ОЦЕНКА ИНФОРМАЦИОННОЙ ЦЕННОСТИ КОНТРОЛИРУЕМЫХ ПОКАЗАТЕЛЕЙ СОСТОЯНИЯ ОБЪЕКТА

канд. техн. наук, доц. Ю.В. Иванчихин, канд. техн. наук, ст. преп. Д.А. Сагайдачный, канд. техн. наук, доц. Я.В. Святкин, асп. Р.О. Корсун, Национальный технический университет "Харьковский политехнический институт", г. Харьков

Идентификация состояния объекта – важнейший элемент процесса его функционирования, во многом определяющий его успех. Парадигма "комплекс контролируемых параметров – состояние объекта" определяет совокупность современных концептуальных установок идентификации этого состояния. К настоящему моменту сформулированы общенаучные идеи и методы решения задачи идентификации систем, функционирующих в условиях неопределенности. Целью дальнейших исследований является развитие общей теории в направлении компьютеризации процесса идентификации состояния объекта исследования с использованием современного математического и программного обеспечения. В соответствии с этим особую важность приобретает задача отбора совокупности показателей, контроль и анализ значений которых целесообразно использовать при идентификации состояния объекта. Одним из традиционных критериев ценности показателей в этом отношении является информационная мера Кульбака-Лейблера [1], определяющая "расстояние" между двумя вероятностными распределениями.

Пусть контролируемый показатель X является случайной величиной с плотностью распределения $f\left(X/H_0\right)$, если объект находится в состоянии H_0, и с плотностью $f\left(X/H_1\right)$, если он находится в состоянии H_1. Тогда, в соответствии с [1], "расстояние" между этими распределениями вычисляется по формуле

$$\eta\left(f\left(X/H_0\right), f\left(X/H_1\right)\right) = \int_{-\infty}^{\infty} f\left(X/H_0\right) \ln \frac{f\left(X/H_0\right)}{f\left(X/H_1\right)} dx . \qquad (1)$$

Чем больше это "расстояние", тем выше информативность параметра x.

Мера Кульбака считается надежным инструментом для адекватной оценки информативности показателей. Вместе с этим, следует отметить существенные недостатки этого подхода к оценке «расстояния» между распределениями случайных величин. Во-первых, соотношение (1), используемое для расчета меры Кульбака, во многих практических случаях сводится к интегралу, не вычисляемому в квадратурах. При этом мера может быть оценена только численно, что неудобно, так как затрудняет возможность анализа зависимости "расстояния" от численных значений параметров распределений. Во-вторых, из (1) следует, что мера Кульбака, равная нулю, если $f\left(X/H_0\right) = f\left(X/H_1\right)$, может принимать произвольно большое, не ограниченное сверху, значение в противном случае, что не информативно. Наконец, самый серьезный недостаток состоит в том, что критерий асимметричен, то есть его численное значение будет разным в зависимости от того, каким образом выражения для сравниваемых распределений входят в конструкцию (1). При этом верхняя и нижняя оценки значений критерия (1) могут существенно различаться. Это свойство делает практически неприемлемым использование меры Кульбака в задаче выбора наиболее информативных показателей при оценке состояния объекта диагностики. Дело в том, что нижняя оценка объективно худшего из пары сравниваемых показателей может оказаться больше по величине критерия различимости, нежели верхняя оценка более информативного показателя. Перечисленные обстоятельства делают целесообразным отыскание другого, симметричного соотношения для расчета меры различения распределений с численным значением, лежащим в конечном интервале, размеры которого инвариантны относительно типа распределений.

Возможный альтернативный вариант построения такого критерия различимости имеет вид:

$$\eta = \hat{J}\left(f\left(X/H_0\right), f\left(X/H_1\right)\right) = 1 - \int_{-\infty}^{\infty}\left(f\left(X/H_0\right)\cdot f\left(X/H_1\right)\right)^{\frac{1}{2}} dx \tag{2}$$

В частном случае различения двух гауссовых распределений значение $\hat{J}\left(f\left(X/H_0\right), f\left(X/H_1\right)\right)$ легко вычисляется:

$$\eta = \hat{J}\left(f\left(X/H_0\right), f\left(X/H_1\right)\right) == 1 - \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma_0\sigma_1}}\exp\left\{\frac{1}{4}\left[\frac{\left(X-m_0\right)^2}{\sigma_0^2} + \frac{\left(X-m_1\right)^2}{\sigma_1^2}\right]\right\} =$$

$$= 1 - \sqrt{\frac{2\sigma_0\sigma_1}{\sigma_0^2+\sigma_1^2}}\cdot\exp\left\{-\frac{\left(m_0-m_1\right)^2}{4\left(\sigma_0^2+\sigma_1^2\right)}\right\}.$$

Дискретный аналог соотношения (2) имеет следующий вид:

$$J_0 = 1 - \sum_{k=1}^{S}\left[P\left(X_k/H_1\right)\cdot P\left(X_k/H_2\right)\right]^{\frac{1}{2}}. \tag{3}$$

Проведем анализ этого соотношения.

1. Формула (3) дает результат, не зависящий от порядка вхождения компонентов $P\left(X_k / H_1 \right)$ и $P\left(X_k / H_2 \right)$ в расчетное соотношение, так как

$$P\left(X_k / H_1 \right) \cdot P\left(X_k / H_2 \right) = P\left(X_k / H_2 \right) \cdot P\left(X_k / H_1 \right).$$

2. Результаты расчета по формуле (4) находятся в интервале [0; 1].

При этом, $J_0 = 0$, если $P\left(X_k / H_1 \right) = P\left(X_k / H_2 \right)$, $k = 1, 2, ..., s$, так как

$$\left[P\left(X_k / H_1 \right) \cdot P\left(X_k / H_2 \right) \right]^{\frac{1}{2}} = P\left(X_k / H_1 \right) = P\left(X_k / H_2 \right),$$

$$\sum_{k=1}^{s} \left[P\left(X_k / H_1 \right) \cdot P\left(X_k / H_2 \right) \right]^{\frac{1}{2}} = \sum_{k=1}^{s} P\left(X_k / H_1 \right) = \sum_{k=1}^{s} P\left(X_k / H_2 \right) = 1.$$

С другой стороны, ясно, что $J_0 = 1$, если $P\left(X_k / H_1 \right) \cdot P\left(X_k / H_2 \right) = 0$, $k = 1, 2, ..., s$, что и требовалось.

В работе рассмотрена задача формирования метода оценки информативности контролируемых показателей состояния объекта. В результате исследования получен метод, улучшающий известный и широко используемый метод, реализующий расчет меры Кульбака. Предложенный метод основан на специальной вычислительной процедуре статистической обработки результатов измерения набора контролируемых показателей.

Список литературі: 1. *Кульбак С.* Теория информации и статистика. М.: Наука, 1967. – 408 с.