

УДК 004.8

DOI: [10.26102/2310-6018/2024.45.2.044](https://doi.org/10.26102/2310-6018/2024.45.2.044)

Многомерный кластерный анализ данных трафика морской акватории для планирования маршрутов судов

В.М. Гриняк^{1✉}, А.В. Артемьев², А.С. Девятисильный³, Д.О. Дудко³,
М.Д. Сазонтова³

¹Владивостокский государственный университет, Владивосток,
Российская Федерация

²Морской государственный университет имени адмирала Г.И. Невельского,
Владивосток, Российская Федерация

³Институт автоматизации и процессов управления Дальневосточного отделения
Российской академии наук, Владивосток, Российская Федерация

Резюме. Работа посвящена проблеме планирования маршрутов судов на акваториях с интенсивным движением. В условиях насыщенного трафика навигационная безопасность может быть обеспечена только при координации движения судов и реализации ими определенной схемы движения. В статье рассматривается задача планирования маршрута таким образом, чтобы он соответствовал практике судоходства, сложившейся в конкретном районе. Предлагаемый в работе метод планирования маршрутов основан на кластеризации данных о движении судов. Выделенные кластеры представляют собой области в трех- или четырехмерном фазовом пространстве с близкими значениями скоростей и курсов судов, на основе которых формируется граф возможных маршрутов. Особенностью подхода к построению графа является уменьшение числа вершин и ребер за счет моделирования выделенных кластеров охватывающими многоугольниками. В работе показано, что во многих случаях могут использоваться не только вогнутые, но и выпуклые многоугольники, что может дополнительно уменьшить мощность графа. В статье дается метрика расстояния между точками в фазовом пространстве, по которой ведется кластеризация данных, обсуждаются проблемы выбора параметров метрики и алгоритма кластеризации. Отмечается перспективность использования плотностного алгоритма пространственной кластеризации (DBSCAN). Работа сопровождается расчетами планируемых маршрутов судов на данных реальной акватории (Сангарский пролив). Приводятся результаты кластеризации данных о движении, выделения местоположения кластеров путем построения охватывающих многоугольников, вычисления маршрута судна. Отмечается, что рассматриваемая задача может быть актуальна в контексте перспективного развития автономного судовождения. В этом случае рассчитанный маршрут судна будет соответствовать движению других судов, находившихся на акватории ранее. Это позволит снизить вероятность возникновения опасных ситуаций при движении автономного судна в общем судопотоке.

Ключевые слова: безопасность судоходства, управление движением судов, система установления путей движения, интенсивное движение, планирование маршрутов, кластеризация, алгоритмы на графах.

Благодарности: Работа выполнена в рамках программы академического стратегического лидерства «Приоритет-2030», проект «Разработка алгоритмов автоматического расхождения судов в соответствии с МППСС-72, оценка их эффективности и безопасности».

Для цитирования: Гриняк В.М., Артемьев А.В., Девятисильный А.С., Дудко Д.О., Сазонтова М.Д. Многомерный кластерный анализ данных трафика морской акватории для планирования маршрутов судов. *Моделирование, оптимизация и информационные технологии*. 2024;12(2). URL: <https://moitvivr.ru/ru/journal/pdf?id=1591> DOI: 10.26102/2310-6018/2024.45.2.044

Multidimensional cluster analysis of vessel traffic data for route planning

V.M. Grinyak¹✉, A.V. Artemiev², A.S. Devyatisilnyi³, D.O. Dudko³, M.D. Sazontova³

¹*Vladivostok State University, Vladivostok, the Russian Federation*

²*Maritime State University named after admiral G.I. Nevelskoy,
Vladivostok, the Russian Federation*

³*Institute of Automation and Control Processes, Far Eastern Branch of the Russian Academy
of Sciences, Vladivostok, the Russian Federation*

Abstract: The work is devoted to the problem of planning ship routes in water areas with heavy traffic. In conditions of heavy traffic, navigational safety can be ensured only if ships adhere to a certain traffic pattern. The paper examines the problem of planning a route in such a way that it corresponds to the shipping practices that have developed in a particular area. The route planning method proposed in this work is based on clustering data on vessel traffic. The selected clusters represent areas in three- or four-dimensional phase space with similar speeds and courses of vessels, on the basis of which a graph of possible routes is formed. A feature of the approach for constructing a graph is the reduction in the number of vertices and edges by identifying the location of the selected clusters by covering polygons. The work shows that in many cases not only concave, but also convex polygons can be used, which can further reduce the power of the graph. The paper provides a metric for the distance between points in phase space, which is used to cluster data, and discusses the problem of choosing metric parameters and the clustering algorithm. The promise of using the density spatial clustering algorithm (DBSCAN) is noted. The work is accompanied by calculations of planned vessel routes based on data from real water areas (Tsugaru Strait). The results of clustering traffic data, identifying the location of clusters by constructing enclosing polygons, and calculating the route of the vessel are presented. It is noted that the problem under consideration may be promising in the context of the future development of autonomous vessels navigation. In this case, the calculated route of the vessel will correspond to the movement of other vessels that were previously in the water area. This will reduce the likelihood of dangerous situations occurring when an autonomous vessel moves in the general traffic flow.

Keywords: navigation safety, vessel traffic control, traffic route establishment system, heavy traffic, route planning, clustering, graph algorithms.

Acknowledgements: This work was supported by the program of academic leadership "Prioritet-2030", project "Development of algorithms for vessels collision avoidance in accordance with COLREGs-72, assessment of their effectiveness and safety".

For citation: Grinyak V.M., Artemiev A.V., Devyatisilnyi A.S., Dudko D.O., Sazontova M.D. Multidimensional cluster analysis of vessel traffic data for route planning. *Modeling, Optimization and Information Technology*. 2024;12(2). URL: <https://moitvvt.ru/ru/journal/pdf?id=1591> DOI: 10.26102/2310-6018/2024.45.2.044 (In Russ).

Введение

В последние два-три десятилетия автоматизация судоходства совершила качественный скачок. Так, в начале нулевых годов на основе накопленного научно-технического задела в повседневную судоводительскую практику была внедрена Автоматическая идентификационная система (АИС). Дальнейшее расширение сервисов АИС вкпе с наступлением цифровой эпохи и ее принципиально новыми средствами накопления, хранения, передачи и обработки данных определило следующий уровень интеграции разнородной судоводительской информации (Е-Навигация) и свежий взгляд на проблему гражданского безэкипажного судоходства [1–4]. В последние годы в отраслевой среде сформировалось направление дальнейшего развития методов

судовождения путем создания технологий, позволяющих исключить человеческий фактор и названное А-Навигацией¹ [5].

Разработка и внедрение безэкипажных транспортных средств стали объектом пристального внимания как со стороны исследователей и инженеров, так и со стороны регулирующих органов и служб. Наряду с понятиями безэкипажного или беспилотного судна было введено более емкое понятие автономного судна и предложена классификация судов по степени их автономности:

- судно с экипажем, на котором имеются средства автоматизации процессов и поддержки принятия решений;
- судно с экипажем на борту, управляемое дистанционно;
- судно без экипажа на борту, с дистанционным контролем оборудования и управляемое дистанционно;
- полностью автономное судно.

Судно может иметь как одну, так несколько степеней автономности, в том числе в течение одного рейса.

