УДК 519.226:519.21:51-73

О КЛАССИФИКАЦИИ ОБЪЕКТОВ ПО РЕАЛИЗАЦИЯМ ИХ ХАРАКТЕРИСТИК НА ПРОДОЛЖИТЕЛЬНОМ ОТРЕЗКЕ ВРЕМЕНИ

Е.С. Кирик*

В отличие от традиционных методов решения задач распознавания образов в основу рассматриваемого в работе подхода к классификации объектов положены методы теории случайных функций. Предлагается классифицировать объект на основе анализа случайных реализаций характеристик данного объекта на продолжительном отрезке времени. Реализация метода представлена на примере решения задачи формальной классификации состояния электролизных ванн.

Поводом для развития излагаемого в данной работе метода классификации объектов явилась следующая прикладная задача. В практике производства алюминия выделяют два основных режима работы электролизных ванн: нормальный режим и расстроенный технологический режим. Ванны, стабильно работающие в нормальном режиме (будем называть их «здоровыми»), имеют высокую производительность. Стабильно плохо работающие ванны (будем называть их «больными») имеют низкую производительность (количество металла, вылитого с ванны за определенный период времени). В этом случае некоторые показатели ванны отклоняются за рамки предписанных «удовлетворительных» диапазонов.

В процессе функционирования электролизные ванны переходят из одного режима в другой. Причем переход ванны из «здорового» режима в «больной» скорее относится к разряду самопроизвольных процессов и вызван разбалансировкой обслуживания ванны. И наоборот, перевод ванны из «больного» состояния в «здо-

---

ровое» – «принудительная», трудоемкая процедура, сопряженная с тщательным обслуживанием. Условимся называть электролизеры, находящиеся в стадии смены режима, «тревожными».

Поскольку нежелателен переход ванн из «здоровых» в «больные», то возникает потребность в мониторинге состояния ванн и выявлении тревожных. Несомненную ценность для технологов электролизных серий представляют *автоматические* системы, позволяющие предсказать ухудшение работы ванны и обратить внимание технолога на тревожные ванны. Для создания экспертных систем используются различные методы – основанные на физической природе происходящих в ваннах процессов и использующие доступные измерения технологических переменных. Например, в [1] предлагается методика, основанная на ежесуточном анализе составляющих энергобаланса каждого электролизера и использовании специально полученных критериев эффективности, которые включают элементы экспертных оценок для определения порогов.

В настоящей работе предлагается метод формальной классификации состояния ванн по их наблюдениям на предыдущем продолжительном временном участке на основе использования методов теории случайных функций. Эти методы находят успешное применение в теории колебаний, акустики, идентификации систем для решения различных задач наземного, водного и воздушного транспорта, строительства, радиотехники, медицины, экономики и др. [2-5].

Таким образом, для классификации выделим три состояния (три класса) ванн (объектов): «здоровое», «больное» – основные классы и «тревожное» – переходный класс. Известны объекты (и значения всех переменных этих объектов, технологически значимых для классификации), которые строго отнесены к одному или другому классу – к классу «здоровых» или классу «больных» электролизеров. Набор этих объектов со всеми своими числовыми показателями $SL=\{x^{(z_m)}(t_i)_{(j)}=x^{(z_m)}(i)_{(j)}\}, i=\overline{1,n}; j=\overline{1,l}; m=\overline{1,k}$ (где n – количество моментов времени, в которые производились наблюдения; l – количество наблюдаемых переменных; k – количество ванн; $x^{(m)}(i)_{(j)}$ – наблюдение j-й переменной m-й ванны в i-й момент времени; z_m – индикатор, показывающий принадлежность ванны к классу «здоровых» ($z_m=+1$) или классу «больных» ($z_m=-1$) электролизеров, $\left|\{z_m=+1\}\right|=k1$, $\left|\{z_m=-1\}\right|=k2$, $k1+k2=k$) составляет *обучающую выборку*. Формально задачу классификации определим следующим образом. Используя обучающую выборку, по измерениям контролируемых технологических переменных электролизной ванны (объекта) $V=\{x(i)_{(j)}\}, i=\overline{1,n}, j=\overline{1,l}$ будем классифицировать режим работы электролизера в $n+1$ момент времени.

