

Методика выявления динамической зависимости между группами событий

Е. В. Зубков, В. М. Белов

В работе рассматривается проблематика использования энтропийного подхода в анализе данных (ЭПАД) применительно к задаче выявления динамической зависимости между различными группами событий. Приводится обоснование применения ЭПАД, описание методики, результаты вычислительного эксперимента.

Ключевые слова: интеллектуальный анализ данных, Data Mining, энтропия, энтропийный подход к анализу данных, события, динамическая зависимость.

1. Введение

Методы, алгоритмы, методики, работы с большими массивами данных представляют и по настоящее время высокий интерес. Как правило, это предмет статистических исследований.

Однако с появлением междисциплинарных подходов к обработке информации, таких как интеллектуальный анализ данных (ИАД), также известный как Data Mining, всё большее число исследователей стремится использовать последние, понимая существующие ограничения теории вероятностей и математической статистики, а также благодаря возможности комплексно подойти к анализу данных со стороны новых методов. Отсюда формируется ряд часто возникающих задач, связанных с обработкой больших массивов данных, которые становятся задачами вышеупомянутых подходов.

Базовым понятием всей теории информации, а в дальнейшем и ЭПАД, является понятие энтропии. Энтропия – мера неопределённости некоторой ситуации. Можно также назвать её мерой рассеяния, и в этом смысле она подобна дисперсии [1]. Но если дисперсия является адекватной мерой рассеяния лишь для специальных распределений вероятностей случайных величин, то энтропия таких ограничений не имеет.

В силу своей универсальности она находит применение в самых различных прикладных областях науки и техники: математика, физика, социология, медицина, биология, информационная безопасность, криптография, экология, геология и т.д. [2–13].

Кроме универсальности энтропия обладает и другими полезными свойствами. Так, если некий опыт имеет n равновероятных исходов, а другой опыт – m равновероятных исходов, то составной опыт имеет $n \times m$ равновероятных исходов. Для используемой меры неопределённости f вполне естественно потребовать, чтобы она, во-первых, росла с ростом числа возможных исходов, а во-вторых, чтобы неопределённость составного опыта была равна просто сумме неопределённости отдельных опытов: $f(nm) = f(n) + f(m)$. Иными словами, мера неопределённостей должна обладать свойством аддитивности. Именно такая удобная величина была введена К. Шеноном [14, 15]:

$$H(X) = - \sum_{i=1}^N P(X_i) \log X_i,$$

где X – дискретная случайная величина с диапазоном изменчивости N , $P(X_i)$ – вероятность i -го значения X .

В теории информации в формуле для энтропии обычно используют двоичные логарифмы, тогда энтропия и информация измеряется в битах. В общем случае выбор основания логарифма – лишь вопрос масштаба, поскольку энтропия безразмерна. Возможная величина энтропии заключена в пределах

$$0 \leq H(X) \leq \log N.$$

Нижняя граница соответствует вырожденному распределению, когда неопределённость величины X отсутствует. Верхняя граница соответствует равномерному распределению, т.е. все N значений X_i встречаются с равной вероятностью.

Если две случайные величины X и Y некоторым образом связаны между собой (например, являются признаками множества исследуемых объектов), то знание значения одной из них, как правило, уменьшает неопределённость значений другой. Это легко продемонстрировать с помощью диаграммы Венна (рис. 1.1):

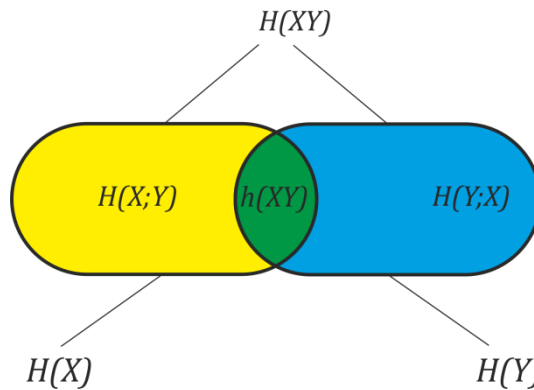


Рис. 1.1. Диаграмма Венна – энтропии зависимых величин

Здесь $H(X)$, $H(Y)$ – значения энтропии величин X и Y соответственно. Энтропия совокупности этих значений $H(XY)$ в общем случае зависит от степени статистической зависимости между X и Y . Так, например, в случае полной связи (т.е. при отношении один к одному) имеем:

$$H(XY) = H(X) = H(Y).$$

Для случая полной независимости вместо этого получим:

$$H(XY) = H(X) + H(Y).$$

Для общего случая справедливы следующие утверждения:

$$H(XY) - H(Y) = H(X; Y),$$

$$H(XY) - H(X) = H(Y; X).$$

Здесь $H(X; Y)$ и $H(Y; X)$ – условные энтропии величин X и Y соответственно. Они показывают, какая часть энтропии остаётся, если становится известно значение второй величины. В случае полной зависимости X и Y условная энтропия равна нулю. Для второго граничного условия при независимости этих величин условная энтропия будет равна полной энтропии, т.е. выполняются:

$$H(X; Y) = H(X),$$

$$H(Y; X) = H(Y).$$

Степень взаимного влияния фактически определяется областью пересечения $h(XY)$. Относительные частные коэффициенты влияния X на Y и Y на X составят соответственно:

$$\frac{h(XY)}{H(X)} \text{ и } \frac{h(XY)}{H(Y)}.$$

Учитывая то, что значение $h(XY)$ симметрично относительно X и Y , для оценки зависимости целесообразно использовать более удобный и эффективный показатель [2]:

$$\frac{h(XY)}{H(XY)},$$

который принимает значения от 0 в случае независимости до 1 в случае их полной связи.

Эффективность применения ЭПАД отмечают в своих работах многие исследователи. Среди отличительных качеств, свойственных энтропии, а следовательно, и ЭПАД, как правило, отмечают:

- независимость от вида распределения;
- универсальность;
- обладание свойством аддитивности.

В настоящей статье рассматривается задача соотнесения между собой групп (последовательностей) событий с тем, чтобы сделать выводы о наличии возможной взаимосвязи между ними. Составляющей частью решения обозначенной задачи ИАД является энтропийный подход к анализу данных. В его основе лежат описанные выше принципы.

2. Методика выявления динамической зависимости с использованием ЭПАД

Постановка задачи. Рассмотрим два множества (класса) событий A и B . Будем считать, что каждое событие класса A переводит некоторую систему A'' в новое состояние. Аналогичным образом каждое событие класса B переводит в новое состояние систему B'' . Тогда, располагая информацией о точном времени наступления всех событий, мы знаем состояния систем a_i и b_j (для систем A'' и B'' соответственно) в произвольные моменты времени t_i . В этом случае считаем, что системы полностью определены. Нам необходимо определить степень согласованности по времени событий, составляющих описанные ранее классы.

Решение. Пусть за анализируемый промежуток времени T произошло n событий класса $A: \{a_1, a_2, a_3, \dots, a_n\}$ и m событий класса $B: \{b_1, b_2, b_3, \dots, b_m\}$. Для класса A получим $n + 1$ интервалов времени (рис. 2.1), для которых выполняется следующее условие:

$$T = \sum_{i=1}^{n+1} dt_i,$$

где $dt_i = t_i - t_{i-1}$ (интервалы времени).



Рис. 2.1. Распределение событий класса A во времени

Вероятность того, что в момент наступления некоторого произвольного события система находится в состоянии a_i , будет равна

$$p_i = \frac{dt_i}{T}. \quad (2.1)$$

Соответственно энтропия класса A на временном отрезке T может быть рассчитана как

$$H_A = -\sum_{i=1}^{n+1} p_i \log_2 p_i.$$

Аналогичным образом, согласно изложенному выше, получаем энтропию класса B :

$$H_B = -\sum_{j=1}^{m+1} p_j \log_2 p_j,$$

$$p_j = \frac{dt_j}{T},$$

где p_j – вероятность наступления события класса B , dt_i – интервалы времени для $m + 1$ событий класса B , T – анализируемый промежуток времени.

При совмещении на временной оси событий обоих классов возможны три случая.

Случай первый. События класса A и класса B происходят строго синхронно (рис. 2.2).



Рис. 2.2. Динамическая зависимость – случай первый

$$H_a = H_b = H_{ab}.$$

Наблюдается максимальная зависимость по времени. Информация о событиях одного класса содержит в себе полную информацию обо всех событиях другого класса.

