

РАЗРАБОТКА БАЗЫ ДАННЫХ ИНФОРМАЦИОННОЙ СИСТЕМЫ ЦИТОМОР ФОЛОГОБИОФИЗИЧЕ СКОЙ ДИАГНОСТИКИ ОРГАНИЗМА ЧЕЛОВЕКА

Е. В. Высоцкая

Кандидат технических наук, доцент*

Контактный тел.: (057) 702-14-64

E-mail: diagnost@kture.kharkov.ua

И. Ю. Панферова

Кандидат технических наук, доцент

Кафедра «Информационные управляющие системы»**

А. П. Порван

Научный сотрудник*

Контактный тел.: (057) 702-14-64

E-mail: diagnost@kture.kharkov.ua

Н. А. Щукин

Инженер 1-ой категории*

Контактный тел.: 093-082-91-92

E-mail: KolaShuk@rambler.ru

*Кафедра биомедицинской электроники**

**Харьковский национальный университет

радиоэлектроники

пр. Ленина 14, г. Харьков, Украина, 61166

Робота присвячена розробці бази даних інформаційної системи цитоморфологобіофізичної діагностики організму людини, яка призначена для покращення ведення документації з відвідувань пацієнтів та швидкого доступу к даним для їхньої візуального відображення

Ключові слова: база даних, діагностика, модель схеми даних

Работа посвящена разработке базы данных информационной системы цитоморфологобиофизической диагностики организма человека, которая предназначена для улучшения ведения документации о посещениях пациентов и быстрого доступа к данным для их визуального отображения

Ключевые слова: база данных, диагностика, модель схемы данных

This article highlights the developing a database information system diagnostics of the human body, which is designed to improve record keeping of patient visits and fast access to data for their visual display

Keywords: a database, diagnostics, model of the scheme of the data

1. Введение

В силу все более широкого распространения персональных компьютеров в медицине широко разрабатываются и внедряются информационные системы, которые позволяют улучшить ведение медицинской документации, проведение статистического учета и анализа медицинской информации, сопровождение наблюдения за пациентами [1]. Одним из аспектов разработки информационных систем в медицине является создание и ведение баз данных (БД), позволяющих хранить данные как о пациенте, так и о медицинских исследованиях.

Для цитоморфологобиофизических медицинских исследований, проводимых врачом цитологом, характерно наличие большого количества как цитоморфологической, так и цитобιοфизической информации, получаемой о пациенте, что усложняет процесс постановки верного диагноза. Автоматизация процесса постановки диагноза по цитоморфологобиофизическим показателям предполагает хранение непротиворечивой консолидированной информации, отражающей все этапы диагностики.

Следовательно, создание БД информационной системы цитоморфологобиофизической диагностики с целью хранения данных о пациенте и его цитомор-

фологобиофизических пробах является актуальным вопросом.

2. Постановка задачи

Существует большое количество моделей данных (на основе которых строятся БД), среди этих моделей выделяют основные: сетевые, реляционные, объектно-ориентированные, объектно-реляционные [2].

В сетевой модели любой объект может быть как главным, так и подчиненным (в сетевой модели главный объект обозначается термином «владелец набора», а подчиненный - термином «член набора»). Один и тот же объект может одновременно выступать и в роли владельца, и в роли члена набора. Это означает, что каждый объект может участвовать в любом числе взаимосвязей. В сетевой модели данных физическая независимость слабая, так как схема зависит от физического описания, и, следовательно, любое физическое изменение пути доступа в той или иной степени влияет на БД. Манипулирование данными в сетевой модели производится с помощью процедурных языков.

В реляционной модели объекты и взаимосвязи между ними представляются с помощью таблиц. Взаимосвязи также рассматриваются в качестве объектов. Каждая таблица представляет один объект и состоит из строк и столбцов. В реляционной базе данных каждая таблица должна иметь первичный ключ (ключевой элемент) - поле или комбинацию полей, которые единственным образом идентифицируют каждую строку в таблице. Благодаря своей простоте и естественности представления реляционная модель получила наибольшее распространение в системы управления базой данных (СУБД) для персональных компьютеров.

В объектно-ориентированной модели данные оформлены в виде моделей объектов, включающих прикладные программы, которые управляются внешними событиями. Результатом совмещения возможностей (особенностей) БД и возможностей объектно-ориентированных языков программирования являются объектно-ориентированные системы управления базами данных. Объектно-ориентированные БД обычно рекомендованы для тех случаев, когда требуется высокопроизводительная обработка данных, имеющих сложную структуру [2].

