УДК 621.391

АНАЛИЗ КОЛИЧЕСТВА ИНФОРМАЦИИ, ТРЕБУЕМОГО ДЛЯ КЛАССИФИКАЦИИ ЛОКАЦИОННЫХ ДАННЫХ С ЗАДАННЫМ СРЕДНИМ РИСКОМ

к.т.н. В.Е. Саваневич, А.В. Пугач
(представил д.т.н., проф. Д.В. Голкин)

В статье изложена методика определения общей и частных средних взаимных информаций между состояниями и решениями, минимально необходимых для стабилизации среднего риска классификации на заданном уровне. Подчеркивается необходимость обеспечения в $\boldsymbol{\varepsilon}$ *- энтропии пропорций между частными информациями, определенными в экстремальном распределении.*

Необходимость разработки локационных систем с заданным качеством принимаемых решений при наименьших затратах делает актуальной задачу синтеза статистических алгоритмов минимальной сложности [1]. Для обеспечения заданного среднего риска $\mathbf{R}^*$ необходимо [1], чтобы средняя взаимная информация (СВИ) между пространствами состояний $\mathbf{X}$ и решений $\mathbf{U}$ была не меньше $\boldsymbol{\varepsilon}(\mathbf{R}^*)$ - энтропии $\mathbf{I}^*_{\mathbf{XU}}$ [2]. Естественно, количество информации в выборке также не должно быть меньше $\mathbf{I}^*_{\mathbf{XU}}$.

Решению задачи определения $\mathbf{R}^*$ - энтропии при классификации локационной информации посвящена данная работа.

Итак, имеют место $(\mathbf{M}+\mathbf{1})$ возможных вариантов (состояний) целевой обстановки (**М** вариантов траекторий и их отсутствие, отметка получена от одной из **М** целей или является ложной и т.д.), составляющих пространство состояний (ПС) $\mathbf{X}$. Каждому состоянию $\mathbf{x}$ с номером $\mathbf{i}$ соответствует априорная вероятность $\mathbf{P_x(x_i)>0}$. Кроме того, имеет место пространство решений (ПР) $\mathbf{U}$, включающее в себя $(\mathbf{M}+\mathbf{1})$ возможных решений **u**, в совокупности, по сути своей, повторяющих ПС.

В виде матрицы $(\mathbf{M}+\mathbf{1})\times(\mathbf{M}+\mathbf{1})$ задана функция потерь с элементами $\boldsymbol{\Pi}_{\mathbf{ij}}$, характеризующими потери, возникающие при принятии решения $\mathbf{u_j}$ в состоянии $\mathbf{x_i}$.

Необходимо найти экстремальные совместные вероятности состояний и решений $\mathbf{P_{xu}(x_i,u_j)}$, обеспечивающие минимум СВИ между ПС и ПР

$$I_{XU}=\sum_{i,j=0}^{M} P_{u/x}(u_j/x_i)P_x(x_i)\log\frac{P_{u/x}(u_j/x_i)}{P_u(u_j)} \qquad (1)$$

и собственно ε - энтропию $I^*_{XU}=\min I_{XU}$ при заданном среднем риске

$$R^*=\sum_{i,j=0}^{M}\Pi_{ij}P_{u/x}(u_j/x_i)P_x(x_i) \qquad (2)$$

и условии неизменности априорных вероятностей целевой обстановки

$$\sum_{j=0}^{M}P_{xu}(x_i,u_j)=P_x(x_i), \qquad (3)$$

где $P_{u/x}(u_j/x_i)$ - вероятность принятия решения u с номером $j\,(j=\overline{0,M})$ в целевой обстановке x с номером $i\,(i=\overline{0,M})$;

$P_u(u_j)=\sum_{i=0}^{M}P_{u/x}(u_j/x_i)P_x(x_i)$ - вероятность принятия решения u_j.

