

УДК 004.02

Этапы и особенности разработки методики повышения качества информационного поиска на ресурсах современных центров обработки данных с использованием нечетких отношений предпочтения и сравнения альтернатив

Stages and features of the development of a methodology for improving the quality of information retrieval on the resources of modern data centers using fuzzy preference and comparison relationships of alternatives

Саяркин / Sayarkin L.

Леонид Андреевич
(leonid.sayarkin@yandex.ru)
ФГКВОУ ВО «Военная академия связи имени Маршала Советского Союза С. М. Буденного» МО РФ (ВАС им. С. М. Буденного), соискатель кафедры автоматизированных систем специального назначения (АС СН).
г. Санкт-Петербург

Владимирова / Vladimirova E.

Елена Сергеевна
(lena_vladimir@mail.ru)
ФГАОУ ВО «Севастопольский государственный университет», старший преподаватель кафедры информационных технологий и компьютерных систем.
г. Севастополь

Паращук / Parashchuk I.

Игорь Борисович
(shchuk@rambler.ru)
доктор технических наук, профессор, заслуженный изобретатель РФ.
ВАС им. С. М. Буденного, профессор кафедры АС СН.
г. Санкт-Петербург

Ключевые слова: центр обработки данных – data center; показатель качества – quality indicator; нечеткость исходных данных – fuzziness of source data; информационный поиск – information search; ресурсы – resources; нечеткие отношения предпочтения – fuzzy preference relationships; анализ альтернатив – analysis of alternatives; методика – methodology; этапы – stages; реализация поисковых запросов – implementation of search queries.

Статья посвящена исследованию перспектив применения методов теории нечетких множеств для устранения неопределенности исходных данных в рамках методики повышения качества информационного поиска на ресурсах современных центров обработки данных, в частности возможности использования для решения подобных задач методов нечетких отношений предпочтения и сравнения альтернатив. Рассмотрены особенности и содержание этапов методики повышения качества реализации поисковых запросов, которая, в свою очередь, позволит повысить качество информационного обеспечения лиц, принимающих решения.

The article is devoted to the study of the prospects of using fuzzy set theory methods to eliminate the uncertainty of the source data within the framework of the methodology for improving the quality of information retrieval on the resources of modern data centers, in particular, the possibility of using methods of fuzzy preference relations and comparing alternatives to solve such problems. The features and content of the stages of the methodology for improving the quality of search query implementation are considered, which, in turn, will improve the quality of information support for decision makers.

Введение

Тревожная военно-политическая и экономическая ситуация, образовавшаяся на современном этапе формирования и развития международных и межнациональных геополитических отношений, например в Ближневосточном регионе и на Украине, стимулирует как зарубежные страны, так и Российскую Федерацию активно строить и повседневно совершенствовать свои системы и комплексы управления народным хозяйством, торговлей, добычей полезных ископаемых, здравоохранением, образованием, а также силами и средствами обеспечения обороноспособности государства.

Именно поэтому как никогда актуальными, по нашему мнению, представляются подходы к решению проблем оперативного доведения важной для всех этих сфер жизнедеятельности России информации до лиц, принимающих решения (ЛПР), во всех органах законодательной и исполнительной власти, департаментах, органах и пунктах управления регионами и отраслями народного хозяйства страны в кратчайшие сроки и с требуемым качеством. Ответственной за эти важные задачи справедливо принято считать ИТ-инфраструктуру государства, которая представляет собой совокупность организационно-технологических структур, призванных обеспечивать и контролировать функционирование информационного пространства России. Это структурно и функционально взаимосвязанная совокупность аппаратных и программных средств и комплексов сбора, обработки и хранения данных, составляющих базис современных центров обработки данных (ЦОД), систем связи и распространения информации, баз знаний и баз данных, а также иных объектов и элементов, технической основой которых являются коммуникационные и компьютерные технологии [1, 2].

При этом очевидно, что качество предоставляемых услуг и эффективность функционирования обеспечивающих эти сервисы аппаратных и программных средств, составляющих основу одного из ключевых компонентов ИТ-инфраструктуры – современных ЦОД, обоснованно претендуют на повышенное внимание со стороны специалистов в данной отрасли, и, на наш взгляд, обязаны контролироваться и управляться непрерывно и на всех уровнях архитектуры сложных объектов такого класса, часто также именуемых дата-центрами [3, 4].

Современные ЦОД, а конкретнее, надежность, безошибочность и оперативность их функционирования, обуславливают качество информационного обеспечения процессов принятия решений и поддержки принятия решений в рамках контроля (мониторинга) как в государственной, правоохранительной, оборонной, так и в промышленно-экономической и общественной (социальной) областях деятельности современного государства [5, 6].

