

## **Распределённая обработка данных в мониторинговых системах и сетях\***

**С.А. АВРОРОВ, М.С. ХАЙРЕТДИНОВ**

Дана общая постановка проблемы локации сейсмических источников и связанной с нею задачи обеспечения высокоточного определения параметров порождаемых источниками сейсмических волн. Рассмотрены вопросы идентификации полезных волн и волн-помех. Дана краткая характеристика программной системы, реализующей предложенные алгоритмы в реальном масштабе времени.

**Ключевые слова:** сейсмолокация источников, сейсмические и акустические волны, волновые формы, вычислительные технологии, идентификация источников, ГИС-системы

### **ВВЕДЕНИЕ**

В мире создан ряд глобальных информационных сетей, предназначенных для мониторинга природных и техногенных процессов. В частности, мировая сеть сейсмического мониторинга IMS (International Monitoring System) предназначена для регистрации землетрясений, цунами, ядерных и промышленных взрывов, происходящих в различных районах Земли [1]. Можно выделить некоторую общую архитектуру построения сетей, включающую три уровня иерархии: нижний уровень – это датчики физических процессов совместно с промежуточными центрами сбора данных (автономные станции), средний – локальные центры сбора и обработки данных на уровне региона, верхний уровень – глобальный центр сбора и обработки данных на уровне страны, мира. Одна из основных задач мониторинга сейсмических событий связана с восстановлением их параметров в очаге, прежде всего географических координат по данным регистрации сейсмических сигналов с помощью сети сейсмических станций либо малой сейсмической группы.

Решения задач мониторинга распределяются между различными уровнями сети: на нижнем уровне в реальном масштабе времени осуществляется первичная обработка регистрируемых колебаний разной физической природы – сейсмических, акустических, гидроакустических и т. д. Обработка проводится с целью определения основных характеристик упругих волн, основными из которых являются времена вступления различных типов волн и их амплитуд. На верхних уровнях сети осуществляется уточненное решение задачи по определению параметров источников.

Влияние сейсмических шумов вызвало потребность в автоматизированной технологии решения задач обнаружения, распознавания сейсмических сигналов и локации сейсмических источников. Такая технология избавляет от множества рутинных операций обработки данных в условиях непрерывного мониторинга событий. Целью данной работы является создание соответствующих алгоритмов и программ для реализации обозначенных целей.

---

\* *Статья получена 17 февраля 2010 г.*

### 1. ПОСТАНОВКА ЗАДАЧИ

Задача оценивания неизвестных параметров источника сводится к решению нелинейной системы условных уравнений [2]:

$$\vec{t} = \vec{\eta}(X, \theta) + \vec{\varepsilon}, \quad (1)$$

где  $\vec{t} = (t_1, t_2, \dots, t_N)$  – вектор времен пробега сейсмических сигналов;  $\vec{\eta}(X, \theta)$  –  $N$ -мерный вектор вычисляемых времен пробега (теоретический годограф) или функция регрессии;  $\vec{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_N)^T$  – вектор невязок;  $\vec{\theta} = (x, y, z, v, t)^T$  –  $m$ -мерный вектор оцениваемых параметров,  $X = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)$  – матрица координат датчиков (или точек излучения);  $N$  – число датчиков (или точек излучения). В качестве оцениваемых параметров выступают пространственные координаты источника  $x, y, z$ , скоростная характеристика среды  $v$  и время в источнике  $t$ . В ряде случаев скорость в среде известна. При оценивании параметров пользуются сведениями о распределении ошибок  $\varepsilon_i = t_i(\vec{x}_i, \vec{\theta}) - \eta(\vec{x}_i, \vec{\theta})$ . Будем в дальнейшем предполагать, что  $\varepsilon_i$  – взаимно независимые случайные величины, имеющие распределение с нулевым средним и заданными дисперсиями:  $E\varepsilon_i = 0$ ,  $E\varepsilon_i\varepsilon_j = \sigma_i^2\delta_{ij}$ ,  $\sigma_i = \sigma(\vec{x}_i)$ , где  $\delta_{ij}$  – символ Кронекера,  $i = 1, 2, \dots, N$ . В случаях затруднений с заданием дисперсий их принимают равными и получают несмещенную оценку дисперсии наблюдения с единичным весом по ходу решения задачи.

