

Алгоритмы иерархического прогнозирования объемов спроса на грузоперевозки

Проект РФФИ 17-20-01212

Московский Центр непрерывного математического образования

Руководитель д.ф.-м.н. В. В. Стрижов¹, докладчик А. Д. Воронов

Участники проекта

- Мотренко Анастасия Петровна — ЗАО “Форексис”
- Воронцов Константин Вячеславович — ФИЦ ИУ РАН
- Ивахненко Андрей Александрович — ЗАО “Анти-плагиат”
- Громов Андрей Николаевич — ФИЦ ИУ РАН
- Кузнецов Михаил Павлович — ФИЦ ИУ РАН
- Борисова Татьяна Игоревна — ФИЦ ИУ РАН
- Замковой Алексей Анатольевич — АО “ИЭРТ”
- Попова Екатерина Витальевна — АО “ИЭРТ”
- Воронов Антон Дмитриевич — ФГБУ ФНКЦ ФХМ ФМБА России

¹strijov@phystech.edu

Цель — создать методологию согласованного прогнозирования иерархических временных рядов.

Решается задача прогнозирования объемов спроса на железнодорожные грузоперевозки с использованием экспертных моделей специалистов РЖД.

- Подготовить базу данных иерархических рядов грузоперевозок.
- Создать методологию построения согласованных прогнозов с учетом собранных специалистами РЖД экспертных оценок.
- Предложить систему критериев оценки качества и достоверности получаемых прогнозов.
- Выполнить вычислительный эксперимент по согласованному прогнозированию иерархических временных рядов.
- Разработать алгоритм согласования для теоретико-игрового подхода ГТОр.
- Подготовить библиотеку экспертных моделей закономерностей и связей между временными рядами.

Цели и задачи, определенные в рамках тем в интересах РЖД

Решается проблема

повышения эффективности транспортировки грузов. Для решения задачи выполняется прогноз потребностей у заказчиков РЖД в узлах погрузки/разгрузки с учетом временных интервалов доставки.

Требуется построить иерархический прогноз объемов спроса

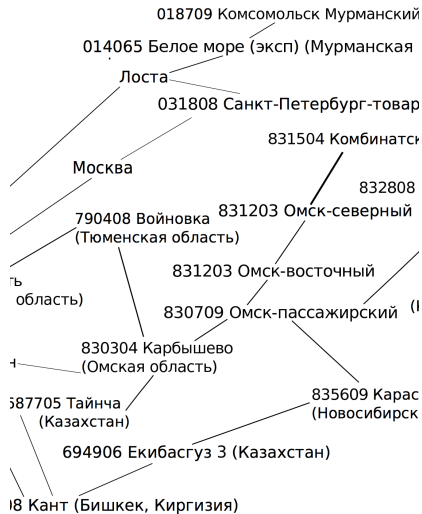
в заданном периоде:

- 1 на месяц посуточно,
- 2 на месяц подекадно,
- 3 на квартал ежемесячно,
- 4 на год ежемесячно,
- 5 на период больше года,

с расположением:

- 1 по группам грузов,
- 2 по парам станций/регионов,
- 3 по комбинированному разложению, учитывающему перечисленные варианты.

Разнородные запросы на прогнозы грузоперевозок от структурных подразделений РЖД



Требуется

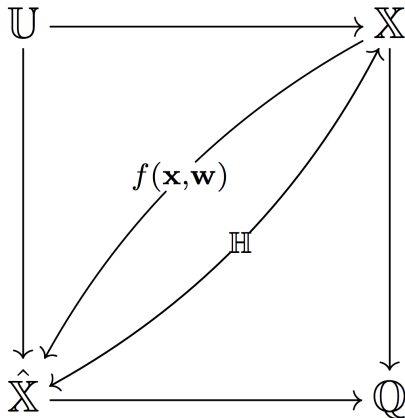
построить прогнозы значений семейства временных рядов, связанных в иерархическую многоуровневую структуру и описывающих объемы погрузки ряда грузов в заданных узлах или на парах узлов с разным уровнем детализации.

