

## **ПРИМЕНЕНИЕ АВТОМАТИЗИРОВАННОГО СИСТЕМО- КОГНИТИВНОГО АНАЛИЗА ДЛЯ СОВЕРШЕНСТВОВАНИЯ МЕТОДОВ КОМПЬЮТЕРНОЙ СЕЛЕКЦИИ ПОДСОЛНЕЧНИКА**

Калустов А.А. – дипломник

Луценко Е.В. – д. э. н., к. т. н., профессор

*Кубанский государственный аграрный университет*

В статье рассматривается применение СК-анализа для компьютерной селекции подсолнечника, что позволяет сэкономить время и другие ресурсы при отборе растений для следующих селекционных поколений, ускорить и повысить качество селекции.

При селекции методом отбора для следующих поколений отбираются лучшие по генотипу растения. Оптимальным считается генотип, детерминирующий наилучшие по заданным критериям потребительские свойства растений.

Однако исследования на уровне генома, дающие информацию о том, какие конкретно гены и их сочетания детерминируют заданные целевые потребительские свойства, весьма дороги, трудоемки, требуют высокого уровня квалификации исследователей, значительного времени и первоклассного оборудования. Все это делает проблематичным и даже практически невозможным отбор конкретных растений для следующих селекционных поколений путем анализа их генома и сравнения его с оптимальным.

Поэтому **традиционным** является отбор растений по их **целевым потребительским** свойствам. В случае подсолнечника – это его морозостойкость, вес семян с одного растения и с гектара, выход масла с единицы

веса семян и с гектара насаждений подсолнечника. Однако и количественная оценка потребительских свойств конкретных растений требует специальных инструментальных измерений, например веса и выхода масла, которые сложно провести в полевых условиях.

Таким образом, **проблема** состоит в разработке полевой неинструментальной экспресс-методики, обеспечивающей оценку того, обладает ли генотип данного конкретного растения заданными целевыми свойствами. При этом не должно осуществляться непосредственное исследование генотипа и не должны использоваться какие-либо специальные инструментальные измерения потребительских характеристик растений.

Идея решения проблемы состоит в следующем.

Генотип растения детерминирует не только его потребительские свойства, но и фенотипические признаки. Поэтому по фенотипическим признакам растения можно судить не только о его генотипе, но и о его потребительских свойствах. Фенотипические же признаки устанавливаются непосредственно визуально, либо их оценка может потребовать простых измерений с помощью рулетки или мерного шеста.

Поэтому предлагается следующая **идея**: если выявить зависимости потребительских свойств растений от их фенотипических признаков, то можно косвенно оценивать, т.е., по сути, измерять эти потребительские свойства по фенотипическим признакам, естественно, с определенной степенью надежности, которую необходимо контролировать.

Необходимо отметить, что на фенотипические признаки, кроме генотипа растения, оказывают влияние также и внешние для растения факторы. Их можно разделить на две основные группы по степени зависимости от воли человека.

1. *Факторы окружающей среды* (прежде всего виды почв и метеорологические факторы) практически не зависят от человека.

2. *Технологические факторы*, т.е. связанные с использованием различных агротехнологий (вспашка, нормы высева, способы удобрения и защиты растений, полив, освещение, севооборот и т.д.), во многом зависят от человека.

Исследование влияния почв и метеорологических факторов на количественные и качественные результаты выращивания плодовых культур проводились И.А. Драгавцевой, Е.В. Луценко и Л.М. Лопатиной [1–8]. По семечковым, в частности яблокам, подобная работа, насколько известно по литературным данным, впервые проведена А.М. Максимовым [11] под научным руководством Е.В. Луценко. Влияние технологических факторов на количественные и качественные результаты выращивания зерновых колосовых изучалось О.А. Засухиной и Е.В. Луценко [9], [10].

Способ учета влияния всех этих внешних факторов один и тот же и не отличается от способа выявления зависимостей между фенотипом и потребительскими свойствами растений. Это означает, что технически мы могли бы исследовать все эти группы факторов в комплексе. Однако в данном исследовании мы этого делать не будем по двум основным причинам:

1) в процессах многолетней селекции подсолнечника морозоустойчивого сорта «Победа» в агрономических журналах *не фиксировались* внешние факторы;

2) внешние факторы *не менялись* в процессе селекции, т.е. выращивание осуществлялось в одной микроне с постоянными метеоусловиями и по одной технологии. Поэтому даже если бы факторы и учитывались в журналах, их влияние на потребительские свойства изучить не представлялось бы возможным из-за практически полного отсутствия вариабельности по этим факторам.

Поэтому в данном исследовании внешними факторами вполне корректно можно пренебречь.

**Актуальность для науки** данной работы определяется также ее *научной новизной*. Как уже упоминалось выше, в какой-то мере сходные исследования и разработки проводились И.А. Драгавцевой, Л.М. Лопатиной, Е.В. Луценко [1–8]; А.М. Максимовым, Е.В. Луценко [11]; О.А. Засухиной, Е.В. Луценко [9], [10]. Однако исследование взаимосвязи фенотипа и потребительских свойств подсолнечника, в частности для морозоустойчивого сорта «Победа», с целями совершенствования методов его компьютерной селекции, по всей видимости, проводится **впервые**, в этом и состоит *научная новизна данной работы*.

**Актуальность для практики** темы работы определяется возможностью применения ее результатов *в практике работы* научно-селекционных и образовательных организаций.

В научно-селекционных организациях разработанные методы компьютерной селекции позволяют существенно снизить затраты различных видов ресурсов, прежде всего времени и финансовых средств, на проведение селекционной работы методом отбора лучших по генотипу растений и использования их для формирования следующего селекционного поколения.

