

АНАЛИТИЧЕСКАЯ СИСТЕМА ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ УПРАВЛЕНИЯ ЗДРАВООХРАНЕНИЕМ НА МУНИЦИПАЛЬНОМ УРОВНЕ

А.В. Целиков

Московский государственный индустриальный университет

Описана компьютерная система учета и обработки медико-статистической информации муниципального уровня, в которой используются различные методы анализа данных, работы с обобщенными показателями, методы кластеризации, прогнозирования и искусственного интеллекта. Система позволяет осуществлять многомерный анализ показателей заболеваемости на различных уровнях управления, основываясь на данных государственной медицинской статистической отчетности.

ВВЕДЕНИЕ

В здравоохранении Российской Федерации применяется технология сбора показателей государственной медицинской статистики, разработанная еще в СССР. Она предполагает передачу заполненных бумажных шаблонов определенного образца, ежегодно утверждаемых Министерством здравоохранения и социального развития РФ, непосредственно от лечебно-профилактических учреждений (ЛПУ) руководящим органам здравоохранения. За достаточно долгий период применения этой технологии выработаны механизмы использования и обработки учетных данных, которые включают в себя отлаженные процедуры сбора, контроля и группировки исходной информации, а также формирование набора отчетной документации. Однако, как замечают специалисты, потенциал этой информации в настоящее время используется далеко не полностью. Причина этого состоит в недостаточном применении современных методов обработки и анализа данных. Именно из-за отсутствия автоматизированной аналитической обработки сведений медико-статистической отчетности на муниципальном уровне, их роль в информационном обеспечении процесса управления здравоохранением весьма занижена [1].

На российском рынке программных продуктов присутствует ряд систем сбора и обработки данных медико-статистической отчетности, однако методы анализа хранимой информации в этих системах практически не реализованы. Поэтому существенная часть трудоемких операций, таких как интеллектуальная обработка и анализ данных, построение тенденций и прогнозов выполняется вручную [1, 2].

Помимо указанных статистических форм, организации управления здравоохранением создают сборники отчетных данных, формирование которых достаточно трудоемко. Кроме того, результаты могут содержать некорректности из-за ошибок переноса и ручного набора данных. Однако именно эта информация является платформой для принятия управленческих решений в области здравоохранения. Таким образом, автоматизация процесса формирования необходимой отчетной информации непосредственно из базы данных компьютерной системы также представляет собой одну из важных задач обработки медицинских данных.

Для автоматизации основных процессов обработки медико-статистических данных автором настоящей статьи разработана «Система сбора и анализа медицинских данных» (далее — система СИАМЕД) — информационный продукт для статистической обработки первоначальных данных, а



также для последующего анализа информации с использованием собственной базы математических моделей и набора соответствующих методов и преобразований [3, 4].

Основные операции в работе системы СИАМЕД:

- автоматизированная загрузка информации в шаблоны форм государственной медицинской статистической отчетности;

- автоматизированный контроль данных всех форм;

- группировка сведений, составление сводных отчетов по различным территориальным образованиям (районам, округам и т. д.), работа с произвольными отчетными показателями;

- использование подсистемы гибкой аналитической отчетности, позволяющей проводить анализ произвольных отчетных показателей, получать динамику их изменения, а также осуществлять визуализацию расчетов;

- автоматизированный процесс формирования необходимой отчетной информации непосредственно из базы данных системы;

- поддержка принятия решений на основе построенной реляционной модели оперативной аналитической обработки данных (On-Line Analytical Processing — OLAP), которая функционирует на платформе специально спроектированного хранилища данных системы [3];

- интегральная оценка показателей заболеваемости, вычисление обобщенного показателя, формирование наборов весовых коэффициентов [4];

- применение методов кластеризации [5];

- прогнозирование уровня показателей с использованием данных предыдущих периодов.

