Текст

Слайд 2 (Введение: постановка задачи)

Будем называть временной ряд однородным, если его структура постоянна. Если под внешним воздействием в какой-то момент времени ряд терпит возмущение, в его структуре появляется разладка и возникает основная задача обнаружения разладки — найти этот момент возмущения. Под структурой ряда будем понимать подпространство сигнала.

Для обнаружения момента возмущения будем использовать функции обнаружения, основанные на скользящих отрезках рассматриваемого временного ряда и сообщать об обнаружении разладки при превышении функцией некоторого порога.

Будем рассматривать синусоидальные временные ряды, которые в момент времени Q меняют свою структуру.

Основной целью моей работы является создание системы с автоматическим выбора порога, способной обнаруживать разладку, порожденную изменением частоты периодики сигнала за указанный промежуток временни.

Слайд 3 (Введение: обозначения)

Для этого, введем следующие обозначения.

Назовем L - длинной окна, характеризующее длину векторов вложения исходного ряда;

$F^{(1)}$ и $F^{(2)}$ - базовые и тестовые подряды ряда F_N длины B и T соответственно;

Обозначим $U_l^{(1)}$ - собственные векторы траекторной матрицы ряда $F^{(1)}$.

Подмножество длины r индексов собственных векторов, принадлежащих сигналу ряда обозначим I большое, где r - ранг ряда.

Подпространство сигнала, являющейся линейной оболочкой, натянутой на собственные векторы с индексами из I большого обозначим готической L с индексами r и 1.

Векторы вложений длины L ряда F^2 обозначим как $X_1, \ldots, X_{K_2}$ с верхним индексом 2, где $K_2 = T - L + 1$.

Слайд 4 (Введение: инструменты поиска неоднородности)

В таких обозначениях, индекс неоднородности задается как сумма нормированных расстояний между векторами вложений тестового ряда и подпространством сигнала и характеризует несоответствие между рядом F^2 и структурой ряда F^1.

Рассматривая в качестве базовых и тестовых рядов все подряды ряда F_N мы получаем матрицу неоднородности, подмножества которой и являются функциями обнаружения неоднородности, а именно:

Строковая, где фиксирован базовый подряд, а тестовые итеративно смещаются;

Столбцовая, где, наоборот, фиксируется тестовый и смещаются базовые.

Диагональная, где базовый и тестовый отрезки смещаются вместе и симметричная, где базовый совпадает с тестовым.

Слайд 5 (Сравнение функций обнаружения)

Имея 4 функции обнаружения, установим, какая из них является лучшей в смысле момента преодоления порога gamma^* и скорости роста в зашумленных рядах.

Для этого сравним их численно и зададим неоднородность ряда F_N четырьмя способами - фазовым сдвигом, выбросом, изменением амплитуды и частоты.

Слайд 6 (Сравнение функций обнаружения)

Для численного эксперимента были выбраны следующие параметры:

Длина ряда N = 700, момент нарушения однородности Q = 301, длина окна L=60, дисперсия шума \sigma^2 = 0.25.

До момента Q частота omega_1 задавалась как 0.1, фазовый сдвиг phi_1 = 0, амплитуда C_1 = 1. После момента Q частота omega_2 равна 0.2, фазовый сдвиг pi/2, а амплитуда C_2 = 2.

Все варианты задания неоднородности рассматривались отдельно.

На данном изображении вид функций обнаружения при одной реализации шума для каждого типа задания разладки.

Численные тесты показали, что лучшими являются диагональная и строковая функция обнаружения. На графиках они соответствуют синим и красным кривым.

Мы будем изучать только строковую поскольку она вычислительно наименее затратна, в силу фиксирования подпространства сигнала и перебора векторов вложений тестового ряда для подсчета индекса неоднородности.

Слайд 7 (Аппроксимация индекса неоднородности)

Рассмотрим ряд F_N с неоднородностью, заданную изменением частоты. Попробуем упростить индекс неоднородности g, чтобы получить в явном виде его зависимость от частот до и после разладки.

В результате анализа я получил формулу, указанную на слайде. Причем для достаточно больших L и K_2 аппроксимация точна.

Слайд 8 (Аппроксимация индекса однородности: точность аппроксимации)

На первом графике показано поведение индекса неоднородности и его аппроксимации при стремлении частот друг к другу. На оси x показаны значения частоты omega_2, на оси y значения индексов. По кривым видно, что полученная формула повторяет поведение индекса неоднородности.

На втором графике показана зависимость значений рассматриваемых индексов от значений L. По кривым видно, что с ростом L оба значения сходятся друг к другу.

Слайд 9 (Аппроксимация переходного интервала)

Поскольку мы получили формулу, аппроксимирующую значения индекса неоднородности по указанным частотам до и после разладки, для построения порога срабатывания разрабатываемой системы, нам надо оценить поведение переходного интервала функции обнаружения.

Было установлено, что при маленьком L, по отношению к T, переходный интервал ведет себя примерно линейно. На графике синяя линия соответствует значению L=10 и на переходном интервале функция неоднородности ведет себя примерно линейно.

Слайд 10 (Система обнаружения момента возмущения)

Принимая во внимание линейность переходного интервала и наличие аппроксимации индекса неоднородности, перейдем к задаче построения системы.