Решение уравнения (1) сводится к решению обратной задачи. При этом точность решения зависит, в первую очередь, от ошибок оценивания вектора времен  $\vec{t} = (t_1, t_2, \dots, t_N)$ , скорости  $v$ , шумов измерения  $\vec{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_N)^T$ , выбора геометрии расстановки датчиков на дневной поверхности Земли. В частности, по отношению к полярной системе координат дисперсия ошибки определения азимута  $Az$  на источник и расстояния  $R$  «источник–приемник» с помощью триады сейсмостанций (случай  $N = 3$ ) определяется соотношениями [3]:

$$\sigma_{AZ}^2 = \sigma_i^2 F_1(\vec{t}, \vec{\psi}), \quad \sigma_R^2 = \sigma_i^2 F_2(\vec{t}, \vec{\psi}), \quad (2)$$

где  $\sigma_i^2$  – ошибки оценивания времён  $\vec{t}$ ,  $\vec{\psi}$  – вектор параметров, характеризующий геометрию расстановки сейсмической группы.

Решение поставленной задачи ведется по этапам:

- 1) обнаружение и измерение времен вступлений волн на фоне внешних шумов;
- 2) идентификация источника;
- 3) расчет параметров источника;
- 4) отображение координат источника на цифровой карте местности.

В данной работе решение задач на перечисленных этапах рассматривается на примерах обработки сигналов от маломощных полигонных взрывов, зарегистрированных на небольших расстояниях от источника (до 2 км).

По отношению к сейсмическим волнам, описываемым импульсными сигналами, ошибки оценивания времен могут быть вычислены в виде [4]

$$\sigma_\tau^2 = \frac{\tau_n}{2\Delta f(2E/N_0)}. \quad (3)$$

Здесь  $\tau_n$  – длительность волнового импульса;  $\Delta f$  – ширина его спектра;  $E/N_0$  – отношение энергии импульса к спектральной плотности внешнего шума. Как следует из (3), уменьшение погрешности можно достичь за счет увеличения соотношения энергий сигнала и шума, расширения спектра частот  $\Delta f$ , занимаемого импульсом, а также за счет сжатия волнового импульса во

времени путем приведения его к  $\delta$ -образному импульсу. Ниже рассматриваются алгоритмы решения задачи минимизации погрешностей оценивания времен с учетом (3).

## 2. ПОВЫШЕНИЕ СООТНОШЕНИЯ СИГНАЛ/ШУМ

Алгоритм повышения соотношения сигнал/шум по отношению к волновым импульсам наиболее эффективно реализуется с помощью вейвлет-фильтрации, основанной на разложении одномерного сигнала по базису, сконструированному из обладающей определенными свойствами солитоноподобной функции (вейвлета) посредством масштабных изменений и переносов. В основе вейвлет-фильтрации лежит интегральное вейвлет-преобразование

$$X(\tau, s) = \int_{-\infty}^{\infty} f(t) \Psi_{\tau, s}(t) dt, \quad (4)$$

где  $f(t)$  – волновой импульс;  $\Psi_{\tau, s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right)$ ,  $\tau, s \in R$ ,  $\psi \in L^2(R)$ .

Базис функционального пространства  $L^2(R)$  может быть построен с помощью непрерывных масштабных преобразований и переносов вейвлета  $\psi(t)$  с произвольными значениями базисных параметров – масштабного коэффициента  $s$  и параметра сдвига  $\tau$ . Наиболее подходящими для аппроксимации сейсмических колебаний являются вейвлеты Добеши [5]. В качестве примера на рис. 1 приведены вейвлеты семейства Добеши 2, 4, 5, 8 и 10-го порядков.

В связи с этим возникает задача подбора порядка материнского вейвлета для обеспечения наилучшего приближения к конкретным видам волновых импульсов.

