

УДК 303.732.4

UDC 303.732.4

РАЗВИТИЕ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ «ЭЙДОС-АСТРА» СНИМАЮЩЕЕ ОГРАНИЧЕНИЯ НА РАЗМЕРНОСТЬ БАЗ ЗНАНИЙ И РАЗРЕШЕНИЕ КОГНИТИВНЫХ ФУНКЦИЙ

DEVELOPMENT OF THE ARTIFICIAL INTELLIGENCE SYSTEM "AIDOS-ASTRA" WITHOUT DATA BASE DIMENSION RESTRICTIONS AND COGNITIVE FUNCTIONS RESOLUTION

Луценко Евгений Вениаминович
д.э.н., к.т.н., профессор
Кубанский государственный аграрный университет, Краснодар, Россия

Lutsenko Evgeny Veniaminovich
Dr. Sci.Econ., Cand. Tech.Sci., professor
Kuban State Agrarian University, Krasnodar, Russia

Трунев Александр Петрович
к. ф.-м. н., Ph.D.
Директор, A&E Trounev IT Consulting, Торонто, Канада

Alexander Trunev
Cand.Phys.-Math.Sci., Ph.D.
Director, A&E Trounev IT Consulting, Toronto, Canada

Трунев Евгений Александрович ,
Разработчик операционных систем, Торонто, Канада

Eugene Trounev,
Software developer, Toronto, Canada

В статье рассматривается новая версия системы искусственного интеллекта «Эйдос-астра» для решения прикладных задач с эмпирическими данными большой размерности. Приложение, написанное на языке JAVA, обеспечивает GUI и позволяет подготовить и выполнить визуализацию матрицы информативностей без ограничений, налаженных реализацией предыдущих версий системы «Эйдос-астра».

The article discusses the expansion of artificial intelligence "Aidos-Astra" for applications with the empirical data of high dimensionality. Application, written in the language JAVA, allows you to prepare and visualize the information content of the matrix without restrictions imposed by the architecture of the system "Aidos-Astra "

Ключевые слова: АВТОМАТИЗИРОВАННЫЙ СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ, БАЗА ЗНАНИЙ, КОГНИТИВНАЯ ФУНКЦИЯ, СЕМАНТИЧЕСКАЯ ИНФОРМАЦИОННАЯ МОДЕЛЬ

Keywords: COMPUTERIZED SYSTEM-COGNITIVE ANALYSIS, KNOWLEDGE BASE, COGNITIVE FUNCTION, SEMANTIC INFORMATION MODEL

Краткая предыстория исследований

Автоматизированный системно-когнитивный анализ /1/ и его инструментарий – базовая система «Эйдос» /2/ получили широкое применение в решении ряда задач, связанных выявлением знаний из эмпирических данных большой размерности и решением на их основе задач идентификации, прогнозирования, принятия решений и исследования предметной области. Как известно из теории управления поведение *любой* системы определяется состоянием самой системы (внутренними факторами и ее предысторией), управляющими (технологическими) факторами, а также воздействием окружающей среды. Не являются исключением и глобальные геосистемы: ноосфера (глобальные социально-экономические процессы), биосфера, геосфера, магнитосфера, атмосфера и другие, для которых внешней средой, с которой эти системы входят в *непосредственное* взаимодействие, является ближайшее космическое окружение нашей планеты, т.е. различные космические тела, входящие в состав Солнечной системы. Задачи математического моделирования влияния небесных тел Солнечной системы на глобальные геосистемы отличаются большой размерностью как по ис-

ходным базам данных, так и по создаваемым в процессе моделирования информационным базам и базам знаний..

Для решения этих задач была создана система «Эйдос-астра» /3-4/, позволяющая *автоматически генерировать различные системы частных моделей (мультимодели), исследовать их на достоверность в разрезе по классам и с учетом этой информации наиболее эффективно использовать их для идентификации и прогнозирования*. Первоначально исследования авторов были посвящены выявлению и исследованию зависимости социального статуса респондентов (которых было около 30 тысяч) от состояния космической среды на момент их рождения /16, 17, 18, 19/. Со временем проблематика исследований расширилась и преобразовалась в целое новое научное направление: *«Применение технологий искусственного интеллекта для исследования влияния космической среды на поведение активных глобальных геосистем»* /5/.

В качестве современной автоматизированной интеллектуальной технологии авторы применили автоматизированный системно-когнитивный анализ (АСК-анализ) и его программный инструментарий – универсальную когнитивную аналитическую систему "Эйдос" /1-2/. В состав этой системы входит подсистема _15, содержащая ряд стандартных программных интерфейсов с внешними источниками данных различных стандартов: текстовых, баз данных (БД) и графических, расширяющих сферу ее применения. Некоторые из подобных интерфейсов при своем развитии преобразовались в системы окружения, которые выполняли больше функций, чем программный интерфейс _152, но использовались совместно с базовой системой «Эйдос». Первоначально авторами исследования проводились с применением именно этого программного интерфейса с исходными базами данных и базовой системы «Эйдос». При этом было разработано много программных интерфейсов с различными внешними базами исходных данных, создано и исследовано большое количество различных моделей. Со временем процесс генерации системы моделей, их исследования на достоверность в разрезе по категориям, а также их использования для решения задач идентификации, был автоматизирован, что и привело к созданию системы окружения «Эйдос-астра» [5], которая включила в себя также все ранее специализированные программные интерфейсы и специальный модуль визуализации когнитивных функций [27].

В настоящее время система «Эйдос-астра» приобрела функциональную полноту и превратилась в самостоятельный инструмент исследования, который может применяться как совместно с базовой системой «Эйдос», так и независимо от нее.

Постановка проблемы, решаемой в работе

Базовая система «Эйдос» [31] была создана еще до создания операционной системы Windows (в 1994 году на нее уже было получено 3 патента). В то время трудно было себе представить, что когда-то в будущем

может потребоваться более 4000 классов распознавания и 4000 признаков и были приняты технические решения, связанные с повышением быстродействия путем организации внутреннего кэширования, которые привели к соответствующему ограничению на размерности создаваемых моделей. Когда создавалась система «Эйдос-астра», то при этом очень широко использовались библиотеки процедур и функций системы «Эйдос» (и не очень много внимания уделялось интерфейсу), что привело к наличию в системе «Эйдос-астра» тех же ограничений на размерности моделей, что и в системе «Эйдос». Эти ограничения создали определенные трудности в анализе задач большой размерности, связанных, например, с анализом влияния небесных тел на сейсмические события, движение полюса Земли и магнитное поле [7-11].

Справедливости ради необходимо отметить, что первоначально система «Эйдос-астра» разрабатывалась как инструмент для проверки научных гипотез о существовании зависимостей между параметрами космической среды и глобальными процессами ноосфере Земли и этой цели она вполне соответствовала. Кроме того указанные выше недостатки базовой системы «Эйдос» и системы «Эйдос-астра» легко преодолеваются (без переписывания исходного текста) путем использования новых версий языка программирования, на котором они были написаны (Alaska xBase++, Arctica).

Идея и путь решения поставленной проблемы

Идея решения сформулированной проблемы состоит в использовании новых перспективных средств разработки новой версии системы «Эйдос-астра».

При этом использование различных инструментальных средств имеет свои плюсы и минусы.

В частности использование Alaska xBase++ или Arctica позволяет максимально использовать уже существующий исходный текст базовой системы «Эйдос» и систем окружения и за счет этого существенно сократить трудоемкость и длительность разработки, а это очень важно, т.к. распечатка этого исходного текста 6-м шрифтом составляет около 800 страниц. Отметим, что базовая система «Эйдос» реализована в *универсальной постановке, не зависящей от предметной области*, а ее привязка к конкретным областям осуществляется на уровне адаптации и пересинтеза приложений, а также с использованием полнофункционального набора программных интерфейсов с внешними базами данных и различных систем окружения [1, 5, 31].

В качестве *пути* решения сформулированной проблемы авторами принято решение о разработке новой версии системы «Эйдос-астра» с использованием языка Java.

Java¹ – это чрезвычайно перспективный язык программирования, использование которого для данной цели вполне оправданно и обоснованно не смотря на то, трудоемкость разработки выше, чем на Alaska xBase++ или Arctica, т.к. приходится не переносить существующий исходный текст, а писать исходный текст практически «с нуля».

Принято решение о реализации в новой версии данной системы всех функций существующей версии, но без ограничений на размерность моделей и в GUI, а затем о реализации в ней качественно новых возможностей, принципиально нереализуемых с применением ранее использовавшегося инструментария. По сути, идет речь не о переносе системы «Эйдос-астра» в существующей постановке, но с преодолением ограничений на размерность моделей и недостатков интерфейса, а о разработке качественно новой версии этой системы с сохранением всего лучшего, что было достигнуто в предыдущих версиях.

Первые версии системы «Эйдос-Java»² уже созданы и тщательно протестированы на задачах, ранее решенных с помощью предыдущих версий системы «Эйдос-астра». Было продемонстрировано полное тождество полученных результатов, а также получены качественно новые результаты, ради которых и создавалась система. В частности были сняты ограничения на размерность модели с M173 до M36000 и по числу входных параметров с 23 до 253. Отметим, что указанные новые ограничения параметров являются условными и определяются только типом выбранных форматов файлов входных данных и оперативной памятью. Расширение системы выполнено в виде отдельного приложения ViewVibe.jar /12/, написанного на языке Java, снабженного блоком графической визуализации расчетных матриц, включая матрицу информативностей. Приложение ViewVibe.jar ориентировано на решение естественно-научных, социально-экономических и психологических задач, связанных с распознаванием событий по астрономическим данным, описанных в работах /7-11, 13-22/ и других.

Далее рассмотрим некоторые **результаты исследований**, проведенных с применением системы «Эйдос-Java».

