

# ОЦЕНКА ЭФФЕКТИВНОСТИ ИСПОЛЬЗОВАНИЯ СЕТЕВОГО СКАНИРУЮЩЕГО РОБОТА ПО РЕЗУЛЬТАТАМ ИСПЫТАНИЙ

В. П. ЗАБОЛОТСКИЙ<sup>1</sup>, Р. М. ЮСУПОВ<sup>2</sup>

Санкт-Петербургский институт информатики и автоматизации РАН

СПИИРАН, 14-я линия ВО, д. 39, Санкт-Петербург, 199178

<sup>1</sup><lai@iias.spb.su>, <sup>2</sup><yusupov@iias.spb.su>

---

УДК 519.8+681.3.06

Заболотский В. П., Юсупов Р. М. Оценка эффективности использования сетевого сканирующего робота по результатам испытаний // Труды СПИИРАН. Вып. 4. — СПб.: Наука, 2007.

**Аннотация.** Рассматривается применение метода индексов для оценивания по результатам испытаний эффективности использования сетевого сканирующего робота для поиска и сбора требуемой информации в информационной среде, включающей различные информационные массивы, которые хранят разнообразную информацию, имеющую различную ценность для лица, в интересах которого осуществляется поиск и сбор информации. — Библ. 5 назв.

UDC 519.8+681.3.06

Zabolotsky V. P., Yusupov R. M. Network Scanning-Bot Using Efficiency Estimation Upon Test Series Results // SPIIRAS Proceedings. Issue 4. — SPb.: Nauka, 2007.

**Abstract.** We consider an application of the indices method to estimating network scanning-bot effectiveness in tasks of searching relevant information. The search is executed in an information environment with different information arrays. These arrays contain information that have different relevance levels for the user. — Bibl. 5 items.

---

## 1. Введение

Концептуальные основы и методология оценивания эффективности использования сетевого сканирующего робота по результатам испытаний базируются на системном подходе, который позволяет всесторонне и комплексно решать поставленные задачи с учетом всех существенных свойств испытываемой системы и процесса её функционирования, а также всех существенных связей и взаимодействий системы с окружающей средой. При этом данный процесс рассматривается как операция (целенаправленный процесс), эффективность которой представляет собой комплексное свойство, основными составляющими которого являются результативность, ресурсоемкость и оперативность.

Для количественного оценивания эффективности использования сетевого сканирующего робота по результатам испытаний применим обобщенные (сводные) показатели или индексы [1-3].

**Индекс** – это показатель, характеризующий уровень или соотношение уровней какого-либо явления, в том числе и в тех случаях, когда составляющие уровней явления или его элементов непосредственно не сопоставимы.

Применяемые индексы подразделяются на абсолютные и относительные обобщенные (сводные) показатели или индексы.

К абсолютным индексам эффективности использования сетевого сканирующего робота, определенной по результатам испытаний, относятся обобщенные (сводные) показатели вида

$$h = \frac{k}{c}, \quad (1)$$

где

$h$  — индекс (показатель) эффективности использования сетевого сканирующего робота, определенной по результатам испытаний;

$k$  — индекс (показатель) результативности использования сетевого сканирующего робота, определенной по результатам испытаний;

$c$  — индекс (показатель) ресурсоемкости использования сетевого сканирующего робота, определенной по результатам испытаний.

В качестве индекса  $c$  ресурсоемкости использования сетевого сканирующего робота, определенной по результатам испытаний, можно выбрать затраты на использование сетевого сканирующего робота в денежном выражении, а базового индекса ресурсоемкости использования сетевого сканирующего робота — базовые (нормативные) затраты в том же выражении на ту же операцию. Кроме того, так как основными затратами, характеризующими ресурсоемкость информационного поиска в сети с помощью сканирующего робота, являются временные затраты на поиск, то в качестве индекса  $c$  ресурсоемкости использования сетевого сканирующего робота могут быть выбраны эти затраты.

К относительным индексам эффективности использования сетевого сканирующего робота, определенной по результатам испытаний, относятся обобщенные (сводные) показатели вида:

$$\delta h = \frac{h - h_0}{h_0} = \frac{h}{h_0} - 1, \quad (2)$$

где

$\delta h$  — индекс изменения эффективности использования сетевого сканирующего робота, определенной по результатам испытаний с учетом исследуемого способа использования сетевого сканирующего робота;

$h$  — индекс эффективности использования сетевого сканирующего робота, определенной по результатам испытаний с учетом исследуемого способа использования сетевого сканирующего робота;

$h_0$  — индекс эффективности использования сетевого сканирующего робота, определенной по результатам испытаний выбранного в качестве базового способа использования сетевого сканирующего робота (базовый индекс эффективности).

