

УДК 550.34

ВОЗМОЖНОСТИ ПРЕДПОДГОТОВКИ СЕЙСМИЧЕСКИХ ДАННЫХ ДЛЯ АНАЛИЗА ГЛУБОКОЙ НЕЙРОННОЙ СЕТЬЮ

© 2020 г. К. В. Кислов^{1, *}, В. В. Гравиров^{1, 2}, Ф. Э. Винберг¹

¹Институт теории прогноза землетрясений и математической геофизики РАН, г. Москва, Россия

²Институт физики Земли им. О.Ю. Шмидта РАН, г. Москва, Россия

*E-mail: kvkislov@yandex.ru

Поступила в редакцию 25.04.2019 г.

После доработки 17.06.2019 г.

Принята к публикации 24.06.2019 г.

Алгоритмы автоматической обработки сейсмических записей постоянно совершенствуются, задачи анализа данных усложняются. Большинство алгоритмов требуют предварительной подготовки данных. Эта обработка является либо очень простой, такой как частотная фильтрация, либо высоко-специализированной, выделяющей специфические особенности сигнала. Удачная предобработка может на порядок повысить эффективность дальнейшего анализа. Однако специфическая предобработка не может использоваться для решения иных задач или с другими алгоритмами постобработки. Мы рассматриваем решения, которые не приводят к значительным потерям информации, такие варианты предобработки, которые можно использовать при решении любых задач. Основными целями предварительной обработки являются снижение уровня шума, устранение помех антропогенного происхождения и уменьшение размерности данных, то есть устранение их избыточности. Мы предполагаем, что для последующей обработки данных используются глубокие нейронные сети той или иной архитектуры, но это не исключает возможности применения других алгоритмов. В данной работе в качестве предварительной обработки сейсмических данных мы рассматриваем вейвлет-преобразование, автоэнкодер и некоторые другие алгоритмы.

Ключевые слова: анализ сейсмических данных, глубокие нейронные сети, предобработка.

DOI: 10.31857/S0002333720010056

ВВЕДЕНИЕ

Объем сейсмических данных увеличивается экспоненциально, создавая потребность в эффективных алгоритмах автоматической обработки. Под обработкой понимается детектирование сигналов землетрясений, идентификация их основных фаз с определением их времен вступления, амплитуд и частотного состава, направления прихода волны. Далее оцениваются эпицентральное расстояние, магнитуда, глубина очага и т.д. В зависимости от целей анализа интерес могут представлять те или иные параметры сейсмического сигнала или шума. В воздухе витает идея создания универсального инструмента анализа сейсмических данных. Мы имеем в виду многозадачную глубокую нейронную сеть (ГНС) [Kisllov, Gravirov, 2018; Ross et al., 2018; Yuan et al., 2018; Meyer et al., 2019]. Ее возможно легко настроить на решение любой задачи. Современные архитектуры сетей включают ячейки долговременной памяти (LSTM), что снимает ограничения на длину сигнала. Однако, в связи со спецификой сейсмических данных и из-за неопределимого заранее круга

задач, потребуется очень глубокая сеть. Кроме того, для первоначального обучения сети мало использовать какой-либо прототип. Надо обучать сеть именно на сейсмических данных, на больших данных, в идеале, на всех доступных данных! Сейсмические данные значительно отличаются от других временных рядов. Временные окна, которые обычно используются для анализа, часто содержат много шума, имеют высокую размерность, данные избыточны, в любом месте окна могут быть расположены импульсные сигналы. Данные могут быть представлены как одним, так и несколькими каналами. Помимо параметров источника сигнала и среды распространения на вид полезного сигнала влияют характеристики сейсмических приборов и средств регистрации. Кроме того, записи сейсмических явлений часто занимают значительный временной интервал, и на записи одновременно могут присутствовать сигналы из разных источников.

Чтобы упростить архитектуру сети и облегчить обучение, естественно проводить предподготовку данных. Обычно она проходит в два этапа. Первый — станционное детектирование сигнала,

фильтрация шума и исключение случайных выбросов. Для детектирования успешно применяют метод STA/LTA [Allen, 1982]. Известны также другие технологии [Madureira, Ruano, 2009; Кислов, Гравиров, 2010; Sheng et al., 2015; Yoon et al., 2015], в том числе нейросетевые [Wang, Teng, 1995; Dai, MacBeth, 1997; Zhao, Takano, 1999; Gentili, Michellini, 2006; Kislov, Gravirov, 2010; Гравиров и др., 2012]. Второй этап зависит от целей анализа. Вычисляется спектрограмма или скалограмма, огибающая сигнала, поляризационная трасса и т.д. [Kislov, Gravirov, 2017; Ганнибал, 2018]. Каждый раз заново разрабатывается новый метод предварительной обработки данных. Удачная предобработка может на порядки повысить эффективность дальнейшего анализа.

ГНС обладает замечательным свойством многозадачности, т.е. обученная сеть может быть легко перестроена на решение другой задачи. Если предполагается возможность адаптации модели к нескольким задачам (multi-task learning) и решение других, пока не определенных задач, следует посмотреть на подготовку данных с другой стороны. Предобработка данных чаще всего препятствует возможности использовать многозадачность ГНС. В результате предобработки может быть потеряна потенциально полезная информация, необходимая для решения других задач, в том числе еще не сформулированных. Необходимо учитывать, что потенциально полезная информация не должна быть потеряна. Естественный вывод, что самой надежной предобработкой является ее отсутствие. Нужно лишь масштабировать амплитуду сигнала в приемлемый для ГНС диапазон, а все проблемы переложить на нейронную сеть. Например, в работе [Perol et al., 2018] на вход сверточной сети подается окно трехкомпонентных данных в виде изображения 1000×3 (1000 отсчетов на три компоненты) без всякой предобработки этих данных. Однако таким методом можно решить лишь ограниченный круг задач. Предварительная обработка данных обязательна для достижения лучшей производительности ГНС [Zheng et al., 2018; Zhang et al., 2016].

Чтобы охарактеризовать десятисекундную сейсмическую запись, вполне достаточно тридцати числовых параметров. Если бы мы могли корректно их измерить, мы бы действительно далее изучали тайны Земли, а не тайны сигналов. Пока это не представляется возможным. Если предобработка выделит особенности сигнала, устранив его избыточность и позволит рассматривать в меньшем пространстве – это, несомненно, упростит алгоритмы последующей обработки и облегчит их обучение [Gravirov, Kislov, 2014]. Предобработка должна снижать аппаратные шумы, импульсные сигналы, ассоциирующиеся с антропогенным воздействием и давать заключение о пригодности сигнала для дальнейшего анализа.

Современные алгоритмы детектирования работают вполне надежно и могут за счет ложных срабатываний быть подстроены на 100-процентный захват цели. Будем считать, что большинство поступивших в предобработку окон временных рядов содержат полезный сейсмический сигнал. Этим мы устраним проблему дисбаланса данных при обучении нейронных сетей. Дисбаланс есть результат того, что один класс (например, землетрясения, предположим, что они образуют один класс) представлен в данных обучения редко по сравнению с другим классом (шумом). Алгоритм обучения нейронной сети создаст гипотезу, соответствующую небольшому количеству примеров, но слишком конкретно (т.е. произойдет переобучение для конкретной задачи или класса).

Надо создать алгоритм предобработки, при котором не будут потеряны информативные особенности сигнала. Его можно будет использовать при любом последующем анализе, хотя в первую очередь имеется в виду многозадачное применение ГНС. Наметим ряд методик, которые, как мы надеемся, будут представлять практический интерес. Подходы, обсуждаемые в этой статье, не предназначены для простого сжатия данных с целью хранения или передачи. Предподготовка должна облегчить дальнейший анализ, упростить архитектуру нейронной сети и ее обучение.

МЕТОДЫ ПРЕДОБРАБОТКИ ДАННЫХ

Предподготовка данных включает [Kong et al., 2018; Kotsiantis et al., 2006] форматирование данных, удаление и/или восстановление данных, извлечение информативных характеристик, масштабирование полученного вектора признаков.

При форматировании сигнал надо привести к единой частоте дискретизации и разрезать поток данных на окна определенной длины (например, по 1000 отсчетов). Возможно и другое форматирование данных. Вектор данных можно, например, формировать из амплитуд сигнала, соответствующих одному моменту времени, полученных по разным каналам регистрации и/или с разных компонент. Вместо амплитуд могут быть использованы отнесенные к одному моменту времени какие-либо представления сигнала. Такой вариант не очень удобен. Сейсмическое событие может занимать очень длинный участок временного ряда данных. Выявление его особенностей потребует особо сложной рекуррентной цепочки. Такое представление данных не приведет к упрощению ГНС, и мы ее рассматривать не будем.

