

УДК 004.6

Б.Л. ГОЛУБ\*, А.В. ГУДЗЬ\*, А.В. БУШМА\*\*

**ИНФОРМАЦИОННАЯ СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЯ  
В ПРОЦЕССЕ ВЫРАЩИВАНИЯ БИОТЕХНИЧЕСКИХ ОБЪЕКТОВ**

\*Национальный университет биоресурсов и природопользования Украины, г. Киев, Украина

\*\*Киевский университет имени Бориса Гринченко, г. Киев, Украина

**Анотація.** Стаття присвячена питанню вирішення проблем, пов'язаних з аналізом стану сільськогосподарських культур у процесі їх вирощування. Серед них виділені проблеми, пов'язані із своєчасним внесенням поживних речовин і виявленням уражених шкідливими речовинами рослин. Запропоновано систему, яка структурно складається з трьох підсистем: первинного збору, зберігання, аналізу. До складу першої підсистеми входять хронофлуорометр і пристрій для вимірювання кількості мікотоксинів. Серед широкого кола параметрів, які характеризують стан сільськогосподарських культур, розглянуто такі: наявність мікотоксинів, вміст хлорофілу в рослині з вимірюванням внесених додаткових хімічних добрив і біологічних добавок та ін. Також як додаткове джерело даних використовується розрахунок індексу вегетації рослини NDVI на основі супутникових знімків. Друга підсистема представлена з технічної точки зору комп'ютером. З інформаційної точки зору – реляційною базою даних і сховищем. Третя підсистема реалізує аналіз даних за допомогою сучасних інформаційних технологій, таких як OLAP і Data Mining. Наведено результати застосування цих технологій для аналізу наявності мікотоксинів. Представлена архітектура системи, основні функції якої – управління проведенням вимірювання параметрів росту рослин за допомогою різних біосенсорів, передача отриманих даних на мобільний пристрій з подальшим накопиченням їх в центральне сховище, аналіз і прогнозування процесів, пов'язаних з вирощуванням культур. Оскільки в системі використовуються методи інтелектуального аналізу даних, це дасть можливість знаходити нові знання або закономірності. Завдяки цьому, передбачається підвищення ефективності вирощування культур.

**Ключові слова:** параметри росту, мікотоксини, біосенсори, хронофлуорометр, хлорофіл, добрива, Android-додаток, сервер, сховище даних, OLAP, Data Mining, KPI.

**Аннотация.** Статья посвящена вопросу решения проблем, связанных с анализом состояния сельскохозяйственных культур в процессе их выращивания. Среди них выделены проблемы, связанные со своевременным внесением питательных веществ и выявлением пораженных вредными веществами растений. Предложена система, структурно состоящая из трёх подсистем: первичного сбора, хранения, анализа. В состав первой подсистемы входят хронофлуорометр и устройство для измерения количества микотоксинов. Среди широкого круга параметров, которые характеризуют состояние сельскохозяйственных культур, рассмотрены такие: наличие микотоксинов, содержание хлорофилла в растении с измерением внесённых дополнительных химических удобрений и биологических добавок и др. Также в качестве дополнительного источника данных используется расчет индекса вегетации растения NDVI на основе спутниковых снимков. Вторая подсистема представлена с технической точки зрения компьютером. С информационной точки зрения – реляционной базой данных и хранилищем данных. Третья подсистема реализует анализ данных с помощью современных информационных технологий, таких как OLAP и Data Mining. Приведены результаты применения этих технологий для анализа наличия микотоксинов. Представлена архитектура системы, основные функции которой – управление проведением измерения параметров роста растений с помощью различных биосенсоров, передача полученных данных на мобильное устройство с дальнейшим накоплением их в центральном хранилище, анализ и прогнозирование процессов, связанных с выращиванием культур. Поскольку в системе используются методы интеллектуального анализа данных, это даст возможность находить новые знания или закономерности. Благодаря этому, предполагается повысить эффективность выращивания культур.

**Ключевые слова:** параметры роста, микотоксины, биосенсор, хронофлуорометр, хлорофилл, удобрения, Android-приложение, сервер, хранилище данных, OLAP, Data Mining, KPI.

