

Н.Я. Гаркави, О.Н. Литвиненко, Н.Г. Нарису, А.В. Рыжов,
А.Н. Сахно

ИДЕНТИФИКАЦИЯ ТЕНДЕНЦИЙ ДИНАМИЧЕСКИХ ПРОЦЕССОВ МЕТОДОМ СОРТИРОВОК И ВИЗУАЛИЗАЦИИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Постановка проблемы в общем виде

Информация, полученная в пассивном эксперименте, зачастую оказывается весьма противоречивой, что объясняется не полной наблюдаемостью исследуемого процесса. Между тем, эту информацию можно рассортировать на соответствующую и не соответствующую каким-то “нулевым” гипотезам [1,2]. Иногда имеет смысл также определить вероятность того, что экспериментальная информация соответствует какой-то (или каждой) “нулевой” гипотезе.

“Нулевые” гипотезы, интересующие практикующего врача при обработке результатов клинических испытаний фармакологических препаратов, в основном перечислены в [2]. При обработке результатов ходовых испытаний подвижного состава [3] “нулевые” гипотезы для сортировок формулируются тогда, когда ходовые испытания уже закончены, а проводить совершенно новые специализированные натурные испытания не целесообразно (частично эти сортировки оговорены в нормативной литературе [4-6]).

Постановка задачи

Данная статья посвящена формулировке критериев для алгоритмирования сортировок накопленной в компьютере информации. Кроме того, показано, что при визуальной оценке рассортированной информации могут быть сформулированы весьма интересные выводы, которые до сортировок и без визуализации сформулировать было бы весьма затруднительно (если вообще возможно).

Изложение основного материала

При формулировании критериев сортировки положим, что в памяти компьютера хранятся реализации $\forall x_i(t)$ зависимостей от аргумента t какого-то показателя x , интересующего обработчика. Здесь i - номер реализации, N - количество реализаций.

© Гаркави Н.Я., Литвиненко О.Н., Нарису Н.Г., Рыжов А.В., Сахно А.Н., 2005

Предполагается, что показатель x зависит, кроме t , еще от ряда аргументов, но обработчику эти аргументы не известны. Область определения показателя x разбивается на три подмножества: $0 \leq t \leq T_1$, $T_1 < t < T_1 + T$, $T_1 + T \leq t \leq T_1 + T + T_2$. Первое и последнее подмножества – интервалы определения стабильных значений показателя, второе подмножество характеризует собственно процесс. Обычно предполагается, что $T_1 \ll T \gg T_2$. Математически критерии сортировок для достаточно простых “нулевых” гипотез [2] можно записать следующими зависимостями.

Гипотеза 1. Понижающий эффект [2] в интересующей врача постановке:

$$2 \cdot \Delta x + \max_{0 \leq t \leq T_2} x_i(t + T_1 + T) < \max_{0 \leq t \leq T_1} x_i(t). \quad (1)$$

Инженера больше интересует понижающий эффект в постановке

$$2 \cdot \Delta x + \max_{0 \leq t \leq T_2} x_i(t + T_1 + T) < \min_{0 \leq t \leq T_1} x_i(t) \quad (2)$$

и понижающий в среднем эффект

$$2 \cdot \Delta x + \frac{1}{T_2} \int_{T_1+T}^{T_1+T+T_2} x_i(t) \cdot dt < \frac{1}{T_1} \int_0^{T_1} x_i(t) \cdot dt. \quad (3)$$

В выражениях (1)-(3) и дальше Δx - допустимая погрешность измерения величин x_i ($i = 1, 2, \dots, N$).

Гипотеза 2. Повышающие [2] эффекты соответственно описываются выражениями

$$\min_{0 \leq t \leq T_2} x_i(t + T_1 + T) > 2 \cdot \Delta x + \min_{0 \leq t \leq T_1} x_i(t), \quad (4)$$

$$\min_{0 \leq t \leq T_2} x_i(t + T_1 + T) > 2 \cdot \Delta x + \max_{0 \leq t \leq T_1} x_i(t), \quad (5)$$

$$\frac{1}{T_2} \int_{T_1+T}^{T_1+T+T_2} x_i(t) \cdot dt > 2 \cdot \Delta x + \frac{1}{T_1} \int_0^{T_1} x_i(t) \cdot dt. \quad (6)$$