Под удалением понимается исключение явно негодных данных, например, из-за превышения сигналом его динамического диапазона (clip). Можно попытаться восстановить или заменить поврежденные данные по тому или иному алго-

ритму. Возможен и несколько другой подход. Если применять ГНС с несколькими входами, каждый из которых обрабатывает свой канал или компоненту, можно при работе не только игнорировать один из входов, но и, используя ГНС в качестве генеративной модели, восстановить данные пропущенного входа с помощью прохода “сверху вниз”.

Извлечение характеристик — это процесс удаления как можно большего количества нерелевантной и избыточной информации. Наш подход использует шаблоны в самом наборе данных. Мы стараемся выделить отличия одной сейсмограммы от другой, а не отличия сейсмограмм от других наборов данных. Эти отличия затем будут предметом дальнейших вычислений или анализа [Guyon, Elisseeff, 2003]. В качестве метода извлечения характеристик может быть использовано преобразование сигнала или фильтрация шума без снижения размерности [Gravirou et al., 2013]. Способ извлечения характеристик может иметь физическую интерпретацию. Этим повышается наглядность метода, улучшается возможность контроля. С другой стороны, признаки, не имеющие четкого физического значения, могут обеспечить лучшую способность распознавания.

Масштабирование — перевод тем или иным способом амплитуд в приемлемый для ГНС диапазон. При этом и абсолютное значение, и соотношения этих величин в разных окнах могут иметь значение, т.е. данные масштабирования тоже должны поступать в сеть.

Для корректной работы ГНС необходимы сведения о частотных характеристиках сеймостанций, представленные, например, нолями и полюсами, а также географические координаты сейсмических датчиков и привязка регистрируемых сейсмограмм к точному времени. Эти метаданные подаются уже на один из последних слоев сети.

ПРИНЦИПЫ ИЗВЛЕЧЕНИЯ ИНФОРМАТИВНЫХ ХАРАКТЕРИСТИК ДАННЫХ

Пусть сигнал — это окно сейсмической записи длиной N отсчетов. Каждое окно w_i принадлежит множеству W^{*N} , в котором представлены только записи сейсмометров (для обозначения размерности используем надстрочные символы). Временные ряды, не имеющие сейсмического происхождения, не представлены в этом пространстве. В результате некоторого преобразования $v_i = f(w_i)$ мы превращаем w_i в другое представление v_i длиной M отсчетов, в котором видны особенности сейсмического сигнала. А все множество возможных окон W^{*N} отображается на множество представлений V^M .

$$V^M = F(W^{*N}). \quad (1)$$

При этом существует и обратное преобразование $\dot{w}_i = g(v_i)$; $W^{*N} = G(V^M)$.

Ошибки преобразований приводят к тому, что \dot{w}_i несколько отличаются от w_i . Это отличие не должно превышать некоторого значения E . Главное, чтобы при допустимой ошибке E восстановленный сигнал отображал то же самое сейсмическое событие, со значениями параметров в соответствующем допуске. Здесь плохо подходит формальная процедура определения ошибки E , например, применение метода наименьших квадратов.

Напомним, что в исходном сигнале содержится шум. Таким образом, реальное множество W^N значительно превосходит идеальное W^* . Плюс к этому, в погоне за компактностью представления v_i мы соглашаемся с некоторой потерей информации при прямом преобразовании. Таким образом, близкие входные сигналы будут иметь одно и то же компактное представление. Различие входных сигналов также не должно превышать E . Это приводит к дополнительному сужению V^M . Прямое и обратное преобразования должны быть стабильны относительно небольших возмущений на входах. “Эквидистантные” во временной области сейсмограммы должны быть “близки” после преобразования. Тогда анализ, проводимый во временной области, может быть в равной степени возможен в области преобразования. Если же вместо сейсмического сигнала w_i на вход преобразования $f(w_i)$ попадет сигнал, не принадлежащий W^N , он тоже отразится в V^M , но при обратном преобразовании восстановленный сигнал будет отличаться от входного с большой ошибкой. Если преобразование сигнала реализуется с помощью машинного обучения, есть смысл вносить в обучающую выборку шум.

Обычно данные искусственно зашумляются гауссовским шумом. Мы считаем этот подход порочным, хотя и наиболее простым. Гауссовский шум довольно легко поддается снижению, так как не является редким сигналом. Трудности возникают с нестационарными колеблющимися, прерывистыми и импульсными шумами. При этом техногенные шумы, которые чаще всего приводят к ошибкам в анализе сигналов, обычно нестационарны и далеки от распределения Гаусса. Искусственное зашумление правильнее производить смещением полезного сигнала с реальными шумовыми трассами [Гравиров, Кислов, 2015], хотя это требует большой дополнительной работы с обучающими данными. Заметим, что при искусственном зашумлении обучающих примеров (и их последующим масштабировании), надо ориентироваться на уровень сигнала. В противном

случае, при его низком уровне мы будем обучать алгоритм на чистом шуме.

Мы не пользуемся понятием “отношение сигнал/шум” (signal to noise ratio – SNR). В работе [Кислов, Гравиров, 2013] обоснована неоднозначность этой характеристики и предложен альтернативный подход к оценке шума. Мы будем просто говорить о том, какой вид помех мы имеем в виду.

Методы и способы решения поставленной задачи могут существенно различаться. Определим, что мы хотим от предварительной подготовки данных:

- выделение информативных особенностей сейсмического сигнала,
- подавление помех разнообразной природы,
- сжатие данных,
- возможность восстановления очищенного от шума сигнала,
- возможность работы алгоритма в реальном времени.

Главной, конечно, является первая процедура. Заметим, что указанные процедуры взаимосвязаны, и, выполнение одних часто приводит, хотя бы частично, к выполнению других. Восстановление сигнала не является обязательным, однако его возможность очень полезна для контроля преобразования данных и может быть необходима при использовании в качестве последующей обработки методов, отличных от нейросетевых.

CS-МЕТОД

Широко известно следствие теоремы В.А. Котельникова (которую теперь почему-то называют теоремой Найквиста–Шеннона) о том, что если максимальная частота в сигнале равна или превышает половину частоты дискретизации, то способ восстановить сигнал из дискретного в аналоговый без искажений не существует. Но наша цель сжать не сигнал, а информацию, которую он содержит [Candes et al., 2006; Donoho, 2006]. И в этом отношении теорема Котельникова не накладывает никаких ограничений.

Предположим, что сжатие сигнала подразумевает выделение, а не запутывание его информативных особенностей (что в общем случае неверно). Все методы сжатия данных разными способами стараются прийти к одному результату – найти такое F -преобразование (1), чтобы $M \ll N$. Тогда для нашего случая (см. ВВЕДЕНИЕ) идеальным вариантом было бы снижение размерности входного вектора с 1000 до 30.

Рассмотрим метод, который в русскоязычной литературе именуется опознанием, а в английской Compressed Sensing, Sparse Sampling или Compressive Sampling и традиционно обозначает-

ся буквами CS. В геофизике этот метод применяется уже около 40 лет [Taylor et al., 1979; Oldenburg et al., 1983; Santosa, Symes, 1986, Vera Rodriguez et al., 2012; Gan et al., 2016].

Пусть w^N – сигнал (исходные данные высокой размерности), который мы хотим проецировать в пространство более низкой размерности. Линейные отображения определяются матрицами $\Phi^{M \times N}$, которые действуют как кодер:

$$y_i = \Phi w_i, \quad (2)$$

Представим сигнал через некоторый базис $w = \Psi x$, где столбцы матрицы $\Psi^{N \times N}$ являются координатами базиса Ψ , а вектор-столбец x^N содержит коэффициенты, которые представляют сигнал w^N в базисе Ψ . Это эквивалентные представления одного и того же сигнала, w во временной области, а x в Ψ -области. Представление x^N будет называться s -разреженным, если x является линейной комбинацией только s базисных векторов, т.е. выполняется условие $\|x\|_{\ell_0} \leq s$ (ℓ_0 -норма определяется как количество ненулевых элементов вектора). Это означает, что все компоненты этого вектора равны нулю кроме s штук. Вектор x называется сжимаемым, если s компонент вектора достаточно велики, а остальные (большинство) малы. На самом деле, CS использует правило, что все естественные сигналы редки, когда они выражены надлежащим образом Ψ [Candès, Wakin, 2008]. После преобразования целевой сигнал должен хорошо аппроксимироваться разреженным вектором [Foucart, Rauhut, 2013]. Можно точно восстановить любой w , если $M \geq N$ и матрица Φ имеет полный ранг. Напротив, при $M < N$, уравнение $y - \Phi z = 0$ имеет бесконечное число решений z , что делает невозможным однозначное восстановление w^N по y^M , если они связаны соотношением (2). Для восстановления исходного сигнала CS использует информацию о том, что x – сжимаемый.

