Географические науки

Бабаханов Аслан Эльдар
аспирант
Института Географии имени академика Г. А. Алиева
НАН Азербайджана
Babakhanov Aslan
Ph.D Candidate of the
Institute of Geography of the
National Academy of Sciences of Azerbaijan

POWER BI КАК ИНСТРУМЕНТ ДЛЯ ГИС-АНАЛИЗА СТИХИЙНЫХ БЕДСТВИЙ НА ТЕРРИТОРИИ АЗЕРБАЙДЖАНА
POWER BI IN GIS ANALYSIS OF NATURAL DISASTERS IN THE TERRITORY OF AZERBAIJAN

Аннотация: *В статье рассматриваются применение программного решения Microsoft © PowerBI для анализа данных и её применение как инструмент для ГИС-анализа и представления отчётов. Приводятся ряд методик для организации данных стихийных бедствий (на примере землетрясений) проявляемые на территории Азербайджана с учётом ведения реестра пространственных зависимостей, а также создание панели отчётов и оперативного анализа поступающих данных. Даётся сравнения с существующими программными продуктами.*

Ключевые слова: *стихийное бедствие, ГИС анализ, PowerBI, отчёт, база данных.*

Summary: *The goal is to review the possibilities of Microsoft Power BI in data analysis and its adopting in GIS analysis and dashboard generation. Listeda number of methods to organize natural hazards data (here, earth quakes) on territory of Azerbaijan by registering referencing geospatial data dependencies,*

making dashboards and operational analysis of incoming data. Providing comparison of existing software.

Key words: *natural disaster, GIS analysis, Power BI, dashboard, data base.*

Решение пространственных задач ГИС всегда упиралось на статистическом анализе табличных данных собранных из различных источников, сгруппированных и классифицированных по тематическим представлениям в табличной форме. Большинство данных сохраняется в электронных таблицах с использованием офисного программного обеспечения, как Microsoft Excel. Более развитые данные представлены в формате баз данных, таких как Microsoft Access, Microsoft SQL Server, Oracle и т.д. Широкое распространение получили формат геобаз данных, реализованным компанией ESRI в программных решениях серии Arc GIS для визуализации, хранения и обработки данных ГИС. Однако, в связи с нарастающей необходимостью прогнозирования поведения и развития стихийных бедствий и их влияний на людей, природу и экономику возникают задачи обобщения исторических данных с последующим анализом. Аналитических программных инструментов на рынке программного обеспечения достаточно, но большинство требует подготовки данных, приведение в свой формат данных, изучение самого продукта, а это в свою очередь оказывает влияние на своевременное и оперативное получение результатов. Свою роль в торможении процесса анализа также вносит и высокая стоимость некоторых решений.

Для решения задачи можно отнести продукт Power BI от компании Microsoft© для анализа и подготовки панели отчётов, настольное решение которого распространяется бесплатно. Все элементы интерфейса интуитивны и просты в освоении, а список поддерживаемых источников данных велик, раскрывая большие возможности.

Систематизация данных СБ на территории Азербайджана. Территория Азербайджана отличается высокой сейсмической активностью,

где за исторический период произошли сильные и катастрофические землетрясения с магнитудой M≥6: Гейгельское 1139 г., Восточно-Кавказское 1668 г., Маштагинское 1842 г, многочисленные Шамахинские землетрясения (1192,1667, 1669, 1828, 1859, 1868, 1872,1902 гг.), Каспийские землетрясения (957, 1812, 1842,1852, 1911, 1935, 1961, 1963, 1986, 1989, 2000) и др.

В истории исследований сейсмичности на территории Азербайджана можно выделить два основных периода:

– Первый период исследований – до инструментальный, исторический, включающий всю информацию с глубокой древности, отраженную в древних арабских хрониках, рукописях, манускриптах, летописях, путевых заметках путешественников.
– Второй период исследований – инструментальный, современный период, включающий информацию о землетрясениях с начала XX века, когда после Шемахинского землетрясения 1902 г., была основана первая в Азербайджане сейсмическая станция "Шемаха" [4].