Помимо этого, учтен ряд не реализованных ранее в аналогичных системах пожеланий непосредственных пользователей программы — сотрудников аналитико-статистических отделов муниципальных департаментов здравоохранения:

- гибкая процедура контроля данных, позволяющая обрабатывать только используемые формы в ЛПУ статистической отчетности, что уменьшает число ошибок и увеличивает скорость контроля;

- возможность выбора списка форм для каждого конкретного ЛПУ; эта операция более точно соответствует реальной ситуации, чем принятая в остальных системах такого рода процедура соответствия списка форм типам ЛПУ, так как сами учреждения внутри типов могут значительно различаться;

- вывод всей отчетной информации в формате MS Excel, что легко позволяет осуществлять дополнительную обработку данных средствами этого популярного табличного редактора.

Использование указанных приемов работы с данными позволит более четко оценивать уровень различных показателей заболеваемости, иметь возможность выявлять причины сложившейся ситуации, сравнивать с показателями других территорий и планировать мероприятия на будущие периоды.

1. КРАТКОЕ ОПИСАНИЕ ИСПОЛЬЗУЕМЫХ ПОКАЗАТЕЛЕЙ СИСТЕМЫ

Основная информация, с которой работает система — данные федерального государственного статистического наблюдения в области здравоохранения, которые предполагают заполнение около 30-ти специально разработанных и утвержденных Госкомстатом России годовых статистических медицинских отчетных форм.

Сбор таких сведений начинается непосредственно с ЛПУ различного уровня подчинения (муниципального, областного и т. п.). По каждому такому учреждению в систему поступает полная информация обо всех необходимых отчетных показателях, приписанных ЛПУ в зависимости от его специфики, медицинской направленности, а также уровня подчинения. Сведения представляют собой информацию о заболеваемости населения, деятельности лечебно-профилактических учреждений, продолжительности лечения, данные о смертности и рождаемости, отчеты о лечении в пансионатах и туберкулезных санаториях, информацию о медицинском наблюдении за спортсменами и т. п.

Таким образом, показателями системы СИАМЕД являются абсолютно все данные государственной медицинской статистической отчетности.

Помимо описанных выше данных в систему вводится информация о территории, населении, зоне обслуживания лечебных учреждений, благодаря которой исходные сведения, собираемые в абсолютных величинах, можно легко представить в величинах относительно зарегистрированного населения, что в свою очередь позволяет проводить сравнительный анализ показателей по различным регионам.

Кратко опишем основные операции, осуществляемые в системе.

2. ПЕРВОНАЧАЛЬНАЯ ОБРАБОТКА ДАННЫХ

Данные государственной медицинской статистической отчетности, а также информация о населении и ЛПУ загружаются в систему автоматически в стандартном формате данных медицинской отчетности «медстат» [2]. При необходимости всю информацию можно ввести вручную или скоррек-

тировать результаты загрузки. После этого все данные проходят обязательную процедуру табличного, межтабличного и межформенного контроля согласно установленным правилам контроля форм статистической отчетности. Такая процедура исправляет большинство ошибок ввода данных, устраняет возможную противоречивость, повышает их надежность.

Первоначальная обработка данных в системе завершается процедурой агрегирования до необходимого отчетного уровня (муниципальный, областной и т. п.) путем выполнения стандартной процедуры свода данных: вычисления относительных и группирующих показателей. После этого этапа система содержит проверенные сведения по всем отчетным направлениям для необходимого отчетного уровня, готовые для обработки и анализа. Инфологическая модель основных процессов системы представлена на рис. 1.

Основная цель применения OLAP-технологий состоит в анализе информации и представлении его результатов в виде, удобном для принятия управленческих решений. Главная идея заключается в построении многомерных кубов, которые будут оперативно доступны для пользовательских запросов [5].

На рис. 2. представлен пример куба, спроектированного системой СИАМЕД. Ячейки куба содержат агрегированные данные по медицинским округам, соответствующие находящимся на осях

Наименование	Всего заболеваний (на 1000 чел. населения)		
	Взрослые	Подростки	Дети
Медицинский округ № 1	908,0	1523,2	1997,8
Медицинский округ № 2	975,6	1483,0	2204,7
Медицинский округ № 3	833,7	1356,4	2204,6
Медицинский округ № 4	996,3	1324,7	2022,8
Медицинский округ № 5	825,4	1342,1	2100,1
Медицинский округ № 6	895,1	1366,7	1764,5
Медицинский округ № 7	1004,6	1197,8	1720,9
Медицинский округ № 8	854,0	988,0	1465,8
Медицинский округ № 9	1122,0	1976,3	3214,8
Медицинский округ № 10	1199,6	1480,6	1973,0
Медицинский округ № 11	949,0	1355,3	2069,1
Медицинский округ № 12	1038,0	1120,2	2121,2

Рис. 2. Пример OLAP-куба

куба значениям параметров SQL-запроса в предложении WHERE. Таким образом, каждая ячейка данного куба отвечает на вопрос: «Какова суммарная заболеваемость на определенном уровне подчинения в установленном году».

