

УДК 550.34+005

АЛГОРИТМ АВТОМАТИЧЕСКОГО ОБНАРУЖЕНИЯ И КЛАССИФИКАЦИИ ПРОМЫШЛЕННЫХ ВЗРЫВОВ НА ОСНОВЕ ЭНТРОПИЙНОГО ОТОБРАЖЕНИЯ СЕЙСМИЧЕСКИХ СИГНАЛОВ

© 2019 г. Р.Ю. Замараев, С.Е. Попов

Институт вычислительных технологий СО РАН, г. Новосибирск, Россия

Представлен алгоритм классификации сейсмических событий на основе предварительных преобразований сигналов в энтропийную модель. Такая модель позволяет обобщать информацию об особенностях волновых форм по трем направлениям измерений, усреднять и сравнивать эти особенности. На основе энтропийной модели рассчитываются характеристические функции событий, используемые затем для оценки подобия сейсмического поля в расчетном окне шаблонам взрывов, землетрясений и абстрактным шаблонам с заданными свойствами. Шаблоны для взрывов и землетрясений получаются путем усреднения характеристических функций множеств соответствующих событий для данной станции на данный регион. Абстрактные шаблоны рассчитываются с помощью функции огибающей энтропийной модели сейсмического поля – суммы энтропийных моделей сигналов по трем направлениям измерений. Подобие выборочной характеристической функции заготовленным шаблонам оценивается с помощью набора мер расстояния между одномерными векторами типа корреляционной дистанции и др. Заключение выдается на основе системы рейтингового голосования по числу минимальных дистанций. Приведен пример классификации сейсмических событий на территории Кемеровской области. Алгоритм ориентирован на региональные сети и системы сейсмического мониторинга и обеспечивает полностью автоматизированный и быстрый процесс обнаружения и классификации заявленных региональных сейсмических событий.

Ключевые слова: региональные сейсмические события, промышленные взрывы, обнаружение событий, быстрая автоматическая классификация, энтропийные модели, шаблоны событий, оценки подобия.

Введение

Актуальной задачей региональных систем сейсмического мониторинга является детектирование и классификация событий природного и искусственного происхождения, которые накладываются по магнитуде и локализации.

Полностью автоматическая классификация сейсмических событий в реальном времени является развитием этой задачи. Она интересна не только научной постановкой, но и обоснована аспектами информационного обеспечения гражданской обороны, спасательных служб и страховых компаний.

Кемеровская область, в которой находится Кузнецкий угольный бассейн (рис. 1), входит в число регионов, где заявленная задача имеет очевидную практическую значимость. На территории области функционирует более 30 предприятий, регулярно проводящих массивные взрывные работы. В непосредственной близости к их промышленным зонам проживает около 2 млн. человек и размещены муниципальные коммуникации. В общей сложности за год здесь регистрируется более 2 тыс. сейсмических событий с магнитудой $1.2 \leq M_L \leq 2.7$, характерной как для региональных землетрясений, так и для типичных (по технологии и мощности) взрывов на угольных разрезах.

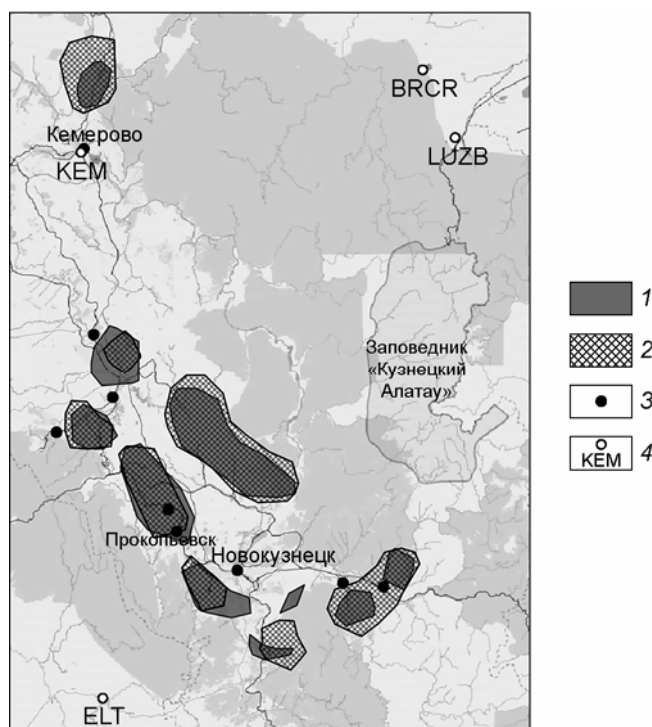


Рис. 1. Зоны концентрации сейсмических событий по протоколам Службы сейсмических наблюдений Агентства по защите населения и чрезвычайным ситуациям Администрации Кемеровской области. 1 – класс “сейсмическое событие”; 2 – класс “промышленный взрыв”; 3 – крупные населенные пункты; 4 – станции сети ASGSR¹ с указанием их международных кодов: КЕМ (г. Кемерово), BRCR (пос. Берчихуль), LUZB (пос. Лужба), ELT (пос. Ельцовка)

Fig. 1. Zones of concentration of seismic events according to the protocols of Service of seismic observations (Agency for the Protection of Population and Emergencies, Administration of the Kemerovskaya Oblast, Russian Federation). 1 – “Seismic event” category; 2 – “Industrial explosion” category, 3 – large settlements, 4 – ASGSR network stations with their international codes: KEM (Kemerovo city), BRCR (Berchikul village), LUZB (Luzhba village), ELT (Yeltsovka village)

Известно множество работоспособных подходов к классификации сейсмических событий, которые опираются на разные комбинации процедур фильтрации, спектрального и вейвлет-анализа сигналов, интегрируемых с байесовским подходом, аппаратом искусственных нейронных сетей и кластерным анализом. Обзор публикаций, представленный в [Замараев, Попов, Логов, 2016], охватывает работы до 2011 г. включительно и позволяет выделить два принципа развития технологии классификации сейсмических событий – разработка оригинальных признаков пространств для популярных алгоритмов классификации и усиление существующих алгоритмов классификации для известных признаков пространств. Среди более поздних публикаций отметим ряд интересных работ, которые также отвечают названным принципам. Так, в статье [Lyubushin et al., 2013] используется линейный Байесовский дискриминатор для нового признака – мультифрактальной оценки параметров мгновенного спектра сейсмического сигнала – и утверждается высочайшая (до 93 %) точность классификации для землетрясений.