В соответствии с теорией CS, кодирование данного сигнала представляет собой простую линейную процедуру $y = Ax$, где матрица $A = \Phi\Psi$ часто называется матрицей измерений, сенсорной или чувствительной матрицей.

Матрица измерений должна удовлетворять RIP-условию (Restricted Isometry Property). Его выполнение с высокой вероятностью может быть достигнуто просто за счет выбора матрицы A со случайными элементами, распределенными, например, по нормальному закону или закону Бернулли (рандомизация процесса). Случайные матрицы в значительной степени некогерентны с любым фиксированным базисом Ψ [Bourgain et al., 2011]. Чем меньше когерентность, тем меньше может быть M . Случайные матрицы не бывают оптимальными или стабильными и зачастую применяются другие варианты [Ji et al., 2008; Gurbuz et al., 2009; Robucci et al., 2010].

Сжатие в некоторой степени снижает шум, который в большинстве случаев не имеет разреженного представления. Под шумом понимается то, что соответствует обнуляемым компонентам вектора. Процедура CS весьма эффективна при небольшом уровне шума [Ben-Haim et al., 2010; Haupt, Nowak, 2006]. Такое шумоподавление ожидается на гипотезе о сигнале малой размерности и большой амплитуды и о шуме большой размерности и маленькой амплитуды. CS может понизить не только гауссовский шум, но и любой симметрично распределенный шум с нулевым средним, не зависящий от элементов вектора проекции.

Следующий этап – выбор алгоритма восстановления. Можно ли корректно восстановить сигнал из сжатого вектора y , зависит от матрицы измерений, самого сжатого вектора y и алгоритма, используемого для восстановления. В CS восстановление сигнала возможно только с помощью нелинейных алгоритмов. В терминах линейной алгебры при $M < N$ имеются меньше уравнений, чем неизвестных, что делает решение в целом некорректным. Обычно в качестве меры соответствия восстановленного сигнала исходному или размера ошибки используется ℓ_1 -оптимизация. Восстановленный сжимаемый сигнал $\hat{x} = \arg \min \|x\|_{\ell_1}$ будет с высокой вероятностью соответствовать исходному при $M \geq C_1 s \log(N/s)$, где C_1 – постоянная, зависящая от способа создания матрицы измерений. Это задача выпуклой оптимизации, которую можно свести к задаче линейного программирования, известной как выбор базиса (basis pursuit) [Foucart, Rauhut, 2013]. Для восстановления применяют и другие алгоритмы [Blumensath, Davies, 2009].

Чтобы учесть шум преобразований и шум, изначально присутствующий в данных используют ℓ_1 -оптимизацию с ослабленными ограничениями: $\min \|\hat{x}\|_{\ell_1} : \|A\hat{x} - y\|_{\ell_2} \leq \epsilon$, где ϵ характеризует шум. Это выпуклая задача (конусная программа второго порядка), и она может быть эффективно решена.

Ошибка восстановления состоит из двух слагаемых: первое пропорционально величине сейсмического шума и шума регистрации, а второе – ошибка аппроксимации, которую можно получить в бесшумном случае. Таким образом, качество реконструкции изрядно ухудшается с увеличением шума в данных. Ни один метод восстановления не может быть принципиально лучше при произвольных возмущениях ϵ .

Мы уже отмечали трудности определения “величины ошибки” E для извлеченных из сейсмических данных характеристик, тем более, не зная заранее, какие характеристики могут понадобиться. Среди них можно упомянуть времена вступления отдельных фаз землетрясения, на-

правления прихода волн этих фаз на сейсмические датчики, частотные полосы фаз, их амплитуды и величины затуханий в среде. Оценить ошибку по этим параметрам возможно лишь, сравнив исходную и восстановленную сейсмограммы (рис. 1).

Очевидно, что сжатые данные имеют меньшую “приведенную погрешность”, чем восстановленные. При этом многие виды анализа возможны без восстановления сигнала [Davenport et al., 2010], например, если CS-процедура построена непосредственно в ГНС [Li et al., 2018]. Поскольку CS – это матричные вычисления, алгоритм может быть легко встроено в нейронную сеть как дополнительный слой. Параметры этого слоя фиксированы и не обновляются во время обучения [Shen et al., 2018].

CS все чаще применяется в сейсморазведке и обработке данных сейсмических сетей, где использование многих датчиков и сжатие их данных практически не ухудшает картину анализа [Yao et al., 2011]. Про сейсморазведку см. также [Hegmann, 2019].

Наибольшими препятствиями применения CS как средства подготовки данных для дальнейшего анализа, являются требование малости шума и необходимость высокой разреженности. Когда измерения искажаются случайным шумом, каждое уменьшение числа отсчетов m в два раза также снижает качество цифрового восстановления в два раза [Davenport et al., 2012]. Уровень s -разреженности сейсмических сигналов может быть недостаточен для существенного сжатия. При этом алгоритмы CS с восстановлением требуют больших вычислительных затрат и их применение затруднительно в системах реального времени.

СЖИМАЮЩЕЕ КОДИРОВАНИЕ С ТРАНСФОРМАЦИЕЙ

Для сжатия звука и видео, например, в форматах JPEG, MP3, используется трансформирующее кодирование, в котором также используются s -разреженные представления. Сигнал w^N преобразуется в вектор $y = \Phi w$, в котором s -компонент обладают большими значениями, а остальные $(N - s)$ малыми значениями, и поэтому они могут игнорироваться. В итоге кодируются сами s -значений и номера их позиций в векторе w .

В качестве трансформирующих преобразований обычно используются дискретные вейвлет-преобразование, Фурье или косинусное преобразование.

При анализе сейсмических данных сжимающее кодирование с трансформацией имеет ряд недостатков по сравнению с CS-методом. Даже при малом s , использование этих алгоритмов требует больших вычислительных затрат. Все N коэффициентов преобразования надо вычислять, хотя от большинства из них (кроме s штук) ника-

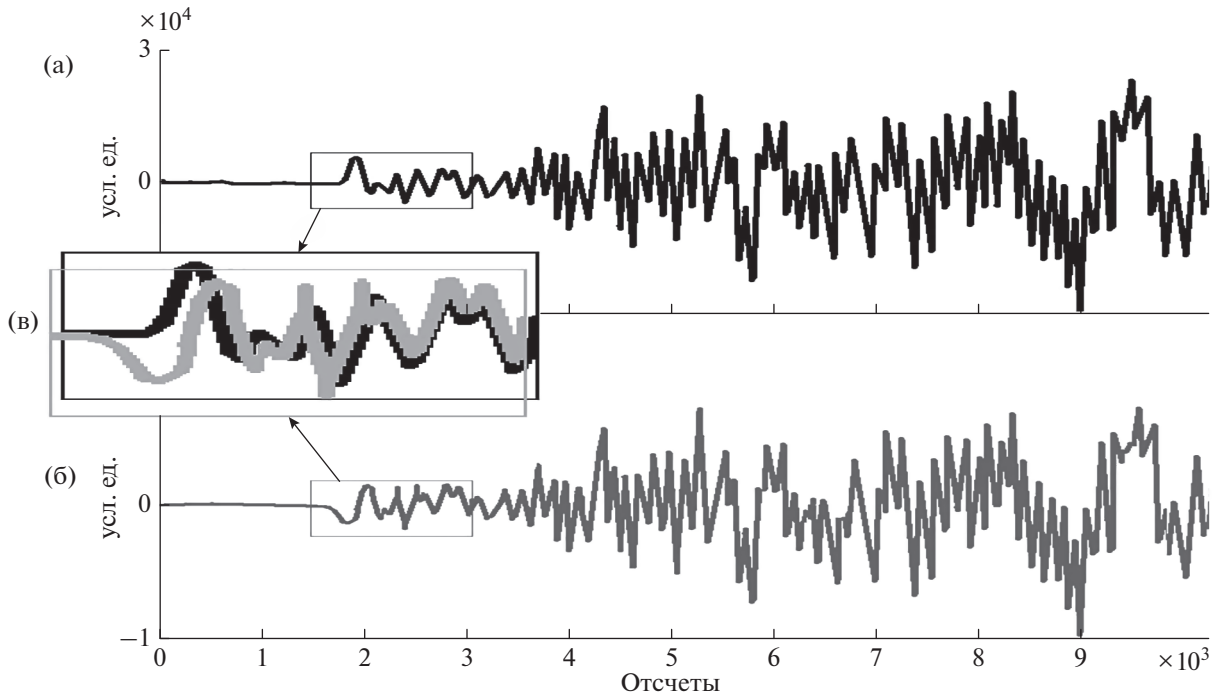


Рис. 1. (а) – запись первого вступления землетрясения (Оахака, Мексика 12.02.2005 г., $M = 5.0$, глубина 20.0 км, сейсмостанция АСАР); (б) – пример восстановления сигнала; сигнал 1024 точки, $M = 60$; (в) – сравнение формы и времени вступления.