## 2. Постановка задачи

Рассмотрим достаточно общий случай применения индексного метода для оценивания эффективности использования сетевого сканирующего робота по результатам испытаний нескольких способов использования сетевого сканирующего робота для поиска и сбора требуемой информации в информационной среде, включающей различные информационные массивы, которые хранят разнообразную информацию, имеющую различную ценность для лица, в интересах которого осуществляется поиск и сбор информации.

Требуется по результатам испытаний оценить эффективность испытываемых способов использования сетевого сканирующего робота и, если это будет необходимо, выбрать среди них наиболее эффективный.

Для получения искомых оценок применим индексный метод.

Все индексы будем измерять в одной и той же Q-балльной шкале, в которой:

- $Q_{\min}$  – наименьшее значение данной шкалы;
- $Q_{\max}$  – наибольшее значение данной шкалы.

Это не накладывает существенных ограничений на применение метода, так как результаты измерений всегда можно привести к одной шкале. Значения используемых индексов при отсутствии необходимой статистики определяются экспертным путем с учетом ценности собранной информации для лица, в интересах которого осуществляется её поиск и сбор.

Введем следующие обозначения:

1.  $M$  — количество испытываемых способов использования сетевого сканирующего робота, подлежащих оцениванию на эффективность.
2.  $N_m$  — количество информационных массивов, подлежащих сканированию  $m$ -м испытываемым способом использования сетевого сканирующего робота,  
 $m = 1(1)M$ .
3.  $R_{mn}$  — количество тематических разделов в  $n$ -м информационном массиве, подлежащих сканированию  $m$ -м испытываемым способом использования сетевого сканирующего робота,  
 $m = 1(1)M$ ,  $n = 1(1)N_m$ .
4.  $S_{mnr}$  — количество информационных документов в  $r$ -м тематическом разделе  $n$ -го информационного массива, сканируемого  $m$ -м испытываемым способом использования сетевого сканирующего робота,  
 $m = 1(1)M$ ,  $n = 1(1)N_m$ ,  $r = 1(1)R_{mn}$ .
5.  $x_{mnr s}$  — индикаторная функция, задающая набор ( $s$ -х) информационных документов ( $r$ -го) тематического раздела ( $n$ -го) информационного массива, сканируемого ( $m$ -м) способом поиска, где

$$x_{mnr s} = \begin{cases} 1 - \text{если конкретный } s\text{-й информационный документ } r\text{-го} \\ \text{тематического раздела } n\text{-го информационного массива} \\ \text{обнаружен } m\text{-м способом поиска,} \\ 0 - \text{в противном случае,} \end{cases}$$

$$m = 1(1)M, \quad n = 1(1)N_m, \quad r = 1(1)R_{mn}, \quad s = 1(1)S_{mnr}.$$

6.  $v_{mnr s}$  — индекс ценности (важности) для лица, в интересах которого осуществляется поиск и сбор информации, сканирования  $s$ -го информационного документа  $r$ -го тематического раздела  $n$ -го информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ ,  $n = 1(1)N_m$ ,  $r = R_{mn}$ ,  $s = 1(1)S_{mnr}$ .

Система индексов, используемых для оценивания эффективности испытываемых способов использования сетевого сканирующего робота для поиска и сбора информации включает, три группы индексов:

- индексы эффективности испытываемых способов использования сетевого сканирующего робота,

- индексы результативности испытываемых способов использования сетевого сканирующего робота,
- индексы ресурсоемкости испытываемых способов использования сетевого сканирующего робота.

В группу индексов эффективности испытываемых способов использования сетевого сканирующего робота входят следующие индексы:

- $h_m$  — индекс эффективности ( $m$ -го) испытываемого способа использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ .
- $h_{mn}$  — индекс эффективности сканирования ( $n$ -го) информационного массива ( $m$ -м) испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ ,  $n = 1(1)N_m$ .
- $h_{mnr}$  — индекс эффективности сканирования ( $r$ -го) информационного документа ( $n$ -го) информационного массива ( $m$ -м) испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ ,  $n = 1(1)N_m$ ,  $r = R_{mn}$ .
- $h_{mnr s}$  — индекс эффективности сканирования ( $s$ -го) информационного документа ( $r$ -го) тематического раздела ( $n$ -го) информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ ,  $n = 1(1)N_m$ ,  $r = R_{mn}$ ,  $s = 1(1)S_{mnr}$ .