В учебном процессе методика разработки и решения перечисленных выше задач может быть использована для полноценной лабораторной работы по дисциплине "Интеллектуальные информационные системы", изучаемой на 5-м курсе специальности 351400 – Прикладная информатика.

Таким образом, **объектом исследования** является изучение взаимосвязи фенотипа и потребительских свойств культурных растений.

**Предмет исследования** состоит в изучении взаимосвязи фенотипа и потребительских свойств подсолнечника сорта «Победа».

**Цель исследования:** выявление причинно-следственных зависимостей между фенотипом и потребительскими свойствами подсолнечника сорта «Победа», и на основе этого разработка неинструментальной поле-

вой методики прогнозирования потребительских свойств и поддержки принятия решений по отбору растений для следующего селекционного поколения.

Основой для выполнения работы являются данные, предоставленные научно-селекционной агрофирмой ООО «Эверест-Олвик», специализирующейся на селекции морозоустойчивых высокотехнологичных сортов подсолнечника, прежде всего, сорта «Победа».

Агрофирму ООО «Эверест-Олвик» создал и уже более 13 лет беспрерывно возглавляет известный ученый, продолжатель дела академика В.С. Пустовойта, ученик академика И.А. Рапопорта, заслуженный деятель науки Кубани, профессор, доктор сельскохозяйственных наук Ашот Андраникович Калайджян – автор 225 научных работ, 3-х монографий и двух новых сортов подсолнечника «Салют» и «Победа». Сегодня возглавляемая им агрофирма – одно из ведущих хозяйств России по выведению новых сортов подсолнечника.

Первичная форма существования исходных данных – это полевые журналы агрономов, в которых они на протяжении 35 лет проведения селекционной работы по сорту «Победа» фиксировали регенеративные (потребительские) и вегетативные (фенотипические) характеристики растений на различных делянках.

Для исследования в данной работе выбраны данные за 10 лет: с 1994 по 2003 годы, по 100 делянок на каждый год. Анализ этих данных показывает, что они отражают не все возможные сочетания значений признаков.

Традиционные методы обработки имеющихся исходных данных не обеспечивают решение поставленных задач, и необходимо использование нового подхода, обеспечивающего *выявление зависимостей между потребительскими свойствами растений подсолнечника и их фенотипическими признаками при неполных (фрагментированных) данных.*

Традиционные статистические модели требуют информации о результатах действия *всех сочетаний исследуемых факторов* («повторности»), что в исследуемой предметной области практически невозможно.

Необходимо особо подчеркнуть, что восполнить отсутствующие данные из опыта не представляется возможным, т.к. объект исследования принципиально не допускает повторения условий прошедших периодов с заданными сочетаниями факторов, которые не зависят от воли человека. Восполнение данных путем интерполяции также некорректно, т.к. в каждой строке и столбце корреляционной матрицы имеется более одного пропуска. Статистические модели очень сложно содержательно интерпретировать, для этого требуется большой труд квалифицированных аналитиков.

Таким образом, можно сделать *вывод*, что для моделирования такого сложного и малоисследованного объекта, каким является подсолнечник в процессе селекции, применение традиционных математических моделей является проблематичным.

По-видимому, решение поставленных задач может быть получено путем применения системно-когнитивного анализа (СК-анализ) [9], [10] – нового перспективного математического метода системного анализа, основанного на теории информации, системном анализе и когнитивном моделировании.

Метод является непараметрическим, позволяет сопоставимо обрабатывать тысячи градаций факторов и будущих состояний объекта управления при неполных (фрагментированных), зашумленных данных различной природы.

Для метода СК-анализа разработаны и методика численных расчетов, и соответствующий программный инструментарий, а также технология и методика их применения. Они прошли успешную апробацию при решении ряда задач в различных предметных областях [1–11].

Наличие *инструментария* СК-анализа позволяет не только осуществить синтез семантической информационной модели (СИМ), но и *периодически проводить адаптацию и синтез ее новых версий*, обеспечивая тем самым отслеживание динамики предметной области и сохраняя высокую адекватность модели в изменяющихся условиях.

В работах [9], [10] приведен перечень этапов системно-когнитивного анализа, которые необходимо выполнить, чтобы осуществить синтез модели и исследование модели объекта управления.

1. Когнитивная структуризация, а затем и формализация предметной области.

2. Ввод данных мониторинга в базу прецедентов за период, в течение которого имеется необходимая информация в электронной форме.

3. Синтез семантической информационной модели (СИМ).

4. Оптимизация СИМ.

5. Проверка адекватности СИМ (измерение внутренней и внешней, дифференциальной и интегральной валидности).

6. Анализ СИМ.

7. Решение задач идентификации состояний объекта управления, прогнозирование и поддержка принятия управленческих решений с применением СИМ.

Важной особенностью СК-анализа является возможность единообразной числовой обработки разнотипных числовых и нечисловых данных. Это обеспечивается тем, что нечисловым величинам тем же методом, что и числовым, приписываются сопоставимые в пространстве и времени, а также между собой, количественные значения, позволяющие обрабатывать их как числовые:

- на первых двух этапах СК-анализа числовые величины сводятся к интервальным оценкам, как и информация об объектах нечисловой приро-

ды (фактах, событиях) (этот этап реализуется и в методах интервальной статистики);

- на третьем этапе СК-анализа всем этим величинам по единой методике, основанной на системном обобщении семантической теории информации А. Харкевича, сопоставляются количественные величины, с которыми в дальнейшем и производятся все операции моделирования (этот этап является уникальным для СК-анализа).

***Учитывая эти этапы СК-анализа, выполним декомпозицию цели работы в ряд задач, решение которых обеспечит ее поэтапное достижение.***

1. Когнитивная структуризация предметной области и формальная постановка задачи, проектирование структуры и состава исходных данных.

2. Получение исходных данных запланированного состава в той форме, в которой они накапливаются в поставляющей их организации (журналы).