кого толка. Дополнительно должны быть закодированы местоположения больших коэффициентов s [Граничин, Павленко, 2010]. Сжимающее кодирование явно проигрывает по сравнению с CS.

АВТОЭНКODЕР

Еще в 2006 г. было известно [Hinton, Salakhutdinov, 2006], что автоэнкодеры (АЕ) могут использоваться для поиска низкоразмерных особенностей сложных наборов данных. В последние годы вновь возрос интерес к этим нейронным сетям [Creswell, Bharath, 2019; Shi et al., 2019]. В работе [Bao et al., 2017], например, представлена сложная структура, состоящая из вейвлет-преобразования сигнала, его сжатия и выделения особенностей с помощью АЕ и последующей обработки рекуррентной нейронной сетью. Были проведены работы по использованию АЕ для сжатия сейсмических сигналов длиной 512 отсчетов в 32-мерный вектор [Valentine, Trampert, 2012].

Самый простой АЕ – это обычная нейронная сеть с одним скрытым слоем, обучаемая по методу обратного распространения ошибки реконструировать свои собственные входы ($y = x$). Входной и выходной слои имеют равное число нейронов. Вход x отображается на скрытый слой z (слой латентных переменных) $z = h_1(Wx + \beta_0)$. Затем z отображается на выходной слой $y = h_2(W'z + \beta_1)$.

Здесь h обозначает поэлементную функцию активации (например, сигмоидальную функцию, ReLU или другую); регулируемые параметры: β – смещения (bias weights, offsets), W – весовая матрица. Тем самым, АЕ также обучается сведению к минимуму ошибки восстановления (например, среднеквадратичной ошибки) $E(x, y) = \|x - y\|^2 = \|x - h_2(W'(h_1(Wx + \beta_0)) + \beta_1)\|^2$. Скрытый слой должен содержать меньше нейронов, чем входной (сокращение размерности). Применяют также разреженную активацию, тогда скрытый слой больше входного, причем количество активных нейронов значительно меньше неактивных, выдающих низкий сигнал. И сокращение размерности, и разреженная активация помогают в процессе обучения выявлять закономерности во входных данных [Кислов, Гравиров, 2017]. АЕ преобразует данные высокой размерности в данные низкой размерности, как и метод главных компонент (principal component analysis, PCA), но при этом преобразование является нелинейным. С соответствующими ограничениями АЕ более перспективны, чем PCA или другие известные методы.

Глубокий АЕ, содержащий много скрытых слоев, может более качественно сжимать данные, выявляя в них особенности. Обозначим выход i -го слоя z_i , $i \in \{1, \dots, q + 1\}$. При этом $z_0 \equiv x$; $y \equiv z_{q+1}$, где q –

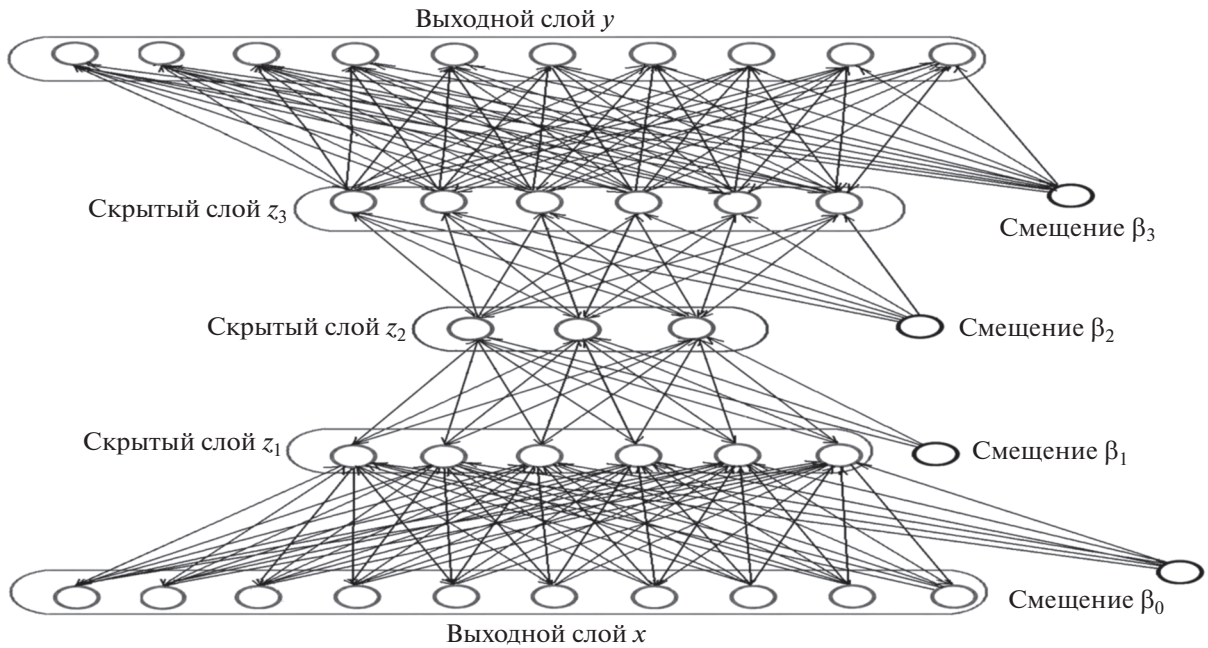


Рис. 2. Автоэнкодер с тремя скрытыми слоями.

количество скрытых слоев. Тогда $z_i = h_i(z_{i-1}W_i + \beta_i)$. Обычно глубокий АЕ содержит 5–7 слоев, иногда до 11-ти (рис. 2). В нашем случае размеры слоев 1000, 500, 250, 60, 250, 500, 1000 нейронов. Средний слой – 60-мерный вектор – и есть сжатое представление сигнала.

При обучении необходимо настроить матрицы W и смещения β так, чтобы минимизировать некоторую функцию потерь $L(x, y)$. Ее можно задавать по-разному, лишь бы сохранялись расстояния между ответами. Обычно используется алгоритм обратного распространения ошибки, хотя он не очень подходит для нескольких скрытых слоев с миллионами параметров. При правильном выборе шага градиентного спуска, он сходится к локальному минимуму. Процесс упрощается, если используются связанные веса $W_i = W_{q-i+2}^T$, то есть весовые матрицы кодирования и декодирования транспонированы друг для друга, что приводит к меньшему числу параметров, а, следовательно, к более быстрому обучению сети. С другой стороны, если отказаться от этого принципа, можно снизить ошибку восстановления сигнала [Majumdar, Tripathi, 2017].

Среди автоэнкодеров нас интересуют АЕ с возможностью сжатия данных с шумоподавлением (Stacked Denoising Autoencoders, SDAE), хотя частично эти функции поддерживаются и другими видами АЕ. SDAE – стохастический автоэнкодер. При его обучении на вход подается искусственно зашумленный сигнал. Зашумление обычно оформляется как добавление гауссовского шума

или случайный процесс обнуления части входов. Это заставляет АЕ выявлять наиболее сильные взаимосвязи, при этом не строя тождественных отображений. SDAE должен удалить шум, чтобы генерировать выходные данные, близкие к входным. Зашумление проводится только при обучении первого скрытого слоя. Для следующих слоев мы просто используем выход из предыдущих слоев.

Обучение глубоких АЕ может быть проведено с жадной послойной подготовкой, проводимой без учителя. Последнее время бытовало мнение, что это излишество и что случайная инициализация весов, использование функции активации ReLU и некоторые ухищрения дают возможность обучить сеть с помощью только точной настройки (fine-tuning). Однако для того, чтобы получить от АЕ максимальный эффект этого мало [Erhan et al., 2010]. Помимо подготовки без учителя, могут быть и более сложные алгоритмы. Например, после каждого этапа подготовки, можно настраивать нижележащие слои.

Чем больше слоев в АЕ, тем лучше выделяются признаки. Однако помимо того, что у глубокого АЕ снижается пропускная способность, значительно увеличивается трудность его обучения. Чрезмерное сжатие данных также ведет к потере информации [Gehring et al., 2013]. В пределе в центральном слое остается один нейрон. Тогда АЕ пронумеровывает все учебные примеры и восстанавливает их по номеру.