Матрица информативностей в задаче о распознавании категорий событий в поле центральных сил

Рассмотрим задачу распознавания категорий по астрономическим данным /5, 7-11/. Имеется множество событий E, которому ставится в соответствие множество категорий K_i. Событиями можно считать, например, землетрясения, происходящие на нашей планете ежедневно, а категориями – число одно- (A), двух- (B) или трехкратных событий (C), магнитуда которых лежит в заданном интервале. Каждое такое событие характеризуется моментом времени и географическими координатами места его происхож-

¹ http://java.com/ru/download/faq/whatis_java.xml

² примем это *рабочее* название для Java-версии системы «Эйдос-астра»

дения. По этим данным можно построить матрицу, содержащую координаты небесных тел, например углы долготы, широты и расстояния. Будем считать, что заданы частотные распределения N_i – число событий, имеющих отношение к данной категории K_i .

Определим число случаев реализации данной категории, которое приходится на заданный интервал изменения астрономических параметров, имеем в дискретном случае:

$$N_{ij}(x_j, k) = N_i w(\tilde{x}_j, k) \Delta x, \quad x_j < \tilde{x}_j < x_j + \Delta x$$

$$1 \leq i \leq n, \quad 1 \leq j \leq 2m, \quad k = 1, \dots, k_0 \tag{1}$$

Здесь w – плотность распределения событий вдоль нормированной координаты. Нормированная переменная определяется через угловую и радиальную координаты следующим образом:

$$x_{jk} = \begin{cases} J_j(k) / 2p, & 1 \leq j \leq m \\ \frac{r(k) - r_{\min}(k)}{r_{\max}(k) - r_{\min}(k)}, & m + 1 \leq j \leq 2m \end{cases}$$

где r_{\min}, r_{\max} – минимальное и максимальное удаление планеты от центра масс системы, k_0 – число небесных тел, используемых в задаче.

Определим матрицу информативностей согласно /1/

$$I_{ijk} = \log_2 \frac{N_{ij} / \sum_j N_{ij}}{\sum_i N_{ij} / \sum_{i,j} N_{ij}}, \quad N_{ij}(x_{jk}) \neq 0$$

$$I_{ijk} = 0, \quad N_{ij}(x_{jk}) = 0,$$

$$dI_{jk} = \sqrt{\frac{1}{n} \sum_i \left(I_{ijk} - \frac{1}{n} \sum_i I_{ijk} \right)^2}$$

$$1 \leq i \leq n, \quad 1 \leq j \leq 2m, \quad 1 \leq k \leq k_0 \tag{2}$$

Первая величина (2) называется информативность признака, а вторая величина является стандартным отклонением информативности или интегральной информативностью (ИИ).

Рассмотрим связь информативности со статистикой хи-квадрат³

Статистика χ^2 представляет собой сумму вероятностей совместного наблюдения признаков и объектов по всей корреляционной матрице или

³ См.: <http://www.machinelearning.ru/wiki/index.php?title=Таблица%20сопряженности>

определенным ее подматрицам (т.е. сумму относительных отклонений частот совместного наблюдения признаков и объектов от среднего) [2, п.3.4]⁴:

$$C^2 = \sum_{j=1}^W \sum_{i=1}^M \frac{(N_{ij} - t)^2}{t} \quad (3)$$

где:

– N_{ij} – фактическое количество встреч i -го признака у объектов j -го класса;
 – t – теоретически ожидаемое количество встреч i -го признака у объектов j -го класса.

$$t = \frac{N_i N_j}{N} \quad (4)$$

Нельзя не обратить внимание на то, что *статистика C^2 математически простым образом связана с количеством информации* в системе признаков о классе распознавания, в соответствии с системным обобщением формулы Харкевича для плотности информации (5):

$$I_{ij} = \text{Log}_2 \left(\frac{N_{ij} N}{N_i N_j} \right)^\Psi \quad (5)$$

а именно из (4) и (5) получаем:

$$I_{ij} = \text{Log}_2 \left(\frac{N_{ij}}{t} \right)^\Psi \quad (6)$$

Выражение (6) для количественной меры знаний, полученное в и используемое в автоматизированном системно-когнитивном анализе (АСК-анализ), является чрезвычайно важным. По сути, его смысл в том, что количественная мера знаний представляет собой результат *сравнения* фактически наблюдаемой абсолютной частоты N_{ij} встреч i -го признака у объектов j -го класса с теоретически ожидаемой частотой его встреч t (4) *в соответствии* с анализом хи-квадрат, которая и выступает в качестве нормы или базы сравнения. Это сравнение в выражении (6) осуществляется путем вычисления *отношения* фактически наблюдаемой абсолютной частоты N_{ij} встреч i -го признака у объектов j -го класса с теоретически ожидаемой частотой его встреч t . Если при использовании критерия хи-квадрат в качестве фильтра предлагается не использовать для дальнейшего анализа фактически наблюдаемые частоты, меньшие теоретически ожидаемых как недос-

⁴ Отметим, что в базовой системе «Эйдос» [31] имеются развитие средства анализа хи-квадрат.

товерные, то в АСК-анализе эти значения *используются* и интерпретируются совершенно иначе: не как недостоверные, а как достоверно отражающие *отрицательную причинно-следственную зависимость*⁵, т.к. дают, как легко видеть из выражения (6) отрицательное количество информации: $I_{ij} < 0$. Поэтому применение фильтра хи-квадрат для отсеивания «недостоверной» информации эквивалентно исключению отрицательных информативностей из базы знаний. Соответственно исчезает возможность построения негативных когнитивных функций [30]. Фактически применяемая в АСК-анализе количественная мера знаний детально показывает как образуется значение критерия хи-квадрат.

Из (6) очевидно:

$$I_{ij} = \Psi(\text{Log}_2 N_{ij} - \text{Log}_2 t) \quad (7)$$

Сравнивая выражения (3) и (7), видим, что числитель в выражении (3) под знаком суммы отличается от выражения (7) только тем, что в выражении (7) вместо значений N_{ij} и t взяты их *логарифмы*, т.е. по сути, *отличаются только единицей измерения*. Поскольку логарифм является монотонно возрастающей функцией аргумента, то введение логарифма не меняет общего характера поведения функции.

Фактически это означает, что:

$$\begin{cases} \text{если } N_{ij} < t \text{ то } c_{ij}^2 > 0, I_{ij} < 0 \\ \text{если } N_{ij} = t \text{ то } c_{ij}^2 = 0, I_{ij} = 0 \\ \text{если } N_{ij} > t \text{ то } c_{ij}^2 > 0, I_{ij} > 0 \end{cases} \quad (8)$$

Если фактическая вероятность наблюдения i -го признака при предъявлении объекта j -го класса равна теоретически ожидаемой (средней), то наблюдение этого признака не несет никакой информации о принадлежности объекта к данному классу. Если же она выше средней – то это говорит в пользу того, что предъявлен объект данного класса, если же ниже – то другого.

Поэтому *наличие статистической связи (информации)* между признаками и классами распознавания, т.е. отличие вероятностей их со-

⁵ Когда i -е значение фактора не способствует, а препятствует переходу объекта в j -е состояние. Отметим, что математическая форма выражения (3) (а именно: *квадрат*) *предусматривает* и учет фактических абсолютных частот, меньших теоретических, как отражающих причинно-следственные связи. Тем ни менее традиционная интерпретация предполагает достоверными только фактические абсолютные частоты, большие теоретических

вместных наблюдений от предсказываемого в соответствии со случайным нормальным распределением, *приводит к увеличению фактической статистики s^2 по сравнению с теоретической величиной.*

Таким образом, применяемая в автоматизированном системно-когнитивном анализе количественная мера силы и направления причинно-следственной связи между факторами и поведением объекта (его переходами в состояния, соответствующие классам), т.е. информативность или количественная мера знаний (2) тесно связана с классическим критерием *наличия* статистически значимой связи хи-квадрат и можно обоснованно утверждать, что она не только не противоречит ему, но и в определенной степени основана на нем и является его развитием. Однако информативность представляется более развитым критерием, т.к. отражает не только *наличие* и силу связи, как хи-квадрат (8), но и ее *знак*, и *главное – ее форму, т.е. сам вид зависимости* [30]. Кроме того, использование в качестве количественной меры силы и направления связи количества информации позволяет привлечь хорошо развитые представления теории информации для исследования причинно-следственных зависимостей.

Связь между критерием хи-квадрат и когнитивными функциями

Частично-редуцированные когнитивные функции [30] строятся по максимуму информативности, т.е. по сути, по наиболее достоверным данным, если для оценки их достоверности использовать критерий хи-квадрат. Поэтому можно обоснованно ожидать, что использование критерия хи-квадрат в качестве фильтра достоверных данных даст результаты, сходные с результатами, приведенными в работе [30], однако приведет к потере существенной информации об отрицательных причинно-следственных зависимостях, визуализируемых в виде негативных когнитивных функций (уменьшится контрастность их визуализации).

Из вышесказанного следует возможность использования в *качестве количественной меры степени выраженности закономерностей в предметной области* использовать не матрицу абсолютных частот и меру χ^2 , а новую меру, основанную на *матрице информативностей и системном обобщении формулы Харкевича для количества информации:*

$$H = \sqrt[2]{\frac{1}{(M \cdot W - 1)} \sum_{j=1}^W \sum_{i=1}^M (I_{ij} - \bar{I})^2} \quad (9)$$

где:

$$\bar{I} = \frac{1}{M \cdot W} \sum_{j=1}^W \sum_{i=1}^M I_{ij} \quad \text{– средняя информативность признаков по матрице информативностей.}$$

Значение данной меры показывает среднее отличие количества информации в факторах о будущих состояниях активного объекта управления от среднего количества информации в факторе (которое при больших выборках близко к 0). По своей математической форме эта мера сходна с мерами для значимости факторов и степени сформированности образов классов и коррелирует с *объемом* когнитивного пространства классов и пространства атрибутов.

Описанная выше математическая модель обеспечивает инвариантность результатов обучения Системы относительно следующих параметров обучающей выборки:

1. Суммарное количество и порядок ввода анкет обучающей выборки.
2. Количество анкет обучающей выборки по каждому классу распознавания.
3. Суммарное количество признаков во всех анкетах обучающей выборки.
4. Суммарное количество признаков по эталонным описаниям различных классов распознавания.
5. Количество признаков и их порядок в отдельных анкетах обучающей выборки.