Основное преимущество технологии многомерных кубов в том, что данные необходимо запросить из реляционной СУБД всего один раз — при построении куба. Поскольку аналитики, как правило, работают с довольно редко изменяемой информацией, каковой и являются данные годовой статистической отчетности, то сформированный куб остается актуальным в течение достаточно продолжительного времени. Это позволяет увеличить скорость выполнения ряда аналитических операций, формирования отчетов, проведения сравнительного анализа путем использования заранее сгруппированной информации.

4. ВЫЧИСЛЕНИЕ ОБОБЩЕННЫХ ПОКАЗАТЕЛЕЙ

Система государственной медицинской статистической отчетности содержит большое число показателей, поэтому решение задачи снижения размерности пространства характеристик актуально как для визуализации данных, возможности сравнительного анализа, так и для управления ситуацией в целом. Для этих целей в системе применяется метод построения сводного показателя, в основе которого лежит решение оптимизационной задачи:

$$\|XD(q) - \mu p e^T\|^2 \rightarrow \min_{\mu, q, p}$$

и системы уравнений [4]:

$$\begin{aligned} Y Y^T p &= \mu p, \\ Y^T Y q &= \mu q, \end{aligned}$$

где X — матрица исходных данных, строки которой соответствуют анализируемым объектам, а столб-

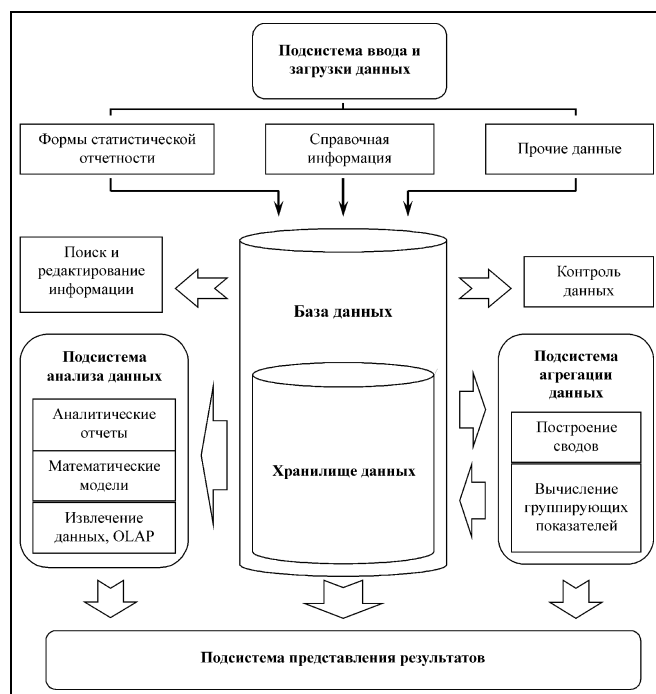


Рис. 1. Инфологическая модель системы

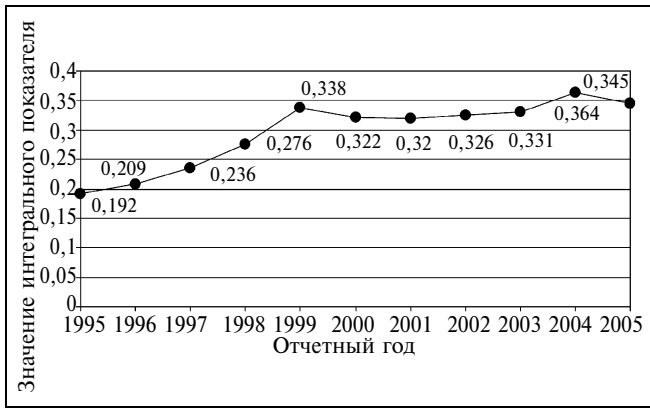


Рис. 3. Изменение интегрального показателя по отчетным годам

цы — их характеристикам; $D(q)$ — диагональная матрица с вектором p на главной диагонали; векторы p и q совпадают с главными собственными векторами матриц YU^T и $Y^T Y$; матрица Y получена из матрицы X путем нормировки столбцов.