Гипотеза 3. Нормализующий эффект [2] в интересующей врача постановке соответствует выражению

$$\left(\left(\begin{array}{c} \exists \\ \mathbf{0} \leq t \leq T_1 \end{array} x_i(t) > x_{\max} + \Delta x \right) \vee \left(\begin{array}{c} \exists \\ \mathbf{0} \leq t \leq T_1 \end{array} x_i(t) < x_{\min} - \Delta x \right) \right) \wedge \left(\begin{array}{c} x_{\min} + \Delta x \leq \\ \forall \\ \mathbf{0} \leq t \leq T_2 \end{array} x_i(t + T_1 + T) \leq x_{\max} - \Delta x \right). \quad (7)$$

Здесь $[x_{\min}, x_{\max}]$ - интервал “нормальных” (технологически допустимых) значений величин x_i ($i = 1, 2, \dots, N$).

Нормализующий в среднем эффект соответствует выражению

$$\left(\left(\begin{array}{c} \frac{1}{T_1} \int_0^{T_1} x_i(t) \cdot dt > x_{\max} + \Delta x \\ \frac{1}{T_1} \int_0^{T_1} x_i(t) \cdot dt < x_{\min} - \Delta x \end{array} \right) \right) \wedge \left(\begin{array}{c} x_{\min} + \Delta x \leq \\ \frac{1}{T_2} \int_{T_1+T}^{T_1+T+T_2} x_i(t) \cdot dt \leq x_{\max} - \Delta x \end{array} \right). \quad (8)$$

Из достаточно простых интересующих инженера тенденций можно также назвать “стабильность” (гипотезу 4)

$$x_{\min} - \Delta x \leq \forall_{\mathbf{0} \leq t < \infty} x_i(t) \leq x_{\max} + \Delta x, \quad (9)$$

и “насыщение” (гипотезу 5)

$$x_m - \Delta x \leq \forall_{\mathbf{0} \leq t \leq T_2} x_i(t + T) \leq x_m + \Delta x. \quad (10)$$

В выражении (10) x_m - “насыщающее” значение величин x_i ($i = 1, 2, \dots, N$).

Более сложные критерии для сортировок формулируются в каждом конкретном случае.

Вероятность (частость) подтверждения гипотезы P_r определится зависимостью

$$P_r = \frac{n}{N}, \quad (11)$$

где n - количество реализаций, в которых гипотеза “г” подтвердилась.

Продемонстрируем выше сказанное на примерах.

Пример 1. Клинические испытания фармакологического препарата (чтобы избежать ненужной рекламы, диагноз и наименование лекарства здесь не указываются).

Каждому больному во взаимно однозначное соответствие поставлен массив с зависимостью температуры больного от длительности лечения. Ниже приводится формат этого массива.

Первая строка массива – идентификатор больного. Обычно идентификатором является фамилия, имя, отчество и год рождения больного.

Затем в массиве следуют строки (по три элемента в каждой строке), соответствующие суткам болезни. В каждой сутке должно быть три измерения температуры. Температура измеряется с точностью до $0,1^{\circ}\text{C}$ (при этом $\Delta x=0$). Одно два или даже все три измерения в каких-то сутках могут быть пропущены. Тогда на месте пропущенного измерения ставится символ "#". В первых и последних сутках хотя бы одно измерение должно быть обязательно. Каждый элемент информационной части массива обозначим $x_{i,j,L}$, где i - идентификатор больного, j - сутки болезни (для первой строки $j=1$, для второй строки $j=2$, для последних суток $j=J_i$), $L \in [1,2,3]$ - порядковый номер измерения в течение суток, J_i - номер последней строки в i -м массиве. $L=1$ соответствует "утреннему", $L=2$ соответствует "дневному", $L=3$ соответствует "вечернему" измерению температуры.