Искомые экстремальные распределения при использовании метода неопределенных множителей Лагранжа представим в виде [4]:

$$P_{xu}(x_i,u_j)=P_x(x_i)P_u(u_j)\exp(-\gamma_i-\beta\Pi_{ij}), \qquad (4)$$

где γ_i,β – неопределённые множители Лагранжа.

Значения неопределенных множителей могут быть найдены из уравнений типа:

$$\sum_{i=0}^{M}P_x(x_i)Q_i\exp(-\beta\Pi_{ij})=1; \qquad (5)$$

$$\sum_{j=0}^{M}P_u(u_j)\exp(-\beta\Pi_{ij})=\exp(\gamma_i), \qquad (6)$$

где

$$Q_i=P_x(x_i)\exp(-\gamma_i) \qquad (7)$$

Значения Q_i для $i=\overline{0,M}$ определяются (при существовании решения) из системы линейных уравнений (5):

$$\begin{cases} Q_0e^{-\beta\Pi_{00}}+Q_1e^{-\beta\Pi_{10}}+\ldots+Q_Me^{-\beta\Pi_{M0}}=1; \\ \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\ Q_0e^{-\beta\Pi_{0M}}+Q_1e^{-\beta\Pi_{1M}}+\ldots+Q_Me^{-\beta\Pi_{MM}}=1, \end{cases}$$

или в матричной форме

$$BQ=\overline{1},$$

где
$$\mathbf{B}=\begin{pmatrix} e^{-\beta\Pi_{00}} & \dots & e^{-\beta\Pi_{M0}} \\ \dots & \dots & \dots \\ e^{-\beta\Pi_{0M}} & \dots & e^{-\beta\Pi_{MM}} \end{pmatrix}; \tag{8}$$

$\mathbf{Q}^T=(\mathbf{Q}_0,\dots,\mathbf{Q}_M)$; $\bar{\mathbf{1}}$ - вектор-столбец, все элементы которого равны 1.

Тем самым
$$\mathbf{Q}=\mathbf{B}^{-1}\bar{\mathbf{1}} \;=\; \sum_{j=0}^{M} b_{xu}(x_i,u_j),$$

где $b_{xu}(x_i,u_j)$ - элементы матрицы $\mathbf{B}^{-1}$.

Далее используя понятие термодинамического потенциала $\Gamma=-H_X-\sum_{i=0}^{M}P_x(x_i)\ln Q_i$ в соответствии с известными тождествами $d\Gamma/d\beta=-R$; $\beta d\Gamma/d\beta-\Gamma=I$ можно вывести выражения для среднего риска и ε-энтропии как функций от неопределенного множителя β [4]:
$$R(\beta)=\sum_{i=0}^{M}P_x(x_i)\frac{\sum_{j=0}^{M}a_{xu}(x_i,u_j)}{\sum_{j=0}^{M}b_{xu}(x_i,u_j)}; \tag{9}$$
$$I(\beta)=H_X-\sum_{i=0}^{M}P_x(x_i)\left[\frac{\beta\sum_{j=0}^{M}a_{xu}(x_i,u_j)}{\sum_{j=0}^{M}b_{xu}(x_i,u_j)}-\ln\left(\sum_{j=0}^{M}b_{xu}(x_i,u_j)\right)\right], \tag{10}$$

где $H_X=-\sum_{i=0}^{M}P_x(x_i)\log P_x(x_i)$ - энтропия пространства состояний;

$$a_{xu}(x_i,u_j)=\frac{\partial b_{xu}(x_i,u_j)}{\partial\beta}.$$

При вычислении производной обратной матрицы достаточно использовать легко доказуемое тождество
$$\frac{d}{dx}\mathbf{B}^{-1}=\mathbf{B}^{-1}\left(\frac{d}{dx}\mathbf{B}\right)\mathbf{B}^{-1}. \tag{11}$$

Для нахождения β следует использовать выражение (9) в виде уравнения $R(\beta)=R^*$.

