СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Все процессы можно разделить на две большие группы: стационарные и нестационарные. Развитие физики и химии начиналось именно с работ со стационарными (равновесными) системами, то есть с такими, чья функция перехода из одного состояния в другое не изменяется или с процессами в равновесном состоянии. Однако и физики, и химики довольно быстро осознали, что такой подход ограничен и что большая часть процессов в мире является все-таки нестационарными. Уже в 1890 году Пуанкаре встретил ”кошмар хаоса”. Новые загадки привели к развитию двух областей: теории стохастических процессов как продолжение теории вероятности и теории динамических систем как продолжение теории дифференциальных уравнений [1].

Одна из основных передовых проблем в науке – это моделирование динамических данных и методы работы с ними. Под динамическими данными понимается любые данные, которые имеют эволюцию во времени, а вариации параметров и/или начальных условий в ней приводит к полному изменению поведения системы.

Например, есть работа, в которой решается проблемы развития клеток, и поверхность для их потенциального развития является динамическая система с бифуркациями — логистическое отображение [2].

Следующая проблема, которая возникает при работе с динамикой – это вопрос об извлечении информации из временного ряда (который еще и обычно многомерный), где каждая переменная имеет свою зависимость от времени и переменные связаны друг с другом. Временные ряды такой сложности возникают повсеместно в физике (гидродинамика), экономике, биологии (генетика). Это проблема также будет обсуждена.

Методы работы с такими данными, которые учитывали бы природу системы, показали результат лишь в последнем десятилетии, и они открыли нам совершенно новый взгляд на мир и эмерджентность. Несмотря на это открытие, у фронтирных исследований не хватает одной очень важной детали – стохастичности и учета неопределенности. Это своего рода наследство от прошлых работ, где природу системы старались всячески игнорировать. На самом же деле мир является статистическим и стохастическим, но взаимодействие элементов приводит иногда к детерминированности, но чаще к возникновению новых свойств.

Новый взгляд на процесс неизбежно должен привести к новому взгляду на аномлию и вообще на процесс. Если раньше ученые работали с системами, которые позволяли усреднять свойства, открывая возможность для использования аппарата либо статистики с ее математическим ожиданием и параметризацией, либо дифференциального исчисления с ее интегральными переходами к макросвойствам в физике (что очень сильно не любят математики).

В общем случае, понимая под аномалией отклонение от тренда, было разработано очень много методов их поиска: различные оценки в статистике и различные декомпозиции из математического анализа и линейной алгебры. Все методы, в сущности, являются декомпозицией и, соответственно, все требуют аксиому о состоятельности разделения сигнала на собственно сам сигнал и шум и аксиому о гомогенности такого распределения (на каждом шаге сигнала существует постоянная форма для функции разложения на сигнал и шум). Однако современные результаты в области динамических систем приводят нас к тому, что шум играет какую-то роль в формировании сигнала и что если усреднять ряд траекторий истинный сигнал не получить. Например, в спектроскопии для того, чтобы избавиться от шума снимают несколько спектров и берут средний, таким образом сглаживая сигнал.

Заполучив **интерпретируемую** модель, мы создали бы не только мощный инструмент для работы с аномлиями (для генерации аномльных временных рядов, для предсказания горизонта аномалий из истории), но и углубили понимание роли того, что мы раньше считали шумом, в процессах вообще.

Цель работы – это решение проблемы предсказания горизонта аномалий в динамических данных с заданным уровнем точности. Именно поэтому большая часть работы будет посвящена динамическим байесовским сетям. Однако существует проблема с обучением таких сетей: оно очень ресурсоемкое. Поскольку данная тема, так или иначе, потребует проверки практикой, мы выбрали задачу крайне нестандартную: предсказывание горизонта возникновения аномалии во временных рядах.

Цель работы: Предсказание горизонта появления аномалий во временных рядах с заданным уровнем точности

Задачи:

1. Сбор и анализ литературы
    а) Анализ способов работы с аномалиями во временных рядах
    б) Анализ проблем в методах
    в) Анализ применения методов работы с динамическими системами к временным рядам
    г) Поиск связи динамических процессов и временных рядов;
2. Предложения мат. аппарата для моделирования детерминированной и случайной составляющей
3. Реализация продвинутых алгоритмов обучения БС, если имеющихся будет недостаточно
4. Комплексная реализация алгоритма предсказания аномалий
5. Экспериментальные исследования