В этой связи все более очевидно проявляется объективная необходимость решения задачи построения оптимальных (быстрых и безошибочных) поисковых платформ, «поисковых машин», «поисковых движков» и поисковых алгоритмов, что, в свою очередь, определяет актуальность формулировки и решения научной задачи повышения качества поиска необходимой ЛПР информации в больших массивах гетерогенных данных, собираемых, хранимых и обрабатываемых на ресурсах современных ЦОД.

В большинстве случаев речь идет о проблемах, связанных с недостаточной полнотой и точностью предоставляемых данных по конкретному поисковому запросу ЛПР, причем данных, предоставляемых из большого массива и за очень короткий интервал времени. Современные исследования, посвященные оптимизации и качеству информационного поиска, трактуют совокупность этих двух ключевых критериев (полнота и точность) как интегральный показатель качества реализации поисковых запросов, характеризующий уровень соответствия данного запроса требованиям ЛПР, и называют этот показатель «релевантностью» поиска [7].

Кроме того, в рамках современных процедур управления и с учетом ограниченного времени на принятие решений довольно часто имеет место неопределенность (нечеткость) критериев и границ информационного поиска на ресурсах ЦОД, а также зачастую – нечеткость («размытость») конкретных целей ЛПР при формулировке поисковых запросов. Более того, существует неопределенность (нечеткость) самих разнородных критериев (показателей) качества поиска и необходимость их компромиссного согласования в рамках реализации запроса с учетом важности весов этих критериев (показателей), например, с учетом их иерархии, а также необходимость разграничения значимости альтернатив в процедурах и механизмах поиска. Все это создает существенные трудности при реализации подобных запросов, делает такого рода задачи нечеткими («размытыми») с точки зрения математики и кибернетики, делает их еще более сложными, но увлекательными и, бесспорно, актуальными [8–10].

Таким образом, практика показывает, что аппаратные и программные средства, используемые для информационного поиска на ресурсах современных ЦОД, не в полной мере удовлетворяют требованиям по своевременности и качеству (релевантности) реализации поисковых запросов. С учетом этого особую актуальность, по нашему мнению, приобретает содержательная формулировка этапов и особенностей методики повышения качества информационного поиска на ресурсах современных ЦОД с использованием методов теории нечетких множеств, например, методов нечетких отношений предпочтения и методов анализа нечетких альтернатив.

Теоретические и методологические аспекты формулировки этапов и особенностей методики повышения качества информационного поиска на ресурсах современных ЦОД с использованием нечетких отношений предпочтения и сравнения альтернатив

Учитывая особенности современных ЦОД, общая цель исследований может быть сформулирована как повышение оперативности и релевантности реализации поисковых запросов ЛПР в распределенных системах хранения данных и дата-центрах с использованием методов интеллектуальной обработки больших массивов гетерогенных данных, методов, основанных на нечетких отношениях предпочтения и алгоритмах анализа нечетких альтернатив.

Подобные методы и алгоритмы достаточно детально изучены, эффективность их практического использования многократно доказана, она затрагивает различные сферы принятия решений, например, процедуры обоснования многоальтернативного выбора сложных технических объектов, как показано в работе [11].

Рассмотрим аспекты, характеризующие физический смысл, сущность и особенности этапов предлагаемого подхода к повышению качества информационного поиска на ресурсах современных ЦОД с использованием методов и алгоритмов принятия решений, предполагающих построение множества недоминирующих альтернатив на основе нечеткого отношения предпочтения.

При этом система показателей качества (СПК) информационного поиска на ресурсах современных ЦОД может быть представлена формально в виде вектора $\vec{Y}_{инп}(k)$, состоящего, например, из N ($n=1, \dots, N$) показателей:

$$\vec{Y}_{инп}(k) = [y_1(k); y_2(k); \dots y_n(k); \dots y_N(k)]^T, \quad (1)$$

где $y_n(k)$ – частные показатели качества реализации поисковых запросов на k -м шаге информационного поиска, определяемые в виде реальных значений параметров, например, безошибочности (достоверности, точности), оперативности, размера (объема) итогов поиска, его полноты, масштабности поиска (количество анализируемых одновременно хранилищ данных), уровня пригодности результатов для дальнейшего использования, релевантности и пертинентности поиска, а также иных переменных, характеризующих отдельные грани качества выполнения запросов такого рода [12–15].

В частности, СПК информационного поиска на ресурсах современных ЦОД может включать не только и столько полноту и точность, характеризующие уровень соответствия данного запроса требованиям ЛПР – собственно релевантность, но и вероятность соответствия между формальной релевантностью (релевантностью, извлекаемой поисковой системой)

и содержательной релевантностью (субъективной релевантностью конкретного ЛПР) – пертинентность [12, 14].