В группу индексов результативности испытываемых способов использования сетевого сканирующего робота для поиска и сбора информации включены следующие индексы:

- $k_m$  — индекс результативности ( $m$ -го) испытываемого способа использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ .
- $k_{mn}$  — индекс результативности сканирования ( $n$ -го) информационного массива ( $m$ -м) испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ ,  $n = 1(1)N_m$ .
- $k_{mnr}$  — индекс результативности сканирования ( $r$ -го) информационного документа ( $n$ -го) информационного массива ( $m$ -м) испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ ,  $n = 1(1)N_m$ ,  $r = R_{mn}$ .
- $k_{mnr s}$  — индекс результативности сканирования ( $s$ -го) информационного документа ( $r$ -го) тематического раздела ( $n$ -го) информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ ,  $n = 1(1)N_m$ ,  $r = R_{mn}$ ,  $s = 1(1)S_{mnr}$ .

В группу индексов ресурсоемкости испытываемых способов использования сетевого сканирующего робота для поиска и сбора информации включены следующие индексы:

- $c_m$  — индекс ресурсоемкости сканирования информационных массивов  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ .
- $c_{mn}$  — индекс ресурсоемкости сканирования ( $n$ -го) информационного массива ( $m$ -м) испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ ,  $n = 1(1)N_m$ .
- $c_{mnr}$  — индекс ресурсоемкости сканирования ( $r$ -го) тематического раздела ( $n$ -го) информационного массива ( $m$ -м) испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ ,  $n = 1(1)N_m$ ,  $r = R_{mn}$ .
- $c_{mnr s}$  — индекс ресурсоемкости сканирования ( $s$ -го) информационного документа ( $r$ -го) тематического раздела ( $n$ -го) информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ ,  $n = 1(1)N_m$ ,  $r = R_{mn}$ ,  $s = 1(1)S_{mnr}$ .

### 3. Вычисление индексов

Индекс результативности  $k_{mnr s}$  сканирования  $s$ -го информационного документа  $r$ -го тематического раздела  $n$ -го информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации будем вычислять по формуле

$$k_{mnr s} = v_{mnr s} x_{mnr s}, \quad (4)$$

$$m = 1(1)M, \quad n = 1(1)N_m, \quad r = 1(1)R_{mn}, \quad s = 1(1)S_{mnr}.$$

Основным ресурсом, затрачиваемым на поиск и сбор информации, будем считать время, затрачиваемое на данную операцию.

Поэтому индекс ресурсоемкости  $c_{mnr s}$  сканирования  $s$ -го информационного документа  $r$ -го тематического раздела  $n$ -го информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации будем вычислять по формуле

$$c_{mnr s} = \Delta t_{mnr s}, \quad (5)$$

$$m = 1(1)M, \quad n = 1(1)N_m, \quad r = 1(1)R_{mn}, \quad s = 1(1)S_{mnr},$$

где

$\Delta t_{mnr s}$  — время, затрачиваемое на поиск и обработку  $s$ -го информационного документа  $r$ -го тематического раздела  $n$ -го информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ ,  $n = 1(1)N_m$ ,  $r = R_{mn}$ ,  $s = 1(1)S_{mnr}$ .

Индекс эффективности  $h_{mnr s}$  сканирования  $s$ -го информационного документа  $r$ -го тематического раздела  $n$ -го информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации вычисляется по формуле

$$h_{mnrs} = \frac{k_{mnrs}}{c_{mnrs}} = \frac{v_{mnrs} x_{mnrs}}{c_{mnrs}} = \frac{v_{mnrs} x_{mnrs}}{\Delta t_{mnrs}}, \quad (6)$$

$$m = 1(1)M, \quad n = 1(1)N_m, \quad r = 1(1)R_{mn}, \quad s = 1(1)S_{mnr}.$$

Индекс результативности  $k_{mnr}$  сканирования  $r$ -го тематического раздела  $n$ -го информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации вычисляется по формуле

$$k_{mnr} = \sum_{s=1}^{S_{mnr}} k_{mnrs} = \sum_{s=1}^{S_{mnr}} v_{mnrs} x_{mnrs}, \quad (7)$$

$$m = 1(1)M, \quad n = 1(1)N_m, \quad r = 1(1)R_{mn}.$$

Индекс эффективности  $h_{mnr}$  сканирования  $r$ -го тематического раздела  $n$ -го информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации, вычисляется по формуле

$$h_{mnr} = \frac{k_{mnr}}{\Delta t_{mnr}} = \frac{\sum_{s=1}^{S_{mnr}} v_{mnrs} x_{mnrs}}{\Delta t_{mnr}}, \quad (8)$$

$$m = 1(1)M, \quad n = 1(1)N_m, \quad r = 1(1)R_{mn},$$

где

$\Delta t_{mnr}$  — время, затрачиваемое на поиск и обработку информационных документов  $r$ -го тематического раздела  $n$ -го информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,

$$m = 1(1)M, \quad n = 1(1)N_m, \quad r = 1(1)R_{mn}.$$

Индекс результативности  $k_{mn}$  сканирования  $n$ -го информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации вычисляется по формуле

