

## ОБРОБКА ІНФОРМАЦІЇ В СКЛАДНИХ ТЕХНІЧНИХ СИСТЕМАХ

УДК 621.39

### ОБОСНОВАНИЕ ПОКАЗАТЕЛЕЙ КАЧЕСТВА ИНФОРМАЦИИ БАЗЫ ДАННЫХ АВТОМАТИЗИРОВАННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ

С.Н. Звиглянич

(Объединенный научно-исследовательский институт Вооруженных Сил, Харьков)

*Обоснован выбор показателей качества информации базы данных автоматизированной информационной системы (АИС). Качество информации базы данных предлагается оценивать на основе времени ее обновления. Данные показатели отражают суть функционирования базы данных, являются критичными и чувствительными по отношению к качеству информации, циркулирующей в автоматизированной информационной системе.*

**качество информации, автоматизированная информационная система**

**Постановка проблемы в общем виде.** Принятие решений при управлении сложными системами предопределяет необходимость целенаправленной переработки значительных объемов информации. В этом аспекте системообразующим элементом в современных АСУ является автоматизированные информационные системы (АИС). На АИС возлагаются функции сбора, хранения, поиска и обработки информации. Требования интеграции информации, циркулирующей в АИС, приводят к необходимости создания баз данных (БД). Решение ряда информационно-расчетных задач, принятие решений с использованием хранящейся в БД информации обуславливают требования учета качества данной информации. Таким образом, возникает необходимость введения показателей, отражающих качество используемой информации БД АИС.

**Анализ последних исследований и публикаций.** Проблема оценки качества функционирования АИС приобретает все большую актуальность в связи с повышением уровня автоматизации процесса принятия решений. Это связано с увеличением объемов обрабатываемой инфор-

мации и сокращением времени на принятие решения [1]. Задачи анализа и синтеза средств обработки информации, рассмотренные в [2], в основном освещают проблему эффективного использования каналов связи, применения оптимизационных алгоритмов маршрутизации. В [3] рассматриваются подходы к построению БД, дается общая оценка различных моделей представления данных. В [4] используются довольно известные подходы к оцениванию степени риска при принятии решений. Данные подходы ориентируются, прежде всего, на безусловное знание вероятностей состояний обстановки. При этом информационная составляющая процесса принятия решения не выделяется и не рассматривается.

Можно отметить, что на сегодняшний день остается актуальным вопрос проведения количественной оценки качества информации, циркулирующей в АИС. При этом необходимо использовать легко вычисляемые, доступные показатели, отражающих цель функционирования данных систем.

**Постановка задачи.** Современные системы управления базами данных (СУБД) ориентированы на реляционную модель данных, которая предусматривает построение БД в виде таблиц. Каждому реальному объекту соответствует объект (запись) БД как некая его информационная модель. Соответственно, каждому параметру реального объекта поставлено в соответствие поле записи БД. Каждый параметр объекта характеризуется вполне определенным временем – периодом обновления информации. Отметим, что использование значения параметра, у которого период обновления превышен, крайне не желательно. Устаревшая информация является источником ошибок различных вычислений, а также увеличивает степень риска при принятии решений. Исходя из выше сказанного, с учетом периодов обновления записей БД, поставим цель – обосновать показатели, характеризующие качество информации БД АИС.

**Обобщенное теоретическое решение задачи.** Введем в рассмотрение матрицу

$$T = (t_{ij}),$$

где  $t_{ij}$  – период обновления  $j$ -го поля  $i$ -й записи.

Определим матрицу

$$V_{mn} = (v_{ij}),$$

где  $v_{ij}$  – действительное время обновления  $j$ -го поля  $i$ -й записи.

Для текущего времени  $t_t$  вычислим

$$\Delta t_{ij} = t_t - v_{ij}.$$

Положим процесс обновления БД стационарным. По своей природе

величины  $\Delta t_{ij}$  – случайные. В дальнейших рассуждениях величины  $\Delta t_{ij}$  заменим на их математические ожидания. Точнее, на практике будем использовать оценки математических ожиданий  $\Delta t_{ij}$ , полученные на основе статистики, собранной в процессе эксплуатации БД. Каждому полю в записях поставим в соответствие характеристическую переменную

$$x_{ij} = \begin{cases} 1 & - \Delta t_{ij} < t_{ij}; \\ 0 & - \text{в противном случае.} \end{cases}$$

С учетом введенной характеристической переменной каждое поле записи может находиться в двух состояниях (рис. 1).

