

УДК 621.39

С.Н. Звиглянич¹, В.Л. Петров²¹ Харьковский университет Воздушных Сил им. И. Кожедуба, Харьков² ЦККБ «Протон», Харьков

ПОКАЗАТЕЛИ КАЧЕСТВА КОНСОЛИДИРОВАННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ

Вводится ряд показателей, характеризующих качество консолидированной информационной системы. Основное внимание уделяется качеству циркулирующей информации, которое оценивается через время обновления данных. Введенные показатели отражают суть функционирования консолидированной информационной системы, являются критичными и чувствительными по отношению к качеству используемой информации при решении практических задач.

Ключевые слова: информация, показатель, консолидированная информационная система.

Введение

Постановка проблемы в общем виде. Принятие решений при управлении сложными системами предопределяет необходимость целенаправленной переработки значительных объемов информации. Требования интеграции информации, циркулирующей в системах поддержки принятия решений, приводят к необходимости создания консолидированных информационных систем. На эти системы возлагаются функции сбора, хранения, поиска и обработки информации. Решение ряда информационно-расчетных задач, принятие решений с использованием хранящейся в таких системах информации обуславливают требования учета качества данной информации. Таким образом, возникает необходимость введения показателей, отражающих качество используемой информации консолидированных информационных системах.

Анализ последних исследований и публикаций. Проблема оценки качества функционирования различного рода информационных систем приобретает все большую актуальность в связи с повышением уровня автоматизации процесса принятия решений. Это связано с увеличением объемов обрабатываемой информации 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 баз данных (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 информационной системе. При этом необходимо использовать легко вычисляемые, доступные показатели, отражающих цель функционирования данных систем.

Постановка задачи. Современные консолидированные информационные системы ориентированы на обработку информации, поступающей от разных источников, и создание специальным образом обработанных баз знаний, адаптированных под информационные потребности определенного круга пользователей. Информация, поступающая от каждого источника, характеризуется вполне определенным временем – периодом обновления информации. Отметим, что использование информации от источника, у которой 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 информации, поступающей от различных источников, поставим цель – обосновать показатели, характеризующие качество информации базы знаний консолидированной информационной системы.

Обобщенное теоретическое решение задачи

Консолидированную информационную систему представим как инфокоммуникационную сеть, где каждый узел (пункт) сети принимает, обрабатывает и хранит информацию от различных источников.

Дальше консолидированная информационная система – сеть.

При этом форма представления информации, как правило, разная.

Пусть каждый i -й источник информации (ИИи) имеет период обновления информации T_i .

Определим

$$\Delta t_i = t_i - t_i^0, \quad (1)$$

где t_i – текущее время, t_i^0 – время последнего обновления.

Положим процесс обновления информации стационарным. По своей природе Δt_i величины случайные. Каждому ИИН данного узла сети поставим в соответствие характеристическую переменную

$$x_i = \begin{cases} 1, & \text{если } \Delta t_i < T_i; \\ 0, & \text{в противном случае.} \end{cases} \quad (2)$$

С учетом введенной характеристической переменной каждый ИИН может находиться в двух состояниях (рис. 1).

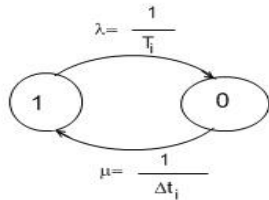


Рис. 1. Граф состояния источника

Тогда, следуя нашим рассуждениям, узел сети, имеющий n ИИН, может находиться в 2^n состояниях. Для определенности пусть узел имеет два ИИН.

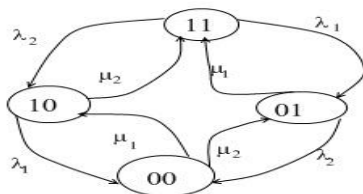


Рис. 2. Граф состояний узла

На рис. 2 представлен его граф состояний, где $\lambda_1 = 1/T_1, \lambda_2 = 1/T_2, \mu_1 = 1/\Delta t_1, \mu_2 = 1/\Delta t_2$.