В исследовательском контексте внедрение в повседневную судоводительскую практику автономных судов потребует новой интерпретации всех классических задач судовождения. Так, остается открытым вопрос о маршрутах автономных судов: должны ли такие суда двигаться в общих с обычными судами судопотоках или для них требуется выделение отдельных зон для движения (например, отдельных фарватеров). С одной стороны, имеющиеся на конкретных акваториях схемы движения судов создавались исходя из удобства реализации их судоводителями, наличия визуальных ориентиров, неформальных традиций судоходства, то есть факторов, не существенных для автономных судов. С другой стороны, формирование таких схем движения происходило с учетом многолетнего коллективного опыта судоходства в конкретных водах, с учетом погодных и гидрологических условий, которые важны для судов всех типов. Поводимому, эта проблема будет решена в будущем по мере накопления реального опыта автономного судовождения.

При планировании маршрута судна в ограниченных водах и в условиях интенсивного движения основными критериями для выбора траектории являются навигационная безопасность движения и следование принятым в конкретном районе правилам судоходства и ограничениям [6–8]. Эти правила зачастую существуют неформально, формируясь на основе судоводительского опыта [9, 10]. Перспективным путем планирования маршрута в последнем случае является использование «исторических» данных о движении судов в выбранном районе. Источником таких данных могут послужить информационные сервисы береговых служб, центров управления движением судов, а в глобальном масштабе – информационные сервисы, работающие с данными АИС (хорошо известен, например, сервис *marinetraffic.com*). Планирование маршрута осуществляется с использованием методов машинного обучения на основе идеи «будем двигаться так, как до нас двигались другие».

В работах, опубликованных ранее, авторы уже исследовали возможность планирования маршрутов на основе ретроспективных данных о движении судов. Например, в статье [11] разрабатывается задача определения характерных значений курсов и скоростей судов на различных участках акваторий. В основе метода решения задачи лежит ее дискретизация небольшими полигональными «доменами». С помощью субтрактивной кластеризации в каждом из доменов оцениваются характерные значения скоростей и курсов судов. Эта информация позволяет в дальнейшем выявлять аномально

¹ а-Навигация. Автономное судовождение. URL: <https://www.a-nav.org/ru/index.html> (дата обращения: 01.06.2024).

движущиеся суда, которые «нарушают» актуальную схему движения акватории. Такие суда способны создавать опасные навигационные ситуации непосредственно или «в развитии» (например, если примут неправильную сторону движения или остановятся на фарватере).

Задача непосредственного планирования маршрута рассматривается в работах [12–14]. В статье [12] акватория разбивается на небольшие полигональные участки и строится граф возможных маршрутов акватории. Вес ребер этого графа назначается на основе выявленных путем кластеризации характерных значений курсов судов. Таким образом учитывается, насколько тот или иной возможный курс судна «желателен» на выбранном участке акватории: для характерных курсов вес ребер уменьшается, для нехарактерных – увеличивается, ребра, соответствующие запрещенным курсам, исключаются из множества. Такой подход позволяет формировать маршруты судов с учетом практики судоходства выбранной акватории. Вместе с тем, результаты решения задачи сильно зависят от разбиения акватории – точки отсчета начала и дискретности разбиения, что является существенным недостатком способа.

В работах [13, 14] граф возможных маршрутов акватории формируется на основе геометрической интерпретации ретроспективных данных трафика. Все маршруты представляют собой ломаные, где точки – это координаты находившихся ранее на акватории судов, а отрезки характеризуют последовательность перехода от точки к точке. Множество вершин и ребер дополняется с учетом пересечения отрезков ломаных. В работе [13] все возможные маршруты были равнозначны, не учитывалась их изолированность или наличие в окрестности похожих маршрутов, что могло приводить к выбору хоть и подходящей, но не типичной траектории. Этот недостаток авторы пытались преодолеть в работе [14], где каждому ребру графа возможных маршрутов приписывается вес, характеризующий близость к нему других ребер. Мера этой близости дается на основе природной аналогии – закона Ампера, определяющего силу взаимодействия проводников с током. Таким образом, ребра графа, лежащие внутри или рядом с «пучком» других ребер, получают меньший вес. Таким образом, преимущество для выбора получают наиболее «популярные» маршруты. Хотя метод и позволяет планировать наиболее типичный для акватории маршрут, его практическая реализация (имея в виду большие объемы ретроспективных данных, на основе которых строится граф возможных маршрутов) характеризуется высокой вычислительной сложностью.

Настоящая работа посвящена дальнейшему развитию методов планирования маршрутов на основе ретроспективных данных о движении судов. Как и в работе [12], предлагаемый метод основан на кластеризации – выделенные кластеры представляют собой области в трех- или четырехмерном фазовом пространстве с близкими значениями скоростей и курсов судов, при этом акватория не разбивается на участки, что позволяет устранить проблемы, связанные с дискретизацией. На основе полученных областей значений формируется граф возможных маршрутов судна, вершины которого соответствуют границам выделенных кластеров и областям их пересечения, а ребра – выделенным характерным курсам судов. Особенность метода состоит в том, что граф возможных маршрутов будет иметь сравнительно небольшое число вершин и ребер даже при значительных объемах исходных данных трафика.

Материалы и методы

В настоящее время наиболее удобный способ получить исходные данные для решения рассматриваемой задачи – это обратиться к сервисам АИС. Предоставляемые ресурсами-агрегаторами траекторные данные включают в себя (наряду с другими) координаты, курсы и скорости судов. Представим их в виде множества записей вида

$\{LAT_i, LON_i, SPEED_i, COURSE_i\}$, где LAT_i и LON_i – географические широта и долгота, $SPEED_i$ и $COURSE_i$ – скорость и курс. Ставится задача выделить участки выбранной акватории, движение судов на которых характеризуется близкими скоростями и курсами, то есть кластеризации исходных данных о движении.

Введем следующую метрику расстояния между точками в четырехмерном фазовом пространстве географических координат, скоростей и курсов:

$$d_{ij}^2 = a_R R^2(LAT_i, LON_i, LAT_j, LON_j) + a_V (SPEED_i - SPEED_j)^2 + a_K (COURSE_i - COURSE_j)^2,$$

где d_{ij}^2 – квадрат расстояния между фазовыми точками с индексами i и j ; $R^2(LAT_i, LON_i, LAT_j, LON_j)$ – квадрат длины дуги Большого круга, соединяющей точку поверхности Земли с географическими координатами LAT_i, LON_i с точкой LAT_j, LON_j ; a_R, a_V, a_K – весовые коэффициенты. Разность $COURSE_i - COURSE_j$ вычисляется так, чтобы учесть периодичность значений угла.

Сущность задачи кластеризации состоит в нахождении подмножеств объектов, лежащих близко друг к другу. Для ее решения вводится метрика расстояния между объектами. Возможны две интерпретации понятия кластера как группы точек: это кластеры – «сгустки» и кластеры – области связности. В первом случае кластеры характеризуются их центрами, а границы определяемого кластером подмножества «размыты». Во втором случае кластеры – это подмножества точек с четко обозначенными границами и входящими в них точками. Для рассматриваемой задачи будем понимать кластер как область связности в определенном фазовом пространстве. Хорошо зарекомендовавший себя метод кластеризации для такого случая – это плотностный алгоритм пространственной кластеризации (DBSCAN). Алгоритм основан на поиске в исходном массиве подмножеств точек, исходя из следующих двух правил:

- каждая точка, входящая в подмножество, имеет «соседнюю» точку, расстояние до которой не выше некоторого наперед заданного значения ε ;
- число точек, входящих в подмножество, не меньше некоторого наперед заданного значения m . В противном случае подмножество не считается кластером.

