

ІТМО

**Предсказание горизонта
появления аномалий в
динамических процессах**

Нетроголов Роман,
магистр 1 курс



Anomaly Detection

- Выявление аномалии post factum
- В основном prediction-based
- Много различных детекторов с разными ядрами (NAB detector, OpenSearch)
- Benchmarks

Early Anomaly Detection

- Сокращение количества точек для детекции аномалии
- В основном prediction-based
- Benchmarks

Anomaly Prediction

- Детекция начала аномалии незадолго до ее начала
- Нерешенная проблема
- Нет данных

Необходимо создать алгоритм, который позволил бы детектировать начало аномального поведения.



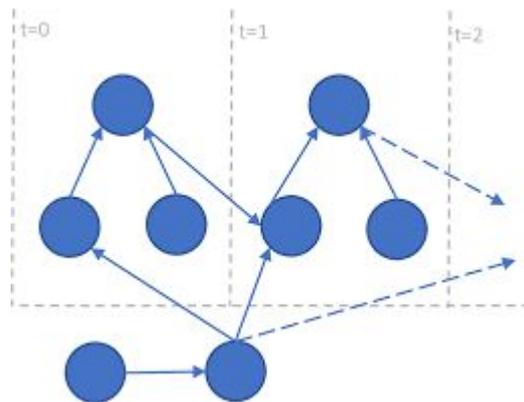
- Учет динамики
 - Внутренние связи между точками ряда(-ов) (корреляции или иные зависимости)
- Учет вариативности рядов
 - Модель должна уметь работать с реализациями процесса, то есть должен быть статистический анализ.

Динамическая байесовская сеть

ДБС - графическая модель (ациклический направленный граф), позволяющая факторизовать многомерное распределение с параметром t .
Главное отличие ДБС от БС – структура.

В каждый момент времени t -р есть:

1. Локальная структура. Точки временных рядов со сдвигом p и связи внутри момента t -р.
2. Глобальная структура. Множество локальных структур и связей между всеми срезами.



Гипотеза



Главная: перед моментом начала аномалии должен наблюдаться резкое изменение вероятности.

Предварительная:
покрытие модели достаточно для баланса между зависимостями нормальными и аномальными.

Постановка задачи (ДБС)



Извлечь изменение вероятности аномалии вблизи ее начала.

Динамическая байесовская сеть (с ограничениями на структуру) позволяет моделировать МРВ вероятности аномалии в момент t .

Чем моделировать структурные связи?

1. Автокорреляция – SVM модели
2. Нелинейные связи – Нелинейные ядра
3. Использование представлений

Суть оптимизационной задачи

notears:

$$\ell(\mathbf{W}, \mathbf{A}) = \frac{1}{2n} \|\mathbf{X} - \mathbf{XW} - \mathbf{YA}\|_F^2.$$



Модель ряда:

Structural Equation Model (SVAR)

Loss:

\mathbf{W} - матрица весов, где ij элемент - вес между i и j узлами

\mathbf{A} - матрица весов, где поблочно соединены матрицы для связи с моментом t

Задача:

Минимизация лосса с регуляризацией + учет ацикличности

1. Датасеты (**все содержат только групповые аномалии**):
 - a. Skoltech Anomaly Benchmarks (SKAB)
 - b. Numenta Anomaly Benchmark (NAB)
 - c. The Application Server Dataset (ASD)
2. Перед обучением для каждой реализации от аномальной зоны делались отступы. Величина этих отступов – гиперпараметры.
3. Локальная структура постоянна.
4. Суммарно гиперпараметры:
 - a. Два регуляризационных параметра (на W и A)
 - b. Границы для W и A (какие значения матрицы весов считать за связь)
 - c. Отступы
 - d. Порядок авторегрессии



1. Для оценки способности реагировать нужно исследовать производную полученного сигнала и действовать на нем пики.
2. Пик должен быть зафиксирован в определенном окне до аномалии (время реагирования).
3. Метрика - время реагирования относительно заданного горизонта.



