

## **ИНТЕЛЛЕКТУАЛЬНАЯ ГЕОИНФОРМАЦИОННАЯ СИСТЕМА АРХЕОЛОГИЧЕСКИХ ОБЪЕКТОВ**

**С.В. Белим**

профессор, д.ф.-м.н., e-mail: sbelim@mail.ru

**Д.М. Бречка**

доцент, к.т.н., e-mail: dbrechkawork@yandex.ru

**Т.А. Горбунова**

доцент, к.и.н., e-mail: GorbunovaTA@omsu.ru

**И.Б. Ларионов**

доцент, к.т.н., e-mail: LarionovIB@omsu.ru

**И.В. Шмидт**

доцент, к.и.н., e-mail: ShmidtIV@omsu.ru

Омский государственный университет им. Ф.М. Достоевского

**Аннотация.** Приведено описание геоинформационной системы археологических объектов ArGIS. Выделены объекты и их свойства. На основе свойств археологических объектов построено пространство признаков. К археологическим объектам как точкам пространства признаков применены методы кластеризации и методы построения ассоциативных правил. Реализованы модули визуализации результатов обработки данных. По результатам обработки данных происходит также выгрузка данных в текстовом виде. Предложенная система позволяет автоматически выявлять закономерности в расположении и датировке археологических фактов.

**Ключевые слова:** геоинформационная система, археологические объекты, Data Mining.

### **Введение**

В археологии объекты исследования — археологические памятники и подъёмный материал — имеют территориальную привязку. Поэтому одним из основных методов исследования является метод картографии. Ещё недавно археологические карты составлялись вручную или с использованием таких графических редакторов как CorelDraw и Photoshop. В настоящее время в археологии все большее распространение получают географические информационные системы (ГИС), которые стали революционной технологией, выходящей за пределы простого составления карт благодаря своему мощному аналитическому инструментарию. Это комплексные системы для сбора, хранения, поиска, анализа и представления пространственных данных любого типа [1]. ГИС позволяют обрабатывать большие массивы данных, применять методы не инвазивного археологического исследования с использованием данных дистанционного зондирования, устанавливать взаимосвязи между археологическими памятниками

и окружающим ландшафтом, прогнозировать обнаружение новых археологических памятников на основе пространственного анализа и многое другое. Кроме того, благодаря ГИС появляется возможность эффективного менеджмента историко-культурного наследия.

Археологические исследования с применением ГИС являются относительно молодым направлением в археологии. Наиболее активно они развиваются в течение последних 20–25 лет, и за это время существенно расширилась тематика исследований в данной области, выросло их количество, сформировалось два основных направления применения ГИС в археологии. Первое может быть обозначено как «управление культурным наследием» («Cultural resource management», CRM), второе — как археология ландшафта («Landscape archaeology»).

Широкое применение теоретические и методические разработки геоинформационного подхода к археологическим исследованиям нашли в региональных археологических проектах, большинство которых сфокусировано на реконструкциях и анализе древних культур и ландшафтов или же территориальном распределении отдельных культурных феноменов, например, образцов древнего искусства [2]. В качестве примеров использования ГИС в археологии можно также отметить такие проекты, как АИС «Археограф», «Археологические памятники Красноярского края» [3], «Археологические памятники Кисловодской котловины» [4, 5], «Палеолит Северной Азии» [6].

Широкое распространение ГИС получили в археологических исследованиях в Германии [7–9], Великобритании [10], Италии [11, 12], Испании [13], Греции [14] и других странах.

Мировыми лидерами в области археологических ГИС сегодня являются ArcGIS от ESRI (США), MapInfo от Pitney Bowes Business Insight (США), GeoMedia от Intergraph (США), Bentley MicroStation от Bentley Systems (США). Среди российских производителей конкуренцию им могут составить продукты компании «Интегро» (Уфа) — ИнГео и Центра Геоинформационных Исследований Института Географии РАН — ГеоГраф (GeoGraph). Большое распространение получают открытые ГИС, такие как ILWIS (Голландия), MapWindow GIS (США), GRASS GIS (США) и, особенно Quantum GIS (международный проект) и SAGA (Германия). Однако большинство открытых ГИС требуют серьёзной доработки для использования в археологических целях.

Однако на сегодняшний день практически отсутствуют работы по интеллектуальному анализу данных, размещённых в ГИС. Тем не менее использование методов анализа данных в качестве поддержки археологических исследований позволит выявить новые закономерности, которые окажут помощь в исторических исследованиях.