Разница между объектно-реляционными и объектными моделями: первые представляют собой надстройку над реляционной схемой, вторые же изначально объектно-ориентированы. Характерные свойства объектно-реляционной модели: комплексные данные, наследование типа, объектное поведение.

Комплексные данные могут быть реализованы через постоянно-храняемые объекты (persistent objects). Иерархия структурных комплексных данных предлагает дополнительное свойство наследование типа. То есть структурный тип может иметь подтипы, которые используют все его атрибуты и содержат дополнительные атрибуты, специфицированные в подтипе.

Объектное поведение закладывается через описание программных объектов. Такие объекты должны быть сохраняемыми и переносимыми в БД, поэтому они называются обычно как постоянные (или долговременные) объекты. Внутри базы данных все отношения с постоянным программным объектом есть отношения

с его объектным идентификатором, универсальным ядром сервера (Oracle, Informix, DB2, – OID).

При создании БД на основе объектно-реляционной модели универсальное ядро (OID) остается ориентированным на работу с реляционными данными, что отрицательно сказывается на производительности, вынуждая СУБД всякий раз производить сборку/разборку объектов при обмене с хранилищем. Еще один существенный недостаток объектно-реляционных БД заключается в том, что добавлений новых типов данных – это, по сути, расширение ядра сервера. Это модификация тщательно отлаженного, оптимизированного механизма, последствия такой операции трудно просчитываются.

Преимуществом реляционной модели перед другими моделями является простая и удобная для пользователя схема данных, представляемая в виде таблицы. Реляционная модель логически и физически независима. Физическая независимость реляционной модели состоит в том, что модель данных не включает никаких физических описаний. Логическая независимость допускает возможность применения одной концептуальной модели различными пользователями. Физическая независимость дает возможность в целях эффективности использования БД модифицировать физическую организацию данных и пути доступа. Например, необходимо добавить или удалить некоторую связь между записями без изменения программы. Для реляционных моделей бессмысленно использовать процедурный язык, поскольку обеспечена физическая независимость данных. С помощью команд процедурного языка программист строит стратегию доступа к данным. Но любое изменение пути доступа приводит к необходимости модификации программы. Реляционная модель позволяет улучшить выражение требований поддержания целостности и защиты данных путем использования языка высокого уровня.

Таким образом, разработка БД на основе реляционной модели, которая по сравнению с другими моделями обладает рядом достоинств, таких как логическая и физическая независимость модели, простая и удобная для пользователя схема данных в виде таблиц, актуальна.

3. Разработка базы данных информационной системы цитоморфологобиофизической диагностики

БД цитоморфологобиофизической диагностики характерна текстовыми и графическими данными, которые описывают данные о пациенте, его состоянии органов и систем организма. Данная БД предназначена для:

- хранения необходимого набора текстовых, числовых и графических данных;
- представления их в удобном, структурированном виде;
- корректировки хранимых данных;
- быстрого доступа к данным для их визуального отображения и проведения операций, связанных с их анализом и классификацией.

Целостность сущностей обеспечивается заданием первичных ключей.

На этапе концептуального проектирования были выделены такие типы сущностей: Primary_inf_patient (первичные данные о пациенте), Second_inf_patient

(второстепенные данные о пациенте, данные, которые заносятся не один раз и могут изменяться), Test (проба), Diagnostics (диагнозы и состояния органов пациента после проведения цитоморфологобио-физической диагностики), Setting (назначение), Citomorphological_parameters (цитоморфологобиофизические параметры пациента), Diuresis (состояние диуреза пациента), Harmful_habits (вредные привычки пациента) Medicinal_preparations (медицинские препараты), Doses_preparations (Дозы препаратов), Time_reception_preparations (время приема препаратов).

Сущность «Primary_inf_patient» предназначена для хранения первичных данных всех обследованных пациентов (например, таких как фотография пациента, дата рождения пациента и др.), которые заносятся один раз и более не изменяются.

Сущность «Second_inf_patient» предназначена для хранения второстепенных данных всех обследованных пациентов (например, таких как жалобы, дата обследования, диурез и др.), которые заносятся не один раз, и могут изменяться.

Сущность «Test» предназначена для хранения изображений проб биопсии эпителия пациентов в графической форме. Поддерживаются такие графические форматы, как JPEG, BMP, TIFF, GIF.