Запишем для этого случая уравнения Колмогорова [5]. Для простоты записи будем считать:

- состоянию «11» соответствует вероятность P_1 ;
- состоянию «01» – P_2 ;
- состоянию «10» – P_3 ;
- состоянию «00» – P_4 .

$$\begin{aligned} \mu_2 P_3 + \mu_1 P_2 - (\lambda_1 + \lambda_2) P_1 &= 0, \\ \lambda_1 P_1 + \mu_2 P_4 - (\mu_1 + \lambda_2) P_2 &= 0, \\ \lambda_2 P_1 + \mu_1 P_4 - (\mu_2 + \lambda_1) P_3 &= 0, \\ \lambda_1 P_3 + \lambda_2 P_2 - (\mu_1 + \mu_2) P_4 &= 0. \end{aligned} \quad (3)$$

Так как в конкретный момент времени узел сети может находиться только в одном состоянии – запишем нормировочное условие

$$P_1 + P_2 + P_3 + P_4 = 1. \quad (4)$$

Решая систему уравнений (3) с учетом условия (4), найдем P_1, P_2, P_3, P_4 .

Неопределенность состояния рассматриваемого узла сети оценим через его энтропию

$$H^u = \sum_{i=1}^4 P_i \log \frac{1}{P_i}. \quad (5)$$

Неопределенность состояния узла сети отражает состояние информации, получаемой от соответствующих ИИН. Можно считать, что качество ин-

формации напрямую зависит от своевременного ее обновления.

Представим сеть в виде графа (рис. 3). Каждая вершина характеризуется, согласно (5), своей энтропией.

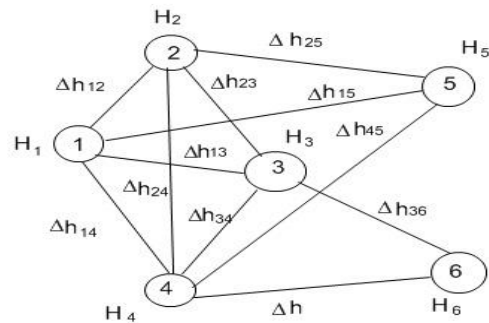


Рис. 3. Граф сети

Найдем суммарную энтропию указанной сети:

$$H_{\text{net}} = \sum_{i=1}^6 H_i^u. \quad (6)$$

Определим нагрузку ребер графа как разности энтропий соответствующих узлов:

$$\begin{aligned} \Delta h_{12} &= |H_1 - H_2|, \\ \Delta h_{13} &= |H_1 - H_3|, \\ \Delta h_{14} &= |H_1 - H_4| \text{ и т.д.} \end{aligned} \quad (7)$$

Используя выражения (7), построим минимальное остовное дерево для графа на рис. 3 [6].

Пусть данное остовное дерево имеет вид, представленный на рис. 4.

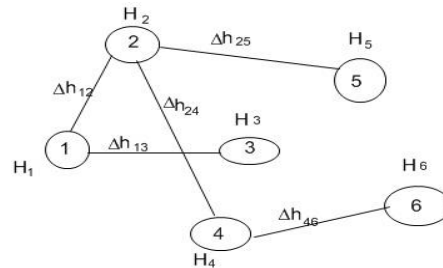


Рис. 4. Минимальное остовное дерево

Просуммируем нагрузку ребер минимального остовного дерева рассматриваемой сети:

$$\Delta h_s = \Delta h_{12} + \Delta h_{13} + \Delta h_{24} + \Delta h_{25} + \Delta h_{46}. \quad (8)$$

С использованием (6) и (8) введем обобщенный показатель, характеризующий опосредовано качество информации, циркулирующей в рассматриваемой сети:

$$Q_{\text{net}} = \frac{\Delta h_s}{H_{\text{net}}}. \quad (9)$$

Данный безразмерный показатель лежит в пределах 0 – 1. Чем больше его значение, тем менее, с точки зрения достоверности циркулирующей информации, сбалансирована рассматриваемая сеть.

Сбалансированность определяется своевременностью обновления информации на всех узлах сети.

Рассмотрим один из узлов сети.

Пусть данный узел имеет N ИИН. Для каждого ИИН интервалы обновления информации случайны.

Сделаем допущение – интервалы обновления образуют простейший поток с интенсивностью λ . Время решения некоторой задачи с использованием информации от данного ИИН строго фиксировано, т.е. интервалы T_p (рис. 5) постоянные.

