УДК 007.057
На правах рукописи

Као Динь Чонг

ИССЛЕДОВАНИЕ И ПРИМЕНЕНИЕ НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ В ЗАДАЧЕ ПРОГНОЗИРОВАНИЯ ЗЕМЛЕТРЯСЕНИЙ

(На примере северо-западного района Вьетнама)

Специальность 05.13.01 - Системный анализ,
управление и обработка информации

Автореферат диссертации на соискание учёной степени
кандидата технических наук

Москва 2012

Работа выполнена в Российском университете дружбы народов.

Научный руководитель: доктор технических наук, профессор
Пупков Константин Александрович

Официальные оппоненты: доктор технических наук, профессор
Дивеев Асхат Ибрагимович

кандидат технических наук, доцент
Тюхов Борис Петрович

Ведущая организация: Институт физики Земли им. О.Ю. Шмидта
(ИФЗ РАН)

Защита диссертации состоится «___» _____ 2013 г. в _____ часов на заседании диссертационного совета Д212.141.02 по специальности 05.13.01 Системный анализ, управление и обработка информации (в технических системах) при Московском государственном техническом университете им. Н.Э. Баумана, по адресу 105005, г. Москва, 2-я Бауманская ул., д. 5.

С диссертацией можно ознакомиться в библиотеке МГТУ им. Н.Э. Бау

Автореферат разослан « ____ » _____ 2012 г.

НТБ МГТУ им. Н.Э. Баумана
2015674
Као Динь ЧонгИсследование

Ученый секретарь
диссертационного совета,
к.т.н., доцент Муратов И.В.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Работа посвящена применению нейросетевых технологий в задаче прогнозирования землетрясений.

В последние годы увеличивается число катастрофических землетрясений, приводящих к колоссальным жертвам и разрушениям, в том числе землетрясение в Италии (2009 г. с магнитудой 6.7), землетрясение в Гаити (2010 г. с магнитудой 7), и землетрясение в Японии (2011 г. с магнитудой 8.9). Поэтому прогнозирование места, времени и мощности землетрясений ещё остаётся одной из наиболее актуальных проблем. Для решения этой проблемы в индустриально развитых странах проводятся различные исследования для совершенствования методов прогноза сейсмических опасностей. Такие методы можно, в принципе, подразделить на инструментальные и интеллектуальные (или аналитические). В диссертации рассмотрена возможность использования одного из методов второй группы.

В таких странах, как Россия, США, Япония или Китай, где сейсмическая характеристика высока, научным исследованиям по прогнозу землетрясений уделяется большое внимание. Появление новых методов прогноза землетрясений у этих стран, как КОЗ у России, М8; КН принадлежат США, научных достижений в сейсмологии очагов у Китая, или новых структур зданий в Японии показывают важность этого научного направления.

Во Вьетнаме в силу экономических и ранее военно-политических обстоятельств долгое время прогнозу землетрясений уделялось относительно малое внимание, хотя здесь происходили землетрясения с магнитудой более 6,5, например, Диен Биен в 1935 г. с 6,8 магнитудой, Туан Зао 1983 года с магнитудой 6,7 и ряд крупных древних землетрясений. Число землетрясений с магнитудой больше 4 особенно увеличилось в Северо-Западном районе Вьетнама в последние годы. Последнее может являться признаком, возможно, крупных землетрясений в ближайшем будущем, вызывает особую тревогу в связи со строящимися гидроэлектростанциями и обсуждаемым вопросом о строительстве атомной электростанции.

Поэтому проблема получения в достаточной мере точных прогнозов землетрясений по времени, месту и мощности является весьма актуальной. Такие современные методы прогноза землетрясений как КН или М8также применены для районов Вьетнама. Получены некоторые новые результаты, но они являются только начальными. Кроме новых методов прогнозирования, важным является исследование входных данных (каталога землетрясений), что является одной из главных задач по прогнозу землетрясений. Применение новых мировых методов прогноза землетрясений, возможно для стран с достаточно высокой сейсмической активностью, имеющих полные каталоги землетрясений, таких как Россия или США. Для районов Вьетнама, где каталог землетрясений является не полным и не достаточным, эти методы дают неточные научные знания и прогнозы.

Одним из направлений улучшения результатов прогноза землетрясений является применение нейросетевых технологий. Поэтому в работе ставится

МГТУ им. Н.Э. Баумана Библиотека

задача исследования и применения таких технологий при использовании каталожных сейсмологических данных для обучения нейронных сетей и прогноза времени, места и мощности события.

Предметом диссертационного исследования является разработка и применение метода прогноза землетрясений, основанного на использовании нейронных сетей, обученных с помощью сейсмических данных каталога землетрясений Вьетнама.

Целью диссертационной работы является применение нейросетевых технологий в задаче прогноза землетрясений и оценка эффективности прогноза.

Для достижения поставленной цели были решены следующие **задачи:**

1. анализ состояния проблемы прогноза землетрясений во Вьетнаме;
2. сравнительный анализ методов прогноза землетрясений и оценка их эффективности;
3. анализ каталога землетрясений во Вьетнаме, выбор достоверного интервала времени и магнитуды для использования в задаче прогноза;
4. разработка и выбор нейронной сети для прогнозирования землетрясений на базе сейсмических данных, полученных из каталога землетрясений для обучения сети;
5. определение с помощью выбранного метода зависимости силы землетрясений от временных интервалов между их проявлениями;
6. использование нейронной сети для прогноза землетрясения и оценка вероятности его появления;
7. рекомендация по предлагаемому методу.