В статье [Kortström, Uski, Tiira, 2016] используется сравнительно новый для сейсмологии метод опорных векторов (SVM) для совокупности 80 признаков –

¹ <http://zor.gsras.ru/www3/stations/net-info.php?lang=ru&info=ASGSR>

оценок энергии в узких частных полосах ($\Delta=2$ Гц) в диапазоне 1–41 Гц спектра сейсмического сигнала; при этом утверждается высокая ($>70\%$) точность классификации в зависимости от тестовой выборки (станции).

На практике на качество реализации любых теоретических подходов негативное влияние оказывает ряд организационных проблем: использование случайно сложившихся, нерегулярных и разреженных сетей сейсмических наблюдений; отсутствие жесткой обратной связи с предприятиями, производящими взрывы, и, как следствие, неполный каталог взрывов; ограниченный доступ к базам данных сейсмических наблюдений для независимых исследователей. Таким образом, подходы с малыми требованиями к объему и качеству исходной информации могут оказаться не лишними.

В подходе, развиваемом авторами работы [Замараев, Попов, Логов, 2016], используется непосредственный признак генезиса сейсмического события – форма сигнала, но приведенная к виду энтропийной модели, которая удобна для манипуляций и сравнения. Этим обеспечивается быстрая разработка шаблонов сейсмических событий на имеющейся сети с минимальными затратами на адаптацию при переходе от одной станции к другой. Использование модифицированной формы сигнала оставляет признаковое пространство в удобно интерпретируемой временной области и позволяет применять в качестве мер близости выборочного события к идентифицированным шаблонам оценки типа дистанции между одномерными векторами.

Совокупность представленных авторами идей, по их мнению, должна обеспечить полностью автоматизированный процесс классификации сейсмических событий с низкой стоимостью развертывания и потокового выполнения.

Характеристическая функция сейсмического поля

Рассмотрим дискретный сигнал $\{x_i\}$, $i=1, \dots, N$, где N – расчетное окно или число отсчетов. Исходные данные сигнала заменяются квадратами размахов, что обеспечивает неотрицательность значений для дальнейших вычислений:

$$d_i = (x_i - x_{i-1})^2, \quad i=1, \dots, N. \quad (1)$$

Сначала вычисляется вектор весов, затем – вектор энтропий $\mathbf{E}$:

$$q_i = \frac{d_i}{\sum_i d_i}, \quad (2)$$

$$E_i = -q_i \ln q_i. \quad (3)$$

Энтропийная модель дискретного сигнала (3) обладает всеми формальными свойствами шенноновской энтропии выборки [MacKay, 2003].

Если принять q_i как вероятность выбора отсчета i по значению d_i , то модель (3) обеспечивает аддитивность элементов, относящихся к одному сигналу и, главное, аддитивность элементов из различных сигналов в синхронных отсчетах.

По изложенному выше алгоритму для трех ортогональных компонент сейсмического поля – сейсмических сигналов x_E, x_N, x_Z со стандартными обозначениями каналов (направлений) измерения E, N, Z – вычисляются энтропийные модели $\mathbf{E}_E, \mathbf{E}_N, \mathbf{E}_Z$, после чего по этим трем каналам (направлениям) находится вектор обобщенной информации

$$\mathbf{H} = \mathbf{E}_E + \mathbf{E}_N + \mathbf{E}_Z \quad (4)$$

и вычисляется характеристическая функция сейсмического поля C в виде

$$C_i = \sum_{j=1}^i H_j. \quad (5)$$

Характеристическая функция C представляет собой кумулятивную последовательность количества информации сейсмических колебаний в точке размещения сейсмостанции, оцененную по трем направлениям и N отсчетам. При её определении три компоненты сейсмического сигнала с высокой долей стохастики сводятся к одномерной стационарной S -образной форме. Из свойств энтропии заведомо известен предел насыщения характеристической функции – $\max(C) = 3 \ln N$, где N – число отсчетов.

Пример трех компонент сейсмограммы типичного карьерного взрыва и результат их трансформации в характеристическую функцию C показаны на рис. 2. Специфическое соотношение амплитуд продольных (P) и поперечных (S) волн по трем направлениям создает в расчетном окне узнаваемую бимодальную форму вектора обобщенной информации H и, соответственно, двухступенчатую характеристическую функцию C . Стационарность C обеспечивает хорошую аппроксимацию по осредненным (сглаженным) данным. Таким образом, развиваемый авторами подход к классификации сейсмических событий может быть усилен разработкой шаблонов с заданными свойствами путем моделирования огибающих вектора обобщенной информации.

Заметим, что в работе [Kortström, Uski, Tiira, 2016], например, используется другой подход – признаком взрыва считается не форма колебаний, а специфический частотный состав двух близко расположенных возмущений.

Исходные данные и шаблоны

Источником первичных данных для авторов служила региональная сеть, включающая несколько принадлежащих разным ведомствам сейсмических станций (см. рис. 1), с которых для анализа были получены достаточно полные и качественные записи. Так, на станции BRСR были отобраны сейсмограммы 35 достоверных промышленных взрывов и 19 региональных землетрясений с магнитудами $1.8 \leq M_L \leq 2.4$.

Станция BRСR удалена от активных сейсмических зон на расстояние от 220 до 430 км. Установленное характерное время прохождения и эффективного затухания сейсмического возмущения из источников на рассматриваемой территории до станции составляет 50–75 с. Для удобства реализации быстрых вычислений было принято расчетное окно шириной в 6144 отсч. (или 61.44 с) при частоте дискретизации сигнала 100 Гц.

Для каждого события по приведенному выше алгоритму строились характеристические функции, из которых отдельно для совокупностей взрывов и совокупностей землетрясений вычислялось по три шаблона: средний шаблон и два, соответствующих границам $S = \pm 0.5\sigma$, где σ – среднеквадратическое отклонение. Вычисленным шаблонам были присвоены используемые далее на рисунках обозначения: B и $BL \pm S$ – для взрывов; EQ и $EQ \pm S$ – для землетрясений.