Таким образом, в выражениях (1), (4), (7) времени T_1 соответствуют первые сутки (одно, два или три измерения температуры): $T_1=1$ сутки. Время T_2 определим следующим правилом. Если в последних сутках нет всех трех измерений, то время T_2 - это двое последних суток; если все три измерения в последних сутках наличествуют, то T_2 - это одни последние сутки:

$$T_2 = \begin{cases} 1 \Leftarrow \bigvee_{L=1}^3 x_{i,J_i,L} \neq \text{"#"}, \\ 2 \Leftarrow \bigwedge_{L=1}^3 x_{i,J_i,L} = \text{"#"} . \end{cases}$$

Время T определяется длиной массива J_i : $T = J_i - T_1 - T_2$.

На рис.1 приведена информация о 18-и больных, выведенная в виде поля точек. Ось абсцисс – длительность лечения, ось ординат – температура больного в каждом измерении температуры (если данное измерение не было пропущено).

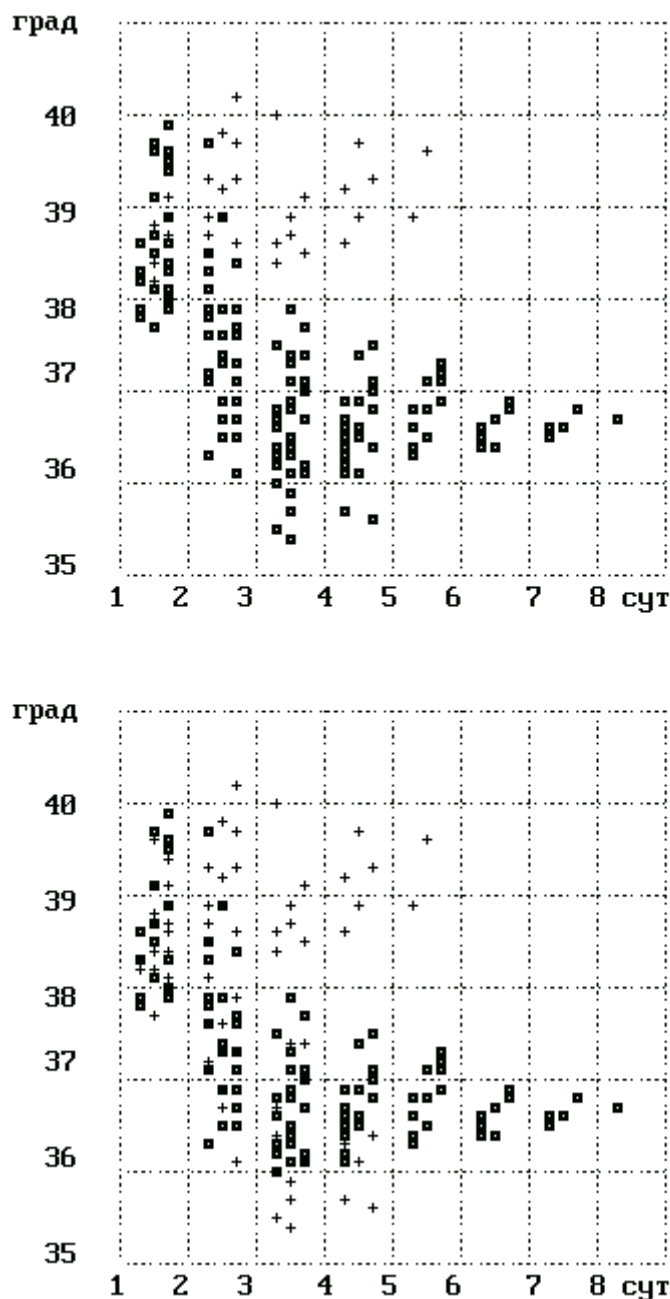


Рисунок 1 Сверху – тенденция “понижения”, внизу – тенденция “нормализации”

Перед выводом информация была рассортирована на укладываемую и не укладываемую в каждую тенденцию (см. (1), (4),

(7)). На рис.1 квадратиками обозначены реализации, удовлетворяющие критерию (1) (левый рисунок) и критерию (7) (правый рисунок), а крестиками – реализации, критерию (1) (левый рисунок) и критерию (7) (правый рисунок) не удовлетворяющие. Аналогичный рисунок можно построить и для информации, удовлетворяющей и не удовлетворяющей критерию (4).

