

# Об одном подходе к обработке временных рядов

С.Н. Куклин, Д.Е. Намиот

**Аннотация**—В статье рассмотрены вопросы, связанные с обработкой данных измерений распределенного датчика акустических воздействий. Предложена модель, которая позволяет с высокой достоверностью определять события, связанные с внешним воздействием на акустический датчик. При этом, предложенная модель позволяет эффективно отсеивать события, связанные с естественным акустическим фоном и эффективно фильтровать ложные сигналы. Предложенная модель реализована в виде переносимого программного обеспечения, основанного на продуктах с открытым кодом. Достоверность модели проверена на ряде практических испытаний.

**Ключевые слова**—временной ряд, акустика, волоконно-оптические системы.

## I. ВВЕДЕНИЕ

В работе рассматривается обработка данных, получаемых от распределенной охранной волоконно-оптической системы [1].

Система представляет собой распределенный датчик акустических воздействий для охранных систем. Чувствительный элемент системы — стандартное телекоммуникационное одномодовое волокно (волоконно-оптический кабель). Система позволяет обнаружить любые виды деятельности, вызывающие колебания воды и почвы, происходящие на расстоянии до 100 метров от кабеля на всем его протяжении.

Преимуществами такой системы являются [2]:

низкая стоимость решения – необходимо всего одно устройство на 40 километровую линию;

устойчивость к электромагнитным воздействиям;

интеграция с охранными системами по протоколу Ethernet: видеонаблюдение, информирование оперативных служб и др.;

использование стандартного оптического волокна как чувствительного элемента (требуется одно волокно в кабеле);

не требует обслуживания, легкость монтажа (не требуется высокая квалификация);

высокая надежность системы (для блока управления — дублированные блоки питания, основные блоки с горячей заменой);

Области применения:

Статья получена 20 января 2015.

С.Н. Куклин – ОИЯИ (e-mail: snkuklin@gmail.com)

Д. Е. Намиот – старший научный сотрудник МГУ им. М.В. Ломоносова (e-mail: dnamiot@gmail.com).

Мониторинг и охрана нефтепроводов и газопроводов - система информирует о приближении к охраняемому объекту и позволяет оперативно реагировать до осуществления врезки в магистраль.

Контроль врезок в нефтепроводы.

Контроль несанкционированных работ возле трубопроводов.

Мониторинг движения внутритрубных снарядов.

Охрана периметра.

Охрана линий связи. Одна из наиболее частых причин выхода из строя ВОЛС обрыв оптического кабеля — случайное повреждение при проведении работ вблизи расположения кабельной линии и вандализм. Охранная система позволяет предупредить аварию, сообщив оператору системы о несанкционированных работах рядом с линией связи.

Система имеет встроенные средства анализа и обработки (рисунок 1).

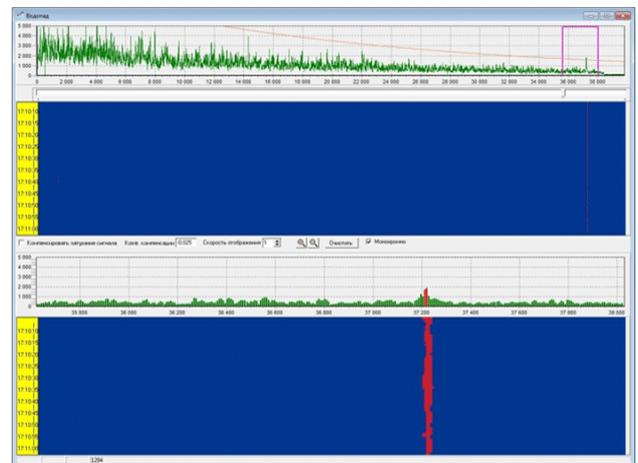


Рис. 1 Графический интерфейс

Вместе с тем, практические испытания выявили достаточно неприятные свойства встроенной обработки. Концептуально, они базируются на некоторой метрике для оцифрованной амплитуды колебаний. Основой для выдачи тревожного сигнала является превышение значения этого показателя над некоторым заданным уровнем. Но чувствительность данного устройства реально весьма высока, он действительно хорошо регистрирует возмущающие воздействия. Если мы имеем дело с некоторой “тихой” обстановкой, то возмущающий сигнал реально выделяется на общем фоне. Для “зашумленной” же местности, система выдает большое количество ложных срабатываний (false