Методы исследования, используемые в диссертации, основываются на системном анализе исходных данных, математической статистике, компьютерных, нейросетевых технологиях, и методах программирования.

Новизна научных результатов, полученных в диссертации, заключается в 1) системном анализе многофакторных данных (геологических, сейсмических), позволяющем на основе функции Гуттенберга-Рихтера определить сейсмическую активность района, 2) разработке алгоритма обработки данных каталога землетрясений, автоматизации процесса обработки данных, программной реализации алгоритма, 3) сравнительном анализе различных типов нейронных сетей и обоснованном выборе наиболее эффективной сети для располагаемого каталога землетрясений, 4) обучении нейронной сети, 5) среднесрочном прогнозе землетрясений (на примере северо-западного района Вьетнама) и оценке вероятности происхождения событий.

Практическая значимость результатов работы заключается в 1) применении метода нейронной сети для прогноза сильных землетрясений во Вьетнаме, 2) сейсмическом районировании территории страны на основе полученных результатов, 3) разработке программы обработки входных данных для метода нейронной сети в задаче прогноза землетрясений и результаты прогноза.

Использование результатов работы. Результаты работы переданы Институту геофизики Вьетнамской Академии Наук и Технологий и органам,

ответственным за сейсмическую безопасность во Вьетнаме, а также внедрены в учебный процесс РУДН по дисциплине "Искусственные нейронные сети в управлении".

Апробация результатов, результаты диссертации подтверждаются докладами на следующих международных научных конференциях и симпозиумах: XXXIV научно-технической конференции преподавателей, сотрудников и аспирантов инженерного факультета РУДН (Москва, 2010), XXXV научно-технической конференции преподавателей, сотрудников и аспирантов инженерного факультета, РУДН (Москва, 2011), VМеждународной научно-практической конференции «Инженерные системы» (Москва, 2012), Международном симпозиуме «Интеллектуальные системы» (INTELS'2012) (Вологда, 2012), Международной конференции «9th General Assembly of Asian Seismological Commission» (Улан-Батор, Монголия, 2012).

Основные положения, выносимые на защиту:

1. Системный анализ многофакторных данных (геологических, сейсмических), позволяющий на основе функции Гуттенберга-Рихтера определить сейсмическую активность района;
2. Алгоритм обработки данных каталога землетрясений, его автоматизация и программная реализация;
3. Выбор и обоснование на основе сравнительного анализа различных типов нейронных сетей наиболее эффективной сети для располагаемого каталога землетрясений;
4. Методика обучения нейронной сети в задаче прогноза землетрясений;
5. Среднесрочный прогноз землетрясений и оценка вероятности происхождения события (на примере северо-западного района Вьетнама).

Результаты диссертации опубликованы в 12 научных трудах.

Структура диссертации. Диссертация состоит из введения, четырёх глав, заключения, списка литературы и приложений. Объем работы – 167 страниц, включая 40 рисунков и 23 таблиц. Список литературы содержит 105 наименований.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы, сформулированы предмет, цель и задачи исследования, методы исследования, новизна научных результатов и практическая значимость полученных результатов, основные положения, выносимые на защиту, приведены данные о структуре и объёме диссертации.

В первой главе рассмотрена конкретно нейронная технология, как одно из главных научных направлений развития 21_ого века, её применение на практике, особенно в области прогноза землетрясений, и основные результаты по прогнозу землетрясений во Вьетнаме.

В задаче прогноза землетрясений нейронные сети используются как для прогнозирования, так и для распознавания образов. Первая попытка применения метода нейронной сети (FBP (Feedforward Back Propagation) сетей) в задаче прогноза землетрясений была предпринята в 1994 году и после этого

такие нейронные сети, как RNN (Recurrent Neural Network) и RBF (Radial Basic Function) тоже использованы в этой задаче. Метод нейронной сети в задаче прогноза землетрясений оказался эффективным во многих странах, но не был использован в России и во Вьетнаме.

В настоящее время во Вьетнаме имеется сеть мониторинга, состоящая из 24 станций.

Можно выделить два направления таких исследований: 1) возможность максимальных землетрясений в конкретных участках и 2) определение времени и магнитуды возможных будущих землетрясений. Рассмотрены такие метода, применимые по первому направлению, как соотношение между силой землетрясений и длинами очагов, выявление закономерностей сейсмичности на основе сочетания геологических и сейсмических данных, оценка сейсмичности по особенностям поверхности Земли или метода Гумбел. Только 2 метода – КН и М8 применены по второму направлению исследований. Такие методы требуют учёта большого числа землетрясений, поэтому прогнозы, которые могут быть получены на их основе, по сравнительно ограниченному каталогу землетрясений Вьетнама не могут считаться достаточно точными.

Во второй главе приведены результаты, полученные автором по выбору нейронной сети и обработке входных данных для решения задачи прогноза землетрясений.

В задаче прогноза землетрясений интервал времени имеет главное значение. Он характеризуется сейсмической активностью каждого района. Существуют 3 типа прогноза землетрясений, краткосрочный, среднесрочный и долгосрочный прогноз. Целью этой работы является среднесрочный прогноз землетрясений на основе сейсмических данных, связанных со временем явления землетрясений. Поэтому автор предложил принцип выбора интервала времени прогноза для среднесрочного прогноза землетрясений методом нейронной сети, который состоит из 2 шагов:

Шаг 1: выбор по каталогу землетрясений подинтервала временного интервала 1 – 10 лет по годам. Количество землетрясений в этом подинтервале должно быть больше или равно двум.