## 1. Описание ArcGIS

Объектами геоинформационной системы являются археологические памятники. Каждый памятник характеризуется набором атрибутов: координаты многоугольника, ограничивающего контуры археологического объекта; наимено-

вание археологического объекта; список работ, выполнявшихся на объекте; документы с результатами выполненных работ; классификационные признаки объекта; степень исследованности объекта. В списке работ, выполнявшихся на объекте, указывается год проведения работ, руководитель работ и виды выполненных работ. Классификационные признаки включают в себя эпоху, к которой относится памятник, вид объекта и др.

Геоинформационная система включает в себя как достаточно простой функционал просмотра и редактирования информации, так и возможность выявления закономерностей методами Data Mining. К простому функционалу можно отнести функции поиска объектов по координатам, степени исследованности, наименованию и классификационным признакам. В качестве дополнительной функции поиска возможна фильтрация объектов в выделенной географической области.

Добавление и редактирование объектов возможно в двух режимах. В стационарном режиме выполняется заполнение полей формы. В мобильном режиме происходит автоматическая привязка по координатам. Также в мобильном режиме возможна загрузка фотографий с мобильного устройства.

Модуль обработки карт даёт возможность привязки как распространённых карт (Bingo, Google, Yandex), так и исторических карт. При загрузке исторических реализована возможность сопоставления с другими картами и взаимная корреляция объектов.

## 2. Построение ассоциативных правил

Для анализа взаимосвязи различных археологических объектов и их свойств используются методы построения ассоциативных правил. Наибольшее распространение метод построения ассоциативных правил получил в экономических науках. Ассоциативные правила позволяют выявлять взаимосвязь между объектами на основе статистических данных. Для построения ассоциативных правил необходимо выделить транзакции. В качестве одной транзакции будем использовать запись об археологическом объекте. Поля в записи об объекте будут играть роль товаров. Причём выбираются только те поля, которые могут быть одинаковыми для любого объекта:

1. Вид объекта.
2. Годы открытия и постановки на учёт.
3. Принадлежность объекта к эпохе.
4. Принадлежность объекта к культуре.
5. Степень разработанности.
6. Виды работ.
7. Результаты раскопок.

Для каждого набора значений  $F_i$  определяется уровень поддержки  $Supp(F_i)$ .

$$Supp(F_i) = \frac{D(F_i)}{N},$$

где  $D(F_i)$  — число транзакций, содержащих набор  $F_i$ ,  $N$  — общее количество транзакций. Из полученных наборов строим ассоциативные правила для каждого встречающегося значения признака  $c_j$  в наборе  $F_i$ :

$$\{c_1, c_2, \dots, c_{j-1}, c_{j+1}, \dots, c_k\} \rightarrow c_j.$$

Для каждого из ассоциативных правил необходимо рассчитать величину, называемую достоверностью:

$$\text{conf}(F_i, c_j) = D(F_i)/D(c_j), \quad (j = 1, \dots, k),$$

где  $D(F_i)$  — количество транзакций, содержащих набор  $F_i$ ,  $D(c_j)$  — число транзакций, содержащих значение  $c_j$ . Достоверность лежит в интервале от 0 до 1.

Оставляя ассоциативные правила с поддержкой не ниже некоторого значения  $\text{Supp}_{min}$ , выбираем наиболее часто встречающиеся сочетания признаков. Из оставленных ассоциативных правил отбираем то, которое имеет наибольшую достоверность.

В системе ArgGIS был реализован поиск ассоциативных правил с подмножеством различной длины. Длины наборов пользователь задаёт самостоятельно.

### 3. Кластеризация данных

Процедура кластеризации позволяет выделить группы объектов, близких по некоторому набору признаков. Для выполнения кластеризации необходимо выделить пространство признаков. Каждому объекту сопоставляется точка в пространстве признаков. Введём следующие координаты объектов:

1. Год выполнения работ ( $x_1$ ).
2. Вид объекта ( $x_2$ ).
3. Эпоха, к которой относится объект ( $x_3$ ).
4. Степень разработанности ( $x_4$ ).
5. Культура, к которой относится объект ( $x_5$ ).
6. Географические координаты ( $x_6, x_7$ ).

Для всех координат в пространстве признаков важным является не абсолютное значение, а степень близости с другими объектами. Год выполнения работ будем отсчитывать от некоторой начальной даты исследований на данной территории. В качестве значения координаты вида объекта будем использовать его номер в полном списке возможных видов объектов. На результатах кластеризации будет сказываться способ упорядочивания видов в полном списке. Чтобы снизить данное влияние, необходимо использовать метод классификации видов объектов, позволяющий проводить однозначное упорядочивание.

Эпохи, к которым относятся объекты, упорядочиваются по возрастанию. В качестве координаты  $x_3$  выбирается номер эпохи, к которой относится объект.