Случай второй. Диапазоны времени событий множества A и множества B не пересекаются (рис. 2.3). Состояния систем A'' и B'' изменяются независимо друг от друга. Энтропия совокупного множества вычисляется как

$$H_{ab} = H_a + H_b + \Delta H.$$

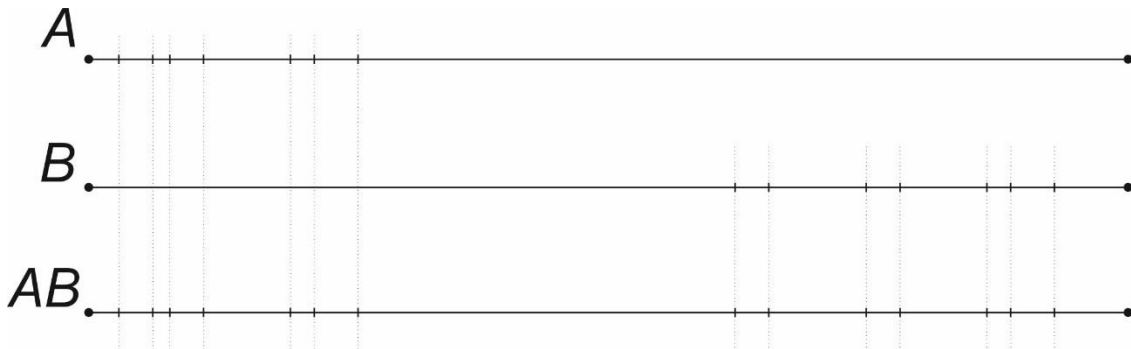


Рис. 2.3. Динамическая зависимость – случай второй

При этом ΔH определяется интервалом времени между последним элементом множества A и первым элементом множества B . Если этот интервал обозначить как t_{a_n, b_1} , то

$$\Delta H = -\frac{t_{a_n, b_1}}{T} \log_2 \frac{t_{a_n, b_1}}{T}.$$

Случай третий. Диапазоны времени событий множества A и множества B пересекаются (рис. 2.4). Это самый общий случай (предыдущие два можно рассматривать в качестве граничных примеров). Изменение состояния систем A и B происходит одновременно.

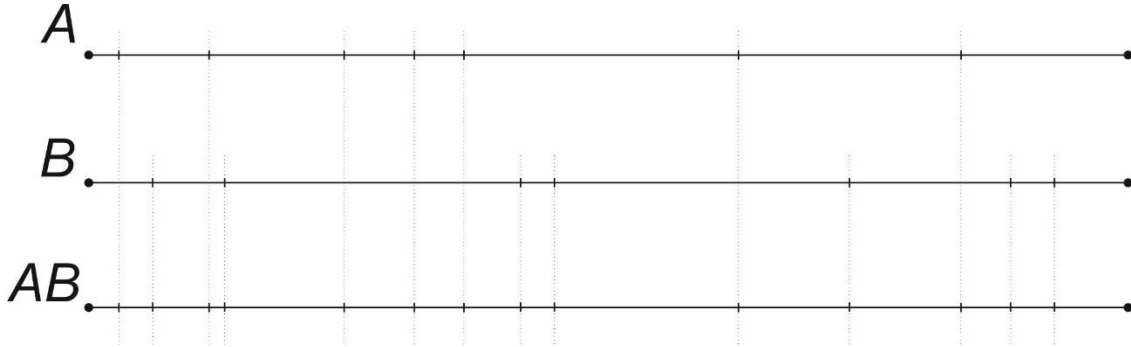


Рис. 2.4. Динамическая зависимость – случай третий

Проекция событий A и B на единую временную ось описывает изменение состояния системы AB , которая является консолидирующей по отношению к системам, её составляющим.

Информация о событиях класса B включает в себя некоторую часть информации о событиях класса A и наоборот. Например, событие b_j наступило в момент времени t_j^b , а событие b_{j+1} , соответственно, в момент времени t_{j+1}^b . Событие a_i наступило в момент времени t_i^a , причем $t_j^b < t_i^a < t_{j+1}^b$ (рис. 2.5).