Компоненты вектора p интерпретируются как обобщенные характеристики (сводные показатели) для строк исходной матрицы X , а компоненты вектора q — как весовые коэффициенты столбцов матрицы X .

В качестве примера на рис. 3 представлен график изменения интегрального показателя общей заболеваемости населения, рассчитанного с помощью описанной методики [4].

Таким образом, интегральный показатель является общей характеристикой, отражающей совокупный уровень заболеваемости для отчетного года, региона или лечебного учреждения. Эта характеристика может быть использована экспертами как для общей оценки числа зарегистрированных заболеваний, так и для сравнительного анализа.

5. ФОРМИРОВАНИЕ ТЕНДЕНЦИЙ И ПРОГНОЗОВ

Определение тенденций развития уровня показателей заболеваемости в будущие периоды представляют собой одну из наиболее важных и интересных задач в управлении здравоохранением. Анализ такого рода служит практическим обоснованием для выработки мер противодействия и профилактики заболеваний. Для этих целей в системе задействован ряд моделей прогнозирования.

Временные ряды. Применяется метод Хольта—Уинтерса, который помимо линейного тренда поддерживает также и сезонные компоненты, выраженные мультипликативно. Прогноз, сделан-

ный в момент t на l тактов вперед, вычисляется по формуле:

$$x(t, l) = [a_0(t) + la_1(t)]\omega(t + l - T),$$

где $\omega(t)$ — коэффициент сезонности, коэффициенты $a_0(t)$ и $a_1(t)$ определяются, исходя из данных предыдущих периодов, T — число временных тактов, содержащихся в полном сезонном цикле [6].

Линейно-гармонические тренды. Для формирования прогноза применяется взвешенная аппроксимация базисом линейно-гармонических функций по методу наименьших квадратов. Коэффициенты аппроксимации вычисляются по набору точек с возможностью корректировки весов.

Искусственные нейронные сети. В системе реализована сеть обратного распространения, обучение которой осуществляется по формуле

$$\Delta\omega_{ij}^{(n)} = -\eta\delta_j^{(n)} x_i^n,$$

с корректировкой весов

$$\omega_{ij}(t) = \eta\delta_j o_i + a\Delta\omega_{ij}(t - 1),$$

где $\Delta\omega_{ij}^{(n)}$ — изменение весов n -го слоя нейронной сети, $\omega_{ij}(t)$ — значения весов нейронной сети на шаге t , η — параметр, определяющий скорость обучения, x_i — значение i -го входа нейрона, δ — значение ошибки для нейрона j в выходном слое n , a — коэффициент инерции, o_i — выходное значение i -го элемента [7].

Все реализованные методы прогнозирования обеспечены наборами настраиваемых параметров, что позволяет применять их для широкого круга исследуемых показателей.

ЗАКЛЮЧЕНИЕ

Разработанная информационно-аналитическая система позволяет интегрировать данные государственной системы медицинской статистической отчетности и современные методы обработки и анализа данных. Применяемые модели позволяют наглядно представлять текущий уровень заболеваемости, оценивать тенденции изменения показателей, проводить информационно-сравнительный анализ, и в целом повысить значение медико-статистической информации при принятии управленческих решений в области здравоохранения. Все необходимые для этого методы реализованы непосредственно в самой системе, что упрощает анализ на любом уровне управления. Программа снабжена рядом настроек, позволяющих автоматизировать выполнение различных опера-

ций. Система реализована для персональных компьютеров в среде Windows и позволяет сетевое многопользовательское применение. В качестве платформы используется бесплатная система управления базой данных Firebird, что существенно для бюджетных организаций.