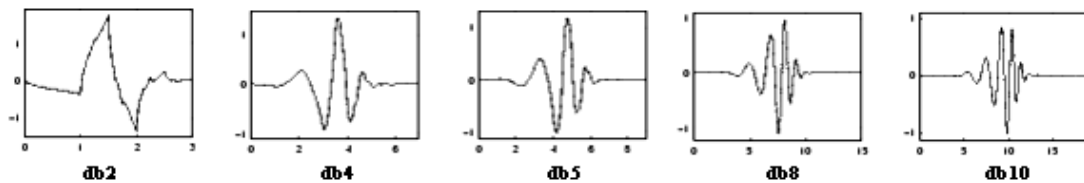


Рис. 1. Семейство вейвлетов Добеши

В качестве иллюстрации рассмотрим волновые импульсы, порождаемые полигонными взрывами (рис. 2, а). По результатам анализа вейвлетов Добеши к использованию был рекомендован вейвлет 4-го порядка. При этом обеспечивается наилучшая аппроксимация формы рассматриваемого волнового импульса (рис. 2, б). Для иллюстрации качества аппроксимации волнового импульса набором вейвлетов на рис. 3 приведены графики зависимости ошибки аппроксимации от глубины вейвлет-декомпозиции. Качество аппроксимации оценивалось по критерию среднего квадрата ошибки. На рис. 3, а это выполнено по отношению к модельному сигналу-импульсу Берлаге, общепринятому для описания взрывных источников [6], на рис. 3, б – по отношению к рассматриваемой реальной записи (рис. 3, б). Действительно, использование вейвлетов порядка выше 3-го в первом случае и выше 4-го не дает эффекта для повышения качества аппроксимации.

Пример фильтрации исходной записи полигонного взрыва, рассмотренной ранее на рис. 2, а, представлен на рис. 2, б. В результате обработки удалось достичь соотношения сигнал/шум, равного 140.

Для вычисления времен вступления волн в автоматическом режиме предварительно рассчитывают огибающую волнового импульса  $A(t)$  через преобразование Гильберта

$$A(t) = \sqrt{f^2(t) + f^{*2}(t)},$$

где  $f^*(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{f(x)}{t-x} dx$  – преобразование Гильберта.

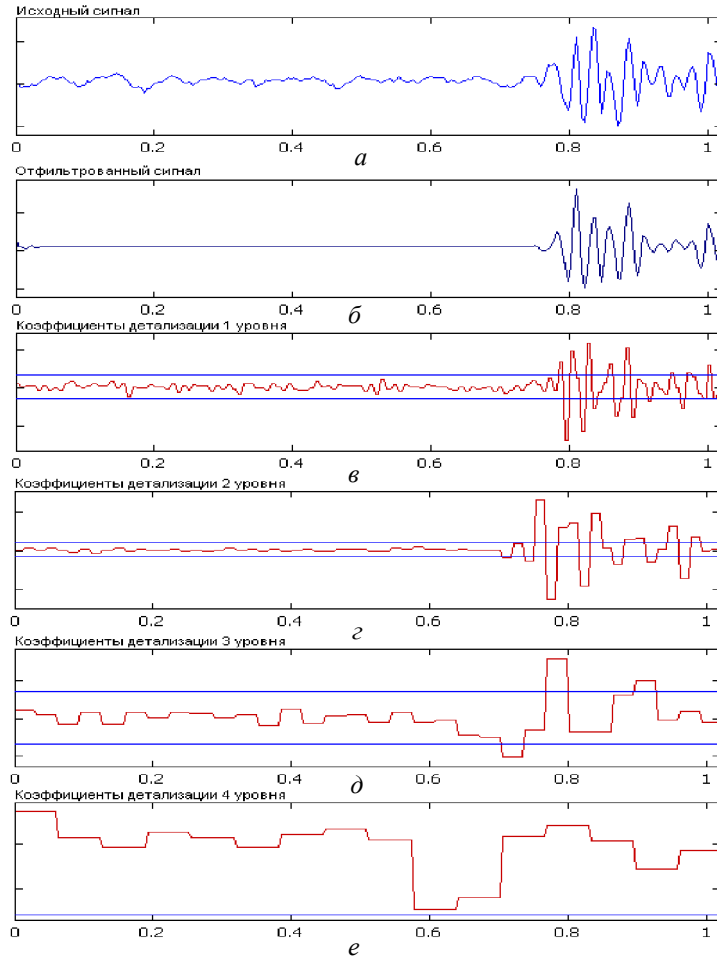


Рис. 2. Пример фильтрации реальной записи полигонного взрыва:  
*a* – исходный сейсмический сигнал, *б* – результата фильтрации,  
*в–е* – вейвлет-коэффициенты на уровнях разложения 1–4 с соответствующими значениями порогов