**Abstract.** The article is devoted to the problem of solving problems associated with the analysis of the state of agricultural plants in the process of their cultivation. Among these problems there were highlighted the problems associated with the timely introduction of nutrients and the identification of plants affected by harmful substances. A system consisting of three subsystems is proposed: primary collection, storage, analysis. The first subsystem includes a chronofluorometer and a device for measuring the number of mycotoxins. Among the wide range of parameters that characterize the condition of agricultural plants, the following are considered: the presence of mycotoxins, the content of chlorophyll in the plant with the measurement of additional chemical fertilizers and biological additives, and etc. Also, as an additional data source, we use the NDVI plant vegetation index calculation based on satellite images. The second subsystem is represented from a technical point of view by a computer. On the informational point of view, relational and data storage. The third subsystem implements data analysis using modern information technologies such as OLAP and Data Mining. The results of the application of these technologies for the analysis of mycotoxins are given. The system architecture is presented, the main functions of which are: managing the measurement of plant growth parameters using various biosensors, transferring the obtained data to a mobile device with further accumulation in the central repository, analyzing and predicting the processes associated with growing crops. Since the system uses data mining methods, it will provide an opportunity to find new knowledge or patterns. Due to this, it is supposed to increase the efficiency of growing crops.

**Keywords:** growth parameters, mycotoxins, biosensors, chronofluorometer, chlorophyll, fertilizers, Android application, server, data storage, OLAP, Data Mining, KPI.

## 1. Введение

Одним из основных факторов экономического развития аграрного сектора государства является эффективная оценка состояния посевов, в том числе, своевременное внесение питательных веществ и выявление пораженных вредными веществами растений. Исследование этих параметров трудно реализовать в связи с отсутствием инструментов проведения тестирования на месте, особенно в полевых условиях, хотя именно они существенно влияют на результат выращивания сельскохозяйственных культур.

Измерение содержания хлорофилла позволяет определить оптимальные дозы химических удобрений и биологических добавок, а обнаружение вредных веществ – уберечь растения от гибели. Одними из наиболее агрессивных вредных веществ являются микотоксины. Микотоксины – биологические контаминанты, природные загрязнители зерна злаковых, бобовых, семян подсолнечника, а также овощей и фруктов. Они могут образовываться как в растениях на стадии вегетации, так и при хранении во многих пищевых продуктах под действием развивающихся в них микроскопических грибов. Продовольственная и сельскохозяйственная организация ООН FAO отмечает, что ежегодные потери пищевых продуктов из-за влияния микотоксинов составляют около 1 млрд тонн, а 4,5 млрд людей во всем мире подвергаются риску воздействия этих веществ.

Также немаловажным является слежение за состоянием вегетации растения на основе определения индекса NDVI [1]. NDVI (Normalized Difference Vegetation Index) – нормализованный относительный индекс растительности – простой количественный показатель количества фотосинтетически активной биомассы (обычно называемый вегетационным индексом). Это один из самых распространенных и используемых индексов для решения задач, использующих количественные оценки растительного покрова. Более точное определение и расчет индекса представлены в [1]. Этот индекс можно определять на основе снимков полей, сделанных как спутниками земли, так и беспилотными летательными аппаратами (дронами).

Целью работы является разработка архитектуры системы поддержки принятия решения при выращивании биотехнических объектов. Основными функциями такой системы

являются управление проведением измерения параметров роста растений с помощью различных биосенсоров и внешних источников, передача полученных данных на мобильное устройство с дальнейшим накоплением их в центральном хранилище, анализ полученных данных и прогнозирование процессов, связанных с выращиванием сельскохозяйственных культур. Предполагается, что внедрение такой системы позволит повысить эффективность выращивания широкого спектра сельскохозяйственных культур.

Проблемная ситуация в процессе определения содержания хлорофилла и микотоксинов в настоящий момент заключается в том, что диагностику можно провести только в лаборатории. Для этого требуется от нескольких дней до нескольких недель. Альтернатива лабораторным исследованиям – использование сенсоров, которые способны провести экспресс-анализ в полевых условиях. Сегодня уже существуют биосенсоры, выполняющие качественный анализ за короткий промежуток времени [2]. Однако большинство таких сенсоров работают автономно, без централизованного накопления данных тестирования, а некоторые вообще не имеют возможности сохранять полученную информацию. Как результат, невозможно быстро оценить статистические данные или провести какой-либо их анализ. Таким образом, результат использования биосенсоров становится неэффективным с точки зрения реализации системы поддержки принятия решений.

Для решения проблемы повышения эффективности использования биосенсоров необходима разработка специальных компьютерных информационных технологий, которые реализовали бы такие функции:

- получение результатов анализов от соответствующего сенсора;
- надежное и долгосрочное хранение данных;
- интеллектуальный анализ полученных данных, в том числе, анализ данных в режиме реального времени.

## **2. Основная часть**

Для реализации поставленных задач разрабатывается система, структура которой состоит из трех подсистем:

- первичного сбора данных;
- хранения полученной информации;
- анализа накопленных результатов.