Шаг 2: вычисление ошибки обучения нейронной сети.

На первом шаге рассматривается каталог землетрясений конкретного района. В каждом интервале времени должна быть характеристика сейсмичности, поэтому количество землетрясений должно больше 2 по интервалу. На втором шаге метод нейронной сети используется как методом оценки точности разделения по интервалу времени. Если выбор правильный, то сети дали точные результаты (точности по обучению сети). Ошибки обучения нейронной сети вычисляются по параметрам R (Regression), MSE (Mean Squared Error) или RMSE (Root Mean Squared Error).

На основе опубликованных данных автором предложена методика выбора нейронной сети в задаче прогноза землетрясений, и определены параметры оценки нейронной сети.

Эта процедура включала следующие шаги:

Шаг 1: выбор возможных нейронных сетей по решению задачи прогноза землетрясений конкретного района;

Шаг 2: определение для каждой сети, полученной по шагу 1, лучшей структуры для решения этой задачи по ошибке обучения по одному из параметров: R (Regression), MSE (Mean Squared Error) или RMSE (Root Mean Squared Error);

Шаг 3: сравнение полученных сетей по шагу 2 по одному из следующих двух принципов (процесс тестирования):

3.1. использование параметров максимальной и средней ошибки обучения при ограниченных данных.

3.2. использование таких параметров, как Rscore или Frequency bias (FB). Значение Rscore и FB вычисляются по формулам:

$$Rscore = \frac{N_{pc}}{N_{pc} + N_{pi}} - \frac{N_{pi}}{N_{pi} + N_{nc}} \text{ и } FB = \frac{N_{pc} + N_{pi}}{N_{pc} + N_{ni}}.$$

где, N_{pc} (predicted-correct) количество тестов, результаты которых по сети совпадают с реальными результатами (примерами); N_{pi} (predicted-incorrect) количество тестов, у которых были результаты по сети, но не были в реальности; N_{nc} (not predicted-correct) количество тестов, у которых не были результаты по сети, и не были в реальности; N_{ni} (not predicted-incorrect) количество тестов, у которых не были результаты по сети, но были в реальности.

До сих пор не определено, каким методом найти число нейронов в скрытном слое. Это число вычисляется для конкретной задачи и по ошибкам обучения сети. По этому процессу есть некоторые замечания. Если число нейронов велико, то процесс обучения замедляется и время обучения тоже долгое. Нахождение структуры нейронной сети является важной задачей. В этой методике шаг 2 и 3 отвечают за это вопрос.

Процесс выбора состоит из двух периодов: ошибки обучения (шаг 2) и ошибки прогноза (тестирования) (шаг 3). Для обучения нейронной сети имеется набор примеров, и эти примеры поступают в сеть по двум процессам: процессу обучения и процессу тестирования.

В нейронные сети по процессу обучения поступают входные данные, и выходные значения вычисляется сетями. На основе различия между этими выходными данными и выходными примерами оценивается качество обученной сети. Эти данные могут оцениваться многими методами с разными показателями (шаг 2), но в этой методике пользуется линейная регрессия с качеством показателей R, MSE или RMSE.

Значение Rscore принадлежит интервалу (-1;1) и является различием между выходными данными сети и реальными значениями обучения. Чем ближе Rscore к 1, тем точнее результаты прогноза. Значение FB принадлежит интервалу $(0,\infty)$ и тоже является различием между числом правильного прогноза и числом прогноза в реальности.

Полнота (достаточность) каталога землетрясений конкретного района во многих зависит от системы сейсмологических станций у этой страны, поэтому каталог землетрясений каждой страны отличается от каталогов других стран. Поэтому обработку каталога землетрясений (данных) необходимо проводить до процесса прогноза. Нахождение интервала времени, на котором количество землетрясения полно зафиксированы, определяется магнитуда с минимальным доверием, которая характеризует мощность землетрясений, что является основной целью обработки каталога землетрясений. Алгоритм этой обработки представлен на рис. 1, и он создан на основе нахождения функции Гуттенберга-Рихтера, который имеет вид $\log(N/T) = a - bM$. N - количество землетрясений во временном интервале исследования T, и необходимо найти значения коэффициентов a, b.