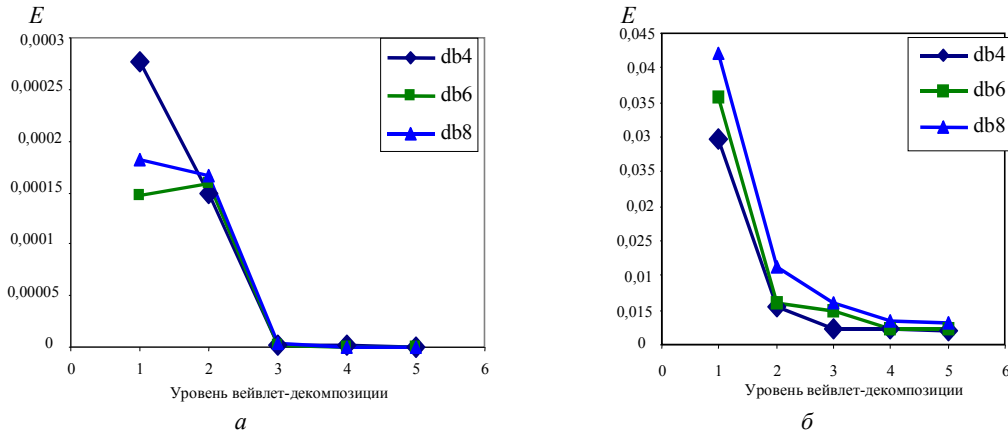


Рис. 3. График зависимости ошибки аппроксимации сейсмического сигнала от глубины вейвлет-декомпозиции при использовании вейвлетов Добеши 4, 6, 8-го порядков, вычисленных для модельного (а) и реального (б) сейсмических сигналов

Для поиска и вычисления времён прихода волновых импульсов использована пороговая обработка по правилу «трёх  $\sigma$ ». В стационарных условиях предшествующий полезному сигналу сейсмический шум описывается нормальным распределением с дисперсией, вычисляемой в поточном режиме с помощью итерационной процедуры вида

$$D[N + 1] = D[N] + \gamma(X[N + 1] - \bar{X}[N])^2, \tag{5}$$

где  $\bar{X}[N] = \bar{X}[N - 1] + \gamma(X[N] - \bar{X}[N - 1])$  – среднее значение, вычисленное на шаге  $N$ ;  $\gamma = \Delta t / T$  – коэффициент сходимости.

Значение порога обнаружения волны выбирается равным  $\Pi = 3\sqrt{D}$ .

В качестве примера успешного использования рассмотренной вычислительной технологии на рис. 4 приведены результаты обнаружения волновых импульсов сейсмической и акустической природы от полигонного взрыва мощностью около 500 г в тротиловом эквиваленте, удалённого на расстояние около 800 м. На записи первая низкоамплитудная волна – сейсмическая, вторая высокоамплитудная – акустическая.

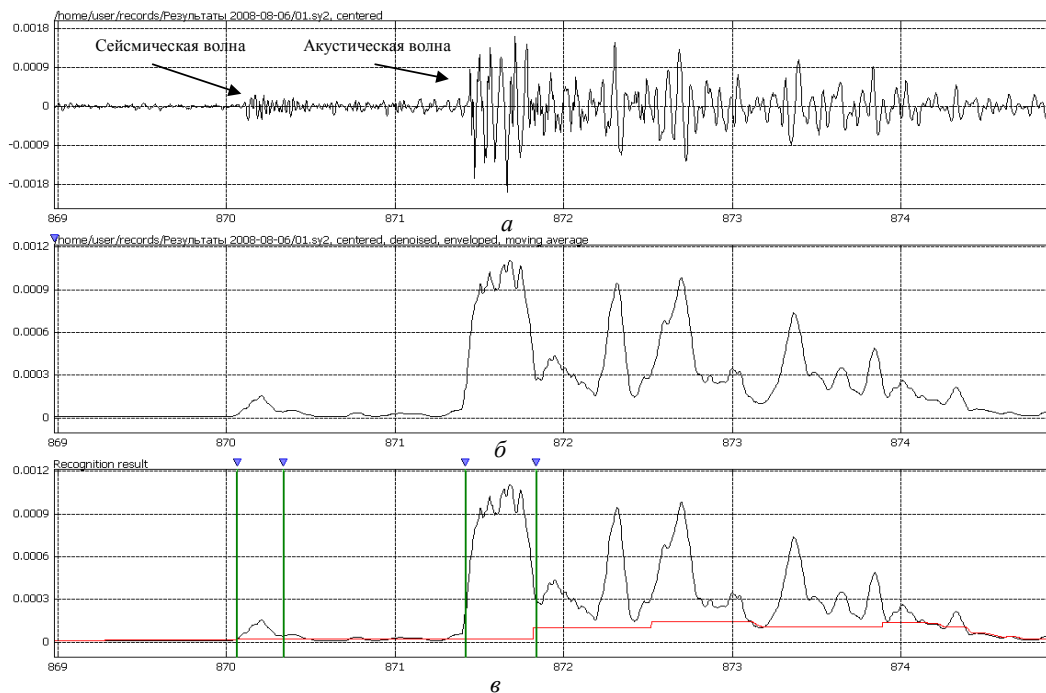


Рис. 4. Пример фильтрации реальной записи полигонного взрыва:

