

СИСТЕМАТИЗАЦИЯ И ОРГАНИЗАЦИЯ ХРАНЕНИЯ ГЕОФИЗИЧЕСКОЙ И СЕЙСМОЛОГИЧЕСКОЙ ИНФОРМАЦИИ С ПРИМЕНЕНИЕМ ТЕХНОЛОГИИ РЕЛЯЦИОННЫХ БАЗ ДАННЫХ

Терещенко М.В., Гриб Н.Н.

ФГАОУ ВПО «Технический институт (филиал) СВФУ им. М.К. Аммосова в г. Нерюнгри», Нерюнгри, Россия (678960, Республика Саха (Якутия), г. Нерюнгри, ул. Кравченко, 16), e-mail: terexa@pochta.ru

В данной статье рассматривается необходимость организации результатов геофизических и сейсмологических наблюдений в реляционную базу данных. На основании анализа накопленных в результате исследований данных определены соответствующие требования к базе данных. Автором предложена схема базы данных в составе комплекса обработки информации и ее программная реализация. Приведен алгоритм импорта экспериментальных данных в базу данных. Рассмотрены недостатки существующего подхода к хранению информации в виде текстовых файлов и электронных таблиц Excel. Обосновано применение СУБД MySQL для решения поставленной задачи. Предложено применение математических фильтров для обработки исходных данных с целью исключения случайных выбросов и результатов системных ошибок. Формирование базы данных и применение математического аппарата для обработки экспериментальных данных позволит осуществить математическое моделирование сейсмических событий по геофизическим предвестникам.

Ключевые слова: математическая модель, база данных, землетрясения, геофизические данные, сейсмологические данные, предвестники землетрясений, прогнозирование сейсмических событий.

THE GEOPHYSICAL AND SEISMOLOGICAL DATA SYSTEMATIZATION AND STORAGE USING RELATIONAL DATABASE TECHNOLOGY

Tereschenko M.V., Grib N.N.

Technical Institute (branch) of North-Eastern Federal University named after M.K. Ammosov, Nerungri, Russia (678960, Nerungri, street Kravchenko, 16), e-mail: terexa@pochta.ru

This article discusses the necessity of organizing the geophysical and seismological observations results in a relational database. The article identifies the relevant requirements to the database based on the analysis of accumulated research data. The author proposed database schema as part of the information processing complex and its software implementation. The article describes an algorithm for experimental data import into the database. It is concluded that the existing approach to storing information in text files and Excel spreadsheets has some significant drawbacks. The author justified the MySQL using to solve the problem. It is proposed to apply mathematical filters to process the source data to prevent accidental emissions and system errors effects. The database creation and mathematical apparatus application will allow the seismic events mathematical modeling using geophysical forerunner.

Keywords: mathematical model, database, earthquakes, geophysical data, seismic data, geophysical forerunner, seismic events prediction.

Среди природных катастроф землетрясения занимают одно из первых мест по наносимому экономическому, экологическому и социальному ущербу. В настоящее время задача определения ключевых параметров возможных сильных землетрясений занимает центральное место при сейсмологических исследованиях. При этом очевидно, что для ее решения необходимо проведение детальных комплексных наблюдений, как за сейсмическими процессами, так и геофизическими полями. Необходимо установить закономерности их динамики во времени и пространстве, аномальное поведение в период подготовки землетрясения в зонах с высокой сейсмической активностью.

Учитывая сложность процесса, недоступность для непосредственного изучения объекта исследования, наличие помехообразующих факторов и ограниченную во времени и объеме достоверную базу данных о сейсмических событиях, прогноз землетрясений носит вероятностный характер. И, таким образом, документирование землетрясений играет значимую роль в научных исследованиях, связанных с прогнозированием сейсмических событий.

Детальные геофизические и сейсмотектонические исследования, в результате которых выделены потенциально опасные сейсмогенерирующие зоны и блоки [3, 6], свидетельствуют о необходимости изучения сейсмической активности Южной Якутии, параметров регистрируемых землетрясений для последующей оценки, районирования и прогнозирования сейсмической опасности региона.

В настоящее время в г. Нерюнгри на базе Технического института (ф) СВФУ функционирует лаборатория мониторинга и прогноза сейсмических событий, оснащенная цифровой сейсмологической станцией, а также комплексом геофизической аппаратуры нового поколения. Сейсмостанция регистрирует все сейсмические события в реальном режиме времени, а комплекс геофизического оборудования – изменения геофизических параметров, связанных с геодинамическими процессами в земной коре.