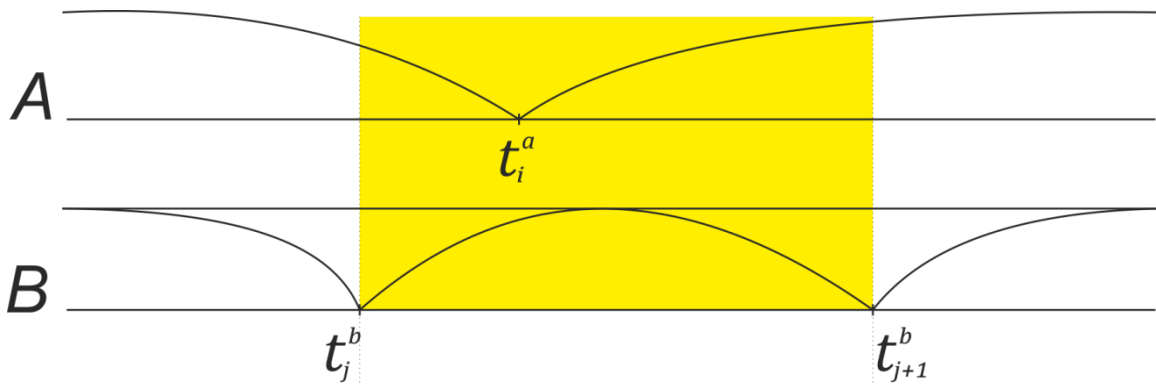


Рис. 2.5. Зависимость между событиями a_i и b_j

Тогда можно сказать, что информация о времени наступления событий b_i и b_{i+1} содержит в себе также информацию и о времени наступления события a_j . Точность этой информации определяется продолжительностью интервала $[t_j^b; t_{j+1}^b]$.

Вероятность того, что в момент наступления произвольного события система будет находиться в состоянии $a_i b_j$, вычисляем:

$$p_k = \frac{dt_k}{T},$$

где $k = i + j$, а t_k и T – вышеопределённые обозначения.

Находим энтропию такой системы:

$$H_{AB} = - \sum_{k=1}^{n+m+1} p_k \log_2 p_k .$$

Относительный частный коэффициент взаимного влияния исходных систем можно определить по формуле:

$$h_{AB} = H_A + H_B - H_{AB} . \quad (2.2)$$

Взаимная динамическая информативность множеств событий A и B составит:

$$I_{AB} = \frac{h_{AB}}{H_{AB}} . \quad (2.3)$$

Если значение взаимной информативности ниже ранее определённого порогового значения, то динамическую зависимость событий таких классов считаем несущественной. Напротив, высокое значение этой характеристики говорит об относительно согласованном характере событий в различных классах (если $I_{AB} = 1$, значит, события в классах A и B происходят строго синхронно). Величина порогового значения определяется эмпирическим путём после проведения серии экспериментов на контрольной выборке.

3. Вычислительный эксперимент

Применим методику из предыдущего раздела для оценки динамической зависимости между событиями классов. Время наступления событий каждого класса можно оценить по представленному графику (рис. 3.1).

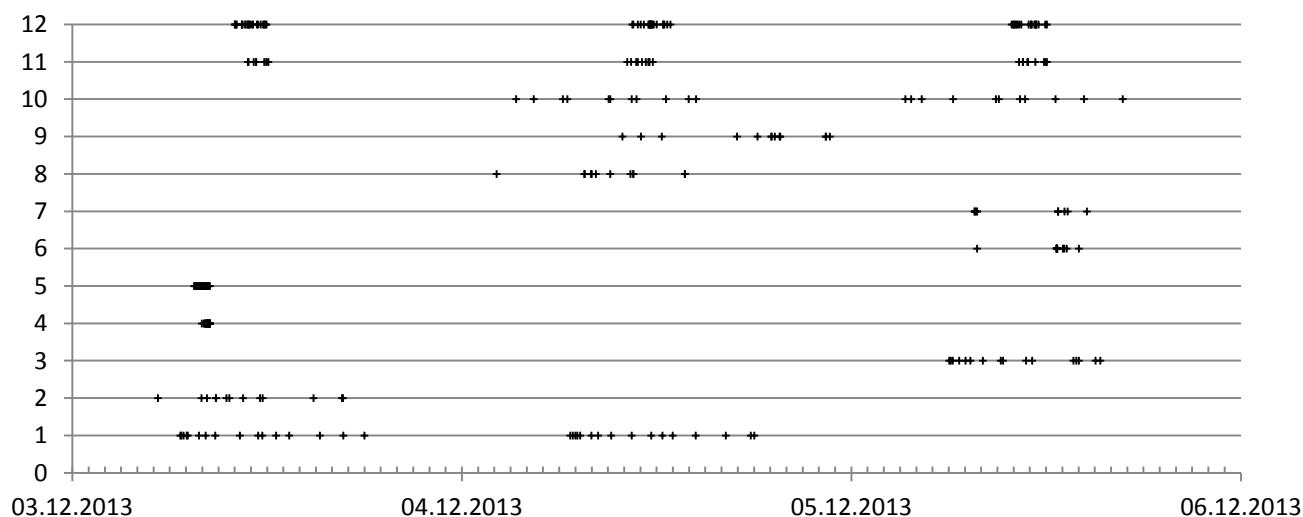


Рис. 3.1. График наступления событий различных классов

Рассмотрим процесс определения динамической зависимости на примере классов № 4 и № 5. Точная хронология событий этих классов, а также их статистические и энтропийные показатели приведены в табл. 3.1–3.3.