1 ЛИТЕРАТУРНЫЙ ОБЗОР

1.1 Динамические системы

1.1.1 Определение

Под динамической системой понимается система, функция перехода в которой является зависимой от времени и от предыдущего состояния системы. Само определение уже порождает сложности, потому что в классических моделях обычно работают с аналитическими объектами, которые могут однозначно поставить для каждого момента времени искомую величину. В динамических системах же приходится работать с функциями на каком-либо интервале (для дифференциального уравнения – на бесконечно малом дифференциале, для разностного – на конечном отрезке динамической оси). Однако же, если удается найти хотя бы вид дифференциального уравнения, то раскрываются удивительные свойства системы, которые поразительно совпадают с реальными явлениями (см. например, [3]). Иными словами, в большинстве случае, аналитический вид функции не существует (исключение, конечно же, решение дифференциального уравнения), а состояние системы можно исследовать только либо построив векторное поле (поток) для непрерывных систем, либо итеративно просчитать точки. В обоих случаях, методом изучения является построяние фазовоого пространства (определение см. далее).

С математической точки зрения под динамической системой вообще понимается любая система, в которой есть отображение, применяемое последовательно и на каждой итерации отображение использует значения из предыдущего шага.

Определение 1.1.1 (из [4]). Пусть X множество, $x \in X$ - точка на множестве, а $\varphi : X \to X$ - отображение множества в себя:

$$
\begin{aligned}
&x_0 = x = \varphi^0(x)\\
&x_1 = \varphi(x)\\
&x_2 = \varphi(x_1) = \varphi \circ \varphi(x) = \varphi^2(x)\\
&\ \ \vdots\\
&x_j = \varphi \circ \ldots \circ \varphi(x) = \varphi^j(x)
\end{aligned}
$$

Данные системы обладают удивительными свойствами, если смотреть на их траектории в фазовом пространстве. У таких систем в фазовом пространстве при определенных условиях появляются аттракторы различных типов, то есть точки, к которым все движение с достаточным количеством шагов сходится.

Определение 1.1.2. Траектория (в некоторых источниках орбита) - последовательность точек $\{x_j\}, j \geq 0$ полученные из последовательного преобразования φ. Множество этих точек находится в фазовом пространстве.

Фазовое пространство – это график зависимости движения точки, где непосредственно оси времени нет, но каждая точка показывает состояние системы в момент времени.

Однако важно отметить, что не всякая динамическая система является истинно таковой. [4] показано, что если использовать в качестве отображение жесткий поворот на окружности, то статистически точки орбиты распределяются равномерно на окружности, и динамическая система сводится к детерминированной. Такие случаи, конечно же, мы в работе не будем рассматривать, однако они показывают, что нужно быть крайне осторожным с выбором данных: не всякие данные дадут динамическую систему. Также данный пример ставит вопрос: какова роль стохастичности в формировании динамики?

Когда разрабатывались динамические системы, вопрос перевода временного ряда к динамике не возникал, поскольку дифференциальные (разностные) уравнения, их описывающие, выводились вручную, в подходе на основе данных не было необходимости. Однако проблему восстановления фазового пространства начали изучать лишь в последнем десятилетии.

Большой прогресс в этой области произошел, когда была сформулирована и доказана теорема Такена [5].

Определение 1.1.3. Пусть M компактное многообразие с размерностью d, тогда гарантировано существует представление Φ с размерностью $2d+1$, которое диффеоморфно M:

$$\Phi : M \to \mathbb{R}^{2d+1} \tag{1}$$

1.1.2 Теорема Такена

Данную теорему в современных исследованиях модифицировали и создали метод time-delayed embedding. Идея метода заключается в сдвиге временного ряда на разные величины, а затем использование каждого ряд как координату в фазовом пространстве. [6] показали пример, в котором диффеоморфизм обучался при помощи нейронных сетей. Другие исследователи [7] использовали time-delayed embedding, чтобы предсказать поведение сложного временного ряда для генетических данных. [8] использовали данные с мозга трансгенных мышей и ряд внешних переменных (диаметр зрачка, движение усиков и собственно активность кальциевых каналов, которые находятся в пресинаптической области и отвечают за передачу спайка). Аттрактор, полученной в этой работе, показан на Рисунке 1.