$$k_{mn} = \sum_{r=1}^{R_{mn}} k_{mnr} = \sum_{r=1}^{R_{mn}} \sum_{s=1}^{S_{mnr}} k_{mnrs} = \sum_{r=1}^{R_{mn}} \sum_{s=1}^{S_{mnr}} v_{mnrs} x_{mnrs}, \quad (9)$$

$$m = 1(1)M, \quad n = 1(1)N_m.$$

Индекс эффективности  $h_{mn}$  сканирования  $n$ -го информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации вычисляется по формуле

$$h_{mn} = \frac{k_{mn}}{\Delta t_{mn}} = \frac{\sum_{r=1}^{R_{mn}} \sum_{s=1}^{S_{mnr}} k_{mnrs}}{\Delta t_{mn}} = \frac{\sum_{r=1}^{R_{mn}} \sum_{s=1}^{S_{mnr}} v_{mnrs} x_{mnrs}}{\Delta t_{mn}}, \quad (10)$$

$$m = 1(1)M, \quad n = 1(1)N_m,$$

где

$\Delta t_{mn}$  — время, затрачиваемое на поиск и обработку информационных документов  $n$ -го информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора ин-

формации,  
 $m = 1(1)M, n = 1(1)N_m, .$

Индекс результативности  $k_m$  сканирования информационных массивов  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации вычисляется по формуле

$$k_m = \sum_{n=1}^{N_m} k_{mn} = \sum_{n=1}^{N_m} \sum_{r=1}^{R_{mn}} k_{mnr} = \sum_{n=1}^{N_m} \sum_{r=1}^{R_{mn}} \sum_{s=1}^{S_{mnr}} v_{mnr s} x_{mnr s}, m = 1(1)M \quad (11)$$

Индекс эффективности  $h_m$  сканирования информационных массивов  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации вычисляется по формуле

$$h_m = \frac{k_m}{\Delta t_m} = \frac{\sum_{n=1}^{N_m} k_{mn}}{\Delta t_m} = \frac{\sum_{n=1}^{N_m} \sum_{r=1}^{R_{mn}} k_{mnr}}{\Delta t_m} = \frac{\sum_{n=1}^{N_m} \sum_{r=1}^{R_{mn}} \sum_{s=1}^{S_{mnr}} v_{mnr s} x_{mnr s}}{\Delta t_m}, m = 1(1)M, \quad (12)$$

где

$\Delta t_m$  — время, затрачиваемое на поиск и обработку информационных документов информационных массивов  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации,  
 $m = 1(1)M$ .

Если один из испытываемых способов использования сетевого сканирующего робота выбрать в качестве базового ( $m_0 \in [1(1)M]$ ), то для вычисления относительных индексов — индексов изменения эффективности и индексов изменения приведенной эффективности использования сетевого сканирующего робота можно применять выражения (5), (6), в которых базовые индексы эффективности, результативности и ресурсоемкости определены по результатам испытания базового способа.

#### 4. Точность вычисления индексов

Точность вычисления индексов эффективности использования сетевого сканирующего робота для поиска и сбора требуемой информации по результатам испытаний можно оценить следующим образом.

Пусть был испытан  $m$ -способ использования сетевого сканирующего робота для поиска и сбора требуемой информации,  $m \in [1(1)M]$ . При этом количество испытаний этого способа равно  $L_m$ .

В качестве оценки  $\tilde{h}_m = \bar{h}_m^* = \tilde{M}[h_m]$  математического ожидания  $M[h_m]$  индекса эффективности  $h_m$   $m$ -го способа использования сетевого сканирующего робота для поиска и сбора требуемой информации, полученного по результатам испытаний, выберем статистическое среднее, которое вычисляется по формуле [4]

$$\tilde{h}_m = \bar{h}_m^* = \tilde{M}[h_m] = \frac{1}{L_m} \sum_{l=1}^{L_m} h_{ml}, m = 1(1)M, \quad (13)$$

где

$h_{ml}$  — значение индекса эффективности  $h_m$   $m$ -го способа использования сетевого сканирующего робота для поиска и сбора требуемой информации, вычисленное по результатам  $l$ -го испытания,  
 $l = 1(1)L_m, m = 1(1)M$ .

Точность оценивания индекса эффективности  $h_m$   $m$ -го способа использования сетевого сканирующего робота для поиска и сбора требуемой информации, полученного по результатам испытаний определим по величине дисперсии  $D[\hat{h}_m]$  и среднего квадратического отклонения (СКО)  $\sigma[\hat{h}_m]$  индекса эффективности  $\hat{h}_m$ .