Весовые коэффициенты a_R, a_V, a_K задаются экспертным способом, исходя из особенностей конкретной выборки данных. Так, например, величина $1/a_R$ может быть равна квадрату характерного расстояния между соседними точками одной траектории; это расстояние зависит от периода обновления данных. С другой стороны, она может быть равна квадрату характерной ширины судопотока (фарватера), где суда движутся с близкими скоростями и курсами. Величина $1/a_V$ равна квадрату разности скоростей, когда они считаются близкими скоростями, например, 2–3 узла. Аналогично, величина $1/a_K$ равна квадрату разности курсов, когда они считаются близкими курсами, например, 5–10 градусов. Исходя из выбранных значений a_R, a_V, a_K определяются характерные значения расстояния между «близкими» фазовыми точками d_{ij} , что дает представление о возможном значении ε . Минимальное количество точек в кластере m также зависит от особенностей данных, в рассматриваемом случае может составлять от десятков до тысяч.

На основе выделенных по результатам кластеризации подмножеств точек формируются охватывающие их многоугольники, задающие границы областей акватории с близкими значениями скоростей и курсов судов. На Рисунке 1 показан схематический пример. Кружками со стрелками показаны координаты судов, их скорость и курс. Цветом выделены те точки, что были отмечены как принадлежащие к

кластерам. На данном рисунке явно выделяются три кластера: суда, движущиеся на юго-запад (зеленые), движущиеся на северо-восток (красные) и движущиеся на восток (синие). Два черных значения внизу рисунка кластера не образуют. В общем случае многоугольники, охватывающие каждый кластер, могут пересекаться. Чаще всего такие многоугольники близки к выпуклым, это обстоятельство позволяет использовать для их формирования более простой алгоритм.

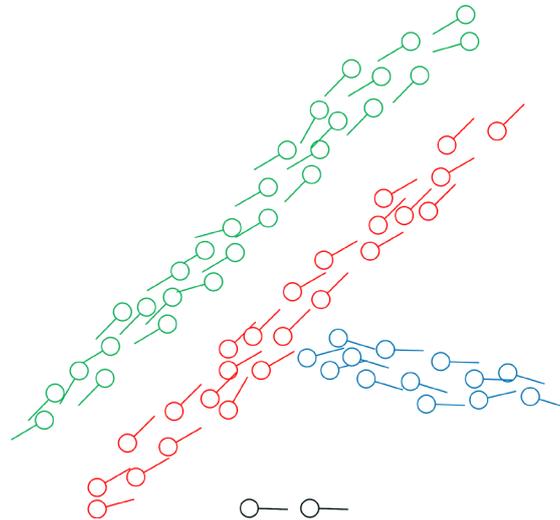


Рисунок 1 – Схематический пример выделения кластеров в данных о движении судов
 Figure 1 – Schematic example of identifying clusters in vessel traffic data

Рассмотрим метод построения графа возможных маршрутов судна на основе данных об охватывающих кластеры многоугольниках. Шаги алгоритма формирования множества вершин графа будут следующими.

1. Включим в множество вершин графа точки вершин охватывающих многоугольников.

2. Если отрезки их сторон слишком велики, включим в множество вершин точки, принадлежащие отрезкам сторон, так, чтобы расстояние между точками не превышало значения ϵ .

3. При наличии пересекающихся охватывающих многоугольников добавим в граф несколько точек, лежащих внутри пересечений, так чтобы можно было хотя бы раз изменить курс при движении от границы к границе.

4. При очень протяженных охватывающих многоугольниках включим в множество вершин несколько точек, лежащих внутри многоугольников, так, чтобы можно было изменить курс при движении внутри многоугольника.

5. Наконец, добавим к множеству вершин точки начала и конца маршрута.

Шаги алгоритма формирования множества ребер графа будут следующими:

1. Если точка начала маршрута не принадлежит ни одному из охватывающих многоугольников – добавляем в множество ребер направленное ребро до ближайшей точки, лежащей на границе многоугольника. Аналогично поступаем с точкой конца маршрута.

2. Считаем, что для каждого охватывающего многоугольника определен диапазон возможных значений курса. Если, выйдя из точки, принадлежащей многоугольнику, и двигаясь возможным курсом можно достичь других точек этого же многоугольника, добавим в множество соответствующие направленные ребра. Прделаем это для всех точек многоугольника – и граничных и внутренних. В некоторых

случаях (при очень протяженных охватывающих многоугольниках) можно ограничить длину ребра.

3. Если имеются непересекающиеся многоугольники, то соединим ребрами их граничные точки по принципу «каждая с каждой», исключив ребра, пересекающие границы многоугольников.

4. Если отрезок, соединяющий точки начала и конца маршрута, не пересекает ни одного из охватывающих многоугольников – добавляем в множество ребер соответствующее направленное ребро.

На Рисунке 2 показан схематический пример, иллюстрирующий построение графа возможных маршрутов судна. Имеется два пересекающихся охватывающих многоугольника, в одном из них суда движутся на восток, в другом – на северо-восток. Красным показана точка начала маршрута, зеленым – конечная точка. Синими точками показаны вершины графа, добавленные в зону пересечения охватывающих многоугольников. Пунктирными линиями со стрелками показаны ребра графа (для наглядности рисунка изображены не все возможные ребра). Из рисунка видна идея планирования маршрута от красной до зеленой точки: сначала достигаем ближайший многоугольник, движемся по одному из ребер графа до вершины, лежащей в зоне пересечения охватывающих многоугольников (на восток), затем меняем курс и движемся по одному из ребер графа до вершины – конечной точки маршрута (на северо-восток).

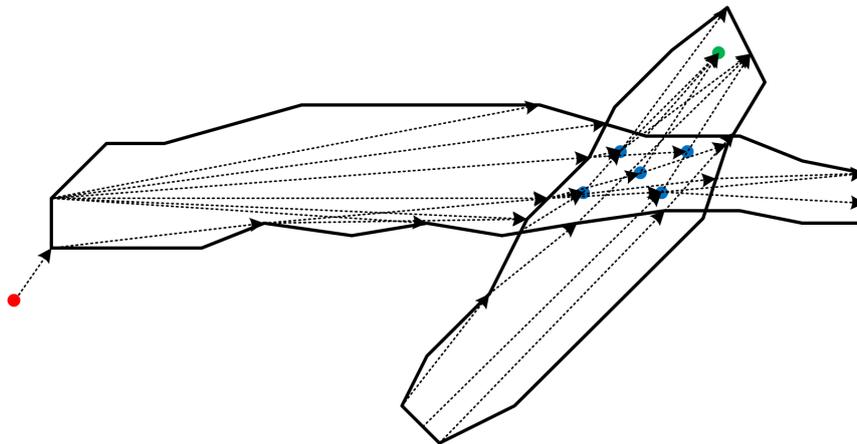


Рисунок 2 – Схематический пример фрагмента графа возможных маршрутов судна
Figure 2 – Schematic example of a graph fragment of possible vessel routes

В самом простом случае вес ребер графа возможных маршрутов принимается равным длине дуги Большого круга, соединяющей инцидентные ребру вершины. Если требуется, чтобы суда стремились двигаться ближе к «середине» судопотока, можно скорректировать вес ребра с учетом наличия в его окрестности других ребер (например, как предложено в работе [14]). Вес ребер также может корректироваться с учетом каких-либо дополнительных условий и эвристик, принятых для движения на конкретной акватории.