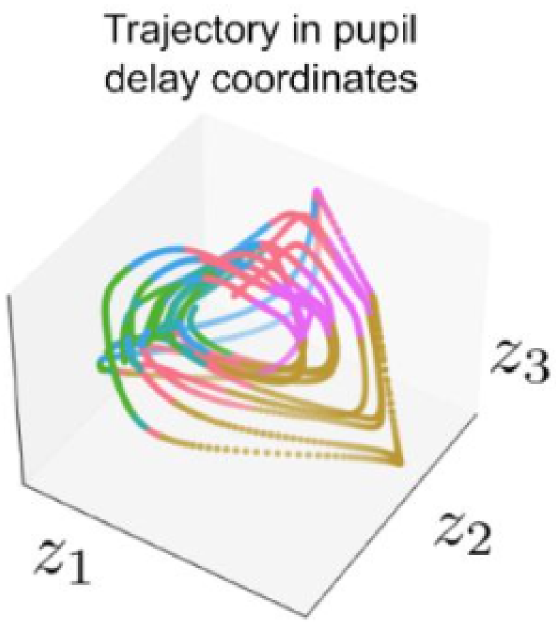

Рисунок 1 — Полный аттрактор, полученный из временного ряда с динамикой зрачка. Источник: [8]

На данный момент основным недостатком такого подхода является отсутствие интерпретируемости (потому что используются энкодеры) и отсуствие учета зависимости между временныи рядами. Имея в распоряжении метод восстановления динамики, по-прежнему не понятно, какие взаимосвязи существуют внутри динамической системы. Однако исследования показывают, что необходимо не так уж много данных для восстановления аттрактора и его использования, что потенциально открывает менее ресурсоемкие методы оптимизации.

Проблема с динамическими данными и шумом не только в том, чтобы понять где сигнал, а где шум, но и понять в какой момент шум становится сигналом. Например, в броуновском движении, полностью хаотическом процессе, нельзя выловить никакой сигнал без ухищрений, все макросвойства определены за счет этого статистического процесса (давление на стенки сосуда, длина пробега и др.). Макросвойства выводятся из движения молекул. Однако если меняются условия (например, молекулы стали очень длинными и негнущимися), то появляются свойства, которые из каждой молекулы не вывести. В ходе нагревания такая система будет обретать новую мезофазу, и на данный момент существует модели работы с такими система лишь после того, как они были сформированы. Такие модели иногда делают эвристические предположения о том, какие факторы влияют на возникновение мезофазы, но общий принцип, который бы связал это с другими похожими системам до сих пор не открыт. В примере с ЖК полимерами, мы можем построить модель (дискретная модель Снорри-Флори), которая даст приемлимый результат, если задача найти условие, при котором реализуются нужные межмолекулярные взаимодействия, однако почему это происходит – вопрос открытый.

Метод Такена был использован в [9], где авторы справедливо отмечают сложность использования метода (вычислительная сложность: $O(n^4)$). Байесовские сети уже показали себя мощным инструментом декомпозиции, понизив сложность моделирования многомерной случайной велчины, поэтому исследователи изучают вопрос о применимости ДБС и time-delayed embedding.

1.2 Стохастические системы

1.2.1 Определение

Стохастический (случайный) процесс представляет собой множество случайных величин, которые каким-либо образом меняются во времени. Для каждой случайной величины существует одно и только одно $t \in T$.

Теория случайных процессов является расширением математической статистики, в котором случайные величины зависят не только от элементарного исхода, но и еще от параметра – времени.

Рассмотрим переход от простой случайной величины к темпоральной и какие понятия такой переход порождает.

$\xi(\omega)$ переходит в $\xi(\omega, t)$, где ω – элементарный исход, ξ – случайная величина, t – динамическая переменная. Если зафиксировать время t_0, то будет снова случайная величина.

Для каждого $t \in T$ существует свое распределение. В классической теории рассматривается только те случаи, когда весь процесс задан одним классом распределений и для каждого t свой набор параметров.

Если из каждого распределения взять по выборке размером 1, то получится временной ряд, который называется реализацией процесса. Множество всех возможных реализаций называется ансамблем.

1.2.2 Ограничения стохастических процессов

Все случайные величины одного процесса должны принадлежать одному пространству состояний. Что само по себе уже является серьезным ограничением, поскольку в реальности, как правило, имеется динамическая ось (например, время, глубина, длина волны) и множество реализаций разных процессов, которые связаны между собой. Моделирование процесса как случайного предполагает существование совместного распределения, которое не является факторизованным, то есть, например, в случае гауссовского процесса вместо рассмотрения одного среднего и одной дисперсии рассматривается матрица. В общем случае, вопрос о независимостях возникает лишь постфактум, когда совместное распредление известно, за исключением марковского процесса, в котором постулируется попарная обусловленность между X_t и X_{t+1} для любого $t \in T$.