Прогнозы должны быть

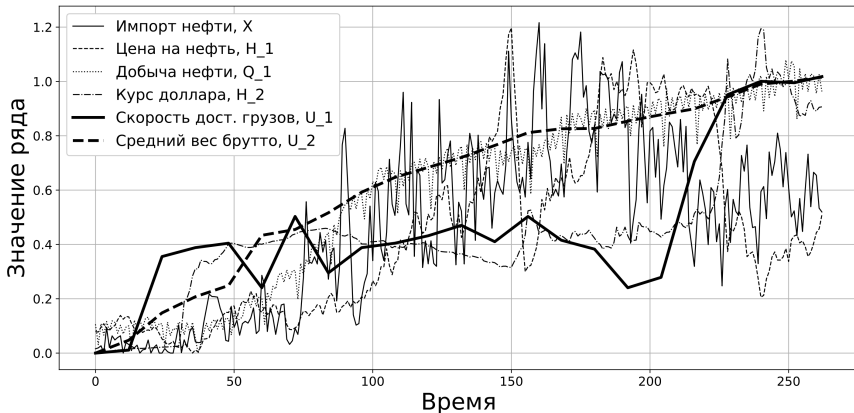
- точны,
- должны удовлетворять структуре иерархии,
- физическим ограничениям,
- должны быть учтены экзогенные факторы.

Учет управляющих воздействий, экзогенных факторов и объемов спроса на грузоперевозки

Наблюдаемое управляющее воздействие U изменяет состояние объекта X . Помимо U на состояние объекта могут влиять экзогенные факторы H . Модель оценки состояния определяет качество текущего состояния объекта Q . На его основе принимается решение о выборе управляющего воздействия. Объектом X являются объемы грузоперевозок. Экспертами РЖД определяются критерии, управляющие воздействия и экзогенные факторы.

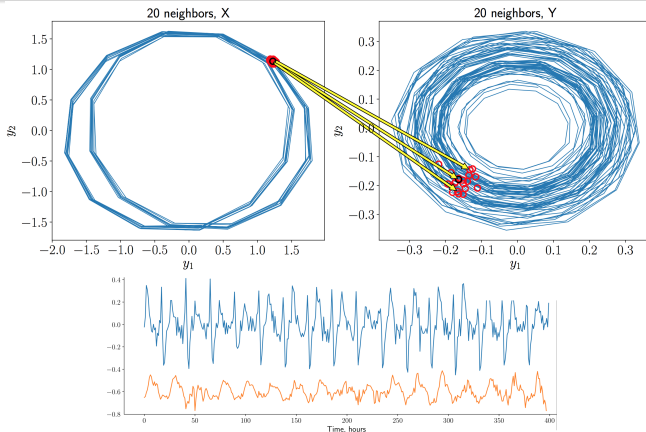


Временные ряды данных о грузоперевозках, собранные специалистами РЖД



Эксперты предоставляют список экзогенных факторов оценки связи между временными рядами.

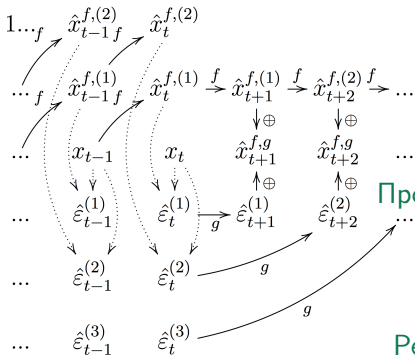
Повышение качества прогноза временного ряда путем учета экзогенных факторов



Считается, что два ряда связаны, если существуют их траекторные подпространства, проекции в которые связаны.

Проекции связаны, если окрестность фазовой траектории одного ряда отображается в окрестность фазовой траектории другого ряда

Выбор оптимальных суперпозиций прогностических моделей

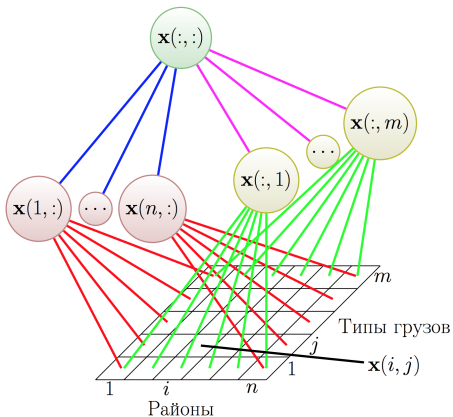


Цель: предложить способ выбора прогностической модели, адекватно описывающей временные ряды с асимметрично распределенным шумом.

Проблема: в зашумленных временных рядах точность прогноза существенно зависит от распределения ошибки.

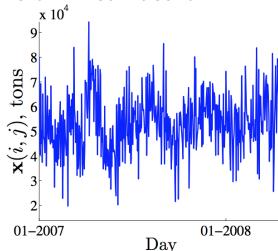
Решение: построить суперпозиции моделей временного ряда и ошибки.