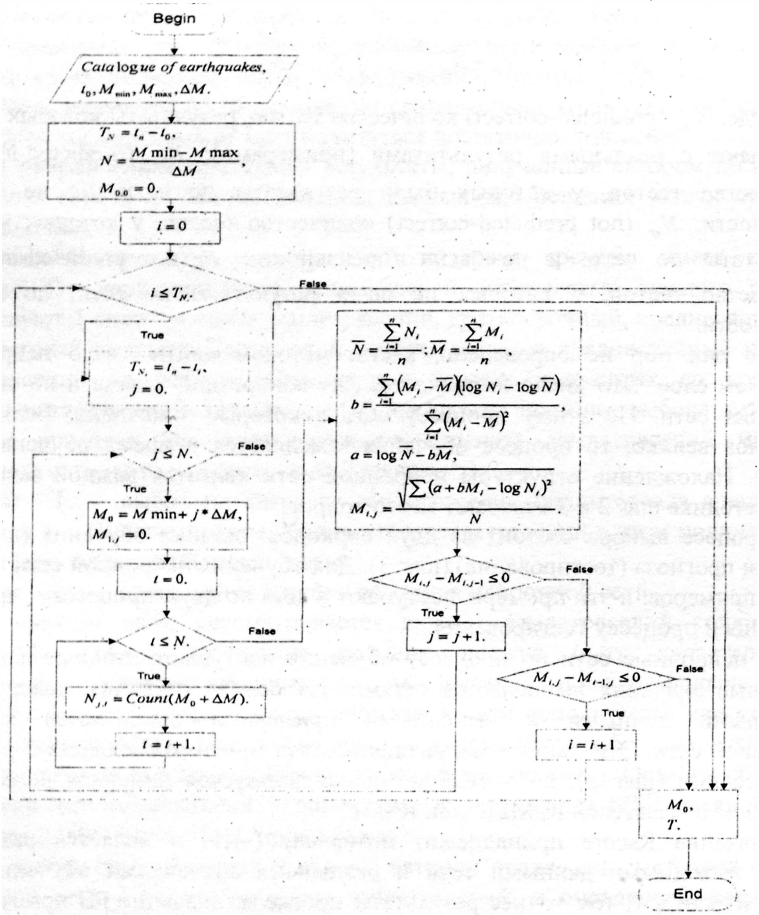


Рис. 1. Блок-схема нахождения интервала времени и магнитуды с минимальным доверием по функции Гуттенберга-Рихтера

Более конкретно надо найти следующие значения:

- магнитуду с минимальным доверием каталога землетрясений (M_0);

- интервал доверенного времени каталога землетрясений (T_N).

Это необходимо, чтобы определить функцию Гуттенберга-Рихтера, характеризующую активность этого района (количество землетрясений).

Разработан алгоритм обработки каталога землетрясений на основе определения минимальной ошибки нахождения функции Гуттенберга-Рихтера. Отсюда можно получить интервал доверенного времени, магнитуду с минимальным доверием каталога землетрясений. Входными данными алгоритма являются: t_0 - самый ранний момент времени по каталогу землетрясений; t_n - последний момент времени по каталогу землетрясений; $T_N = t_n - t_0$ - интервал времени каталога землетрясений;

ΔM - шаг разделения интервала магнитуды и определяется для того, чтобы деление $\frac{M\max - M\min}{\Delta M}$ являлось чётным числом, где $M_{\min}$, $M_{\max}$ соответственно минимальная и максимальная мощность землетрясений по каталогу.

Алгоритм обработки каталога землетрясений применяется для каталога землетрясений северо-западному району Вьетнама.

Таблица 1.

Значения среднеквадратичной минимальной ошибки (MSE) по вычислению функции Гуттенберга-Рихтера по интервалу магнитуд

Время t_0 / Интервал магнитуда	1960	1961	...	1974	1975	1976	1977
3.1-6.5	0.0109	0.0102	...	0.0149	0.0175	0.0174	0.0160
3.2-6.5	0.0084	0.0082	...	0.0125	0.0146	0.0146	0.0134
3.3-6.5	0.0081	0.0080	...	0.0122	0.0143	0.0142	0.0130
3.4-6.5	0.0071	0.0071	...	0.0110	0.0130	0.0130	0.0117
3.5-6.5	0.0051	0.0050	...	0.0085	0.0104	0.0104	0.0094
3.6-6.5	0.0060	0.0061	...	0.0098	0.0117	0.0118	0.0104

Из табл. 1 видно, что магнитуда с минимальным доверием каталога землетрясений северо-западного района Вьетнама (M_0) равна 3.5.

Интервал времени, на котором полное количество землетрясений начинается с 1976_ого года, близок к полученному интервалу в данной работе. Начиная с 1976 года, система сейсмических станций во Вьетнаме зарегистрировала магнитуды землетрясений больше 3.0. Функция Гуттенберга-Рихтера для этого района по этому алгоритму имеет вид $\log(N/T) = 3{,}76 - 0{,}89 * M$

Плотность распределения частоты землетрясений по северо-западному району Вьетнама по годам представляется графиками, которые используются для проверки точности алгоритма (рис. 2, 3).

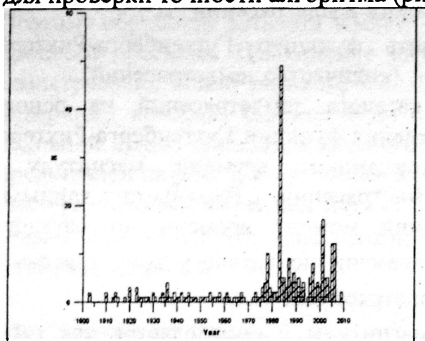


Рис. 2. Плотность распределения землетрясений по северо-западному району Вьетнама по годам

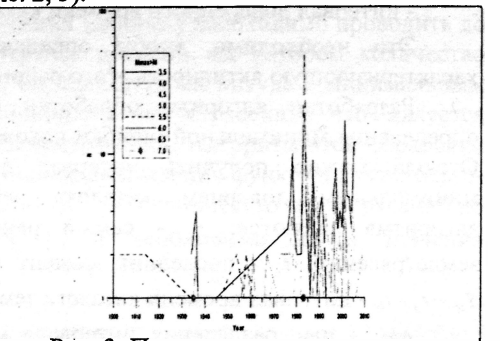


Рис. 3. Плотность распределения землетрясений по северо-западному району Вьетнама по интервалам магнитуд

По графику плотности распределения землетрясений по северо-западному району Вьетнама по интервалам магнитуд видно, что землетрясения с магнитудой ниже 3.5 не имели значения, особенно с 1976 года. В результате применения алгоритма 3.5 является доверенной магнитудой каталога Северо-западного района Вьетнама.