Для этого применим формулы [5]:

$$\tilde{D}[\hat{h}_m] = \frac{1}{L_m - 1} \sum_{l=1}^{L_m} (h_{ml} - \bar{h}_m^*)^2, \quad (14)$$

$$\sigma[\hat{h}_m] = E_{L_m-1} \sqrt{\tilde{D}[\hat{h}_m]} = E_{L_m-1} \sqrt{\frac{1}{L_m - 1} \sum_{l=1}^{L_m} (h_{ml} - \bar{h}_m^*)^2}, \quad (15)$$

$$E_{L_m-1} = \sqrt{\frac{E_{L_m-1}}{2}} \cdot \frac{\Gamma\left(\frac{E_{L_m-1}}{2}\right)}{\Gamma\left(\frac{E_{L_m}}{2}\right)}, \quad (16)$$

где

$\tilde{D}[\hat{h}_m]$  — состоятельная, несмещенная и асимптотически эффективная оценка дисперсии  $D[\hat{h}_m]$  индекса эффективности  $\hat{h}_m$   $m$ -го способа использования сетевого сканирующего робота для поиска и сбора требуемой информации, полученного по результатам испытаний этого способа,  $m = 1(1)M$ .

$\sigma[\hat{h}_m]$  — состоятельная, несмещенная и асимптотически эффективная оценка среднего квадратического отклонения (СКО) индекса эффективности  $\hat{h}_m$   $m$ -го способа использования сетевого сканирующего робота для поиска и сбора требуемой информации, полученного по результатам испытаний этого способа,  $m = 1(1)M$ .

$\Gamma(x)$  — гамма-функция.

$E_{L_m-1}$  — поправочный коэффициент, обеспечивающий несмещенность среднего квадратического отклонения СКО индекса эффективности  $\hat{h}_m$   $m$ -го способа использования сетевого сканирующего робота для поиска и сбора требуемой информации, полученного по результатам испытаний этого способа,  $m = 1(1)M$ .

$h_{ml}$  — значение индекса эффективности  $h_m$   $m$ -го способа использования сетевого сканирующего робота для поиска и сбора требуемой информации, вычисленное по результатам  $l$ -го испытания,  
 $l = 1(1)L_m, m = 1(1)M$ .

Значения поправочного коэффициента  $E_{L_m-1}$  для нескольких значений функции  $E_{L_m-1}$ , взятые из [4], приведены в табл. 1.



Значения поправочного коэффициента  $E_{L_m-1}$ 

$L_m-1$	2	3	4	5	6	9	12	18	24
$E_{L_m-1}$	1,128	1,085	1,064	1,051	1,042	1,028	1,021	1,014	1,010

Так как значения индексов эффективности использования сетевого сканирующего робота для поиска и сбора требуемой информации, вычисленные по результатам испытаний, являются случайными величинами, то наиболее полно эффективность использования сетевого сканирующего робота с данной целью может быть оценена с помощью соответствующих законов распределения.

Пусть все значения случайной величины  $\hat{q}$  расположены в интервале  $[q_{\min}, q_{\max}]$ , а само распределение можно считать унимодальным. Тогда для случайной величины  $\hat{q}$  в качестве аппроксимирующего закона может быть выбрано бета-распределение.

Выражения для плотности распределения  $\varphi(\hat{q}; q)$  и параметров  $\alpha$  и  $\gamma$  этого закона в этом случае имеют вид:

$$\varphi(\hat{q}; q) = \begin{cases} C(q - q_{\min})^\alpha (q_{\max} - q)^\gamma, & q \in (q_{\min}, q_{\max}), \\ 0, & q \notin (q_{\min}, q_{\max}), \end{cases} \quad (17)$$

где

$$C = [(q_{\max} - q_{\min})^{\alpha + \gamma + 1} B((\alpha + 1, \gamma + 1))]^{-1},$$

$$\alpha = \frac{(\bar{q} - q_{\min})\gamma - (q_{\max} + q_{\min}) + 2\bar{q}}{q_{\max} - \bar{q}},$$

$$\gamma = \frac{q_{\max} - \bar{q}}{q_{\max} - q_{\min}} \left[ \frac{(q_{\max} + q_{\min})\bar{q} - q_{\min}q_{\max} - \bar{q}^2}{D[\hat{q}]} + \frac{q_{\max} + q_{\min} - 2\bar{q}}{q_{\max} - \bar{q}} - 3 \right],$$

$q_{\min}$  — наименьшее значение случайной величины  $\hat{q}$ ,

$q_{\max}$  — наибольшее значение случайной величины  $\hat{q}$ ,

$\bar{q}$  — математическое ожидание случайной величины  $\hat{q}$ ,

$D[\hat{q}]$  — дисперсия случайной величины  $\hat{q}$ ,

$B(\alpha + 1, \gamma + 1)$  — бета-функция.