Однако, картографирование СБ до 1920 года в целом не получило своего широкого развития. Имевшиеся карты в основном относились к землетрясениям, так из них можно отметить карту землетрясений в Шемахе и Эрзуруме, составленную академиком Г. В. Абихом. Первая после 1920 года карта землетрясений составлена Е. И. Бюсом. Она посвящена землетрясениям в Закавказье, произошедшим в течение столетия (1830-1930). Первая карта селевых явлений в Закавказье подготовлена 1931 г. И. Рощиным. В последующие годы по отдельным видам СБ были составлены ряд карт. Часть из них, посвященных республике, помещена в Атласе Азербайджанской ССР (1963). Среди них:

- Карта сейсмического районирования (Ш.С. Рагимова, В.П. Кузнецова, Э.Ш. Шихалибейли, З.З. Султанова);
- Карта числа дней с градом в год (А.А. Мадатзаде);
- Карта зоны суховеев (Э. М. Шихлинский).

По отдельным территориям составлены многочисленные карты, относящиеся к оползням (Б.Будагов, Н.Керимов), некоторым климатическим явлениям (Э. М. Шихлинский, А. А. Мадатзаде), сейсмическим явлениям (Э.Ш. Шихалибейли, З.З. Султанова) [1].

Однако, в целом карт, посвященных СБ в Азербайджане, недостаточно, и они в полной мере не обеспечивают нужды государства. Табличные данные обычно составлялись исходя из собранных данных, в некоторых источниках применялось обыкновенное представление текста с форматированием основных параметров и расстановкой абзацев [2].

Инструментальные исследования сейсмичности в Азербайджане ведутся с 50-х годов прошлого столетия. С 1996-го г. исследования проводились цифровыми телеметрическими системами ISS южноафриканского производства. С целью изучения активных тектонических процессов, в 2003-ем г. в Институте геологии НАНА, в рамках помощи Фонда Развития и Гражданских исследований, создан Центр сейсмического мониторинга и геодинамических исследований. В настоящее время на территории Азербайджана землетрясения регистрируются 14 аналоговыми сейсмическими станциями. С 2003-го г., согласно указу общенационального лидера Гейдара Алиева, мониторинг сейсмичности в Республиканском Центре Сейсмологической Службы РЦСС проводится с помощью цифровой системы Kinemetriks, состоящей из 14 станций. Передача информации осуществляется в режиме реального времени с использованием космических средств связи.

С момента создания в 1999 г. мировым сейсмологическим сообществом современной Карты глобальной сейсмической опасности (Global Seismic Hazard Assessment Program - GSHAP), в Азербайджане проведены исследования в области геодинамики и расширения сейсмической сети. Эти работы более интенсивно развивались после землетрясения, произошедшего 25 ноября 2000 г. с M= 6.4 - в 60 км юго-восточнее Баку - в Каспийском море.

Анализ состояния изученности активных разломов и сейсмической опасности для территории Азербайджана показал, что до сегодняшнего дня не

составлены геоинформационные электронные карты активных разломов и сейсмической опасности, а также не использована новая информация о сейсмичности и геодинамическом режиме [4].

Источники данных. В последнее время делается упор на глобализацию собранных данных в интернете, создаются веб порталы для доступа к открытым данным для последующего анализа и визуализации. Эта тенденция будет продолжаться и набирает обороты. Уже к 2016-году количество перенесённых данных в веб пространство увеличилось на 35% по сравнению с предыдущим годом. Для решения нашей задачи мы будем рассматривать несколько источников данных о землетрясениях, зарегистрированных на территории Азербайджана и на Каспийском море.