Ситуация с полнотой передачи информации и возможностью адаптации метода АЕ к реальному шуму во многом сходна с проблемами алгоритма

CS. С другой стороны, поскольку АЕ обучается на большом количестве реальных данных, взаимосвязи отдельных особенностей сигнала сохраняются полнее. Как и у CS сжатые данные имеют меньшую “приведенную погрешность”, чем восстановленные и их выгоднее использовать при дальнейшем анализе сигнала. Обученный АЕ работает на порядок быстрее CS.

ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЕ

В качестве возможных способов подготовки сейсмических записей рассмотрим технику вейвлет-анализа и фильтрации [Стадник, 2004; Saad et al., 2019]. Вейвлет-преобразование позволяет перевести исследуемый сейсмический сигнал из временного представления в частотно-временное [Chen et al., 2019]. Выбор конкретного вида вейвлета зависит от класса решаемых задач. Для получения оптимальных алгоритмов преобразования разработаны определенные критерии, но они, как правило, неполно учитывают все разнообразие сейсмических сигналов [Комаров, 2010].

Вейвлет-преобразование представляет собой свертку анализируемого сигнала $x(t)$, $t \in (-\infty, \infty)$ с вейвлет-функциями $\psi\left(\frac{t-b}{a}\right)$, $a, b \in (-\infty, \infty)$, $a \neq 0$. В основе большинства применяемых компьютерных алгоритмов лежит непрерывное вейвлет-преобразование, комплексной функции $x(t)$ в комплексную функцию $w(a, b)$:

$$w(a, b) = \frac{1}{|a|^{1/2}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt, \quad (3)$$

где: $a, b \in R$; $a \neq 0$; $x(t)$ -сигнал; $\psi(t)$ – оконная функция, называемая вейлетом. Параметр a – масштаб (scale) – определяет размер вейвлет-функции $\psi\left(\frac{t-b}{a}\right)$. Его аналогом в Фурье-анализе является период гармонического колебания. Величина b – сдвиг по оси времени функции $\psi(t)$ – позволяет соотносить результат преобразования со шкалой времени $t \in (-\infty, \infty)$. Поскольку сейсмические данные $x(t)$ представляют собой временной ряд следующих друг за другом с постоянным шагом Δt значений: $x_k = x(t_k)$, $t_k = \Delta t k$, $k = 0, 1, \dots, N-1$, их дискретное вейвлет-преобразование задается следующим выражением:

$$W(a, b) = \frac{1}{n(a, b)} \sum_{k=0}^{N-1} x_k \psi\left(\frac{t_k - b}{a}\right), \quad (4)$$

где

$$n(a, b) = \sum_{k=0}^{N-1} e^{-\frac{1}{b} \left(\frac{t_k - b}{a}\right)^2}. \quad (5)$$

При переходе от функции (3) к (5) из знаменателя формулы (3) убирается множитель $|a|^{1/2}$ с его

заменой $\int_{-\infty}^{\infty} e^{-\frac{(t-b)^2}{a^2 B}} dt = a\sqrt{B\pi}$. Это устраняет зависимость амплитуд гармонических компонент от параметра a , что обычно мешает правильно оценить их относительные интенсивности по представлению вейвлет-спектров.

В качестве входа ГНС удобно использовать элементы локального спектра энергии или скалограммы, способной описывать распределение энергии по масштабам.

$$S(a, b) = |W_A(a, b)|^2. \quad (6)$$

Так как это распределение локализовано во времени с помощью параметра сдвига b , то (6) можно назвать локальной скалограммой (рис. 3). В нижней части скалограммы расположены коэффициенты с малыми номерами, дающие детальную картину сигнала, а сверху – с большими номерами, дающие огрубленную картину сигнала.

На рис. 3а хорошо видно усложнение спектра по краям, так называемые краевые разрывы. Нами разработан алгоритм, который реализует адаптивный выбор полос пропускания фильтра, позволяющий проводить обработку и фильтрацию входных цифровых сейсмических данных [Gravirou et al., 2014]. Можно выделить мешающую низкочастотную составляющую и всплески, ассоциируемые с техногенными помехами и удалить их из исходного процесса (см. рис. 3б) [Lin et al., 2013]. В ходе экспериментов были опробованы различные типы вейвлетов: Хаара, Симлета различных порядков, Мейера и другие. Мы остановились на вейвлетах Добеши четвертого порядка “db4”. При их использовании форма восстановленного сигнала наименее отличалась от оригинала.

Легко увидеть, что полученная информация избыточна. Обычно для дальнейшей обработки выбирают 5–8 наиболее характерных уровней разложения (которые имеют максимальную амплитуду, и про которые известно, что они не насквозь шумовые).

Спектральные полосы могут быть выбраны на основе спектра мощности [Altunkaynak, 2014]. Разложение по каждой полосе – своеобразная карта признаков сигнала [Hu, Wang, 2018]. Восстановление сигнала по характерным картам дает очищенный от шума сигнал. Восстановление по одному из порядков вейвлет-разложения дает возможность рассмотреть выделяемые этой картой признаки. На рис. 4 представлены исходная запись горизонтальной составляющей землетрясения, происшедшего 12.02.2005 г. у побережья Мексики на глубине 20 км и восстановленный по порядку № 64 этот же сигнал. Запись сейсмической станции АСАР Калифорнийского универси-

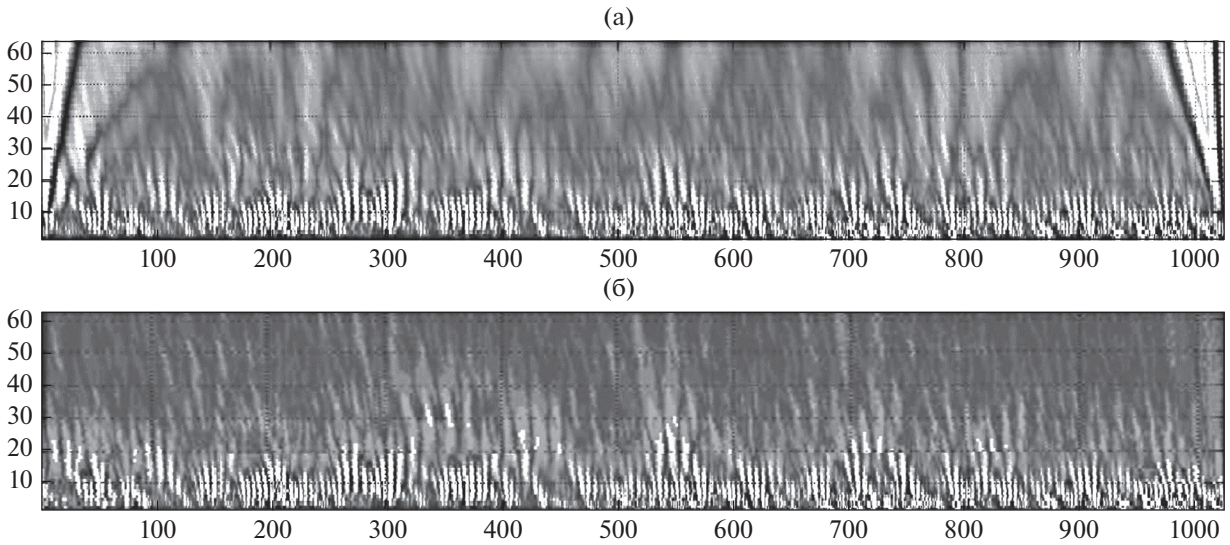


Рис. 3. Пример скалограммы тестового сейсмического сигнала с глубиной разложения 64: (а) – скалограмма окна сейсмической записи, длиной 1024 отсчета; (б) – обработанная скалограмма того же участка записи.

