

Вероятностное прогнозирование траекторий движения судов на основе кластеризации АИС-данных и регрессионных моделей на основе гауссовских процессов

Ерофеевский Д.В., СПбГУ, Санкт-Петербург

Аннотация

В работе предложен алгоритм, реализующий вероятностное прогнозирование движения судов на основе их исторических АИС-данных. В исходных данных выделяются потоки маршрутов методом кластеризации (DBSCAN и DTW-соединение), что позволяет выделить типичные маршруты и их «средние» траектории. Затем по проекциям векторов скорости и бокового дрейфа вдоль полученных маршрутов строится регрессионная модель на основе гауссовского процесса (GP), дающая не только предсказанную величину движения, но и доверительный интервал прогноза. Такой подход учитывает стохастичность наблюдений и позволяет заранее выявлять зоны повышенного риска. Проведен анализ свойств метода: показано, что GP-модель адекватно аппроксимирует шумные данные и даёт ширину доверительных интервалов, растущую при удалении от обучающих точек. Результаты эксперимента на синтетических данных демонстрируют эффективность предлагаемого подхода. Предложенный метод гибок и может быть применен для различных типов судов и водных путей, обеспечивая точные прогнозы с оценкой неопределенности.

Введение

Прогнозирование траекторий судов является важной задачей для повышения безопасности и эффективности судоходства. Современные исследования показывают, что наибольшую практическую ценность имеют не только модели, выдающие точку будущего положения, но и подходы, оценивающие неопределенность прогноза и учитывающие неоднородность движения судов в разных условиях. Особенно это важно для маршрутов смешанного типа, где морской участок характеризуется высокой вариативностью траекторий, а речной – жесткими геометрическими ограничениями фарватера. В литературе также отмечается, что

кластеризация AIS-данных является базовым этапом построения интеллектуальных систем анализа маршрутов, поскольку позволяет выделять типовые потоки движения и снижать влияние шумовых наблюдений.

Цель работы состоит в разработке единого алгоритмического конвейера, который объединяет кластеризацию АИС-данных, выравнивание траекторий, восстановление центролинии маршрута и вероятностное прогнозирование динамики движения. Предложенная идея заключается в том, чтобы рассматривать прогноз траектории как часть целостной интеллектуальной системы, ориентированной на обработку сложных навигационных процессов.

Методика

Пусть для каждого рейса судна задана дискретная траектория:

$X(i) = \{x1i, x2i, \dots, xTi(i)\}$, где $xt(i) = \{\varphi ti, \lambda ti, vt i, \dots, \psi t(i)\}$ включает широту, долготу, скорость и курс.

Требуется построить отображение:

$$F: X_{1:t}^{(i)} \rightarrow p(X_{t+1:t+H} | X_{1:t}),$$

то есть по наблюдаемой части траектории оценить вероятностное распределение будущего движения.

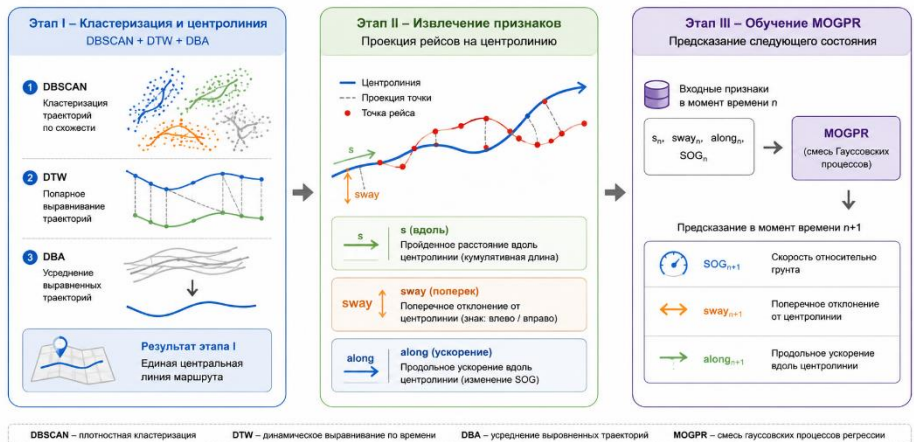


Рисунок 1. Этапы работы алгоритма

На рисунке 1 представлены этапы работы алгоритма. На первом этапе выполняется кластеризация исторических AIS-траекторий. Метод DBSCAN

позволяет выделить плотносно устойчивые группы маршрутов и автоматически отделить выбросы, не задавая число кластеров заранее. Это особенно полезно для судовых данных, где траектории могут иметь произвольную форму и заметно отличаться по длине и времени прохождения. Далее в пределах каждого кластера применяется алгоритм динамической трансформации временной шкалы (DTW), который обеспечивает сопоставление траекторий различной длины и темпа движения. По набору выровненных траекторий строится центролиния маршрута – усреднённый эталонный путь, относительно которого затем анализируются отклонения. Подобная комбинация кластеризации и DTW соответствует современным исследованиям, в которых именно группировка потоков данных АИС рассматривается как основа последующего прогнозирования.