3. Разработка Excel-формы для представления исходных данных.

4. Преобразование исходных данных в Excel-форму.

5. Контроль достоверности исходных данных и исправление ошибок.

6. Использование программного интерфейса для преобразования исходных данных из формы по датам в стандартную форму по фенофазам.

7. Использование программного интерфейса для преобразования исходных данных из стандартной формы по фенофазам в базы данных, используемые в инструментарии системно-когнитивного анализа (СК-анализ) – когнитивной аналитической системе "Эйдос" (система "Эйдос").

8. Синтез семантической информационной модели (СИМ).

9. Оптимизация СИМ.

10. Измерение адекватности СИМ.

11. **Задача 1:** выявление причинно-следственных зависимостей между фенотипическими признаками подсолнечника и его потребительскими свойствами.

12. **Задача 2:** разработка методики *прогнозирования* потребительских свойств растений подсолнечника на основе анализа их фенотипических признаков.

13. **Задача 3:** разработка методики *поддержки принятия решений* по отбору растений для селекции не по их потребительским свойствам, а на основе анализа фенотипических признаков.

14. Разработка принципов оценки экономической эффективности разработанных технологий при их применении:

- в научно-селекционных организациях;
- в образовательных учреждениях.

15. Исследование ограничений разработанной технологии и перспектив ее развития.

**Задача 1: выявление причинно-следственных зависимостей между фенотипическими признаками подсолнечника и его потребительскими свойствами.** Мерой причинно-следственных зависимостей является количество информации (которое содержится в факте действия определенного значения фактора) о том, что растение подсолнечника будет иметь определенное потребительское свойство. Значение информативности может быть по модулю различной величины и положительным и отрицательным по знаку, что означает, соответственно, величину и направление влияния данного значения фактора на данное потребительское свойство.

В полном виде все зависимости содержатся в матрице информативностей. Вертикальная шапка этой матрицы содержит градации описательных, а горизонтальная шапка – градации классификационных шкал.

**Задача 2: разработка методики прогнозирования потребительских свойств растений подсолнечника на основе анализа их феноти-**

**пических признаков.** Данная задача решается на основе задачи 1. Суть методики прогнозирования состоит в следующем. Из матрицы информативностей известно, какое количество информации содержится в определенном фенотипическом признаке конкретного растения подсолнечника о том, что данное растение будет обладать каждым из исследованных в модели потребительских свойств. Если о растении известно, что оно обладает определенным набором фенотипических признаков, то естественно считать, что оно будет обладать теми потребительскими свойствами, о которых в данном наборе содержится наибольшее количество информации. Данный интегральный критерий предложен и обоснован в [9].

Исходные данные для прогнозирования вводятся в 1-м режиме 4-й подсистемы системы «Эйдос» в форме распознаваемой выборки, само прогнозирование осуществляется во 2-м режиме этой же подсистемы, а результаты выводятся в 3-м режиме в двух разрезах:

- один объект – много классов;
- один класс – много объектов.

Результаты прогнозирования выводятся системой в обобщенной форме, каждая строка которой соответствует классу, с которым данный объект имеет наибольшее сходство, и в детализированной форме карточек прогнозирования (распознавания). Примеры этих карточек приведены на рисунке 1.