Это обеспечивает высокое качество решения задач системой распознавания на неполных и разнородных (в вышперечисленных аспектах) данных как обучающей, так и распознаваемой выборки, т.е. при таких статистических характеристиках потоков этих данных, которые чаще всего и встречается на практике и которыми невозможно или очень сложно управлять.

Каждой категории можно сопоставить вектор информативности астрономических параметров размерности $2mk_0$, составленный из элементов матрицы информативности, путем последовательной записи столбцов, соответствующих нормированной координате, в один столбец, т.е.

$$c_{is} = I_{ijk} \Big|_{jk=s}, \quad 1 \leq s \leq 2mk_0 \quad (10)$$

С другой стороны, процесс идентификации, распознавания и прогнозирования может рассматриваться как разложение вектора распознаваемого объекта в ряд по векторам категорий (классов распознавания) /1, 5, 7-11/. Этот вектор, состоящий из единиц и нулей, можно определить по координатам небесных тел, соответствующих дате и месту происхождения события l в виде

$$a_{ls} = \begin{cases} 1, & (j-1)\Delta x \leq x_{jk}(l) \leq j\Delta x, \quad jk = s \\ 0, & 1 \leq s \leq 2mk_0 \end{cases} \quad (11)$$

Таким образом, если нормированная координата небесного тела из данных по объекту исследуемой выборки попадает в заданный интервал, элементу вектора придается значение 1, а во всех остальных случаях – значение 0. Перечисление координат осуществляется последовательно, для каждого небесного тела. В качестве астрономических параметров в работах /7-11/ и других были использованы долгота, широта и расстояние от Земли до десяти небесных тел – Солнца, Луны, Марса, Меркурия, Венеры, Юпитера, Сатурна, Урана, Нептуна и Плутона, и долгота Северного Узла Луны.

В некоторых задачах возникает необходимость исследования совместного влияния небесных тел и местных параметров, типа координат и угловой скорости движения полюса Земли, индукции магнитного поля по измерениям на разных станциях, гравитационного потенциала, числа пятен на Солнце и т.п. на сейсмические события /7-12/.

В работах /13-15/ была построена модель линейной регрессии, описывающая движение полюса Земли, с использованием комбинаций астрономических параметров, характеризующих влияние каждого небесного тела в виде:

$$P_{i1} = \frac{\sin LAT_i}{R_i} - \frac{\cos J_e}{R_i}$$

$$P_{i2} = k_i \frac{\cos LAT_i \sin LON_i}{R_i} - \frac{\sin J_e}{R_i} \quad (12.1)$$

$$P_{i3} = k_i \frac{\cos LAT_i \cos LON_i}{R_i}$$

$$k_i = \sin J_e \cos LAT_i \sin LON_i + \cos J_e \sin LAT_i, \quad i = 1, 2, \dots, 10$$

$$\begin{aligned}
 P'_{i1} &= Q_i (-\cos J_e \cos LAT_i \sin LON_i + \sin J_e \sin LAT_i) / R_i^2 \\
 P'_{i2} &= Q_i \cos J \frac{\cos LAT_i \sin LON_i}{R_i^2} \\
 P'_{i3} &= Q_i \sin J \frac{\cos LAT_i \cos LON_i}{R_i^2}
 \end{aligned}
 \tag{12.2}$$

$$i = 1, 2, \dots, 10$$

Здесь долгота (LON), широта (LAT) и расстояние (R) определяется для каждого из 10 небесных тел, $J_e = 23,439291^0$ - угол наклона земной оси относительно нормали к орбитальной плоскости, Q_i – заряды небесных тел.

Отметим, что данные для расстояний от Земли до небесных тел вычисляются в формулах (12) в астрономических единицах. Параметры (12.1), (12.2) были использованы в настоящей работе наряду с астрономическими параметрами.

При создании моделей в настоящей работе были использованы данные ежедневного числа солнечных пятен по наблюдениям американских астрономов /23/ – параметр RADAILY, данные по индукции магнитного поля Земли /24/, данные по сейсмическим событиям /25/, а также координаты географического полюса – X, Y /26/. Данные по индукции магнитного поля были взяты со следующих 23 станций (приведены только IAGA коды): GNA, GUA, IRT, KAK, MMB, RES, THL, DRV, HER, FUG, ABG, HON, CLF, LRV, SOD, AAE, AAA, TAN, SJG, AIA, TUC, BNG, MBO.

Решение прямой задачи включает в себя нормирование входных параметров и приведение их к одному масштабу изменения в интервале (0;360), разбиение интервалов на M частей, $2 < M < 36200$, вычисление матрицы абсолютных частот и информативности, в соответствии с формулами (1-2). Вопросы визуализации матрицы информативностей и построение различных когнитивных функций рассмотрены в работах /27-28/ и других.

Описание программы

Программа ViewVibe.jar написана на языке Java и может быть использована на компьютерах с ОС типа Linux, Windows XP и выше. Скриншот основного меню программы приведен на рис. 1. Работа начинается с загрузки данных (кнопка Load Data Table) из файла типа книги Excel с двумя страницами. На первой странице файла данных перечислены имена категорий, а на второй странице задаются номера и имена строк, категории и столбцы входных параметров - рис. 2.

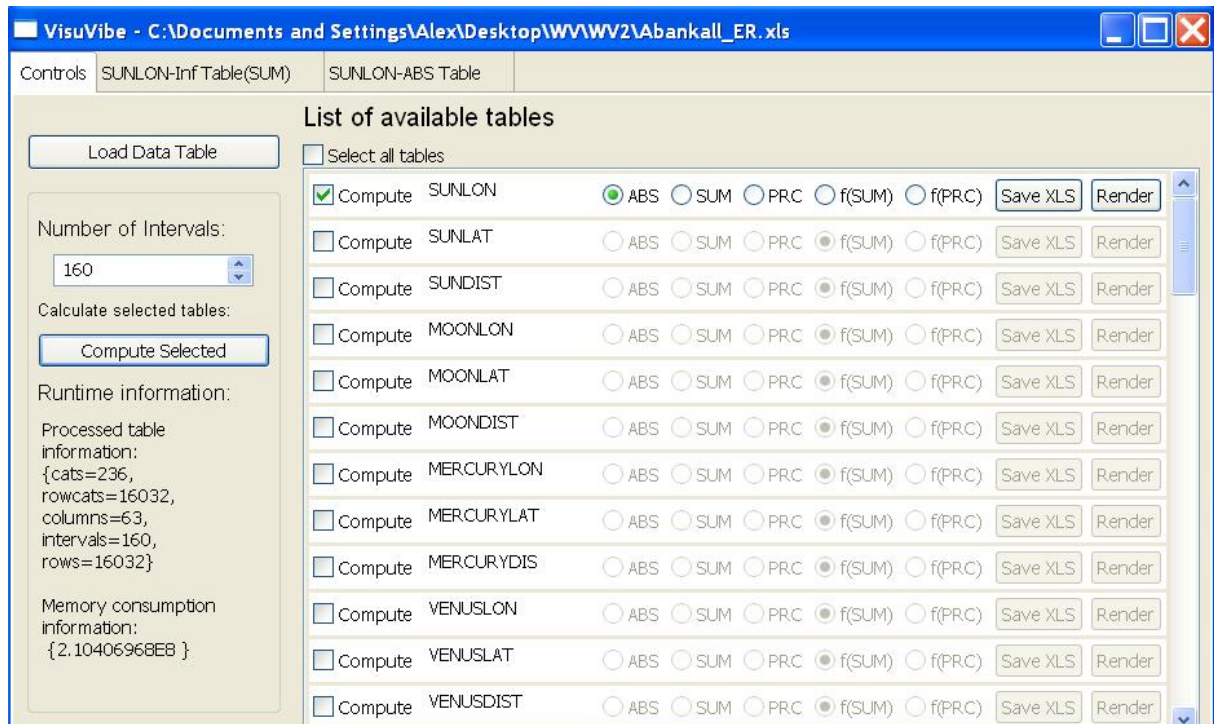
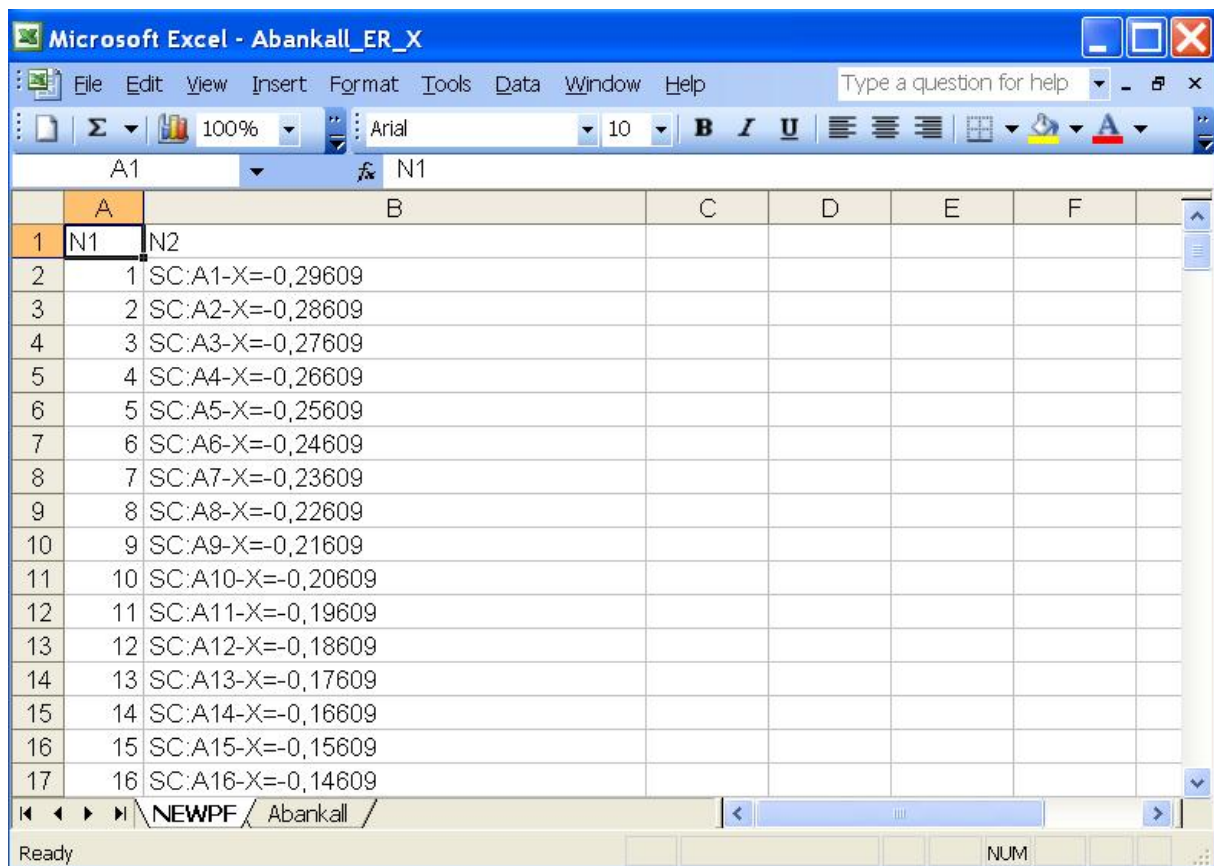