Поиск маршрута на сформированном таким образом графе может осуществляться популярными алгоритмами поиска кратчайшего пути на взвешенном графе, например, Дейкстры или Беллмана-Форда.

Таким образом, алгоритм планирования маршрутов состоит в следующем:

1. Формирование выборки ретроспективных данных о движении судов на конкретной акватории, подбор параметров метрики расстояния a_R , a_V и a_K .

2. Подбор параметров ε и m алгоритма кластеризации DBSCAN. Выделение подмножеств точек по результатам кластеризации.

3. Формирование для выделенных кластеров охватывающих многоугольников в пространстве параметров LAT_i и LON_i (географические широта и долгота). Эти многоугольники задают границы областей акватории с близкими значениями скоростей и курсов судов.

5. Построение взвешенного направленного графа возможных маршрутов судна.

6. Поиск кратчайшего пути на графе возможных маршрутов известными алгоритмами.

Результаты

Для проведения вычислительного эксперимента по планированию маршрутов из открытых источников (marinetraffic.com) с помощью специально созданной программной системы [15, 16] были собраны наборы данных о реальном движении судов на различных акваториях. Ниже показаны результаты расчетов маршрутов судов в районе Сангарского пролива. На этой акватории хорошо выделяется характерное движение в направлениях «север-юг» через пролив и «запад-восток» между Японским морем и Тихим океаном.

На Рисунке 3 показан исходный набор данных о движении судов. Данные были собраны в течение одной недели в летний период и содержали около 100 тыс. записей.



Рисунок 3 – Исходные данные о движении судов (Сангарский пролив)
Figure 3 – Initial data on vessel traffic (Tsugaru Strait)

На Рисунках 4–7 показаны результаты кластеризации траекторных данных. Кластеризация выполнялась описанным выше способом при следующих параметрах алгоритма: весовые коэффициенты $a_R = 10^{-4}$ (что соответствует характерному расстоянию между точками траекторий 100 м), $a_V = 0,04$ (соответствует характерной разнице скоростей 5 м/с), $a_K = 0,025$ (соответствует характерной разнице курсов 20°),

расстояние между соседними точками ϵ равно 0,3, минимальное число точек в кластере m равно 50.

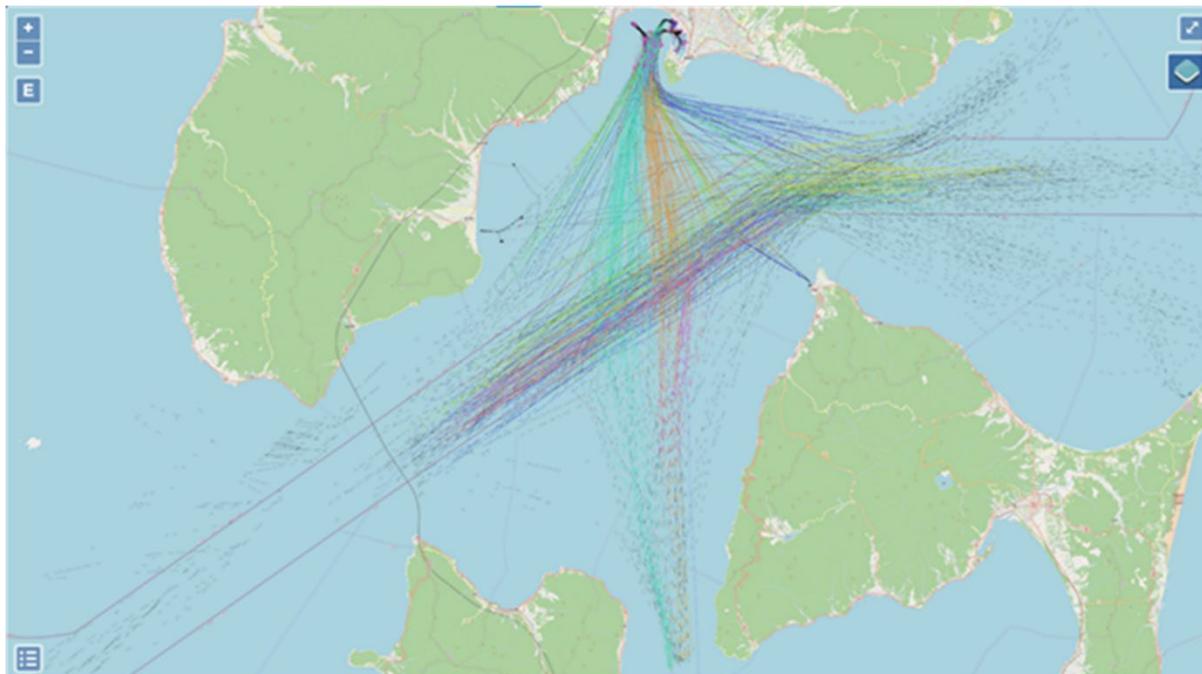


Рисунок 4 – Результаты кластеризации данных о движении
Figure 4 – Traffic data clustering results



Рисунок 5 – Результаты кластеризации данных о движении
Figure 5 – Traffic data clustering results

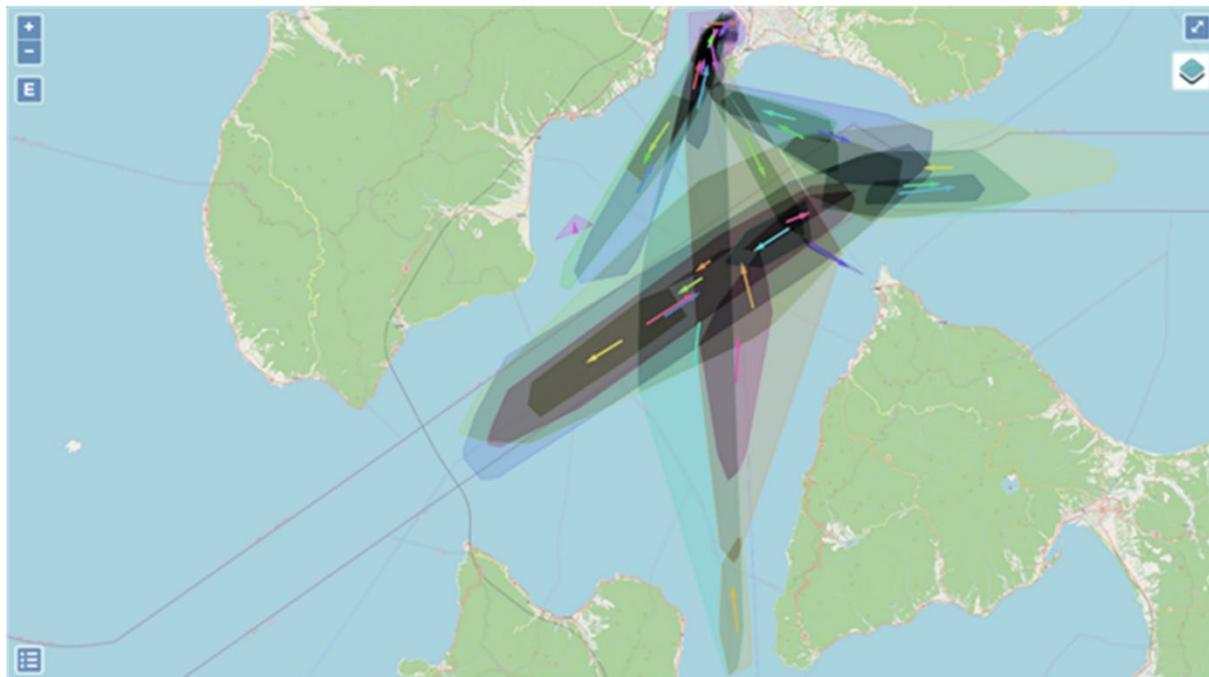


Рисунок 6 – Результаты выделения кластеров как охватывающих многоугольников: выпуклые многоугольники