Данные в систему попадают от биосенсорных устройств под управлением агрономов. Также входной информацией служит расчет индекса NDVI по снимкам со спутников или дронов. Кроме того, внешние источники могут давать климатические, погодные и другие данные о различных параметрах окружающей среды.

С технической точки зрения, подсистема хранения состоит из компьютерных устройств серверного типа. С программной точки зрения, она представлена реляционными базами данных (БД) и хранилищем данных. Поскольку необходимо агрегировать данные из различных источников, общая схема информационного обеспечения системы будет выглядеть следующим образом. В центре подсистемы находится хранилище собранных данных, а их источниками служит набор реляционных БД.

Подсистема анализа данных решает задачи по трем основным направлениям: информационно-поисковому, анализу в режиме реального времени, а также реализует интеллектуальную обработку накопленной информации.

В результате аналитик с помощью системы сможет получить статистические данные, новую информацию или гипотезу, а также подтвердить или отвергнуть уже существующие гипотезы. Кроме того, аналитик сможет подготовить для руководства ряд рекомендаций или отчетов, которые станут основой для принятия правильных решений при выращивании биотехнических объектов.

Общая структура системы поддержки принятия решений при выращивании биотехнических объектов показана на рис. 1.

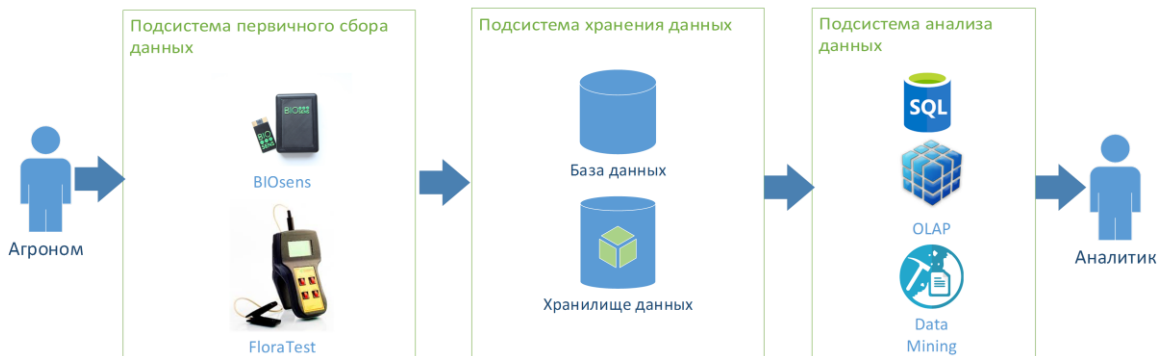


Рисунок 1 – Структура системы поддержки принятия решений при выращивании биотехнических объектов

### 2.1. Подсистема первичного сбора данных

Один из наиболее эффективных источников данных для разрабатываемой системы хронофлуорометр «Флоратест» позволяет быстро оценить степень воздействия природной окружающей среды и загрязнений на живые растения [3]. Работа прибора основана на измерении кривой индукции флуоресценции хлорофилла в реальном времени, что позволяет определить оптимальные дозы химических удобрений и биологических добавок. Значения этой кривой отображают текущее общее состояние растения, то есть его здоровье. Соответственно, для здорового растения может вообще не понадобится вносить удобрения, а, с другой стороны, агроном может своевременно обратить внимание на определенное растение и начать вносить нужные ему удобрения.