Иерархическое прогнозирование временных рядов



Задача: построить прогнозы семейства временных рядов, связанных в иерархическую многоуровневую структуру и описывающих объемы погрузки ряда грузов в заданных узлах РЖД с разным уровнем детализации.

Проблема согласования прогнозов: прогнозы, полученные для каждого временного ряда независимо, могут не быть согласованными.



Поставленные в проекте задачи выполнены полностью.

- Предложены методы согласования сложноорганизованных иерархических временных рядов, включающих объемы спроса на железнодорожные грузоперевозки, экзогенные макроэкономические показатели, управляющие воздействия, и экспертные оценки.
- Предложен и разработан математический аппарат построения прогностических моделей и их суперпозиций.
- Предложен способ построения иерархических прогнозов. Выполнен ряд вычислительных экспериментов, исследующих предлагаемые методы и модели.
- Разработан макет программного обеспечения для построения прогнозов.

- 1 Воронов А.Д., Громов А.Н., Инякин А.С., Замковой А.А.: Прогнозирование объемов спроса на грузоперевозки для стационарных временных рядов // **Машинное обучение и анализ данных**, Т.4 №1, 2018.
- 2 Воронов А.Д., Громов А.Н., Инякин А. С., Замковой А. А.: Верификация экспертных оценок при выявлении релевантных экзогенных факторов, оказывающих влияние на объемы спроса на грузоперевозки // **Машинное обучение и анализ данных**, Т.4 №1, 2018.
- 3 Уваров Н. Д., Кузнецов М.П., Малькова А.С., Рудаков К.В., Стрижов В.В.: Выбор суперпозиции моделей при прогнозировании грузовых железнодорожных перевозок // **Вестник московского университета, сер.15. Вычислительная математика и кибернетика**, 2018.
- 4 Усманова К.Р., Кудияров С.П., Мартышкин Р.В., Замковой А.А., Стрижов В.В. Анализ зависимостей между показателями при прогнозировании объема грузоперевозок // **Системы и средства информатики**, 28(3), 2018.

В течение второго года планируется выполнить следующие работы по проекту

- Исследовать свойства метода построения суперпозиций прогностических моделей.
- Разработать процедуру попарного сравнения результатов несогласованных прогнозов.
- Провести вычислительный эксперимент по согласованному прогнозированию иерархических временных рядов.
- Протестировать теоретико-игровой алгоритм согласования на модельных данных.
- Для разработанных алгоритмов согласования прогнозов предложить методику поиска взаимосвязи между гетерогенными временными рядами.
- Модифицировать теоретико-игровой алгоритм для повышения качества и надежности прогнозов гетерогенных временных рядов в ходе процедуры их согласования.
- Экспериментально исследовать свойства разработанных алгоритмов, сравнить их с аналогами.

Усманова К.Р., Кудяров С.П., Мартышкин Р.В., Стрижов В.В.
Выявление взаимосвязей между гетерогенными временными рядами с помощью кросс-структурных моделей // Информатика и ее применения, 2019.

При прогнозировании сложноорганизованных временных рядов, зависящих от экзогенных факторов и имеющих множественную периодичность, требуется решить задачу выявления связанных рядов. Предполагается, что добавление этих рядов в модель повышает качество прогноза. В данной работе для обнаружения связей между временными рядами предлагается использовать метод сходящегося перекрестного отображения. При таком подходе два временных ряда считаются зависимыми, если существует отображение окрестности фазовой траектории из одного пространства в другое. Решение этой задачи продемонстрировано на двух парах рядов: “потребление электроэнергии и температура” и “объем железнодорожных перевозок нефти и цена на нефть”.

Stenina, M.M., Popova M., Strijov, V.V. Reconciliation of hierarchical time series // **Data Mining and Knowledge Discovery**, 2019.

Исследуется проблема одновременных предсказаний иерархических временных рядов. Предлагается алгоритм GTOp (теоретико-игровой алгоритм оптимального соответствия) для согласования иерархических временных рядов, точность согласованных оценок которого лучше или равна качеству независимых прогнозов. Также предлагается модификация алгоритма GTOp, которая учитывает надёжность независимых предсказаний, корректируя самые надежные прогнозы во время процедуры приведения в соответствие. Качество предложенного алгоритма оценивалось на данных об объеме грузоперевозок РЖД и данных о потреблении электроэнергии. Полученная оценка качества сравнивалась с качеством уже существующих алгоритмов согласования. Показано, что GTOp дает лучшие согласованные прогнозы.