Таким образом, найдя элементы матриц $\mathbf{B}$, $\mathbf{B}^{-1}$ и $\dfrac{\mathbf{dB}^{-1}}{\mathbf{d\beta}}$, а затем и множитель β, можно найти ε-энтропию $\mathbf{I}^{*}_{\mathbf{XU}}$ классификатора локационной информации в соответствии с выражением (10).

При решении задачи обнаружения ($\mathbf{x = 0, 1}$) выражения для элементов матриц $\mathbf{B}^{-1}$, $\dfrac{\mathbf{dB}^{-1}}{\mathbf{d\beta}}$ существенно упрощаются:

$$\mathbf{B}^{-1} = \frac{1}{\mathbf{D}}\begin{pmatrix} \exp(-\beta\Pi_{11}) & -\exp(-\beta\Pi_{10}) \\ -\exp(-\beta\Pi_{01}) & \exp(-\beta\Pi_{00}) \end{pmatrix}; \qquad (12)$$

$$\frac{\mathbf{dB}^{-1}}{\mathbf{d\beta}} = \frac{1}{\mathbf{D}^2}\begin{pmatrix} -(\Pi_{11}\mathbf{D}+\mathbf{C})\exp(-\beta\Pi_{11}) & (\Pi_{10}\mathbf{D}+\mathbf{C})\exp(-\beta\Pi_{10}) \\ (\Pi_{01}\mathbf{D}+\mathbf{C})\exp(-\beta\Pi_{01}) & -(\Pi_{00}\mathbf{D}+\mathbf{C})\exp(-\beta\Pi_{00}) \end{pmatrix}, \qquad (13)$$

где $\mathbf{D} = \exp(-\beta(\Pi_{00}+\Pi_{11})) - \exp(-\beta(\Pi_{01}+\Pi_{10}))$;

$$\mathbf{C} = (\Pi_{01}+\Pi_{10})\exp(-\beta(\Pi_{01}+\Pi_{10})) - (\Pi_{00}+\Pi_{11})\exp(-\beta(\Pi_{00}+\Pi_{11})).$$

Выражение для среднего риска обнаружителя принимает вид

$$\mathbf{R} = \frac{1}{\mathbf{D}}\left[\mathbf{q}\frac{((\Pi_{01}\mathbf{D}+\mathbf{C})\exp(-\beta\Pi_{01}) - (\Pi_{11}\mathbf{D}+\mathbf{C})\exp(-\beta\Pi_{11}))}{(\exp(-\beta\Pi_{11}) - \exp(-\beta\Pi_{01}))} + \right.$$
$$\left. + \mathbf{p}\frac{((\Pi_{10}\mathbf{D}+\mathbf{C})\exp(-\beta\Pi_{10}) - (\Pi_{00}\mathbf{D}+\mathbf{C})\exp(-\beta\Pi_{00}))}{\mathbf{D}(\exp(-\beta\Pi_{00}) - \exp(-\beta\Pi_{10}))}\right], \qquad (14)$$

где $\mathbf{p}$, $\mathbf{q = 1 - p}$ - априорные вероятности наличия и отсутствия объекта.

Общий вид зависимости среднего риска обнаружителя от минимально необходимой средней взаимной информации ((9), (10)) $\mathbf{I}^{*}_{\mathbf{XU}}$ приведен на рис.1.