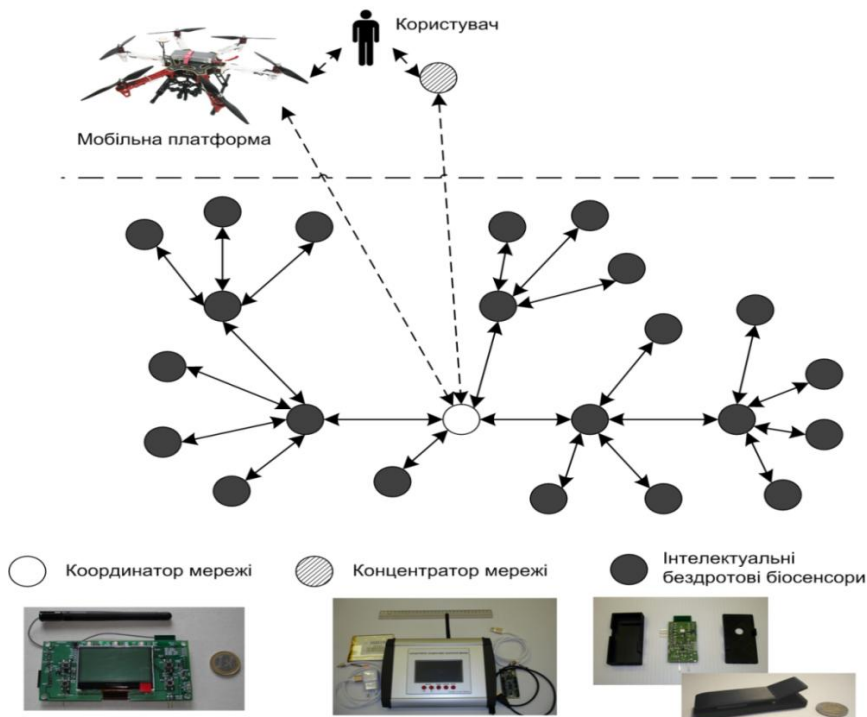


Рисунок 2 – Схема беспроводной системы для устройства «Флоратест» [4]

С позиции ведения точного земледелия, экологического мониторинга и охраны окружающей среды, мобильные и традиционные технологии сбора данных играют ключевую роль. Поэтому в Институте кибернетики им. В.М. Глушкова НАНУ создана беспроводная система сбора данных, которая в реальном времени получает информацию о состоянии растений на территории сельскохозяйственных угодий или зеленых насаждений [4].

В соответствии со схемой размещения, представленной на рис. 2, беспроводная система сбора данных состоит из:

- интеллектуальных беспроводных биосенсоров «Флоратест»;
- координаторов сети;
- концентраторов сети;
- мобильных платформ.

Функционально данные от биосенсоров поступают в координаторы сети. При потребности, данные от координаторов через концентратор сети или мобильную платформу поступают непосредственно в локальную БД пользовательского компьютера. В результате пользователь на своем компьютере может увидеть актуальную информацию о состоянии растений в режиме реального времени.

Такая беспроводная система в режиме реального времени может собирать данные о состоянии растений и передавать их по различным каналам связи в подсистему хранения информации.

Для анализа микотоксинов в пищевых продуктах и кормах существует ряд тестовых решений. Более подробно эти методы, их плюсы и минусы описаны в [5]. Эти решения варьируются от быстрых тестов, которые легко проводить, до трудоемких базовых методов, обеспечивающих более информативные результаты. Основные доступные быстрые и базовые методы перечислены в табл. 1.

Таблица 1 – Основные доступные методы определения микотоксинов [5]

Быстрые методы	Базовые методы
LateralFlowTest	ThinLayerChromatography (TLC)
Enzyme-Linked Immunosorbent Assay (ELISA)	Gas Chromatography
Fluorometry	High Performance Liquid Chromatography (HPLC)
	Liquid Chromatography – Mass Spectrometry (LC/MS)

В мире на данный момент разрабатывается множество устройств, которые могут быстро проводить анализы. Но, в основном, в таких устройствах можно просто увидеть результат текущего анализа. Однако в дальнейшем эта информация нигде централизованно не сохраняется и не обрабатывается для получения новых знаний о микотоксинах.

Одним из устройств для проведения быстрого анализа на наличие микотоксинов является прибор «BIOsens» [6]. Он предназначен для определения количества микотоксинов в растениях и продуктах. Устройство состоит из контроллера, модуля передачи и приема данных по протоколу Bluetooth, а также модуля для считывания вольтамперных характеристик с чувствительной поверхности.

Расчет значения NDVI происходит на основе фотоснимков. По этой причине важно иметь снимки высокого качества, максимально свободные от различных помех, таких как тени облаков и туч. Решения по очистке изображений описаны в [7]. Также высокий показатель индекса не всегда свидетельствует о хорошем развитии культуры растений. Это связано с тем, что на его величину влияют показатели сорняков. Большое количество сорняков повышает индекс, но, в целом, это не указывает на хорошее развитие растений. В

таких случаях необходимо разрабатывать отдельный алгоритм интеграции данных в систему, который будет отфильтровывать некорректные результаты.

Эти проблемы вызывают необходимость определять параметры роста сельскохозяйственных культур в полевых условиях. Поэтому в качестве устройства для сбора и первоначальной обработки информации целесообразно использовать любое мобильное устройство типа смартфона или планшета.

Для управления и получения результатов от биосенсорного датчика «BIOsens» было разработано мобильное приложение на основе операционной системы Android с использованием языка программирования Java [8].