Сущность «Diagnostics» предназначена для хранения возможных диагнозов и состояний органов пациента при проведении цитоморфологобиофизической диагностики.

Сущность «Setting» предназначена для хранения информации о назначении лекарственных средств пациенту.

Сущность «Medicinal_preparations» предназначена для хранения названий лекарственных средств.

Сущность «Doses_preparations» предназначена для хранения информации о дозах лекарственных средств при назначении.

Сущность «Time_reception_preparations» предназначена для хранения информации о времени приема назначенных лекарственных средств.

Сущность «Citomorphological_parameters» предназначена для хранения информации о цитоморфологобиофизических параметрах биопсии эпителия пациента.

Сущность «Harmful_habits» предназначена для хранения информации о вредных привычках пациента на данный момент.

Сущность «Diuresis» предназначена для хранения информации о состоянии диуреза пациента на данный момент.

Между сущностями (таблицами) БД могут существовать взаимосвязи различных типов: один к одному, один ко многим, много к одному.

Сущность БД «Primary_inf_patient» имеет следующие атрибуты:

- id_primary_patient – номер первичной информации о пациенте по порядку, целочисленного типа;
- ФИО_пациент – фамилия, имя, отчество пациента, символьного типа;
- Фото_пациент – фотография пациента, графического типа;
- Домашний_адрес_телефон – домашний адрес и телефон пациента, символьного типа;
- Работа – место работы пациента, символьного типа;

– Генетическая предрасположенность – генетическая предрасположенность пациента к тому или иному заболеванию.

Сущность БД «Second_inf_patient» имеет следующие атрибуты:

- id_second_patient – номер вторичной информации о пациенте по порядку, целочисленного типа;
- Дата_обследования – дата обследования пациента, в формате даты (DD.MM.YYYY);
- Жалобы – жалобы, с которыми обратился пациент, символьного типа;
- Операции – операции, перенесенные пациентом, символьного типа;
- Травмы – перенесенные пациентом травмы, символьного типа.

Сущность БД «Test» имеет следующие атрибуты:

- id_test – номер исследования по порядку, целочисленного типа;
- Фото_пробы – фотография пробы биопсии эпителия пациента, графического типа;
- ЭОЯ_% – электроотрицательность ядер (ЭОЯ %) пациента, символьного типа.

Сущность БД «Diagnostics» имеет следующие атрибуты:

- id_diagnostics – номер расшифровки диагностики по порядку, целочисленного типа;
- Желудок_и_12_ти_перстная_кишка – состояние желудка пациента и состояние его 12-ти перстной кишки, символьного типа;
- Желчный_пузырь – состояние желчного пузыря пациента, символьного типа;
- Кишечник – состояние кишечника пациента, символьного типа;
- Молочные_железы – состояние молочных желез у пациентки, символьного типа;
- Иммунная_реакция – иммунная реакция организма пациента, символьного типа;
- Мочевыделительная_система – состояние мочевыделительной системы пациента, символьного типа;
- ОДС – состояние опорно-двигательной системы пациента;
- Паразитарно_инфекционный_фон – наличие той или иной инфекции и паразита у пациента, символьного типа;
- Печень – состояние печени пациента, символьного типа;
- Поджелудочная_железа – состояние поджелудочной железы пациента, символьного типа;
- Половые_органы – состояние половых органов пациента, таких как: яичники и трубы, тело матки, бартолининовая железа, внешние половые органы, предстательная железа, яичники, уретра, вагина, символьного типа;
- Сердечно_сосудистая_система – состояние сердечнососудистой системы пациента, символьного типа;
- ЦНС – состояние центральной нервной системы пациента, символьного типа;
- Щитовидная_железа – состояние щитовидной железы пациента, символьного типа;
- Дыхательная_система – состояние дыхательной системы, символьного типа;
- Другое – другие возможные состояния организма и органов пациента, символьного типа.

Сущность БД «Citomorphological_parameters» имеет следующие атрибуты:

- id_citomorphological_parameters – номер параметров пациента по порядку, целочисленного типа;
- Форма_клетки – содержит информацию о формах клеток букального эпителия пробы пациента, символического типа;
- Целостность_мембраны_клетки – содержит информацию о целостности мембран клеток букального эпителия пробы пациента, символического типа;
- Цвет_цитоплазмы – содержит информацию об окрасе цитоплазмы клеток букального эпителия пробы пациента, символического типа;
- Включения_в_цитоплазму – содержит информацию о наличии тех или иных включений в цитоплазму клеток букального эпителия пробы пациента, символического типа;
- Форма_ядер – содержит информацию о формах ядер клеток букального эпителия пробы пациента, символического типа;
- Целостность_мембраны_ядра – содержит информацию о целостности мембран ядер клеток букального эпителия пробы пациента, символического типа;

- Окрас_ядра – содержит информацию об окрасе ядер клеток букального эпителия пробы пациента, символического типа.