Решение

Рассматриваемая задача классификации состояния электролизных ванн относится к задачам распознавания образов (РО). Традиционно задача РО состоит в отнесении предъявляемого для экзамена объекта к одному из классов. Формируют набор значимых для классификации прямых признаков объекта (непосредственно измеряемых) или косвенных признаков объекта (вычисляемых через прямые) в виде вектора значений этих признаков. Тогда M классам отвечает набор из M областей в s-мерном пространстве параметров, описывающих объект, а исследуемому объекту – точка в этом пространстве. Классы характеризуются тем, что принадлежащие им объекты обладают некоторым сходством, что выражается в близости значений характеристик, описывающих объекты. Для решения задачи отнесения объекта $\vec{x}^*=(x^*_1,...,x^*_s)$ к одному из классов каким-либо методом находят уравнение поверхностей, делящих области (классы) между собой, так называемых разделяющих поверхностей (РП) (как правило, для нахождения РП используется обучающая выборка). И затем определяют положение $\vec{x}^*$ относительно этой поверхности.

Обучающая выборка LS содержит переменные электролизных ванн, значимые с точки зрения классификации режимов работы ванн. И тогда, казалось бы, все множество элементов обучающей выборки LS можно поместить в l-мерное пространство переменных, в котором окажется $n\cdot k$ точек. Один из методов – восстановить РП. Затем по значению переменных тестируемой ванны в какой-то момент времени $(x(i)_{(1)},...,x(i)_{(l)}\}$ можно определять, к какому классу относится режим этой ванны. Однако особенность задачи в том, что, вообще говоря, объектом исследования является процесс, текущий во времени, причем инертный. Поэтому природа проходящих в электролизерах процессов, свойственная им динамичность, различные подходы в обслуживании ванн из разных классов, а также ошибки измерений и неизбежные недостатки измерительных процедур накладывают ограничения на информативность, представляемую данными, которые относятся к i-му моменту времени с точки зрения рассматриваемой задачи классификации. Существует и другая особенность. Так, одной из значимых переменных электролизера служит так называемое криолитовое отношение (КО). Технология такова, что КО не должно выходить на допустимые границы, но при этом залогом хорошей работы ванны считают стабильное КО. Преимущественно «здоровые» ванны имеют малые вариации значений КО по сравнению с «больными», при примерно одинаковом среднем значении. Другая переменная – добавка AlF_3, является управляющим параметром. «Здоровые» ванны в среднем равномерно «потребляют» AlF_3. Потребление AlF_3 «больными» ваннами менее равномерно, поскольку он расходуется не только на поддержание КО. Следствие этого – следующая ситуация. Среди значимых для классификации переменных имеются такие, что их средние значения, вычисленные по всему множеству соответствующих данных в каждом классе, близки, а среднеквадратический разброс значений переменных в классе «больных» электролизеров в несколько раз превосходит разницу между средними значениями в классах. В итоге такие переменные оказываются неинформативными для классификации при использовании традиционного подхода решения задач РО, в то время как с точки зрения технологии они значимые.

В связи с этим предлагается следующий подход, основанный на исследовании ближайшей к моменту тестирования «истории» ванны. Реализации технологических переменных $\left\{x^{(z_m)}{}_{(j)} = (x^{(z_m)}(1)_{(j)},...,x^{(z_m)}(n)_{(j)})\right\}$, $j=\overline{1,l}; m=\overline{1,k}$, содержащихся в обучающей выборке, можно считать наблюдениями случайных процессов, поскольку измерения содержат случайные ошибки разной природы. Физическая природа исследуемых технологических переменных позволяет высказать гипотезу об эргодичности соответствующих им случайных процессов на продолжительных временных отрезках. (В работе рассматривались такие переменные: $x^{(z_m)}{}_{(1)}$ – уровень металла (UrMet), $x^{(z_m)}{}_{(2)}$ – криолитовое отношение (КО), $x^{(z_m)}{}_{(3)}$ – рабочее напряжение электролизера (U), $x^{(z_m)}{}_{(4)}$ – содержание CaF_2 в электролите, $x^{(z_m)}{}_{(5)}$ – добавка AlF_3, $x^{(z_m)}{}_{(6)}$ – вылитый металл (Met)). Таким образом, предлагается анализировать не *l*-мерное пространство переменных, содержащее $n \cdot k$ различных точек, а множество (мощности $l \cdot k$) случайных процессов длинной *n*.