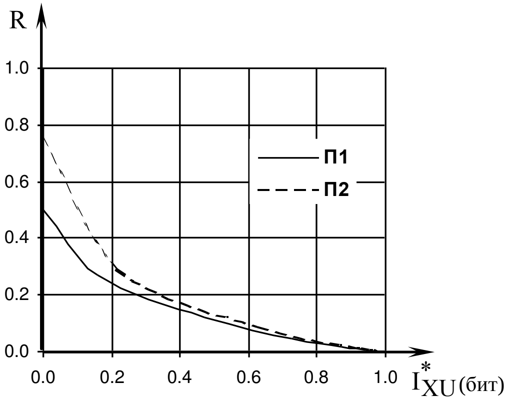

Рис.1. Зависимость среднего риска от средней информации между состояниями и наблюдениями

Кривые соответствуют вероятности наличия цели 0.5 и отличаются матрицами потерь: в первом случае платы за ошибочные решения равны $\Pi 1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, во втором же плата за пропуск цели выше, чем за ложную тревогу – $\Pi 2 = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$. Видно, что с ростом СВИ до максимального значения (соответствует энтропии ПС) средний риск уменьшается до нуля.

Обеспечение общей СВИ между ПС и ПР на уровне $\mathbf{I}^*_{\mathbf{XU}}$ является необходимым, но никак не достаточным условием стабилизации среднего риска на уровне $\mathbf{R}^*$. Дополнительно необходимо соблюдение качественного состава СВИ, определенного в экстремальном распределении. При этом недостаток хотя бы по одной из частных СВИ приводит к большему, чем требуется, риску $\mathbf{R}$ даже, иногда, при существенном избытке общей информации $\mathbf{I}_{\mathbf{XU}}$ (рис. 2). Кроме того, для обеспечения одного и того же среднего риска может быть «израсходовано» существенно различное количество информации. На рис. 3 приведены кривые, все точки которых соответствуют $\mathbf{R}^* = \mathbf{0.25}$ при четырех различных функциях потерь:

$$\mathbf{П1} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix};\ \mathbf{П2} = \begin{pmatrix} \mathbf{0} & \mathbf{4} \\ \mathbf{1} & \mathbf{0} \end{pmatrix};$$

$$\mathbf{П3} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{4} & \mathbf{0} \end{pmatrix};\ \mathbf{П4} = \begin{pmatrix} \mathbf{0} & \mathbf{4} \\ \mathbf{4} & \mathbf{0} \end{pmatrix}.$$

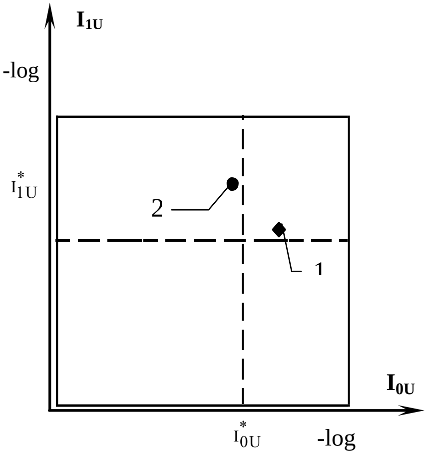

Рис. 2. Отображение ε-энтропии и вариантов формирования выборок в плоскости частных средних информаций

Видно, что при неблагоприятных ситуациях количество СВИ, необходимое для стабилизации среднего риска на уровне $\mathbf{R}^*$, может более чем вдвое превышать минимально необходимое.

Итак, для нахождения экстремальных вероятностей необходимо из определения $\mathbf{Q_i}$ (7) найти непосредственные значения множителей

$$\gamma_\mathbf{i} = \mathbf{ln}\,\mathbf{P_x}(\mathbf{x_i}) - \mathbf{ln}(\mathbf{Q_i})\,. \qquad (15)$$

Далее безусловные вероятности принятия решений определяются из системы линейных уравнений типа (6):

$$\mathbf{B}^\mathbf{T}\,\mathbf{P_u} = \mathbf{E}_\gamma \Rightarrow \mathbf{P_u} = \left(\mathbf{B}^\mathbf{T}\right)^{-1}\mathbf{E}_\gamma\,, \qquad (16)$$

где $\mathbf{P_u}$ - вектор-столбец безусловных вероятностей решений; $\mathbf{E}^\mathbf{T}_\gamma = \left(\mathbf{e}^{\gamma_0}, \ldots, \mathbf{e}^{\gamma_M}\right)$.