Одной из важнейших задач – это определение критериев для самих данных и способов из чтения. Для землетрясений важными параметрами визуализации и анализа являются время, координаты, глубина и магнитуда. Также, необходимо учесть и способ чтения самих данных: это формат текстового представления с разделителями полей, форматы JSON, GeoJSON. Эти форматы являются стандартом де-факто при передачи данных из источника (базы данных) получателю (программы для визуализации, анализа) используя интернет соединение. Таким образом, мы определили критерии необходимые для анализа:

- Данные: время, координаты, глубина и магнитуда;
- Территория: Азербайджан и часть Каспийского моря;
- Формат: текстовая, с разделителями полей;
- Режим выгрузки: автоматическая, без разделения на страницы, с поддержкой веб -запросов сервер.

Республиканский Центр Сейсмологической Службы при Национальной Академии Наук Азербайджана создало систему поиска и мониторинга землетрясений. Данные поступают в центр из различных станций, обрабатываются и размещаются на веб странице по запросу пользователя в интерактивном режиме. По умолчанию, показываются данные за две недели,

разделённые на страницы. Удобство этой системы в том, что страница не нагружена дополнительной информацией, есть карта, показывающая эпицентры и сейсмологические станции. Однако, неудобство в отображении результатов (разделение на станицы) и последующая обработка после загрузки усложняют процесс автоматизации обработки. Также, отсутствует режим выгрузки в текстовый формат. Интерфейс подбора параметров фильтра не позволяет выбрать все источники на территории Азербайджана и Каспийского моря одновременно.

Большинство сайтов с интерактивной картой землетрясений утилизируют данные из сервисов компаний Google, EMSC и USGS. После детальных рассмотрений этих сервисов, USGS удовлетворил наши критерии для загрузки данных. [5]

Более того, сервис USGS позволяет выбрать прямоугольный регион на карте, таким образом решая задачу выбора территории. Т.к. границы территорий стран не являются прямоугольными, то при выборе региона, невольно будут попадать данные из соседних стран. Задачу фильтрации по территории можно решить в системе Power BI после выгрузки данных.Следующий веб запрос позволяет выгрузить историю землетрясений с сервиса USGS начиная с 1900 года по текущее время, регион Азербайджан и Каспийское море, минимальная магнитуда 2.5.

Результатом запроса будет текстовая информация с разделителями для полей и отсортированных по времени возникновения землетрясений.

Как уже было сказано выше, система Power BI позволяет обрабатывать результаты веб запросов и переводить данные в табличную форму, удобную для последующего анализа и отображения на карте. Для анализа воздействия СБ необходимо определиться с объектами и их метаданными с пространственной привязкой. Следующие объекты можно отнести к классу рисковых, имеющие важное, государственное значение.

- природные объекты (лесной фонд);
- производственные объекты;

- непроизводственные объекты;
- социально значимые объекты;
- объекты сельского хозяйства;
- транспортная сеть и инфраструктура.

Первое из требований по логическому согласованию (соответствие названий) должно обеспечиваться в процессе создания данных, а далее могут быть сгенерированы после создания путем применения специализированных процедур обработки.

Цифровое семантически-атрибутивное описание объектов осуществляется посредством использования кодов объектов, кодов характера их локализации, кодов характеристик, указанных в классификаторе для каждого вида объектов, а также значений

или кодов значений характеристик. При этом особая роль при цифровом описании объектов принадлежит характеристикам.

По назначению характеристики подразделяются на качественные и количественные. К ним относятся характеристики, содержащие информацию о свойствах объектов. По составу содержательной части характеристики подразделяются на числовые и символьные. В содержательной части числовой характеристики может быть только одно число (целое или вещественное, положительное или отрицательное). Для числовых характеристик (кроме представляемых в символьном виде) в качестве разделителя между целой и дробной частями используется "." (точка). Содержательная часть символьной характеристики может состоять из любых символов. Содержательной частью количественных характеристик являются их количественные значения для соответствующих объектов, выражающиеся вещественными или целыми числами [3].