Подставляя в (17) значения математических ожиданий и дисперсий, вычисленные с помощью приведенных выше выражений (9, 10), а также значения наибольшей  $h_{\max}$  и наименьшей оценок  $h_{\min}$ , получаем формулы для соответствующих законов распределения.

Мода случайных величин, распределенных по этим законам, определяется выражением

$$\text{Mo}[\hat{q}] = \frac{\alpha q_{\max} + \gamma q_{\min}}{\alpha + \gamma}. \quad (18)$$

Следует отметить, что мода может быть также использована в качестве оценки индекса эффективности  $h_m$   $m$ -го способа использования сетевого сканирующего робота. Для ее вычисления достаточно применить выражение (18) к соответствующим результатам испытаний.

Применение индексного метода для оценивания эффективности использования сетевого сканирующего робота рассмотрим на следующем иллюстративном примере.

## 5. Пример

### Постановка задачи

Пусть индексный метод применяется для оценивания по результатам испытаний эффективности двух ( $M = 2$ ) способов использования сетевого сканирующего робота для поиска и сбора требуемой информации в информационной среде, которая включает различные информационные массивы, хранящие разнообразную информацию, имеющую различную ценность для лица, в интересах которого осуществляется поиск и сбор информации.

Требуется по результатам испытаний оценить эффективность испытываемых способов использования сетевого сканирующего робота и выбрать среди них наиболее эффективный.

### Исходные данные

Пусть для решения поставленной задачи имеются следующие исходные данные.

1. Для каждого способа проведено три ( $L_m = 3, m = 1,2$ ) испытания. При этом каждым способом был просканирован один ( $N_m = 1, m = 1,2$ ) информационный массив и выявлены  $S_{ml}, (m = 1,2, l = 1,2,3)$  информационных документов.

Количество  $S_{ml}$  выявленных при каждом испытании документов приведено в табл. 2.

Таблица 2

Количество  $S_{ml}$  информационных документов, выявленных при испытании способов использования сетевого сканирующего робота для поиска и сбора информации в соответствии с запросом  $m = 1,2, l = 1,2,3$

Испытание	Способ	
	$m = 1$	$m = 2$
$l = 1$	3	2
$l = 2$	1	3
$l = 3$	2	2

2. Для каждого информационного документа  $s_{ml}, m = 1,2, l = 1,2,3$ , выявленного при проведении испытаний, заданы (вычислены или определены экспертным путем) значения следующих показателей:

–  $v_{mls}, m = 1,2, l = 1,2,3, s = 1(1)S_{lm}$  – индекса ценности (важности) для лица, в интересах которого осуществляется поиск и сбор информации, сканирования  $s$ -го информационного документа информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации в  $l$ -м испытании,

–  $c_{ml}, m = 1,2, l = 1,2,3$  – индекса ресурсоемкости, в качестве которого выбрано время  $\Delta t_{ml}$ , затрачиваемое на поиск и обработку информационных документов информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота в  $l$ -м испытании, т. е.

$$c_{ml} = \Delta t_{ml}, m = 1,2, l = 1,2,3 \quad (19)$$

Эти значения представлены в таблице 3.

Исходные данные для расчета эффективности испытываемого способа использования сетевого сканирующего робота для поиска и сбора информации

$$m = 1, 2, \quad l = 1, 2, 3, \quad s_{ml} = 1(1)S_{ml}$$

Индекс		Значения индексов					
		$m = 1$			$m = 2$		
		$l = 1$	$l = 2$	$l = 3$	$l = 1$	$l = 2$	$l = 3$
$V_{mIs}$	$s = 1$	10	0	0	30	20	10
	$s = 2$	50	70	60		40	30
	$s = 3$	5	0	20	40	10	0
$C_{ml} = \Delta t_{ml}$		100	90	120	500	300	100

### Решение задачи

1. Индекс результативности  $k_{mIs}$  сканирования  $s$ -го информационного документа  $r$ -го тематического раздела  $n$ -го информационного массива  $m$ -м испытываемым способом использования сетевого сканирующего робота для поиска и сбора информации в соответствии с формулой (4) примем равным

$$k_{mIs} = v_{mIs}, \quad (20)$$

$$m = 1, 2, \quad l = 1, 2, 3, \quad s = 1(1)S_{ml}.$$

2. Индекс результативности  $k_{ml}$   $m$ -го способа использования сетевого сканирующего робота для поиска и сбора информации в  $l$ -м испытании вычисляется по формуле

$$k_{ml} = \sum_{s=1}^{S_{ml}} k_{mIs}, \quad m = 1, 2, \quad l = 1, 2, 3. \quad (21)$$

3. Индекс эффективности  $h_{ml}$   $m$ -го способа использования сетевого сканирующего робота для поиска и сбора информации в  $l$ -м испытании вычисляется по формуле

$$h_{ml} = \frac{k_{ml}}{\Delta t_{ml}} = \frac{\sum_{s=1}^{S_{ml}} k_{mIs}}{\Delta t_{ml}}, \quad m = 1, 2, \quad l = 1, 2, 3, \quad (22)$$

где

$\Delta t_{ml}$  — время, затрачиваемое на поиск и обработку информационных документов  $m$ -м способом использования сетевого сканирующего робота в  $l$ -м испытании,  $m = 1, 2, \quad l = 1, 2, 3$ .