Таким образом, СПК информационного многокритериального и многоальтернативного поиска (с учетом комплексирования нескольких альтернатив оценки реализации или, точнее, итогов реализации поисковых запросов) на ресурсах современных ЦОД, по нашему мнению и с учетом обозначений, введенных в выражении (1), может включать, например, 4 (четыре) частных показателя качества:

$y_1(k) = y_{рел}(k)$ – релевантность, т. е. уровень соответствия данного запроса текущим требованиям ЛПР на k -м шаге информационного поиска на ресурсах ЦОД;

$y_2(k) = y_{опер}(k)$ – оперативность поиска, т. е. время реализации конкретного текущего поискового запроса ЛПР на k -м шаге информационного поиска (интервал времени поиска);

$y_3(k) = y_{адек}(k)$ – адекватность, т. е. уровень достоверности (безошибочности) результатов поиска;

$y_4(k) = y_{раз}(k)$ – размерность (объем) данных, которые необходимо найти в рамках конкретного текущего поискового запроса ЛПР на k -м шаге информационного поиска на ресурсах ЦОД.

Тем самым определена СПК информационного поиска на ресурсах современных ЦОД в виде вектора $\vec{Y}_{инп}(k)$, состоящего, для нашего примера, из четырех показателей

$$\vec{Y}_{инп}(k) = [y_{рел}(k); y_{опер}(k); y_{адек}(k); y_{раз}(k)]^T. \quad (2)$$

С точки зрения теоретических и методологических аспектов формулировки задачи исследования очевидно, что в рамках реализации информационного поиска на ресурсах современных ЦОД на основе рассмотренных критериев (показателей) должна быть решена задача принятия решения, заключающегося в определении (оценке, идентификации) степени соответствия результатов реализации поисковых запросов совокупности (комплексу) требуемых значений этих показателей, зачастую заданных недостоверно (нечетко) [15].

Это вариант типовой задачи многокритериального (комплексного) принятия решения в условиях неопределенности, нечеткости исходных данных.

В этой связи рассматриваемая и предлагаемая к практическому использованию методология информационного поиска может и должна, по нашему мнению, опираться на алгоритмы принятия решений в условиях неопределенности такого типа, основанные на построении множества недоминирующих альтернативных критериев (показателей) качества поиска с использованием нечетких отношений предпочтения [11, 16–20].

Теоретическими (математическими) и методологическими аспектами этапов решения данной задачи

выступают особенности комплексирования критериев (показателей) качества информационного поиска с использованием нечетких отношений предпочтения и сравнения альтернатив, которые можно представить в виде последовательности шагов реализации поискового запроса.

На первом шаге задаются значения (или интервал значений) сформулированных в виде выражения (2) альтернатив для критериев (показателей) качества информационного поиска на ресурсах современных ЦОД. В контексте терминов теории нечетких множеств данная процедура носит название «определение нечетких переменных для показателей $y_1(k) \dots y_4(k)$ » [8, 9].

Так, например, для $y_1(k) = y_{\text{рел}}(k)$ – релевантности информационного поиска, значениями, т. е. возможными альтернативами уровня соответствия данного запроса текущим требованиям на k -м шаге реализации поисковых запросов на ресурсах современных ЦОД могут выступать 6 (шесть) допустимых альтернатив, характеризующих в рамках нашего примера потенциальную релевантность для: $\{s_1\}$ – мультимедийных запросов, $\{s_2\}$ – транзакционных запросов, $\{s_3\}$ – информационных запросов, $\{s_4\}$ – навигационных запросов, $\{s_5\}$ – картографических запросов и, наконец, $\{s_6\}$ – для смешанных (неопределенных) запросов ЛПР.

На втором шаге каждый из этих показателей качества информационного поиска формулируется в виде некоторой переменной, причем значения данных переменных задаются на конкретном базовом (универсальном) множестве значений, а также формируется тип функции принадлежности данной переменной.

Допустим, H – то самое базовое множество, h – элемент базового множества H , а W – определенное свойство, какая-либо грань качества информационного поиска. В этом случае подмножество S базового множества H , элементы которого находятся в соответствии со свойством W , в терминах теории нечетких множеств определяются как множество упорядоченных пар:

$$\tilde{S} = \{\mu_{\tilde{s}}(h) | (h)\}, \quad (3)$$

где $\mu_{\tilde{s}}(h)$ – функция принадлежности. Физический смысл функции принадлежности – это характеристическая функция, принимающая значение 0, если h полностью не удовлетворяет свойству W , и 1 – в противном случае, т. е. полностью удовлетворяет. При этом конкретные значения альтернатив (которые сами являются нечеткими множествами) по всем рассмотренным ранее четырем показателям качества информационного поиска $y_1(k) \dots y_4(k)$ находят с использованием характеризующих их функций принадлежности.