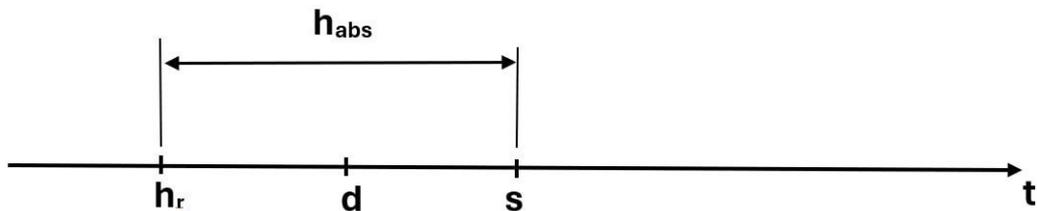
h_{abs} - абсол. горизонт

h_r - относи. горизонт

d - точка детекции

s - начало аномалии

$$\alpha = (s - d) / h_{abs}$$



Результат (SKAB)

ИТМО

Среднее значение метрики:
0.77

Абсолютный горизонт:
400 единиц времени

Модель для предсказания:
logistic regression

Порядок: 5

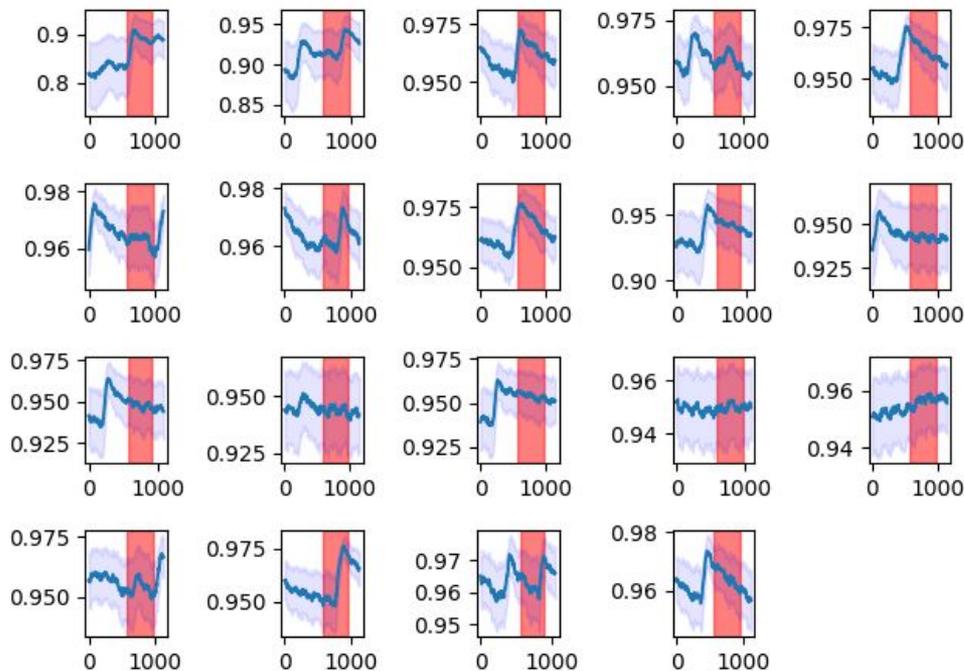


Рис.1. Динамика вероятности на тестовом датасете, красные поля - зоны аномалий.

Результат (SKAB)