*a* – исходные сейсмическая и акустическая волны; *б* – результата вейвлет-фильтрации и построения огибающей; *в* – результат автоматического порогового вычисления времён вступлений сейсмических и акустических волн с последующим определением длительности импульсов

Погрешность вычисления времён вступления волн оценивалась по результатам внимательной ручной работы человека-оператора.

### 3. ПРОЦЕДУРА РАСПОЗНАВАНИЯ ТИПА СОБЫТИЯ

В реальных условиях стационарный сейсмический шум осложнен присутствием ложных импульсов-помех, имеющих техногенную природу (транспортные шумы, движение пешеходов и т.д.). Во избежание ошибочной интерпретации шумовых всплесков как полезных сигналов дальнейшим этапом обработки становится идентификация источника сейсмических сигналов, результатом которой будет отнесение всплесков к одному из двух классов: взрыв или помеха. Построение системы идентификации сейсмических сигналов подразумевает выбор системы информативных признаков, на основе которой будет производиться классификация. В результате изучения особенностей сейсмических записей взрывов была выбрана система из четырёх признаков:

- соотношение амплитуд сейсмических и акустических волн. Исходные значения амплитуд волн представлены на рис. 5, *a*;
- соотношение средних частот этих волн  $F_s/F_a$ . Сами значения средних частот представлены на рис. 5, *б*;
- соотношения длительностей волн (рис. 5, *в*);
- задержка между временами вступления сейсмических и акустических волн  $T_a-T_s$  (рис. 5, *г*).

На рис. 5, *a* и *б* каждому маркёру графика соответствует отдельная зарегистрированная волна: квадратным маркёрам – сейсмическая, ромбовидным – акустическая. Время обнаружения соответствующей волны отображено на оси абсцисс.

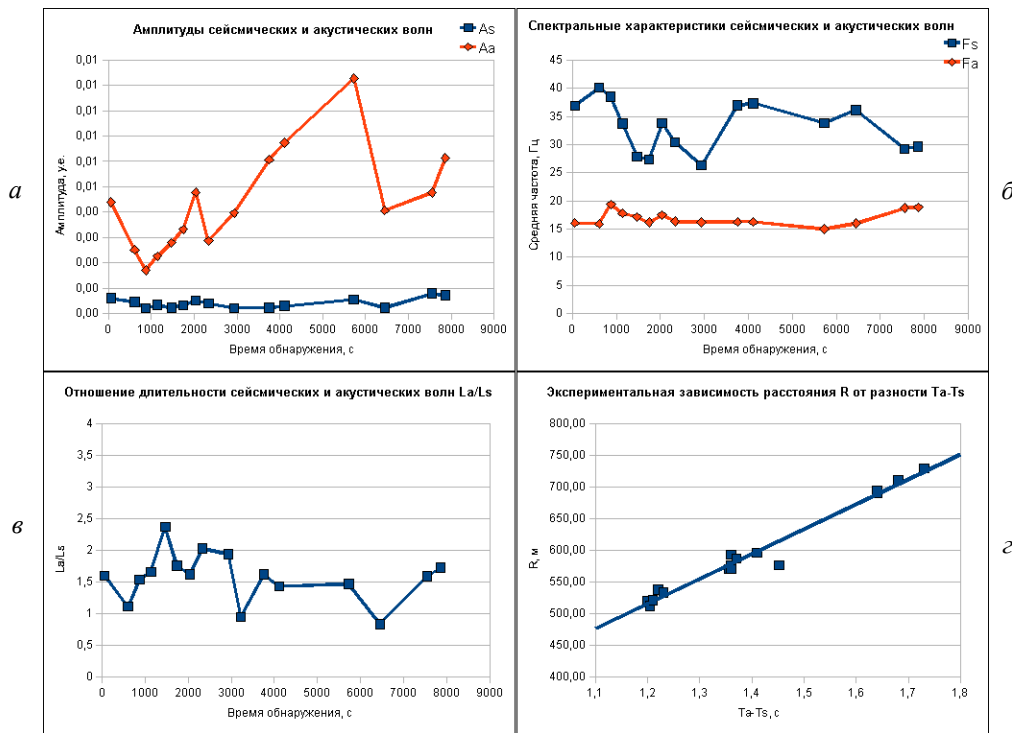


Рис. 5. Экспериментальные данные, отражающие:

*a* – соотношение амплитуд сейсмических и акустических волн  $Aa/As$ ; *б* – доминирующие частоты сейсмических  $Fs$  и акустических волн  $Fa$ ; *в* – длительности волн; *г* – задержки между временами вступления сейсмических и акустических волн  $Ta-Ts$

На рис. 5, *a* видно, что амплитуда сейсмических волн менее вариабельна по сравнению с акустическими в диапазоне мощностей взрывов (мощность взрывов в процессе эксперимента варьировалась от 200 до 2600 г в тротиловом эквиваленте). На рис. 5, *б* отражены значения средних спектров для обоих типов волн. Видно, что частота акустических сигналов лежит в диапазоне 15...18 Гц, сейсмических волн – в диапазоне 25...40 Гц. На рис. 5, *г* показана экспериментальная зависимость расстояния «источник–приёмник»  $R$  от разности времён прихода сейсмических и акустических волн. Разность времён вступления сейсмических и акустических волн (так же как разность времен вступлений продольных и поперечных волн, распространяющихся в упругой среде) есть функция расстояния от источника волн до точки регистрации. Располагая данными о скоростях распространения волн для конкретного региона, задав необходимый временной интервал разностей времён прихода  $\Delta t$ , можно обнаружить событие нужного диапазона расстояний. Это свойство используется как дополнительный признак в процессе обнаружения и распознавания сейсмических сигналов и учитывается на этапе определения времён вступления сейсмических и акустических волн.

На основе статистического анализа набора данных по полевым взрывам мощностью от 200 до 2600 г в тротиловом эквиваленте, зарегистрированных на расстояниях от 472 до 750 м, была построена программная система, реализующая возможности поточного обнаружения и определения параметров волн, а также распознавания типа события (полевой взрыв или помеха). В качестве решающей функции использовано логическое решающее правило

$$T = A(Aa/As) \ \& \ F(Fs/Fa) \ \& \ T(Ta - Ts) \ \& \ L(La/Ls),$$

где операция  $\&$  («логическое И») означает одновременность выполнения заданных условий:

Если  $T = 1$ , то полезный сигнал  $T = 0$  – реакция на шум.

#### 4. ЭКСПЕРИМЕНТАЛЬНОЕ ОЦЕНИВАНИЕ НЕИЗВЕСТНЫХ ПАРАМЕТРОВ ИСТОЧНИКА

С использованием системы алгоритмов и программ обнаружения и распознавания сейсмических и акустических сигналов был обработан набор записей 19 полевых взрывов мощностью от 200 до 2600 г в тротиловом эквиваленте. Диапазон расстояний «источник-приемник» лежал в пределах от 570 до 750 м. Обработка данных производилась в автоматическом режиме. Из 19 взрывов в автоматическом режиме было обнаружено 16, при этом к пропущенным событиям относятся полевые взрывы, проведенные в особых условиях (в болоте и воздухе).

В обоих случаях вступления сейсмической волны обнаружить не удалось, в то время как акустическая волна проявлялась очень четко. Результаты определения координат источников в полярной системе представлены в таблице.

#### Результаты вычислений расстояний «источник-приёмник» $R$ , азимутальных направлений и сопоставления с истинными значениями

№ события	Расстояние $R$ , м			Азимут, град		Погрешность, м	
	Автоматическое	Ручное	Истинное	Автоматическое	Истинное	Автоматическое	Ручное
1	508,94	499,48	512,37	256,92	253,49	30,76	32,91
2	613,77	569,65	576,28	264,35	260,16	57,42	42,41
3	573,19	559,34	570,52	262,79	261,60	12,17	16,21
4	574,89	563,39	570,70	262,79	263,09	5,15	7,89
5	507,25	509,71	519,66	269,00	269,30	12,70	10,31
6	511,48	509,20	521,51	270,18	270,48	10,39	12,60
7	519,93	520,02	533,53	282,59	281,20	18,66	18,60
8	515,71	526,79	538,05	283,57	283,25	22,54	11,65
9	579,11	575,23	586,68	295,16	295,16	7,57	11,45
10	574,89	581,15	593,02	296,80	296,27	18,92	13,05
11	710,15	709,82	710,77	300,77	300,02	9,32	9,35
12	731,29	731,04	729,54	300,77	300,93	2,69	2,53
13	693,24	684,20	690,42	299,98	300,26	4,41	7,07
14	693,24	690,71	694,36	298,40	298,08	4,03	5,32
15	574,89	574,00	576,54	256,60	256,63	1,68	2,56
16	596,02	591,12	596,18	257,25	257,25	0,16	5,06