Сущность БД «Setting» имеет следующие атрибуты:

- id_setting – номер назначения по порядку, целочисленного типа;
- Назначен_е_лекар_е_препараты – назначение лекарственных средств пациенту, символического типа;
- Дозы_препаратов – дозировка назначенных препаратов пациенту, символического типа;
- Время_приема_препаратов – время приема назначенных препаратов пациенту, символического типа.

Сущность БД «Medicinal_preparations» имеет следующие атрибуты:

- id_medicinal – номер назначения по порядку, целочисленного типа;
- Лекарственные_препараты – оптимальный список лекарственных препаратов, символического типа;

Сущность БД «Doses_preparations» имеет следующие атрибуты:

- id_doses – номер назначения по порядку, целочисленного типа;
- Дозы_препаратов – возможные дозы принятия лекарственных средств, символического типа;

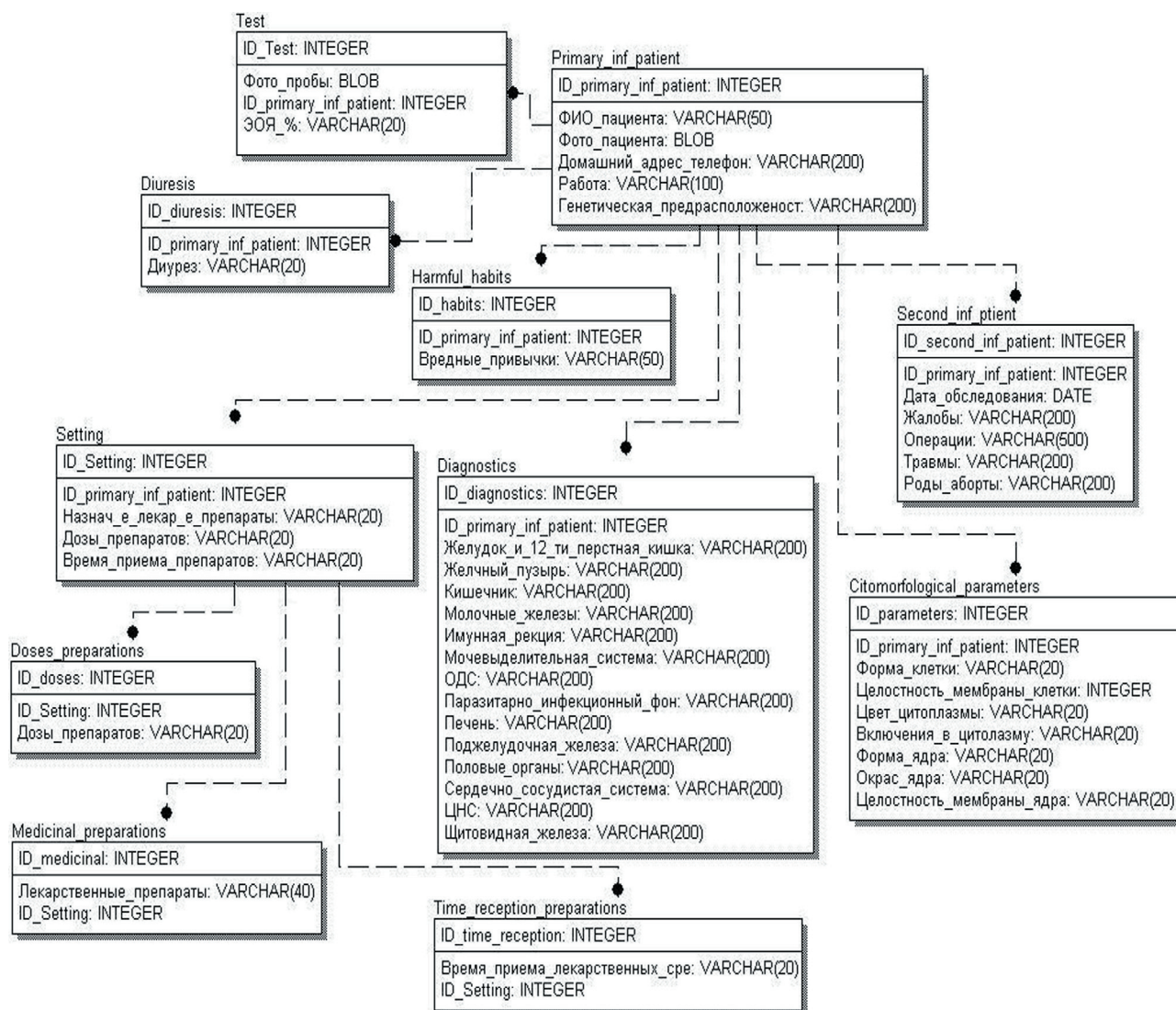


Рис. 1. Физическая модель базы данных

Сущність БД «Time_reception_preparations» має наступні атрибути:

- id_time_reception – номер призначення по порядку, цілочисленного типу;
- Время_приема_лекарственных_средств – можливі варіанти приєму лікарських засобів по часу, символічного типу.

Сущність БД «Harmful_habits» має наступні атрибути:

- id_harmful_habits – номер шкідливих звичок пацієнта по порядку, цілочисленного типу;
- Вредные привычки – шкідливі звички пацієнта на даний момент, символічного типу.

Сущність БД «Diuresis» має наступні атрибути:

- id_Diuresis – номер шкідливих звичок пацієнта по порядку, цілочисленного типу;
- Диурез – Стан діурезу у пацієнта на даний момент, символічного типу.

На етапі логічного проектування БД інформаційної системи цитоморфологіо-фізическої діагностики була розроблена фізическа модель схеми даних, заснована на реляційній моделі (рис. 1).

Виходячи з вимог до інформаційного забезпечення, вибрана фізическа модель, орієнтована на СУБД MS SQL Server 2008, так як вона показує найвищий швидкість роботи [2].

4. Выводы