Для описания случайных процессов традиционно используют следующие числовые характеристики: автокорреляционная функция, математическое ожидание и среднеквадратичное отклонение.

Для эргодического процесса для вычисления оценки математического ожидания *Mx* по наблюдениям процесса $\{x(1), x(2),..., x(n)\}$ используется формула [6]

$$\hat{M}x = \frac{1}{n}\sum_{i=1}^{n} x(i). \tag{1}$$

Выборочная несмещенная оценка среднеквадратического отклонения σ [6]:

$$\hat{\sigma} = \sqrt{\hat{D}x} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x(i)-\hat{M}x)^2}. \tag{2}$$

Автокорреляционная функция K_{xx} характеризует внутреннюю структуру случайного процесса $x(t)$, которую не отражают математическое ожидание и дисперсия. Она отражает степень зависимости между сечениями случайной функции, которые относятся к различным моментам времени. Для эргодических процессов автокорреляционная функция вычисляется по следующей формуле [6]:

$$\hat{K}_{xx}(\tau) = \frac{1}{n-\tau}\sum_{i=1}^{n-\tau}(x(i)-\hat{M}x)(x(i+\tau)-\hat{M}x), \quad \tau=\overline{1, n-lag}, \tag{3}$$

где τ – разница между моментами времени, lag – так называемая задержка, обусловленная особенностью вычислительной формулы (1.3) для K_{xx}. Заметим, что $\hat{K}_{xx}(0) = \hat{D}x$.

Множество (мощности $l \cdot k$) случайных процессов длиной *n* можно разбить на $2 \cdot l$ подмножеств, состоящих из реализации каждой переменной отдельно для обоих классов. То есть каждое подмножество содержит элементы вида либо $(x^{(+1)}(1)_{(j)},..., x^{(+1)}(n)_{(j)})$, либо $(x^{(-1)}(1)_{(j)},..., x^{(-1)}(n)_{(j)})$. Введем обозначения: $X^{(+1)}{}_{(j)} = \left\{(x^{(z_m)}(1)_{(j)},..., x^{(z_m)}(n)_{(j)})\right\}_{z_m=+1} = \left\{(x^{(z_m)}{}_{(j)})\right\}_{z_m=+1}$ – множество, состоящее из реализаций *j*-й переменной для «здоровых» ванн ($\left|X^{(+1)}{}_{(j)}\right| = k1$); $X^{(-1)}{}_{(j)} = \left\{(x^{(z_m)}(1)_{(j)},..., x^{(z_m)}(n)_{(j)})\right\}_{z_m=-1} = \left\{(x^{(z_m)}{}_{(j)})\right\}_{z_m=-1}$ – множество, состоящее из реализаций j-й переменной для «больных» ванн ($\left|X^{(-1)}{}_{(j)}\right| = k2$). Каждую пару множеств $X^{(+1)}{}_{(j)}$ и $X^{(-1)}{}_{(j)}$ можно интерпретировать как набор эталонных «здоровых» и «больных» реализаций *j*-й переменной соответственно. Будем использовать эти эталонные реализации для вычисления средних эталонных числовых характеристик $\hat{M}x$, $\hat{\sigma}$, $\hat{K}_{xx}$ для каждой *j*-й переменной. Средние эталонные характеристики для переменных «здоровых» ванн:

$$\hat{M}x^{(+1)}{}_{(j)} = \frac{1}{k1} \sum_{x^{(z_m)}{}_{(j)} \in X^{(+1)}(j)} \hat{M}x^{(z_m)}{}_{(j)},\ j = \overline{1,l}, \tag{4}$$