Несмотря на ограничение выше, такая теория оказывается приближенной к ограниченной (но не упрощенной) реальности. Например, можно рассмотреть количество частиц от распада радиоактивного вещества, которое регистрирует датчик. Вероятность столкновения с датчиком крайне мала, но частиц много, поэтому из закона редких событий делается предположение о пуассоновском распределении частиц на интервале времени. Весь процесс в таком случае будет пуассоновским.

Другая крайне важная проблема кроется в связи между случайностью и самой реальностью. Обычно, со случайностью работают, полагая, что при большом количестве реализаций системы можно, усреднив их, получить истинную картину. В данной работе мы будем рассматривать системы, которые имеют под собой сложное и нелинейное поведение.

Понимая динамику исключительно как ансамбль возможных состояний, мы попадаем в ловушку статистической физики. Рассмотрим эту проблему с двух крайностей.

Концентрация. Когда химик говорит, что раствор карбоната натрия имеет концентрацию 1 моль/литр, это означает, что каким бы образом он ни отобрал 0.001 л (1 мл) из этого раствора, его концентрация будет пропорциональна исходному раствору. Количество частиц настолько велико, что флуктуации концентрации будут незначительны. Однако же представим себе концентрацию очень маленькую, например 10 молекул карбонат натрия на 6 * 1023 молекул воды (10 молекул / 1 моль воды). Рассуждения по аналогии выше будут уже неверны, ведь флуктуации концентрации будут столь велики, что среднее по ансамблю за тот же интервал времени будет уже не равномерным. Естественно (как любят делать химики), обычно такой маленькой концентрацией просто пренебрегают и получают нужный результат из расчетов, однако наша же задача показать, что среднее по ансамблю не всегда будет работать.

Сложное поведение. Допустим, что количество частиц теперь достаточно, чтобы брать среднее по ансамблю, и случай, кажется, хороший. Но почему-то в растворе частицы перестают вести себя статистически, как материальные точки. Например, химик работает с раствором какой-то длинной молекулы (но не с полимером). И неожиданно раствор вместо того, чтобы проявлять свойства статистические, начинает проявлять свойства динамические: при нагревании возникают оптические свойства, которые никак не выводятся из свойств отдельной молекулы. Для нас это означает, что рассматривая только стохастический процесс, мы рискуем упустить очень важный элемент – динамику. Иными словам, нам следует задать вопрос: каково взаимодействие между ансамблями? Действительно ли набора случайных переменных достаточно, чтобы описать все возможные ансамбли?

Процесс в стохастической системе лучше всего понимать как игру: в каждый момент времени есть свое распределение значений состояния, и система берет значение из этого распределения. Таким образом, в ходе нескольких шагов формируется реализация или траектория процесса. Математический аппарат стохастических систем был разработан для процессов.

Аппарат стохастических систем широко используется в обучении с подкреплением. Авторы работы полагают, что методы обучения с подкреплением (или хотя бы постановка задачи в форме игры) могут быть применены к динамическим байесовским сетям. Например, для стационарных (обычных) байесовских сетей такие методы уже существуют [10].

Из данной области можно почерпнуть много идей для понимания процесса: является ли всякий процесс стохастичным? Когда его можно свести к детерминированному?

1.3 Графические вероятностные модели. Динамические байесовские сети

1.3.1 Определение

Графовые вероятностные модели – это модели, которые позволяют моделировать многомерное распределение в виде ненаправленного ациклического графа. В байесовских сетях узлы показывают признаки как случайные переменные, а связи отображают статистические зависимости между признаками. Такие модели позволяют получать условное распределение, если дан ряд значений предшественников.

Основная идея байесовских сетей заключается в декомпозиции многомерной случайной величины на факторы X_j (которые представлены данными, поиск факторов осуществляется исследователем), которые зависят только от родителей $Pa(X_j)$):

$$\mathbb{P}(X_1, \ldots, X_p) = \prod_{j=1}^{p} \mathbb{P}(X_j | Pa(X_j)) \tag{2}$$

Несмотря на высокую интерпретируемость, такие модели имеют ряд недостатков: их крайне сложно обучать. Классические методы с использованием оценивающей функций осуществляют комбинаторную оптимизацию, и сложность растет экспоненциально с ростом количества узлов. В последнее время появляются непрерывные методы структурного обучения, которые используют регрессионные методы [11].