Приведены результаты автоматического и ручного вычислений расстояний «источник-приёмник»  $R$ , азимутальных направлений на источник и погрешности определения расстояния  $R$  при ручной и автоматической обработках с учётом угла азимута. Погрешность вычисляется как расстояние между координатами вычисленного эпицентра и полученного на основе высокоточных данных GPS. Средние значения погрешностей для ручной и автоматической обработок составляют 13,66 и 13,06 м соответственно, чему соответствует точность около 5 %.

#### 6. ОСОБЕННОСТИ РЕАЛИЗАЦИИ ПРОГРАММНОЙ СИСТЕМЫ

На основе разработанных методов решения задачи локации сейсмоакустических источников была разработана программная система, предназначенная для функционирования в режиме реального времени и включающая следующее:



1) модуль вейвлет-фильтрации поступающих на вход сигналов в процессе непрерывного мониторинга. Реализован на основе алгоритма быстрого вейвлет-преобразования библиотеки GNU GSL;

2) модуль обнаружения импульсных сигналов и вычисления времён прихода волн пороговым методом;

3) модуль распознавания типа источника на фоне шумов;

4) модуль вычисления координат источника методом расчета коэффициента направленного действия сейсмической антенны;

5) модуль сопряжения с ГИС-системой. Для отображения координат источника на цифровой карте местности был использован унифицированный язык на основе XML для представления трёхмерных геопространственных данных под названием KML (Keyhole Markup Language – язык разметки Keyhole) [8]. На выходе программы создаётся KML-файл, используемый в целом ряде профессиональных ГИС-систем, таких как ESRI ArcGIS Explorer, GRASS, MAPinfo, OpenStreetMap, Карты Google, Adobe PhotoShop, AutoCAD, Yahoo! и др. Таким образом, благодаря универсальности открытого формата представления пространственных данных дальнейшие (более высокие) уровни обработки и анализа сейсмических данных могут производиться привычным способом. Пример отображения источников и приемников с использованием ГИС Google Earth приведен на рис. 6.

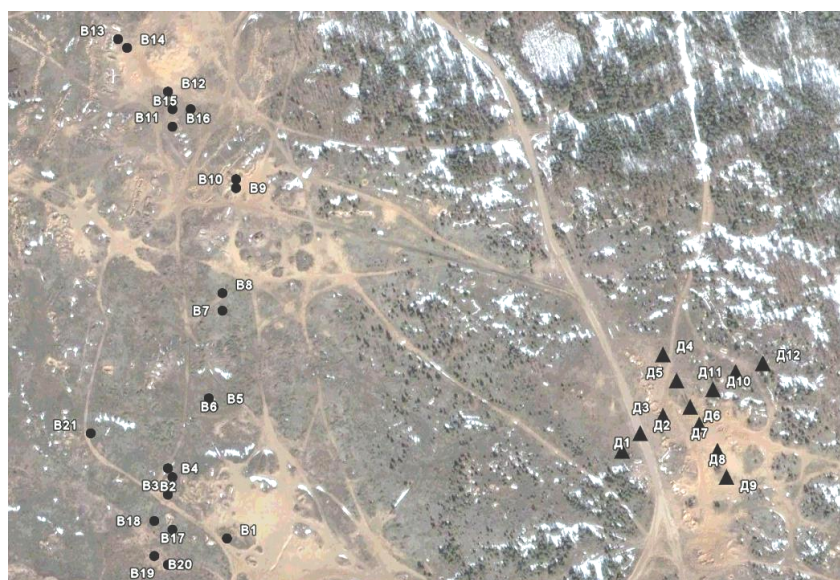


Рис. 6. Пример отображения сейсмических источников и регистрирующей группы (расстановка типа «Крест») с использованием ГИС Google Earth

Система алгоритмов реализована на языке Си. По результатам замеров время полной обработки отрезка записи длительностью 2100 с (при частоте дискретизации 250 Гц) составило 8 с. Таким образом, уложившись в заданные временные рамки, можно создать значительный запас ресурсов на дополнительное математическое обеспечение.