Для реализации потоковой автоматической классификации необходим был учет фрагментарного попадания события в расчетное окно N , для чего использовалась функция

$$f(t) = A \left(\frac{t}{Tn^2} \right)^n \exp \left(n - \frac{t}{Tn} \right). \quad (6)$$

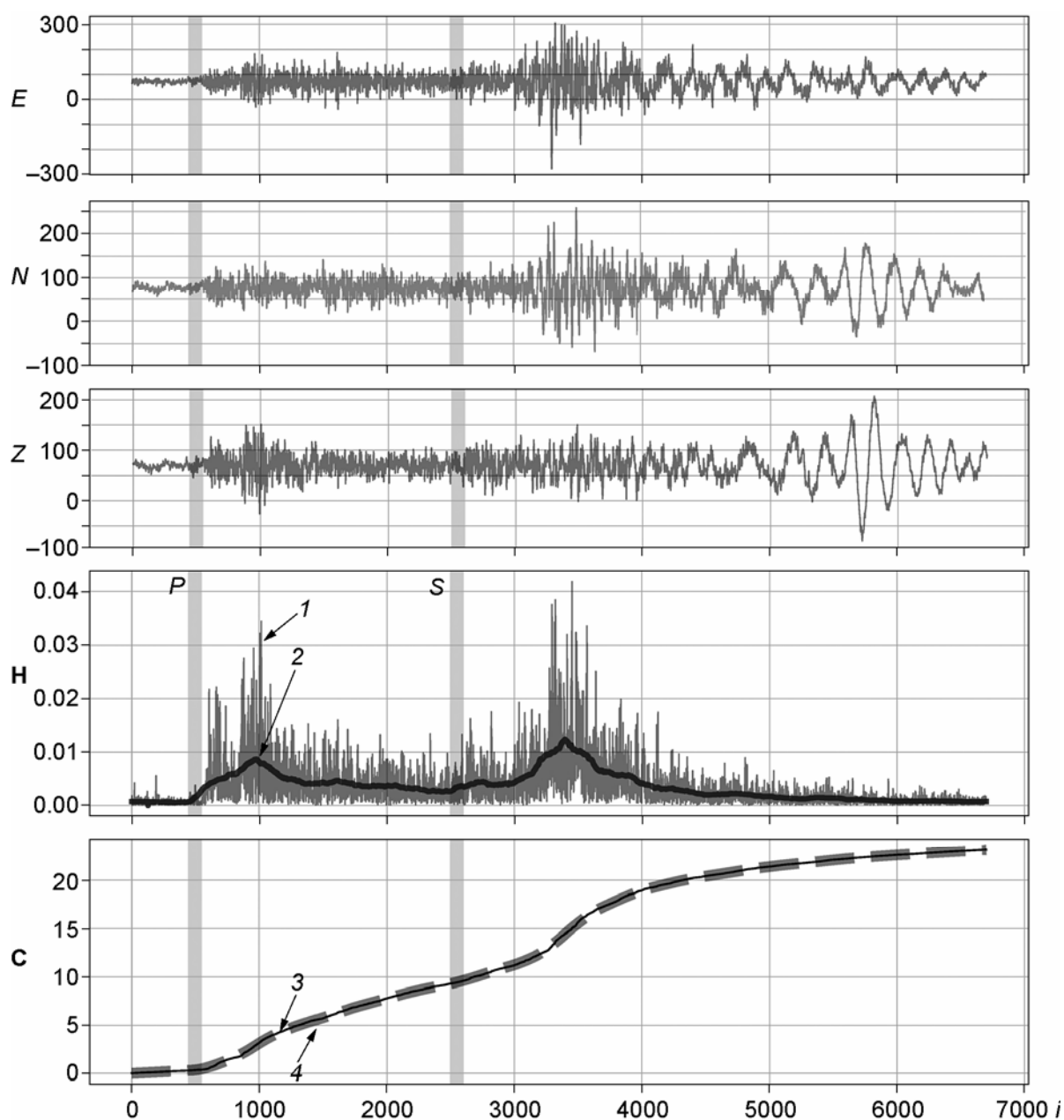


Рис. 2. Сейсмограммы компонент E , N , Z анализируемого сигнала (вверху), графики вектора обобщенной информации $\mathbf{H}$ (по центру) и характеристической функции $\mathbf{C}$ (внизу). 1 – график вектора $\mathbf{H}$, рассчитанный по сигналам, приведенным вверху; 2 – тот же график, сглаженный скользящим средним по 128 отсч.; 3 – график функции $\mathbf{C}$, рассчитанный из исходного вектора $\mathbf{H}$; 4 – тот же график, рассчитанный из сглаженного вектора $\mathbf{H}$

Fig. 2. Seismograms of E , N , Z components of the analyzed signal (on the top), graphs of generalized information vector $\mathbf{H}$ (in the middle) and characteristic function vector $\mathbf{C}$ (in the bottom). 1 – graph of vector $\mathbf{H}$, calculated from the signals on the top; 2 – the same graph, smoothed by moving average through samples 128; 3 – characteristic function vector $\mathbf{C}$, calculated from the original vector $\mathbf{H}$; 4 – the same graph calculated from the smoothed vector $\mathbf{H}$

Путем изменения значений параметров T и n в (6) при $t=0, \dots, N$ был сформирован набор одномодальных огибающих для вектора обобщенной информации $\mathbf{H}$ (рис. 3).

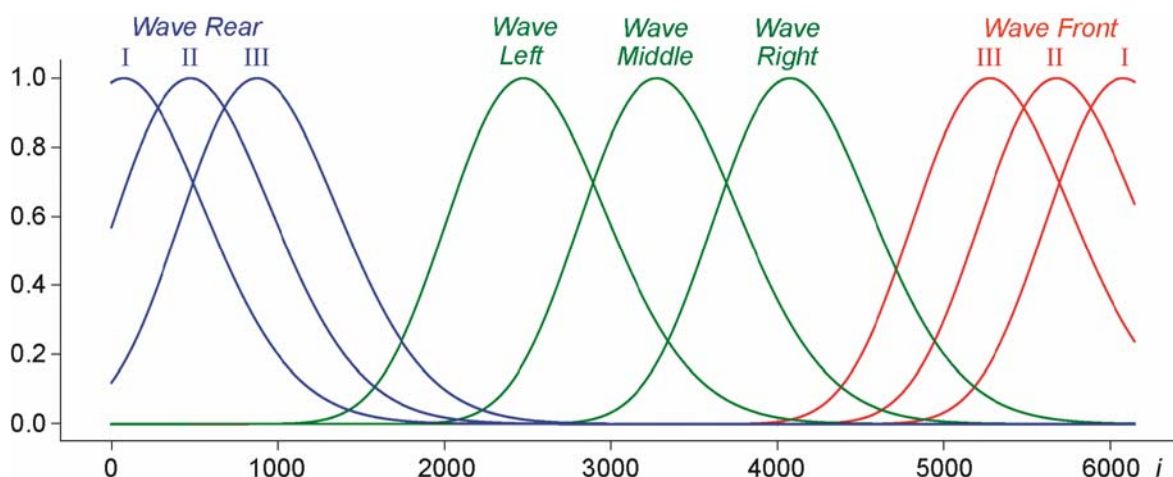


Рис. 3. Одномодальные огибающие вектора обобщенной информации $\mathbf{H}$, использованные для построения абстрактных шаблонов. Комментарии в тексте