positive, в классической формулировке). В данном случае имеются в виду не шумы измерительного устройства, а реальная акустическая обстановка в месте измерений. Если по природе решаемых задач мы имеем дело с длительными по времени воздействиями (например, работа строительной техники в течение десятков минут), то время воздействия можно использовать как дополнительный параметр для правил срабатывания, основанных на величине амплитуды воздействия. Но, например, для систем охраны периметра время воздействия будет всегда небольшим (пересечение кабеля). В таком случае множественные ложные срабатывания практически исключают применение устройства. Исходя из этого, было принято решение о разработке новой модели анализа измерений.

## II. ПРОГРАММНОЕ ВЗАИМОДЕЙСТВИЕ С ИЗМЕРИТЕЛЬНЫМ УСТРОЙСТВОМ

Измерительное устройство записывает результаты оцифровки (значения амплитуды) во внутреннюю базу данных MySQL. Каждые 500 мс с каждого метра волоконно-оптического кабеля получается 24 значения амплитуды. Оцифровка осуществляется для трех разных диапазонов частот (три частотных фильтра). Общая архитектура системы изображена на рисунке 2.

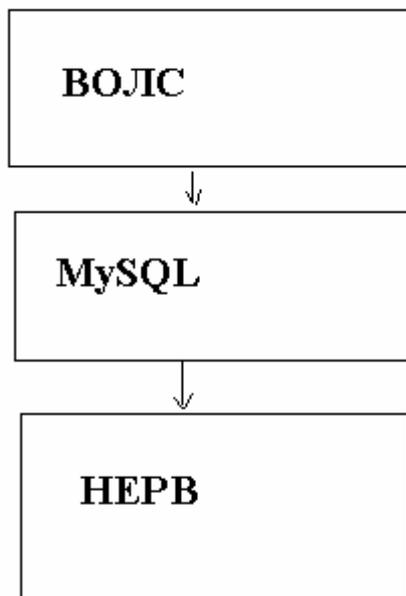


Рис. 2 Архитектура системы

Здесь:

ВОЛС – акустический датчик

MySQL – рабочая станция, аккумулирующая измерения

НЕРВ – это разработанная нами система.

Логически, акустический датчик и рабочая станция для получения результатов измерений представляют собой единый комплекс.

НЕРВ представляет собой следующий набор компонент:

Веб-сервер Apache

База данных MySQL

Набор веб-приложений (CGI-скриптов) на PHP.

Веб-приложения отвечают за анализ и интерпретацию результатов измерений, предоставляют пользовательский интерфейс для системы анализа, а также программный интерфейс (API) для внешних приложений.

В такой форме, приложение получается полностью переносимым. Веб-интерфейс обеспечивает дополнительные удобства пользователям – нет необходимости загружать отдельную программу для просмотра (мониторинга). Приложение для мониторинга автоматически поддерживает доступ с мобильных устройств (через мобильный браузер). Поскольку система анализа включает в свой состав собственный веб-сервер, легко обеспечить доступ к приложениям с произвольной машины как локальной, так и глобальной сети. Собственный веб-сервер позволяет организовать программный интерфейс (API) для сторонних приложений на базе HTTP запросов (REST модель) [3]. В целом, легко представить даже облачную модель обработки, построенную на основе данной архитектуры.

Ответ на вопрос о скорости обработки (скорости считывания данных) зависит, естественно, от мощности рабочей станции с системой НЕРВ. Профилирование системы и практические испытания показывают, например, что производительности рабочей станции с процессором Intel Core 2 Duo CPU E4600 2\*2/4GHz хватает для чтения и обработки в реальном времени данных измерений с 2500 метров оптического кабеля. Очевидно также, что подобная архитектура легко допускает горизонтальное масштабирование. Можно просто добавлять новые станции, отвечающие за обработку данных со своего линейного участка кабеля.