$$\hat{\sigma}^{(+1)}{}_{(j)} = \sqrt{\frac{1}{k1} \sum_{x^{(z_m)}{}_{(j)} \in X^{(+1)}(j)} \hat{D}x^{(z_m)}{}_{(j)}} = \sqrt{\frac{1}{k1} \sum_{x^{(z_m)}{}_{(j)} \in X^{(+1)}(j)} \hat{K}_{xx}{}^{(z_m)}(0)_{(j)}},\ j = \overline{1,l}, \tag{5}$$

$$\hat{K}_{xx}{}^{(+1)}(\tau)_{(j)} = \frac{1}{k1} \sum_{x^{(z_m)}{}_{(j)} \in X^{(+1)}(j)} \hat{K}_{xx}{}^{(z_m)}(\tau)_{(j)},\ \tau = \overline{0, n-lag},\ \overline{j = 1, l}, \tag{6}$$

где $\hat{M}x^{(z_m)}{}_{(j)}$ – математическое ожидание j-й переменной ванны z_m из множества $X^{(+1)}{}_{(j)}$, $\hat{D}x^{(z_m)}{}_{(j)} = (\sigma^{(+1)}{}_{(j)})^2$ – дисперсия j-й переменной ванны z_m из множества $X^{(+1)}{}_{(j)}$, $K_{xx}{}^{(+1)}(\tau)_{(j)}$ – значение корреляционной функции в точке τ j-й переменной ванны z_m из множества $X^{(+1)}{}_{(j)}$.

Средние эталонные характеристики для переменных «больных» ванн:

$$\hat{M}x^{(-1)}{}_{(j)} = \frac{1}{k2} \sum_{x^{(z_m)}{}_{(j)} \in X^{(-1)}(j)} \hat{M}x^{(z_m)}{}_{(j)},\ j = \overline{1,l}, \tag{7}$$

$$\hat{\sigma}^{(-1)}{}_{(j)} = \sqrt{\frac{1}{k2} \sum_{x^{(z_m)}{}_{(j)} \in X^{(-1)}(j)} \hat{D}x^{(z_m)}{}_{(j)}} = \sqrt{\frac{1}{k2} \sum_{x^{(z_m)}{}_{(j)} \in X^{(-1)}(j)} \hat{K}_{xx}{}^{(z_m)}(0)_{(j)}},\ j = \overline{1,l}, \tag{8}$$

$$\hat{K}_{xx}{}^{(-1)}(\tau)_{(j)} = \frac{1}{k2} \sum_{x^{(z_m)}{}_{(j)} \in X^{(-1)}(j)} \hat{K}_{xx}{}^{(z_m)}(\tau)_{(j)},\ \tau = \overline{0, n-lag},\ \overline{j = 1, l}, \tag{9}$$

где $\hat{M}x^{(z_m)}{}_{(j)}$ – математическое ожидание j-й переменной ванны z_m из множества $X^{(-1)}{}_{(j)}$, $\hat{D}x^{(z_m)}{}_{(j)} = (\hat{\sigma}^{(-1)}{}_{(j)})^2$ – дисперсия j-й переменной ванны z_m из множества $X^{(-1)}{}_{(j)}$, $K_{xx}{}^{(-1)}(\tau)_{(j)}$ – значение корреляционной функции в точке τ j-й переменной ванны z_m из множества $X^{(-1)}{}_{(j)}$.

По наблюдениям переменных эталонных ванн можно вычислить так называемые ожидаемые значения каждой j-й переменной на периоде длиной n для «здоровых» и «больных» ванн. Это будут средние арифметические следующего вида соответственно:

$$\overline{x}^{(+1)}(i)_{(j)} = \frac{1}{k1} \sum_{x^{(z_m)}{}_{(j)} \in X^{(+1)}(j)} x^{(z_m)}(i)_{(j)},\ i = \overline{1,n},\ j = \overline{1,l}, \tag{10}$$

$$\overline{x}^{(-1)}(i)_{(j)} = \frac{1}{k2} \sum_{x^{(z_m)}{}_{(j)} \in X^{(-1)}(j)} x^{(z_m)}(i)_{(j)},\ i = \overline{1,n},\ j = \overline{1,l}. \tag{11}$$