Из рис.1 можно сделать следующий вывод. Лекарственный препарат обладает понижающим действием, но помогает не всем (14 из 18-и, т.е. $\approx 78\%$). Если на третьи сутки лечения понижающего эффекта не было, то дальнейшее лечение этим препаратом бессмысленно. Препарат в некоторых случаях имеет нормализующий эффект (10 из 18-и, т.е. $\approx 56\%$), но возможны не только случаи, когда препарат не действует (4 из 18-и, т.е. $\approx 22\%$), но и случаи, когда действие препарата остается понижающим ниже нормы (4 из 18-и, т.е. $\approx 22\%$). Эффекта, понижающего температуру до нормы или ниже нормы, удавалось достичь не более, чем за пять суток ($\max_i T = 5$).

При сортировках по критерию (7) полагали $x_{\min} = 36,0^\circ$, $x_{\max} = 37,0^\circ$.

Пример 2. Во время ходовых испытаний мотор-вагонного подвижного состава с датчиков в компьютер в течение всего времени испытаний с различными частотами квантования вводится [7] разнообразная информация (в том числе с частотой квантования 100 Гц вводится информация об изменении во времени коэффициентов вертикальной динамики $K_{ДВ}$ рессорного подвешивания экипажа [3,6]). Согласно требованиям нормативной литературы [6] введенная информация разбивается на реализации длительностью τ . Каждой реализации ставится в соответствие тип пройденного участка пути (прямая, кривая, стрелки), средняя скорость на этом участке, оценка состояния пути ("отл", "хор", "уд", "неуд") и т.п. Длительность τ определяется типом участка пути. На прямых участках $10 \leq \tau \leq 15$ секунд. В кривых и в стрелках начало и конец реализации должны быть записаны соответственно на предшествующей и последующей прямой. Таким образом, длительность реализаций в кривых и в стрелках может варьироваться в весьма широких пределах. В каждой реализации определяется наибольший по модулю $K_{ДВ}$. Документ [8] нормирует

максимально допустимое значение $K_{ДВ}$ на пути отличного, хорошего и удовлетворительного содержания. Если по результатам испытаний окажется, что $K_{ДВ}$ в каких-то реализациях превышал допустимый уровень и оценка состояния данного участка пути выше “неуд”, то предприятию-изготовителю будет предложено либо устранить этот недостаток, либо понизить заявленную конструкционную скорость.

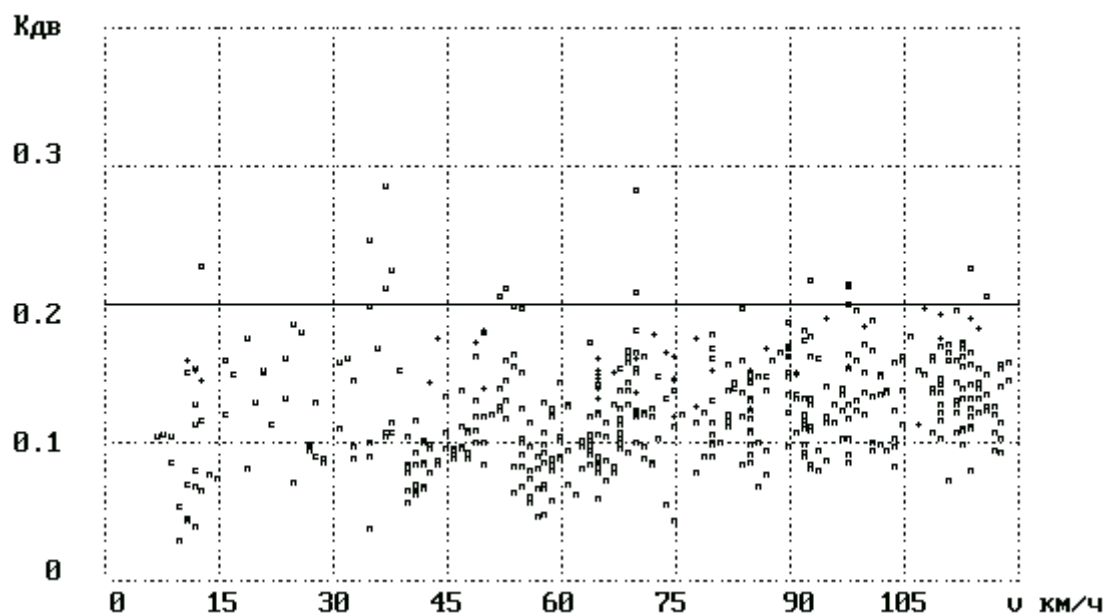


Рисунок 2 Результаты ходовых испытаний железнодорожного экипажа.