Рис. 1. Скриншот программы ViewVibe.jar



	A	B	C	D	
1	ID	NAME	CATS1	SUNLON	SUNLAT
2	1	09:02:1963	A13:B32:C40:D37:	320,0163486987710	-0,0
3	2	10:02:1963	A13:B32:C40:D37:	321,0281621052010	-0,0
4	3	11:02:1963	A13:B32:C41:D37:	322,0396282292110	-0,0
5	4	12:02:1963	A13:B32:C40:D37:	323,0507556897180	-0,0
6	5	13:02:1963	A13:B32:C41:D37:	324,0615515920590	-0,0
7	6	14:02:1963	A14:B32:C41:D37:	325,0720209176110	-0,0
8	7	15:02:1963	A14:B32:C41:D37:	326,0821660960550	-0,0
9	8	16:02:1963	A14:B33:C41:D36:	327,0919866910990	-0,0
10	9	17:02:1963	A14:B33:C41:D37:	328,1014791444830	-0,0
11	10	18:02:1963	A14:B33:C41:D37:	329,1106364828810	-0,0
12	11	19:02:1963	A15:B33:C41:D36:	330,1194478509500	-0,0
13	12	20:02:1963	A15:B33:C41:D37:	331,1278978932240	-0,0
14	13	21:02:1963	A15:B33:C41:D36:	332,1359663387060	-0,0
15	14	22:02:1963	A15:B33:C41:D36:	333,1436284262530	-0,0
16	15	23:02:1963	A15:B34:C41:D37:	334,1508566036650	-0,0
17	16	24:02:1963	A16:B34:C41:D36:	335,1576232309430	-0,0

Рис. 2. Вид первой и второй страниц файла загрузки

При загрузке данные нормируются в соответствии с формулами (1) для приведения данных к одной шкале. После этого программа готова к вычислениям. На этом этапе необходимо задать число интервалов (номер модели) в окне Number of Intervals. По формулам (2) вычисляется матрица частот и матрица информативности, а также служебные матрицы, используемые в алгоритмах СИМ1 и СИМ2 /1/. Вычисления можно производить для всех или только выбранных столбцов входных параметров путем пометки соответствующего окна и запуска Compute Selected.

После вычислений активируются окна основных и служебных матриц ABS, SUM, PRC, f(SUM), f(PRC). Каждую из матриц можно сохранить в формате XLS, путем нажатия кнопки Save XLS, а также создать на ее основе изображение процесса, путем нажатия кнопки Render. Путь ко всем созданным программой изображениям сохраняется на верхней панели - рис. 1. Таким образом, можно сравнивать изображения, полученные на основе различных матриц и в различных моделях – рис. 3. При переходе по любому из путей открывается изображение вместе с панелью редактирования и кнопкой возврата в основное меню (Controls). На рис 3 представлен скриншот рабочего окна программы визуализации матриц ABS, SUM, PRC, f(SUM), f(PRC). Окно содержит две полосы прокрутки для просмотра изображений большого размера и кнопку изменения масштаба изображения Zoom Image. Изображение можно раскрасить - кнопки Red, Green, Blue, Gray, сгладить – кнопка Smooth, сделать правый и левый поворот на

90 градусов - Rotate Left, Rotate Right. Изображение сохраняется при нажатии кнопки Save Image по выбору в формате BMP, JPG, PNG с исходным размером, который задается размером фрагмента матрицы информативностей, т.е. размером=(число интервалов)×(число категорий) пиксель.

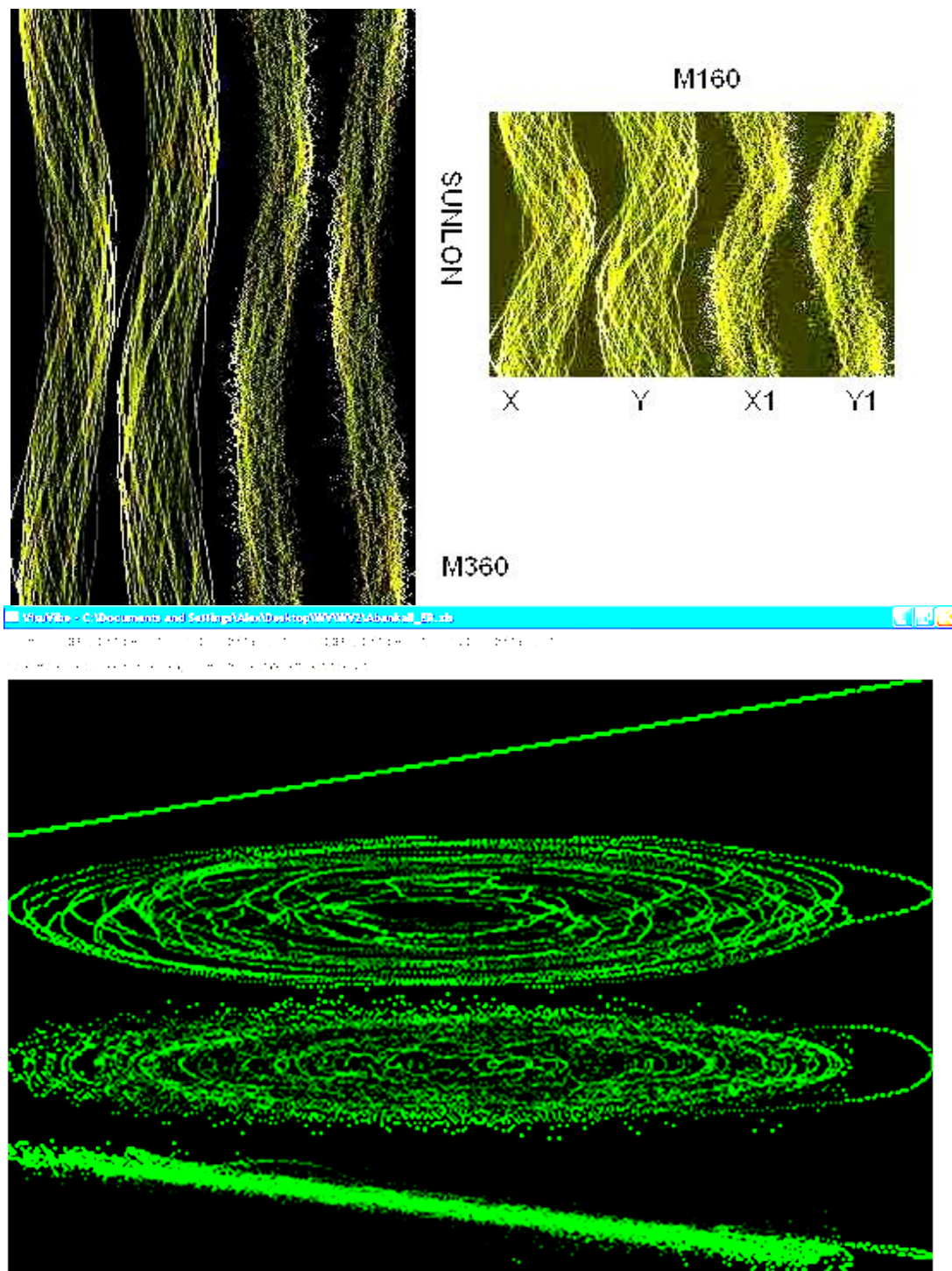


Рис. 3. Два изображения координат и угловой скорости движения полюса Земли в зависимости от долготы Солнца, созданных на основе матриц информативностей в моделях M360 и M160 и скриншот рабочего окна программы визуализации

Визуализация матриц семантических информационных моделей движения полюса Земли

Для решения прикладных задач в АСК-анализе проводится последовательное повышение степени формализации исходных данных до уровня, обеспечивающего их обработку на компьютере в программной системе. После выполнения когнитивной структуризации и формализации предметной области осуществляется синтез модели. Он включает в себя расчет на основе эмпирических данных, представленных в исследуемой выборке, следующих матриц:

- матрицы абсолютных частот ABS, которые используются во многих статистических системах;
- матрицы условных и безусловных процентных распределений PRC, которые используются в некоторых статистических системах;
- матрицы информативностей или матрицы знаний INF, которые используются только в АСК-анализе.