Таблица 3.1. Информация о событиях класса № 4

i	Время событий	Интервал, с	P	dH_i
	03.12.2013 00:00:00	0	0	0
1.	03.12.2013 7:58:22	28702	0.11073	0.35156
2.	03.12.2013 8:06:31	489	0.00189	0.01707
3.	03.12.2013 8:07:41	70	0.00027	0.00320
4.	03.12.2013 8:10:59	198	0.00076	0.00791
5.	03.12.2013 8:11:00	1	0.00000	0.00007
6.	03.12.2013 8:12:20	80	0.00031	0.00360
7.	03.12.2013 8:13:06	46	0.00018	0.00221
8.	03.12.2013 8:14:07	61	0.00024	0.00284
9.	03.12.2013 8:16:12	125	0.00048	0.00531
10.	03.12.2013 8:17:42	90	0.00035	0.00399
11.	03.12.2013 8:18:23	41	0.00016	0.00200
12.	03.12.2013 8:20:11	108	0.00042	0.00468
13.	03.12.2013 8:20:32	21	0.00008	0.00110
14.	03.12.2013 8:20:45	13	0.00005	0.00072
15.	03.12.2013 8:21:13	28	0.00011	0.00142
16.	03.12.2013 8:21:43	30	0.00012	0.00151
17.	03.12.2013 8:22:03	20	0.00008	0.00105
18.	03.12.2013 8:22:27	24	0.00009	0.00124
19.	03.12.2013 8:22:58	31	0.00012	0.00156
20.	03.12.2013 8:23:14	16	0.00006	0.00086
21.	03.12.2013 8:23:32	18	0.00007	0.00096
22.	03.12.2013 8:24:04	32	0.00012	0.00160
23.	03.12.2013 8:24:27	23	0.00009	0.00119
24.	03.12.2013 8:24:45	18	0.00007	0.00096
25.	03.12.2013 8:24:59	14	0.00005	0.00077
26.	03.12.2013 8:25:17	18	0.00007	0.00096
27.	03.12.2013 8:25:38	21	0.00008	0.00110
28.	03.12.2013 8:26:01	23	0.00009	0.00119
29.	03.12.2013 8:26:11	10	0.00004	0.00057
30.	03.12.2013 8:26:25	14	0.00005	0.00077
31.	03.12.2013 8:26:47	22	0.00008	0.00115
32.	03.12.2013 8:27:08	21	0.00008	0.00110
33.	03.12.2013 8:27:33	25	0.00010	0.00129
34.	03.12.2013 8:27:46	13	0.00005	0.00072
	06.12.2013 00:00:00	228734	0.88246	0.15919
Σ		259200	1	0.58742

Таблица 3.2. Информация о событиях класса № 5

i	Время событий	Интервал, с	P	dH_i
	03.12.2013 00:00:00	0	0	0
1.	03.12.2013 7:29:22	26962	0.10402	0.33963
2.	03.12.2013 7:31:45	143	0.00055	0.00597
3.	03.12.2013 7:33:39	114	0.00044	0.00490
4.	03.12.2013 7:35:28	109	0.00042	0.00472
5.	03.12.2013 7:37:37	129	0.00050	0.00546
6.	03.12.2013 7:39:44	127	0.00049	0.00539
7.	03.12.2013 7:41:53	129	0.00050	0.00546
8.	03.12.2013 7:44:32	159	0.00061	0.00655
9.	03.12.2013 7:48:21	229	0.00088	0.00896
10.	03.12.2013 7:50:15	114	0.00044	0.00490
11.	03.12.2013 7:52:12	117	0.00045	0.00502
12.	03.12.2013 7:54:23	131	0.00051	0.00553
13.	03.12.2013 7:56:34	131	0.00051	0.00553
14.	03.12.2013 7:58:22	108	0.00042	0.00468
15.	03.12.2013 8:00:05	103	0.00040	0.00449
16.	03.12.2013 8:02:11	126	0.00049	0.00535
17.	03.12.2013 8:04:24	133	0.00051	0.00561
18.	03.12.2013 8:06:31	127	0.00049	0.00539
19.	03.12.2013 8:08:37	126	0.00049	0.00535
20.	03.12.2013 8:10:35	118	0.00046	0.00505
21.	03.12.2013 8:12:42	127	0.00049	0.00539
22.	03.12.2013 8:14:39	117	0.00045	0.00502
23.	03.12.2013 8:16:50	131	0.00051	0.00553
24.	03.12.2013 8:19:01	131	0.00051	0.00553
25.	03.12.2013 8:20:58	117	0.00045	0.00502
26.	03.12.2013 8:23:07	129	0.00050	0.00546
27.	03.12.2013 8:25:20	133	0.00051	0.00561
28.	03.12.2013 8:27:25	125	0.00048	0.00531
	06.12.2013 00:00:00	228755	0.15909	2.6476
Σ		259200	1	0.64591