Степень разработанности численно определяется с помощью метода экспертных оценок. Проводится попарное сравнение степени разработанности объектов. После чего строится матрица парных сравнений и определяется удельный вес каждого объекта, который используется как значение координаты  $x_4$ .

Для характеристики культурной составляющей объекта используется тот же подход, что и для эпохи.

Географические координаты  $(x_6, x_7)$  задаются как широта и долгота центра масс многоугольника, задающего границы объекта.

Для устранения доминирования одной координаты над остальными необходимо провести масштабирование и привести интервалы значений всех координат к отрезку  $[0, 1]$ . Меру близости двух точек в пространстве координат будем вычислять как евклидово расстояние между ними.

Для кластеризации объектов используем метод построения минимального остовного дерева. Для этого сопоставим множеству археологических объектов полносвязный взвешенный граф. Вершинами графа служат точки в пространстве признаков. Вес ребра определяется расстоянием между точками. Кластеризация сводится к разбиению графа на подграфы. Для данного графа ищем минимальное остовное дерево, используя алгоритм Прима. Найдём среднее значение и дисперсию длин рёбер минимального остовного дерева. Разобьём дерево на поддеревья. Для этого удалим наиболее длинные ребра. Каждое поддерево будет отдельным кластером. Для выбора удаляемых ребер может быть использован один из двух подходов. В первом случае удаляем ребра, длина которых больше среднего значения более чем в  $p$  раз ( $p > 1$ ). Во втором случае удаляем  $k\%$  самых длинных рёбер ( $0 < k < 100$ ). Варьирование параметрами  $p$  и  $k$  позволяет менять размеры кластеров. Возможна кластеризация в любом подпространстве полного пространства признаков. В этом случае не учитывается влияние некоторых параметров.

Визуализация результатов кластеризации состоит в отображении на карте объектов, входящих в различные кластеры, разными цветами.

#### 4. Статистический анализ

При анализе исторических процессов на основе археологических объектов важную роль играет информация о распределении признаков объектов на карте. Наиболее важным фактором является частота встречаемости того или иного признака. Для нахождения распределения частот всю исследуемую территорию необходимо разбить на области.

Важным вопросом является выбор формы и размеров образуемых областей. Будем определять площадь областей в зависимости от общего количества археологических объектов  $N$ . Проведём разбиение так, чтобы в одну область попало в среднем  $n$  объектов. Таким образом, общее количество областей будет не менее  $N/n$ . Наложим на карту территории прямоугольную сетку. Найдём минимальное число  $m$  такое, что  $m^2 \geq N/n$ . Пусть максимальное расстояние с севера на юг между точками исследуемой территории равно  $d_y$ . Максимальное расстояние с запада на восток обозначим  $d_x$ . Вся территория может быть вписана в прямоугольник  $d_x \times d_y$ . Построим прямоугольную сетку с шагом по горизонтали  $d_x/m$ , а по вертикали  $d_y/m$ . Частоту встречаемости признаков археологических объектов будем вычислять по клеткам наложенной сетки.

Выбор параметра  $n$  осуществляет пользователь. Значение  $n$  влияет на размер ячеек сетки. Исследования могут быть реализованы методом последовательных приближений. Выбор большого  $n$  позволяет выделить крупные области, представляющие интерес. Уменьшая  $n$  можно получить более детальную информацию.

Визуализация результатов частотного анализа осуществляется с помощью цветовой раскраски сетки. Частотное распределение одного признака визуализируется градациями одного цвета. Палитра каждого цвета содержит  $m + 1$  оттенок в интервале  $[0, m]$ . Отображая одновременно частотное распределение нескольких признаков, можно визуально выявить зависимости между ними.

## 5. Заключение

Таким образом, представление базы археологических объектов в виде географической информационной системы позволяет не только визуализировать данные, но и проводить их автоматизированный анализ. Взаимосвязь между свойствами объектов выявляется с помощью ассоциативных правил. Дальнейшему ручному анализу подлежат ассоциативные правила с высоким уровнем достоверности. Такие правила позволяют обнаружить новые закономерности в хорошо разработанных объектах. Также появляется возможность прогнозировать новые результаты для недостаточно разработанных объектов. Кластеризация позволяет сгруппировать схожие объекты. Визуализация результатов кластеризации играет роль предварительной обработки информации для дальнейшего ручного анализа.

Разработанная ГИС вместе с модулями интеллектуального анализа данных представляет собой мощный инструмент поддержки археологических исследований.

Система тестировалась для археологических объектов Омской области. Были выявлены закономерности, связанные с процессом исследования объектов. Также удалось выявить ряд закономерностей расположения объектов, относящихся к различным эпохам и культурам.