В третьей главе описаны результаты применения метода нейронной сети, используемого как для распознавания образов и для прогноза, в задаче прогноза землетрясений для северо-западного района Вьетнама. Входными данными являются образы или значения сейсмических параметров в прошлом. Метод нейронной сети применяется по 2 направлениям, для определения возможности максимальных землетрясений и землетрясениям по магнитуде и времени.

В рамках работы проведены эксперименты по применению нейронных сетей для прогноза максимальной силы землетрясений Северо-Западного района Вьетнама. Данные, использованные для прогноза: 1 - Значение плотности линеаментов; 2 - Значение градиента Буге гравитационного поля; 3 - Градиент аномалий магнитной поля; 4 - Градиент вертикального сдвига современной коры тектоники; 5 - Градиент толщины осадочного чехла; 6 - Градиент глубины кристаллического фундамента; 7 - Градиент толщины земной коры.

Эти параметры влияют на возникновение землетрясений, которые частично использовались в прежних исследованиях. В работе использован каталог землетрясений Вьетнама, который составлен Институтом геофизики до 2010 года и обновлён международными и историческими данными. Такие значения вычисляются по сеткам с величиной 0.2 по координатам.

Северо-западный район разделён на 225 квадратных одинаковых частей с интервалом 20 км. Для каждого квадрата вычисляются средние значения по 7 данным. Выбирается представленное землетрясение для каждого квадрата. И выходными данными сети является максимальное значение землетрясений по каждой сетке. Геологические параметры вычисляются по сеткам.

Обучение сети водится на выборке, содержащей 70% данных, 15% данных для проверки состояния сети и последние 15% для проверки работоспособности сети. Для каждого интервала данных вычисляются значения R и последнее значение, которое показывает точность сети, вычисляется по средним значениям R. Примерами для обучения сети, для землетрясений больше 4.5, будут подрайоны с возможностью землетрясений с магнитудами больше или равными 4.5.

Из опыта выбора сети и по теории нейронных сетей выбираются только такие сети, у которых есть 1 скрытый слой. Проведём эксперименты с 3 типами сетей, FBP, RNN и RBF. Параметром сравнения сетей (по главе 2) является значение R (Regression). Чем ближе R к 1, тем точнее полученная сеть.

Таблица 2.

Ошибки по разным нейронным сетям

Нейронные сети	Regression (R)
FBP	0.9006
RNN	0.553
RBF	0.604

Из табл. 2 видно, что сеть FBP является самой лучшей обученной сетью. Она используется для прогноза возможно максимальных землетрясений по северо-западному району Вьетнама (рис. 4).

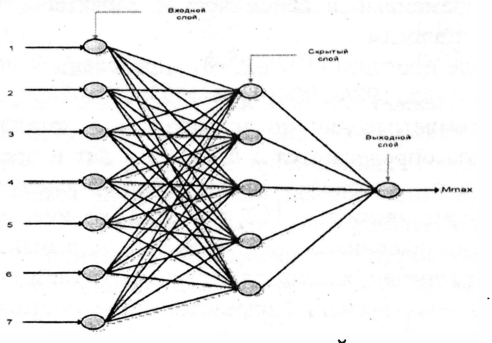


Рис. 4. Нейронная сеть для прогноза землетрясений северо-западного района Вьетнама

Полученное значение R равно 0.9006, ошибка нейронной сети не превышалась 0.1. Этим результатам подтверждают эффективность применения НС с тремя слоям, с 7 различными значениями на входе, с 5 нейронами у

скрытного слоя и одного нейрона на выходе, который даёт нам возможно максимальные землетрясения северо-западного района Вьетнама.

Результаты прогноза показаны на рис. 5. Они совпадают с положениями очагов этого района. В результате выделены пять наиболее сейсмических участков Северо-Западного района Вьетнама.

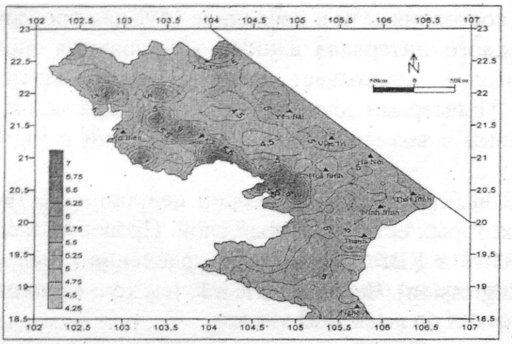


Рис. 5. Карта результатов прогноза значений максимальных землетрясений для северо-западного района Вьетнама по каталогу с 1900 г.

По прогнозу землетрясений надо ответить на 3 вопроса: где, когда и с какой магнитудой будут землетрясения. Поэтому задача прогноза возможно максимальных землетрясений даёт только решения по характеристике сейсмичности района или очагов, эти результаты нуждаются в дальнейших научных исследованиях. Метод нейронной сети позволяет определить вероятное время и магнитуды будущих землетрясений. Для этих целей были использованы временные и сейсмические характеристики землетрясений по региональным каталогам.