Fig. 3. Single-mode enveloping vectors of generalized information $\mathbf{H}$ used for constructing abstract templates. Comments in the text

По формуле (5) из этих огибающих получен набор абстрактных характеристических функций (шаблонов), отображающих последовательное прохождение одномодального сейсмического возмущения через расчетное окно (рис. 4): три шаблона на входе в расчетное окно – *Wave Front* (*WF I*, *WF II*, *WF III*); три на выходе из него – *Wave Rear* (*WR I*, *WR II*, *WR III*); три внутри расчетного окна – *Wave Middle* (по середине, *WM*), *Wave Left* (смещение влево, *WL*), *Wave Right* (смещение вправо, *WRt*). Дополнительно был введен абстрактный шаблон *White Noise* (*WN*), отображающий сейсмический фон и возможные мультимодальные малоамплитудные возмущения; этот шаблон получен как характеристическая функция одноуровневого сигнала $f(t)=1$.

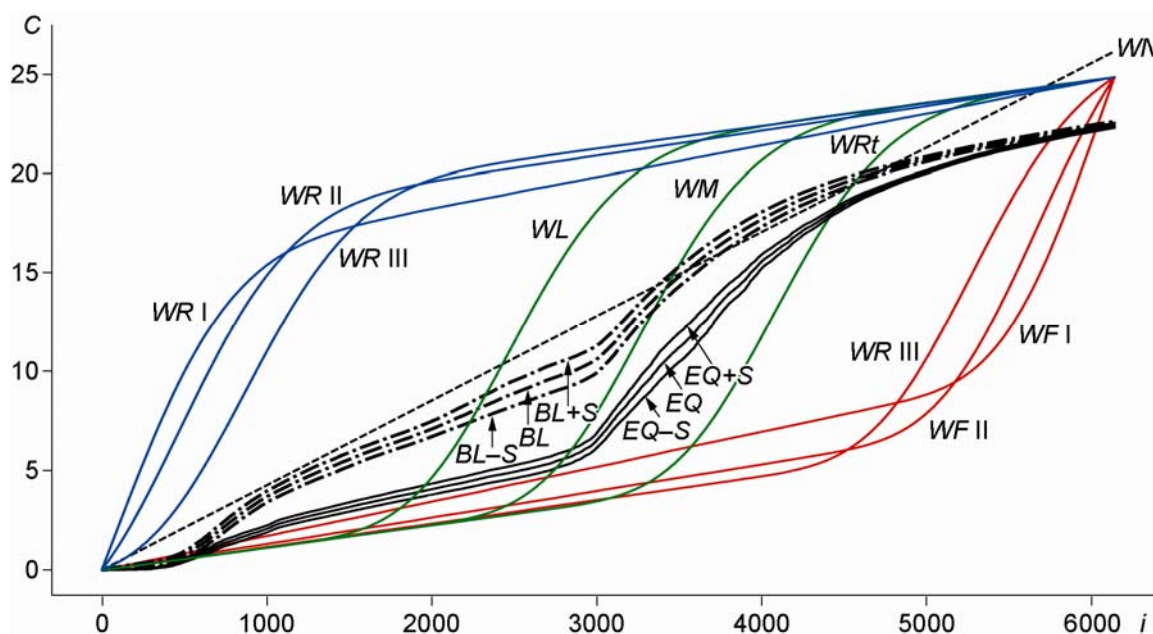


Рис. 4. Полный набор именных шаблонов сейсмических событий

Fig. 4. A full set of nominal templates of seismic events

Диагностическая матрица и алгоритм принятия решения

Согласно алгоритму вычисления характеристической функции, все шаблоны находятся в одном метрическом пространстве и имеют общий предел насыщения. Тогда, полагая шаблоны признаками, а отсчеты объектами (независимыми наблюдениями), можно дополнить их набор выборочной характеристической функцией и вычислить аналог диагностической матрицы по Байесу путем стандартизации в объектах:

$$\mu_i = \frac{1}{K} \sum_k C_{i,k}; \quad \sigma_i = \sqrt{\frac{1}{K} \sum_k (C_{i,k} - \mu_i)^2};$$

$$i = 1, \dots, N; \quad k = 1, \dots, K + 1,$$
(7)

где k – номера шаблонов; K – число шаблонов. Теперь в матрице (7) можно оценить подобие выборочной характеристической функции каждому шаблону из набора как дистанцию между двумя признаками (одномерными векторами). Априорных суждений о преимуществах тех или иных дистанций не существует, поэтому в текущей версии алгоритма используются все пригодные для номинальных признаков [Guojun Gan, Chaoping Ma, Jianhong Wu, 2007] с различными вариациями; всего используется 12 дистанций.

После этого можно реализовать простейшую систему голосования, в которой каждая дистанция имеет один голос, отдаваемый шаблону с минимальной дистанцией до выборочной характеристической функции. Простым суммированием голосов для каждого шаблона определяется его рейтинг R_k , $k=1, \dots, K$, на основании которого делается вывод о ближайшем шаблоне и оценивается надежность этого вывода по следующей схеме:

1. Заполняется таблица дистанций $K \times L$, где L – число дистанций.
2. Подсчитывается рейтинг каждого шаблона R_k – число случаев, когда выбранная дистанция до выборочной характеристической функции минимальна.
3. Определяется номер шаблона с максимальным рейтингом – $k_{\max} = \operatorname{argmax}(R)$.
4. Если значение $k_{\max}$ не является единственным, т.е. существует еще один шаблон, получивший такое же количество голосов, то выборочная характеристическая функция остается без класса.
5. Если значение $k_{\max}$ единственное, то классом выборочной характеристической функции признается шаблон с номером $k_{\max}$, после чего устанавливается надежность вывода:

- 5.1. при $k_{\max} \geq 3/4L$ (абсолютное большинство) вывод считается достоверным;
- 5.2. при $k_{\max} \geq 2/3L$ (квалифицированное большинство) вывод почти достоверный;
- 5.3. при $k_{\max} \geq 1/2L + 1$ (простое большинство) вывод допустимый;
- 5.4. при $k_{\max} < 1/2L + 1$ выборочная характеристическая функция остается без класса.