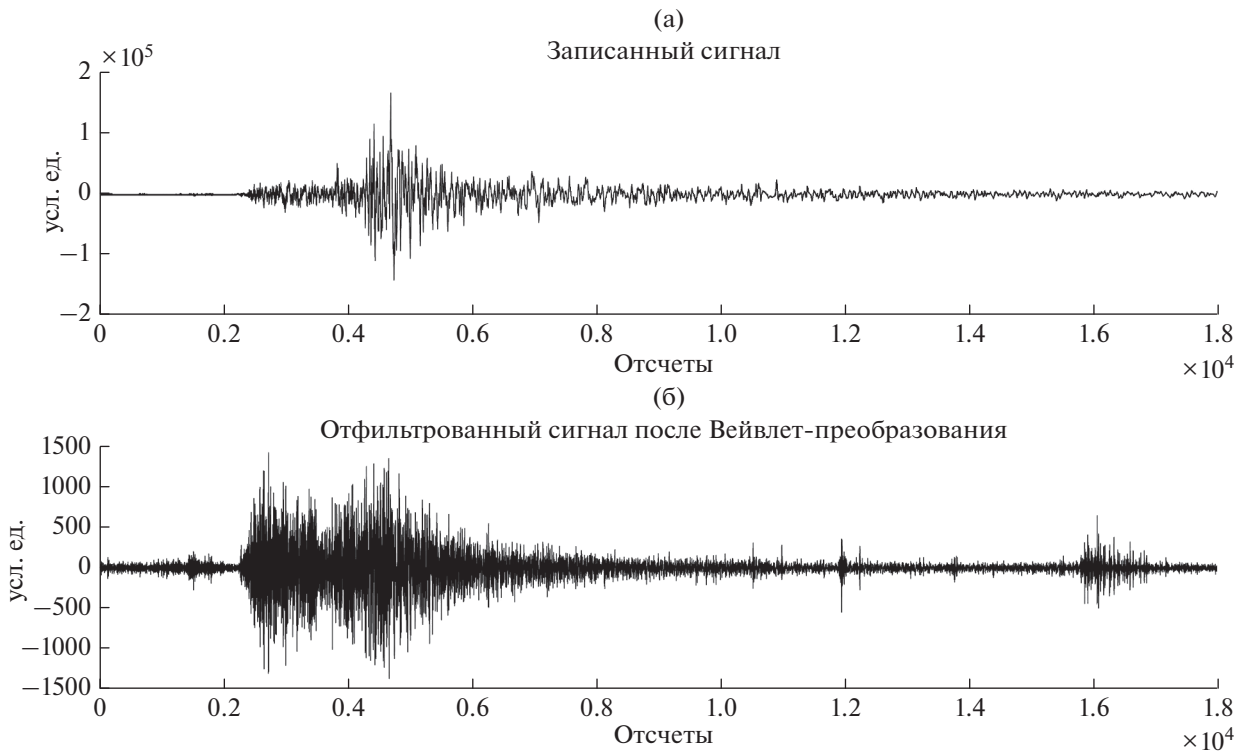


Рис. 4. (а) – запись землетрясения на сейсмической станции АСАР. Магнитуда землетрясения $M = 5.0$ (Оахака, Мексика, 12.02.2005 г., $M = 5.0$, глубина 20.0 км); (б) – тот же сигнал, восстановленный после вейвлет-преобразования по порядку № 64.

тета в Лос-Анджелесе, расстояние от эпицентра 2850 км.

В нашем примере при использовании восьми коэффициентов для трехкомпонентного сейсмометра, на вход ГНС мы будем подавать такой мас-

сив: $(1024 \text{ отсчетов}) \times (8 \text{ коэффициентов}) \times (3 \text{ компоненты сейсмометра}) = 24576 \text{ числовых значений}$. При использовании сверточного слоя в качестве входа ГНС это вполне реально. Учитывая, что признаки заранее разложены по коэффи-

циентам (и, естественно, по компонентам), это должно привести к выигрышу в сложности архитектуры сети и времени обучения. Использование вейвлетов также позволит отследить и подчеркнуть характерные места изменения спектральной структуры сигнала [Gravirov et al., 2013].

ОБСУЖДЕНИЕ

В этой работе рассматривались различные методы предварительной обработки сейсмических данных для моделей глубокого обучения. Эти методы преобразуют сегменты сейсмических записей в различные типы представлений, из которых ГНС могут извлекать необходимые для анализа характеристики сигнала. Мы исследовали наиболее перспективные методы, хотя возможны и другие подходы [Lv et al., 2017]. Сделать корректные заключения о преимуществе того или иного метода не представляется возможным в силу разнообразия сейсмологических задач, тем более что могут появиться и новые задачи. Разные методы подготовки преследуют несколько разные цели. Общие для них – выделение особенностей сигнала, снижение шума и сжатие данных – очень взаимосвязаны. Кодированная часть АЕ хорошо сжимает данные, но проигрывает в шумоподавлении. Вейвлет-подготовка увеличивает входные данные ГНС, но хорошо выделяет особенности сигнала, CS хорошо устраняет “симметричный” шум и хорошо сжимает данные, но требует дополнительного контроля сохранения информации.

Каждый метод подразумевает и возможность восстановления сейсмических записей, очищенных от шумов. Однако при этом добавляется ошибка восстановления, и увеличиваются вычислительные затраты. Восстановление лучше использовать только для визуализации и контроля, а дальнейшую работу проводить только с преобразованными данными. Вычисления могут производиться в низкоразмерном пространстве, дающем преимущества по времени вычислений. Преобразование должно выделять информацию, подсвечивать особенности данных, а не запутывать их, иначе только после распаковки данные будут пригодны для обработки. Удалению подлежат только шумовые составляющие записей: аппаратные шумы и техногенные помехи. Алгоритмы предобработки должны допускать возможность дальнейшей специализированной обработки полезного сигнала, нацеленной на решение разных сейсмологических задач.

ВЫВОДЫ

Проблемы сейсмологии обуславливают необходимость разработки новых гибких инструментов анализа, легко перестраиваемые на решение

новых задач. Внедрение нейросетевых технологий в практику сейсмических наблюдений может значимо увеличить их достоверность и информативность, автоматизировать процесс выделения и обработки сигналов сейсмических событий. Многие зависят от качества данных. Если данные содержат постороннюю и нерелевантную информацию, ГНС могут давать менее точные и менее понятные результаты. Поэтому предварительная обработка сейсмических записей является важным шагом в процессе машинного обучения.

Этап предварительной обработки необходим для устранения избыточности данных, определение их пригодности, решения проблем шума. Предобработка необходима не только для анализа сейсмических записей с помощью ГНС, но и для других методов анализа. Она не приводит к потере информации, а способствует увеличению количества полезной информации на единицу данных.

Какой же из рассмотренных методов предобработки наиболее перспективен?

CS – очень эффективный метод. Но его основные преимущества скажутся, когда сейсмические приемники будут осуществлять рандомные наблюдения. Это дело будущего. Тогда, проводя анализ сжатых данных (без восстановления) можно получить значительный эффект в снижении вычислительных затрат и упрощении алгоритмов.

Качественно обученный АЕ с расширенной версией шумоподавления может решать многие важные задачи сейсмологии. Однако он не обеспечивает эффективного подавления многообразных техногенных помех.

Нам кажется, что при современном уровне развития ГНС, предпочтительным методом предобработки сейсмических записей является вейвлет-преобразование. Помимо отличной способности этого метода фильтровать шумы разнообразной природы, это очень наглядный метод, позволяющий четко локализовать изменения сигнала во времени.

Целью данной работы было сравнение различных подходов к предварительной обработке данных для задач анализа сейсмических записей и предоставление справочных материалов для будущих исследований.

СПИСОК ЛИТЕРАТУРЫ

- Ганнибал А.Е. О возможности применения искусственных нейронных сетей в задачах сейсмического мониторинга // Сейсмические приборы. 2018. Т. 54. № 3. С. 5–21.
<https://doi.org/10.21455/si2018.3-1>
- Гравиров В.В., Кислов К.В. “DataCollector – Программа формирования выходных массивов данных для обучения нейронных сетей в системе Matlab”. Св-во гос. регистр. программ для ЭВМ № 2015611797. 2015.