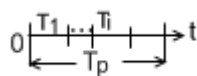


Рис. 5. Время решения фиксировано

С учетом того, что поток T_i простейший, вероятность хотя бы одного обновления на интервале T_p :

$$P_{\{t^o\}} = 1 - e^{-\lambda t_p}. \quad (10)$$

Математическое ожидание случайного интервала обновления T_i

$$M[T_i] = 1/\lambda. \quad (11)$$

Введем характеристическую переменную

$$x = \begin{cases} 1, & \text{если } M[T] < T_p; \\ 0, & \text{в противном случае.} \end{cases} \quad (12)$$

Математическое ожидание данной случайной величины

$$M[X] = P_{\{t^o\}}. \quad (13)$$

В целом, для рассматриваемого узла сети введем обобщенный показатель как приведенный уровень обновления информации от ИИН –

$$O^u = \frac{\sum_{i=1}^N M[x_i]}{N}, \quad (14)$$

где i – номер ИИН узла, N – число ИИН узла.

O^u лежит в пределах $0 - 1$, где единице соответствует полная согласованность периода обновления, т.е. к моменту использования – информация данного узла от всех ИИН обновлена.

Введем в рассмотрение индекс рассогласования для двух узлов сети как

$$\Delta z_{ij} = |O_i^u - O_j^u|. \quad (15)$$

Составим матрицу рассогласования для сети.

$$z = \begin{pmatrix} 0 & \Delta z_{12} & \dots & \Delta z_{1m} \\ \Delta z_{2m} & 0 & \dots & \Delta z_{2m} \\ \dots & \dots & \dots & \dots \\ \Delta z_{m1} & \Delta z_{m2} & \dots & 0 \end{pmatrix}, \quad (16)$$

где m – количество узлов сети.

Запрос на использование информации из сети представим в виде вектора

$$R^T = \{r_1, r_2, \dots, r_i, \dots, r_m\}, \quad (17)$$

где

$$r_i = \begin{cases} 1, & \text{если используется информация } i\text{-го узла} \\ 0, & \text{в противном случае.} \end{cases}$$

Рассмотрим произведение $Z \times R = K$. Полученный вектор K характеризует уровень рассогласования по времени обновления информации для запроса, представленного вектором R^T .

Для практического использования уровень рассогласования можно оценить модулем вектора K :

$$|K| = \sqrt{k_1^2 + k_2^2 + \dots + k_m^2}. \quad (18)$$

Модуль вектора K позволяет количественно сравнить между собой запросы к базе данных консолидированной информационной системы с учетом достоверности используемой информации.

Выводы

Введенные показатели качества консолидированной информационной системы позволяют количественно оценить состояние базы данных, определить пути повышения достоверности циркулирующей в системе информации. Также использование предложенных показателей дает возможность опосредованно оценить степень риска при использовании информации из базы данных при проведении расчетов, принятии решений.

Список литературы

1. Ларичев О.И. Теория и методы принятия решений / О.И. Ларичев. – М.: Логос, 2002. – 392 с.
2. Королёв А.В. Адаптивная маршрутизация в корпоративных сетях / А.В. Королёв, Г.А. Кучук, А.А. Пашнев. – Х.: ХВУ, 2003. – 224 с.
3. Конолли Т. Базы данных: проектирование, реализация и сопровождение. Теория и практика / Т. Конолли, К. Бегг, А. Страчан. – 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. Теория случайных процессов и ее инженерные приложения / Е.С. Вентцель, Л.А. Овчаров. – М.: Наука, 1991. – 384 с.
6. Ирвин Дж. Передача данных в сетях: инженерный подход: Пер.с англ. / Дж. Ирвин, Д. Харль. – СПб.: БХВ-Петербург, 2003. – 448 с.

Поступила в редколлегию 8.08.2012

Рецензент: д-р техн. наук, проф. В.А. Краснобаев, Полтавский национальный технический университет им. Ю. Кондратюка, Полтава.

ПОКАЗНИКИ ЯКОСТІ КОНСОЛІДОВАНОЇ ІНФОРМАЦІЙНОЇ СИСТЕМИ

С.М. Звиглянич, В.Л. Петров

Вводиться ряд показників, що характеризують якість консолідованої інформаційної системи. Основна увага приділяється якості циркулюючої інформації, яка оцінюється через час оновлення даних. Введені показники відбивають суть функціонування консолідованої інформаційної системи, є критичними і чутливими по відношенню до якості виконаної інформації при рішенні практичних завдань.

Ключові слова: інформація, показник, консолідована інформаційна система.

INDEXES OF QUALITY OF THE CONSOLIDATED INFORMATIVE SYSTEM

S.N. Zviglyanich, V.L. Petrov

The article considers some parameters that characterize quality of consolidated system. They are critical and essential for practical cases of consolidated system usage. The object of research is a quality of information and its assessment by the data refreshing time parameters.

Keywords: *information, index, consolidated informative system.*