1.3.2 Временные ряды и ГВМ

Существуют попытки обучать байесовские сети на временных рядах, такие модели называются динамические байесовские сети (Рис.2). Их обучение крайне проблематично: для каждого временного среза нужно хранить структуру графа, что резко повышает сложность из-за роста узлов; сверх этого необходимо также обучать связи между временными срезами, то есть пространство поиска еще больше (по сравнению со стационарными). Все это превращает обучение ДБС в серьезное испытание для исследователя.

Изначально, ДБС разрабатывали статистики, и они понимали процесс исключительно как первопричину и не делали никакого акцента на системе, которая этот процесс порождает. Очевидно, это привело к крайне сложным методам, которые пытаются учесть все и везде. Один из таких методов описан, например, здесь [12]. Исследователи использовали методы стркутрной авторегрессии, которые полагают, что сигнал в каждый момент времени можно представить как линейную комбинацию истории этого сигнала и истории других сигналов, получая в итоге матрица коэффициентов.

Были обнаружены работы, в которых данные содержат не так много признаков. [13] использовали minimum description length (MDL) scoring function, LL score и динамические переменные вместе со статичными. Однако даже им пришлось наложить ограничения на связи внутри временного среза. Данная работа по неизвестной причине содержит лишь 8 предсказанных шагов и не содержит вид полученного графа, хоть и качество работы оказалось крайне высоким.

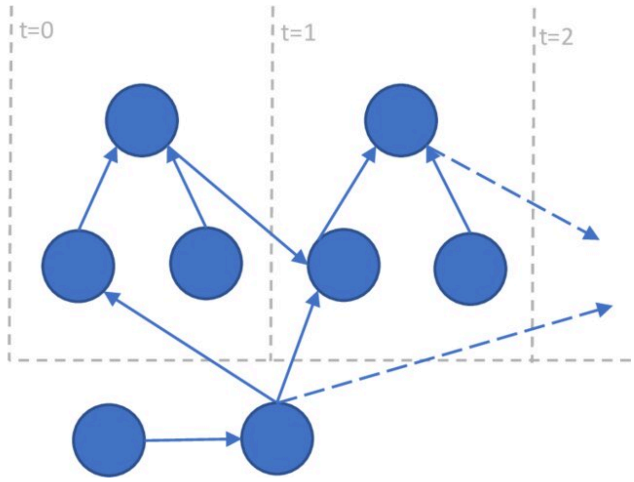

Рисунок 2 — Пример динамической сети. Снизу узлы-шаги во временном ряду. Источник: [14]

Исследовательская работа автора направлена в первую очередь на расширение методов обучения ДБС с использованием знания динамики.

1.3.3 Области для исследования

Как было уже показано, временные ряды могут образовывать не только временную динамику, но и еще пространственную. Безусловно, вопрос о существовании пространственно-временных представлений временного ряда в природе остается открытым, однако он вне рамок данной работы. Для авторов важно, что временной ряд сложной природы не просто линейная комбинация истории, а более сложное явление. Из этого вытекает явная необходимость учета еще и пространственной состовляющей, ведь, как было показано, теорема Такена позволяет с поразительной эффективностью работать со сложными и непредсказуемыми данными.

Более того, делатальное изучение могло бы пролить свет на динамические системы, позволив, наприрмер, находить точки возникновения новых свойств как аттракторы, хотя последнее явно нужнается в подкреплении результатами экспериментов.

1.4 Современное состояние проблемы предсказания аномалий

В ходе изучения было обнаружено небольшое терминологическое разногласие в проблемах. Так, например, говоря про anomaly prediction и anomaly detection авторы подразумевают иногда совершенно разные понятия, а иногда взаимозаменяемые. Природа этого разногласия будет обсуждена в этой главе, а границы и договренности установлены.

1.4.1 Anomaly Detection

Наиболее простая и наиболее развится область изучения, в которой решается задача обнаружения аномалий post factum. По сути своей этой либо задача классификации исходя из истории, либо задача кластеризации на базе какого-нибудь расстояния. Мы не будем обсуждать в рамках данной работы эти методы, потому что они не подходят под поставленную задачу.

1.4.2 Early Classification of Time Series

На практике куда ценнее не просто понять, в какой момент времени произошла аномалия, но и еще делать это заранее, анализируя поток, а не историю. С этой целью была выделена область early classification. Мы поместили этот класс задач рядом с anomaly detection, потому что в этих задачах просто добавляется новый параметр delay, который также подвергается минимизации, в связи с этим эта задача не является анализом потока, но в ней минимизируются размер истории. Способов решить задачу примерно столько же, сколько и придуманных классификаторов (см. Таблицу 2 в [15]). Их классификация представлена на рисунке 3.