В интервале прогноза (1976-2010) определены 7 подинтервалов времени: 1976-1980; 1981-1985; 1986-1990; 1991-1995; 1996-2000; 2001-2005 и 2006-2010. Минимальное землетрясение по магнитудам в каталоге 3.5. Подинтервалы времени прогноза определяются с шагом в 5 лет и предыдущий подинтервал используется для прогноза землетрясения следующего подинтервала. Входными данными являются: 1/ T значения (количество дней исследования); 2/ $Mmean$ средние значения; 3/ $dE^{1/2}$ значения; 4/ b значения; 5/ η значения; 6/ ΔM значения. Эти значения вычисляются по формулам 1-10.

1. T значение: Интервал времени T определяется следующей формулой:

$$T = t_n - t_1 \qquad (1)$$

где, t_1 - момент происхождения первого землетрясения; t_n - момент происхождения n_ого землетрясения; n - количество землетрясений в разделенном интервале времени.

2. $Mmean$ значение: определяется по следующей формуле:

$$Mmean = \sum M_i / n; \qquad (2)$$

M_i : мощность i_ого землетрясения по магнитудам, $i = \overline{1,n}$.

3. $dE^{1/2}$ значение: считается по формуле:

$$dE^{1/2} = \sum E^{1/2}/T \quad (3)$$

E вычисляется по формуле Рихтера:

$$E = 10^{(11,8+1,5M)} \quad (4)$$

4. b значение определяется по уравнению Гуттенберга-Рихтера:

$$logN = a - bM \quad (5)$$

где, N количество землетрясений, имеющих мощности по магнитудам $\geq M$;

a, b: вычисляются по формулам;

$$b = (n\sum(M_i log_{10}N_i) - \sum M_i \sum log_{10}N_i)/((\sum M_i)^2 - n(\sum M_i{}^2)) \quad (6)$$

$$a = \sum(logN_i + bM_i)/n \quad (7)$$

N_i- количество землетрясений, имеющих мощности по магнитудам $\geq M_i$.

5. η значениеопределяется по функции Гуттенберга-Рихтера:

$$\eta = \sum(logN_i - (a - bM_i))^2/(n-1) \quad (8)$$

6. ΔM значениеопределяется по формуле:

$$\Delta M = Mmax(\text{наблюдения}) - Mmax(\text{прогноза}) \quad (9)$$

где, $Mmax$(наблюдения): значение максимального землетрясения по магнитудам среди n землетрясений; $Mmax$(прогноза): значение максимального землетрясения в прогнозе по магнитудам по функции Гуттенберга-Рихтера;

$$Mmax(\text{прогноза}) = a/b \quad (10)$$

Проведём эксперименты с 3 типами сетей, FBP, RNN и RNN. Параметром сравнения сетей являются максимальная ошибка и средняя ошибка.

Таблица 3.

Ошибки по разным нейронным сетям

Тип нейронной сети	Максимальная ошибка	Средняя ошибка
FBP	1.47	0.34
RNN	2.95	0.94
RBF	0.2	0.03

Из табл. 3 видно, что у нейронной сети RBF есть самый лучший результат обучения, но значение прогноза землетрясений этого района отличается от предыдущих результатов, приведённых в публикациях. Неточный результат прогноза объясняется малым числом данных для обучения сети, и нейронная сеть RBF не смогла найти сейсмический закон этого района. Малое количество данных также является причиной невозможности применения сети RNN. Нейронная сеть FBP с 5 нейронов скрытого слоя и передаточная функция tansig применима в нашем случае.

Применим нейронную сеть FBP для прогноза возможно максимальных землетрясений по северо-западному району Вьетнама (рис. 6). Используется нейронная сеть FBP с 3 слоям: входной (6 нейронов), скрытый (5 нейронов), и выходной (1 нейрон) для прогноза землетрясений северо-западного Вьетнама со значением R = 0.72.

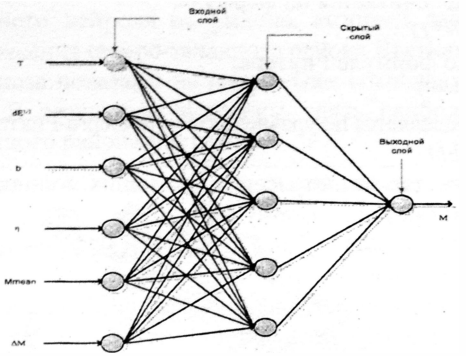


Рис. 6. Нейронная сеть для прогноза землетрясений северо-западного района Вьетнама

Выходным значением является мощность будущего землетрясения, активационная функция нейронов - logsig. Разница между реальными и прогнозными значениями не большая (табл. 5). Таким образом, возможность применения метода нейронной сети для прогноза землетрясений северо-западного района Вьетнама (время, место и мощность будущего землетрясения) подтверждена моделированием в Matlab с модулем нейронной сети (Neural Network Toolbox 6.0.4/nntool).

Таблица 4.

Входные данные для прогноза землетрясений северо-западного Вьетнама

Подинтервалы прогноза (годы)	T	$dE^{1/2}$	b	η	$Mmean$	ΔM
1976.1980	1716	1.21581e+07	0.72819	0.00676	4.01111	0.34626
1981.1985	1358	1.219722e+08	0.79910	0.00779	3.93152	0.78071
1986.1990	1484	2.248179e+07	0.69068	0.00135	4.03929	0.74227
1991.1995	1674	1.87813e+07	0.78829	0.01469	4.16528	0.42182
1996.2000	1555	1.15112e+07	0.97625	0.00336	3.92381	0.12019
2001.2005	1725	3.01123e+07	0.76267	0.01445	4.15556	0.41732

Таблица 5.