Таким образом, сканируя сигналы с заданным шагом по времени (числу отсчетов) и применяя предложенный алгоритм, можно зафиксировать появление обособленного события и классифицировать его.

Тестирование и обсуждение результатов

Рассмотрим возможности и ограничения разработанного алгоритма на примере типичных ситуаций, возникающих при потоковой классификации сейсмограмм с его использованием.

Ниже в таблице приведен фрагмент протокола региональной Службы сейсмических наблюдений Агентства по защите населения и чрезвычайным ситуациям

Фрагмент протокола Службы сейсмических наблюдений Агентства по защите населения и чрезвычайным ситуациям
Администрации Кемеровской области

№ события	Наименование события	Время по Гринвичу	Местное время	с.ш.	в.д.	K	M	Пункт наблюдения	Местоположение	Наименование разреза	Объем ВВ, т
2398	промышленный взрыв	07:15:12:83	14:15:12:83	53.43	89.11	7.1	1.7	Кемерово	Республика Хакасия	Прокопьевский	33
2399	промышленный взрыв	07:41:54:65	14:41:54:65	53.93	86.63	7.1	1.7	Кемерово	Прокопьевский район	Виноградовский	86
2400	промышленный взрыв	07:47:21:98	14:47:21:98	54.32	86.85	7.6	2.0	Кемерово	Беловский район	Междуречье	104.6
2401	промышленный взрыв	08:25:52:23	15:25:52:23	53.66	88.00	7.6	2.0	Кемерово	Междуреченский район	Сибиргинский	120
2402	промышленный взрыв	08:27:46:01	15:27:46:01	53.63	87.82	7.1	1.7	Кемерово	Междуреченский район	Черниговец	414
2403	промышленный взрыв	08:41:35:17	15:41:35:17	55.62	86.07	9.2	2.9	Кемерово	Кемеровский район	Салек	49
2404	промышленный взрыв	09:21:43:03	16:21:43:03	54.26	86.80	7.1	1.7	Кемерово	Прокопьевский район	Виноградовский	179
2405	промышленный взрыв	09:29:28:48	16:29:28:48	54.35	86.83	7.4	1.9	Кемерово	Беловский район		
2406	промышленный взрыв	09:49:44:20	16:49:44:20	53.77	88.18	7.2	1.8	Кемерово	Междуреченский район		
2407	промышленный взрыв	09:59:12:83	16:59:12:83	53.67	87.04	7.2	1.8	Кемерово	Новокузнецкий район		
2408	далекое землетрясение	11:46:09:91	18:46:09:91	28.76	82.29		5.0	Кемерово	Непал		
2409	сейсмическое событие	15:35:36:58	22:35:36:58	54.32	87.73	7.6	2.0	Кемерово	Новокузнецкий район		
2410	сейсмическое событие	16:06:30:71	23:06:30:71	54.30	86.15	6.5	1.4	Кемерово	Беловский район		
2411	сейсмическое событие	16:33:51:41	23:33:51:41	54.36	87.79	7.2	1.8	Кемерово	Новокузнецкий район		
2412	региональное землетрясение	17:23:27:60	00:23:27:60	50.33	86.93	12.6	4.8	Кемерово	Республика Алтай		
2413	далекое землетрясение	23:59:12:86	06:59:12:86	24.04	122.06		5.6	Кемерово	Тайвань		

Администрации Кемеровской области, в котором содержится информация о рассматриваемых в работе событиях.

Основными параметрами регистрируемых событий при включении их в протокол являются дата и время события, магнитуда и координаты (широта/долгота). Если какое-либо заметное возмущение произошло в течение рабочего дня и его координаты попали в границы горного отвода предприятия, ведущего взрывные работы, то на это предприятие Службой делается запрос. При подтверждении проведения взрыва в указанное время зарегистрированное событие получает наименование “промышленный взрыв”; в протоколе указывается наименование разреза, на котором он выполнен, и объем взрывчатки.

Однако на практике существуют обстоятельства, вносящие неопределенность в протокол. Так, не все предприятия сотрудничают со Службой сейсмических наблюдений. В случае проведения массированных взрывных работ (несколько взрывов, разнесенных во времени и по бортам разреза) подтверждается только факт их проведения и объем взрывчатки на всю серию. Зачастую предоставляется не фактическая, а плановая информация по взрывным работам, вследствие чего возникают отклонения по времени, которые могут составлять десятки минут от фактического момента подрыва.

В качестве примера на рис. 5 представлена развертка потоковой классификации сейсмограммы для временного интервала от 7:35 до 10:05 GMT, в течение которого, как это следует из таблицы, в протоколе зарегистрировано девять событий.