Таблица 3.3. Сводная информация о событиях классов № 4 и № 5

i	Время событий		Интервал, с	P	dH_i
	Класс №4	Класс №5			
	03.12.2013 00:00:00		0	0	0
1.		03.12.2013 7:29:22	26962	0.10402	0.33963
2.		03.12.2013 7:31:45	143	0.00055	0.00597
3.		03.12.2013 7:33:39	114	0.00044	0.00490
4.		03.12.2013 7:35:28	109	0.00042	0.00472
5.		03.12.2013 7:37:37	129	0.00050	0.00546
6.		03.12.2013 7:39:44	127	0.00049	0.00539
7.		03.12.2013 7:41:53	129	0.00050	0.00546
8.		03.12.2013 7:44:32	159	0.00061	0.00655
9.		03.12.2013 7:48:21	229	0.00088	0.00896
10.		03.12.2013 7:50:15	114	0.00044	0.00490
11.		03.12.2013 7:52:12	117	0.00045	0.00502
12.		03.12.2013 7:54:23	131	0.00051	0.00553
13.		03.12.2013 7:56:34	131	0.00051	0.00553
14.	03.12.2013 7:58:22		0	0.00000	0.00000
15.		03.12.2013 7:58:22	108	0.00042	0.00468
16.		03.12.2013 8:00:05	103	0.00040	0.00449
17.		03.12.2013 8:02:11	126	0.00049	0.00535
18.		03.12.2013 8:04:24	133	0.00051	0.00561
19.	03.12.2013 8:06:31		0	0.00000	0.00000
20.		03.12.2013 8:06:31	127	0.00049	0.00539
21.	03.12.2013 8:07:41		70	0.00027	0.00320
22.		03.12.2013 8:08:37	56	0.00022	0.00263
23.		03.12.2013 8:10:35	118	0.00046	0.00505
24.	03.12.2013 8:10:59		24	0.00009	0.00124
25.	03.12.2013 8:11:00		1	0.00000	0.00007
26.	03.12.2013 8:12:20		80	0.00031	0.00360
27.		03.12.2013 8:12:42	22	0.00008	0.00115
28.	03.12.2013 8:13:06		24	0.00009	0.00124
29.	03.12.2013 8:14:07		61	0.00024	0.00284
30.		03.12.2013 8:14:39	32	0.00012	0.00160
31.	03.12.2013 8:16:12		93	0.00036	0.00411
32.		03.12.2013 8:16:50	38	0.00015	0.00187
33.	03.12.2013 8:17:42		52	0.00020	0.00246
34.	03.12.2013 8:18:23		41	0.00016	0.00200
35.		03.12.2013 8:19:01	38	0.00015	0.00187
36.	03.12.2013 8:20:11		70	0.00027	0.00320
37.	03.12.2013 8:20:32		21	0.00008	0.00110
38.	03.12.2013 8:20:45		13	0.00005	0.00072
39.		03.12.2013 8:20:58	13	0.00005	0.00072
40.	03.12.2013 8:21:13		15	0.00006	0.00081
41.	03.12.2013 8:21:43		30	0.00012	0.00151
42.	03.12.2013 8:22:03		20	0.00008	0.00105
43.	03.12.2013 8:22:27		24	0.00009	0.00124
44.	03.12.2013 8:22:58		31	0.00012	0.00156
45.		03.12.2013 8:23:07	9	0.00003	0.00051
46.	03.12.2013 8:23:14		7	0.00003	0.00041
47.	03.12.2013 8:23:32		18	0.00007	0.00096
48.	03.12.2013 8:24:04		32	0.00012	0.00160
49.	03.12.2013 8:24:27		23	0.00009	0.00119
50.	03.12.2013 8:24:45		18	0.00007	0.00096