## 2.2. Подсистема хранения данных

Общая схема информационного обеспечения состоит из одного центрального хранилища и множества БД, которые служат источниками информации для хранилища. Кроме того, источниками могут служить и различные другие форматы данных, например, таблицы «Excel» или данные в формате xml-файла. В основном, для накопления и длительного хранения информации от датчиков используются реляционные БД. При этом применяется технология OLTP. На рис. 3 представлена структура БД, предназначенная для накопления информации для решения задачи определения уровня микотоксинов.

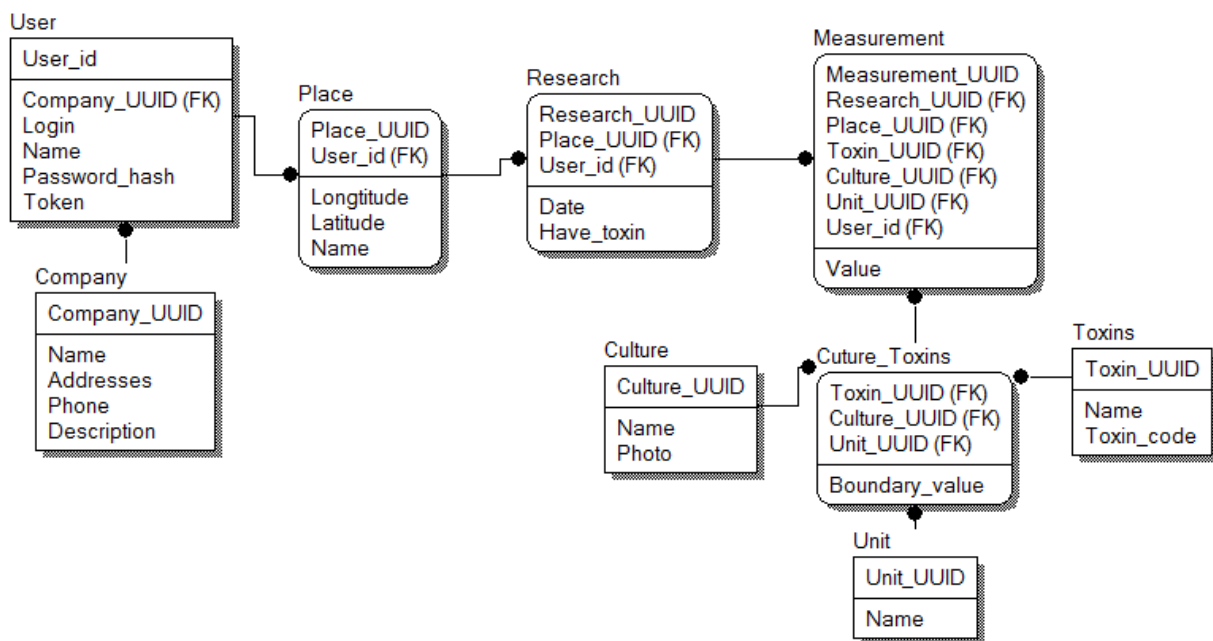


Рисунок 3 – Структура БД

БД работает в среде PostgreSQL. Ее структура (рис. 3) хранит информацию, полученную от биосенсора, измеряющего уровень микотоксинов. Реляционная БД приведена в третью нормальную форму (3НФ) [9]. Отношение находится в 3НФ, когда находится во 2НФ и каждый неключевой атрибут нетранзитивно зависит от первичного ключа. Иными словами, второе правило требует выносить все неключевые поля, содержимое которых может относиться к нескольким записям таблицы, в отдельные таблицы.

Витрина данных – срез хранилища данных, представляющий собой массив тематической, узконаправленной информации, ориентированный, например, на пользователей одной рабочей группы или департамента. Для хранилища данных сформирован гиперкуб, управляемый MS SQL Server. На рис. 4 представлена витрина данных, которая описывает структуру хранения данных о результатах анализа на наличие микотоксинов.

Видно, что структура состоит из таблицы фактов, то есть мер, которые хранят количественные показатели результатов анализа на микотоксины, и таблиц измерений, которые описывают данные. Измерения представляют качественными показателями:

- культура растения;
- вид микотоксина;
- погодные условия;
- пользователь системы;
- предприятие;
- регион;
- страна;
- дата.

При этом обязательной мерой для любого хранилища (витрины) данных есть мера времени. В представленной витрине мерой времени есть таблица “DimDate”.