Интересно сравнить изображения, полученные на основе этих трех матриц в задаче о движении полюса Земли – рис. 4. В данной задаче движение полюса характеризуется координатами X , Y и угловой скоростью $X1$, $Y1$ – таблица 1. Как следует из данных, приведенных на рис. 4, при сравнении четырех зависимостей $Y(X)$, $Y(Y)$, $Y(X1)$, $Y(Y1)$ на основе матрицы абсолютных частот достаточно достоверно выявляется только линейная зависимость $Y=Y$. При визуализации на основе матрицы процентных распределений выявляются все зависимости, однако детали траектории движения полюса $Y=Y(X)$, размыты из-за сильного контраста данных с высокой и низкой частотой. **И только при визуализации на основе матрицы информативностей оказывается возможным проследить все детали движения полюса** – см. правый рис. 4 и рис. 3, на котором в увеличенном масштабе представлены зависимости $X(X)$, $X(Y)$, $X(X1)$, $X(Y1)$.

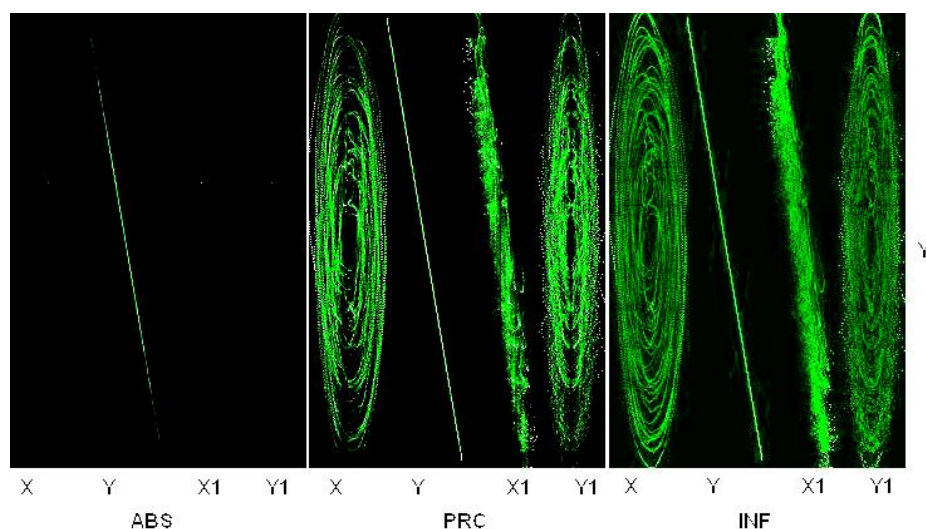


Рис. 4. Визуализация матрицы абсолютных частот (слева), матрицы процентных распределений (в центре) и матрицы информативностей (справа)

В будущих версиях системы «Эйдос-Java» планируется использование критерия хи-квадрат в качестве *фильтра*, не пропускающего сомнительные с точки зрения их достоверности данные при визуализации матриц ABS, PER, INF, что, возможно, позволит уменьшить шум в изображениях когнитивных функций.

Теоретическое объяснение целесообразности использования именно матрицы информативностей (базы знаний) для визуализации когнитивных функций и решения других задач (идентификации, прогнозирования, поддержки принятия решений и исследования предметной области) состоит в следующем. На основе сравнения абсолютных частот можно делать какие-либо выводы только в том случае, когда по разным классам приведено одинаково количество примеров. В общем случае это количество примеров по разным классам *всегда разное*, поэтому матрица абсолютных частот сама по себе непригодна для решения вышеперечисленных задач. Но на практике иногда встречается ситуация, когда это количество мало отличается по разным классам и тогда использование этой матрицы не дает большой ошибки и позволяет составить более-менее достоверное представление о предметной области. Получить модель предметной области, **инвариантную** относительно различий в количестве примеров по классам, можно перейдя от матрицы абсолютных частот к матрице условных и безусловных процентных распределений⁶. Использование матрицы условных и безусловных процентных распределений позволяет получить такой же уровень достоверности выводов о предметной области, какой матрица абсолютных частот *в случае* равного количества примеров по разным классам. Однако для того, чтобы получить эти выводы необходимо сравнивать условные процентные распределения друг с другом. При небольших размерностях моделей это сравнение может быть проведено вручную, правда лишь на качественном уровне, однако при реальных встречающихся на практике размерностях вручную это сделать не представляется возможным. Поэтому в АСК-анализе принято решение **автоматизировать** количественное сравнение условных процентных распределений. При этом возник вопрос о том, как именно их сравнивать: друг с другом или с какой-то базой сравнения. Если в модели всего два класса, то приемлем вариант сравнения условных процентных распределений по ним друг с другом. Но

⁶ При увеличении объема выборки относительные частоты (выраженные в процентах) стремятся к вероятностям, как пределам. При достаточно больших объемах выборки различие относительных частот от вероятностей невелико и на практике приемлемо вместо терминов «относительные и абсолютные частоты» использовать термины «условные и безусловные вероятности». Под процентными распределениями понимаются именно частоты, а не вероятности.

когда классов всего три, то уже не очень понятно как это делать, если же их сотни, тысячи или десятки тысяч, то это становится вообще непонятным. В АСК-анализе этот вопрос решен путем сравнения условных процентных распределений по классам с безусловным процентным распределением по всей выборке, которое и выступает базой (нормой) сравнения. Это решение соответствует принятому в статистике методу средних и отклонений от средних и представляет собой косвенное или опосредованное сравнение условных процентных распределений друг с другом, т.к. база сравнения рассчитывается с их использованием. Таким образом, *матрица информативностей представляет собой результат автоматизированного сравнения условных процентных распределений признаков по классам с их безусловным процентным распределением по всей выборке*. По сути это результат *нормировки* условных процентных распределений с использованием в качестве нормы безусловного процентного распределения. Поэтому использование матрицы информативностей освобождает исследователя-аналитика от необходимости выполнения огромной рутинной работы (которую он как правило и не может выполнить вручную, а когда может, то лишь на качественном, а не на количественном уровне) по сравнению условных процентных распределений друг с другом или с какой-либо базой сравнения и обеспечивает более высокий уровень достоверности выводов, чем использование матрицы условных и безусловных процентных распределений.

С увеличением количества интервалов оказывается возможным установить такие закономерности движения полюса Земли, которые никак не просматриваются при меньших разбиениях. На верхнем рис. 3 представлены зависимости координат и угловой скорости полюса от долготы Солнца. На левом рисунке, полученном в модели М360, на данных для угловой скорости видны колебания, видимо, обусловленные влиянием Луны, которые в модели М160 представляются как случайные выбросы.

На рис. 5 представлен фрагмент матрицы информативностей, демонстрирующий зависимость координат (X,Y) и угловой скорости (X1, Y1) движения полюса Земли – всего 244 категории, перечисленные в таблице 1, от долготы Солнца в 1963-2006 годов по данным /26/. Как следует из данных, приведенных на рис. 5, координаты полюса в зависимости от долготы Солнца образуют *жгуты*, которые формируются из отдельных годичных движений на протяжении многих лет. Отметим, что категории угловой скорости не образуют достаточно четких нитевидных структур, подобных тем, что образуют категории координат.