Для построения математической модели сейсмологических событий необходимо выявить закономерности в статистике эмпирических сейсмологических и геофизических данных, отражающих изменение сейсмического режима региона. При этом предполагается наличие детальной информации о значениях и состояниях, которые принимали те либо иные характеристики изучаемого явления, объекта. Отсутствие данных за определенный период может стать причиной снижения статистической значимости полученных результатов.

Таким образом, качество модели и точность оценки параметров прогнозируемых сейсмических событий зависит как от методов и средств регистрации сейсмических событий, так и от качества информации, хранимой в накопленной базе данных, и методов ее первичной обработки.

Указанные обстоятельства обуславливают актуальность разработки базы данных (БД) сейсмических событий и реализации методик обработки экспериментальных геофизических данных, содержащихся в указанной БД (рис.1). Реализация указанных задач позволяет исследователю варьировать набором параметров с целью повышения адекватности модели реальным геодинамическим условиям региона и разработать математическую модель сейсмических событий на основании определенного им самостоятельно набора предвестников и характеристик сейсмического процесса.

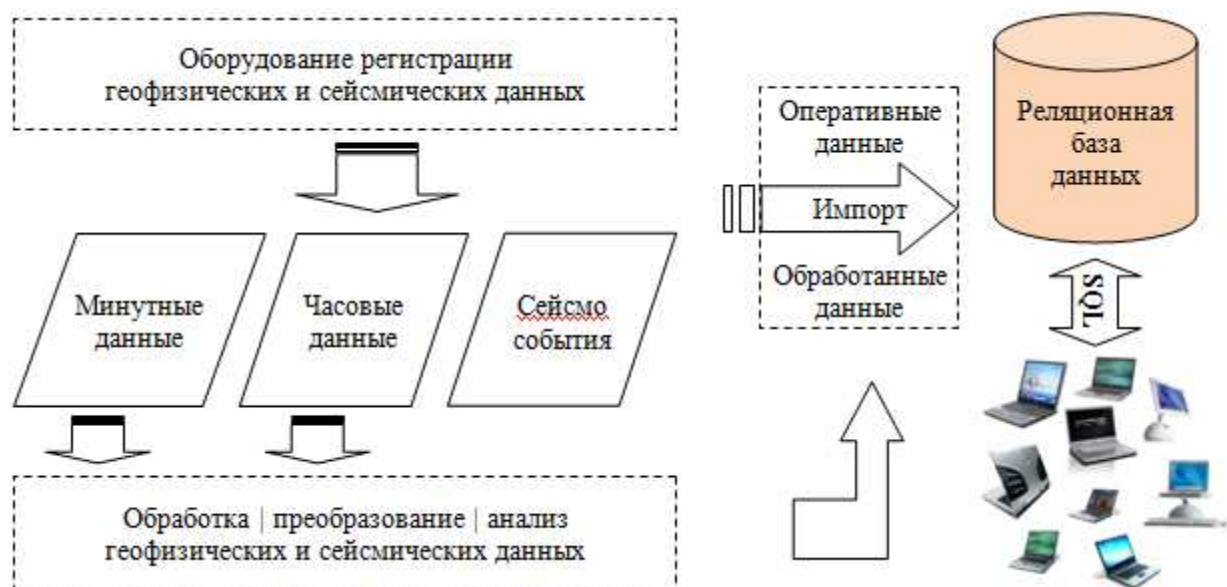


Рис.1. База данных в составе комплекса обработки информации

Систематизация и хранение геофизической и сейсмологической информации с применением технологии реляционных баз данных позволяет обеспечить:

- многопользовательский режим работы с экспериментальными данными;
- удаленный режим работы с данными различных исследователей;
- оперативную обработку информации за счет устранения избыточности информации, применения оптимальных методов обработки табличных данных, специализированного языка программирования SQL.

Проведение многолетних регулярных наблюдений за вариациями геофизических полей (гравиметрические и магнитометрические наблюдения, электромагнитные, барометрические, температурные) определило необходимость изменения подхода к организации хранения экспериментальных данных.

Применение текстовых файлов на раннем этапе позволило неким образом сгруппировать исходную информацию, обеспечить универсальность доступа и устойчивость данных. Вместе с тем данный формат обладает низкой энтропией при неэффективных способах доступа к информации и ее обработке [7].