i	Время событий		Интервал, с	P	dH _i
	Класс №4	Класс №5			
51.	03.12.2013 8:24:59		14	0.00005	0.00077
52.	03.12.2013 8:25:17		18	0.00007	0.00096
53.		03.12.2013 8:25:20	3	0.00001	0.00019
54.	03.12.2013 8:25:38		18	0.00007	0.00096
55.	03.12.2013 8:26:01		23	0.00009	0.00119
56.	03.12.2013 8:26:11		10	0.00004	0.00057
57.	03.12.2013 8:26:25		14	0.00005	0.00077
58.	03.12.2013 8:26:47		22	0.00008	0.00115
59.	03.12.2013 8:27:08		21	0.00008	0.00110
60.		03.12.2013 8:27:25	17	0.00007	0.00091
61.	03.12.2013 8:27:33		8	0.00003	0.00046
62.	03.12.2013 8:27:46		13	0.00005	0.00072
	06.12.2013 00:00:00		228734	0.88246	0.15919
Σ			259200	1.00000	0.65496

Расчётный период составил три дня или 259200 секунд ($3 \times 60 \times 60 \times 24$). Следуя принятому в разделе 2 допущению, вероятность того, что в момент наступления произвольного события система будет находиться в заданном состоянии, определяется по формуле 2.1. В данном случае формула примет следующий вид:

$$P = \frac{t_i}{259200}.$$

Энтропию вычисляем по формуле:

$$H = \sum_{i=1}^N dH_i,$$

где N – количество событий в исследуемом классе, а $dH_i = -P_i \log_2 P_i$.

В результате получим:

- энтропия для событий 4 класса: $H_4 = 0.58742$;
- энтропия для событий 5 класса: $H_5 = 0.64591$;
- совместная энтропия событий 4 и 5 классов: $H_{4,5} = 0.64591$.

Согласно 2.2 и 2.3, взаимная динамическая зависимость между классами № 4 и № 5 соответственно составит:

$$I = \frac{H_4 + H_5 - H_{4,5}}{H_{4,5}} = \frac{0.58742 + 0.64591 - 0.65496}{0.65496} = 0.883072.$$

Вычисление динамических зависимостей для остальных пар классов выполняется аналогичным образом. Результаты вычислений приведены в табл. 3.4. Можно наблюдать, что чем меньше значение величины I , тем более разнесены по времени события данной пары классов. Определив пороговое значение для динамической зависимости как $I_{min} = 0.6$, можно выделить пять пар классов, для которых выполняется условие $I \geq I_{max}$. Как видно из графика, события этих классов действительно коррелируют по времени.

Таблица 3.4. Значения динамической зависимости между классами событий

i	j	I
1	2	0.3817
1	3	0.0944
1	4	0.1696
1	5	0.178
1	6	0.0904
1	7	0.0902
1	8	0.458
1	9	0.4123
1	10	0.3666
1	11	0.4527
1	12	0.436
2	3	0.0398
2	4	0.3789
2	5	0.3771
2	6	0.0408
2	7	0.0409
2	8	0.1566
2	9	0.1007
2	10	0.0986
2	11	0.2679
2	12	0.2674
3	4	0.0259
3	5	0.0252
3	6	0.6609
3	7	0.6511
3	8	0.1139
3	9	0.1723
3	10	0.3709
3	11	0.2791
3	12	0.2872
4	5	0.8831
4	6	0.0283
4	7	0.0286

i	j	I
4	8	0.0878
4	9	0.0586
4	10	0.0512
4	11	0.1552
4	12	0.1586
5	6	0.0273
5	7	0.0276
5	8	0.0855
5	9	0.0572
5	10	0.0503
5	11	0.1519
5	12	0.1549
6	7	0.9055
6	8	0.114
6	9	0.1692
6	10	0.3118
6	11	0.261
6	12	0.2752
7	8	0.1142
7	9	0.1693
7	10	0.3039
7	11	0.2644
7	12	0.2792
8	9	0.4668
8	10	0.4778
8	11	0.3821
8	12	0.3983
9	10	0.4003
9	11	0.3914
9	12	0.4044
10	11	0.4605
10	12	0.4596
11	12	0.8404

Таблица 3.5. Динамически зависимые классы

i	j	I
3	6	0.6609
3	7	0.6511
4	5	0.8831
6	7	0.9055
11	12	0.8404

4. Выводы

Выявление скрытых закономерностей в больших информационных массивах относят к основным задачам ИАД. Расчёт динамической зависимости между последовательностями событий – частный случай этого класса задач. Основным результатом применения предлагаемого ЭПАД заключается в возможности соотнести между собой различные группы событий. Результаты вычислительного эксперимента наглядно подтверждают соответствие расчётных значений интуитивно ожидаемым. Степень взаимной связи выражается в виде числовой величины от 0 до 1. Она является мерой согласованности событий по времени и может быть

использована на следующих этапах ИАД в качестве эквивалента расстояния между объектами.