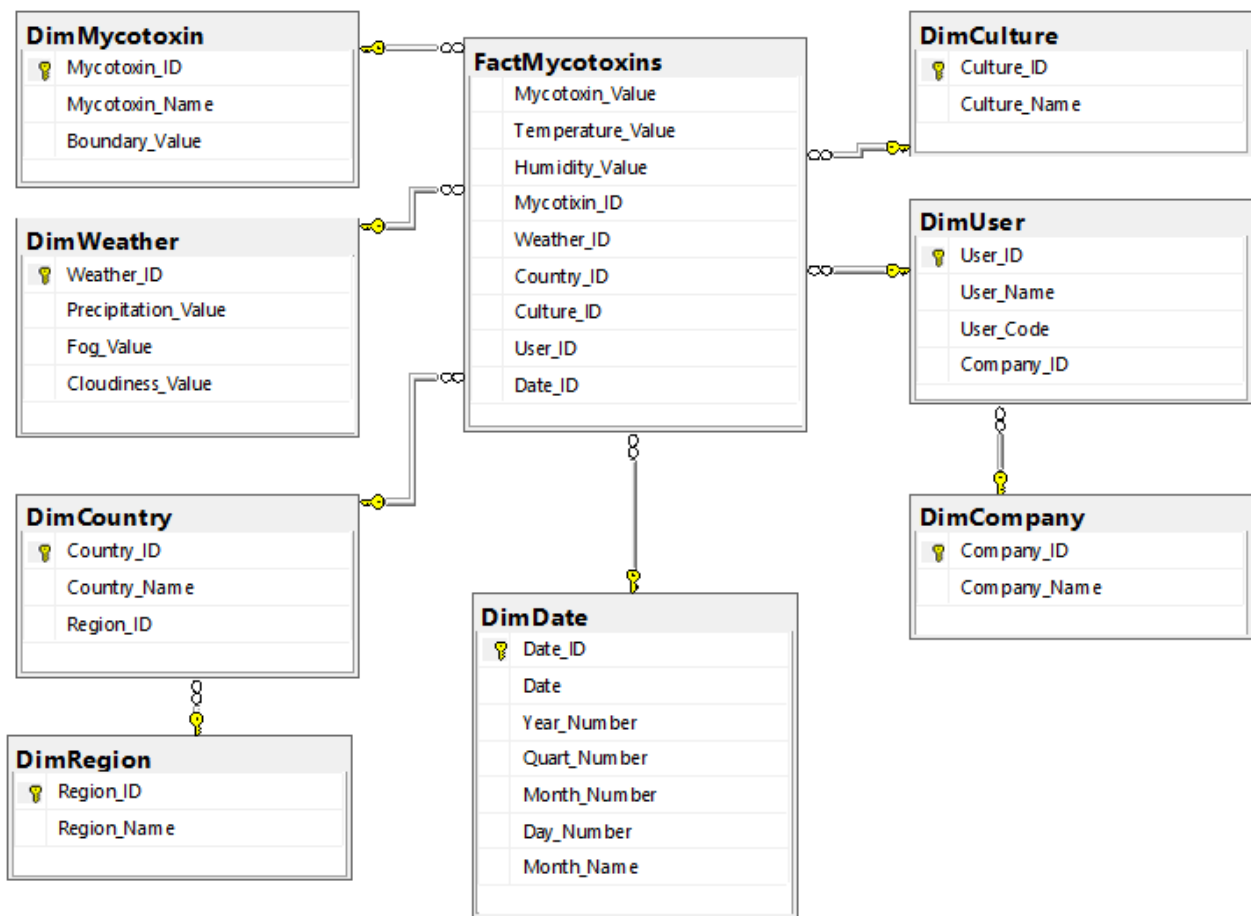


Рисунок 4 – Многомерная структура витрины данных

В системе существуют ряд внешних источников и хранилище данных. Для того, чтобы данные через определенный промежуток времени собирались из источников и поступали в хранилище, необходимо провести их интеграцию. Для обеспечения этого процесса в среде BI MS SQL Server создан ETL-пакет (Extract, Transform, Load) – процесс, который используется для организации перемещения и хранения данных в хранилище. Этот процесс включает три основных этапа: извлечение данных из источников (Extract), их преобразование (Transform) и загрузку в хранилище (Load). На рис. 5 представлена общая схема взаимодействия источников данных. Видно, что вначале все данные из источников проходят ETL-процесс. В результате ETL-процесса формируются достоверные и очищенные данные, которые загружаются в хранилище.

Интеграция данных настраивается для каждого источника отдельно. В этом случае нужно правильно соотнести данные, которые хранятся в ячейках реляционной БД с аналогичными ячейками для этих данных в самом хранилище. При этом, возможно, понадобится провести ряд преобразований для выполнения корректной интеграции. Немаловажным этапом есть очистка данных, поскольку пустые или некорректные значения могут повлиять на достоверность конечных результатов анализа. Для очистки данных разрабатываются отдельные алгоритмы, которые в ходе интеграции необходимо постоянно совершенствовать.