## III. ПРЕДСТАВЛЕНИЕ ИЗМЕРЕНИЙ

Для иллюстрации, покажем, как выглядят реальные измерения. На рисунке 3 изображены значения амплитуды по времени и расстоянию для одного из фильтров в “тихой” местности (участок вдоль грунтовой дороги на окраине леса вне населенных пунктов)

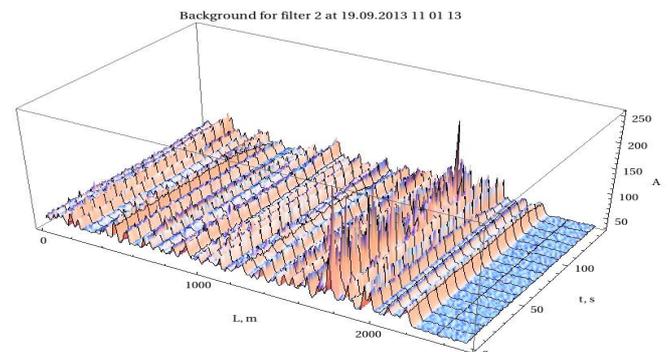


Рис. 3 “Тихий” участок

На рисунке 4 показаны измерения на отрезке кабеля в

городских условиях (Москва, 50м от загруженной улицы, 10-15м от линии метро)

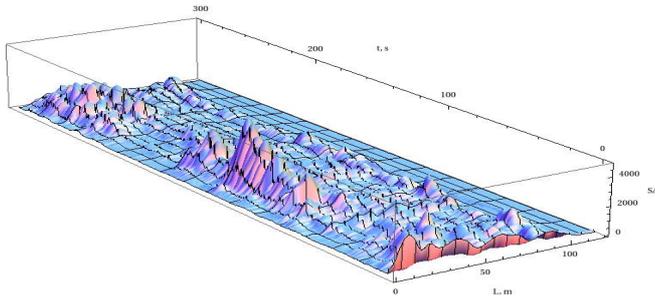


Рис. 4 “Шумный” участок

Оба графика построены в системе Wolfram Mathematica [4], которая и являлась основным рабочим инструментом на этапе отработки алгоритмов.

#### IV. АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

Данные, которые в итоге считываются из регистрирующего устройства, представляют собой типичный временной ряд. Точнее – три временных ряда (по числу фильтров оцифровки). Мы исходили из того, что не удастся подобрать такую комбинацию предельных уровней, которая подходила бы для всех условий функционирования. Это означает, что система должна обучаться, получать информацию о “естественном” фоне. А уже потом, с учетом имеющейся априорной информации осуществлять классификацию измерений. Эта концепция соответствует другим работам в области охранных систем, использующих сейсмологические данные [5]. В целом, определение событий в таких системах основано на сравнении текущих измерений с некоторой историей [6].

Из анализа данных измерений (раздел III) было отмечено, что появление внешнего воздействия, в первую очередь, отражается на скорости роста амплитуды. На картинках, подобных приведенным в разделе III, пики становятся более крутыми. Хотя по абсолютному значению могут и не превышать имеющегося фона. Это наблюдение привело к идее о том, что анализировать историю нужно с точки зрения разностной модели. Нужно снабдить систему знаниями о возможных скоростях изменения.

В соответствии с этим, на этапе обучения для каждого метра линейной дистанции получались последовательные значения амплитуды за два такта (такт – 500мс). На каждом такте измерительное устройство выдает 24 значения. Вычислялось среднее этих измерений и частное средних:

$$\Delta = \frac{Avg(t)}{Avg(t-1)} \quad (1)$$

Далее, для каждого метра накапливались три значения:

$Min(\Delta)$

$Avg(\Delta)$

$Max(\Delta)$

На практике мы использовали обучение в интервале времени 5-10 минут.

Вычисленные значения сохранялись в собственной базе данных системы НЕРВ. Далее в режиме анализа система последовательно считывает значения за каждые два такта. Также как и при обучении – значения амплитуд усредняются и вычисляется частное средних. Это полученное значение сравнивается с историческим максимумом (для данного метра). Если оно превышает максимум, то это может быть искомым сигналом. Для заключения о достоверности используется следующее предположение. Сигнал не может быть точечным (привязанным к одному метру). Мы должны увидеть группу таких превышений на некотором метрическом интервале. Важно, что речь идет именно об интервале. Это не обязательно несколько метров подряд. Соответственно, для принятия решения о выдаче сигнала система оперирует парой интерактивно задаваемых пользователем параметров:

$L$  – длина интервала в метрах

$N$  – количество выходов частных амплитуд за верхнюю границу исторического интервала

Типичные значения в практических тестах: 3 превышения на 5 метрах.

