

М. М. С о л о в ь е в (Ульяновск, УлГУ). **Задача эффективного поиска в сети.**

Работа, представленная данным сообщением, посвящена задаче нахождения наиболее благоприятного момента остановки поиска, при котором затраченное на него время и полученный результат наиболее эффективны. Приводятся ее теоретический анализ и сравнение с реальным экспериментом.

Пусть есть некоторая поисковая структура, которая, приняв заявку, начинает просматривать объекты, удовлетворяющие ее условию. Системе требуется время, чтобы обработать поступивший запрос и классифицировать его как подходящий под заданные требования.

Формально задачу можно представить следующим образом. Обозначим M_ε границу, превысив которую, система останавливается, считая результат найденным. Пусть $t \in [1, \dots, T]$ — временной интервал, в котором работает поисковик. Если за время T не было найдено ничего, удовлетворяющего требованиям, то поиск прекращается. Далее $\{X_i\}_{i=1, \dots}$ — значения параметра объекта поиска в момент времени i , причем $\{X_i\}_{i=1, \dots}$ — независимые, одинаково распределенные случайные величины с функцией распределения

$$F_{X_i}(x) = \begin{cases} 1 - e^{-\gamma x}, & x > 0, \\ 0, & x \leq 0, \end{cases} \quad (1)$$

где γ — положительный коэффициент, и пусть $\tau_i = \inf\{t > 0: \pi_t = i\}$ — моменты времени, в которые поисковая система получает объект для анализа, при $\tau_0 = 0$.

Решается задача определения τ — момента нахождения первого удовлетворяющего объекта условию $\tau_i = \inf\{t: Y_t \geq M_\varepsilon\}$, где процесс наблюдений $Y = (Y_t)_{t \geq 0}$ имеет вид $Y_t = Y_0 + \int_0^t (X_{\pi_s} - Y_{s-}) d\pi_s$, т. е. Y_t имеет кусочно-постоянные траектории, которые равны X_{π_s} в моменты скачка π_s . Здесь π_t — пуассоновский процесс с интенсивностью наблюдений $\lambda > 0$: $\pi_t = \lambda t + m_t$, в моменты скачка которого будут появляться объекты в поисковой системе для анализа.

Таким образом, задача сводится к минимизации функционала $\Phi(\varepsilon) = \alpha \mathbf{E} \tau + \varepsilon \rightarrow \min_\varepsilon$, где правая часть в сумме — это среднее всех моментов остановки, α — левая оптимальная граница. Заметим, что при уменьшении ε величина $\mathbf{E} \tau$ возрастает. Оптимальные значения ε и регулируются коэффициентом $\alpha > 0$ при $\mathbf{E} \tau$.

Результаты имитационного моделирования и оценивания по реальным данным привели к близким результатам. Задача была рассмотрена на примере нахождения времени наибольшей угрозы сбой работы сервера. Расчеты для имитационного моделирования были выполнены в среде Borland Delphi 7.

Автор выражает благодарность за внимание к работе профессору А. А. Бутову. Работа выполнена в рамках ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009–2013, Государственный контракт 02.740.11.0610, а также при поддержке РФФИ, проект № 08–01–97009.