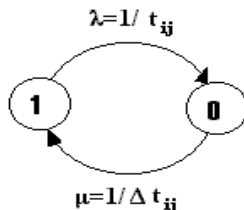


Рис. 1. Граф состояний поля записи

Тогда, следуя нашим рассуждениям, объект, представленный записью в БД и имеющий  $n$  полей, может находиться в  $2^n$  состояниях. Для определенности пусть объект имеет два поля. На рис. 2 представлен его граф состояний, где  $\lambda_{11} = 1/t_{11}$ ;  $\lambda_{12} = 1/t_{12}$ ,  $\mu_{11} = 1/\Delta t_{11}$ ;  $\mu_{12} = 1/\Delta t_{12}$ .

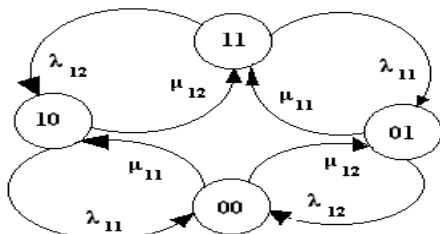


Рис. 2. Граф состояний объекта с двумя полями

Запишем для этого случая уравнения Колмогорова [5] (для простоты записи будем считать, что состоянию «11» соответствует вероятность  $P_1$ ; состоянию «01» –  $P_2$ ; состоянию «10» –  $P_3$ ; состоянию «00» –  $P_4$ ):

$$\begin{aligned} \mu_{12}P_3 + \mu_{11}P_2 - (\lambda_{11} + \lambda_{12})P_1 &= 0; & \lambda_{11}P_1 + \mu_{12}P_4 - (\lambda_{12} + \mu_{11})P_2 &= 0; \\ \lambda_{12}P_1 + \mu_{11}P_4 - (\lambda_{11} + \mu_{12})P_3 &= 0; & \lambda_{11}P_3 + \lambda_{12}P_2 - (\mu_{11} + \mu_{12})P_4 &= 0. \end{aligned} \quad (1)$$

Так как в конкретный момент времени объект может находиться только в одном состоянии, то запишем нормировочное условие как

$$P_1 + P_2 + P_3 + P_4 = 1. \quad (2)$$

Решая систему уравнений (1) с учетом условия (2), найдем  $P_1, P_2, P_3, P_4$ .

Неопределенность состояния  $i$ -го объекта БД оценим через его энтропию [6]

$$H_i = \sum_{j=1}^n P_j \log \frac{1}{P_j}. \quad (3)$$

Представим  $H_i$  как показатель качества информации об объекте в БД, который отражает ее полноту относительно реального объекта.

Учитывая аддитивность энтропии, общая неопределенность БД, отражающая качество хранящейся в ней информации, равна

$$H_{\text{БД}} = \sum_{i=1}^m H_i. \quad (4)$$

На основе (3) введем в рассмотрение вектор достоверности информации БД

$$D = (H_1, H_2, \dots, H_m). \quad (5)$$

Каждой информационно-расчетной задаче или принятию решения с использованием информации БД можно поставить в соответствие частный вектор достоверности информации  $d$ , элементы которого есть энтропия используемых объектов. При малых значениях энтропии по значению вероятностей нахождения объекта в крайних состояниях можно с достаточной степенью уверенности судить о практической целесообразности использования данной информации. Во всех же других случаях использование информации должно проводиться с учетом обеспечения минимального значения суммы элементов соответствующего частного вектора достоверности.

**Выводы.** Введенные показатели качества информации БД АИС (3) – (5) позволяют количественно оценить состояние БД, определить пути повышения достоверности циркулирующей в системе информации, а также оценить степени риска при проведении расчетов, принятии решений с использованием АИС.

## ЛИТЕРАТУРА

1. Ларичев О.И. Теория и методы принятия решений. – М.: Логос, 2002. – 392 с.
2. Королёв А.В., Кучук Г.А., Пашнев А.А. Адаптивная маршрутизация в корпоративных сетях. – Х.: ХВУ, 2003. – 224 с.
3. Конолли Т., Бегг К., Страчан А. Базы данных: проектирование, реализация и сопровождение. Теория и практика. – М.: Вильямс, 2000. – 1120 с.
4. Системный анализ в управлении: Учеб. пособие /В.С. Анфилатов, А.А. Емельянов, А.А. Кукушкин / Под ред. А.А. Емельянова. – М.: Финансы и статистика, 2002. – 368 с.
5. Вентцель Е.С., Овчаров Л.А. Теория случайных процессов и ее инженерные приложения. – М.: Наука, 1991. – 384 с.
6. Ирвин Дж., Харль Д. Передача данных в сетях: инженерный подход: Пер.с англ. – С.-Пб.: БХВ-Петербург, 2003. – 448 с.

Поступила 31.01.2005

**Рецензент:** доктор военных наук, профессор И.О. Кириченко,