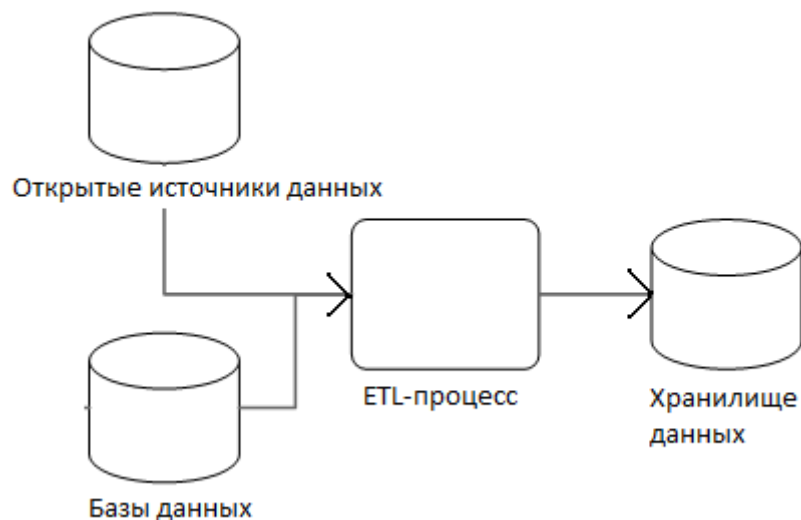


Рисунок 5 – Схема потоков данных

Представляет интерес разработка прямой интеграции данных, собранных беспроводной сетью биосенсоров «Флоратест», данных индекса развития растений и данных климатических условий от внешних источников.

### 2.3. Подсистема анализа данных

Эта подсистема представлена тремя технологиями:

- информационно-поисковый анализ на основе использования языка SQL;
- анализ данных в режиме реального времени OLAP;
- интеллектуальный анализ данных DataMining [10].

Первая из этих технологий позволяет, используя возможности языка SQL, реализовать отбор и обработку данных с целью получения некоторой статистической информации, актуальной в момент отбора. Технология OLAP позиционируется как метод получения важных характеристик исследуемого процесса, основная из которых – величина KPI (Key Performance Indicator) [11]. Кроме того, с помощью OLAP могут быть сформированы отчеты любой сложности и получен тренд параметров исследуемого процесса. Технология DataMining позиционируется как совокупность инструментария, позволяющая проводить интеллектуальный анализ данных [12]. В этой системе используется метод поиска ассоциативных правил, метод для выявления определенных шаблонов в наборе данных.

При работе над поиском правил аналитик также может задать значение минимальной поддержки. Это дает возможность получить так называемый частый набор (large itemset), который состоит из правил, являющихся более точными среди интересных для аналитика.

Однако довольно часто возникает потребность не только определять частоту вхождения объекта в транзакции, но и их последовательность в транзакциях. Поиск правил с



определенной последовательностью называется сиквенциальный анализ. Такой анализ может обеспечить прогнозирование наступления события при выявлении определенной последовательности событий. Так, при проведении поиска ассоциативных правил было определено, что правилом с наибольшей вероятностью наступления (почти 100%) является следующее: при влажности воздуха более, чем 64%, концентрация микотоксинов составляет не менее 1,24 мг/кг. При этом поддержка наборов, входящих в это правило, является одной из самых высоких и составляет 15. Учитывая проведенный анализ, выявленную гипотезу можно использовать как правило.

Разработанная система находится в режиме тестирования, поэтому данные для проведения анализа являются тестовыми и не представлены реальными показателями.

Для подтверждения или опровержения гипотезы был проведен расчет ключевого показателя эффективности, который позволяет оценить эффективность выполняемых действий – KPI.

При расчете система использует ряд выражений:

- выражение значения – для отображения текущей величины KPI;
- выражение цели – определяет значение, которому должна соответствовать его текущая величина при идеальных условиях;
- выражение статуса – определяет статус текущего значения в соответствии с целевым значением;
- выражение тренда – определяет условия, необходимые для расчета тренда.

В качестве выражения значения используется средняя величина содержания микотоксинов, где целевым является условие, для которого при среднем значении влажности более 40% содержание микотоксинов должно быть больше, чем 0,1 мг/кг [13, 14].