Воспользуемся примером применения метода нечеткого отношения предпочтения для принятия

оптимального решения по выбору сложного технического объекта [11].

Речь идет не только о наглядной демонстрации применимости, но и о практической проверке пригодности предлагаемых алгоритмов для решения нашей задачи комплексного (оптимального компромиссного сравнения и сочетания альтернативных категорий «важности» для различных показателей качества) информационного поиска на ресурсах современных ЦОД с использованием нечетких отношений предпочтения и метода сравнения альтернатив.

Таким образом, на третьем шаге реализации поискового запроса для рассматриваемого (в качестве примера) показателя $y_1(k) = y_{\text{рел}}(k)$ – релевантности информационного поиска (уровня соответствия данного запроса текущим требованиям ЛПР на k -м шаге реализации поисковых запросов на ресурсах ЦОД), анализируемое нечеткое множество может иметь вид:

$$\mu_{y_1}(k) = \{0,1 | 2; 0,2 | 10; 1 | 20; 1 | 50; 1 | 70; 0,2 | 90; 0,05 | 100\}, \quad (4)$$

где значения функции принадлежности изменяются от 0 до 1 в соответствии с мнением экспертов (0,1; 0,2; 1; 1; 1; 0,2; 0,05) о потенциальном уровне релевантности, который можно условно измерить в процентах, например от 0 до 100 (2, 10, 20, 50, 70, 90, 100). Подобным же образом формируются нечеткие множества $\mu_{y_2}(k)$, $\mu_{y_3}(k)$ и $\mu_{y_4}(k)$ для соответствующих показателей качества $y_1(k) \dots y_4(k)$ информационного поиска.

Имея эти исходные данные, поисковой системе необходимо отыскать рациональную (в идеале – оптимальную) многокритериальную альтернативу качественного информационного поиска с максимальной степенью недоминируемости.

На четвертом шаге, опираясь на исходные данные (например, для $y_1(k) = y_{\text{рел}}(k) = \{0,1; 0,2; 1; 1; 1; 0,2; 0,05\}$, выражение (4)), для всех четырех показателей качества информационного поиска должны быть составлены матрицы нечетких отношений предпочтения W (соответственно W_1, W_2, W_3 и W_4). Эти матрицы, в свою очередь, определены на множестве решений нечеткого подмножества \tilde{S}

$$\tilde{S} = \{\tilde{s}_1; \tilde{s}_2; \tilde{s}_3; \tilde{s}_4; \tilde{s}_5; \tilde{s}_6\}, \quad (5)$$

и содержат соответствующие нечеткие отношения предпочтения $\mu_{w_1}(k)$, $\mu_{w_2}(k)$, $\mu_{w_3}(k)$ и $\mu_{w_4}(k)$.

На пятом шаге приступают к построению нечеткого отношения \tilde{Z}_1 :

$$\tilde{Z}_1 = W_1 \cap W_2 \cap W_3 \cap W_4. \quad (6)$$

Здесь нули либо единицы, характеризующие значения совместной функции принадлежности нескольких i -х и j -х нечетких переменных ($\tilde{s}_i; \tilde{s}_j$)

в рамках данного нечеткого отношения, выступают элементами матрицы нечеткого отношения \tilde{Z}_1 :

$$\mu_{\tilde{Z}_1} = \{\tilde{s}_i; \tilde{s}_j\}. \quad (7)$$

Шестой шаг поиска рациональной многокритериальной альтернативы качественного информационного поиска с максимальной степенью недоминируемости характеризуется определением подмножества недоминирующих альтернатив на нечетком множестве $\{\tilde{S}, \mu_{\tilde{Z}_1}\}$ по всем i -м и j -м нечетким переменным $(\tilde{s}_i; \tilde{s}_j)$, где $i \neq j$:

$$\mu_{\tilde{Z}_1}(\tilde{s}_i) = 1 - \sup_{i,j} (\mu_{\tilde{Z}_1}(\tilde{s}_j, \tilde{s}_i) - \mu_{\tilde{Z}_1}(\tilde{s}_i, \tilde{s}_j)). \quad (8)$$

