 2, 3 и 4 приведены фрагменты (подмножества) для множеств  $G_j$  и  $K_j$  ( $j=2$ ) и соответствующие их элементам значения  $B$ -энтропии. Значения элементов этих множеств вычислялись относительно второй орбиты. Как видно из Таблицы 2, значения всех элементов множества  $G_j$  больше единицы (условие геометрической регулярности – это равенство всех элементов единице), тем самым отражены фрактальные свойства пространственной структуры потока орбит группы 7. Логические аналогии для элементов множества  $K_j$  (условие однородности геометрии потока – это равенство всех элементов нулю) указывают на то, что «дырявая» пространственная геометрия потока траекторий обладает неоднородной плотностью. На основе элементов пятого столбца была определена орбита, которая является «центром притяжения» потока, т. е. относительно нее наиболее плотно сгруппированы орбиты тел группы. Элементы этой орбиты на момент наблюдений имеют следующие значения:  $a=2.211545$  а.е.,  $i=7.50626^0$ ,  $e=0.5894668$ ,  $\omega=123.26392^0$ ,  $\Omega=313.24332^0$ . Этот этап обработки и анализа данных в логической цепочке информационных технологий позволяет получить оценки количественной меры фрактальной природы и фрактальной геометрии потока орбит рассматриваемой группы тел, а также определить центр потока. Вторым этапом в логической цепочке информационных технологий обработки и анализа являются пространственное представление и

описание распределения малых тел исследуемой группы в локальной области гелиоцентрической прямоугольной системы координат. Гелиоцентрическая система координат  $HOYZ$  определялась с началом  $O$  в барицентре Солнечной системы. Плоскость  $HOY$  – плоскость эклиптики. Ось  $OX$  направлена в точку  $\Upsilon$  (точка весеннего равноденствия). Ось  $OY$  перпендикулярна ей. Ось  $OZ$  выбрана так, чтобы система векторов  $OX, OY, OZ$  образовала правую тройку.

Проиллюстрируем результаты, полученные на этом этапе, на примере объектов-тел группы 1. Разбиение объема, в котором локализованы тела на момент наблюдения, определяется ограниченной трехмерной решеткой  $Z^3$  следующего масштаба: по оси  $X$  – 15 узлов, по оси  $Y$  – 14 узлов, по оси  $Z$  – 12 узлов. Процедура пространственной кластеризации на решетке  $Z^3$  позволила получить следующие результаты.

Во-первых, был выделен 31 кластер: один кластер типа перколяционного фрактала и тридцать типа фрактального агрегата. Размер перколяционного фрактала составлял 473 узла, а размеры фрактальных агрегатов варьировались от одного до нескольких узлов.

Во-вторых, перколяция обнаружена в плоскости  $HOY$ .

В-третьих, степень заполнения узлов решетки  $Z^3$  составляла в пределах двадцати одного процента: около восемнадцати процентов занимает перколяционный фрактал, а остальное – фрактальные агрегаты.

Аналогичные расчеты для тел группы 7 с теми же размерами решетки  $Z^3$  показали следующие результаты: не было обнаружено ни одного перколяционного фрактала, а только кластеры типа фрактального агрегата в количестве 31 (размеры от 1 до 9 узлов), т. е. объекты этой группы не образуют компактные пространственно-протяженные образования. Процессы перколяции малых тел на узлах решетки  $Z^3$  достаточно полно отражают и описывают пространственную и эволюционную связанность этих объектов на их орбитах. Эти процессы наиболее характерны для объектов группы 7, указывая на то, что пространственно-временная геометрия этих объектов в пространстве возможных состояний имеет фрактальную природу, тем самым определяя тип их фрактальной динамики. Для этого типа динамики характерна наиболее нерегулярная пространственно-временная фрактальная геометрия. Такой геометрией с небольшими пространственно-временными масштабами обладают малые тела типа метеорных тел и ряда других. Фрактальная агрегация в пространстве возможных состояний наиболее ярко проявляется в других группах малых тел. Такие фрактальные структуры отражают другой тип фрактальной динамики, проявляемой в различных пространственно-временных масштабах фрактальной геометрии по-разному: при больших – это более регулярный, а при небольших – менее регулярный тип динамики.

**Таблица 1** Классификация малых тел по элементам кеплеровских орбит

Группа 1	Главный пояс астероидов: $e < 1/3$ ; $i < 20^\circ$ ; $2.1 < a < 3.5$ а. е.
Группа 2	Короткопериодические кометы и метеорные тела (включая астероиды группы Аполлона – Амура): $1/3 < e < 0.95$ ; $i < 30^\circ$ ; $a < 15$ а. е.
Группа 3	Долгопериодические кометы и метеорные тела: $e > 0.95$ ; $i$ – случайное; $a > 15$ а. е.
Группа 4	Троянцы (захваченные Юпитером и колеблющиеся относительно его передней и задней лагранжевых точек либрации): $a \approx 5.2$ а. е.
Группа 5	Астероиды группы Гильды: $e \approx 0.2$ ; $i \approx 10^\circ$ ; $a \approx 3.95$ а. е.
Группа 6	Астероиды группы Венгрии: $e \approx 0.1$ ; $i \approx 25^\circ$ ; $a \approx 1.9$ а. е.
Группа 7	Малые тела, орбиты которых пересекают орбиту Земли в окрестности радиуса сферы ее влияния.