## ЗАКЛЮЧЕНИЕ

В работе получены следующие результаты.

1. Разработаны алгоритмы и программная система, предназначенные для использования в иерархических сетях мониторинга природных и техногенных сейсмических событий. Основные задачи, решаемые с помощью разработанных средств, связаны с обнаружением и измере-

нием времён вступлений сейсмических волн на фоне внешних шумов, идентификацией типов источников и расчетом их географических координат, отображением координат источника на цифровой карте местности. Высокая эффективность созданных средств доказана применительно к импульсным источникам в виде полигонных взрывов по интегральному критерию. Последний оценивает значение невязки между истинными и расчетными координатами серии взрывов. Полученное среднее значение невязки по 150 взрывам составило около 1,5 %.

2. Высокая производительность обработки записей доказана результатами тестирования производительности: искомые записи длительностью 1 ч обрабатываются за 2 с, что соответствует ускорению в 1500 раз. Это говорит о возможности применения разработанной системы для обработки данных в режиме реального времени в полевых условиях.

3. Разработан интерфейс для сопряжения программной системы с современными ГИС-системами с целью отображения обнаруженных сейсмоисточников на цифровой карте местности.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Зербо Л. Возможности подготовительной Комиссии ОДВЗЯИ отслеживать соответствие Договору о всеобъемлющем запрещении ядерных испытаний: докл. на 5-й Междунар. конф. «Мониторинг ядерных испытаний и их последствий», 4–8 авг. 2008 / Боровое (Казахстан).
- [2] Khairetdinov M.S., Omelchenko O.K. et al. A computing technology to determine the parameters of the borehole bottom and medium // Bull. Nov. Comp. Center. Math. Model. in Geoph. – 2005. – Iss.10. – P. 19–26.
- [3] Kijko A., Ungvarai J. The accuracy of teleseismic events location by a tripartite seismic network in Hungary // Acta Geodaet. – Tomus (Hungary), 2002. – P. 239–246.
- [4] Тихонов В.И. Оптимальный прием сигналов. – М.: Радио и связь, 1983.
- [5] Oonincx P.J. Automatic phase detection in seismic data using the discrete wavelet transform // CWI Report PNA-R9811, October 1998.
- [6] Гурвич И.И., Боганик Г.Н. Сейсмическая разведка: учебник для вузов. – 3-е изд., перераб. – М.: Недра, 1980.
- [7] Авроров С.А., Хайретдинов М.С. Автоматизированная процедура поточного обнаружения и идентификации сейсмических событий // Науч.-техн. журн. Национального ядерного центра республики Казахстан. – 2008. – Вып. 2. – С. 70–75.
- [8] Введение в документацию KML. Электронное справочное руководство. <http://code.google.com/intl/ru/apis/kml/documentation/>

*Авроров Сергей Александрович*, аспирант кафедры сетевых информационных технологий Новосибирского государственного технического университета. Основное направление научных исследований – вычислительные технологии в сейсмологии. Имеет 11 публикаций.

E-mail: sergey.avrorov@gmail.com

*Хайретдинов Марат Саматович*, доктор технических наук, профессор, зав. кафедрой сетевых информационных технологий Новосибирского государственного технического университета. Основное направление научных исследований – методы цифровой обработки данных, интеллектуальный анализ данных и распознавание образов, математическое моделирование в геофизике, геоинформационные системы. Имеет 180 публикаций.

E-mail: marat@opg.ss.cc.ru

**S.A. Avrorov, M.S. Khairetdinov**

*Distributed data processing in monitoring systems and monitoring networks*

In the work the problem of creating a technology for automatic stream-handling detection and recognition of seismic events which are generated by natural and man-caused processes is considered. The overall definition of seismic sources location problem is formulated as well as related problem of calculation of high-precision seismic waves parameters under background noise. The questions of desired waves and noisy signals identification are considered. A brief overview of the program system including all the algorithms developed and performing in a real-time mode is presented. The efficiency of this approach is illustrated with respect to polygon explosions location problem.

**Key words:** seismic sources location, seismic and acoustic waves, wave shapes, processing technologies, sources identification, GIS-systems.