Для нечеткой переменной $\{\tilde{s}_1\}$, характеризующей в нашем случае одну из альтернатив значений релевантности поиска (альтернативу релевантности поиска при реализации мультимедийных запросов ЛППР), подмножество недоминирующих (non-dominant, n-d) альтернатив на нечетком множестве $\{\tilde{S}, \mu_{\tilde{Z}_1}\}$ определяется по формуле:

$$\begin{aligned} \mu_{\tilde{Z}_1}^{n-d}(\tilde{s}_1) = 1 - \sup_{1, \dots, 6} (\mu_{\tilde{Z}_1}(\tilde{s}_2, \tilde{s}_1) - \mu_{\tilde{Z}_1}(\tilde{s}_1, \tilde{s}_2), \mu_{\tilde{Z}_1}(\tilde{s}_3, \tilde{s}_1) - \\ - \mu_{\tilde{Z}_1}(\tilde{s}_1, \tilde{s}_3), \mu_{\tilde{Z}_1}(\tilde{s}_4, \tilde{s}_1) - \mu_{\tilde{Z}_1}(\tilde{s}_1, \tilde{s}_4), \mu_{\tilde{Z}_1}(\tilde{s}_5, \tilde{s}_1) - \\ - \mu_{\tilde{Z}_1}(\tilde{s}_1, \tilde{s}_5), \mu_{\tilde{Z}_1}(\tilde{s}_6, \tilde{s}_1) - \mu_{\tilde{Z}_1}(\tilde{s}_1, \tilde{s}_6)) = 1. \end{aligned} \quad (9)$$

Искомые подмножества недоминирующих альтернатив $\mu_{\tilde{Z}_1}^{n-d}(\tilde{s}_2)$, $\mu_{\tilde{Z}_1}^{n-d}(\tilde{s}_3)$, $\mu_{\tilde{Z}_1}^{n-d}(\tilde{s}_4)$, $\mu_{\tilde{Z}_1}^{n-d}(\tilde{s}_5)$ и $\mu_{\tilde{Z}_1}^{n-d}(\tilde{s}_6)$ для соответствующих нечетких переменных $\{\tilde{s}_2\}$, $\{\tilde{s}_3\}$, $\{\tilde{s}_4\}$, $\{\tilde{s}_5\}$ и $\{\tilde{s}_6\}$ определяются аналогично.

По аналогии с результатами, полученными в работе [11], в итоге получаем функцию принадлежности из нечеткого множества недоминирующих альтернатив $\mu_{\tilde{Z}_1}^{n-d}$:

$$\mu_{\tilde{Z}_1}^{n-d} = \{1; 1; 1; 1; 1; 1\}. \quad (10)$$

Седьмой шаг поиска рациональной многокритериальной альтернативы качественного информационного поиска с максимальной степенью недоминируемости – строится отношение в соответствии с выражением (11):

$$\mu_{\tilde{Z}_2}(\tilde{s}_i, \tilde{s}_j) = \sum_{l=1}^4 \Omega_l \mu_{\mu_{\tilde{Z}_1}}(\tilde{s}_i, \tilde{s}_j). \quad (11)$$

Здесь используемые для построения отношения (11) индексы Ω_l – коэффициенты относительной «важности» каждого из четырех введенных нами ранее ($l = 1, \dots, 4$) показателей качества информационного поиска на ресурсах современных ЦОД, должны быть определены заблаговременно на основе экспертных оценок, например, $\Omega_1 = 0,59$; $\Omega_2 = 0,14$; $\Omega_3 = 0,18$; $\Omega_4 = 0,09$.

В результате находим нечеткое отношение \tilde{Z}_2 .

На следующем шаге по всем i -м и j -м нечетким переменным ($i \neq j$) определяется подмножество недоминирующих альтернатив на нечетком множестве $\{\tilde{S}, \mu_{\tilde{Z}_2}\}$:

$$\mu_{\tilde{Z}_2}(\tilde{s}_i) = 1 - \sup_{i,j} (\mu_{\tilde{Z}_2}(\tilde{s}_j, \tilde{s}_i) - \mu_{\tilde{Z}_2}(\tilde{s}_i, \tilde{s}_j)). \quad (12)$$

На основе выражения (12) с учетом введенных индексов (коэффициентов) Ω_l и по аналогии с нечетким отношением \tilde{Z}_1 для нечеткого отношения \tilde{Z}_2 определяются подмножества недоминирующих альтернатив $\mu_{\tilde{Z}_2}^{n-d}(\tilde{s}_2)$, $\mu_{\tilde{Z}_2}^{n-d}(\tilde{s}_3)$, $\mu_{\tilde{Z}_2}^{n-d}(\tilde{s}_4)$, $\mu_{\tilde{Z}_2}^{n-d}(\tilde{s}_5)$ и $\mu_{\tilde{Z}_2}^{n-d}(\tilde{s}_6)$ для соответствующих нечетких переменных $\{\tilde{s}_2\}$, $\{\tilde{s}_3\}$, $\{\tilde{s}_4\}$, $\{\tilde{s}_5\}$ и $\{\tilde{s}_6\}$.

В результате имеем функцию принадлежности из нечеткого множества недоминирующих альтернатив $\mu_{\tilde{Z}_2}^{n-d}$.