БД інформаційної системи цитоморфологіо-фізическої діагностики розроблена на основі реляційної моделі, яка має просту і зручну для користувача схему даних у вигляді таблиць. Вона дозволяє зберігати необхідний набір текстових, числових і графічних даних про пацієнта, представити їх у зручному, структурованому вигляді, з можливістю корекції цих даних, а також має можливість швидкого доступу до даних для їх візуального відображення і проведення операцій, пов'язаних з їх аналізом і класифікацією.

Литература

1. Проблема формалізації текстових даних в універсальних медических інформаційних системах / М.Ю. Болгов, Д.А. Микитенко // Український журнал телемедицини і медическої телематики. – 2006. – Т.4, №2. – С.171-176.
2. Коннолли, Т. Базы данных. Проектирование, реализация и сопровождение. Теория и практика. Database Systems: A Practical Approach to Design, Implementation, and Management Third Edition 3-е изд. / Т. Коннолли, К. Берг – М.: "Вильямс" 2003. – 1436 с. – ISBN 0-201-70857-4.

УДК 001.891:65.011.56

АЛГОРИТМИ ПОШУКУ ПЛАГІАТУ

А.А. Чижова

Харківський національний університет радіоелектроніки
 просп. Леніна, 14, м. Харків, Україна, 61166
 Контактний тел.: 097-847-29-63
 E-mail: Chijova_alla@ukr.net

Дана загальна характеристика алгоритму. Досліджено основні алгоритми пошуку плагіату тексту. Описано основні принципи роботи алгоритмів

Ключові слова: алгоритм, реалізація, програма, пошук

Дана общая характеристика алгоритма. Исследованы основные алгоритмы поиска плагиата текста. Описаны основные принципы работы алгоритмов

Ключевые слова: алгоритм, реализация, программа, поиск

A general characteristic of the algorithm. The basic algorithms for finding plagiarism of text. The basic principles of the algorithms

Keywords: algorithm, implementation, program search

1. Введення

Плагіат - навмисне привласнення авторства чужого твору науки чи думок або мистецтва чи винаходи. Плагіат може бути порушенням авторсько-правового законодавства та патентного законодавства і в їх якості може спричинити за собою юридичну відповідаль-

ність. З іншого боку, плагіат можливий і в областях, на які не поширюється дія яких-небудь видів інтелектуальної власності, наприклад, в математиці та інших фундаментальних наукових дисциплінах.

Алгоритм - це система правил виконання обчислювального процесу, що обов'язково приводить до розв'язання певного класу задач після скінченного