## ЛИТЕРАТУРА

1. Gaffney V., Stancic Z. GIS Approaches to Regional Analysis: A Case Study of the Island of Hvar, Ljubljana, Yugoslavia. Znanstveni instituti Filozofske fakultete, 1991.
2. Engur — history and culture. URL: [http://engur.ru/ru\\\_RU/map-prehistoric-art](http://engur.ru/ru\_RU/map-prehistoric-art) (дата обращения: 13.06.2016).
3. Артемьев Е.В., Дроздов Н.И., Зайцев Н.К., Шапарев Н.Я., Якубайлик О.Э., Шахматов А.В. Создание геоинформационной системы «Археологические памятники Красноярского края» // Вычислительные технологии. 1998. Т. 3, № 5. С. 5–10.
4. Афанасьев Г.Е., Савенко С.Н., Коробов Д.С. Древности Кисловодской котловины. М. : Мир науки, 2004. 240 с.

5. Гарбузов Г.П. Археология ландшафта и геоинформатика: теоретические аспекты взаимоотношений [Электронный ресурс]. / Археология и геоинформатика. М. : АГИС, ИА РАН, 2007. Вып. 4. (CD-ROM).
6. Деревянко А.П., Холюшкин Ю.П., Бердников Е.В., Воронин В.Т. ГИС «Палеолит Северной Азии» // Информационные технологии в гуманитарных исследованиях. Вып. 6. Новосибирск : Ред.-изд. Центр НГУ. 2003. С. 21–29.
7. Geldener R., Kreienbrink F. Erfassung ortsgebundener Denkmale und Fundstellen // Ausgrabungen in Sachsen. 2010. No. 2. Arbeits- und Forschungsberichte zur sochsische Bodendenkmalpflege. Beiheft 21. P. 189–198.
8. Grunwald S. Das ergab aber ein so buntes wenig eindrucksvolles Bild. Zu den Anfängen der archeologischen Kartographie in Deutschland (1870–1914) // EAZ. 2012. No. 53. P. 5–34.
9. Stuble H. Braunkohlen- und Trassenarchaologie: eine Herausforderung mit Tradition // Ausgrabungen in Sachsen. 2010. No. 2. Arbeits- und Forschungsberichte zur sochsische Bodendenkmalpflege. Beiheft 21. P. 67–82.
10. Conolly J., Lake M. Geographical Information Systems in Archaeology. Cambridge Manuals in Archaeology. Cambridge : Cambridge University Press. 2006.
11. Moscati P. GIS application in Italian archeology // Archeologia e Calcolatori. 1998. No. 9. P. 191–236.
12. Tantillo M.D. GIS application in archaeological site of Solunto // XXI International CIPA Symposium, 01–06 October 2007, Athens, Greece.
13. Mejuto J., Castao J.G., Rodriguez-Caderot G. GIS Techniques in Archaeology: An Archaeoastronomical Approach // Archaeology, New Approaches in Theory and Techniques, Dr. Imma Ollich-Castanyer (Ed.), InTech, 2012.
14. Bevan A., Conolly J. GIS and Archaeological Survey Data: Four Case Studies in Landscape Archaeology from the Island of Kythera (Greece) // Postprint of a 2004 paper in Journal of Field Archaeology. 2004. No. 29. P. 123–138.

## INTELLECTUAL GEOGRAPHIC INFORMATION SYSTEM OF ARCHAEOLOGICAL OBJECTS

**S.V. Belim**

Dr.Sc.(Phys.-Math.), Professor, e-mail: sbelim@mail.ru

**D.M. Brechka**

Ph.D. (Eng.), Associate Professor, e-mail: dbrechkawork@yandex.ru

**T.A. Gorbunova**

Ph.D. (Hist.), Associate Professor, e-mail: GorbunovaTA@omsu.ru

**I.B. Larionov**

Ph.D. (Eng.), Associate Professor, e-mail: LarionovIB@omsu.ru

**I.V. Schmidt**

Ph.D. (Hist.), Associate Professor, e-mail: ShmidtIV@omsu.ru

Dostoevsky Omsk State University

**Abstract.** The description of the geographic information system of archaeological objects ArGIS is provided. The objects and their properties are selected. Based on

the properties of archaeological objects the feature space is built. Clustering methods and methods of construction of association rules are applied to the archaeological objects as points of the feature space. The modules of visualization of data processing results are implemented. According to the data processing results, the data unloading in a text format occurs. The proposed system can automatically identify patterns in the distribution and dating of archaeological facts.

**Keywords:** GIS, archaeological objects, Data Mining.