



МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«МИРЭА - Российский технологический университет»
РТУ МИРЭА

Институт Информационных Технологий
Кафедра Инструментального и прикладного программного обеспечения

Практическая работа №8
по дисциплине
«Проектирование информационных систем»

Студент группы ИКБО-04-22

Егоров Л.А.
(Ф.И.О. студента)

Принял

Ткаченко Д.И.
(Ф.И.О. преподавателя)

СОДЕРЖАНИЕ

| | |
|--|---|
| 1 РАСЧЁТ ИНФОРМАЦИОННОЙ ЭНТРОПИИ ПРОЕКТИРУЕМОЙ СИСТЕМЫ | 3 |
| 1.1 Цель и задачи работы | 3 |
| 1.2 Описание ЭСЕ и наполнение системы | 3 |
| 1.3 Расчёт математического ожидания системы | 4 |
| 1.4 Расчёт дисперсии системы | 4 |
| 1.5 Расчёт среднеквадратического отклонения системы | 5 |
| 1.6 Расчёт энтропии системы | 5 |
| 1.7 Параметры проектируемой ИС | 5 |
| ЗАКЛЮЧЕНИЕ | 6 |

1 РАСЧЁТ ИНФОРМАЦИОННОЙ ЭНТРОПИИ ПРОЕКТИРУЕМОЙ СИСТЕМЫ

1.1 Цель и задачи работы

Цель работы: изучить методы расчета информационной энтропии проектируемой системы и применить их для анализа степени неопределенности данных в рамках проектирования информационной системы.

Задачи:

- освоить теоретические основы информационной энтропии, включая формулу Шеннона;
- провести анализ данных проектируемой системы для определения ключевых параметров, влияющих на энтропию;
- выполнить расчет энтропии системы на основе выявленных параметров;
- проинтерпретировать полученные результаты и оценить их влияние на проектирование информационной системы.

1.2 Описание ЭСЕ и наполнение системы

Элементарная семантическая единица (ЭСЕ) – неделимая единица информации, используемая в ИС. ЭСЕ представляет собой завершенную контекстную конструкцию, вызываемую в результате поиска по различным атрибутам или в результате тех или иных команд в виде отклика или отчета. В случае исследования системы «Электронный университет» за ЭСЕ выбрано количество долгов по изучаемым дисциплинам.

Система наполнена 100 ЭСЕ, но в рамках ограничения объёма данной работы приведены только первые десять записей в Таблице 1.1.

Таблица 1.1 — Список элементарных семантических единиц

| Наименование дисциплины | Количество долгов |
|---|-------------------|
| Проектирование баз данных | 560 |
| Математическая логика и теория алгоритмов | 1940 |
| Информатика | 733 |

Продолжение Таблицы 1.1

| | |
|---|------|
| Математический анализ | 456 |
| Моделирование сред и разработка приложений виртуальной и дополненной реальности | 976 |
| Проектирование информационных систем | 1034 |
| Объектно-ориентированное программирование | 287 |
| Структуры и алгоритмы обработки данных | 254 |
| Философия | 156 |
| Безопасность жизнедеятельности | 1 |

Для оценки данных вероятностей принято решение разбить весь диапазон значений на 10 дискретных величин с шагом в 200 (Таблица 1.2).

Таблица 1.2 — Список элементарных семантических единиц

| Диапазон | x | $P(x)$ |
|-------------|-------|--------|
| 0 — 200 | 78,5 | 0,2 |
| 200 — 400 | 270,5 | 0,2 |
| 400 — 600 | 508 | 0,2 |
| 600 — 800 | 733 | 0,1 |
| 800 — 1000 | 976 | 0,1 |
| 1000 — 1200 | 1034 | 0,1 |
| 1200 — 1400 | 0 | 0 |
| 1400 — 1600 | 0 | 0 |
| 1600 — 1800 | 0 | 0 |
| 1800 — 2000 | 1940 | 0,1 |

1.3 Расчёт математического ожидания системы

Расчет математического ожидания распределения долгов выполняется по Формуле 1.1.

$$M_{x_i} = \sum_{i=0}^n [p_i \cdot x_i] \quad (1.1)$$

Рассчитанное значение математического ожидания равно 639,7 [долгов].

1.4 Расчёт дисперсии системы

Расчет дисперсии распределения долгов выполняется по Формуле 1.2.

$$D(x_i) = M(x_i^2) - M^2(x_i) \quad (1.2)$$

Рассчитанное значение дисперсии равно 290525,31 [долгов²].

1.5 Расчёт среднеквадратического отклонения системы

Расчет среднеквадратического отклонения распределения долгов выполняется по Формуле 1.3.

$$\sigma_{x_i} = \sqrt{D_{x_i}} \quad (1.3)$$

Рассчитанное значение среднеквадратического отклонения равно 539 [долгов].

1.6 Расчёт энтропии системы

Энтропия системы – это сумма произведений вероятностей различных состояний системы на логарифмы этих вероятностей, взятая с обратным знаком (Формула 1.4).

$$H(x) = - \sum_{i=0}^n [p_i \cdot \log_a p_i] \quad (1.4)$$

За основание логарифма a взято двоичное основание. Используя данные из Таблицы 1.2, получено значение энтропии, равное 2,722 [бит].

1.7 Параметры проектируемой ИС

В Таблице представлены результаты расчётов основных характеристик проектируемой ИС:

Таблица 1.3 — Параметры проектируемой ИС

| Параметр | Значение |
|---------------------------------|----------------------------------|
| Математическое ожидание | 639,7 [долгов] |
| Дисперсия | 290525,31 [долгов ²] |
| Среднеквадратическое отклонение | 539 [долгов] |
| Энтропия | 2,722 [бит] |

ЗАКЛЮЧЕНИЕ

В рамках практической работы изучена теория информационной энтропии как меры неопределенности данных. На основе анализа параметров проектируемой системы проведены расчеты по формуле Шеннона. Результаты расчетов позволили количественно оценить уровень хаотичности информации в системе и выявить области, требующие оптимизации. Полученные данные могут быть использованы для улучшения структуры базы данных, минимизации избыточности информации и повышения эффективности обработки данных в проектируемой информационной системе.