- Гравиров В.В., Кислов К.В., Винберг Ф.Э.* Выделение информативного сигнала из нестационарных сильноза- шумленных сейсмических данных с использованием нейросетевых классификаторов // Пром. АСУ и контроллеры. 2012. № 12. С. 55–59.
- Граничин О.Н., Павленко Д.В.* Рандомизация получения данных и 11-оптимизация // Компьютерные инструменты в образовании. 2010. № 1. С. 4–14.
- Кислов К.В., Гравиров В.В.* Глубокие искусственные нейронные сети как инструмент анализа сейсмических данных // Сейсмические приборы. 2017. Т. 53. № 1. С. 17–28.
<https://doi.org/10.21455/si2017.1-2>
- Кислов К.В., Гравиров В.В.* Использование искусственных нейронных сетей в классификации зашумленных сейсмических сигналов // Сейсмические приборы. 2016. Т. 52. № 2. С. 46–64.
- Кислов К.В., Гравиров В.В.* Исследование влияния окружающей среды на шум широкополосной сейсмической аппаратуры // Вычислительная сейсмология. Вып. 42. М.: Красанд. 2013. 240 с. ISBN: 978-5-396-00626-3
- Кислов К.В., Гравиров В.В.* Распознавание вступления землетрясения на фоне техногенных шумов // Сейсмические приборы. 2010. Т. 46. № 2. С. 26–46.
- Комаров И.Э.* Выбор параметров дискретного вейвлет-преобразования для различных классов сигналов. Дис. ... канд. тех. наук. Омск: Омский гос. тех. ун-т. 2010. 128 с.
- Стадник А.В.* Использование искусственных нейронных сетей и вейвлет-анализа для повышения эффективности в задачах распознавания и классификации. Дис. ... канд. физ.-мат. наук. Иваново: Иван. гос. ун-т. 2004. 88 с.
- Чичеев М.А.* Быстрые алгоритмы дискретного косинусного преобразования. Самара: изд-во Самар. гос. аэрокосм. ун-та. 2006. 56 с. ISBN 5-7883-04-97-0
- Allen R.* Automatic phase pickers: their present use and future prospects // Bull. Seismol. Soc. Am. 1982. 72(6B) P. 225–242.
- Altunkaynak A.* Predicting Water Level Fluctuations in Lake Michigan-Huron Using Wavelet-Expert System Methods // Water Resources Management. 2014. V. 28(8). P. 2293–2314.
<https://doi.org/10.1007/s11269-014-0616-0>
- Bao W, Yue J, Rao Y.* A deep learning framework for financial time series using stacked autoencoders and long-short term memory // PLoS ONE. 2017. V. 12(7). e0180944, 16 p.
<https://doi.org/10.1371/journal.pone.0180944>
- Ben-Haim Z., Eldar Y.C., Elad M.* Coherence-based performance guarantees for estimating a sparse vector under random noise // IEEE Trans. Signal Process. 2010. V. 58(10). P. 5030–5043.
<https://doi.org/10.1109/TSP.2010.2052460>
- Blumensath T., Davies M.* Iterative hard thresholding for compressive sensing // Appl. Comput. Harmon. Anal. 2009. V. 27(3). P. 265–274. (<https://arxiv.org/pdf/0805.0510.pdf>).
- Bourgain J., Dilworth S., Ford K., Konyagin S., Kutzarova D.* Explicit constructions of rip matrices and related problems // Duke Math. Journal. 2011. V. 159. P. 145–185. (arXiv: 1008.4535v3).
- Candès E., Romberg J., Tao T.* Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information // IEEE Trans. Inform. Theory. 2006. V. 52(2). P. 489–509.
<https://doi.org/10.1109/TIT.2005.862083>
- Candès E.J., Wakin M.B.* An Introduction to Compressive Sampling [A sensing/sampling paradigm that goes against the common knowledge in data acquisition] // IEEE Signal Processing Magazine. 2008. V. 25(2). P. 21–30.
<https://doi.org/10.1109/MSP.2007.914731>
- Chen X., Kopsaftopoulos F., Wu Q., Ren H., Chang F.-K.* A Self-Adaptive 1D Convolutional Neural Network for Flight-State Identification // Sensors. 2019. V. 19. pii: E275.
<https://doi.org/10.3390/s19020275>
- Claerbout J.F., Muir F.* Robust modeling with erratic data // Geophysics. 1973. V. 38(5). P. 826–844.
- Creswell A., Bharath A.A.* Denoising Adversarial Autoencoders // IEEE Trans. Neural Netw. Learn. Syst. 2019. V. 30(4). P. 968–984.
<https://doi.org/10.1109/TNNLS.2018.2852738>
- Dai H., MacBeth C.* The application of back propagation neural network to automatic picking seismic arrivals from single component recordings // J. Geophys. Res.-Sol. Ea. 1997. V. 102(B7). P. 15105–15113.
- Davenport M.A., Boufounos P.T., Wakin M.B., Baraniuk R.G.* Signal Processing With Compressive Measurements // IEEE Journal of Selected Topics in Signal Processing. 2010. V. 4(2). P. 445–460.
<https://doi.org/10.1109/JSTSP.2009.2039178>
- Davenport M.A., Laska J.N., Treichler J.R., Baraniuk R.G.* The pros and cons of compressive sensing for wideband signal acquisition: noise folding versus dynamic range // IEEE Trans. Signal Process. 2012. 60(9). P. 4628–4642.
<https://doi.org/10.1109/TSP.2012.2201149>
- Donoho D.* Compressed sensing // IEEE Trans. Inform. Theory. 2006. V. 52(4). P. 1289–1306.
<https://doi.org/10.1109/TIT.2006.871582>
- Erhan D., Bengio Y., Courville A., Manzagol P.-A., Vincent P., Bengio S.* Why Does Unsupervised Pre-training Help Deep Learning? // J. Machine Learning Research. 2010. V. 11. P. 625–660.
- Foucart S., Rauhut H.* A mathematical introduction to compressive sensing. N.Y. Springer. 2013. xviii + 625 p., ISBN 978-0-8176-4948-7
- Gan S., Wang S., Chen Y., Chen X., Huang W., Chen H.* Compressive sensing for seismic data reconstruction via fast projection onto convex sets based on seislet transform // J. Applied Geophysics. 2016. V. 130. P. 194–208.
<https://doi.org/10.1016/j.jappgeo.2016.03.033>
- Gehring J., Miao Y., Metze F., Waibel A.* Extracting deep bottleneck features using stacked auto-encoders / Proceedings of the ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing. 2013. Vancouver, BC, Canada.
<https://doi.org/10.1109/ICASSP.2013.6638284>
- Gentili S., Michelini A.* Automatic picking of P and S phases using a neural tree // J. Seismol. 2006. V. 10(1). P. 39–63.
<https://doi.org/10.1007/s10950-006-2296-6>
- Gravirou V.V., Kislov K.V.* The seismic data preparation program for training of an artificial neural network for ultrashort earthquake warning system. Book of Abstracts, 10th International Conference “Problems of Geocosmos”, St. Petersburg, 2014. P. 86–87. (http://geo.phys.spbu.ru/geocosmos/book_of_abstracts.pdf).