Значения индексов результативности  $k_{mIs}$  и  $k_{ml}$ , а также индексов эффективности  $h_{ml}$ , определенные в соответствии с (20), (21) и (22), приведены в табл. 4.

Таблица 4

Расчетные значения индексов результативности  $k_{m/s}$ ,  $k_{ml}$  и эффективности  $h_{ml}$  испытываемого способа использования сетевого сканирующего робота для поиска и сбора информации  
 $m = 1, 2$ ,  $l = 1, 2, 3$ ,  $s = 1(1)S_{ml}$

Индекс		Значения индексов					
		$m = 1$			$m = 2$		
		$l = 1$	$l = 2$	$l = 3$	$l = 1$	$l = 2$	$l = 3$
$k_{m/s}$	$s = 1$	10	–	–	30	20	10
	$s = 2$	50	70	60	–	40	30
	$s = 3$	5	–	20	40	10	–
$k_{ml}$		65	70	80	70	70	40
$c_{ml} = \Delta t_{ml}$		100	90	120	500	300	100
$h_{ml}$		0,65	0,78	0,67	0,14	0,23	0,40

4. Оценка  $\tilde{h}_m = \bar{h}_m^* = \tilde{M}[\hat{h}_m]$  математического ожидания  $M[\hat{h}_m]$  индекса эффективности  $h_m$   $m$ -го способа ( $m = 1, 2$ ) использования сетевого сканирующего робота для поиска и сбора требуемой информации, полученного по результатам испытаний, вычисляется по формуле (13)

$$\tilde{h}_m = \bar{h}_m^* = \tilde{M}[\hat{h}_m] = \frac{1}{3} \sum_{l=1}^{L_m} h_{ml}, \quad (23)$$

где

$h_{ml}$  — значение индекса эффективности  $h_m$   $m$ -го способа использования сетевого сканирующего робота для поиска и сбора требуемой информации, вычисленное по результатам  $l$ -го испытания,  $m = 1, 2$ ,  $l = 1, 2, 3$ .

Для рассматриваемого примера

$$\tilde{h}_1 = \bar{h}_1^* = \frac{1}{3} \sum_{l=1}^3 h_{1l} = \frac{1}{3} (0,65 + 0,78 + 0,67) = 0,7.$$

$$\tilde{h}_2 = \bar{h}_2^* = \frac{1}{3} \sum_{l=1}^3 h_{2l} = \frac{1}{3} (0,14 + 0,3 + 0,40) = 0,26.$$

5. Выберем в качестве базового способа использования сетевого сканирующего робота второй способ ( $m = 2$ ).

Используя в качестве базового индекса эффективности  $h_0$  оценку  $\tilde{h}_2 = \bar{h}_2^* = \tilde{M}[\hat{h}_2]$  математического ожидания  $M[\hat{h}_2]$  индекса эффективности  $h_2$  второго способа использования сетевого сканирующего робота для поиска и сбора требуемой информации, получаем

$$h_0 = \tilde{h}_2.$$

Индекс эффективности  $h_1$  первого способа использования сетевого сканирующего робота для поиска и сбора требуемой информации примем равным

оценке  $\tilde{h}_1 = \bar{h}_1^* = \tilde{M}[\hat{h}_1]$  математического ожидания индекса эффективности этого способа использования сетевого сканирующего робота для поиска и сбора требуемой информации.

$$h_1 = \tilde{h}_1 = \bar{h}_1^*.$$

Индекс  $\delta h_1$  изменения эффективности при применении первого ( $m = 1$ ) способа использования сетевого сканирующего робота, определенной по результатам испытаний, вычислим по формуле (2).

$$\delta h_1 = \frac{h_1 - h_0}{h_0} = \frac{0,7 - 0,26}{0,26} = 1,7.$$

6. Точность вычисления оценки  $\tilde{h}_m$  математического ожидания  $M[\hat{h}_m]$  индекса эффективности  $h_m$   $m$ -го способа ( $m = 1, 2$ ) использования сетевого сканирующего робота для поиска и сбора требуемой информации, полученного по результатам испытаний, определяется по величине дисперсии  $D[\hat{h}_m]$  или среднего квадратического отклонения (СКО)  $\sigma[\hat{h}_m]$  индекса эффективности  $\hat{h}_m$ , которые вычисляются по формулам (14), (15) с использованием таблицы 1.