Таблица 1 – ЕЖЕДНЕВНАЯ ЧАСТОТА ПОЯВЛЕНИЯ КАТЕГОРИЙ КООРДИНАТ И УГЛОВОЙ СКОРОСТИ ДВИЖЕНИЯ ПОЛЮСА ЗЕМЛИ В 1963-2006 ГОДАХ

Категория X	ABS	Категория Y	ABS	Категория X1	ABS	Категория Y1	ABS
A1-X=-0,29609	11	B1-Y=-0,01292	30	A1-X1=-0,006	1	B2-Y1=-0,00554	1
A2-X=-0,28609	8	B2-Y=-0,00292	55	A2-X1=-0,0058	2	B4-Y1=-0,00514	6
A3-X=-0,27609	8	B3-Y=0,00708	114	A3-X1=-0,0056	1	B5-Y1=-0,00494	4
A4-X=-0,26609	7	B4-Y=0,01708	139	A4-X1=-0,0054	5	B6-Y1=-0,00474	20
A5-X=-0,25609	17	B5-Y=0,02708	116	A5-X1=-0,0052	7	B7-Y1=-0,00454	42
A6-X=-0,24609	70	B6-Y=0,03708	95	A6-X1=-0,005	11	B8-Y1=-0,00434	46
A7-X=-0,23609	99	B7-Y=0,04708	104	A7-X1=-0,0048	23	B9-Y1=-0,00414	93
A8-X=-0,22609	140	B8-Y=0,05708	123	A8-X1=-0,0046	32	B10-Y1=-0,00394	130
A9-X=-0,21609	125	B9-Y=0,06708	217	A9-X1=-0,0044	47	B11-Y1=-0,00374	137
A10-X=-0,20609	194	B10-Y=0,07708	248	A10-X1=-0,0042	67	B12-Y1=-0,00354	169
A11-X=-0,19609	199	B11-Y=0,08708	253	A11-X1=-0,004	136	B13-Y1=-0,00334	204
A12-X=-0,18609	188	B12-Y=0,09708	207	A12-X1=-0,0038	150	B14-Y1=-0,00314	258
A13-X=-0,17609	173	B13-Y=0,10708	247	A13-X1=-0,0036	202	B15-Y1=-0,00294	345
A14-X=-0,16609	238	B14-Y=0,11708	274	A14-X1=-0,0034	265	B16-Y1=-0,00274	315
A15-X=-0,15609	378	B15-Y=0,12708	256	A15-X1=-0,0032	309	B17-Y1=-0,00254	390
A16-X=-0,14609	269	B16-Y=0,13708	314	A16-X1=-0,003	354	B18-Y1=-0,00234	434
A17-X=-0,13609	272	B17-Y=0,14708	317	A17-X1=-0,0028	356	B19-Y1=-0,00214	433
A18-X=-0,12609	269	B18-Y=0,15708	346	A18-X1=-0,0026	383	B20-Y1=-0,00194	446
A19-X=-0,11609	340	B19-Y=0,16708	375	A19-X1=-0,0024	335	B21-Y1=-0,00174	440
A20-X=-0,10609	354	B20-Y=0,17708	451	A20-X1=-0,0022	434	B22-Y1=-0,00154	449
A21-X=-0,09609	271	B21-Y=0,18708	427	A21-X1=-0,002	445	B23-Y1=-0,00134	473
A22-X=-0,08609	299	B22-Y=0,19708	432	A22-X1=-0,0018	412	B24-Y1=-0,00114	500
A23-X=-0,07609	302	B23-Y=0,20708	422	A23-X1=-0,0016	398	B25-Y1=-0,00094	586
A24-X=-0,06609	342	B24-Y=0,21708	341	A24-X1=-0,0014	422	B26-Y1=-0,00074	542
A25-X=-0,05609	385	B25-Y=0,22708	372	A25-X1=-0,0012	449	B27-Y1=-0,00054	581
A26-X=-0,04609	379	B26-Y=0,23708	478	A26-X1=-0,001	403	B28-Y1=-0,00034	515
A27-X=-0,03609	515	B27-Y=0,24708	417	A27-X1=-0,0008	510	B29-Y1=-0,00014	587
A28-X=-0,02609	406	B28-Y=0,25708	374	A28-X1=-0,0006	595	B30-Y1=0,00006	521
A29-X=-0,01609	460	B29-Y=0,26708	340	A29-X1=-0,0004	515	B31-Y1=0,00026	614
A30-X=-0,00609	421	B30-Y=0,27708	332	A30-X1=-0,0002	574	B32-Y1=0,00046	612
A31-X=0,00391	441	B31-Y=0,28708	333	A31-X1=0	656	B33-Y1=0,00066	575
A32-X=0,01391	336	B32-Y=0,29708	338	A32-X1=0,0002	554	B34-Y1=0,00086	468
A33-X=0,02391	347	B33-Y=0,30708	356	A33-X1=0,0004	575	B35-Y1=0,00106	427
A34-X=0,03391	347	B34-Y=0,31708	400	A34-X1=0,0006	546	B36-Y1=0,00126	403
A35-X=0,04391	369	B35-Y=0,32708	427	A35-X1=0,0008	458	B37-Y1=0,00146	503
A36-X=0,05391	419	B36-Y=0,33708	369	A36-X1=0,001	473	B38-Y1=0,00166	554
A37-X=0,06391	469	B37-Y=0,34708	370	A37-X1=0,0012	503	B39-Y1=0,00186	456
A38-X=0,07391	382	B38-Y=0,35708	341	A38-X1=0,0014	469	B40-Y1=0,00206	438
A39-X=0,08391	414	B39-Y=0,36708	379	A39-X1=0,0016	485	B41-Y1=0,00226	427
A40-X=0,09391	402	B40-Y=0,37708	450	A40-X1=0,0018	428	B42-Y1=0,00246	328
A41-X=0,10391	410	B41-Y=0,38708	372	A41-X1=0,002	338	B43-Y1=0,00266	271
A42-X=0,11391	350	B42-Y=0,39708	351	A42-X1=0,0022	343	B44-Y1=0,00286	306
A43-X=0,12391	371	B43-Y=0,40708	289	A43-X1=0,0024	344	B45-Y1=0,00306	246
A44-X=0,13391	416	B44-Y=0,41708	301	A44-X1=0,0026	353	B46-Y1=0,00326	197
A45-X=0,14391	267	B45-Y=0,42708	335	A45-X1=0,0028	257	B47-Y1=0,00346	170
A46-X=0,15391	280	B46-Y=0,43708	207	A46-X1=0,003	329	B48-Y1=0,00366	132
A47-X=0,16391	284	B47-Y=0,44708	214	A47-X1=0,0032	239	B49-Y1=0,00386	99
A48-X=0,17391	258	B48-Y=0,45708	241	A48-X1=0,0034	192	B50-Y1=0,00406	60
A49-X=0,18391	257	B49-Y=0,46708	229	A49-X1=0,0036	175	B51-Y1=0,00426	37
A50-X=0,19391	250	B50-Y=0,47708	282	A50-X1=0,0038	132	B52-Y1=0,00446	15
A51-X=0,20391	300	B51-Y=0,48708	181	A51-X1=0,004	106	B53-Y1=0,00466	5
A52-X=0,21391	256	B52-Y=0,49708	131	A52-X1=0,0042	83	B54-Y1=0,00486	8
A53-X=0,22391	294	B53-Y=0,50708	125	A53-X1=0,0044	65	B55-Y1=0,00506	3
A54-X=0,23391	181	B54-Y=0,51708	115	A54-X1=0,0046	41	B56-Y1=0,00526	4
A55-X=0,24391	170	B55-Y=0,52708	122	A55-X1=0,0048	15	B59-Y1=0,00586	2
A56-X=0,25391	190	B56-Y=0,53708	132	A56-X1=0,005	11	B62-Y1=0,00646	1
A57-X=0,26391	79	B57-Y=0,54708	147	A57-X1=0,0052	9	B64-Y1=0,00686	1
A58-X=0,27391	92	B58-Y=0,55708	103	A58-X1=0,0054	5	B65-Y1=0,00706	1
A59-X=0,28391	78	B59-Y=0,56708	68	A59-X1=0,0056	1		
A60-X=0,29391	56	B60-Y=0,57708	14	A60-X1=0,0058	2		
A61-X=0,30391	72						
A62-X=0,31391	23						

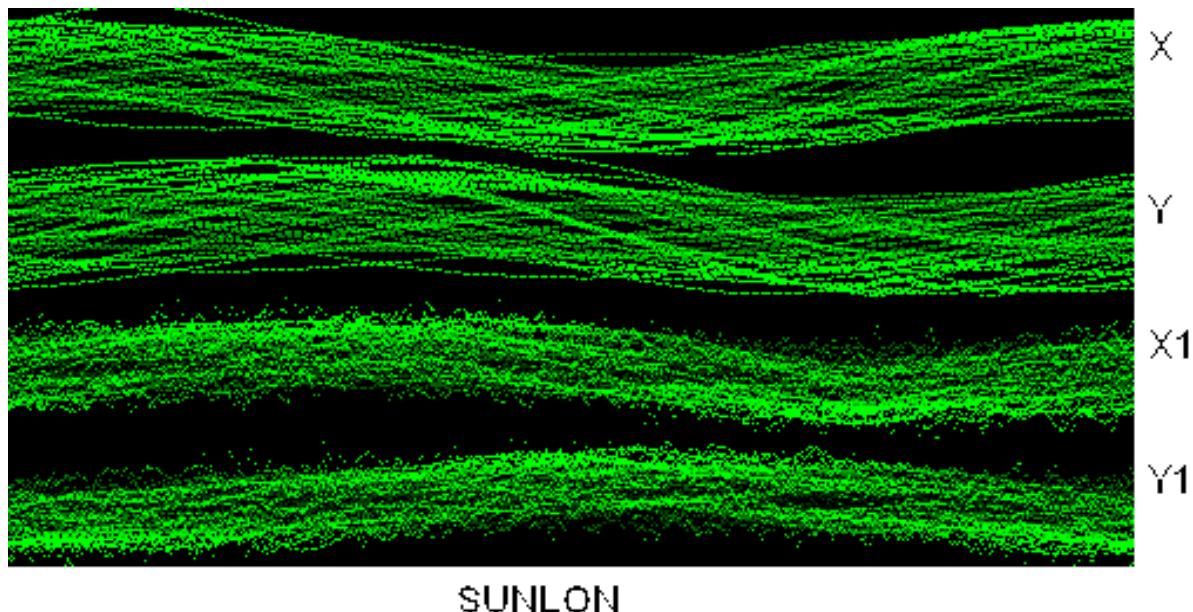


Рисунок 5. Зависимость категорий координат и угловой скорости полюса Земли от долготы Солнца в модели M480.

Согласно существующим представлениям, такое поведение угловой скорости полюса в зависимости от долготы Солнца обусловлено наличием случайной составляющей, связанной с движением атмосферы и океана относительно земной коры.

Визуализация матриц семантических информационных моделей сейсмических событий

Исследуемая база данных сейсмических событий была сформирована на основе базы данных Международного сейсмологического центра – ISC /25/, содержащей 20489816 записей регистрации различными сейсмо-станциями событий землетрясений, произошедших на нашей планете в период с 1 января 1961 года по 31 декабря 2006 г. Из исходной базы было образовано несколько различных БД для исследования влияния астрономических параметров на магнитуду и глубину гипофокуса, на ежедневное число землетрясений, а также на средние параметры сейсмической активности. В работах /9-10/ исследована совокупность 128320 событий землетрясений с магнитудой $m_b \geq 4$, произошедших на нашей планете в период с 9 февраля 1963 года по 31 декабря 2006 г (всего 16032 дня).

В исходной БД сейсмические события характеризуются магнитудой m_b , которой можно сопоставить категорию магнитуды – таблица 2. Поскольку события с одной и той же магнитудой могут повторяться в один день, каждому значению магнитуды сопоставляется несколько типов категорий, а именно:

- А – событие с магнитудой m_b повторяется один раз;
- В – событие с магнитудой m_b повторяется два раза;
- С – событие с магнитудой m_b повторяется три раза.

Кроме того, можно рассмотреть случай, когда, например, категория А усекается, путем отбрасывания некоторых событий. Таким образом, были образованы категории А66, В59 и С53.