Теперь задачу классификации объекта в $n+1$ момент времени, имея набор реализаций V его переменных на предыдущем временном отрезке длиной n, можно решать следующим образом. Для каждой переменной объекта (электролизной ванны) находят числовые характеристики – математическое ожидание, среднеквадратическое отклонение, корреляционную функцию, по формулам (1)-(3) соответственно. Затем вычисляют отклонения этих характеристик от аналогичных эталонных характеристик. Отклонения от «здоровых» эталонов вычисляют по формулам:

$$\Delta\hat{M}x^{(+1)}{}_{(j)} = \hat{M}x^{(+1)}{}_{(j)} - \hat{M}x_{(j)},\ j = \overline{1,l}, \tag{12}$$

$$\Delta\hat{\sigma}^{(+1)}{}_{(j)} = \hat{\sigma}^{(+1)}{}_{(j)} - \hat{\sigma}_{(j)},\ j = \overline{1,l}, \tag{13}$$

$$\Delta\hat{K}_{xx}{}^{(+1)}{}_{(j)} = \sqrt{\frac{1}{n-lag+1} \sum_{\tau=0}^{n-lag} (\hat{K}_{xx}{}^{(+1)}(\tau)_{(j)} - \hat{K}_{xx}(\tau)_{(j)})^2},\ j = \overline{1,l}. \tag{14}$$

$$Md^{(+1)}{}_{(j)} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\overline{x}^{(+1)}(i)_{(j)} - x(i)_{(j)})^2},\ j = \overline{1,l}, \tag{15}$$

где $\hat{M}x_{(j)}$ – выборочное математическое ожидание j-й переменной тестируемой ванны, $\hat{\sigma}_{(j)}$ – выборочное среднеквадратичное отклонение j-й переменной тестируемой ванны, $\hat{K}_{xx}(\tau)_{(j)}$ – выборочное значение корреляционной функции j-й переменной тестируемой ванны в точке τ. Величина (15) есть среднее уклонение j-й переменной на периоде длиной n от ожидаемого «здорового» значения.

Отклонения от «больных» эталонов вычисляются по формулам:

$$\Delta\hat{M}x^{(-1)}{}_{(j)} = \hat{M}x^{(-1)}{}_{(j)} - \hat{M}x_{(j)},\ j = \overline{1,l}, \tag{16}$$

$$\Delta\hat{\sigma}^{(-1)}{}_{(j)} = \hat{\sigma}^{(-1)}{}_{(j)} - \hat{\sigma}_{(j)},\ j = \overline{1,l}, \tag{17}$$

$$\Delta\hat{K}_{xx}{}^{(-1)}{}_{(j)} = \sqrt{\frac{1}{n-lag+1}\sum_{\tau=0}^{n-lag}(\hat{K}_{xx}{}^{(-1)}(\tau)_{(j)} - \hat{K}_{xx}(\tau)_{(j)})^2},\ j = \overline{1,l}. \tag{18}$$

$$Md^{(-1)}{}_{(j)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\bar{x}^{(-1)}(i)_{(j)} - x(i)_{(j)})^2},\ j = \overline{1,l}. \tag{19}$$

Величина (19) есть среднее уклонение j-й переменной на периоде длиной n от ожидаемого «больного» значения.

Сравнение соответствующих отклонений позволяет делать вывод о *тяготении* характеристики j-й переменной тестируемого объекта (электролизной ванны) к «больному» или «здоровому» эталону. Например, если $\Delta\hat{M}x^{(-1)}{}_{(j)} > \Delta\hat{M}x^{(+1)}{}_{(j)}$, то математическое ожидание j-й переменной ванны тяготеет к «здоровому» эталону; если $\Delta\hat{M}x^{(-1)}{}_{(j)} < \Delta\hat{M}x^{(+1)}{}_{(j)}$, то математическое ожидание j-й переменной ванны тяготеет к «больному» эталону. Аналогичный анализ имеет место для трех других характеристик: $\Delta\hat{\sigma}^{(-1)}{}_{(j)}$ и $\Delta\hat{\sigma}^{(+1)}{}_{(j)}$, $\Delta\hat{K}_{xx}{}^{(-1)}{}_{(j)}$ и $\Delta\hat{K}_{xx}{}^{(+1)}{}_{(j)}$, $Md^{(-1)}{}_{(j)}$ и $Md^{(+1)}{}_{(j)}$.

