



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

Институт информационных технологий
Кафедра инструментального и прикладного программного обеспечения

Отчет по практической работе №8
по дисциплине
«Проектирование информационных систем»

Тема работы:
«Создание полного текстового описания, глоссария и расчет
параметров проектируемой информационной системы»

Выполнил: студент группы ИВБО-02-19

К. Ю. Денисов

Принял: ассистент

А. А. Русляков

Москва 2022

1 Создание диаграммы состояний

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 настоящей системы за элементарную семантическую единицу была выбрана одна из характеристик поиска, а именно, количество файлов, удовлетворяющим пользовательским критериям поиска. В нашем примере эта величина меняется случайным образом в пределах от 0 до 49999 [файлов].

1.1 Наполнение системы

Проектируемая информационная система «Электронный сборник лабораторных работ» может быть наполнена практически любым количеством элементов базы данных. Их количество ограничиваются только параметрами сервера.

В рамках данной практической работы система была наполнена 50000 ЭСЕ. Было проведено сто экспериментов, в ходе которых стало известно количество файлов, удовлетворяющих пользовательским параметрам, заданным в модуле поиска Подсистемы хранения ИС «Электронный сборник лабораторных работ». Список результатов измерений приведен в приложении 1.

1.2 Математические расчеты

Для дальнейшего исследования проектируемой ИС необходимо рассчитать вероятности, с которыми ЭСЕ принимает то или иное значение. Для оценки этих вероятностей было принято решение разбить весь диапазон значений на 10 дискретных величин с шагом в 5000. Расчеты вероятности ведутся с помощью следующей формулы 1.

$$P(\xi) = \frac{n}{N} \quad (1)$$

В данной формуле n — благоприятное число исходов (в данном случае число найденных файлов, попавших в данный диапазон), а N — общее число исходов.

В таблице 1 приведены возможные значения, принимаемые ЭСЕ и их вероятности.

Таблица 1 — Ряд распределения

№	x_i	$P(x)$
1	2499.50	150.15
2	7499.50	100.10
3	12499.50	80.08
4	17499.50	90.09
5	22499.50	110.11
6	27499.50	130.13
7	32499.50	100.10
8	37499.50	90.09
9	42499.50	60.06
10	47499.50	90.09

1.2.1 Расчет математического ожидания информационного блока системы

Математическим ожиданием случайной величины называется сумма произведений всех возможных значений случайной величины на вероятности этих значений.

Рассчитаем математическое ожидание для нашей системы, взяв за случайную величину число файлов. Расчёт математического ожидания информационного блока на примере 10 записей:

$$M_{x_i} = \sum_{i=0}^n [p_i \cdot x_i] \quad (2)$$

Используя данные, полученные в таблице 1, получаем:

$M(10) = 23199.50$ файлов, следовательно, наиболее вероятное количество файлов в ответе на запрос находится в районе 23199.50 [файлов].

1.2.2 Расчет дисперсии информационного блока системы

Рассчитаем дисперсию информационного блока системы по формуле 3.

$$D_{x_i} = \sum_{i=0}^n [p_i \cdot (x_i)^2] - \left[\sum_{i=0}^n (p_i \cdot x_i) \right]^2 \quad (3)$$

Используя данные, полученные в таблице 1, получаем:

$$D(10) = 206010000 \text{ файлов}^2$$

1.2.3 Расчет среднеквадратического отклонения

Рассчитаем среднеквадратическое отклонение по формуле 4.

$$\sigma x_i = \sqrt{Dx_i} = \sqrt{206010000} = 14353.05 \text{ файлов} \quad (4)$$

1.2.4 Расчет энтропии системы

Информационная энтропия — мера неопределённости некоторой системы (в статистической физике или теории информации), в частности, непредсказуемость появления какого-либо символа первичного алфавита. В последнем случае при отсутствии информационных потерь энтропия численно равна количеству информации на символ передаваемого сообщения

Вычислим энтропию системы по формуле

$$H(x) = - \sum_{i=1}^n [p_i \cdot \log_a p_i] \quad (5)$$

За основание логарифма a возьмем двоичную систему счисления (формула Шеннона). Энтропия фрагмента информационного наполнения в размере 10 ЭСЕ: Используя данные, полученные в таблице 1, получаем: $H(10) = 3.28$ бит.

1.3 РЕЗУЛЬТАТЫ РАСЧЕТОВ

В данной главе был осуществлен расчет основных характеристик проектируемой ИС, и получены следующие результаты:

Таблица 2 — Параметры проектируемой ИС

Математическое ожидание информационного блока	$M(10) = 23199.50$ файлов
Допустимый разброс значений смысловых информационных блоков (дисперсия)	$D(10) = 206010000$ файлов ²
Среднеквадратичное отклонение	$\sigma x_i = 14353.05$ файлов
Энтропия информационного наполнения	$H(10) = 3.28$ бит

Приведем таблицу с исходными сгенерированными данными (см. таблицу 3).

Таблица 3 — Исходные значения

№	Значение	№	Значение	№	Значение	№	Значение
1	19956	26	1983	51	3986	76	47044
2	30937	27	26452	52	20469	77	26520
3	49379	28	899	53	14592	78	35967
4	32802	29	41888	54	1352	79	21749
5	3998	30	8182	55	32704	80	18295
6	31012	31	2176	56	40990	81	21813
7	38782	32	28627	57	29975	82	13467
8	15895	33	25433	58	10325	83	23783
9	27847	34	9457	59	3700	84	32821
10	35964	35	2373	60	24083	85	17607
11	3827	36	10010	61	33412	86	36274
12	24247	37	48230	62	43067	87	10648
13	39557	38	26061	63	49735	88	19679
14	2321	39	5810	64	40199	89	49114
15	29616	40	25460	65	20796	90	43305
16	24485	41	30058	66	27099	91	21310
17	46153	42	859	67	36430	92	31711
18	35865	43	46295	68	18874	93	18424
19	44427	44	1274	69	6761	94	6493
20	26134	45	37563	70	8725	95	5029
21	2133	46	47881	71	25199	96	9433
22	3270	47	15273	72	21981	97	38202
23	27499	48	1346	73	48131	98	33052
24	33954	49	11422	74	13900	99	24332
25	17943	50	11215	75	6764	100	5224