Таблица 2 – Частота повторения категорий сейсмической активности

Категория	ABS	Категория	ABS	Категория	ABS
A40-Mb=4,0	1362	B40-Mb=4,0	446	C40-Mb=4,0	130
A41-Mb=4,1	1580	B41-Mb=4,1	660	C41-Mb=4,1	259
A42-Mb=4,2	1796	B42-Mb=4,2	835	C42-Mb=4,2	362
A43-Mb=4,3	2224	B43-Mb=4,3	955	C43-Mb=4,3	436
A44-Mb=4,4	2744	B44-Mb=4,4	1099	C44-Mb=4,4	482
A45-Mb=4,5	3358	B45-Mb=4,5	1223	C45-Mb=4,5	468
A46-Mb=4,6	4119	B46-Mb=4,6	1455	C46-Mb=4,6	515
A47-Mb=4,7	4768	B47-Mb=4,7	1612	C47-Mb=4,7	501
A48-Mb=4,8	4954	B48-Mb=4,8	1817	C48-Mb=4,8	450
A49-Mb=4,9	5008	B49-Mb=4,9	1636	C49-Mb=4,9	447
A50-Mb=5	4904	B50-Mb=5	1428	C50-Mb=5	356
A51-Mb=5,1	4582	B51-Mb=5,1	1206	C51-Mb=5,1	293
A52-Mb=5,2	4134	B52-Mb=5,2	936	C52-Mb=5,2	166
A53-Mb=5,3	3563	B53-Mb=5,3	617	C53-Mb=5,3-6,0	105
A54-Mb=5,4	3010	B54-Mb=5,4	422		
A55-Mb=5,5	2367	B55-Mb=5,5	261		
A56-Mb=5,6	1940	B56-Mb=5,6	180		
A57-Mb=5,7	1460	B57-Mb=5,7	93		
A58-Mb=5,8	1179	B58-Mb=5,8	73		
A59-Mb=5,9	864	B59-Mb=5,9-6,4	69		
A60-Mb=6,0	656				
A61-Mb=6,1	453				
A62-Mb=6,2	319				
A63-Mb=6,3	202				
A64-Mb=6,4	137				
A65-Mb=6,5	87				
A66-Mb=6,6-7,0	68				

На рис. 6 представлены данные визуализации матрицы абсолютных частот в задаче распознавания сейсмических событий по астрономическим данным /9-10/, а также с учетом влияния магнитного поля и движения полюса Земли /14/ в модели M720. Анализируя данные, приведенные на рис. 6, можно сделать вывод о том, что во всех рассмотренных случаях события группируются вблизи некоторых особых точек отображения (1) – см. /20/, образуя своеобразный линейчатый спектр. Такое поведение обнаружено впервые, благодаря использованной технике визуализации.

Другой эффект, обнаруженный на моделях большой размерности, заключается в том, что отображение сейсмических событий в зависимости от вертикальной компоненты индукции магнитного поля имеет вид своеобразных «пальцев» или жезлов, которые по виду *аналогичны* тем, что получают при визуализации матрицы абсолютных частот сейсмических событий в зависимости от расстояния до Урана – рис. 7.

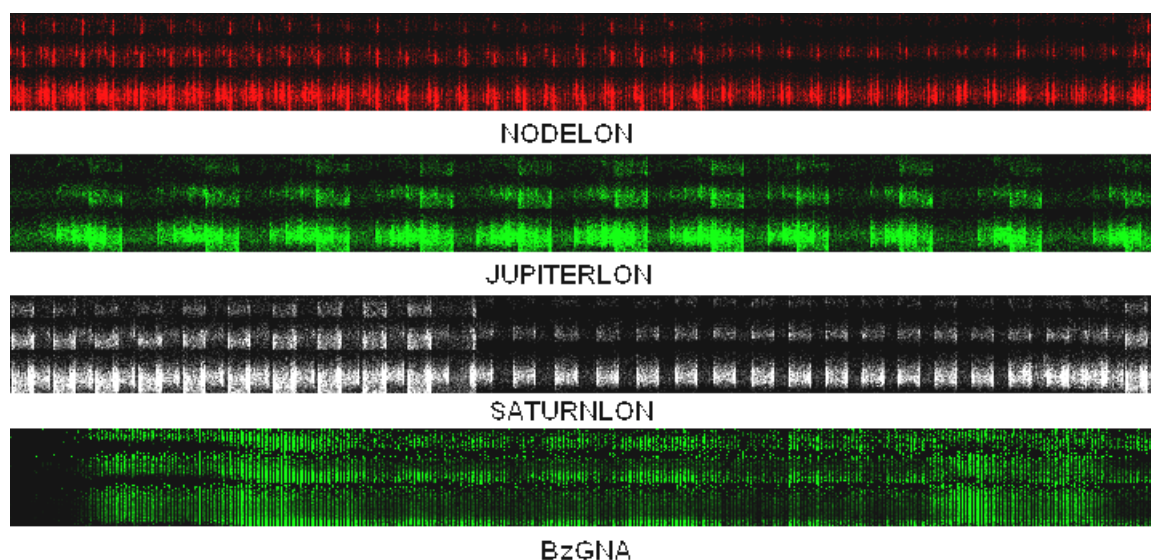


Рисунок 6. Зависимость категорий сейсмических событий А,В,С (в каждом рисунке категории отложены по вертикали снизу вверх) от долготы Лунного Узла, Юпитера и Сатурна, и от вертикальной компоненты индукции магнитного поля на станции GNA в модели M720. Для визуализации использована матрица абсолютных частот и условных и безусловных процентных распределений в случае магнитного поля

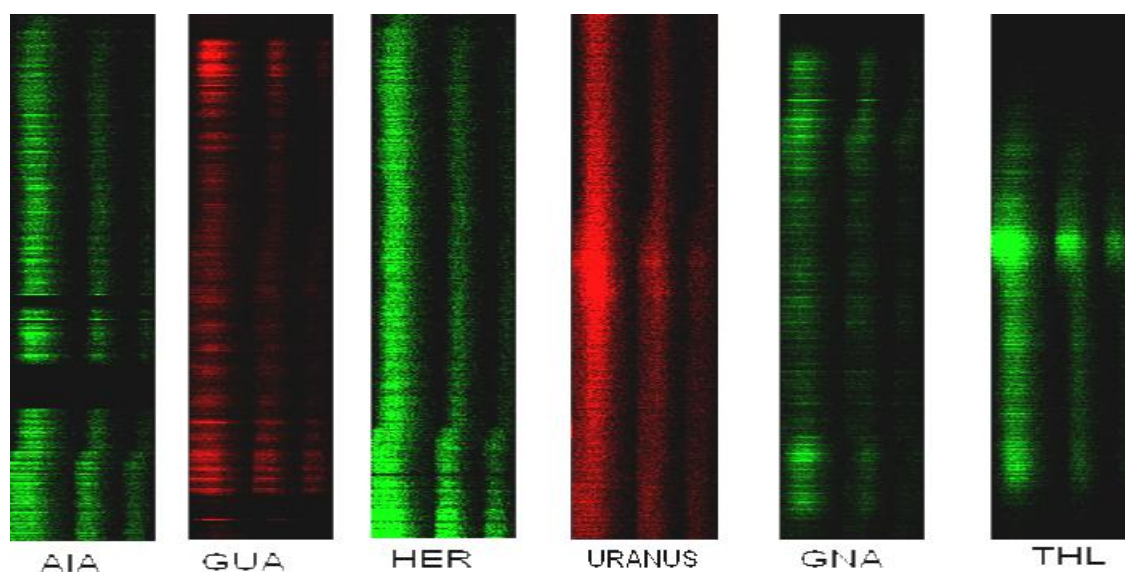


Рисунок 7. Зависимость категорий сейсмических событий А,В,С (в каждом рисунке категории отложены по горизонтали слева направо) от вертикальной компоненты индукции магнитного поля на станциях AIA, GUA, HER, GNA, THL и от расстояния от Земли до Урана в модели M360. Для визуализации использована матрица абсолютных частот.

Эта аналогия, впервые обнаруженная в работе /9/, послужила основой для создания моделей влияния небесных тел на геомагнитное поле и на движение полюса Земли /11, 13-15/.

Сравнивая картины визуализации трех матриц в модели M360, описывающей зависимость категорий сейсмических событий А,В,С от расстояния до Урана – рис. 8, находим, что *матрица информативностей наиболее точно отражает все детали этой зависимости*. И хотя в этом случае корреляция событий является весьма слабой, тем не менее,

благодаря использованию АСК-анализа /1, 29/ удастся выделить эту зависимость и найти комбинации астрономических параметров в форме (5.1)-(5.2), которые в наибольшей степени влияют на сейсмические события, движение полюса и магнитное поле Земли /13-15/.

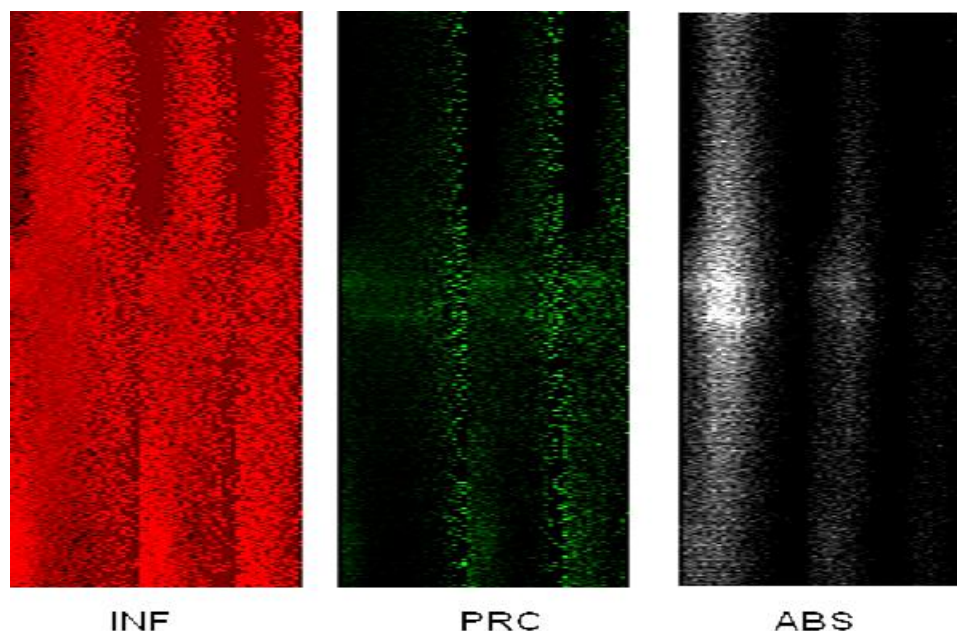


Рисунок 8. Визуализация матрицы информативностей (слева), матрицы процентных распределений (в центре) и матрицы абсолютных частот (справа) в модели М360 демонстрирует зависимость категорий сейсмических событий А,В,С (в каждом рисунке категории отложены по горизонтали слева направо) от расстояния до Урана (вертикальная координата, направленная сверху вниз).