Литература:

1. *Коротаев С. М.* Энтропия и информация – универсальные естественные понятия. [Электронный ресурс], URL: http://www.chronos.msu.ru/old/RREPORTS/korotaev_entropia.pdf (дата обращения: 10.10.2015).
2. *Акопян С. Ц.* Применение метода сейсмической энтропии для решения технологических задач в нефтегазовой отрасли // *Rogtec. Российские нефтегазовые технологии.* 2014. № 38. С. 56–63.
3. *Арутюнян О. М., Яворский А. Г., Гулешов В. А., Селезнев М. Н., Бунятян А. А.* Контроль глубины анестезии с помощью энтропии при кардиохирургических операциях // *Вестник анестезиологии и реаниматологии.* 2011. № 6. С. 17–23.
4. *Арутюнян О.М., Яворский, А.Г., Гулешов В.А., Дутикова Е.Ф., Федулова С.В., Бунятян А.А.* Применение нейромониторинга на основе энтропии для определения гипоперфузии головного мозга при операциях на сонных артериях // *Патология кровообращения и кардиохирургия.* 2011. № 3. С. 51–57.
5. *Бородин А. А.* Энтропия и применение энтропии для анализа систем // *Вестник РосНОУ.* 2010. № 3. С. 33–36.
6. *Королев О. Л.* Применение энтропии при моделировании процессов принятия решений в экономике. Монография / О. Л. Королев, М. Ю. Кусый, А. В. Сигал / Под ред. доц. А. В. Сигала. Симферополь : Издательство «ОДЖАКЪ», 2013. 148 с.
7. *Корольков Б. П.* Термодинамические основы самоорганизации: монография / Б. П. Корольков. Иркутск : ИрГУПС, 2011. 120 с.
8. *Матвеева В. С.* Энтропия и ее использование для решения задач информационной безопасности // *Безопасность информационных технологий.* 2014. № 3. С. 30–36.
9. *Мёллер Ф.* Роль энтропии в номинальной классификации / Ф. Мёллер // *Математика в социологии. Моделирование и обработка информации* / Под ред. Аганбегяна А., Блейлока Х., Бородкина Ф., Р. Будона, Капекки В. М. : Мир, 1977. 385 с.
10. *Осипов А. И., Уваров А. В.* Энтропия и ее роль в науке // *Соросовский образовательный журнал.* 2004. Т. 8, № 1. С. 70–79.
11. *Прангишвили И. В.* Энтропийные и другие системные закономерности: Вопросы управления сложными системами / И.В. Прангишвили; Ин-т проблем управления им. В. А. Трапезникова. М. : Наука, 2003. 428 с.
12. *Сен Н. Д., Котляров В. П., Григорьев Я. Ю.* Применение оценок на основе энтропии для сравнения криптостойкости алгоритмов шифрования // *Современные наукоемкие технологии.* 2013. № 2. С. 105.
13. *Щитиков В. К., Розенберг Г. С.* Оценка биоразнообразия: попытка формального обобщения / Структурный анализ экологических систем. Количественные методы экологии и гидробиологии (Сборник научных трудов, посвященный памяти А. И. Баканова). Тольятти : СамНЦ РАН, 2005. С. 91–129.
14. *Шеннон К. Э.* Работы по теории информации в кибернетике. / под ред. Добрушина Р. Л., Лупанова О. Б; пер. с англ. М. : Издательство иностранной литературы, 1963. 830 с.
15. *Shenon C. E.* A Mathimatical Theory of Communication // *Bell System Technical Journal.* 1948. V. 27. P. 379–423, 623–656.

Зубков Евгений Валерьевич

аспирант кафедры «Безопасность и управление в телекоммуникациях» СибГУТИ,
тел.+7-913-798-01-48, e-mail: evz.nsk@gmail.com

Белов Виктор Матвеевич

д.т.н., профессор, профессор кафедры «Безопасность и управление в телекоммуникациях» СибГУТИ, тел.+7-963-906-84-83, e-mail: vmbelov@mail.ru

Technique for dynamic dependence detection between groups of events

E. Zubkov, V. Belov

The article considers the problems of entropy data analysis (EDA) application used for detection of dynamic dependence between different groups of events. Reasons for EDA usage, description of a technique, results of calculating experiment are presented.

Keywords: data mining, entropy, entropy data analysis, events, dynamic dependence.