Определив эталон, к которому тяготеет числовая характеристика, с помощью следующих коэффициентов определяется степень тяготения числовой характеристики к этому эталону:

$$k3_{(j)} = \frac{\Delta\hat{M}x^{(-1)}{}_{(j)}}{\Delta\hat{M}x^{(+1)}{}_{(j)}},\ j = \overline{1,l}\text{, если } \Delta\hat{M}x^{(-1)}{}_{(j)} > \Delta\hat{M}x^{(+1)}{}_{(j)}, \tag{20a}$$

$$k3_{(j)} = \frac{\Delta\hat{M}x^{(+1)}{}_{(j)}}{\Delta\hat{M}x^{(-1)}{}_{(j)}},\ j = \overline{1,l}\text{, если } \Delta\hat{M}x^{(-1)}{}_{(j)} > \Delta\hat{M}x^{(+1)}{}_{(j)}, \tag{20b}$$

$$k4_{(j)} = \frac{\Delta\hat{\sigma}x^{(-1)}{}_{(j)}}{\Delta\hat{\sigma}x^{(+1)}{}_{(j)}},\ j = \overline{1,l}\text{, если } \Delta\hat{\sigma}x^{(-1)}{}_{(j)} > \Delta\hat{\sigma}x^{(+1)}{}_{(j)}, \tag{21a}$$

$$k4_{(j)} = \frac{\Delta\hat{\sigma}x^{(+1)}{}_{(j)}}{\Delta\hat{\sigma}x^{(-1)}{}_{(j)}},\ j = \overline{1,l}\text{, если } \Delta\hat{\sigma}x^{(+1)}{}_{(j)} > \Delta\hat{\sigma}x^{(-1)}{}_{(j)}, \tag{21b}$$

$$k1_{(j)} = \frac{\Delta\hat{K}_{xx}{}^{(-1)}{}_{(j)}}{\Delta\hat{K}_{xx}{}^{(+1)}{}_{(j)}},\ j = \overline{1,l}\text{, если } \Delta\hat{K}_{xx}{}^{(-1)}{}_{(j)} > \Delta\hat{K}_{xx}{}^{(+1)}{}_{(j)}, \tag{22a}$$

$$k1_{(j)} = \frac{\Delta\hat{K}_{xx}{}^{(+1)}{}_{(j)}}{\Delta\hat{K}_{xx}{}^{(-1)}{}_{(j)}},\ j = \overline{1,l}\text{, если } \Delta\hat{K}_{xx}{}^{(+1)}{}_{(j)} > \Delta\hat{K}_{xx}{}^{(-1)}{}_{(j)}, \tag{22b}$$

$$k2_{(j)} = \frac{Md^{(-1)}{}_{(j)}}{Md^{(+1)}{}_{(j)}},\ j = \overline{1,l}\text{, если } Md^{(-1)}{}_{(j)} > Md^{(+1)}{}_{(j)}, \tag{23a}$$

$$k2_{(j)} = \frac{Md^{(+1)}{}_{(j)}}{Md^{(-1)}{}_{(j)}},\ j = \overline{1,l}\text{, если } Md^{(+1)}{}_{(j)} > Md^{(-1)}{}_{(j)}. \tag{23b}$$

Результаты классификации можно представить в таблице следующего вида:

	$n = 21$, $lag = 5$	K_{xx}	K1	Md	K2	M_x	K3	D_x	K4	Res
Ванна	$x_{(1)}$ = UrMet	+1	6,7	-1	1,5	-1	1,5	+1	2,64	-1
	$x_{(2)}$ = КО	-1	2,0	-1	1,2	-1	4,8	-1	3,6	-1
	$x_{(3)}$ = U	+1	1,7	+1	1,1	+1	34	-1	1,5	+1
	$x_{(4)}$ = CaF2	+1	1,1	-1	1,3	-1	2,2	-1	7,8	-1
	$x_{(5)}$ = AlF3	-1	1,2	+1	1,1	+1	1,1	-1	1,2	-1
	$x_{(6)}$ = Met	-1	1,3	-1	1,3	-1	4,9	-1	1,1	-1
	Res									-1

Крайний правый столбец этой таблицы содержит результат классификации по каждой из шести рассматриваемых переменных в целом, а нижняя клетка – итоговый результат классификации тестируемого объекта.

Если классификация по всем четырем числовым характеристикам для всех переменных совпадает, тогда решение о классификации каждой переменной и объекта в целом выносится однозначным образом. Такого рода «единогласие» свидетельствует о сильном тяготении объекта к определенному классу.

Но возможны ситуации, когда для тестируемого объекта классификация по числовым характеристикам переменных (как, например, по компонентам $x_{(1)}$, $x_{(3)}$, $x_{(4)}$, $x_{(5)}$ в приведенной таблице) и/или по переменным в целом не «единогласная». Тогда формализация правил принятия решения определяется предметной областью задачи и физической природой переменных, вследствие чего числовым характеристикам присваиваются разные веса. Определяющими являются коэффициенты (20-23). Так, уже упоминалось, что залог хорошей работы электролизной ванны – стабильное КО ($x_{(2)}$). Поэтому первоочередную важность при вынесении решения по КО имеют классификация по корреляционной функции и дисперсии и соответствующие коэффициенты. Аналогично для $x_{(5)}$ = AlF3. Классификация по переменным $x_{(1)}$, $x_{(3)}$, $x_{(4)}$ определяется, прежде всего, классификацией по математическому ожиданию и среднему уклонению от ожидаемого значения.

«Не единогласная» классификация и значения коэффициентов представляют богатую информацию о тестируемом объекте в целом и его отдельных составляющих. Приведем некоторые примеры. Учитывая физическую суть переменной, проанализируем возможные результаты классификации по численным характеристикам переменной $x_{(6)}$ – вылитый металл. Для нее наиболее значимо «выборочное математическое ожидание», поскольку оно отражает в среднем количество выливаемого металла за период между наблюдениями. В то же время здесь нельзя без внимания оставлять характеристики, которые дают качественное представление о течении процесса: $\hat{K}_{xx}$ и $\hat{\sigma}$. Для «здоровых» ванн существует корреляция между классификацией по $\hat{M}x$ и $\hat{K}_{xx}$. Действительно, не может быть так, что ванна имеет высокую производительность (то есть согласно классификации по $\hat{M}x$ тяготеет к «здоровому» эталону), при этом корреляционная функция вылитого металла соответствует режиму слива «больных» ванн. Такое сочетание скорее характеризует ванну как «больную» и указывает на то, что с ванны единовременно было слито очень много металла (что и дает высокое значение $\hat{M}x$ и положительную классификацию по $\hat{M}x$). Но, если ванна по средней выливке ($\hat{M}x$) тяготеет к «больным», режим ее слива (что отражает $\hat{K}_{xx}$) вполне может тяготеть к режиму слива «здоровых» ванн. В этом случае «больную» ванну от «здоровых» как раз и отличает количество сливаемого металла. Исходя из этих соображений формулируются правила принятия решения по этой переменной.

Если имеет место «не единогласная» классификация по парным числовым характеристикам, $\hat{K}_{xx}$ и $\hat{\sigma}x$ или Md и $\hat{M}x$ какой-либо переменной свидетельствуют об имевших место всплесках значений переменной тестируемой ванны на рассматриваемом временном отрезке. Близкие к «единице» значения коэффициентов говорят об одинаковом удалении значения характеристики от обоих эталонов.