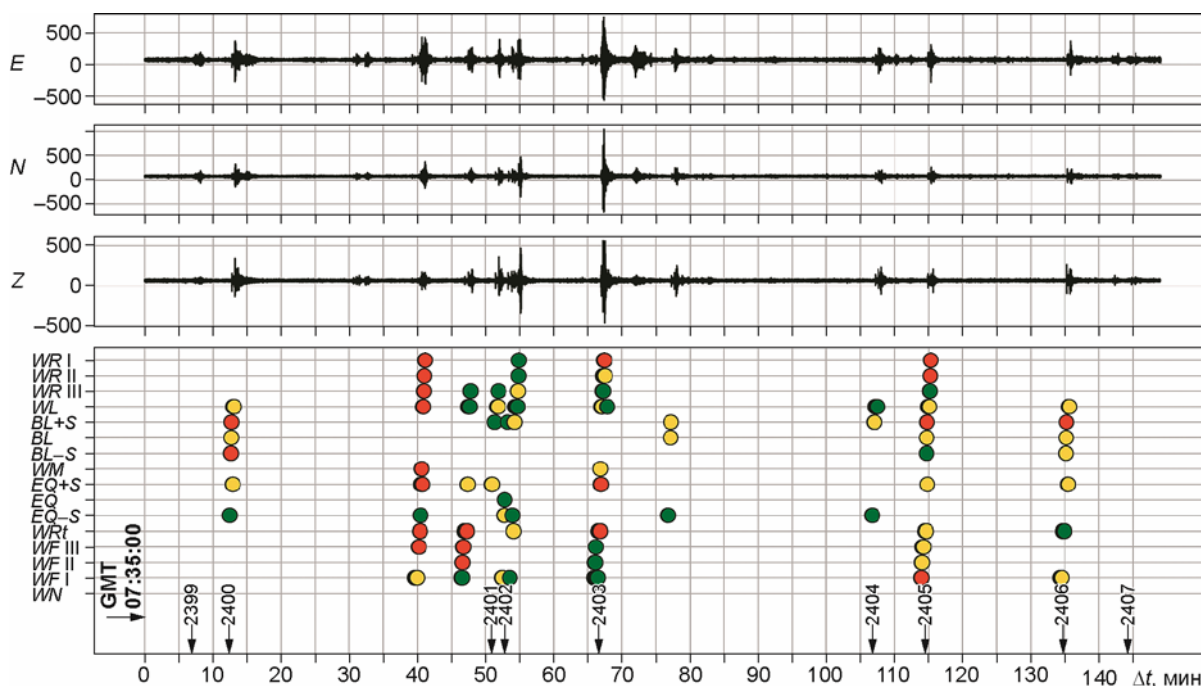


Рис. 5. Трехкомпонентная сейсмограмма длительностью 2 ч 30 мин с шагом 1 с (вверху) и развертка ее автоматической классификации (внизу). Кружки – моменты успешной классификации: красные – достоверный вывод; желтые – почти достоверный; зеленые – допустимый. Δt , мин – время от начала рассматриваемого интервала (фрагмента); стрелками отмечены моменты сейсмических событий с указанием их номеров. На вертикальной оси – имена шаблонов

Fig. 5. Three-component seismogram with a duration of 2 h 30 min in 1-second increments (on the top) and sweep of its automatic classification (in the bottom). Circle markers indicate the moments of successful classification, their color indicates output reliability: red – reliably; yellow – almost sure; green – acceptably. Δt , min – time from the beginning of the considered interval (fragment); arrows mark moments of seismic events with their numbers. On the vertical axis are given the template names

Можно видеть, что из семи подтвержденных в протоколе взрывов (см. табл., события №№ 2399–2405) классификатором правильно определены пять с номерами №№ 2400–2402, 2404, 2405. Промышленный взрыв на разрезе “Прокопьевский”, внесенный в протокол как событие № 2399, на сейсмограмме проявился слабо и классификатором замечен не был.

Два отсутствующих в протоколе события, одно из которых зафиксировано при $\Delta t \approx 40$ мин (рис. 6, 1), а другое при $\Delta t \approx 46$ мин (рис. 6, 2), были классифицированы как природные. На соответствующем рисунке можно видеть, что на сейсмограмме первого события отсутствуют признаки, характерные для взрывов, в связи с чем оно достоверно отнесено к землетрясениям.

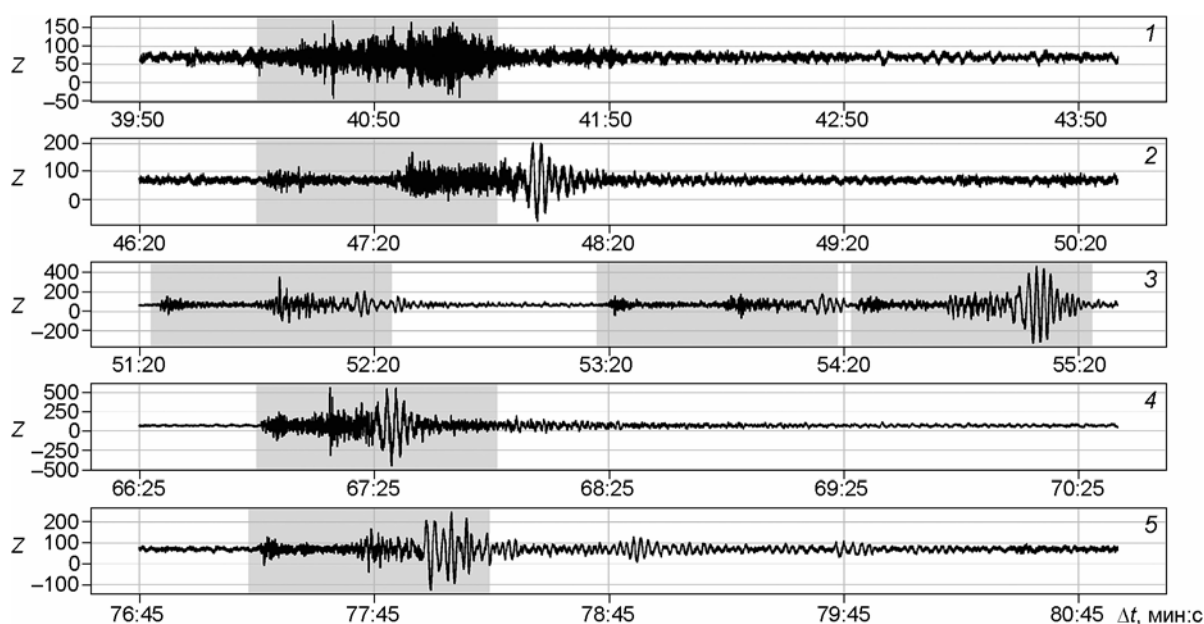


Рис. 6. Фрагмент приводимой на рис. 5 сейсмограммы компоненты Z в увеличенном масштабе; заливкой выделена ширина расчетного окна 61.44 с

Fig. 6. Zoomed view of the fragment of seismogram given in Fig. 5 (component Z). A 61.44 seconds wide calculation window is highlighted

Сейсмограмма второго события (см. рис. 6, 2) позволяет считать его взрывом, но заметно удаленным от средней дистанции шаблона – эффективное затухание этого события произошло за пределами ширины расчетного окна, что не позволило его верно классифицировать.

Сложное возмущение от двух подтвержденных взрывов (события №№ 2401, 2402), произведенных в соседних разрезах, отмечено вблизи $\Delta t \approx 51$ мин. При увеличении масштаба развертки классификации по оси времени (рис. 7) стало видно, что классификатор обнаружил в этом интервале времени три взрыва, наличие которых подтверждается приведенной на рис. 6, 3 сейсмограммой – промежуточное маломощное событие хорошо укладывается в расчетное окно и имеет все визуальные признаки взрыва.