Реальные значения и полученные результаты прогноза нейронной сетью FBP

	1981-1985	1986-1990	1991-1995	1996-2000	2001-2005	2006-2010
M реальные	6,6	6,4	5,1	4,7	5,0	4,8
$Mmax$прогноза	6,7	4,9	5,0	4,8	5,3	4,7

Использован каталог землетрясений северо-западного района Вьетнама в интервале времени 2006- 2010, входные значения приведены в табл. 6.

Таблица 6.

Входные значения для прогноза будущего землетрясения северо-западного района Вьетнама в интервале времени 2011-2015

Интервал времени	T	$dE^{1/2}$	b	η	M_{mean}	ΔM
2006.2010	1815	6.341305e+06	0.66698	0.00348	3.57368	0.26739

Максимальное прогнозируемое значение землетрясения северо-западного района Вьетнама равно 6.7 по магнитудам для интервала времени 2011-2015 получено для нейронной сети FBP. Это значение совпадает с результатом применения алгоритма M8, определявшегося прогнозируемое значение землетрясения для района больше 6.0. Таким образом, подтверждена высокая вероятность будущего землетрясения этого района.

Рассмотрим выбор интервала прогноза в задаче среднесрочного прогноза землетрясений по методике, предложенной во второй главе.

По количеству землетрясений северо-западного района Вьетнама в задаче среднесрочного прогноза для этого района применён только подинтервал с 3 до 5 лет. Интервал времени прогноза определяется интервалом времени анализа каталога землетрясений. Этот принцип был использован для выбора интервала прогноза и результаты (табл. 7) показывают, что интервал времени 3 года достаточный для прогноза землетрясений по 5.0-6.0 магнитудам, для 4 лет для прогноза землетрясений по 6.0-7.0 магнитудам. Прогноз получен для интервала времени (2011-2015) при помощи нейронной сети FBP.

Таблица 7.

Результаты выбора интервала прогноза

Параметры / Интервал времени	R	MSE	M прогноз
5 лет	0.72	0.3740	6.6991
4 года	0.89	0.1750	6.6997
3 года	0.92	0.0307	5.1573

В четвертой главе описаны метод вычисления вероятности появления землетрясения и применение его для вычисления вероятности землетрясения с магнитудой более 6 для северно-западного района Вьетнама. Кроме этого, представлена программа обработки данных для задачи прогноза землетрясений методом нейронной сети.

Имеются некоторые методы вычисления вероятности явления землетрясений такие, как метод на основе распределения Пуассона или нормальное распределение логарифма деления между временами явления землетрясений в реальности и в теории по статистике. В этой работе предлагается метод вычисления по нормальному распределению.

Вероятность появления землетрясений вычисляется как условная вероятность между временами последнего землетрясения этого типа до возможного момента. Вычисления осуществляются по формуле.

$$P\left(\frac{T^*}{T}\right) = \frac{P(TT^*)}{P(T)}, \quad (11)$$

где, $P(TT^*) = P(T \le t \le T^*)$; $P(T) = P(t \ge T) = 1 - P(t < T)$.

$\log(\frac{T}{T_{lt}})$ имеет нормальное распределение, это значит, что плотность распределения иметь вид:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, x \in R,$$

где, a - центральный момент; σ^2- дисперсия.

Функция распределения имеет вид:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(x-a)^2}{2\sigma^2}} dx. \quad (12)$$

Для вычислений необходимо преобразить нормальное распределение к стандартному распределению. Пусть $z = \frac{x-a}{\sigma}$, тогда (12) становится: $F(x) = \frac{1}{2} + \phi(\frac{x-a}{\sigma})$, где ϕ - функция Лапласа: $\phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz$.

По теории вероятности имеется формула вычисления вероятности в интервале.

$$P(\alpha \le X < \beta) = F(\beta) - F(\alpha) = \phi(\frac{\beta-a}{\sigma}) - \phi(\frac{\alpha-a}{\sigma}) \quad (13)$$

Тогда (11) становится:

$$P = \left[\phi\left(\frac{T^*-a}{\sigma}\right) - \phi\left(\frac{T-a}{\sigma}\right)\right] / [1 - \phi(\frac{T-a}{\sigma})] \quad (14)$$

где, $T^* = \log[(t + \Delta t)/T_{lt}]$ и $T = \log[t/T_{lt}]$.

В итоге по формуле (14) вычисляется вероятность появления землетрясения.

Процесс вычислений состоит из следующих шагов: Выбираются землетрясения с магнитудой более 3.5 из каталога. Количество таких землетрясений будет 21; Определение значения $Mmin$ для этого района; Определение данных для линейной регрессии T, $Mmin$, Mp; Выбор данных с шагом 3, между которыми был интервал времени более 1 года и землетрясения после этого имеют магнитуды более Mp. Получены данные 23 землетрясений; Вычисление параметров $\log(T)$, $Mmin$ и Mp.

В результате: a $= -0.0138$; b $= 2.121$; c $= -1.5136$ и дисперсия равна 2.6. Вероятность вычисляется по значениям $Mmin$=5 и Mp= 6.1. Тогда, $T = 23$ и Δt считается интервалом от 2011 года до 2015 года (5 лет).