Решение о классификации объекта в целом выносится большинством голосов. Если решение выносится «не единогласно», то «выпадающие» переменные следует проанализировать отдельным образом. Если количество голосов в пользу того и другого решения равное, тогда объект (электролизер) нельзя с достаточной долей уверенности отнести к одному из классов. Его следует относить к промежуточной категории (категории тревожных). А изучение текущих результатов классификации по каждой переменной и «истории» результатов классификации поможет определить, в направлении какого класса «двигается» данный объект.

Суммируя сказанное, отметим, что данный подход позволяет проводить не только «глобальную» классификацию (в данном случае режима работы алюминиевого электролизера), но и разносторонний анализ отдельных переменных объекта. Такая возможность представляет особую ценность для технологических динамических объектов, поскольку позволяет давать рекомендации по обслуживанию на ближайшее будущее.

Результаты применения данного алгоритма на реальных данных таковы. Из 15 предложенных для тестирования «здоровых» ванн 13 были отнесены к категории «здоровых», 2 ванны – к категории «тревожных». Из 12 протестированных «больных» ванн 9 были отнесены к категории «больных», 3 ванны – к категории «тревожных». Ни одна ванна не была классифицирована ошибочно (то есть отнесена к противоположному классу). Отметим, что по большей части решение о классификации «здоровых» ванн выносилось единогласно, в то время как в результатах классификации по различным переменным и их числовым характеристикам у «больных» ванн наблюдается разногласие. Поэтому подробный анализ результатов классификации ванны по каждой переменной, а также «истории» результатов классификации – полезны с точки зрения выявления особенностей ванны, определения средств «лечения» ванны или предотвращения грозящего перехода из класса «здоровых» в класс «больных».

Если рассматриваемые объекты (и соответствующие им процессы) нестационарны, тогда естествен периодический выбор новых эталонных объектов и пересчет эталонных характеристик.

Суммируя сказанное, отметим, что данный подход позволяет решать *задачу классификации* объектов (в данном случае режимов работы алюминиевых электролизеров) по периодическим (еженедельным) наблюдениям их значимых переменных. Результатом классификации является отнесение объекта (режима ванны) либо к одному из классов. Если определенного вывода сделать не удается, объект относится к промежуточной категории (категории «тревожных»). По итогам тестирования можно говорить о высоком качестве классификационных свойств алгоритма.

Изучение текущих результатов классификации по каждой переменной и «истории» результатов классификации помогает определить, в сторону какого класса «двигается» данный объект. Такая возможность представляет особую ценность для технологических динамических объектов, поскольку позволяет давать рекомендации по их обслуживанию на ближайшее будущее.

Автор выражает признательность ОАО «Красноярский алюминиевый завод» и ИТЦ «РУСАЛ» за предоставленную информацию.

СПИСОК ЛИТЕРАТУРЫ

1. Rodnov O.O., Poliakov P.V., Berezin A.I., Stont P.D., Mezhubovsky I.V. Estimation of a technological condition of the Aluminium Reduction Cells on the basis of its daily energy balance. Light Metals 2003. – p.457-462.
2. Antoulas A. C. Mathematical System Theory. – Springer-Verlag, Berlin, 1991.
3. Söderström T., and Stoica, P. System identification. – London, Prentice-Hall, 1989.
4. Бендат Дж. Применение корреляционного и спектрального анализа / Дж.Бендат, А.Пирсол. – М.: Мир, 1983.
5. Бессонов А.А. Методы и средства идентификации динамических объектов / А.А.Бессонов, Ю.В.Загашвили, А.С.Маркелов. – Л.: Энергоатомиздат, 1989.
6. Пугачев В.С. Теория случайных функций / В.С. Пугачев. – М.: Физматгиз, 1960.

ON OBJECTS CLASSIFICATION USING REALIZATION OF THEIR CHARACTERICTIC ALONG PREVIOUS PERIOD OF TIME

E.S. Kirik

The paper deals with the new approach to the solution of the pattern recognition problem. Proposed method is based on the random function theory. It is suggested to classify objects using realizations of their characteristics along previous period of time. Realization of this method is introduced on example of formal classification of the technological condition of the aluminium reduction cells.