Например, для нашего случая по аналогии с работой [11] в результате вычислений с использованием введенных ранее индексов Ω_l получим значения нечеткого множества недоминирующих альтернатив $\mu_{\tilde{Z}_2}^{n-d}$:

$$\mu_{\tilde{Z}_2}^{n-d} = \{0,998; 0,733; 0,601; 0,6; 0,355; 0,544\}. \quad (13)$$

На заключительном шаге поиска рациональной многокритериальной альтернативы качественного информационного поиска с максимальной степенью недоминируемости определяется множество недоминирующих альтернатив как пересечение множеств $\mu_{\tilde{Z}_1}^{n-d}$ и $\mu_{\tilde{Z}_2}^{n-d}$:

$$\begin{aligned} \mu_{\tilde{Z}_1}^{n-d} \cap \mu_{\tilde{Z}_2}^{n-d} = \{(1; 1; 1; 1; 1; 1) \cap (0,998; 0,733; 0,601; 0,6; 0,355; 0,544)\} = \\ = \{(0,998; 0,733; 0,601; 0,6; 0,355; 0,544)\}. \end{aligned} \quad (14)$$

Иными словами, благодаря применению метода нечеткого отношения предпочтения, выяснен тот факт, что в рамках решения задачи комплексного (оптимального компромиссного сравнения и сочетания альтернативных категорий «важности» для различных показателей качества) информационного поиска на ресурсах современных ЦОД, рациональным выбором следует считать альтернативу $\{s_1\}$. Именно данная альтернатива получила максимальную степень недоминируемости наиболее близкую (в идеале – равную) к единице.

Физический смысл данного результата для нашего примера состоит в том, что рациональный выбор предпочтений при информационном поиске на ресурсах современных ЦОД с точки зрения показателя качества реализации поисковых запросов $y_1(k) = y_{\text{рел}}(k)$ – релевантности поиска (уровня соответствия данного запроса текущим требованиям ЛППР), должен быть ориентирован на альтернативу $\{s_1\}$ – потенциальная релевантность поиска данных при мультимедийных запросах, т. е. реализация мультимедийного запроса потенциально будет иметь в нашем случае наибольшую

степень соответствия данного запроса текущим требованиям лица, принимающего решения. Это будет потенциально самый релевантный результат информационного поиска.

Предлагаемая совокупность подходов к реализации процедур информационного поиска на ресурсах современных ЦОД, основанная на методах нечетких отношений предпочтения, методах анализа и сравнения альтернатив, позволяет осуществить переход к комплексной и оптимальной реализации поисковых запросов и в конечном итоге к повышению качества поисковых алгоритмов в целом.

Содержание этапов и обзор потенциальных результатов разработки методики повышения качества информационного поиска на ресурсах современных ЦОД с использованием нечетких отношений предпочтения и сравнения альтернатив

Таким образом, разнообразие и разноплановость типов запросов ЛПР, а также связанные с этим трудности единого (комплексного) целеуказания для поисковых алгоритмов, зачастую присутствующая неопределенность (нечеткость) формулировки целей информационного поиска, многокритериальность оценки и контроля качества реализации поисковых запросов предопределили необходимость поиска новых подходов к решению задачи повышения качества информационного поиска на ресурсах современных ЦОД.

Эффективному решению подобных задач могут, по нашему мнению, помочь современные методологические подходы, основанные на нечетких отношениях предпочтения и на сравнении альтернатив.

В этой связи в рамках решения подобных задач предлагается рассмотреть научные вопросы и осуществить ряд этапов исследования:

1. Обобщение опыта создания и применения, а также синтез оптимальной системы показателей качества информационного поиска на ресурсах современных ЦОД, обобщение опыта создания и применения современных и перспективных методов реализации многокритериальных поисковых запросов.

2. Анализ современных и разработку новых методов повышения качества реализации информационного поиска на ресурсах современных ЦОД на основе математического моделирования процессов (процедур) реализации поисковых запросов.

3. Разработку методики оценивания показателей качества информационного поиска на ресурсах современных ЦОД с использованием методов теории нечетких множеств (нечетких отношений предпочтения) и сравнения альтернатив [20, 21].

4. Создание методики повышения качества информационного поиска на ресурсах современных ЦОД с использованием нечетких отношений предпочтения и сравнения альтернатив (с учетом полученных оценок

– результатов автоматизированного анализа качества поиска с использованием нечетких отношений предпочтения), с учетом требований по релевантности и оперативности этого поиска, а также с учетом возможного изменения характеристик поисковых систем современных ЦОД при воздействии на них деструктивных факторов.

5. Синтез частных алгоритмов адаптивного управления качеством информационного поиска на ресурсах современных ЦОД с использованием нечетких отношений предпочтения и сравнения альтернатив.

6. Разработку научно-технических предложений по практической реализации методов, алгоритмов и средств управления качеством информационного поиска на ресурсах современных ЦОД с использованием нечетких отношений предпочтения и сравнения альтернатив.