Отметим одну из особенностей работы алгоритма. Возмущение классифицируется по мере его протекания через расчетное окно, поэтому промышленный взрыв сначала порождает ложные классификации землетрясения. И только когда возмущение от взрыва “заходит” в расчетное окно полностью, событие классифицируется верно. Понимание

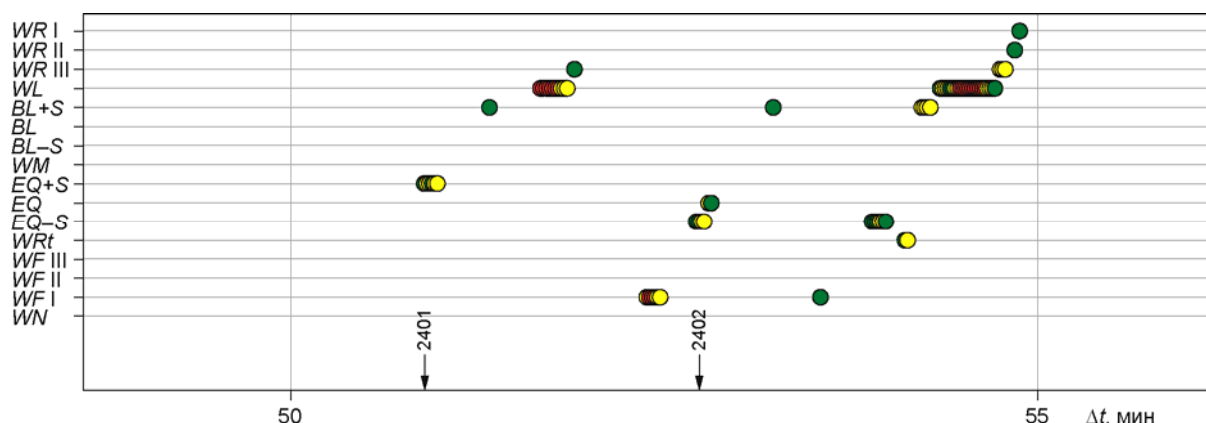


Рис. 7. Развертка автоматической классификации для фрагмента сейсмограммы, приведенного на рис. 6, 3. Условные обозначения см. на рис. 5. Стрелки на горизонтальной оси – моменты событий, отнесенных в протоколе к промышленным взрывам

Fig. 7. Sweep of automatic classification for the fragment of seismogram shown in Fig. 6, 3. See Fig. 5 for legend. Arrows on the time axis show moments of events referred to industrial explosions in the protocol

этой особенности необходимо на стадии синтаксического разбора протокола классификации, с тем чтобы сформулировать верное итоговое заключение.

Самый мощный на рассматриваемом временном интервале (фрагменте протокола) взрыв, произведенный на разрезе “Черниговец” (событие № 2403), не был верно классифицирован. Это объясняется тем, что названный разрез расположен севернее г. Кемерово (см. рис. 1), в стороне от основной массы разрезов, и интервал эффективного затухания события оказался заметно короче ширины расчетного окна. По этой причине взрыв и не мог быть классифицирован верно.

Еще одно незарегистрированное событие (см. рис. 5, $\Delta t \approx 76$ мин) классифицировано как взрыв и по форме сигнала (см. рис. 6, 5) полностью ему соответствует.

У события № 2406 классификация подтверждена; событие № 2407 на выбранной станции проявилось слабо и классификатором не было замечено.

В нерабочее время с 15:30 до 17:40 GMT в протоколе зарегистрировано три сейсмических события №№ 2409–2411 и одно недалекое региональное землетрясение № 2412. Событие № 2410 на сейсмограмме не проявляется и классификатором не было замечено. Остальные события правильно классифицированы как природные (рис. 8).

Тестовые классификации показали, что их качество в наибольшей мере зависит от начальной энергии сейсмического возмущения, отраженной в магнитуде события M_L , а также от уровня сейсмического шума, отклонения от средней дистанции шаблонов и технологии взрыва. Успешность классификации оценивалась на основании протокола региональной Службы сейсмических наблюдений по тем взрывам, которые были подтверждены предприятиями, т.е. имелись сведения об объеме взрывчатки, а также по неклассифицированным событиям, которые произошли в нерабочее время, т.е. полагались заведомо природными. Для совокупности зон ведения горных работ на территории Кузбасса (см. рис. 1) граница надежного срабатывания рассматриваемого алгоритма составляет $M_L \geq 1.8$ – при такой энергии событий число успешных классификаций достигает 90 %. При меньших энергиях ($M_L \approx 1.6$) качество снижается до 80 % и падает до неудовлетворительного уровня уже при $M_L \leq 1.4$.

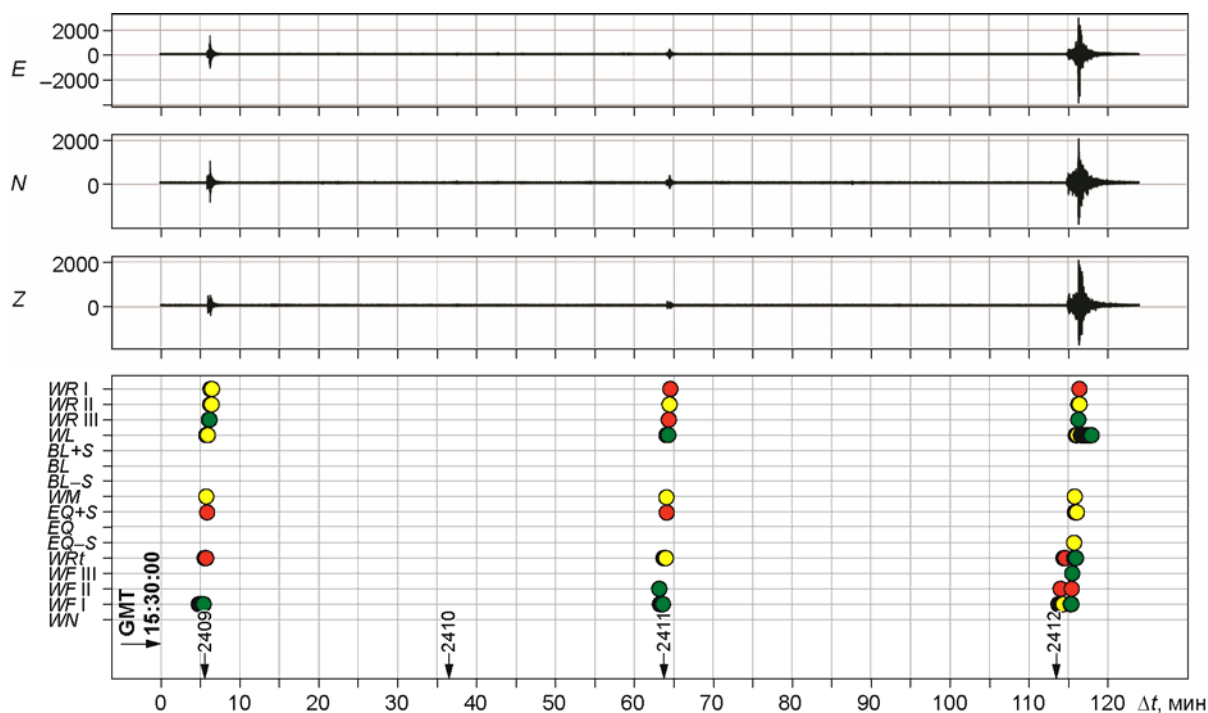


Рис. 8. Развертка автоматической классификации трехкомпонентной сейсмограммы длительностью 2 ч 10 мин с шагом в 1 с. Условные обозначения см. на рис. 5

Fig. 8. Sweep of automatic classification of the three-component seismogram with a duration of 2 h 10 min in 1-second increments. See Fig. 5 for legend

Понятно, что использование записей более близких станций может, с одной стороны, улучшить распознавание слабых взрывов, с другой – появляется опасность повышения уровня сейсмического шума при приближении к промышленным зонам.