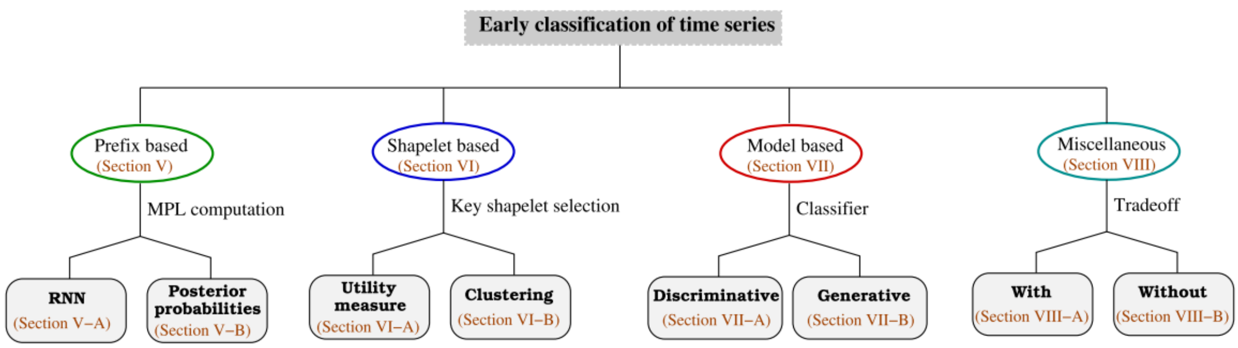

Рисунок 3 — Класификация алгоритмов Early Classification of Time Series. Источник: [15]

Основная идея этого семейства алгоритмов в подборе такой границы, при которой истории окажется достаточной для классификации аномалии.

Однако, данный подход страдает из-за отсутствия интепретируемости. Дизайн метода совершенно не учитывает внутреннюю динамику системы, а значит очень легко переобучается. В связи с этим, например, [15] отмечают, что данный подход не учитывает ”Concept drift момент когда система ”эволюционирует более точно - меняет свое состояние или фазу. Это основной недосток вообще всех существующих подходов, которой, по мнению авторов, лежит в понимании временного ряда как достаточной репрезентации системы, хотя для последнего нет никаких оснований.

1.4.3 Anomaly Prediction

Большая часть работ, которая содержит хоть какую-то взаимосвязь с anomaly prediction оказывается простым продолжением задачи time-series forecasting (предсказывание временного ряда). Авторы статей применяют один из методов предсказывания временного ряда, а затем, если предсказание не удалось, то считают, что произошла аномалия. В сущности своей, такой подход не решает действительно задачу anomaly prediction, ведь само название предполагает предсказывание не временного ряда, а момента возникновения аномалии, либо же интервал аномального поведения. В связи с этой ошибкой, некоторые работы содержат очень комичные заявления. Например, в [16] авторы использовали VAE, обучили скрытые представления, а затем через скрытые представления предсказывали сигнал. Однако в момент оценки алгоритма, авторам приходится написать: ”By comparing the predictions of our method with the actual future signals, the effectiveness of the proposed model in detecting and predicting anomalies could be assessed.”. В общем случае, почти во всех статьях авторы полагаются на обобщение моделью всего домена, что в принципе является невозможным в силу no free lunch теоремы.

Задача Anomaly Prediction является фронтирной и осваивается сейчас лишь поверхностно, поскольку никакого иного подхода, нежели как к классификации, к данной проблеме не было найдено.

1.4.4 Классические методы и SOTA

Единственное, что сейчас разработали следующую схему. Дан сигнал, где есть разметка с аномалиями. Задача поместить эту последовательность в какое-то скрытое представление модели, предсказать подпоследовательность, а затем провести классификацию либо с использованием предсказанных значений [17], либо с использованием значений loss функции (или score) [18].

Есть довольно старые методы, которые до сих пор считаются SOTA для предсказания: различные стрктурные методы (SVAR), иногда даже переводят в задачу Linear Regression и считают не момент предсказания аномалии, а долю аномалии в определенном интервале (anomaly rate). Среди подходов глубокого обучения опробованы и оценены почти все существующие модели: GAN + LSTM [19], ConvLSTM [], STG2Seq [], LSTM + SAX [], VAE [], Attention-Based Networks [].