Решение этих научных задач, на наш взгляд, позволит получить ряд новых теоретических и практических результатов, к основным из которых следует отнести:

1. Новые теоретические аспекты и методы применения существующих и перспективных механизмов оценки и управления качеством информационного поиска на ресурсах современных ЦОД, основанные на использовании нечетких отношений предпочтения и сравнения альтернатив и позволяющие повысить оперативность и релевантность реализации поисковых запросов ЛПР с учетом возможных изменений характеристик поисковых систем дата-центров при воздействии на них деструктивных факторов в различных условиях обстановки:

– теоретическое обобщение известных методов контроля и управления качеством информационного поиска;

– алгоритмы синтеза (порядок формирования) системы показателей, отражающих качество информационного поиска;

– математические модели и алгоритмы, используемые для описания процессов, протекающих в поисковых системах современных ЦОД при реализации поисковых запросов, ограничения их применения при описании процессов, происходящих в системах такого класса;

– теоретические основы и методы исследования алгоритмов анализа качества информационного поиска на ресурсах современных ЦОД;

– методологические основы оценки и управления качеством информационного поиска на ресурсах современных ЦОД с использованием нечетких отношений предпочтения и сравнения альтернатив: общие принципы управления качеством информационного поиска; методика текущего и прогностического оценивания показателей качества реализации поисковых запросов;

– методы оценки эффективности механизмов управления качеством информационного поиска на

ресурсах современных ЦОД с использованием нечетких отношений предпочтения и сравнения альтернатив.

2. Алгоритмическая структура управления качеством информационного поиска на ресурсах современных ЦОД с использованием нечетких отношений предпочтения и сравнения альтернатив, а также частные алгоритмы решения задач классификации и быстрого нахождения необходимых пользователям данных.

3. Предложения по технической реализации системы управления качеством информационного поиска на ресурсах современных ЦОД с использованием нечетких отношений предпочтения и сравнения альтернатив.

Заключение

Таким образом, исследованы теоретическая возможность и перспективы практического применения методов нечетких отношений предпочтения и сравнения альтернатив для устранения неопределенности (нечеткости) исходных данных в рамках единой методики повышения качества информационного поиска на ресурсах современных ЦОД.

Сформулированы общие подходы и намечены конкретные пути решения актуальных задач, нацеленных на повышение качества реализации поисковых запросов ЛППР. Рассмотренные новые подходы и практическая реализация предложенных этапов оценки и управления информационным поиском с использованием нечетких отношений предпочтения и сравнения альтернатив позволяют повысить качество информационного обеспечения должностных лиц органов управления исполнительной власти и различных ведомств нашей страны, использующих системы хранения данных и дата-центры большой размерности.

Предложены и обоснованы физическая и математическая сущность, а также общее содержание этапов методики повышения качества информационного поиска на ресурсах современных ЦОД в условиях априорной неопределенности (нечеткости) исходных данных, устраняемой с использованием методов нечетких отношений предпочтения и сравнения альтернатив.

Практическое применение предложенной методики позволит преодолеть трудности единого (комплексного) целеуказания для поисковых алгоритмов, позволит компенсировать неопределенность (нечеткость) формулировки поисковых «меток» и текущих задач реализации информационного поиска. Это, в свою очередь, позволит повысить качество (релевантность и оперативность результатов) реализации информационного поиска по запросам лиц, принимающих решения, повысить качество их информационного обеспечения с учетом возможных изменений характеристик поисковых систем ЦОД при воздействии на них деструктивных факторов.