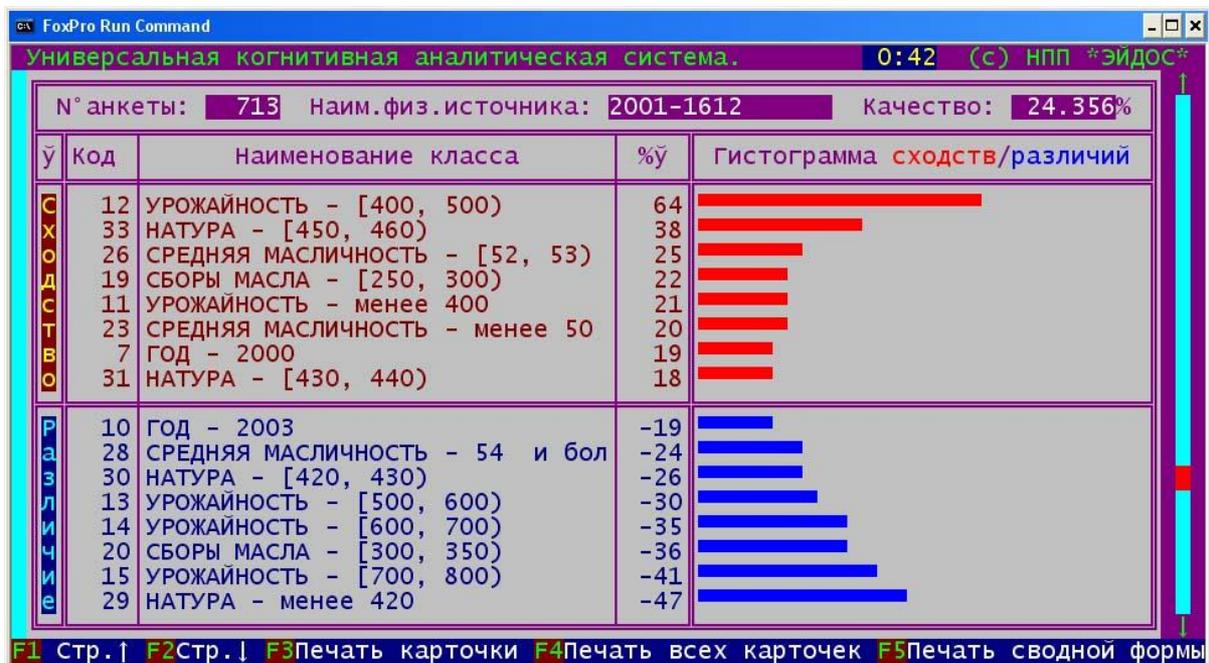
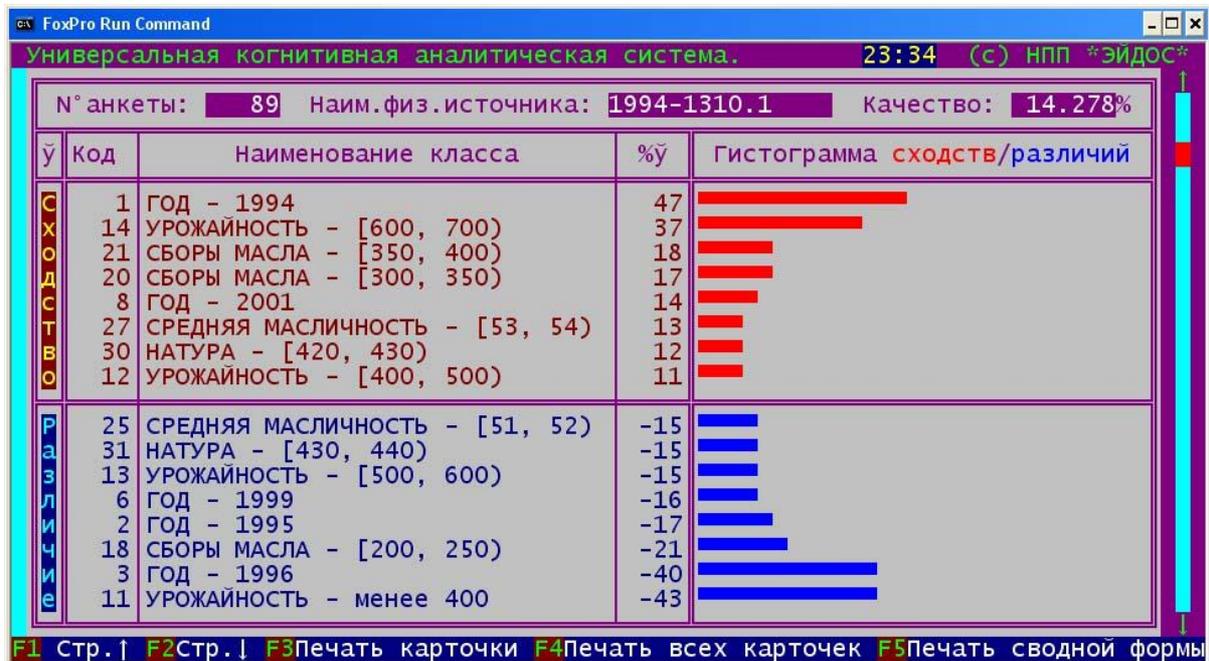


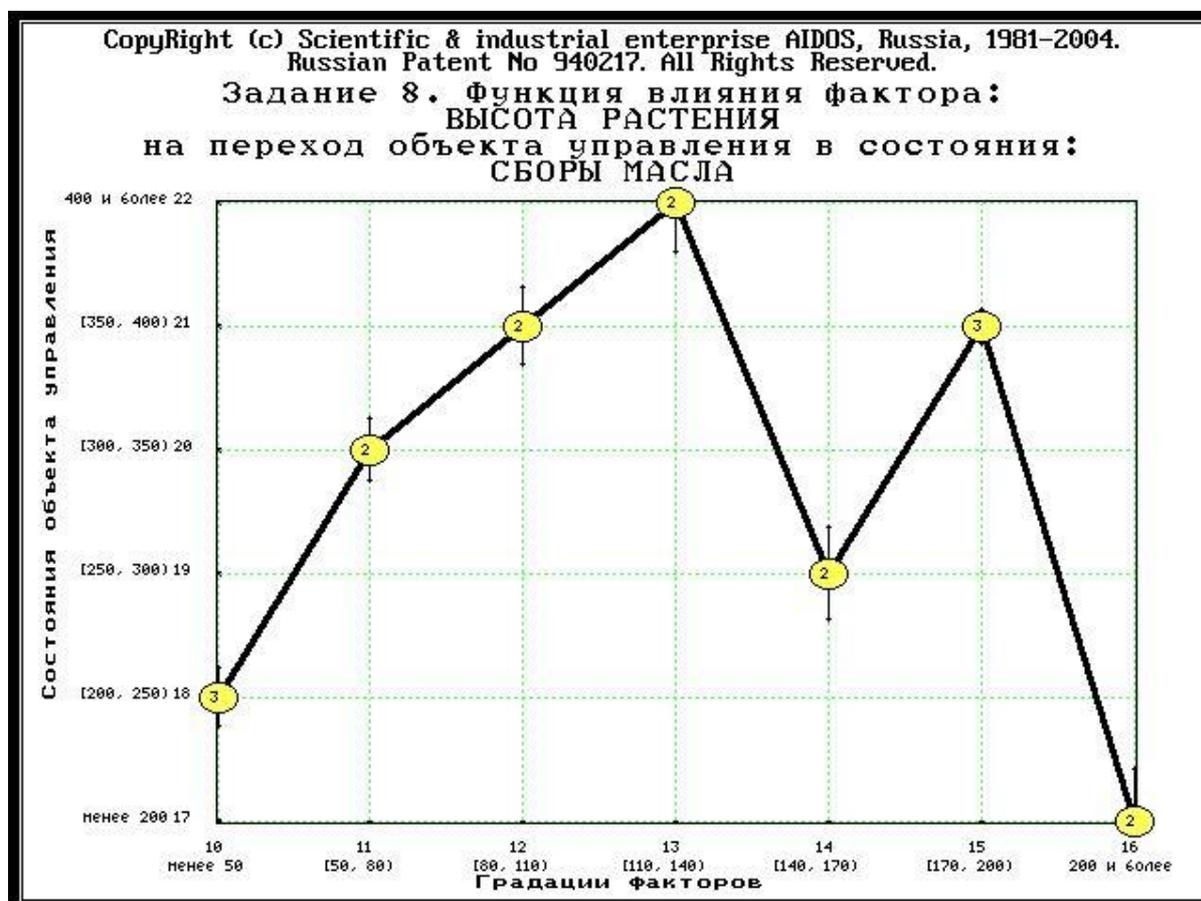
Рисунок 1 – Примеры карточек прогнозирования (экранные формы)

Карточка разделена на две части. В верхней части приведены классы, с которыми данный объект имеет наивысшее сходство в порядке его убывания, а в нижней – классы, от которых данный объект максимально отличается. Задача идентификации может решаться в поле с использованием ноутбука, на котором установлена система «Эйдос».