- Gravurov V.V., Kislov K.V., Gravurova L., Vinberg F.* The Use of Wavelet Transformation Techniques in Structure of an Artificial Neural Network for Recognition of Early Arrival of Earthquakes on Strongly Noisy Seismic Records. Book of Abstracts CTBT: Science and Technology. Vienna, Austria. 2013. T3-P132, 2013. P. 139. (http://www.ctbto.org/fileadmin/user_upload/SnT2013/bookofabstracts.pdf).
- Gravurov V.V., Kislov K.V., Vinberg F.E.* Wavelet transform as a tool for processing and analysis of seismograms // Proceedings of the 10th International Conference Problems of Geocosmos – 2014, St. Petersburg State University. St. Petersburg, Petrodvorets. 2014. P. 168–172. http://geo.phys.spbu.ru/materials_of_a_conference_2014/S2014/30_Gravurov.pdf
- Gurbuz A.C., McClellan J.H., Scott W.R.* A compressive sensing data acquisition and imaging method for stepped frequency GPRs // IEEE Trans. Signal Process. 2009. V. 57(7). P. 2640–2650. <https://doi.org/10.1109/TSP.2009.2016270>
- Guyon I., Elisseeff A.* An Introduction to Variable and Feature Selection // JMLR Special Issue on Variable and Feature Selection. 2003. № 3. P. 1157–1182. (<http://www.jmlr.org/papers/volume3/guyon03a/guyon03a.pdf>).
- Haupt J., Nowak R.* Signal reconstruction from noisy random projections // IEEE Trans. Inform. Theory. 2006. V. 52(9). P. 4036–4048. <https://doi.org/10.1109/TIT.2006.880031>
- Herrmann F.J.* Compressive Sensing. Seismic Laboratory for Imaging and Modeling. <https://www.slim.eos.ubc.ca/research/compressive-sensing> (последний просмотр. 01.04. 2019).
- Herrmann F.J., Wang D., Hennenfent G., Moghaddam P.* Curvelet-based seismic data processing: A multiscale and nonlinear approach // Geophysics. 2008. V. 73(1). 5 p. <https://doi.org/10.1190/1.2799517>
- Hinton G., Salakhutdinov R.* Reducing the dimensionality of data with neural networks // Science. 2006. V. 313. P. 504–507. <https://doi.org/10.1126/science.1127647>
- Hu R., Wang Y.* A first arrival detection method for low SNR microseismic signal // Acta Geophysica. 2018. V. 66(5). P. 945–957. <https://doi.org/10.1007/s11600-018-0193-3>
- Ji S., Xue Y., Carin L.* Bayesian compressive sensing // IEEE Trans. Signal Process. 2008. V. 56(6). P. 2346–2356. <https://doi.org/10.1109/TSP.2007.914345>
- Kislov K.V., Gravurov V.V.* Deep Artificial Neural Networks as a Tool for the Analysis of Seismic Data // Seismic Instruments. 2018. V. 54(1). P. 8–16. <https://doi.org/10.3103/S0747923918010073>
- Kislov K.V., Gravurov V.V.* The use of artificial neural networks for classification of noisy seismic signal // Seismic Instruments. 2017. V. 53(1). P. 87–101. <https://doi.org/10.3103/S0747923917010054>
- Kislov K.V., Gravurov V.V.* Neural network techniques for earthquake detection in high noise // Исследовано в России. 2010. № 070e. С. 837–848.
- Kong Q., Trugman D.T., Ross Z.E., Bianco M.J., Meade B.J., Gerstoft P.* Machine Learning in Seismology: Turning Data into Insights // Seismological Research Letters. 2018. 90(1). P. 3–14. <https://doi.org/10.1785/0220180259>
- Kotsiantis S.B., Kanellopoulos D., Pintelas P.E.* Data Preprocessing for Supervised Learning // International J. Computer Science. 2006. V. 1(1). ISSN 1306-4428. P. 111–117. (<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.104.8413&rep=rep1&type=pdf>).
- Li Y., Song B., Kang X., Du X., Guizani M.* Vehicle-Type Detection Based on Compressed Sensing and Deep Learning in Vehicular Networks // Sensors. 2018. V. 18(12). pii: E4500. <https://doi.org/10.3390/s18124500>
- Lin H., Li Y., Yang B., Ma H.* Random denoising and signal nonlinearity approach by time-frequency peak filtering using weighted frequency reassignment // Geophysics. 2013. V. 78(6). P. V229–V237. <https://doi.org/10.1190/geo2012-0432.1>
- Lv L., Zhao D., Deng Q.* A Semi-Supervised Predictive Sparse Decomposition Based on Task-Driven Dictionary Learning // Cognitive Computation. 2017. V. 9(1). P. 115–124. (<https://doi.org/>). <https://doi.org/10.1007/s12559-016-9438-0>
- Madureira G., Ruano A.E.* A Neural Network Seismic Detector // Acta Technica Jaurinensis. 2009. V. 2(2). P. 159–170. <https://doi.org/10.3182/20090921-3-TR-3005.00054>
- Majumdar A., Tripathi A.* Asymmetric stacked autoencoder. International Joint Conference on Neural Networks (IJCNN). 2017. <https://doi.org/10.1109/IJCNN.2017.7965949>
- Meyer M., Weber S., Beutel J., Thiele L.* Systematic identification of external influences in multi-year microseismic recordings using convolutional neural networks // Earth Surf. Dynam. 2019. V. 7. P. 171–190. <https://doi.org/10.5194/esurf-7-171-2019>
- Oldenburg D.W., Scheuer T., Levy S.* Recovery of the acoustic impedance from reflection seismograms // Geophysics. 1983. V. 48(10). P. 1318–1337. <https://doi.org/10.1190/1.1441413>
- Perol T., Gharbi M., Denolle M.* Convolutional Neural Network for Earthquake Detection and Location // Science Advances. 2018. V. 4(2). E1700578. 8 p. <https://doi.org/10.1126/sciadv.1700578>
- Robucci R., Gray J.D., Chiu L.K., Romberg J., Hasler P.* Compressive sensing on a CMOS separable-transform image sensor // Proc. IEEE. 2010. V. 98(6). P. 1089–1101. <https://doi.org/10.1109/JPROC.2010.2041422>
- Ross Z.E., Meier M.-A., Hauksson E.* P Wave Arrival Picking and First-Motion Polarity Determination With Deep Learning // J. Geophys. Res.-Sol. Ea. 2018. V. 123. P. 5120–5129. <https://doi.org/10.1029/2017JB015251>
- Saad O.M., Shalaby A., Sayed M.S.* Automatic discrimination of earthquakes and quarry blasts using wavelet filter bank and support vector machine // J. Seismol. 2019. V. 23(2). P. 357–371. <https://doi.org/10.1007/s10950-018-9810-5>
- Santosa F., Symes W.W.* Linear inversion of band-limited reflection seismograms // SIAM J. Sci. Statist. Comput. 1986. V. 7(4). P. 1307–1330. <https://doi.org/10.1137/0907087>
- Shen Y., Han T., Yang Q., Yang X., Wang Y., Li F., Wen H.* CS-CNN: Enabling Robust and Efficient Convolutional Neural Networks Inference for Internet-of-Things Applications // IEEE Access, Special section on multimedia analysis for internet-of-things. 2018. № 6. P. 13439–13448. <https://doi.org/10.1109/ACCESS.2018.2810264>

- Sheng G., Li Z., Wang W., Lan G.* A new automatic detection method of microseismic events based on wavelet decomposition and high-order statistics // *Geophys. Prospect. Petrol.* 2015. V. 54(4). P. 388–395.
- Shi Y., Lei M., Ma R., Niu L.* Learning Robust Auto-Encoders With Regularizer for Linearity and Sparsity // *IEEE Access.* 2019. № 7. P. 17195–17206.
<https://doi.org/10.1109/ACCESS.2019.2895884>
- Taylor H.L., Banks S.C., McCoy J.F.* Deconvolution with the l1 norm // *Geophysics.* 1979. V. 44(1). P. 39–52.
<https://doi.org/10.1190/1.1440921>
- Valentine A.P., Trampert J.* Data space reduction, quality assessment and searching of seismograms: autoencoder networks for waveform data // *Geophys. J. Int.* 2012. V. 189(2). P. 1183–1202.
<https://doi.org/10.1111/j.1365-246X.2012.05429.x>
- Vera Rodriguez I., Sacchi M., Gu Y.* Simultaneous recovery of origin time, hypocentre location and seismic moment tensor using sparse representation theory // *Geophys. J. Int.* 2012. V. 188. P. 1188–1202.
<https://doi.org/10.1111/j.1365-246X.2011.05323.x>
- Wang J., Teng T.-L.* Artificial neural network-based seismic detector // *Bull. Seismol. Soc. Am.* 1995. V. 85. P. 308–319.
- Yao H., Gerstoft P., Shearer P.M., Mecklenbräuker C.* Compressive sensing of the Tohoku-Oki Mw 9.0 earthquake: Frequency-dependent rupture modes // *Geophys. Res. Lett.* V. 38. L20310.
<https://doi.org/10.1029/2011GL049223>
- Yoon C.E., O'Reilly O., Bergen K.J., Beroza G.C.* Earthquake detection through computationally efficient similarity search // *Science Advances.* 2015. V. 1(11). E1501057. 14 p.
<https://doi.org/10.1126/sciadv.1501057>
- Yuan S, Liu J, Wang S, Wang T., Shi P.* Seismic waveform classification and first-break picking using convolution neural networks // *IEEE Geosci. Remote Sens. Lett.* 2018. V. 15(2). P. 272–276. ISSN 1545-598X.
<https://doi.org/10.1109/LGRS.2017.2785834>
- Zhang L., Zhang L., Du B.* Deep learning for remote sensing data: A technical tutorial on the state of the art // *IEEE Geosci. Remote Sens. Mag.* 2016. V. 4(2). P. 22–40.
<https://doi.org/10.1109/MGRS.2016.2540798>
- Zhao Y., Takano K.* An artificial neural network approach for broadband seismic phase picking // *Bull. Seismol. Soc. Am.* 1999. V. 89(3). P. 670–680.
- Zheng X., Wang M., Ordieres-Meré J.* Comparison of Data Preprocessing Approaches for Applying Deep Learning to Human Activity Recognition in the Context of Industry 4.0 // *Sensors.* 2018. V. 18(7). 2146. 13 p.
<https://doi.org/10.3390/s18072146>

CAPABILITIES OF SEISMIC DATA PREPARATION FOR ANALYSIS OF DEEP NEURAL NETWORK

K. V. Kislov^{a,*}, V. V. Gravirov^{a,b}, and F. E. Vinberg^a

^a*Institute of Earthquake Prediction Theory and Mathematical Geophysics of the Russian Academy of Sciences (IEPT RAS), Moscow, Russia*

^b*The Schmidt Institute of Physics of the Earth of the Russian Academy of Sciences (IPE RAS), Moscow, Russia*

^{*}*E-mail: kvkislov@yandex.ru*

Algorithms of automatic processing of seismic recordings are constantly being improved, data analysis tasks are becoming more complex. Most algorithms require preliminary data preparation. This preparation is either very simple, such as frequency filtering, or highly specialized, highlighting the specific features of the signal. An adequate preprocessing can increase efficiency of the further analysis by orders. However, specific preprocessing cannot be used for other tasks or with other postprocessing algorithms. We consider the decisions that do not result in the significant loss of information and that can be used with any tasks. The main objectives of the preprocessing are noise reduction, elimination of anthropogenic noise and the reduction of dimensionality of data. We assume that deep neural networks are used for further data processing. Nevertheless, that should not exclude the possibility of other algorithms application. In this paper, we consider wavelet transform, autoencoder, and some other algorithms as preprocessing of seismic data.

Keywords: seismic data analysis, deep neural networks, preprocessing.