Figure 6 – Results of identifying clusters as enclosing polygons: convex polygons

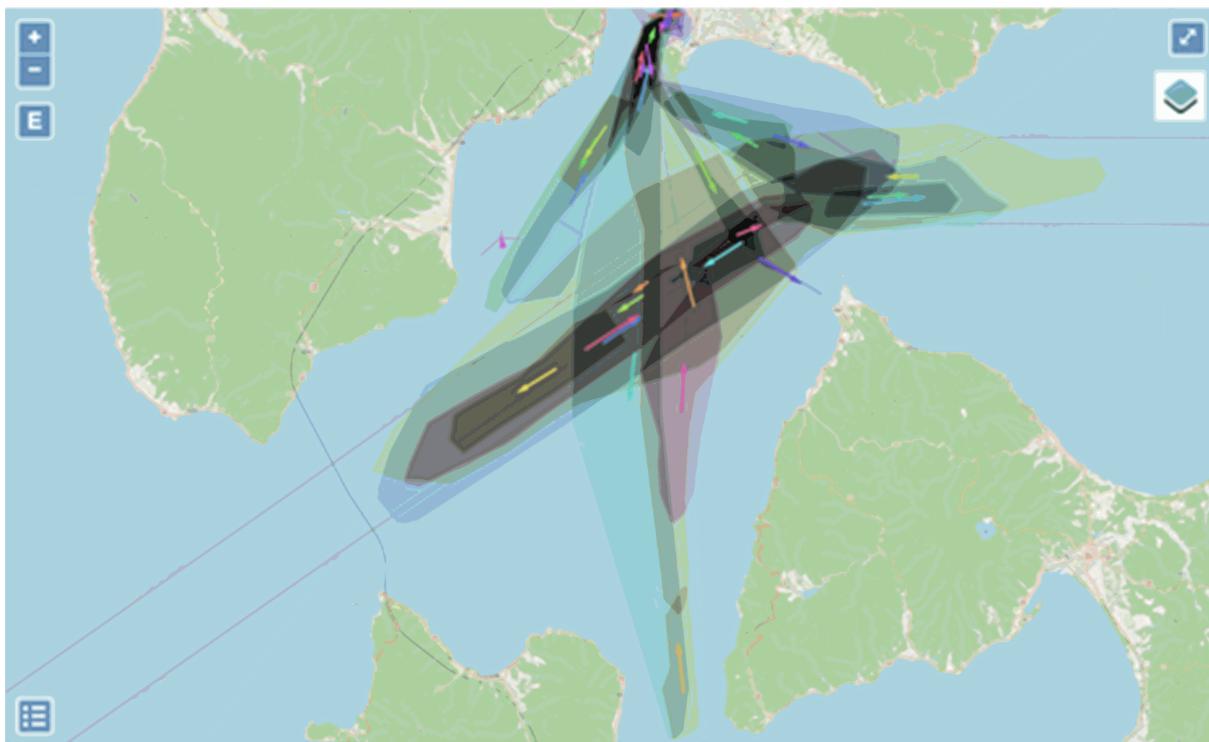


Рисунок 7 – Результаты выделения кластеров как охватывающих многоугольников: вогнутые многоугольники

Figure 7 – Results of identifying clusters as enclosing polygons: concave polygons

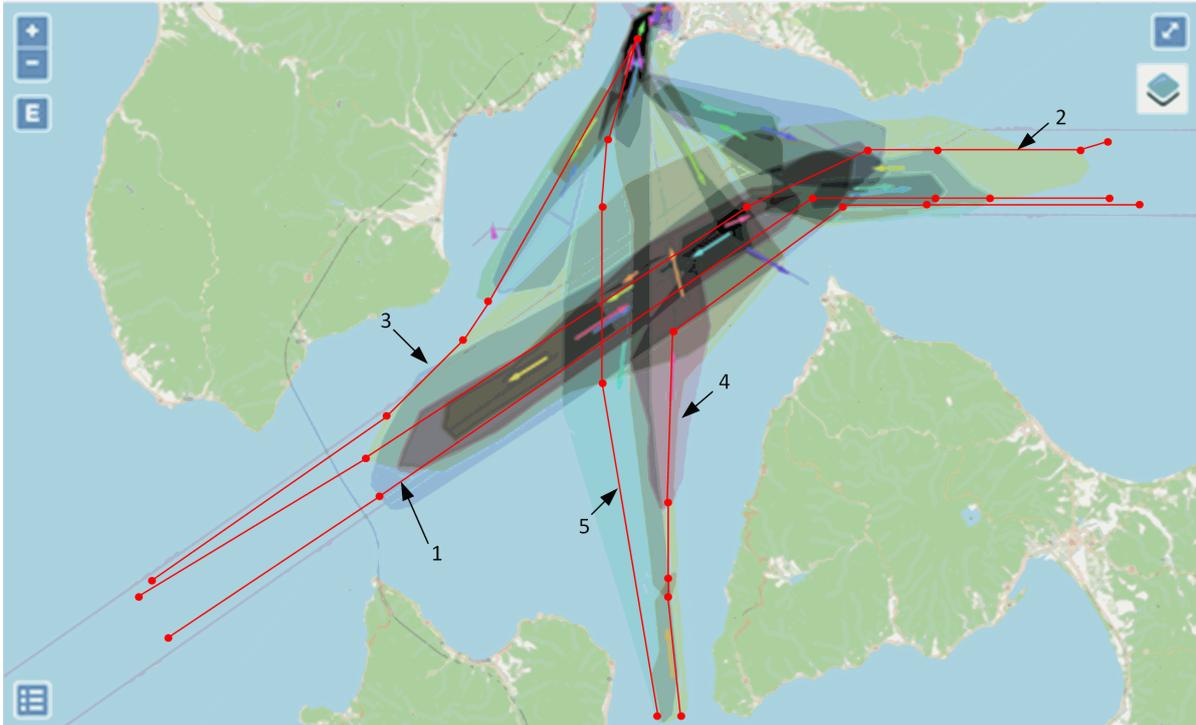


Рисунок 8 – Результаты планирования маршрутов судов
Figure 8 – Vessel route planning results

Было выделено 32 кластера. На Рисунках 4 и 5 показаны результаты разделения точек на кластеры: разные кластеры выделены разными цветами, стрелками показаны значения векторов скорости обрабатываемых записей. На Рисунке 4 показаны данные по всей акватории, на Рисунке 5 – данные по ее южной части (в увеличенном масштабе). Хорошо видны выделенные подмножества исходных данных, соответствующие различным характерным траекторным параметрам движения.