**Задача 3: разработка методики поддержки принятия решений по отбору растений для селекции не по их потребительским свойствам, а на основе анализа фенотипических признаков.** Задача принятия решений является обратной по отношению к задаче прогнозирования. Если при прогнозировании мы по фенотипическим признакам определяем возможные потребительские свойства, то при принятии решений – наоборот, по заданным потребительским свойствам определяем, какими фенотипическими признаками должны обладать растения, имеющие эти свойства. В системе «Эйдос» эта задача решается в 1-м режиме 5-й подсистемы, который позволяет генерировать и отображать так называемые «информационные портреты классов». Эти информационные портреты показывают системе детерминации будущих состояний объекта управления, в нашем случае – потребительских свойств подсолнечника.

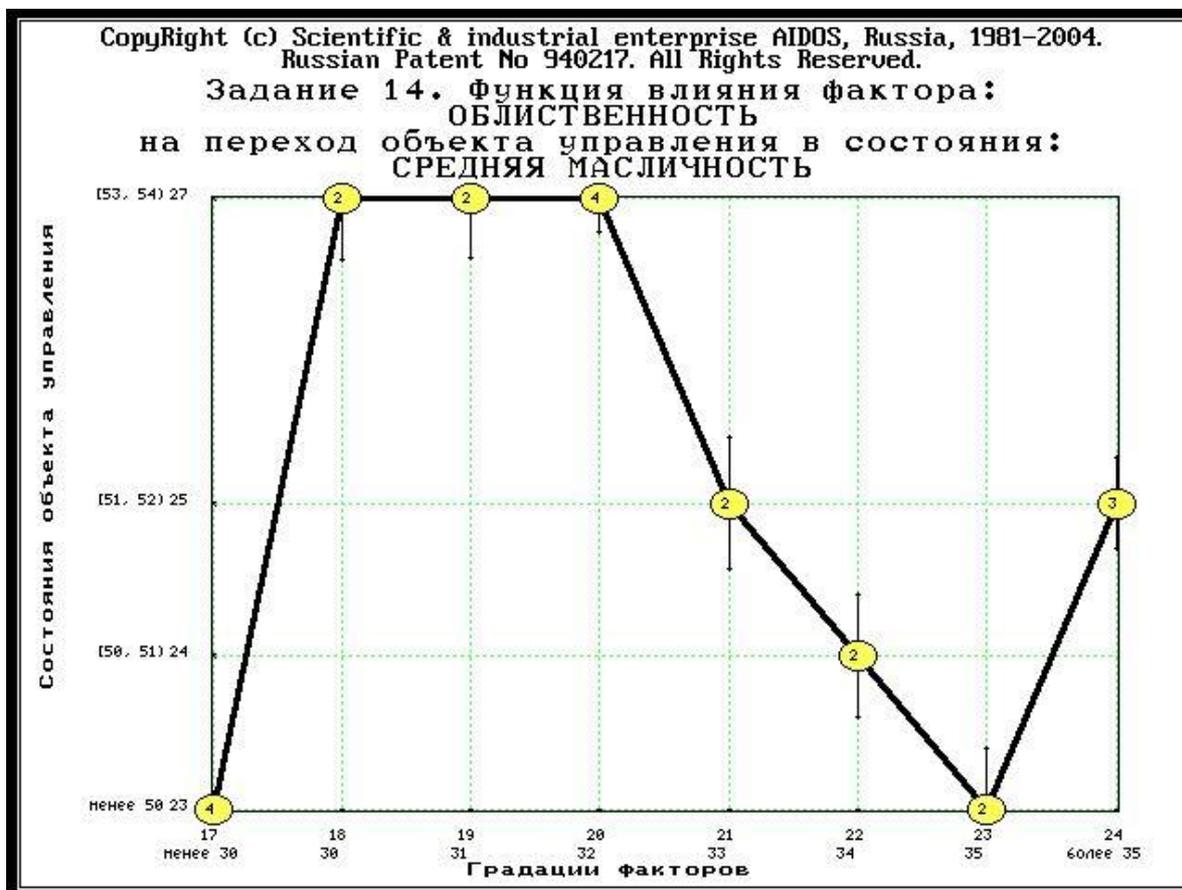
Они могут непосредственно использоваться при отборе растений подсолнечника для формирования следующего поколения при селекции.

Аппарат СК-анализа позволяет строить функции взаимосвязи между описательными и классификационными шкалами. С учетом целей ресурсосберегающих технологий возделывания подсолнечника [12] для селекции особый интерес представляет взаимосвязь между высотой растения и его потребительскими свойствами, прежде всего, сбором масла с одного растения (рис. 2).



**Рисунок 2 – Взаимосвязь между высотой растения  
и сбором с него масла**

Из этого графика очевидно, что оптимальной высотой растения является 110–140 см, а не 200, как считалось ранее. Объяснение причин этого явления сходно с содержательной интерпретацией взаимосвязи между количеством листьев на растении и масличностью собранных с него семян, представленной на рисунке 3.