На рис.2 приведено поле точек, соответствующее зависимости $K_{ДВ}$ от скорости движения. Символом “о” обозначены $K_{ДВ}$, зафиксированные на пути неудовлетворительного содержания. Символом “п” обозначены $K_{ДВ}$, зафиксированные в прямых; символом “+” – в кривых; символом “с” – в прямых стрелках; символом “∪” – в стрелках на боковой путь. Из рис.2 следует, что $K_{ДВ}$ остается нормативным в прямых участках пути, в кривых и в прямых стрелках во всем диапазоне разрешенных для данных участков скоростей (не превышающих конструкционную скорость 120 км/час). Однако, $K_{ДВ}$ превышает нормативный уровень $x_{\max} = 0,2$ в стрелках на боковой путь при скоростях, больших $v = 30$ км/час. Таким образом, в результате испытаний данному типу экипажей будет сформулировано эксплуатационное ограничение скоростей по стрелкам на боковой путь 25 км/час. Пре-

вышение нормативного уровня $K_{ДВ} x_{\max} = 0,2$ на пути с оценкой состояния "неуд" браковочным признаком экипажа не является.

При сортировках каждой (i -той) реализации ставились в соответствие x_i - наибольшее по модулю значение $K_{ДВ}$, v_i - средняя скорость на данном участке, тип участка $t_i \in ["п", "к", "я", "б"]$ ("п" – прямой участок пути, "к" – кривая, "я" – прямые стрелки, "б" – стрелки на боковой путь) и оценка состояния пути $ц_i \in ["отл", "хор", "уд", "неуд"]$. Символ, которым на рис.2 обозначены соответствующие x_i , определяется следующими правилами

$$S(x_i) = \begin{cases} "о" \Leftarrow ц_i = "неуд", \\ "п" \Leftarrow (ц_i \neq "неуд") \wedge (t_i = "п"), \\ "+" \Leftarrow (ц_i \neq "неуд") \wedge (t_i = "к"), \\ "с" \Leftarrow (ц_i \neq "неуд") \wedge (t_i = "я"), \\ "у" \Leftarrow (ц_i \neq "неуд") \wedge (t_i = "б"). \end{cases}$$