Переход от хранения в файлах текстового формата к использованию электронных таблиц Excel (рис. 2) позволил представить информацию более наглядно, но проблема работы со значительными выборками данных, тем не менее, осталась не решенной. Это связано с:

- дублированием информации в пределах как одного, так и разных файлах;
- ограничением наглядности представления информации;
- ограничением возможностей многопользовательского доступа;
- скоростью обработки информации;

- отсутствием контроля применения единого формата при обозначении параметров и способов их представления. Например, как и в случае txt файлов время регистрации может быть представлено и в кратком формате даты-время, и в любом другом виде;
- ограниченным объемом файла электронной таблицы в более ранних версиях.

№ п/п	Дата	Tr	Ts	s-p	Δ°	км	To	tp	ZNE	NE	Y(N)	$\lambda(E)$	h	to	K
п/п	д м	ч мин с	ч мин с	мин с	о	км	ч мин с	мин с мс	ZNE	NE	Y(N)	$\lambda(E)$	км	ч мин с	
3	6.02	+p04:17:07.5	04:17:11.7	0-04	0.569	5	04:16:55.2		37.1	36.8	56.67	123.86	5	04:16:53.7	5.4
13	27.03	p04:15:40.7	04:15:42.9	0-02	0.134			7.04	50.7	50.0	55.81	125.04		04:15:17.7	
14	27.03	p04:16:20.7	04:16:22.7	0-02	0.165		04:16:11.2	5.24	37.3	37.3	56.65	124.42		04:16:11.1	
15	29.03	p04:29:11.0	04:29:14.3	0-03	0.217			5.01	219.9	218.4	56.89	124.92			

Рис. 2. Пример организации каталога сейсмологических событий в формате Excel

Помимо этого, отметим необходимость обеспечения совместимости информации, представленной в файлах более ранних версий. Зачастую это вызывает ошибки, системные вопросы при открытии и сохранении файлов, а значит, и снижение быстродействия. А также необходимость ручного контроля соответствия полей в разных файлах, типов данных и наименований информационных файлов.

Анализ накопленной информации, а также технических характеристик применяемого в исследованиях оборудования позволил сформулировать требования к организации хранения экспериментальных геофизических и сейсмологических данных:

- соблюдение требований теории реляционных баз данных;
- устранение дублирования информации;
- возможность корректировки экспертным методом хранимой информации;
- обеспечение необходимой точности хранения информации;
- гарантия сохранности информации в течение длительного времени;
- масштабируемость и переносимость на другие платформы;
- расширяемость, универсальный доступ с использованием специализированного широко распространенного стандартизированного языка доступа к данным SQL.

Импорт ретроспективных данных (рис.3) и оперативной информации в реляционную базу данных, спроектированную с применением правил нормализации, позволяет получить оптимальную структуру для хранения данных, а также возможность организации их графической интерпретации и оперативной обработки.



Рис. 3. Блок-схема обработки экспериментальных данных

Применение языка SQL обеспечивает минимизацию времени обработки пользовательских запросов, что с учетом объема накопленной информации дает значительную экономию временных ресурсов.

База данных (рис. 4) реализована в системе управления базами данных MySQL (СУБД). Соответствующий выбор определен тем фактом, что указанная СУБД предоставляет полноценные возможности организации хранения значительных объемов данных, многопользовательский доступ к информации, обеспечивает взаимодействие с различными WEB-серверами, свободна к распространению для использования в образовательных и научных целях и характеризуется оптимальным быстродействием [1].