Наконец, заметим, что исследованные в настоящей работе зависимости относятся к классу когнитивных функций /27-28, 30/, визуализация которых позволяет получить новые знания о поведении сложных систем, зависящих от множества параметров большой размерности. Приведенные выше примеры показывают существенное различие в стандартном классическом описании таких систем, основанном на использовании матрицы абсолютных частот, и с использованием АСК-анализа /1/.

Таким образом, по результатам работы можно сделать обоснованный **вывод** о том, что *сформулированная проблема решена*, т.е. создана новая версия системы искусственного интеллекта «Эйдос-астра» для решения прикладных задач с эмпирическими данными большой размерности. Приложение, написанное на языке JAVA, обеспечивает GUI и позволяет подготовить и выполнить визуализацию матрицы информативностей без ограничений, налагаемых реализацией предыдущих версий системы «Эйдос-астра».

В качестве *перспективы* авторы рассматривают реализацию в новой версии всех функций предыдущей версии, а также реализацию качественно новых возможностей,

Кроме того, авторы приняли решение предоставить заинтересованным читателям возможность бесплатно скачивать новую версию системы «Эйдос-астра» /12/ с исходными базами данных /32-33/ и исследовать примеры, приведенные в данной статье.

Библиографический список

1. Луценко Е.В. Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем): Монография (научное издание). – Краснодар: КубГАУ. 2002. – 605с⁷.
2. Луценко Е.В. Универсальная когнитивная аналитическая система "ЭЙДОС". Пат. № 2003610986 РФ. Заяв. № 2003610510 РФ. Оpubл. от 22.04.2003.
3. Patent 2008610097⁸, Russia, System for Typification and Identification of the Social Status of Respondents Based on the Astronomical Data at the Time of Birth - "AIDOS-ASTRO" / E.V. Lutsenko, A.P. Trunev, V.N. Shashin; Application № 2007613722, January 9,2008.
4. Луценко Е.В. Artificial intelligence system for identification of social categories of natives based on astronomical parameters / Е.В. Луценко, А.П. Трунев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2008. – №03(37). – Режим доступа: <http://ej.kubagro.ru/2008/03/pdf/07.pdf>
5. Луценко Е.В. «Эйдос-астра» – интеллектуальная система научных исследований влияния космической среды на поведение глобальных геосистем / Е.В. Луценко, А.П. Трунев // Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2010. – №07(61). – Режим доступа: <http://ej.kubagro.ru/2010/07/pdf/17.pdf>
6. Патент РФ⁹: Луценко Е.В., Трунев А.П., Шашин В.Н., Бандык Д.К. Интеллектуальная система научных исследований влияния космической среды на глобальные геосистемы «Эйдос-астра» (ИСНИ «Эйдос-астра»). Пат. № 2011612054 РФ. Заяв. № 2011610345 РФ 20.01.2011. Оpubл. от 09.03.2011.
7. Трунев А.П. Прогнозирование сейсмической активности и климата на основе семантических информационных моделей / А.П. Трунев, Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №09(53). – Шифр Информрегистрa: 0420900012\0098. – Режим доступа: <http://ej.kubagro.ru/2009/09/pdf/09.pdf>
8. Трунев А.П. Системно-когнитивный анализ и прогнозирование сейсмической активности литосферы Земли, как глобальной активной геосистемы / А.П. Трунев, Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2010. – №01(55). – Шифр Информрегистрa: 0421000012\0001. – Режим доступа: <http://ej.kubagro.ru/2010/01/pdf/22.pdf>
9. Трунев А.П. Семантические информационные модели глобальной сейсмической активности при смещении географического и магнитного полюса / А.П. Трунев, Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2010. – №02(56). – Шифр Информрегистрa: 0421000012\0023. – Режим доступа: <http://ej.kubagro.ru/2010/02/pdf/15.pdf>
10. Трунев А.П. Прогнозирование землетрясений по астрономическим данным с использованием системы искусственного интеллекта / А.П. Трунев, Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного универси-

⁷ Для удобства читателей ряд работ из списка приведен на сайте автора: <http://lc.kubagro.ru/>

⁸ <http://lc.kubagro.ru/aidos/2008610097.jpg>

⁹ <http://lc.kubagro.ru/aidos/2011612054.jpg>

- тета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №08(52). – Шифр Информрегистра: 0420900012\0086. – Режим доступа: <http://ej.kubagro.ru/2009/08/pdf/13.pdf>
11. Трунев А.П. Автоматизированный системно-когнитивный анализ влияния тел Солнечной системы на движение полюса Земли и визуализация причинно-следственных зависимостей в виде когнитивных функций / А.П. Трунев, Е.В. Луценко, Д.К. Бандык // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №01(65). – Режим доступа: <http://ej.kubagro.ru/2011/01/pdf/20.pdf>
 12. <http://chaosandcorrelation.org/AS/ViewVibe.jar>
 13. Трунев А.П. Моделирование электромагнитного и гравитационного влияния небесных тел солнечной системы на смещение географического полюса и магнитное поле Земли / А.П. Трунев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2010. – №07(61). – Шифр Информрегистра: 0421000012\0152. – Режим доступа: <http://ej.kubagro.ru/2010/07/pdf/16.pdf>
 14. Трунев А.П. Семантические информационные модели влияния солнечных пятен на сейсмическую активность, движение полюса и магнитное поле Земли / А.П. Трунев, Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №02(66). С. 546 – 571. – Режим доступа: <http://ej.kubagro.ru/2011/02/pdf/46.pdf>, 1,625 у.п.л.
 15. Трунев А.П. Моделирование влияния небесных тел на движение полюса Земли / А.П. Трунев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2010. – №10(64). – Режим доступа: <http://ej.kubagro.ru/2010/10/pdf/22.pdf>
 16. Трунев А. П., Луценко Е. В. Астросоциотипология: Монография (научное издание). – Краснодар: КубГАУ, 2008, – 279 с.
 17. Трунев А. П., Луценко Е. В. Системно-когнитивный анализ взаимосвязей между астрономическими и социальными событиями в астросоциотипологии /8-я Международная ФАМ конференция по финансово-актуарной математике и смежным вопросам, Красноярск, 2009.
 18. Луценко Е.В. Типизация и идентификация респондентов в социологии по их астрономическим показателям на момент рождения. / Е.В. Луценко, А.П. Трунев, В.Н. Шашин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2007. – №01(25). – Шифр Информрегистра: 0420700012\0014. – Режим доступа: <http://ej.kubagro.ru/2007/01/pdf/14.pdf>
 19. Луценко Е.В. Астросоциотипология и спектральный анализ личности по астросоциотипам с применением семантических информационных мультимodelей / Е.В. Луценко, А.П. Трунев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2008. – №01(35). – Шифр Информрегистра: 0420800012\0002. – Режим доступа: <http://ej.kubagro.ru/2008/01/pdf/10.pdf>
 20. Трунев А.П. Проблема распознавания событий в поле центральных сил и прогнозирование курсов валют / А.П. Трунев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №06(50). – Шифр Информрегистра: 0420900012\0057. – Режим доступа: <http://ej.kubagro.ru/2009/06/pdf/08.pdf>
 21. Трунев А.П. Прогнозирование курсов валют по астрономическим данным с использованием системы искусственного интеллекта / А.П. Трунев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №07(51). – Шифр Информрегистра: 0420900012\0068. – Режим доступа: <http://ej.kubagro.ru/2009/07/pdf/15.pdf>

22. Трунев А.П. Корреляция фондового индекса s & p 500 с астрономическими и геофизическими параметрами (Системно-когнитивный анализ взаимосвязи ноосферы, литосферы, магнитосферы и космической среды) / А.П. Трунев, Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2010. – №03(57). – Шифр Информрегистра: 0421000012\0039. – Режим доступа: <http://ej.kubagro.ru/2010/03/pdf/13.pdf>
23. National Geophysical Data Center/ <http://www.ngdc.noaa.gov/ngdc.html>
24. World Data Centre for Geomagnetism (Edinburgh)/ <http://www.wdc.bgs.ac.uk/>
25. International Seismological Center/ <http://www.isc.ac.uk/>
26. Earth orientation centre / <http://hpiers.obspm.fr/eop-pc/>
27. Луценко Е.В. Метод визуализации когнитивных функций – новый инструмент исследования эмпирических данных большой размерности / Е.В. Луценко, А.П. Трунев, Д.К. Бандык // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №03(67). С. 240 – 282. – Режим доступа: <http://ej.kubagro.ru/2011/03/pdf/18.pdf>
28. Трунев А.П. Автоматизированный системно-когнитивный анализ влияния тел Солнечной системы на движение полюса Земли и визуализация причинно-следственных зависимостей в виде когнитивных функций / А.П. Трунев, Е.В. Луценко, Д.К. Бандык // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №01(65). С. 232 – 258. – Шифр Информрегистра: 0421100012\0002. – Режим доступа: <http://ej.kubagro.ru/2011/01/pdf/20.pdf>
29. Луценко Е.В. АСК-анализ как метод выявления когнитивных функциональных зависимостей в многомерных зашумленных фрагментированных данных / Е.В. Луценко // Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2005. – №03(11). – Режим доступа: <http://ej.kubagro.ru/2005/03/pdf/19.pdf>
30. Луценко Е.В. Когнитивные функции как адекватный инструмент для формального представления причинно-следственных зависимостей / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2010. – №09(63). – Режим доступа: <http://ej.kubagro.ru/2010/09/pdf/01.pdf>
31. Луценко Е.В. 30 лет системе «Эйдос» – одной из старейших отечественных универсальных систем искусственного интеллекта, широко применяемых и развивающихся и в настоящее время / Е.В. Луценко // Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №10(54). – Шифр Информрегистра: 0420900012\0110. – Режим доступа: <http://ej.kubagro.ru/2009/10/pdf/04.pdf>
32. http://chaosandcorrelation.org/AS/Abankall_ER.rar
33. <http://chaosandcorrelation.org/AS/AbankallS3.rar>