На Рисунках 6 и 7 показаны соответствующие кластерам охватывающие многоугольники, для случая выпуклых (Рисунок 6) и вогнутых (Рисунок 7) многоугольников, стрелками показаны направления и значения средних по кластеру векторов скорости. Видно, что выпуклая и вогнутая оболочки кластеров часто близки друг другу, а сами подмножества параметров движения устойчиво находятся в характерных местах, образуя систему пересекающихся фигур.

Наконец, на Рисунке 8 представлен результат планирования маршрутов на основе вогнутых оболочек кластеров. Показаны следующие маршруты:

1. Японское море – Тихий океан.
2. Тихий океан – Японское море.
3. Хакодате – Японское море.
4. Залив Муцу – Тихий океан.
5. Хакодате – залив Муцу.

Красными точками показаны начальная и конечная точки маршрута, а также вершины графа, через которые проходит маршрут (как правило находятся на границах охватывающих кластеры многоугольников). Красные линии – ребра графа. Видно, что найденные маршруты вполне соответствуют характерным траекториям акватории: маршрут пролегает по зонам интенсивного движения, суда принимают нужную сторону движения (в судовождении принято правостороннее движение). Также стоит отметить сравнительно небольшое число ребер, из которых состоят найденные маршруты, что

упрощает их реализацию на практике, позволяет минимизировать число изменений курса судна.

Обсуждение

Предлагаемый способ построения графа возможных маршрутов на основе многомерной кластеризации траекторных данных показал свою эффективность и конструктивную разрешимость в рассматриваемой задаче планирования маршрутов. В отличие от кластеризации данных по курсу и скорости на каждом небольшом участке акватории отдельно (как в работе [12]), рассмотренный в настоящей работе подход предполагает выделение протяженных областей, характеризующихся близкими значениями скоростей и курсов судов. Это позволяет выделять, например, зоны фарватеров или зоны якорных стоянок и планировать маршруты судов с учетом сложившейся практики судоходства конкретной акватории.

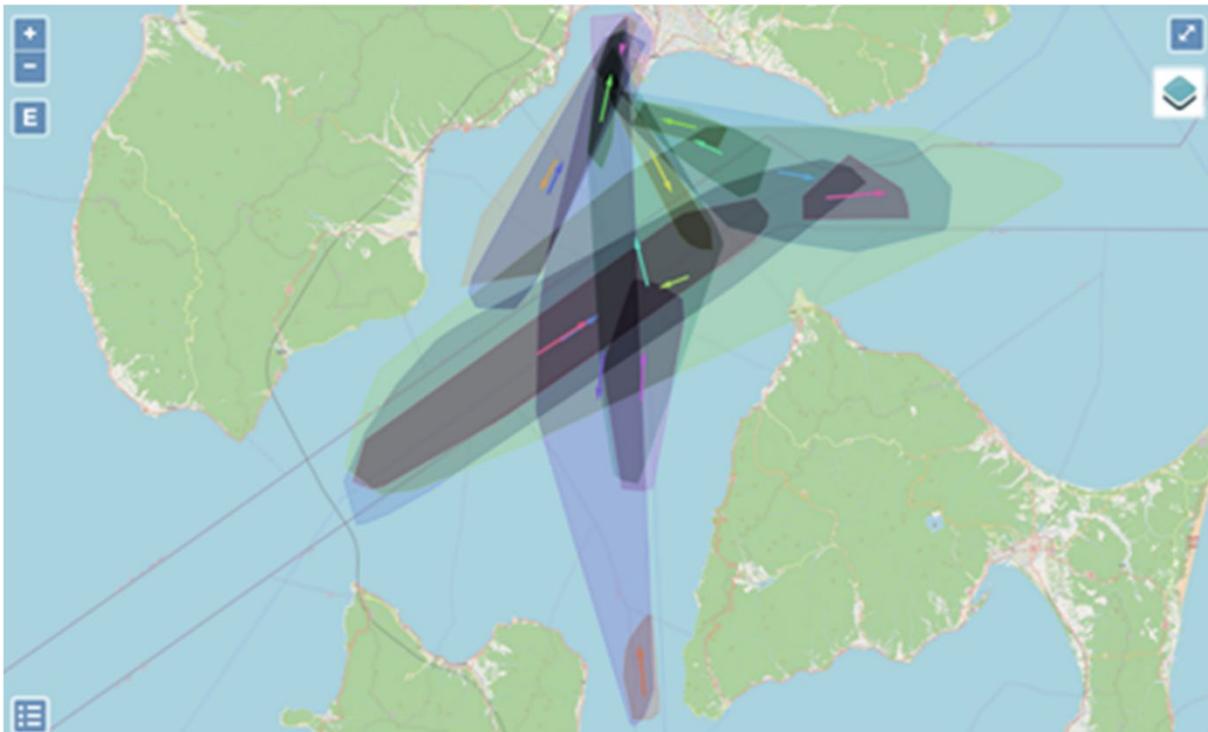


Рисунок 9 – Результаты выделения кластеров как охватывающих многоугольников: выпуклые многоугольники, случай $m=100$

Figure 9 – Results of identifying clusters as enclosing polygons: convex polygons, case $m=100$

Основной проблемой предлагаемого подхода является подбор параметров кластеризации: максимального расстояния между точками кластера ϵ , минимального числа точек в кластере m и коэффициентов метрики расстояния a_R , a_V , a_K . Проведенное исследование показывает, что эти параметры существенно влияют на результат выделения кластеров. Так, на Рисунке 9 показан результат кластеризации при $m=100$ (выпуклые многоугольники, аналог Рисунка 6).

Из Рисунка 9 видно, что при увеличении минимального числа точек в кластере их геометрические размеры растут, а их число уменьшается. При практической реализации рассматриваемой задачи подбор подходящих параметров кластеризации следует осуществлять на основе результатов численного исследования имеющихся данных, в

значительной мере – интуитивно, опираясь на опыт работы с данными такого типа. Вопрос о том, подходят ли параметры кластеризации данных одной акватории для других акваторий остается открытым и требует дополнительного исследования.

Рассматриваемая задача планирования маршрутов представляет собой совокупность трех подзадач: кластеризации данных алгоритмом DBSCAN, выделения охватывающих кластеры многоугольников, построения графа возможных маршрутов, поиска кратчайшего пути на взвешенном графе. Первые три задачи характеризуются высокой вычислительной сложностью, особенно для акваторий с интенсивным разнонаправленным трафиком и сложной географией. Так, например, данные о движении в районах некоторых крупных портов Азии в течение суток содержат несколько миллионов записей. Работа с такими данными требует использования специализированных вычислительных платформ. Вместе с тем, сама по себе высокая вычислительная сложность алгоритмов и большой объем данных не являются препятствием для практического использования рассмотренного подхода к планированию маршрутов, так кластеризация данных и формирование графа возможных маршрутов могут проводиться в течение длительного времени заранее. Оперативного же решения требует только последняя задача – поиска кратчайшего маршрута на взвешенном графе. При формировании графа описанным способом маршруты характеризуются сравнительно небольшим числом ребер, для их поиска можно использовать известные детерминированные алгоритмы.

