

ХАРАКТЕРИСТИКИ ТА КІЛЬКІСНІ ОЦІНКИ СТРУКТУРИЗОВАНОЇ ІНФОРМАЦІЇ

Возна Н.Я.

Тернопільський національний економічний університет, к.т.н., доцент

I. Постановка проблеми

Теорія, методологія та інформаційна технологія структуризації поліфункціональних даних є фундаментальною основою синтезу та аналізу складних комп'ютерних та комп'ютеризованих систем. Важливим аспектом цього напрямку досліджень є чітка формалізація та інтегрована оцінка структурної складності поліфункціональних даних, пов'язана з теоретичними засадами моделювання та інформаційної побудови широкого класу моделей. При цьому потребує ефективного вирішення актуальна задача диференціації форм представлення інформації та їх відображення відповідними моделями.

II. Структуризація характеристик інформаційних потоків

Основними формами представлення структуризованої інформації (СІ) є: символічна, текстова та графічна [1]:

- символічна форма застосовується при передаванні сигналів і базується на використанні системи символів: знаки, цифри, букви, аббревіатури, коди та ін.;
- текстова форма є основана на системному, змістовому впорядкуванні цифр, літер, математичних знаків та ін. При цьому інформація, впорядкована таким чином, несе певне змістове навантаження;
- графічна форма є найбільш ємнісною і складною, представлена структурними моделями і великим об'ємом інформаційних даних.

Усю різноманітність поліфункціональних даних можна згрупувати за наступними ознаками [2]: за джерелом виникнення: біологічна, технічна, соціальна; за способом сприйняття та передавання: аудіальна, візуальна, машинна; за суспільним призначенням: особиста (призначена для конкретної особи), масова (призначена для будь-якого користувача).

Для структуризованої інформації характерні наступні атрибути (рис.1).



Рисунок 1 - Класифікація атрибутів теорії структуризації.

Як видно з рисунка стартовим атрибутом є джерело інформації, що собою представляє матеріальний об'єкт або суб'єкт інформації, який реалізує функції генерування, зберігання, накопичення та перетворення СІ у вигляді повідомлень або сигналів різної фізичної природи. При цьому існують два типи джерел структуризованої інформації: без пам'яті, в якому імовірність формування СІ не залежить від попередніх символів чи даних; з накопичувальною пам'яттю попередніх СІ, які відповідають Марківському типу джерел інформації.

Існують три типи джерел СІ, які відрізняються статистичними характеристиками: стаціонарні, квазістаціонарні та нестаціонарні, які ідентифікуються статистичними, кореляційними, спектральними, логіко-статистичними, кластерними, ентропійними та образно-кластерними моделями. Властивості та моделі різних типів джерел інформації та їх системні характеристики детально описані в роботі [3].

Визначити кількісну оцінку структуризованої інформації досить складно, оскільки в її основу входять репрезентативність, адекватність, змістовність, точність, достовірність, достатність, доступність, своєчасність, цінність, актуальність, стійкість та захищеність [4].

Репрезентативність визначає правильність структурного відображення заданих властивостей об'єкта, а також критерії відбору інформаційних даних, вибраної концепції, кодування станів об'єкта, методики вимірювання, формування та перетворення СІ.

Адекватність СІ задає певний рівень відповідності змісту отриманої інформації образу реального об'єкта. Адекватність СІ може виражатися у трьох формах: синтаксичній, семантичній та прагматичній. Синтаксична адекватність є відтворенням формально-структурних характеристик відображення станів об'єкта. На синтаксичному рівні враховують тип носія та способу представлення СІ, швидкість передавання та опрацювання СІ, формати кодів, надійність та точність перетворення СІ. Таку СІ, яка формалізується із синтаксичних позицій ідентифікують в якості даних. Семантична адекватність враховує зміст інформації і змістовні зв'язки між кодами представлення СІ. Прагматична адекватність відображає корисні властивості інформації для вироблення управлінських рішень.

Змістовність СІ визначається відношенням:

$$S = \frac{I_c}{V_d},$$

де I_c - кількість семантичної інформації;

V_d - об'єм даних.

Показник інформативності характеризується:

$$Y = \frac{I}{V_d},$$

де I - кількість синтаксичної інформації;

V_d - об'єм даних.

Точність СІ визначає ступінь відхилення одержаної інформації від реального стану об'єкта.

Достовірність СІ характеризується ймовірністю відсутності помилок при відображенні станів об'єктів з необхідною точністю. Достовірність СІ визначається виразом:

$$D = \{ \Delta \in [\Delta_f] \},$$

Δ - реальна точність відбиття параметра,

$[\Delta_f]$ - діапазон необхідної точності відбиття параметра.

Висновок

Теоретичне обґрунтування реалізації функціоналів, їх моделювання та аналіз, а також оцінка ефективності методів формування та організації руху ідентифіковано-структуризованих даних в РКС є актуальним предметом дослідження сучасних КС.

Створення фундаментальних основ теорії структуризації даних стає потужним інструментом підвищення ефективності та розширення функцій сучасних інформаційних систем та успішного вирішення завдань переходу до інформаційного суспільства.

Список використаних джерел

1. Згуровський М.З. Вступ до комп'ютерних інформаційних технологій: навч. посібник / М.З.Згуровський, І.І.Коваленко, В.М.Михайленко. – К.: Вид-во Європ. ун-ту, 2002. – 256 с.
2. Павлиш В.А. Основи інформаційних технологій і систем: навч. посібник / В.А.Павлиш, Л.К.Гліненко. – Львів: Видавництво Львівської політехніки, 2013. – 500с.
3. Николайчук Я. М. Теорія джерел інформації. / Видання друге, виправлене // - Тернопіль: ТзОВ «Терно-граф», 2010. – 536 с.
4. Николайчук Я.М. Проектування спеціалізованих комп'ютерних систем / Я.М. Николайчук, Н.Я. Возна, І.Р. Пітух // - навч. посібник для вузів. - Т.: ТзОВ «Терно-граф», 2010. - 392 с.