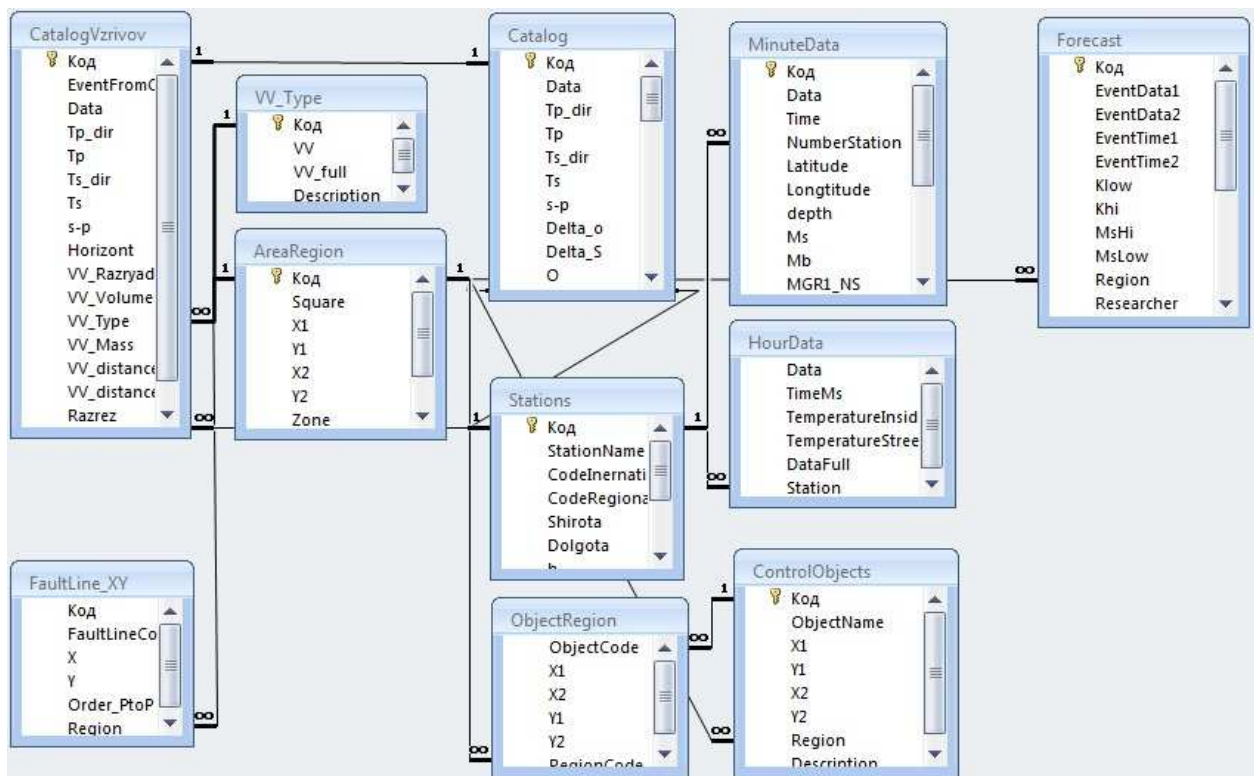


Рис. 4. Фрагмент схемы базы данных SeismicEvents

Предлагаемая структура обеспечивает хранение необходимых регистрируемых параметров сейсмических событий с заданной точностью, целостность данных и отсутствие дублирования информации.

Отметим, что при формировании базы данных на основании проводимого сейсмического и геофизического мониторинга получаемая последовательность значений может содержать как систематические, случайные ошибки, так и выбросы [4]. Выбросы могут быть вызваны влиянием внешних источников: статический разряд, неисправность регистрирующей аппаратуры, перепады напряжения и т.п. Перед обработкой данных такие значения необходимо устранить, поскольку иначе это приводит к некорректному описанию данных одной функцией распределения, искажению результатов статистического анализа [5].

Применение математических фильтров, усреднения экспериментальных данных с определенным шагом на временном интервале позволяет исключить значения, относящиеся к случайным выбросам, системным ошибкам; а также установить тренды в хранимых временных рядах. Вместе с тем нельзя исключать факт значимости каждого отдельного зафиксированного значения в выявлении и формулировке зависимости параметров сейсмических событий. На основании этого в описываемой базе данных обеспечено сохранение данных значений для последующей оценки после накопления более значительного массива информации (рис.5).