Развитие средств хранения данных и вычислений в последние два десятилетия обусловило техническую возможность анализа больших выборок ретроспективных данных трафика для решения различных задач судовождения, в том числе – задачи планирования маршрутов. Исторически первым серьезным исследованием на эту тему с примерами расчетов на реальных данных была работа [17], где показаны особенности траекторных данных, предоставляемых сервисами Автоматической идентификационной системы, и обосновывается перспективность их использования в исследованиях и на практике. В работе рассмотрены постановки целого ряда судоводительских задач: определения аномально движущихся судов, планирования маршрутов (как в «большом» варианте, т. е. между портами, так и в «малом», т. е. в ограниченных водах локальных акваторий), оценки конфигурации судопотоков на акватории и паттернов опасных ситуаций и т. д.; приводится множество примеров. Для решения указанных задач используются методы статистики, а также только начавшие в то время оформляться в современном понимании методы анализа больших объемов данных: кластерного анализа и машинного обучения.

В работе [18] рассматривается задача планирования маршрута на локальной акватории на основе ретроспективных данных о движении. Акватория разбивается на небольшие участки, после чего методами математической статистики для каждого участка определяются характерные скорости и курсы судов. Далее по этим данным строится маршрут судна. Как показывают представленные в статье примеры расчетов на реальных данных, получаемый таким образом маршрут может быть достаточно «изломанным» и не пригодным для практики, требует его дополнительной обработки и «сглаживания».

В статьях [8] и [19] для выделения характерных маршрутов судов на акватории используются алгоритмы кластеризации DBSCAN и K-medoids. В результате кластеризации формируются своего рода «пучки ломаных», представляющие собой «паттерны движения» конкретной акватории. Сравнение движения судна с такими характерными маршрутами позволяет выделять аномально движущиеся суда. В статьях

также отмечается возможность планирования маршрутов судов на основе этих данных. Похожая задача решается и в работе [20].

Предложенный в настоящей работе метод планирования маршрутов судов является дальнейшим развитием существующих в настоящее время подходов.

Заключение

В работе рассматривается задача планирования маршрутов судов в конкретной акватории с интенсивным движением на основе ретроспективных данных о ее трафике. Сформулированы общие модельные представления задачи планирования маршрутов. Отмечено, что перспективным способом решения задачи является построение графа возможных маршрутов судна с учетом сложившейся конфигурации судопотоков акватории. Формализация особенностей судопотоков делается путем кластеризации данных в многомерном пространстве «координаты – курс – скорость».

В работе предложена метрика расстояния между точками, на основе которой происходит кластеризация. В качестве метода кластеризации используется алгоритм DBSCAN. Предложен способ формирования множества вершин и ребер графа возможных маршрутов судна на основе информации о выделенных кластерах. Дается оценка сложности алгоритмов кластеризации, формирования графа и поиска на нем маршрута. Отмечено, что при формализации выделенных кластеров охватываемыми многоугольниками число вершин и ребер графа возможных маршрутов будет невелико, что дает возможность реализации предложенного подхода вычислительными системами общего назначения.

Главными результатами работы являются:

- форма представления данных о движении судов, ориентированная на кластеризацию;
- обоснование использования метода кластеризации DBSCAN и рекомендации по выбору его параметров, а также параметров метрики расстояния для кластеризации;
- метод уменьшения размера данных, характеризующих выделенные кластеры путем их представления охватываемыми многоугольниками;
- алгоритм построения графа возможных маршрутов судна.

Результаты моделирования задачи на реальных данных (в настоящей работе представлены результаты моделирования в Сангарском проливе) дают хорошее представление о ее характерных особенностях: числе кластеров, их размерах, значениях коэффициентов метрики расстояния, параметрах алгоритма кластеризации, числе ребер и вершин графа возможных маршрутов, свойствах выделенных маршрутов судов. Приведенный пример планирования маршрутов свидетельствует о конструктивной разрешимости задачи и пригодности использования разработанных алгоритмов в практике судовождения.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Ардельянов Н.П. Промежуточные результаты концепции е-навигации. *Вестник государственного морского университета имени адмирала Ф.Ф. Ушакова*. 2022;(2):8–11.
Ardelyanov N. Intermediate results of the e-navigation concept. *Vestnik gosudarstvennogo morskogo universiteta imeni admirala F.F. Ushakova*. 2022;(2):8–11. (In Russ.).
2. Ривкин Б.С. е-Навигация. Прошло 5 лет. *Гироскопия и навигация*. 2020;28(1):101–120. <https://doi.org/10.17285/0869-7035.0026>

- Rivkin B.S. e-Navigation: Five Years Later. *Gyroscopy and Navigation*. 2020;11(2):176–187. <https://doi.org/10.1134/S2075108720020066>
3. Коренев А.С., Хабаров С.П., Шпекторов А.Г. Формирование траекторий движения безэкипажного судна. *Морские интеллектуальные технологии*. 2021;(4-1):158–165. <https://doi.org/10.37220/MIT.2021.54.4.047>
Korenev A.S., Khabarov S.P., Shpectorov A.G. A route calculation for unmanned vessel. *Morskie intellektual'nye tekhnologii = Marine Intellectual Technologies*. 2021;(4-1):158–165. (In Russ.). <https://doi.org/10.37220/MIT.2021.54.4.047>
 4. Дыда А.А., Пушкарев И.И., Чумакова К.Н. Алгоритм обхода статических препятствий для безэкипажного судна. *Вестник государственного университета морского и речного флота имени адмирала С.О. Макарова*. 2021;13(3):307–315. <https://doi.org/10.21821/2309-5180-2021-13-3-307-315>
Dyda A.A., Pushkarev I.I., Chumakova K.N. Static obstacles avoidance algorithm for unmanned ship. *Vestnik gosudarstvennogo universiteta morskogo i rechnogo flota imeni admirala S.O. Makarova*. 2021;13(3):307–315. (In Russ.). <https://doi.org/10.21821/2309-5180-2021-13-3-307-315>
 5. Пинский А.С. Автономное судовождение. *Морской вестник*. 2021;(2):101–105.
Pinskii A.S. Avtonomnoe sudovozhdenie. *Morskoi vestnik*. 2021;(2):101–105. (In Russ.).
 6. Tsolakakis A., Benders D., De Groot O., Negenborn R.R., Reppa V., Ferranti L. COLREGs-aware Trajectory Optimization for Autonomous Surface Vessels. *IFAC-PapersOnLine*. 2022;55(31):269–274. <https://doi.org/10.1016/j.ifacol.2022.10.441>
 7. Wang H.-B., Li X.G., Li P.F., Veremey E.I., Sotnikova M.V. Application of Real-Coded Genetic Algorithm in Ship Weather Routing. *The Journal of Navigation*. 2018;71(4):989–1010. <https://doi.org/10.1017/S0373463318000048>
 8. Zhao L., Shi G. Maritime Anomaly Detection using Density-based Clustering and Recurrent Neural Network. *The Journal of Navigation*. 2019;72(4):894–916. <https://doi.org/10.1017/S0373463319000031>
 9. Таратынов В.В. Целесообразность разделения морских путей. *Морской флот*. 1969;(9):19–20.
Taratynov V.V. Tselesoobraznost' razdeleniya morskikh putei. *Morskoi flot*. 1969;(9):19–20. (In Russ.).
 10. Zhang B., Hirayama K., Ren H., Wang D., Li H. Ship Anomalous Behavior Detection Using Clustering and Deep Recurrent Neural Network. *Journal of Marine Science and Engineering*. 2023;11(4). <https://doi.org/10.3390/jmse11040763>
 11. Гриняк В.М., Девятисильный А.С., Иваненко Ю.С. Поддержка принятия решений при обеспечении безопасности движения судов на основе кластеризации траекторий. *Вестник государственного университета морского и речного флота имени адмирала С.О. Макарова*. 2020;12(3):436–449. <https://doi.org/10.21821/2309-5180-2020-12-3-436-449>
Grinyak V.M., Devyatisilnyi A.S., Ivanenko Yu.S. Decision support for marine traffic control based on route clustering. *Vestnik gosudarstvennogo universiteta morskogo i rechnogo flota imeni admirala S.O. Makarova*. 2020;12(3):436–449. (In Russ.). <https://doi.org/10.21821/2309-5180-2020-12-3-436-449>
 12. Гриняк В.М., Шуленина А.В. Кластеризация данных траекторий морских судов для планирования маршрутов через акватории с интенсивным движением. *Информационные технологии*. 2021;27(11):607–615. <https://doi.org/10.17587/it.27.607-615>