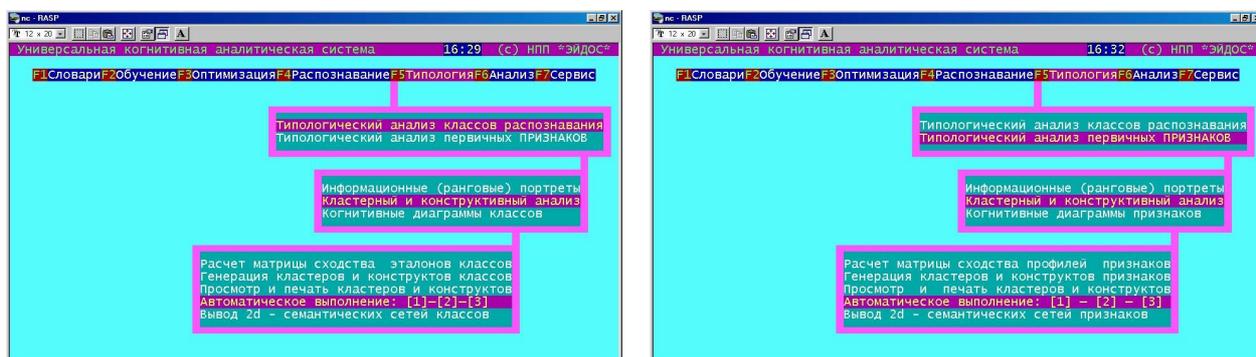


**Рисунок 3 – Взаимосвязь между количеством листьев на растении  
и масличностью собранных с него семян**

При малом количестве листьев растение не имеет необходимых ресурсов для генерации масла в семенах, при количестве листьев 30–32 наблюдается максимальная масличность, при увеличении количества листьев масличность закономерно падает, что, по-видимому, можно объяснить тем, что ресурсы растения переориентируются с формирования регенеративных качеств растения на увеличение его вегетативной массы. Этот эффект, наблюдаемый у подсолнечника, полностью аналогичен «бройлерному эффекту», известному в птицеводстве, и впервые был обнаружен методом СК-анализа эмпирических данных в 1993 году О.А. Засухиной и Е.В. Луценко при исследовании интенсивных технологий возделывания зерновых колосовых [9].

Все функции влияния, отражающие взаимосвязи между всеми описательными и классификационными шкалами и градациями (во всех сочетаниях), приведены в приложении 4. Однако необходимо отметить, что их содержательная интерпретация является делом ученого-селекционера и по этой причине в данной работе не рассматривается.

**Кластерно-конструктивный и системно-когнитивный анализ результатов выращивания и факторов** обеспечивается режимами 5-й подсистемы "Типология" системы «Эйдос» (рис. 4).



**Рисунок 4 – Режимы подсистемы типологического анализа классов и факторов**

Кластерный анализ классов показывает, какие качественные и количественные результаты выращивания различных сортов детерминируются (вызываются) сходными системами факторов и могут быть получены одновременно, а какие противоположными, несовместимыми и одновременно недостижимыми. В графической форме любые заданные фрагменты матрицы сходства отображаются в виде семантических сетей (рис. 5, 6 и 7).