Каждый из приведённых классов объект можно охарактеризовать как отдельный слой данных с уникальным идентификационным номером в одной единой таблице. Использование единой таблицы позволяет фильтровать

данные для обработки по их классу или же произвести анализ по нескольким, объединив их по гетерогенности.

Например, объекты нефтегазовой промышленности, станции по перекачки нефти и газа, наземные узловые элементы магистральных трубопроводов, хранилища нефтепродуктов. Задав атрибут гетерогенности, можно получить связанную, логически непрерывную (семантически) и согласованный объект. Дополнительные атрибуты отдельных объектов можно перенести в логические таблицы, таким образом нормализовать связи между логическими объектами риска и их дополнительными параметрами. Например, такие объекты как магистральные трубопроводы и наземные узловые элементы (вентили, клапан, насосы) относятся к одному типу глобального объекта - трубопровод, однако отдельные атрибуты могут отличаться. Поэтому целесообразно перенести эти данные в отдельные таблицы. Ключевым элементами привязки атрибутивных таблиц и логических элементов является уникальные идентификационные поля в обеих таблицах. Таким образом достигается гранулярность данных с взаимосвязями для гибкого анализа данных в системе Power BI.

Организация данных. Широкие возможности одновременного подключения к различным источникам данных даёт преимущество Power BI перед другими продуктами. Например, в системе ArcMap Desktop компании ESRI возможно подключение к геобазам, векторным, текстовым файлам, а также к электронным таблицам Microsoft Excel и баз данных Microsoft Access. Однако подключение к онлайн данным является проблематичным, требует дополнительного программного пакета и его настройки (ArcGIS Server). После того, как будут организованы пространственные данные объектов риска, нужно будет перенести их в один из общедоступных форматов.

Выбор формата зависит от имеющихся на данный момент у исследователя программных решений по хранению данных. Следующие критерии предъявляются для хранения данных:

- источник данных должен быть обновляемым;

- иметь табличный формат;
- основные атрибуты согласованы с пространственными данными;
- как менее один атрибут должен быть привязан с другими логическими категориями;
- возможность редактирования и добавления данных.

Приведём несколько практических решений удовлетворяющих выше предложенные критерии:

1. Microsoft Excel;
2. Microsoft Access;
3. Системы управления базами данных;
4. Онлайн хранилища, облачные сервисы.

Первые два решения имеют преимущества ввода и обработки данных пользовательских машинах и минимальные возможности по разделению данных между пользователями по сети.

Третье решение является универсальным решением, давая доступ ко многим пользователям. Это решение имеет преимущество в организациях со внутренней, закрытой сетью, где данные доступны исследователям внутри одной организации, вне зависимости от месторасположения. Однако необходимо учесть, то что это только система хранения, для обновления/изменения данных потребуется пользовательский программный интерфейс для подключения для работы с базами данных в административном режиме.

Четвёртое решение схожи с третьим, только с той разницей, что данные полностью децентрализованы и к ним можно получить доступ вне зависимости от месторасположения и места работы исследователя, таким образом давая возможность мобильного анализа данных. Так же как и в третьем решении, здесь также понадобиться пользовательский интерфейс для изменения данных.Изменение данных является прерогативой администратор.

Здесь необходимо учитывать делегацию ролей в управлении и разделении данными - это пользователи и администраторы. Только при таком разделении ролей можно будет добиться оптимального управления и организации непрерывных данных.