- Grinyak V.M., Shulenina A.V. Marine Traffic Data Clustering for Ships Route Planning. *Informatsionnye tekhnologii = Information Technologies*. 2021;27(11):607–615. (In Russ.). <https://doi.org/10.17587/it.27.607-615>
13. Гриняк В.М., Девятисильный А.С. Планирование маршрутов судов на основе ретроспективных данных о движении на морской акватории. *Транспорт: наука, техника, управление. Научный информационный сборник*. 2022;(10):34–40. <https://doi.org/10.36535/0236-1914-2022-10-6>
Grinyak V.M., Devyatisilny A.S. Vessel route planning based on historical traffic data of marine area. *Transport: nauka, tekhnika, upravlenie. Nauchnyi informatsionnyi sbornik = Transport: science, equipment, management. Scientific Information Collection*. 2022;(10):34–40. (In Russ.). <https://doi.org/10.36535/0236-1914-2022-10-6>
14. Гриняк В.М., Прудникова Л.И., Артемьев А.В., Левченко Д.М. Планирование маршрутов судов по ретроспективным данным о движении на основе модельных представлений вычислительной геометрии. *Моделирование, оптимизация и информационные технологии*. 2023;11(2). <https://doi.org/10.26102/2310-6018/2023.41.2.015>
Grinyak V.M., Prudnikova L.I., Artemiev A.V., Levchenko D.M. Vessel route planning based on historical traffic data and model representations of computational geometry. *Modelirovanie, optimizatsiya i informatsionnye tekhnologii = Modeling, Optimization and Information Technology*. 2023;11(2). (In Russ.). <https://doi.org/10.26102/2310-6018/2023.41.2.015>
15. Гриняк В.М., Шурыгин А.В.; правообладатель Федеральное государственное бюджетное образовательное учреждение высшего образования «Владивостокский государственный университет экономики и сервиса». *Программа сбора траекторных данных о движении судов из открытых интернет источников: опубл. 19.07.2018*. Свидетельство о государственной регистрации программы для ЭВМ № 2018618729 Российская Федерация. Зарегистрировано в реестре программ для ЭВМ.
Grinyak V.M., Shurygin A.V.; rightholder Federal State Budgetary Educational Institution of Higher Education «Vladivostok State University of Economics and Service». *Programma sbora traektornykh dannykh o dvizhenii sudov iz otkrytykh internet istochnikov: publ. 19.07.2018*. The Certificate on Official Registration of the Computer Program № 2018618729 the Russian Federation. This product is registered in the registry of the computer programs.
16. Головченко Б.С., Гриняк В.М. Информационная система сбора данных о движении судов на морской акватории. *Вестник государственного университета морского и речного флота имени адмирала С.О. Макарова*. 2014;(2):156–162.
Golovchenko B.S., Grinyak V.M. Information system for vessels traffic data capture. *Vestnik gosudarstvennogo universiteta morskogo i rechnogo flota imeni admirala S.O. Makarova*. 2014;(2):156–162. (In Russ.).
17. Pallotta G., Vespe M., Bryan K. Vessel Pattern Knowledge Discovery from AIS Data: A Framework for Anomaly Detection and Route Prediction. *Entropy*. 2013;15(6):2218–2245. <https://doi.org/10.3390/e15062218>
18. Naus K. Drafting Route Plan Templates for Ships on the Basis of AIS Historical Data. *The Journal of Navigation*. 2020;73(3):726–745. <https://doi.org/10.1017/S0373463319000948>
19. Zhen R., Jin Y., Hu Q., Shao Zh., Niktakos N. Maritime Anomaly Detection within Coastal Waters Based on Vessel Trajectory Clustering and Naïve Bayes Classifier.

- The Journal of Navigation*. 2017;70(3):648–670. <https://doi.org/10.1017/S0373463316000850>
20. Tang H., Wei L., Yin Y., Shen H., Qi Y. Detection of Abnormal Vessel Behaviour Based on Probabilistic Directed Graph Model. *The Journal of Navigation*. 2020;73(5):1014–1035. <https://doi.org/10.1017/S0373463320000144>

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Гриняк Виктор Михайлович, доктор технических наук, доцент, кафедра Информационных технологий и систем, Владивостокский государственный университет, Владивосток, Российская Федерация.

e-mail: victor.grinyak@gmail.com

ORCID: [0000-0003-0549-230X](https://orcid.org/0000-0003-0549-230X)

Victor M. Grinyak, Doctor of Technical Science, associate professor, Information Technologies and Systems Department, Vladivostok State University, Vladivostok, the Russian Federation.

Артемьев Андрей Владимирович, кандидат технических наук, доцент, кафедра судовождения, Морской государственный университет имени адмирала Г.И. Невельского, Владивосток, Российская Федерация.

e-mail: artemyev@msun.ru

Andrey V. Artemiev, Candidate of Technical Science, associate professor, Department of Navigation, Maritime State University named after admiral G.I. Nevelskoy, Vladivostok, the Russian Federation.

Девятисильный Александр Сергеевич, доктор технических наук, профессор, сектор Управления и навигации, Институт автоматизации и процессов управления Дальневосточного отделения Российской академии наук, Владивосток, Российская Федерация.

e-mail: devyatis@dvo.ru

Alexander S. Devyatisilnyi, Doctor of Technical Science, Professor, Navigation and Control Department, Institute of Automation and Control Processes, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, the Russian Federation.

Дудко Денис Олегович, аспирант, сектор Управления и навигации, Институт автоматизации и процессов управления Дальневосточного отделения Российской академии наук, Владивосток, Российская Федерация.

e-mail: dudko@dvo.ru

Denis O. Dudko, post graduate student, Navigation and Control Department, Institute of Automation and Control Processes, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, the Russian Federation.

Сазонтова Мария Дмитриевна, аспирант, сектор Управления и навигации, Институт автоматизации и процессов управления Дальневосточного отделения Российской академии наук, Владивосток, Российская Федерация.

e-mail: sazontova@dvo.ru

Maria D. Sazontova, post graduate student, Navigation and Control Department, Institute of Automation and Control Processes, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, the Russian Federation.

Статья поступила в редакцию 30.05.2024; одобрена после рецензирования 14.06.2024; принята к публикации 21.06.2024.

The article was submitted 30.05.2024; approved after reviewing 14.06.2024; accepted for publication 21.06.2024.