Задача ставилась следующая — создать систему, способную обнаружить разладку в синусоидальных временных рядах за заданный промежуток времени. Причем разладку, порожденную изменением частоты сигнала. Сигнал о моменте возмущения подается при превышении функции обнаружения порога gamma^*.

Описание системы представлено на слайде.

На вход подается ряд F_N, промежуток k за который нужно определить разладку и минимальное для обнаружения раздладки отклонение от начальной частоты delta_{min}. Результатом работы системы должна быть оценка момента возмущения, содержащаяся в промежутке от Q до Q+k.

Введем также требование к ряду - его первая четверть должна быть однородной. Можно расценивать это как наличие неких исторических данных, на которых оценивается начальная частота ряда и нижняя граница порога gamma^*.

Оценивать нижнюю границу порога необходимо для минимизации ложных срабатываний системы, так как значения индекса неоднородности до разладки зависят от дисперсии шума. Я решил взять значение третьего квартиля.

В качестве значения самого порога gamma^* берем значение прямой в точке k, соединяющей значение нижней границы gamma^* и аппроксимации индекса неоднородности g с нижним индексом a от частот omega_1 и omega_{min}. omega_{min} задается как сумма начальной частоты ряда, оцененное на его однородной части и входного значения Delta_{min}. Численно проверено что значение аппроксимации индекса однородности для omega_{min}, заданной суммой и разностью при достаточно большом L равны.

На слайде показан график примера работы системы. Кривая a с индексом n-1 до момента Q равна значению gamma_{min}, после переходного интервала значению g с индексом a от частот omega_1 и omega_{min}. Переходный интервал - прямая, соединяющая значения до переходного интервала и после.

При параметрах, указанных на изображении, значение строковой функции неоднородности превзошло порог gamma^* за указанный промежуток времени k и система справилась с поставленной задачей.

Слайд 11 (Оценка качества системы)

Чтобы исключить случайность из результатов работы системы, оценим свободные для выбора параметры B, T и L.

Для этого введем введем следующие характеристики.

Будем считать, что произошло ложноположительное обнаружение неоднородности если Q с крышкой меньше Q . Если Q с крышкой лежит в отрезке от Q до Q+k, то у нас точное обнаружение . Если же Q с крышкой больше Q+k, то произошло ложноотрицательное обнаружение неоднородности.

Промоделировав 200 реализаций шума будем характеризовать систему с точки зрения вероятности ложноположительного, точного и ложноотрицательного обнаружений.

Слайд 12 (Оценка системы: параметр T)

По введенным метрикам оценим параметр T.

На слайде представлены графики оценки системы. На оси x указаны значения порогов от 0 до 1. Вертикальная пунктирная линия соответствует выбранному системой порогу gamma^*.

На верхнем графике значение параметра T = 70, на нижнем 130.

По графикам видно, что при уменьшении параметра T вероятность точного обнаружения, характеризуемая зеленой кривой, увеличивается.

Слайд 13 (Оценка системы: T - L)

При дальнейшем исследовании зависимости работы системы от параметра T было обнаружено, что при маленьком значении разности T и L скорость роста функции обнаружения на переходном интервале становится быстрее линейной, что обуславливает повышение вероятности точного обнаружения.

Так, на верхнем графике, при разности T и L равной 10, синяя кривая превосходит оранжевую практически на всем переходном интервале, а так как порог gamma^* выбирается исходя из значений прямой, аппроксимирующей переходный интервал, требование к обнаружению за промежуток k будет удовлетворено. На нижнем графике, при разности в 70, наблюдается обратная картина.

Слайд 14 (Оценка системы: параметр B)

Параметр B также влияет на устойчивость системы и его уменьшение может увеличить вероятность ложного срабатывания. Чем больше B, тем точнее определяется подпространство сигнала и до момента Q значения строковой функции обнаружения неоднородности имеют меньший разброс. Это мы и наблюдаем на графиках, где изображены строковые функции обнаружения для разных реализаций шума. На верхнем графике значение параметра B = 90, на нижнем 200.

Слайд 15 (Проблемы)

Основная проблема при тестировании системы для разных частот и разной дисперсии шума это корректность аппроксимации переходного интервала прямой линией. Обратите внимание на первый график на слайде. Если на вход системе поступит маленькое значение k, значения функции обнаружения на переходном интервале будут меньше значений аппроксимации на первых k точках после момента Q. Это частично нивелируется уменьшением разности параметров T и L. Однако мне встретились параметры сигнала, на которых даже в случае отсутствия шума вероятность ложноотрицательного обнаружения разладки была равна 1. Функция разладки данной ситуации приведена на втором графике, где частота периодики равна 1/7. На графике видно, что функция неоднородности начинает быстро расти с некоторой задержкой и это поведение нуждается в дальнейшем исследовании.

Однако, при достаточно большом значении k, например 30, при стандартном отклонении шума, не превышающем половину амплитуды сигнала, вероятность точного обнаружения для частот от 1/3 до 1/9 составляла не менее 98 процентов. А при k равном 45 - эта вероятность составляла 1.

Таким образом, при достаточно большом значении k, работа системы удовлетворяет требованиям поставленной задачи.

Спасибо за внимание.