Самое максимальное землетрясение с магнитудой 6.7 произошло в 1983 году. С 1983 года до 2011 года будет 28 лет. Тогда, $P = 0.89$.

В результате получено, что вероятность появления землетрясения более 6 в районе северо-западного Вьетнама до 2015 года равна 89%.

Результат вычисления вероятности проверяется с применением другого метода вычислений. Вычисляется частота появления землетрясений по

функции Гуттенберга-Рихтера, имеющий вид $\log(\frac{N}{T}) = 3{,}76 - 0{,}89 * M$, для этого района. Значение T необходимо определять по магнитуде M. M – это минимальная магнитуда землетрясений. Результаты показаны, что до 2020 года будет, хотя бы одного землетрясения с магнитудой более 6. Поэтому вероятность 89% имеет реальный смысл.

Для решения задачи прогноза землетрясений методом нейронной сети программа обработки данных написана на языке высшего уровня C#. При помощи программы решены 3 типа задач: обработки входных данных для двух описанных выше задач прогноза землетрясений, нахождение доверенного интервалы и магнитуды.

ЗАКЛЮЧЕНИЕ

В результате проведённого научного исследования были получены следующие результаты:

1. На основе системного анализа многофакторных данных каталога землетрясений Вьетнама, с использованием функции Гуттенберга-Рихтера, определена сейсмическая активность районов страны. Определено, что наибольшей сейсмической активностью обладает северо-западный район;
2. Разработка, исследование и реализация алгоритма обработки данных каталога землетрясений;
3. Обоснованный выбор наиболее эффективной нейронной сети (FBP)с обратным распространением ошибки для среднесрочного прогноза землетрясений на основе сравнительного анализа различных типов нейронных сетей;
4. Методика обучения нейронной сети в задаче прогноза землетрясений;
5. Применение нейронной сети для среднесрочного прогноза землетрясений в северо-западном районе Вьетнама и результаты прогноза. Показано, что в северо-западном районе Вьетнама в период с 2010 г. по 2015 г. возможно землетрясения с вероятностью 0.89, магнитудой 6.7;
6. Обоснование достаточности данных каталога землетрясений с магнитудой более 3.5, произошедших в северо-западном районе Вьетнама с 1975 года для определения вероятности землетрясений с магнитудой более 6;
7. Разработка и реализация полной методики прогноза землетрясений с использованием нейросетевых технологий, исходя из системного анализа каталожных данных по землетрясениям, и получение результатов прогноза.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Прогноз возможно максимальных землетрясений в северо-западном районе Вьетнама / К.А. Пупков [и др.] // Вестник Российского университета дружбы народов. Инженерные исследования. 2011. № 3. С. 74-79.
2. Пупков К.А., Као Динь Чонг. Применение нейросетевых технологий в задачах прогнозирования землетрясений (на примере северо-западного района Вьетнама) // Вестник Московского государственного технического университета имени Н.Э. Баумана. Приборостроение. 2012. № 2 (87). С. 70-78.

3. Пупков К.А., Као Динь Чонг. Принципы многослойных нейронных сетей для прогноза землетрясений (на примере северо-западного района Вьетнама) // Интеллектуальные системы (INTELS'2012). Вологда, 2012. С. 137-144.

4. Као Динь Чонг. Применение метода нейтронной сети в области среднесрочного прогноза землетрясений // Инженерные системы: Тез. докл. Меж. науч.-практ. конф. Москва, 2012. С. 48.

5. Као Динь Чонг. The application of neural network for earthquake prediction in North-West Vietnam // 9th General Assembly of Asian Seismological Commission: Extended abstract of the International Conference. Ulaanbaatar, 2012. P. 149-151.

6. Cao Dinh Trieu, Pham Nam Hung, Cao Dinh Trong. Establishing the earthquake catalogue and intermediate term earthquake prediction in Vietrnam on the basis of CN algorithm // Journal of Geology. Series A. 2007. Vol. 5-6, № 300. P. 35-49.

7. Cao Dinh Trieu, Le Van Dung, Cao Dinh Trong. Initial results of seismic hazard essessments in Vietnam on the basic of neo-determinictic algorithm // Journal of Sciences of the Earth. 2007. Vol. 29, № 4. P. 333-341.

8. Cao Dinh Trieu, Le Van Dung, Cao Dinh Trong. Seismic hazard essessment in Vietnam on the basic of neodeterministic algorithm // Proceeding of National Symposium on Geohazards and Mitigation. Hanoi, 2008. P. 31-43.

9. Cao Dinh Trieu, Thai Anh Tuan, Cao Dinh Trong. Some seismotectonic features of Southeast Asia // Journal of Geology. Series A. 2008. Vol. 5-6, № 306. P. 3-13.

10. Seismic hazard assessment of Vietnam territory on the basis of deterministic approach / Cao Dinh Trieu [et al.] // Journal of Geology. Series B. 2008. №31. P. 220 - 230.

11. Some new outcomes of the intermediate term earthquake prediction in Vietnam / Cao Dinh Trieu [et al.] // Journal of Geology. Series B. 2008. № 31. P. 231 - 240.

12. Research on middle-earthquake predictions in Vietnam / Cao Dinh Trieu [et al.] // Journal of Geology. Series A. 2009. № 314. P. 47 - 55.