В последнее время выходят алгоритмы, которые отмечают недостаток существующих техник: исопльзование лишь временного домена. Вместо этого, они предлагают использовать как пространственный, так и временной домен, но (вероятнее всего, пока что) в задаче anomaly detection (например, Graph Attention Network для пространственного домена и LSTM для временного) [].

1.4.5 Применением ГВМ для предсказания аномалий

Моделирование случайной величины приводит к вполне очевидному способу детектирования аномалий: имея набор значений, которые принимают факторы, мы можем получить условное распределение и оценить правдоподобие и вероятность данного занчения, и, на основе какой-нибудь установленной границы, классифицировать значение. К задачам предсказывания аномалий ГВМ не применялись.

1.5 Заключение

Агрегируя все подходы к пониманию и работе с временными рядами, автор пришел к тому, что существует некоторое разногласие в понимании природы временного ряда. Первая позиция утверждает, будто ряд это первопричина, а породившая его система не имеет значения (назовем это инвариантность выборки, то есть какую бы выборку из набора распределений вы бы не взяли, свойства по ансамблю не меняются). Эти методs хороши, если мы работаем с системой, свойства которой изотропны. В таком случае, мы вполне можем полагаться на статистические характеристики и считать, что корреляция даст нам каузацию. Первые, кто стал применять это в своей области, - это инженеры, поэтому очень много примеров связано с цифоровой обработкой сигнала. В этом подходе разработаны методы декомпозиции, позволяющие разложить сигнал либо аддитивно, либо мультипликативно. Ключевая идея здесь в том, что из-за однородности и хорошо перемешанной системы, свойства на микроуровне легко обощаются на макроуровень путем интегрирования [20; 21].

Однако, как было показано, сложный процесс не описать через линейную комбинацию: его природа нелинейная.

В исследовательской работе мы будем полагать следующее:

1. Существование стохастичности является следствием большой вариативности системы. Однако после реализации система становится детерминированной.
2. Вопрос о возникновении новых свойств динамической системы: будут учитываться лишь те динамические системы, функциональный вид преобразования в которой не меняется. То есть, например, спектры химических соединений нам не подходят: на каждом волном числе преобразование изменяетя.
3. В данной работе будут установлены ограничения на данные: нас интересуют лишь те случаи, когда система проявляет эмерджентные свойства (свойства, которые не вытекают из элементов системы).
4. Либо в фазовом пространстве должен быть странный аттрактор (именно на такие системы нацелена работа). Либо, если метод позволит вписать и другие аттракторы, то в работу будут включены эксперименты с ними.
5. Временной ряд должен иметь сложную природу. Это значит, что методы с декомпозицией не должны быть эффективными для таких рядов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. *Strogatz S.H.* Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering. — Westview Press, 2000.
2. *Wang J.*, *Zhang K.*, *Xu L.*, *Wang E.* Quantifying the Waddington landscape and biological paths for development and differentiation // Proceedings of the National Academy of Sciences. — 2011. — Vol. 108, no. 20. — P. 8257–8262. — eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.1017017108. — URL: https://www.pnas.org/doi/abs/10.1073/pnas.1017017108.
3. *Yadav S.*, *Saha S.K.*, *Kar R.* An application of the Kalman filter for EEG/ERP signal enhancement with the autoregressive realisation // Biomedical Signal Processing and Control. — 2023. — Vol. 86. — P. 105213. — URL: https://www.sciencedirect.com/science/article/pii/S1746809423006468.
4. *Moser J.*, *Zehnder E.*, *Mathematical Sciences C.I. of*, *Society A.M.* Notes on Dynamical Systems. — Courant Institute of Mathematical Sciences, New York University, 2005. — (Courant Lecture Notes Series). — URL: https://books.google.fr/books?id=IJVJngEACAAJ.
5. *Takens F.* Detecting strange attractors in turbulence // Dynamical Systems and Turbulence, Warwick 1980 / ed. by D. Rand, L.-S. Young. — Berlin, Heidelberg : Springer Berlin Heidelberg, 1981. — P. 366–381.
6. *Bakarji J.*, *Champion K.*, *Kutz J.N.*, *Brunton S.L.* Discovering Governing Equations from Partial Measurements with Deep Delay Autoencoders. — 2022. — arXiv: 2201.05136 [cs.LG].
7. *Mitra M.*, *Taylor P.*, *Hutchison C.*, *McLeish T.*, *Chakrabarti B.* Delayed self-regulation and time-dependent chemical drive leads to novel states in epigenetic landscapes // Interface. — 2014. — Nov. — Vol. 11, no. 100. — URL: https://eprints.whiterose.ac.uk/154995/ ; © 2014 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

8. *Raut R.V.*, *Rosenthal Z.P.*, *Wang X.*, [et al.]. Arousal as a universal embedding for spatiotemporal brain dynamics // bioRxiv. — 2023. — eprint: https://www.biorxiv.org/content/early/2023/11/08/2023.11.06.565918.full.pdf. — URL: https://www.biorxiv.org/content/early/2023/11/08/2023.11.06.565918.