Результаты расчета индекса следующие:

- среднее значение содержания микотоксинов равно 0,34 мг/кг;
- текущее значение соответствует целевому;
- статус положительный;
- тренд нулевой.

Для того, чтобы пользователь системы мог оперативно получать информацию о состоянии растений или результаты анализа данных, разрабатывается Web-интерфейс. Использование Web-технологий позволяет вывести результаты с помощью любого браузера в удобном для пользователя виде, в том числе и на мобильных устройствах.

### 3. Выводы

Разрабатываемая система поддержки принятия решения при выращивании биотехнических объектов многофункциональна. Она позволяет фиксировать параметры роста растений в полевых условиях, передавать полученные результаты на мобильные устройства и отображать их на экране. Поскольку важно иметь доступ к полученным результатам в удаленном режиме, используются Web-технологии. Для проведения анализа данных, который возможен при достаточно большом объеме накопленной информации, используются хранилище данных и соответствующие технологии – OLAP и DataMining.

В ходе тестирования системы с помощью поиска ассоциативных правил был получен ряд гипотез, подтвержденных с помощью OLAP-технологии. Разрабатываемая система нуждается в дальнейшем комплексном тестировании и практической оптимизации, апробации других методов интеллектуального анализа данных, накоплении результатов измерений, но с концептуальной точки зрения она может быть рекомендована для внедрения на аграрных предприятиях.

## СПИСОК ИСТОЧНИКОВ

1. NDVI – теория и практика. URL: <http://gis-lab.info/qa/ndvi.html> (дата звернення: 01.10.2018).
2. Starodub N.F., Pilipenko I.V., Pilipenko L.N., Katsev A.M. Express Control of Toxicity and Content of Patulin by Optical Biosensors. *Nanotechnology 2010: Bio Sensors, Instruments, Medical, Environment and Energy*. 2010. P. 137–140.
3. Флоратест. URL: <http://www.dasd.com.ua/floratest.php?lang=2> (дата звернення: 05.10.2018).
4. Романов В.О., Палагін О.В., Галелюка І.Б., Вороненко О.В. Безпроводна сенсорна мережа для прецизійного землеробства та екологічного моніторингу. *Комп'ютерні засоби, мережі та системи*. 2014. № 13. С. 53–62.
5. MYCOTOXIN ANALYSIS. URL: <http://www.mycotoxins.info/analysis/mycotoxin-analysis/> (дата звернення: 25.09.2018).
6. BIOsens. URL: <http://sens.bio/> (дата звернення: 06.10.2018).
7. Басараб Р.М. Індуктивні методи попередньої обробки даних ДЗЗ (Січ-2): автореф. ... канд. техн. наук: 05.07.12. К.: Державне космічне агенство України. Інститут космічних досліджень, 2015. 20 с.
8. Гудзь О.В. Розробка програмного забезпечення для системи проведення аналізу наявності мікотоксинів. *Інтернет-конференції НУБІП України. «Глобальні та регіональні проблеми інформатизації в суспільстві і природокористуванні 2016»*. 2016. URL: <http://econference.nubip.edu.ua/index.php/grpi/grpi16/paper/view/406> (дата звернення: 01.09.2018).
9. Третья нормальная форма. URL: <https://goo.gl/dpq2tE> (дата звернення: 05.09.2018).
10. Барсегян А.А., Куприянов М.С., Холод И.И., Тесс М.Д., Елизаров С.И. Технологии анализа данных: Data Mining, Text Mining, Visual Mining, OLAP. 2-е изд. СПб.: БХВ-Петербург, 2009. 512 с.
11. KPI. URL: <https://goo.gl/N1XA1q> (дата звернення: 06.09.2018).
12. Голуб Б.Л. Анализ деятельности предприятия. *Наука и жизнь Израйля*. 2015. № 8. URL: <http://nizi.co.il/author/bella-golub> (дата звернення: 01.08.2018).
13. Гудзь О.В. Information technology insolving problems the of contentanally sisofmycotox insincrops. *Інтернет-Конференції НУБІП України. VII міжнар. наук.-практ. конф. молодих вчених «Інформаційні технології: Економіка, Техніка, Освіта»*. 2016. URL: <http://econference.nubip.edu.ua/index.php/itete/vii/paper/view/649> (дата звернення: 25.09.2018).
14. Голуб Б., Гудзь А. Использование компьютерных технологий при решении задачи анализа содержания вредных веществ в сельскохозяйственных культурах. *Сборник трудов междисциплинарной научной конф. «Многофакторные подходы к формированию комфортной среды»*. Нетания, Израиль, 2017. С. 225–230.

Стаття надійшла до редакції 27.09.2018