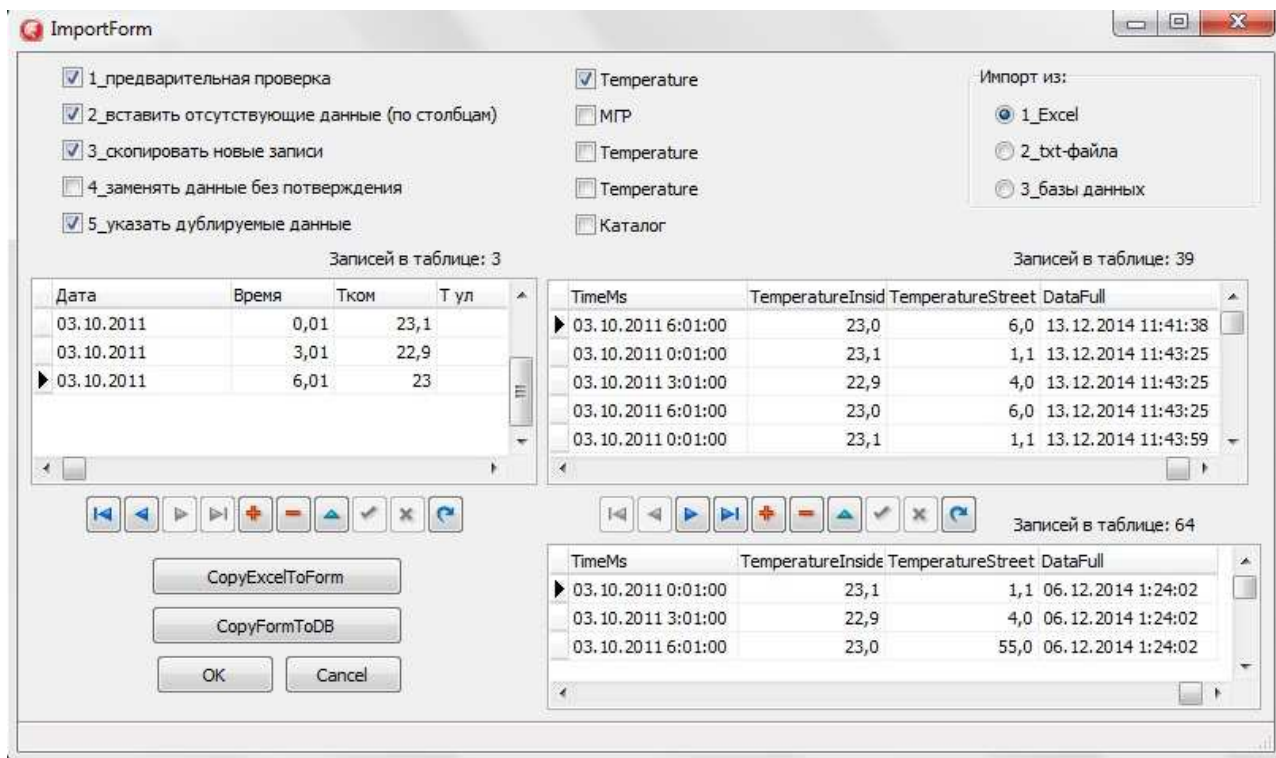


Рис.5. Форма импорта исходных данных модуля SeismicAnalysis

Формирование указанной базы данных направлено на создание условий для автоматического выявления аномалий, предвещающих сейсмические события, или надежных предвестников [2], с целью последующего моделирования сейсмических событий.

Список литературы

1. Велинг Л., Томсон Л. MySQL. Учебное пособие: пер. с англ. – М.: Вильямс, 2005. – С.20.
2. Гриб Н.Н., Терещенко М.В. Динамика сейсмического режима и геофизических полей в Ю.-Якутском регионе // Фундаментальные исследования. – 2014. – № 9, часть 4. – С.784-788.
3. Имаева Л.П., Имаев В.С., Козьмин Б.М. Сейсмогеодинамика Алдано-Станового блока // Тихоокеанская геология. – 2012. – Т. 31. – № 1. – С. 5-17.
4. Косарев Е.Л. Методы обработки экспериментальных данных. – М.: Физматлит, 2008. – С.34.
5. Кулаичев А.П. Методы и средства комплексного анализа данных: учебное пособие / А.П. Кулаичев. – 4-е изд., перераб. и доп. – М.: Форум: Инфра-М, 2014. – С.75.
6. Ovsyuchenko, A.N., Trofimenko, S.V., Marakhanov, A.V. et. al. Detailed Geological–Geophysical Studies of Active Fault Zones and the Seismic Hazard in the South Yakutia Region. *Russian Journal of Pacific Geology*, 2009, Vol. 3, No. 4, pp. 356–373. DOI: 10.1134/S1819714009040046
7. Wikipedia. Свободная энциклопедия. URL: https://ru.wikipedia.org/wiki/Текстовый_файл.

Рецензенты:

Имаев В.С., д.г.-м.н., профессор, гл. научный сотрудник Института земной коры СО РАН, г. Иркутск;

Трофименко С.В., д.г.-м.н., профессор, ведущий научный сотрудник лаборатории сейсмологии и сеймотектоники ФГБУН Институт тектоники и геофизики им. А.Ю. Косыгина Дальневосточного отделения Российской академии наук, г. Хабаровск.