9. *Галкин В.А. Гавриленко Т.В. Д.И.* Применимость теоремы Такенса об обнаружении «Странных аттракторов» для биологических систем // Сложность. Разум. Постнеклассика. — 2016. — URL: https://cyberleninka.ru/article/n/primenimost-teoremy-takensa-ob-obnaruzhenii-strannyh-attraktorov-dlya-biologicheskih-sistem%7C%D0%BD%D0%BE%D0%BC%D0%B5%D1%80=3.

10. *Zheng Z.*, *Wang C.*, *Gao X.*, *Chen G.* RBNets: A Reinforcement Learning Approach for Learning Bayesian Network Structure // Machine Learning and Knowledge Discovery in Databases: Research Track / ed. by D. Koutra, C. Plant, M. Gomez Rodriguez, E. Baralis, F. Bonchi. — Cham : Springer Nature Switzerland, 2023. — P. 193–208.

11. *Zheng X.*, *Aragam B.*, *Ravikumar P.*, *Xing E.P.* DAGs with NO TEARS: Continuous Optimization for Structure Learning. — 2018. — arXiv: 1803.01422 [stat.ML].

12. *Pamfil R.*, *Sriwattanaworachai N.*, *Desai S.*, [et al.]. DYNOTEARS: Structure Learning from Time-Series Data. — 2020. — arXiv: 2002.00498 [stat.ML].

13. *Leão T.*, *Madeira S.C.*, *Gromicho M.*, *de Carvalho M.*, *Carvalho A.M.* Learning dynamic Bayesian networks from time-dependent and time-independent data: Unraveling disease progression in Amyotrophic Lateral Sclerosis // Journal of Biomedical Informatics. — 2021. — Vol. 117. — P. 103730. — URL: https://www.sciencedirect.com/science/article/pii/S1532046421000599.

14. *Rothmund S.*, *Tengesdal T.*, *Brekke E.*, *Johansen T.* Intention modelling and inference for autonomous collision avoidance at sea. — 10/2021.

15. *Gupta A.*, *Gupta H.P.*, *Biswas B.*, *Dutta T.* Approaches and Applications of Early Classification of Time Series: A Review // IEEE Transactions on Artificial Intelligence. — 2020. — Vol. 1, no. 1. — P. 47–61.

16. *Chalongvorachai T.*, *Woraratpanya K.* LSCR: Latent Space Coordination Relation for Anomaly Prediction // 2023 15th International Conference on Information Technology and Electrical Engineering (ICITEE). — 2023. — P. 13–18.

17. *Lee M.-C.*, *Lin J.-C.*, *Gran E.G.* RePAD: Real-time Proactive Anomaly Detection for Time Series. — 2023. — arXiv: 2001.08922 [cs.LG].

18. *Li T.*, *Comer M.L.*, *Delp E.J.*, [et al.]. Anomaly Scoring for Prediction-Based Anomaly Detection in Time Series // 2020 IEEE Aerospace Conference. — 2020. — P. 1–7.

19. *Wang W.*, *Peng Z.*, *Wang S.*, [et al.]. IFP-ADAC: A Two-stage Interpretable Fault Prediction Model for Multivariate Time Series // 2021 22nd IEEE International Conference on Mobile Data Management (MDM). — 2021. — P. 29–38.

20. *Kirchgässner G.*, *Wolters J.* Introduction to Modern Time Series Analysis. — Berlin, Heidelberg : Springer-Verlag Berlin Heidelberg, 2007. — Online–Ressource. — URL: http://aleph.bib.uni-mannheim.de/F/?func=find-b&request=276336909&find_code=020&adjacent=N&local_base=MAN01PUBLIC&x=0&y=0.

21. *Montgomery D.*, *Jennings C.*, *Kulahci M.* Introduction to Time Series Analysis and Forecasting. — Wiley, 2011. — (Wiley Series in Probability and Statistics). — URL: https://books.google.ru/books?id=-qaFi0oOPAYC.