Следующий момент – это возможность дополнительной интеграции определенных таким образом событий. Мы использовали для такой интеграции термин тревога. Идея состоит в том, что можно не выдавать сразу информацию о событии, дожидаться его развития (подтверждения) во времени. Для выдачи тревоги требуется наличие нескольких событий в течение заданного интервала времени. Оба параметра (временной интервал и количество событий) в интерактивном режиме задаются пользователем.

В итоге получился потоковый алгоритм. Не нужно накапливать большой объем данных перед принятием решения, на минимальном уровне агрегации событие можем быть определено за 2 такта (1 секунда). Алгоритм может быть легко масштабирован (горизонтальное масштабирование) и также легко распараллелен.

Практические испытания показали хорошие результаты детектирования событий в охранной системе при практически полном отсутствии ложных

срабатываний.

В подобном виде система вполне может применяться, например, в проектах Smart Cities [7].

## V. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ

На рисунке 5 проиллюстрирована программная реализация пользовательского интерфейса. Три ленты событий – по числу фильтров. В силу потокового характера алгоритма, все параметры (определение событий, их агрегация) могут быть заданы

интерактивно. Они становятся актуальными, начиная со следующего такта работы.

Обучение системы не требует вмешательства от пользователя. Система просто читает данные из потока до явного указания момента остановки. Легкость и простота процесса обучения позволяет запускать его по необходимости. Например, перед началом рабочей смены, при изменении погодных условий и т.п.