1. Проблемы со стабильностью
2. Большая неопределенность

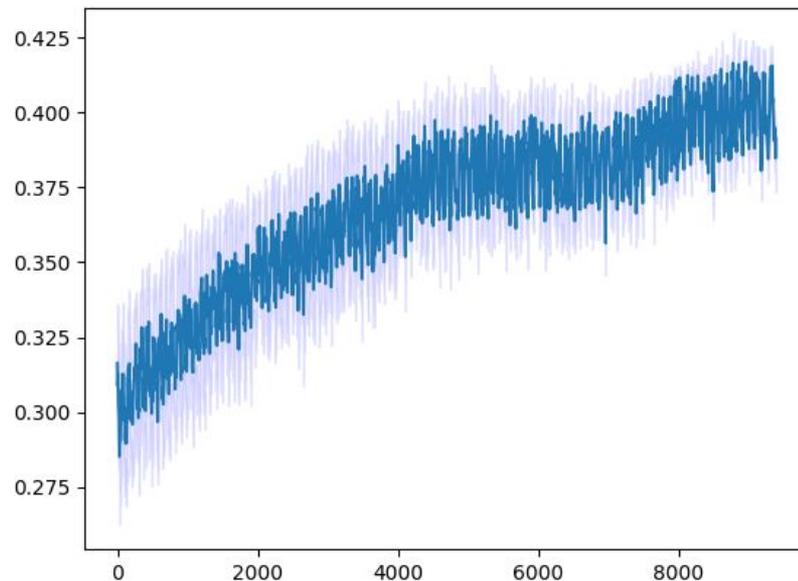


Рис.2. Динамика на датасете без аномалий

Запуски на других датасетах

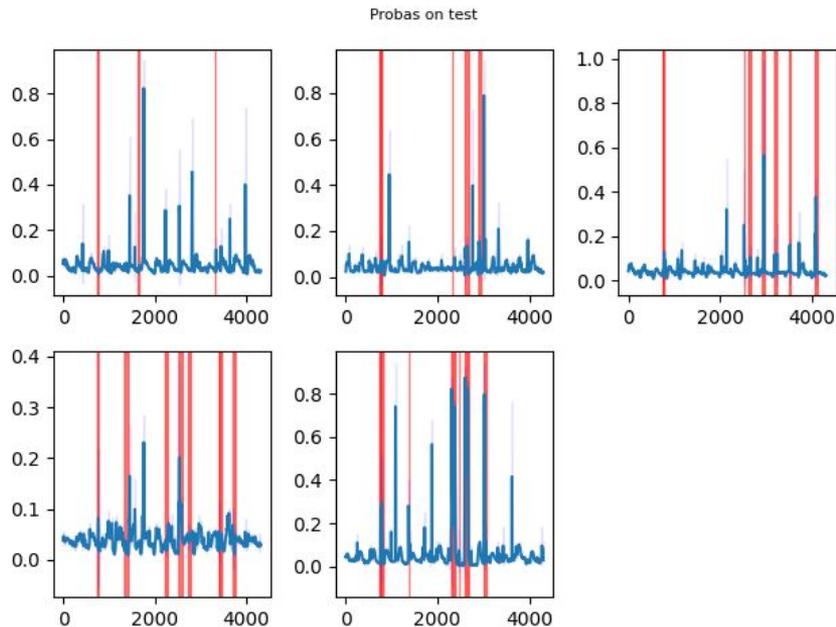


Рис.3 Результаты запуска AWS

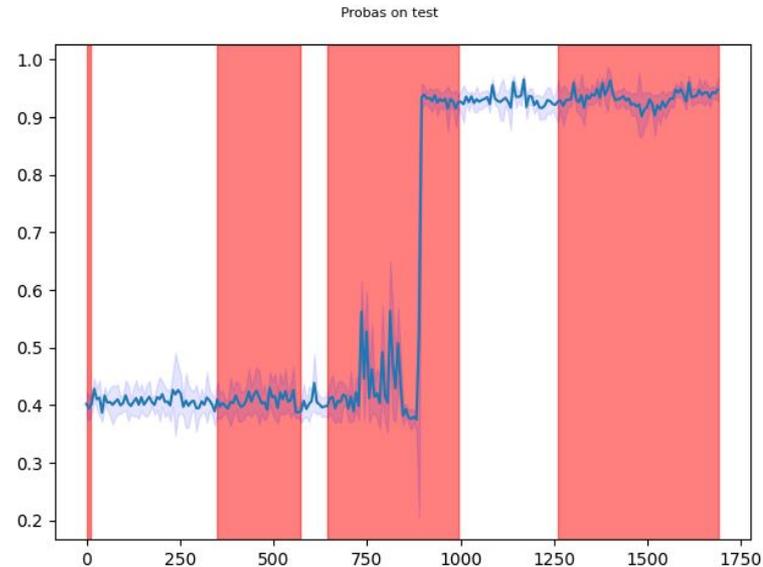


Рис.4 Результаты запуска NAB

Общие вопросы



1. Как обосновать пик перед началом аномалии?
2. Как гарантировать покрытие моделью всей структуры ряда?
3. Модель ряда – какую правильно использовать?

Данные

1. Где найти хорошие данные?

Причем тут Open Source?

1. Код к статьям авторы, как правило, не выкладывают. Если кто-то воспроизводит эксперимент (со своими изменениями), приходится писать код с нуля.
2. Если код есть и если качество кода низкое, котрибьюторы могли бы это исправить.
3. Open Source and Science.
4. (Open Source) Benchmarks.



1. Для решения задачи предсказания аномалии лучше использовать методы обучения без учителя.
2. Для моделирования ряда нужен инструмент, которые бы моделировать динамическую составляющую.
3. ДБС требуют очень тонкой настройки гиперпараметров и сильно переобучаются.
4. ДБС и прикладные примеры должны быть часть open source (будет реализовано в ВАМТ).



**Спасибо
за внимание!**

ITMO *re than a*
UNIVERSITY

Ваши контакты