Затем совместные вероятности состояний и решений определяются вы-

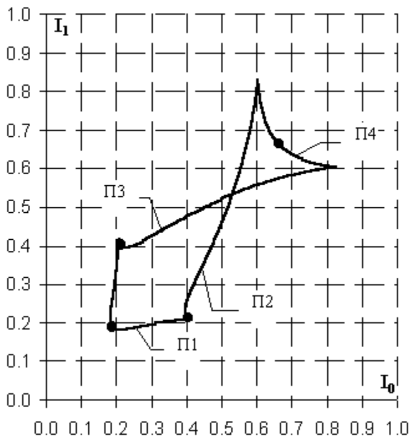

Рис. 3. Взаимные зависимости частных средних информаций, обеспечивающих средний риск 0.25, для обнаружителей с различными матрицами потерь (точки на кривых соответствуют минимально необходимой СВИ I^*_{XU})

ражением (4). В соответствии с определением условной вероятности события вероятности принятия решения $\mathbf{u_j}$ при состоянии $\mathbf{x_i}$ имеют вид

$$\mathbf{P_{u/x}(u_j / x_i) = P_{xu}(x_i, u_j) / P_x(x_i)}. \qquad (17)$$

В свою очередь, искомые частные $\boldsymbol{\varepsilon}$-энтропии находятся по определению

$$\mathbf{I_{xU} = \sum_{j=0}^{M} P_{u/x}(u_j / x_i) \log \frac{P_{u/x}(u_j / x_i)}{P_u(u_j)}}. \qquad (18)$$

Для обнаружителя качество обнаружения часто задают условными вероятностями правильного обнаружения (УВПО) $\mathbf{D^*}$ и ложной тревоги (УВЛТ) $\mathbf{F^*}$. При этом необходимые частные средние взаимные информации в соответствии с определением (1) находятся наиболее просто:

$$\begin{aligned} \mathbf{I^*_{0U}} &= \mathbf{\left(1 - F^*\right) \log \frac{1 - F^*}{q\left(1 - F^*\right) + p\left(1 - D^*\right)} + F^* \log \frac{F^*}{qF^* + pD^*}}; \\ \mathbf{I^*_{1U}} &= \mathbf{\left(1 - D^*\right) \log \frac{1 - D^*}{q\left(1 - F^*\right) + p\left(1 - D^*\right)} + D^* \log \frac{D^*}{qF^* + pD^*}}. \end{aligned} \qquad (19)$$

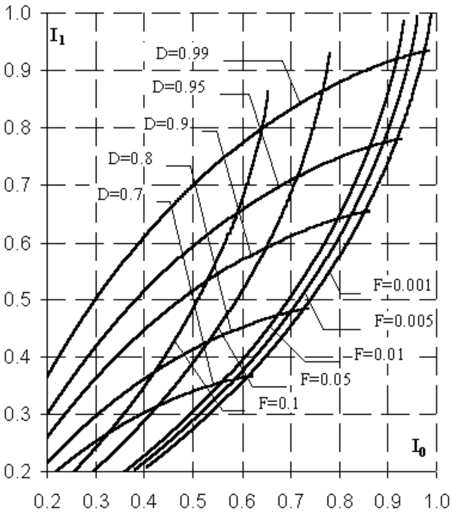

Рис.4. Линии «уровней» УВПО и УВЛТ в плоскости частных средних информаций

Максимальное возможное значение частных СВИ в данном случае может быть получено из выражений (19) путем предельного перехода $\mathbf{D^* \to 1}$, $\mathbf{F^* \to 0}$ и составляет $\mathbf{I_{0U} = -\log q}$, $\mathbf{I_{1U} = -\log p}$. В общем же случае связь между УВПО и УВЛТ, с одной стороны, и средним количеством информации о наличии и отсутствии цели, с другой, имеет вид, приведенный на рис. 4.