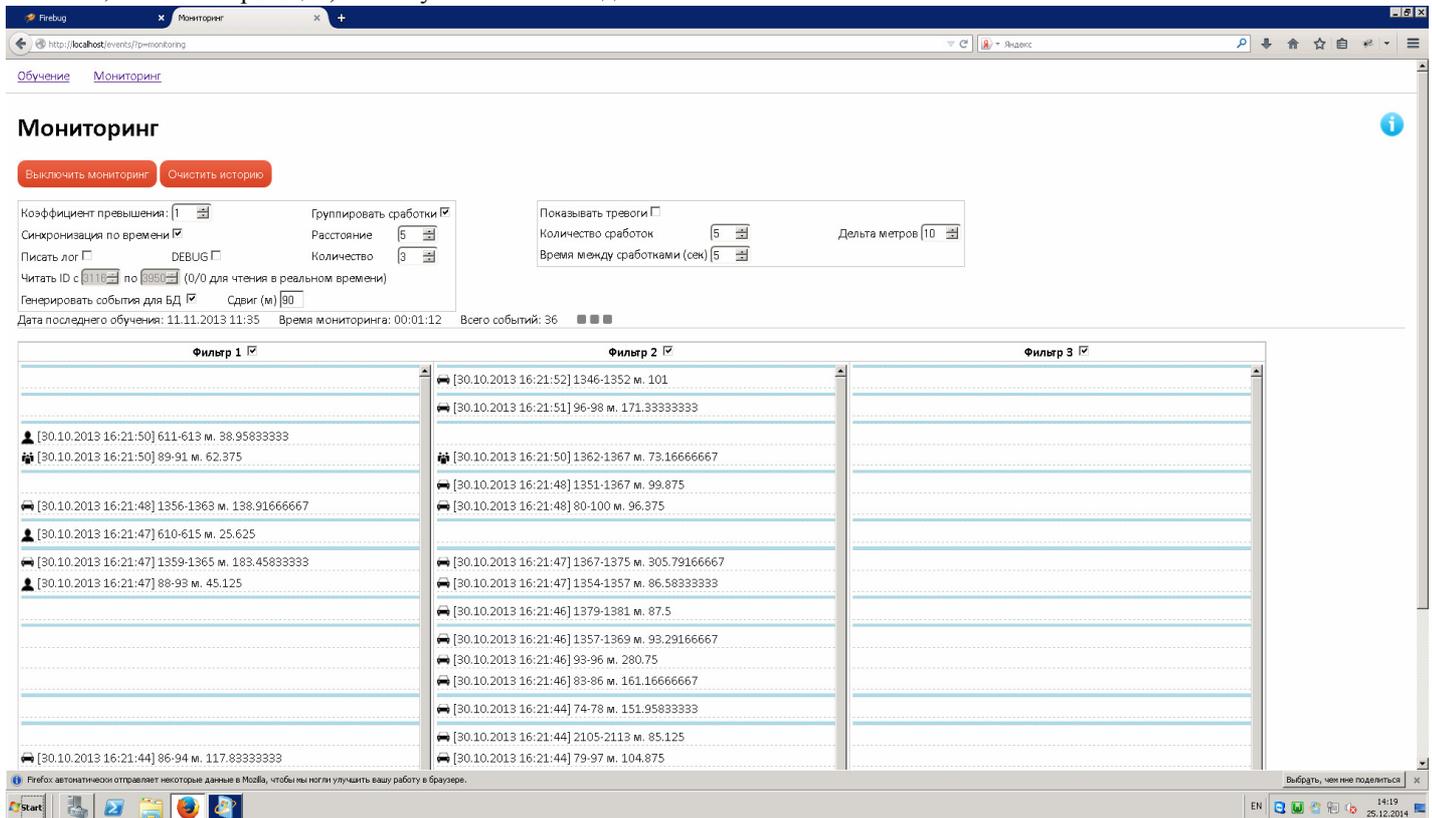


Рис.5 Пользовательский интерфейс

При выдаче события в ленту указывается метрическое расстояние (вдоль кабеля) и классификатор. Вопросы классификации событий – это тема отдельного исследования (статьи).

Открытый программный интерфейс системы НЕРВ позволяет подключать внешние приложения (использовать информацию о событиях во внешних приложениях). Например, в качестве такого внешнего приложения может выступать гео-портал, где отметки событий будут отображаться на карте.

## БЛАГОДАРНОСТИ

Авторы выражают признательность компаниям Т8 (Трещиков В.Н.) и Авиатех (Нагаев А.М.) за предоставленный доступ к оборудованию и испытательному стенду.

## БИБЛИОГРАФИЯ

[1] Нестеров Е.Т., Марченко К.В., Трещиков В.Н., Леонов А.В. Волоконно-оптическая система мониторинга протяженных

объектов (нефтепроводов) на основе когерентного рефлектометра // T-Comm . 2014. №1. С.25-28.

- [2] Распределенная охранная волоконно-оптическая система «ДУНАЙ» [http://t8.ru/?page\\_id=221](http://t8.ru/?page_id=221) Retrieved: Dec, 2014
- [3] Namiot D., Snepc-Sneppe M. On IoT Programming //International Journal of Open Information Technologies. – 2014. – Т. 2. – №. 10. – С. 25-28.
- [4] Wolfram, Stephen. "Mathematica." Wolfram Research, Champaign (2009).
- [5] Langmann, Benjamin, et al. "MOVEDTECT–Secure Detection, Localization and Classification in Wireless Sensor Networks." Internet of Things, Smart Spaces, and Next Generation Networking. Springer Berlin Heidelberg, 2013. 284-297.
- [6] Havskov, J.: Instrumentation in Earthquake Seismology (Modern Approaches in Geophysics). Springer (2004)
- [7] Волков А. А., Намиот Д. Е., Шнепс-Шнеппе М. А. О задачах создания эффективной инфраструктуры среды обитания //International Journal of Open Information Technologies. – 2013. – Т. 1. – №. 7. – С. 1-10.

# An Approach to Time Series Analysis

S.N. Kuklin, D.E. Namiot

***Abstract***— The article deals with issues related to the processing of measurement data distributed acoustic sensor influences. We propose the model, which allows high reliability to determine events related to external influences on the acoustic sensor. At the same time, the proposed model can effectively filter out the events associated with the natural acoustic background and effectively filter out false signals. The proposed model is implemented as a portable software, based on open source products. The accuracy of the model was validated in a number of practical tests.

***Key words***— time series, acoustics, fiber-optic system.