В настоящее время система проходит опытную эксплуатацию в ряде учреждений управления муниципальным здравоохранением.

ЛИТЕРАТУРА

1. Тюков Ю.А., Ползик Е.В. Официальная медицинская статистика как основа управления здоровьем населения: возможности и ограничения // Экономика здравоохранения. Медицинская статистика. — 2000. — № 7.
2. Целиков А.В., Ткаченко Л.А. Автоматизированная обработка медико-статистической отчетности в анализе заболеваемости населения и влияния состояния окружающей среды //

Техника, технологии и перспективные материалы: Сб. науч. тр. — М.: 2005. — С. 300—304.

3. Целиков А.В. Информационно-аналитическая система сбора и анализа медико-статистических данных // Труды XXXIII междунар. конф. «Гагаринские чтения» / МАТИ — М., 2007. — Т. 6. — С. 263—264.
4. Куренков Н.И. Особенности анализа многомерных данных. — М.: ЗАО «Научно-исследовательский центр распознавания образов», 2001. — 14 с.
5. Методы и модели анализа данных: OLAP и Data Mining / А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод — СПб.: БХВ-Петербург, 2004. — 336 с.
6. Айвазян С.А. Прикладная статистика. Основы эконометрики. Т. 2. — М.: Юнити, 2001. — 432 с.
7. Головкин В.А. Нейронные сети: обучение, организация и применение. — М.: ИПРЖР, 2002. — 256 с.

e-mail: atselikov@gmail.com

Статья представлена к публикации членом редколлегии
В.Н. Новосельцевым. □

III Всероссийская молодежная научная конференция по проблемам управления (МКПУ-2008)

Москва, Институт проблем управления им. В.А. Трапезникова РАН,
7—9 апреля 2008 г.

III Всероссийская молодежная научная конференция по проблемам управления продолжает традиции проведения молодежных научных конференций по современным проблемам теории и практики управления (в 2006 году ИПУ РАН совместно с рядом ведущих ВУЗов и научных центров провел первую такую конференцию на базе СГАУ (г. Самара), в 2007 г. — вторую конференцию на базе ВГАСУ (г. Воронеж)).

Приглашаются к участию молодые ученые (студенты и аспиранты, кандидаты наук в возрасте до 35 лет, доктора наук — в возрасте до 40 лет).

Основные цели МКПУ-2008:

— предоставить возможность молодым сотрудникам и аспирантам ИПУ РАН и других научных центров и вузов выступить с докладами по современным проблемам теории и практики управления и обсудить свои результаты с коллегами и ведущими учеными;

— выявить наиболее интересные доклады и рекомендовать их авторов для участия в IV конференции молодых ученых, которая состоится в июне 2008 г. на базе Казанского государственного технического университета, и V Всероссийской школе-семинаре молодых ученых "Управление большими системами", которая состоится в октябре 2008 г. на базе Липецкого государственного технического университета;

— отобрать наиболее интересные и целостные с точки зрения научных результатов доклады для публикации результатов в ведущих научных журналах. По итогам конференции будет подготовлен специальный выпуск сборника "Управление большими системами". Отобранные доклады будут рекомендованы к печати в ведущих журналах по теории управления: "Автоматика и телемеханика", «Известия РАН. Серия "Теория и системы управления"», "Проблемы управления", "Системы управления и информационные технологии" и др.;

— рекомендовать наиболее интересные доклады для участия в конкурсе на лучшую работу молодых ученых по теории управления и ее приложениям (итоги очередного конкурса будут подведены в июне 2008 г.).

Основные направления работы конференции:

- теория систем управления;
- управление в социальных и экономических системах;
- технические средства автоматизации и вычислительной техники;
- системы управления технологическими процессами;
- разработка программного обеспечения систем управления;
- автоматизированные системы организационного управления и обработки данных;
- управление подвижными объектами и навигация.

Оргкомитет готов рассмотреть предложения по организации секций в рамках проблематики конференции, а также обращается к Вам с просьбой передать информацию о конференции Вашим коллегам. Чтобы получить более подробную информацию, задать вопросы или подать заявку на участие в конференции, посетите сайт <http://www.ipu-conf.ru/kmu>.