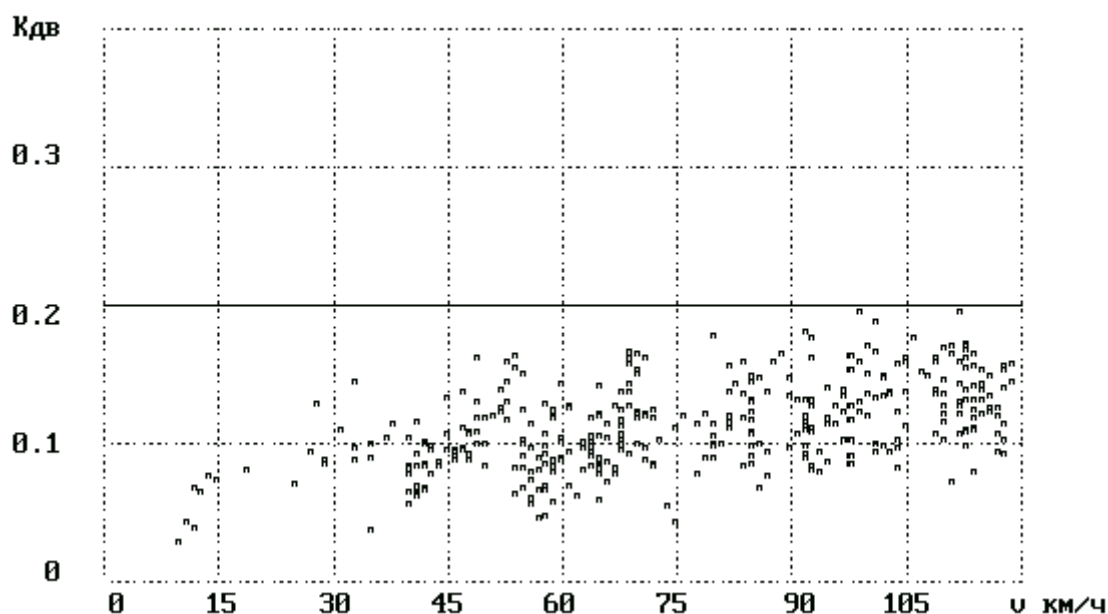


Рисунок 3. Зависимость $K_{ДВ}$ от скорости испытательного поезда на прямых участках пути отличного, хорошего и удовлетворительного содержания

Результатом сортировки $\forall x_i$ на удовлетворение гипотезы 4 при $x_{\min} = 0$ и $x_{\max} = 0,2$, если $S(x_i) \neq "о"$, будет массив, каждый элемент которого должен содержать значения $x_i > x_{\max}$, $S(x_i)$, v_i . Выведенная

на рис.2 информация (зависимость x_i от v_i с учетом $S(x_i)$ и линия $y(v) = x_{\max}$) в данном случае оказывается более информативной, чем поле точек, не удовлетворяющих гипотезе 4. На рис. 3 выведено поле точек, соответствующее $S(x_i) = "п"$. Согласно рис.3 $K_{ДВ}$ имеет тенденцию к росту с ростом скорости (подтверждается гипотеза 2, зависимость (6) при $T_1 \geq 15$ км/час, $T_2 = 15$ км/час, $T > 0$, $\Delta x = 0$).

Вывод

Сортировки с последующей визуализацией накопленной в экспериментах информации являются мощным средством идентификации тенденций исследуемых процессов.

ЛИТЕРАТУРА

1. Александровская Л.Н., Лисов А.А., Смирнов В.В. Идентификация моделей деградиционных процессов старения при эксплуатации изделий. // Методы менеджмента качества. Надежность и контроль качества. 7'99. – М.: РИА “Стандарты и качество”, 1999. – С.49-54.
2. Сахно А.Н. Профилактика послеоперационных осложнений у больных хроническими заболеваниями желчевыводящих путей. // Диссертация на соискание научной степени кандидата медицинских наук. – Харьков: Харьковский институт усовершенствования врачей, 1985.
3. Блохин Е.П., Манашкин Л.А., Данович В.Д., Коротенко М.Л., Грановский Р.Б., Мямлин С.В., Гаркави Н.Я., Федоров Е.Ф. Обработка результатов ходовых испытаний железнодорожных вагонов с целью определения их динамических качеств. // Вісті Академії інженерних наук України. Спецвипуск. – “Машинобудування”. №4(17), 2002. – С.52-61.
4. РД 24.050.37-95. Вагоны грузовые и пассажирские. Методы испытаний на прочность и ходовые качества. - 101 с.
5. СТ ССФЖТ ЦТ 15-98. Стандарт системы сертификации на федеральном железнодорожном транспорте. Тяговый подвижной состав. Типовая методика динамико-прочностных испытаний локомотивов. - М.: МПС России. Введен в действие 15.02.99 г. указанием МПС России №Г-165у. - 26 с.
6. СТ ССФЖТ ЦТ 16-98. Стандарт системы сертификации на федеральном железнодорожном транспорте. Тяговый подвижной состав.

Типовая методика динамико-прочностных испытаний электропоездов и дизель-поездов. - М.: МПС России. Введен в действие 17.11.99 г. указанием МПС России №А-2641у. - 25 с.

7. Блохин Е.П., Мямлин С.В., Грановский Р.Б., Горобец В.Л., Глухов В.В., Гаркави Н.Я., Федоров Е.Ф. О вводе в компьютер информации от датчиков при натурных испытаниях железнодорожного подвижного состава. // Вісті Академії інженерних наук України. Спецвипуск - "Машинобудування". №4(17), 2002. - С.19-28.
8. Типовая методика испытаний подвижного состава по воздействию на путь после изготовления или перед вводом в эксплуатацию. - М.: ВНИИЖТ МПС (ИЦЖТ), 1990. - 20 с.

Получено 24.10.2005 г.