Настройка системы классификации для новой станции требует перерасчета абстрактных шаблонов по формуле (6) для другой ширины расчетного окна при существенном изменении расстояния. Также потребуется некоторое количество представительных сигналов от взрывов и землетрясений для построения их шаблонов. В остальном же алгоритм классификации является полностью автоматическим и не требует вмешательства оператора ни на стадии обнаружения сейсмического события, ни на стадии получения вывода.

Заключение

Разработан алгоритм классификации сейсмических событий, в котором используется оригинальный классификационный признак – характеристическая функция сейсмического поля. Она производится из сейсмического сигнала с помощью энтропийной модели и обобщает особенности сейсмических колебаний по трем направлениям измерений. Характеристическая функция используется для оценки подобия сейсмического поля в расчетном окне шаблонам взрывов и региональных землетрясений.

Алгоритм ориентирован на региональные сети и системы сейсмического мониторинга и призван решить задачу быстрого различения взрывов и региональных (микро) землетрясений, которые перекрываются по магнитуде и локализации в зонах интенсивных горных работ.

Алгоритм прост с точки зрения программной реализации и организации потоковых вычислений, быстро адаптируется к новым сейсмическим станциям, обеспечивая полностью автоматизированный процесс обнаружения и классификации сейсмических событий.

Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований в рамках научного проекта № 18-07-00013А.

Литература

- Замараев Р.Ю., Попов С.Е., Логов А.Б. Алгоритм классификации сейсмических событий на основе энтропийного отображения сигналов // Физика Земли. 2016. № 3. С.31–37.
- Guojun Gan, Chaoqun Ma, Jianhong Wu. Data clustering: theory, algorithms, and applications (ASA-SIAM series on statistics and applied probability). Society for Industrial and Applied Mathematics Philadelphia, PA, USA. 2007. 466 p. ISBN: 978-0-898716-23-8.
- Kortström J., Uski M., Tiira T. Automatic classification of seismic events within a regional seismograph network // Computers & Geosciences. 2016. N 87. P.22–30.
- Lyubushin A.A., Kaláb Z., Lednická M., Haggag H.M. Discrimination of earthquakes and explosions using multi-fractal singularity spectrum properties // J. Seismol. 2013. V. 17, N 3. P.975–983.
- MacKay D. Information Theory, Inference, and Learning Algorithms. Cambridge: Cambridge University Press, 2003. 640 p. ISBN 0-521-64298-1.

Сведения об авторах

ЗАМАРАЕВ Роман Юрьевич – кандидат технических наук, старший научный сотрудник, Институт вычислительных технологий СО РАН. 630090, Новосибирск, просп. академика Лаврентьева, д. 6. Тел.: +7(384) 221-14-00. E-mail: zrukem@gmail.com

ПОПОВ Семен Евгеньевич – кандидат технических наук, старший научный сотрудник, Институт вычислительных технологий СО РАН. 630090, Новосибирск, просп. академика Лаврентьева, д. 6. Тел.: +7(384) 221-14-00. E-mail: popov@ict.sbras.ru

AUTOMATIC DETECTION AND CLASSIFICATION ALGORITHM OF INDUSTRIAL BLASTS BASED ON THE ENTROPY MAPPING OF SEISMIC SIGNALS

R.Yu. Zamaraev, S.E. Popov

Institute of Computational Technologies SB RAS, Novosibirsk, Russia

Abstract. The seismic event classification algorithm based on the preliminary transformations of signals into the entropy model is presented. It allows summarizing the information about the waveform features from three directions of measurements, averaging and comparing these features. The events' characteristic functions based on entropy model are calculated for estimating the similarity of the seismic field in the calculation window to templates of explosions, earthquakes and abstract templates with given properties. Templates for explosions and earthquakes are obtained by averaging characteristic functions of the sets of corresponding events for a given station in a given region. Abstract templates are calculated using the envelope function of the seismic field's entropy model, which is a sum of entropy models of signals in three dimensions. The similarity of the sample characteristic function to premade templates is estimated using a set of distance measures between one-dimensional vectors such as correlation distance, etc. Conclusion is issued on the basis of the rating voting system according to the amount of minimum distances, which is used in. An example of a seismic events classification in the Kemerovo region territory is given. The algorithm is oriented to regional networks and seismic monitoring systems

and provides a fully automated and fast process for detecting and classifying the declared regional seismic events.

Keywords: regional seismic events, industrial blasts, event detection, rapid automatic classification, entropy models, events' templates, similarity estimations.

References

- Guojun Gan, Chaoqun Ma, Jianhong Wu. Data clustering: theory, algorithms, and applications (ASA-SIAM series on statistics and applied probability). Society for Industrial and Applied Mathematics Philadelphia, PA, USA, 2007. 466 p. ISBN: 978-0-898716-23-8.
- Kortström, J., Uski, M., Tiira, T. Automatic classification of seismic events within a regional seismograph network, *Computers & Geosciences*, 2016, no. 87, pp. 22-30.
- Lyubushin, A.A., Kaláb, Z., Lednická, M., and Haggag, H.M. Discrimination of earthquakes and explosions using multi-fractal singularity spectrums properties, *J. Seismol.*, 2013, vol. 17, no. 3, pp. 975-983.
- MacKay David. Information Theory, Inference, and Learning Algorithms. Cambridge: Cambridge University Press, 2003. 640 p. ISBN 0-521-64298-1.
- Zamaraev R.Yu., Popov S.E., and Logov A.B. The algorithm for classifying seismic events based on the entropy mapping of signals, *Izv., Phys. Solid Earth*, 2016, vol. 52, no. 3, pp. 364-370.