Из таблицы 1 для  $L_m - 1 = 2$  находим

$$E_2 = 1,128.$$

Для  $m = 1$  по формулам (14), (15) получаем

$$\begin{aligned} \tilde{D}[\hat{h}_1] &= \frac{1}{2} \sum_{l=1}^3 (h_{1l} - \bar{h}_1^*)^2 = 0,5[(0,65-0,7)^2 + (0,78-0,7)^2 + (0,67-0,7)^2] = \\ &= 0,5(0,0025+0,0064+0,0009) = 0,0049. \end{aligned}$$

$$\tilde{\sigma}[\hat{h}_1] = E_2 \sqrt{\tilde{D}[\hat{h}_1]} = 1,128 \sqrt{0,0049} = 0,08.$$

Для  $m = 2$  по формулам (14), (15) получаем

$$\begin{aligned} \tilde{D}[\hat{h}_2] &= \frac{1}{2} \sum_{l=1}^3 (h_{2l} - \bar{h}_2^*)^2 = 0,5[(0,14-0,26)^2 + (0,23-0,26)^2 + (0,40-0,26)^2] = \\ &= 0,5(0,0144+0,0009+0,0169) = 0,016. \end{aligned}$$

$$\tilde{\sigma}[\hat{h}_2] = E_2 \sqrt{\tilde{D}[\hat{h}_2]} = 1,128 \sqrt{0,016} = 0,14.$$

### **Результаты решения задачи**

Результаты расчетов и исходные данные поставленной задачи сведены в табл. 5.

Анализ полученных результатов показывает, что эффективность  $h_1$  первого способа использования сетевого сканирующего робота для поиска и сбора требуемой информации выше, чем эффективность  $h_2$  второго способа в 1,7 раза.

Примененный индексный метод оценивания эффективности использования сетевого сканирующего робота для поиска и сбора требуемой информации имеет следующие достоинства:

– позволяет относительно просто с помощью абсолютных и относительных индексов оценить и сравнить между собой по эффективности различные способы **использования сетевого сканирующего робота** информации;

- не требует для проведения расчетов большого количества исходных данных;
- результаты расчетов относительно легко интерпретируются.

Таблица 5

Исходные данные и результаты расчетов при оценивании эффективности испытываемого способа использования сетевого сканирующего робота для поиска и сбора информации по результатам испытаний  
( $m = 1, 2$ ,  $m = 1, 2, 3$ ,  $s = 1(1)S_{ml}$ )

Индекс		Значения индексов					
		$m = 1$			$m = 2$		
		$l = 1$	$l = 2$	$l = 3$	$l = 1$	$l = 2$	$l = 3$
<b>Исходные данные</b>							
$v_{mls}$	$s = 1$	10	–	–	30	20	10
	$s = 2$	50	70	60	–	40	30
	$s = 3$	5	–	20	40	10	–
$c_{ml} = \Delta t_{ml}$		100	90	120	500	300	100
<b>Результаты расчетов</b>							
$k_{mls}$	$s = 1$	10	–	–	30	20	10
	$s = 2$	50	70	60	–	40	30
	$s = 3$	5	–	20	40	10	–
$k_{ml}$		65	70	80	70	70	40
$c_{ml} = \Delta t_{ml}$		100	90	120	500	300	100
$h_{ml}$		0,65	0,78	0,67	0,14	0,23	0,40
$h_m = \tilde{h}_m = \bar{h}_m^*$		0,7			0,26		
$\tilde{D}[\hat{h}_m]$		0,0049			0,016		
$\tilde{\sigma}[\hat{h}_m]$		0,08			0,14		
$h_0 = \tilde{h}_2 = \bar{h}_2^*$		–			0,26		
$\delta h_m$		1,7			0		

## Литература

1. Общая теория статистики / Ф. Г. Долгушевский, В. С. Козлов, П. И. Полушин, Я. М. Эрлих. Изд-е 2-е, М.: Статистика, 1967. С. 329–331.
2. Козлов Т. И., Овсиенко В. Е., Смирнский В. И. Курс общей теории статистики. Изд-е 2-е. М.: Статистика, 1965. С. 262.
3. Статистика / Под ред. С. Г. Струмилина // Изд-е 2-е, М.: Статистика, 1969. С. 207.
4. Статистические методы обработки результатов наблюдений / Под ред. Р. М. Юсупова. Л.: Изд-во МО СССР. 1984. 563 с.
5. Юсупов Р. М., Заболотский В. П. Научно-методологические основы информатизации. СПб.: Наука, 2000. 455 с.