**Таблица 2** Значения для элементов множеств  $G_j$  и  $K_j$  ( $j=2$ ) и соответствующие этим элементам значения  $B$ -энтропии [5]

№ №	$B$ -энтропия	Элементы множества $G_j$	Элементы множества $K_j$	Элементы множества $S_{orb}$
1.	0.4750731	1.2494696	0.20632162	0.026309
2.	0	1	0	0.036631
3.	0.4301281	1.2460074	0.18680227	0.033831
4.	0.4807579	1.2096293	0.20879053	0.025871
5.	0.5052235	1.1657263	0.21941579	0.028186
6.	0.4058040	1.2416164	0.17623847	0.036397
7.	0.4985102	1.1521853	0.21650024	0.026056
8.	0.4528685	1.2466675	0.19667829	0.022999
9.	0.4474228	1.1198819	0.19431325	0.030783
10	0.4823662	1.2204674	0.20948901	0.026293
11	0.4808495	1.2688664	0.20883028	0.022580
12	0.5008954	1.2177231	0.21753612	0.031327

### 3 Выводы и некоторые обобщения

Результаты обработки и анализа данных наблюдений для групп малых тел, обозначенных выше, позволяют сделать ряд выводов и обобщений следующего характера. Во-первых, фрактальные методы в информационных технологиях обработки и анализа больших потоков астрономических данных на основе логических схем когнитивной аналитики раскодирования сокрытой в них информации являются перспективной и уникальной парадигмой в области разработки информационных технологий нового поколения для широкого класса задач не только современной астрономии. Во-вторых, потоки данных астрономических наблюдений можно обрабатывать, используя различные процессы и методы теории фракталов и генетических данных как

для получения совокупностей и популяций выборочных данных, так и для их анализа. Эти методы и процессы отражают и определяют особенности получаемых оценок фрактальных мер и размерностей, а также область применения выводов, которые можно сделать на основе этих данных. В этом случае используются два типа выборочности – генетическая и статистическая. Статистическая выборочность связана с определением пространственных масштабов решетки  $Z^3$ , а генетическая – с распределением информации и объектов на узлах этой решетки.

В широком аспекте фундаментальных астрономических исследований проблемы образования и эволюции планетных систем результаты данной работы впервые позволили показать, как и в чем проявляется синергия геометрии пространственной структуры и динамической эволюции объектов обозначенных систем и как это можно описать и объяснить в рамках фрактальной парадигмы. Можно ли провести такие аналогии в рамках традиционных моделей, алгоритмов, схем и др.? Если да, то необходимо показать результаты обозначенных аналогий и сформулировать тренды их теоретического развития и практического продолжения.

Прикладные аспекты результатов работы тесно связаны с решением задач астероидно-кометно-метеорной безопасности и проблемой космического мусора. С одной стороны, предложены методы фрактальной теории решения сложных нелинейных задач обработки, анализа и интерпретации результатов динамической эволюции объектов космического пространства с нерегулярной пространственно-вре-менной фрактальной геометрией. С другой стороны, разработана и реализована новая IT-технология в тренде DAMDID обработки, анализа и классификации орбитальных данных малых тел Солнечной системы (для программной реализации IT-технологий использованы данные с порталов MPC ([www.cfa.harvard.edu](http://www.cfa.harvard.edu)) и NASA ([www.nasa.gov](http://www.nasa.gov))).

## Литература

- [1] Гусева, И.С., Лих, Ю.С.: Статистический анализ орбит комет. Известия ГАО РАН, 220, сс. 219-224 (2012)
- [2] Кочетова, О.М., Кузнецов, В.Б., Медведев, Ю.Д., Шор, В.А.: Каталог элементов орбит нумерованных астероидов ИПА РАН. Известия ГАО РАН, 220, сс. 255-258 (2012)
- [3] Малкин, З.М.: Некоторые результаты статистического анализа определений галактического расстояния Солнца. Известия ГАО РАН, 220, сс. 401-406 (2012)
- [4] Брюно, А.Д., Варин, В.П.: О распределении астероидов по средним движениям. Астрономический вестник, 45 (1), сс. 334-340 (2011)
- [5] Мышев, А.В.: Метрологическая теория динамики взаимодействующих объектов в информационном поле нейросети и нейрона. Информационные технологии, 4, сс. 52-63 (2012)
- [6] Павлов, А.Н., Онищенко, В.С.: Мультифрактальный анализ сложных сигналов. УФН, 7 (8), сс. 859-876 (2007)
- [7] Федер, Е.: Фракталы. М.: Мир (1991)
- [8] Емельяненко, В.В., Нароенков, С.А., Шустов, Б.М.: Распределение околоземных объектов. Астрономический вестник, 45 (6), сс. 512-517 (2011)
- [9] Гафтонюк, Н.М., Горькавый, Н.Н.: Астероиды со спутниками: анализ наблюдательных данных. Астрономический вестник, 47 (3), сс. 213-220 (2013)
- [10] Нароенков, С.А.: Хранение и обработка астрометрических и фотометрических данных об АЗС: настоящее и будущее в России. Космические исследования, 48 (5), сс. 467-470 (2010)
- [11] Альвен, Х., Аррениус, Г.: Эволюция Солнечной системы. М.: Мир (1979)