Литература

1. Олейник, А. И. ИТ-инфраструктура : учебно-методическое пособие / А.И. Олейник, А.В. Сизов. – Москва : Изд. дом Высшей школы экономики, 2012. – 134 с.
2. Емельянов, В. А. ИТ-инфраструктура организации : учебное пособие / В.А. Емельянов. – Москва : КноРус, 2022. – 146 с.
3. Докучаев, В. А. Архитектура центров обработки данных / В.А. Докучаев, А.А. Кальфа, В.В. Маклачкова. – Москва : Горячая линия-Телеком, 2020. – 240 с.
4. Паращук, И. Б. Особенности построения и анализа качества дата-центров как базовых элементов ИТ-инфраструктуры / И.Б. Паращук, Н.В. Михайличенко // Перспективные направления развития отечественных информационных технологий : материалы IV межрегиональной научно-практической конференции (Севастополь, 18–22 сентября 2018 г.). – Севастополь : СевГУ, 2018. – С. 28–29.
5. Прохоров, А. Н. Центры обработки данных: анализ, тренды, мировой опыт : корпоративное издание / А.Н. Прохоров, С.А. Рахматуллин ; научное редактирование : К. Королев, И. Дорофеев. – Москва : АльянсПринт, 2021. – 414 с.
6. Паращук, И. Б. Многопараметрические системы хранения данных, дата-центры и электронные библиотеки: способ контроля параметров технического состояния и анализа качества / И.Б. Паращук, Н.В. Михайличенко, Е.С. Крюкова // Региональная информатика (РИ-2020). XVII Санкт-Петербургская международная конференция : материалы конференции (Санкт-Петербург, 28–30 октября 2020 г.). – Т. 1. – Санкт-Петербург : Региональная общественная организация «Санкт-Петербургское Общество информатики, вычислительной техники, систем связи и управления», 2020. – С. 102–104.
7. Тарнбулл, Д. Релевантный поиск с использованием Elasticsearch и Solr / Д. Тарнбулл, Д. Берримен. – Москва : ДМК Пресс, 2017. – 410 с.
8. Чернов, В. Г. Нечеткие множества. Основы теории и применения : учеб. пособие / В.Г. Чернов. – Владимир : Изд-во ВлГУ, 2018. – 156 с.
9. Паращук, И. Б. Нечеткие множества в задачах анализа сетей связи / И.Б. Паращук, И.П. Бобрик. – Санкт-Петербург : ВУС, 2001. – 80 с.
10. Моделирование и проектирование систем. Учебник / И.Б. Саенко, И.Б. Паращук, Ю.В. Кондрашов [и др.]. – Санкт-Петербург : ВАС, 2023. – 472 с.
11. Винокуров, А. С. Использование метода нечеткого отношения предпочтения для принятия оптимального решения по выбору цифрового фотоаппарата / А.С. Винокуров, И.В. Белов, Р.И. Баженов // Современная техника и технологии. – 2014. – №11. – URL: <https://technology.snauka.ru/2014/11/4868> (дата обращения: 25.11.2023).
12. Белов, В. В. Повышение пертинентности поиска в современных информационных средах / В.В. Белов, А.А. Терехов, В.И. Чистякова. – Москва : Горячая линия-Телеком, 2012. – 158 с.
13. Крюкова, Е. С. Основные направления оценки качества поисковых информационно-справочных систем и анализ

перспектив их применения для управления и технического обеспечения электронных библиотек / Е.С. Крюкова, И.Б. Паращук // V Межвузовская научно-практическая конференция «Проблемы технического обеспечения войск в современных условиях» : труды конференции (Санкт-Петербург, 20 марта 2020 г.). – Санкт-Петербург : ВАС, 2020. – С. 179–183.

14. Миронов, А. Л. Соотношение релевантности и pertinентности поисковой выдачи / А.Л. Миронов, Г.В. Поздышева // Актуальные проблемы агроинженерии в XXI веке : материалы Национальной научно-практической конференции с международным участием (п. Майский, 1 декабря 2022 г.). – ФГБОУ ВО Белгородский ГАУ, 2023. – С. 257–260.

15. Розенберг, И. Н. Комплексность информационного поиска / И.Н. Розенберг // Образовательные ресурсы и технологии. – 2017. – № 1 (18). – С. 41–48.

16. Рыжов, А. П. Модели поиска информации средствами теории нечетких множеств : монография / А.П. Рыжов. – Москва : Изд-во МГУ им. М.В. Ломоносова, 2004. – 96 с.

17. Орловский, С. А. Проблемы принятия решений при нечеткой исходной информации / С.А. Орловский. – Москва : Наука, 1986. – 208 с.

18. Пронина, В. А. Семантический текстовый поиск, основанный на теории нечетких множеств / В.А. Пронина, Л.А. Панкова // Искусственный интеллект и принятие решений. – 2013. – № 3. – С. 19–23.

19. Целых, А. Н. Метод информационного поиска на основе нечеткого сходства ситуаций / А.Н. Целых, Э.М. Котов, А.А. Целых // Известия Южного федерального университета. Технические науки. – 2014. – № 6 (155). – С. 74–78.

20. Трушина, В. П. Информационные технологии многокритериального выбора альтернатив на основе нечеткого отношения предпочтения / В.П. Трушина, А.Л. Осипов // Современные инструментальные системы, информационные технологии и инновации : сборник научных трудов XII Международной научно-практической конференции в 4-х томах (Курск 19–20 марта 2015 г.). – Курск : Университетская книга, 2015. – Т. 4. – С. 150–153.

21. Kotenko, I. V. Analysis of Threats to Information Security of Industrial Automation Systems Using Euclidean and Hamming Distances between Fuzzy Sets / I.V. Kotenko, I.B. Parashchuk // 2023 International Russian Automation Conference (RusAutoCon-2023): IEEE Xplore Digital Library: Browse Conferences, (Sochi, 10–16 September 2023). – 2023. – Vol. 10272922. – P. 13–18.