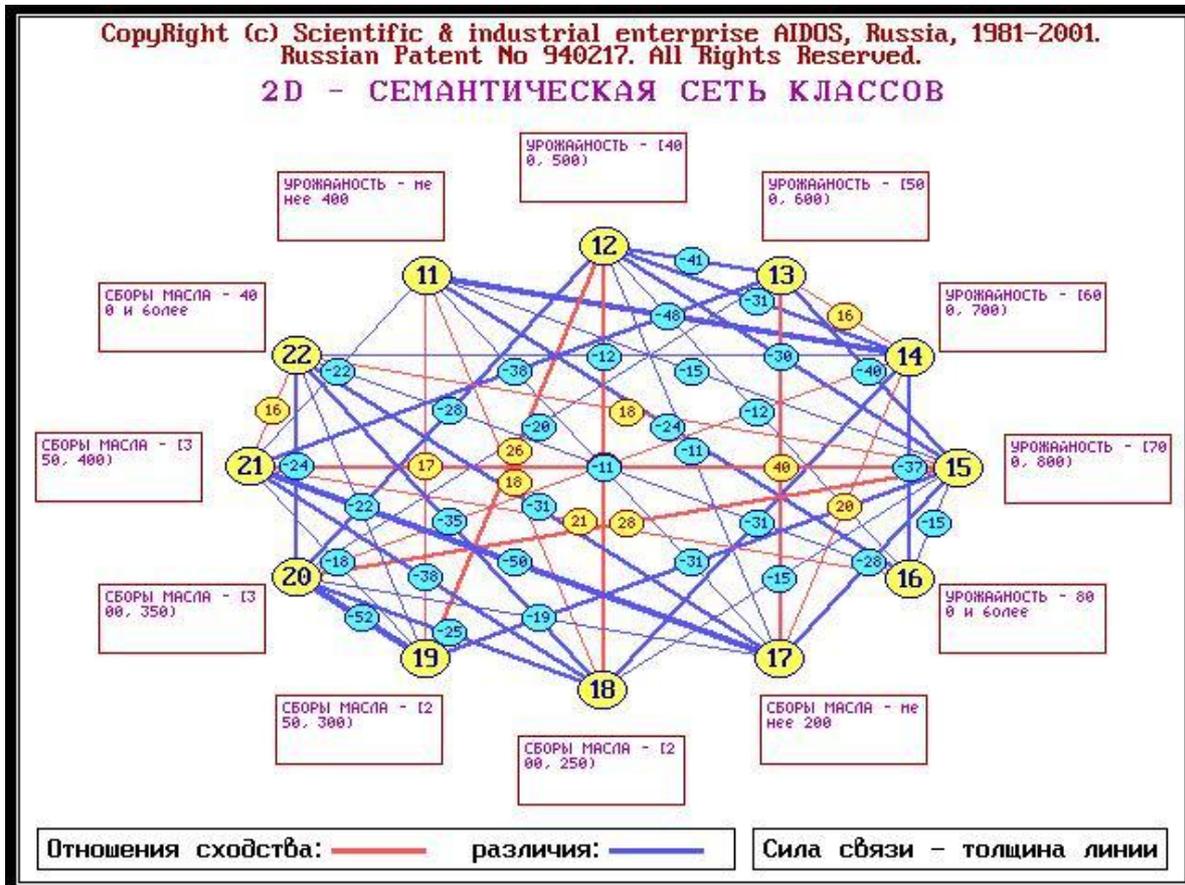
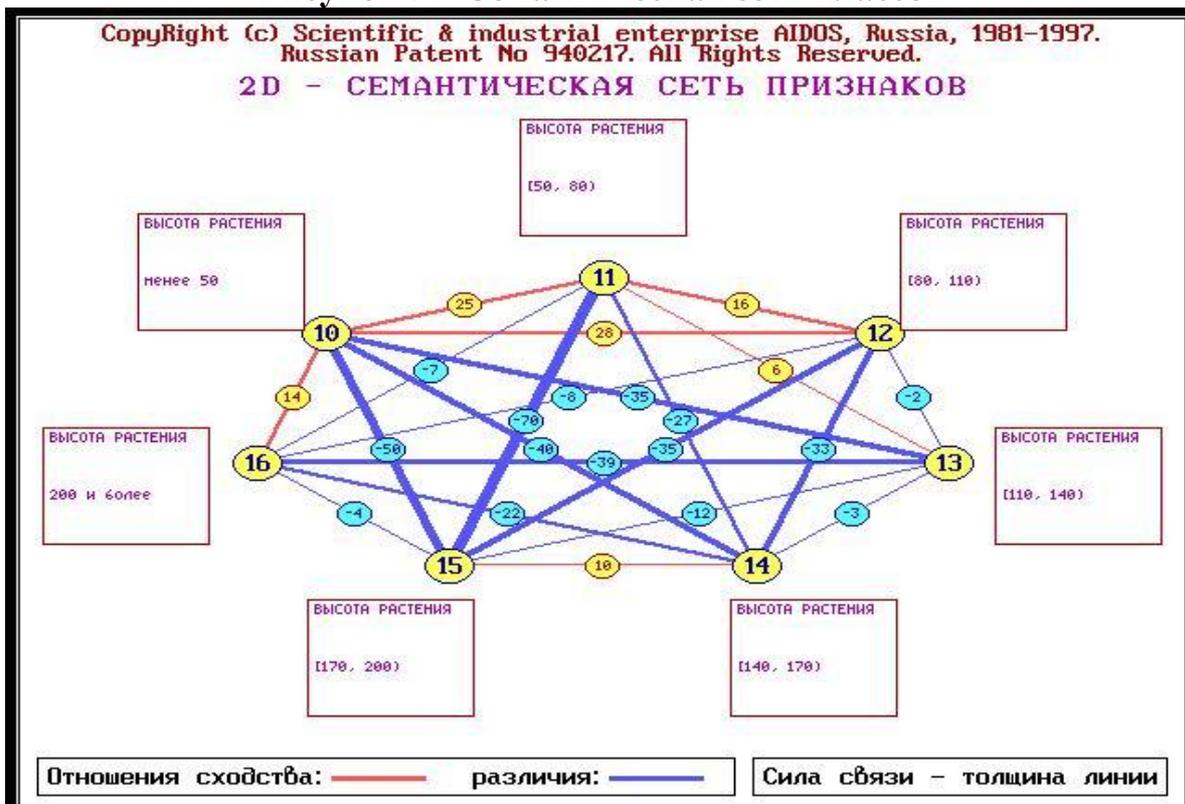
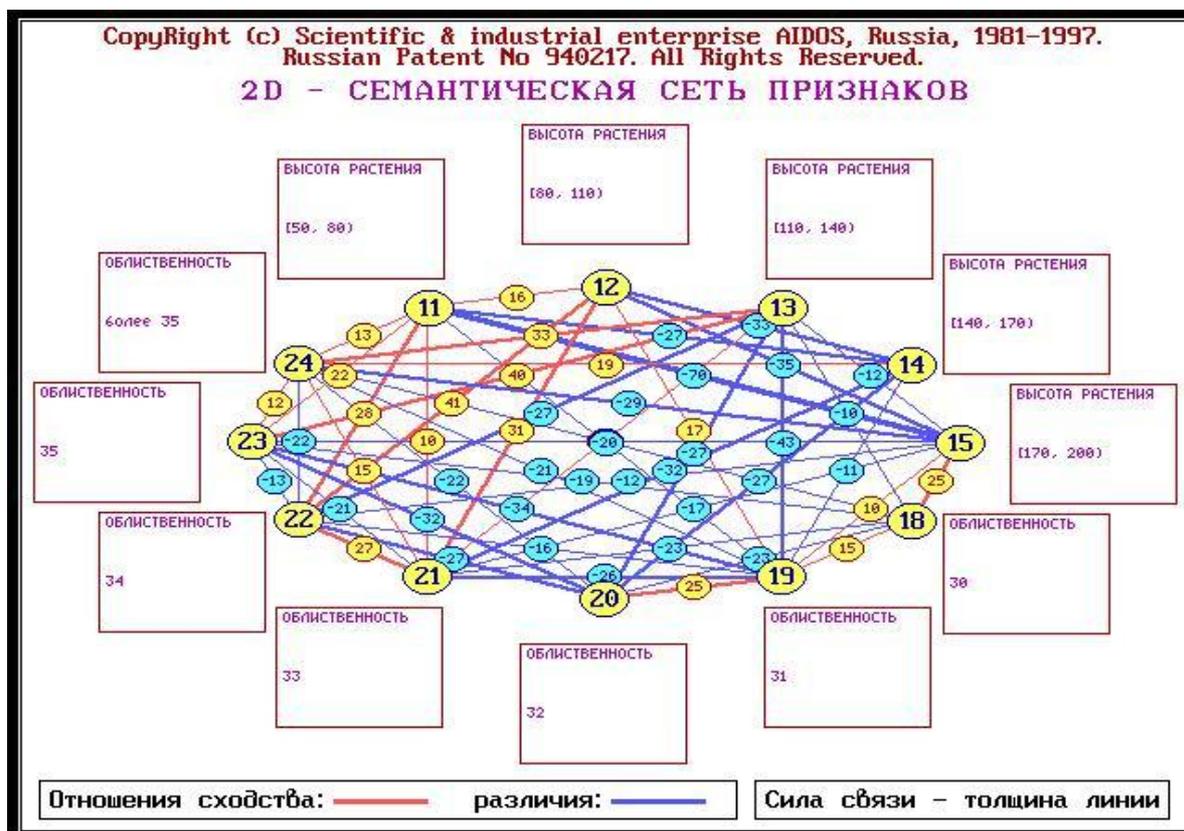