Тем самым, для обеспечения заданных УВПО $\mathbf{D^*}$ и УВЛТ $\mathbf{F^*}$ при заданной условной вероятности наличия цели $\mathbf{p}$, необходимо, чтобы выборка в плоскости $\mathbf{I_{0U} : I_{1U}}$, по крайней мере, попала в правый верхний квадрант относительно точки характеризующей $\boldsymbol{\varepsilon}$-энтропию (рис. 2). Например, проверяется простая гипотеза о наличии цели против простой альтернативы об ее отсутствии по бинарным сигналам. При этом $\mathbf{p = 0.5}$, $\mathbf{D = 0.9}$, $\mathbf{F = 0.1}$, $\mathbf{D^*} = 0.89$, $\mathbf{F^*} = 0.05$. Задано два варианта последовательного правила обнаружения объекта, которые схематически приведены на рис. 5.

Их средние частные информации о состоянии в виде точек 1 и 2 представлены на рис. 2. Видно, что вторая последовательность однозначно не удовлетворяет требованиям по качеству принимаемых решений. Вместе с тем, оба варианта имеют одинаковую среднюю информацию между пространствами состояний и наблюдений, составляющую 0.684.

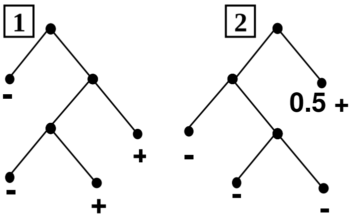

Рис. 5. Варианты последовательных правил обнаружителя

Итак, для нахождения минимально необходимой (для обеспечения среднего риска $\mathbf{R}^*$) средней взаимной информации между пространствами состояний и решений $\mathbf{I}^*_{\mathbf{XU}}$ в качественном составе, определенном частными $\mathbf{R}^*$-энтропиями $\mathbf{I}^*_{\mathbf{XU}}$ необходимо выполнить следующую последовательность операций: в соответствии с выражениями (8), (11) определить элементы матриц $\mathbf{B}$, $\mathbf{B}^{-1}$ и $\frac{\mathbf{dB}^{-1}}{\mathbf{d}\beta}$; найти неопределенный множитель β, исходя из заданного среднего риска $\mathbf{R}^*$ в соответствии с выражением (9) и множители $\gamma_{\mathbf{x}}$, используя выражение (15); затем рассчитать экстремальные значения условных и безусловных вероятностей принятия решений (16), (4), (17), после чего, используя выражения (1), (10), (18), необходимо определить общую и частные ε-энтропии.

В свою очередь, при выборе пространства наблюдений для классификации анализа общего количества информации содержащегося в выборке явно недостаточно. Имеет место необходимость контроля всего вектора частных информаций на предмет достижения его элементами значений, определенных экстремальным распределением.

ЛИТЕРАТУРА

1. *Саваневич В.Е. Информационный подход к синтезу статистических алгоритмов с минимальной сложностью // Системи обробки інформації. – Х. : НАНУ, ПАНМ, ХВУ. – 2000. – Вип. 3(9). – С. 123 - 128.*
2. *Колмогоров А.Н. Теория информации и теория алгоритмов. – М.: Наука, 1987. – 304 с.*
3. *Колесник В.Д., Полтырев Г.Ш. Курс теории информации. – М.: Наука, 1982. – 416 с.*
4. *Стратонович Р.Л. Теория информации. – М.: Сов. радио, 1975. – 424 с.*

Поступила 8.04.2002

Саваневич Вадим Евгеньевич*, канд. техн. наук, доцент, докторант ХВУ. В 1986 году окончил ХВУРЭ. Область научных интересов – обработка локационной информации,*

информметрия.

***Пугач Андрей Витальевич**, адъюнкт Харьковского военного университета. В 1997 году окончил ХВУ. Область научных интересов – обработка локационной информации.*