После этого траектория переводится в локальные координаты относительно основной линии движения, то есть относительно типовой траектории маршрута. Такой переход особенно важен для маршрутов «река–море»: в море значимы естественная вариативность и ширина коридора движения, а на реке – боковое отклонение от фарватера. Это делается для того, чтобы анализировать не просто положение на карте, а именно насколько судно продвинулось по маршруту и насколько оно отклонилось от него.

Вектор состояния можно представить, как:

$$z_t = (s_t, d_t, v_t, a_t)^T,$$

Где s_t – продольная координата, d_t – поперечное смещение, v_t – скорость, a_t – ускорение. Это делает задачу прогнозирования не только геометрической, но и навигационно содержательной.

Центральная линия траекторий движения судов может быть задана координатами долготы и широты и используется для описания судовой динамики. На рисунке 2 показаны параметры бокового смещения судна, продольного движения и угла дрейфа вдоль центральной линии траектории судна (см. красную линию на рис. 2, справа). Положительное значение угла дрейфа соответствует смещению судна на левый борт, тогда как отрицательное значение указывает на смещение в противоположную сторону.

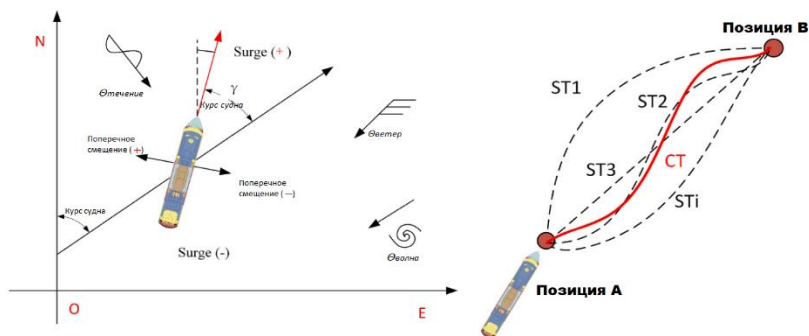


Рисунок 2. Динамика движения судна в реальных эксплуатационных условиях.

Для вероятностного прогноза используется регрессия гауссовских процессов.

Пусть:

а скрытая функция $f(x)$ априорно $\overset{y}{f}$ распределена как:

$$f(x) \sim GP(\mathbf{m}(x), k(x, x'))$$

Тогда апостериорный прогноз в новой точке имеет вид нормального распределения со средним и дисперсией:

$$\bar{f}(x_*) = K_*^T (K + \sigma_n^2 I)^{-1} y, \quad \text{Var}(f(x_*)) = K_{**} - K_*^T (K + \sigma_n^2 I)^{-1} K_*$$

Именно это свойство делает GP особенно удобным для навигационных задач: модель возвращает не только ожидаемое значение, но и количественную оценку неопределенности.

В современных публикациях отмечается, что GP-модели хорошо подходят для AIS-данных, поскольку они устойчивы к шуму и позволяют строить вероятностные траектории. Вместе с тем их вычислительная сложность требует sparse-приближений и аккуратной настройки гиперпараметров.

Результаты

В вычислительных экспериментах на маршруте типа «река–море» получены устойчивые результаты. Для скорости, бокового смещения и продольного ускорения модель показала очень малые ошибки и коэффициенты детерминации, близкие к единице; доверительные

интервалы 95% полностью покрывали тестовые наблюдения. Это означает, что GP-модель не только точно аппроксимирует среднюю динамику движения, но и адекватно оценивает неопределенность. Для морского участка траектории интервалы оказались шире, что соответствует большей вариативности движения, а для речного – более компактными, что согласуется с ограниченностью фарватера. Полученная схема тем самым объединяет кластеризацию, геометрическое выравнивание и вероятностное прогнозирование в единую интеллектуальную систему, пригодную для анализа навигационного риска.

Отдельно следует отметить, что в научной литературе подтверждается высокая эффективность подходов, в которых группировка маршрутов по схожести используется для повышения точности краткосрочного прогнозирования движения судов, а разреженные модели регрессии гауссовских процессов с подбором оптимальных параметров настройки позволяют существенно уменьшить вычислительные затраты без заметного снижения точности и надежности прогноза. Это делает предложенную архитектуру практически применимой для задач мониторинга и поддержки принятия решений в судовождении.

Заключение

Предложен метод вероятностного прогнозирования траекторий движения судов на основе АИС-данных, объединяющий DBSCAN, DTW, восстановление центролинии маршрута и многовыходную GP-регрессию. Показано, что такой подход позволяет учитывать различия между морским и речным режимами движения, формировать прогноз с доверительными интервалами и выявлять участки повышенного риска. По данным эксперимента, модель обеспечивает высокую точность предсказания и хорошую калибровку неопределенности. Научная новизна состоит в интеграции кластеризации и вероятностной регрессии в единый алгоритм, а практическая значимость – в возможности применения метода для навигационного мониторинга и поддержки принятия решений.

Литература

1. Медведев О. Use case: отладка реализации RISC процессора для FPGA / Медведев О. // Материалы 2-й межвузовской научной конференции

п
о

п
р
о
б
л
е
м
а
м