Рисунок 5 – Семантическая сеть классов

Рисунок 6 – Семантическая сеть фенотипических признаков:  
«Высота растения»



**Рисунок 7 – Семантическая сеть фенотипических признаков: «Облиственность» и «Высота растения»**

Таким образом, можно сделать вывод о том, что предложенная семантическая модель и технология СК-анализа обеспечили решение следующих задач.

**Задача 1** – выявление причинно-следственных зависимостей между фенотипическими признаками подсолнечника и его потребительскими свойствами.

**Задача 2** – разработка методики *прогнозирования* потребительских свойств растений подсолнечника на основе анализа их фенотипических признаков.

**Задача 3** – разработка методики *поддержки принятия решений* по отбору растений для селекции не по их потребительским свойствам, а на основе анализа фенотипических признаков.

## Список литературы

1. Драгавцева И.А. Применение системного анализа для прогнозирования успешности выращивания сельскохозяйственных культур (на примере плодовых) / И.А. Драгавцева, Е.В. Луценко, Л.М. Лопатина, Н.Е. Луценко // В сб. "Формы и методы повышения эффективности координации исследований для ускорения процесса передачи реальному сектору экономики завершенных разработок". – Краснодар: СКЗНИИСиВ, 2002. – С.62–67.
2. Драгавцева И.А. Автоматизация системного анализа продуктивности плодовых культур Юга России / И.А. Драгавцева, Е.В. Луценко, Л.М. Лопатина // Научные труды Российской академии сельскохозяйственных наук. – 2002. – С.11–14.
3. Драгавцева И.А. Применение автоматизированного системного анализа для прогноза продуктивности плодовых культур на Юге страны / И.А. Драгавцева, Е.В. Луценко, Л.М. Лопатина, Н.Е. Луценко // Материалы Всероссийской научно-практической конференции. –Краснодар: СКЗНИИСВ, 2002. – С.8–11.
4. Егоров Е.А. Интенсивные технологии возделывания плодовых культур / Е.А. Егоров, И.А. Драгавцева, Е.В. Луценко, Л.М. Лопатина и др. // Монография (научное издание): СКЗНИИСиВ. – Краснодар. –2004. – 394 с.
5. Драгавцева И.А., Луценко Е.В., Лопатина Л.М. Автоматизированная система мониторинга, анализа и прогнозирования развития сельхозкультур "ПРОГНОЗ-АГРО". Пат. №2003610433 РФ. Заяв. № 2002611927 РФ. Оpubл. от 18.02.2003.
6. Драгавцева И.А., Луценко Е.В., Лопатина Л.М. База данных автоматизированной системы мониторинга, анализа и прогнозирования развития сельхозкультур "ПРОГНОЗ-АГРО". Пат. №2003620035 РФ. Заяв. №2002620178 РФ. Оpubл. от 20.02.03.
7. Лопатина Л.М. Создание автоматизированной системы мониторинга, анализа, прогноза и управления продуктивностью сельскохозяйственных культур / Л.М. Лопатина, Е.В. Луценко // Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2003. – №02(2). – Режим доступа: <http://ej.kubagro.ru/2003/02/07/p07.asp>.
8. Лопатина Л.М. Концептуальная постановка задачи: "Прогнозирование количественных и качественных результатов выращивания заданной культуры в заданной точке" / Л.М. Лопатина, Е.В. Луценко, И.А. Драгавцева // Научный журнал Куб-

- ГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2004. – №05(7). – Режим доступа: <http://ej.kubagro.ru/2004/05/08/p08.asp>.
9. Луценко Е.В. Интеллектуальные информационные системы: Учебное пособие для студентов специальности: [351400 "Прикладная информатика \(по отраслям\)"](#). – Краснодар: КубГАУ. 2004. – 633 с.
  10. 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. – 605 с.
  11. Максимов А.М. Прогнозирование продуктивности и качества плодовых культур на основе применения систем искусственного интеллекта (по данным ЗАО Агрофирмы "Сад Гигант" Славянского района Краснодарского края): Дипломная работа. – Краснодар: КубГАУ, 2004. – 100 с.
  12. Калайджан А.А. Индуцированные макромутанты подсолнечника в биологизации защиты растений / А.А. Калайджан, Н.Г. Малюга // Трансгенные растения – новое направление в биологической защите растений. – Краснодар. – 2003. – С.233–238.