Теперь, когда определены источники данных, можно легко соединить и связать все данные воедино и отобразить результаты в панели отчётов в системе Power BI. (рис.1)

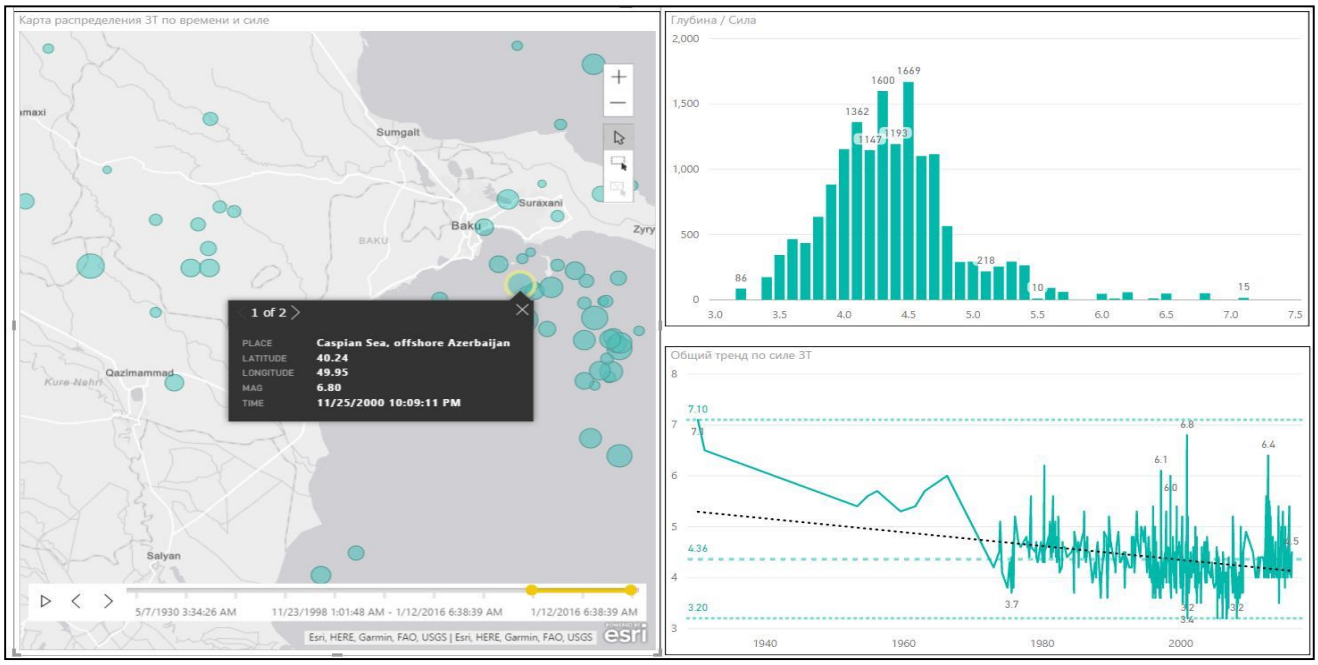


Рис. 1. Интерактивный панель отчёта землетрясений в системе PowerBI

Выводы. Систему подготовки отчётов и анализа данных PowerBI можно смело использовать при решении ГИС задач, так, переход из информации к действиям происходит в считанные минуты. Одним из важных особенностей можно отметить возможность подключения к любым источникам данных, отображение географических карт, кластерный и динамический анализ. Соединив источники данных в реальном режиме в купе с пространственными таблицами можно получить мощный инструмент для создания отчётов и опубликовать их для пользователей своей организации, что в свою очередь позволяет оперативно реагировать на происходящие процессы, анализировать исторические стихийные бедствия и принимать важные решения.

Литература

1. Бабаханов Н.А. Стихийные Бедствия: Их проявления в Азербайджане, наносимые ущербы, меры защиты от них. Баку: Изд «Элм», 2013, с.70-73.
2. Заида Султанова, "Землетрясения Азербайджана", 1969. - с. 3-75.
3. Руководство по региональной оценке риска стихийных бедствий на территории Республики Таджикистан. с. 48-50, 2011http://untj.net/files/Publications/DRMP/resource_page/RA_Methodology_RUSS.pdf
4. Сейсмичность Азербайджана, Кадиров Ф.А., Агаева С. https://www.gia.az/view.php?lang=ru&menu=45&id=575
5. USGS,"Earth Quakes database" https://earthquake.usgs.gov/earthquakes/search/