

А. Н. Алфимцев, В. В. Девятков,  
С. А. Сакулин

## ПЕРСОНАЛИЗАЦИЯ В ГИПЕРТЕКСТОВЫХ СЕТЯХ НА ОСНОВЕ РАСПОЗНАВАНИЯ ДЕЙСТВИЙ ПОЛЬЗОВАТЕЛЕЙ И НЕЧЕТКОГО АГРЕГИРОВАНИЯ

*Рассмотрен метод персонализации в гипертекстовых сетях, основанный на распознавании действий пользователей и нечетком агрегировании. Достоинствами метода являются возможность использования двух типов нечеткого интеграла без необходимости экспертного задания нечетких мер, а также охват всех этапов персонализации, начиная с агрегирования параметров запроса одиночного пользователя и заканчивая агрегированием индивидуальных профилей пользователей в единый параметр группы пользователей.*  
**E-mail: [alfim@bmstu.ru](mailto:alfim@bmstu.ru); [deviatkov@iu3.bmstu.ru](mailto:deviatkov@iu3.bmstu.ru);  
[ss141291@yandex.ru](mailto:ss141291@yandex.ru)**

**Ключевые слова:** персонализация в Веб, нечеткое агрегирование, нечеткий интеграл Шоке, нечеткий интеграл Сугено, нечеткая мера.

История гипертекстовых сетей (Веб) в некотором смысле началась с эссе Ванневары Буша “As We May Think” (Как мы можем думать), опубликованного им в 1945 г. в журнале “Atlantic Monthly” [1]. В этом эссе Буш описывал гипотетическую машину “Мемекс”, работающую на основе принципов гипертекста.

В настоящее время Веб находится в стадии экспоненциального роста, и это явление превратилось в нечто большее, чем просто гипертекст. Веб стал приобретать элементы интеллектуальности. Одним из таких элементов является адаптация под конкретного пользователя или персонализация. Например, поисковая машина Google поддерживает персонализацию поиска на основе истории запросов конкретного пользователя [2]. Торговля и реклама в Веб также становятся все более индивидуальными, направленными непосредственно на каждого отдельного пользователя (покупателя) [3]. Персональный подход в области безопасности позволит интернет-провайдерам или владельцам ресурсов в Веб договариваться со своими клиентами о приемлемой в каждом отдельном случае политике безопасности [4]. Разработчики интерфейсов также начали принимать во внимание индивидуальные особенности пользователей путем построения персональной модели, основанной на возможностях индивидуума [5].

На сегодняшний день основными методами, применяемыми в целях персонализации Веб, являются различные методы извлечения данных, которые используют алгоритмы кластеризации. Кластеризация широко используется в таких областях, как машинное обучение, поиск информации и т.д. В соответствии с данными работы [6], проблема

идентификации групп пользователей по своей природе опирается на использование методов кластеризации. Кластеризация данных может быть также использована для генерации профилей пользователей на основе информации о действиях каждого пользователя, а затем для формирования групп пользователей на основе их профилей. Однако современные методы кластеризации не принимают во внимание многие факторы и не пригодны для успешного использования во множестве реальных задач. Поэтому для решения сложных практических задач применяют методы распознавания в совокупности с интенсивным машинным обучением.

Адаптация под конкретного пользователя — достаточно сложная задача, поскольку для ее успешного решения необходимо принимать во внимание присущие человеку неопределенность, спонтанность и даже противоречивость, свойственные ему от природы. Множество неопределенностей различных типов присутствует в глобальной сети: например, Веб-страницы появляются и исчезают, ключевые слова в поиске могут иметь различное значение в зависимости от контекста и т.п. Поэтому поиск в сети, по сути, является нечетким. Для того чтобы работать с этими неопределенностями с использованием компьютера, они должны быть формализованы тем или иным способом. Нечеткая логика позволяет принимать во внимание различные типы неопределенностей. Нечеткая кластеризация профилей пользователей может применяться для создания нечетких правил и заключений для модификации запросов. Результаты этой кластеризации могут применяться для извлечения знаний из профилей пользователей, например в целях рекламы [7].

Однако простое применение нечеткой логики для персонализации в Веб может не принести желаемых результатов, поскольку при этом реальная картина предпочтений пользователя значительно упрощается. Одним из таких упрощений является применение для агрегирования нескольких критериев средневзвешенного оператора, когда каждому критерию ставится в соответствие вес, в то время как не учитывается явление взаимодействия между критериями [8]. Применение в качестве операторов агрегирования нечетких интегралов (операторов) Сугено и Шоке по отношению к нечеткой мере позволяет учитывать такое взаимодействие [9], но в этом случае возникает проблема идентификации нечетких мер. Эта проблема связана не только с экспоненциально возрастающей сложностью нечетких мер, зависящей от числа критериев, но и с отсутствием ясного понимания сути концепции нечеткой меры у большинства практических специалистов [10]. Обычно эту проблему решают с помощью значительного упрощения модели и применения различных вспомогательных средств для идентификации

[11, 12] либо с помощью методов идентификации, предусматривающих отсутствие эксперта. Например, в работе [13] описано применение нечеткого интеграла Сугено с идентификацией меры без эксперта для агрегирования результатов поиска, полученных с помощью различных поисковых машин в Веб.

В настоящей статье рассмотрен подход к персонализации в гипертекстовых сетях, базирующийся на методе распознавания на основе агрегирования источников информации с помощью интегралов Сугено и Шоке. Отличие этого подхода от существующих состоит в возможности применения обоих типов интегралов без необходимости экспертного задания нечетких мер (что практически невозможно в реальных условиях), а также в том, что этот подход охватывает все стадии персонализации, начиная от агрегирования параметров запросов отдельного пользователя и заканчивая агрегированием параметров профилей отдельных пользователей в единый параметр группы пользователей.

**Нечеткие меры и интегралы.** Нечеткие интегралы, или нечеткие операторы агрегирования, позволяющие учесть взаимозависимость функций принадлежности, используют нечеткую меру. Нечеткой мерой называется функция  $g : 2^R \rightarrow [0, 1]$ , где  $R$  — множество каких-либо параметров, характеризующих некоторый объект. Нечеткая мера  $g(Q_i)$  характеризует совокупную значимость параметров, входящих в множество  $Q_i$ . Нечеткая мера удовлетворяет ряду условий [14]: в частности,  $g(\emptyset) = 0$ ,  $g(Y) = 1$ ; если  $Q, P \in Y$  и  $Q \subset P$ , то  $g(Q) \leq g(P)$ .

Если  $R$  есть множество всех подмножеств множества источников информации (о каком-либо пользователе или о группе пользователей)  $Y = \{Y_1, \dots, Y_m\}$ , то операторы (интегралы) агрегирования могут быть записаны следующим образом.

Нечеткий оператор Сугено

$$A_k = A_k^C = \max_{i=1}^{i=m} [\min(\mu_i^k, g(Q_i))], \quad (1)$$

где  $\mu_1^k(y_1) \geq \mu_2^k(y_2) \geq \dots \mu_m^k(y_m)$ ,  $Q_i = \{Y_1, \dots, Y_i\}$ ,  $i = 1, \dots, m$ .

Нечеткий оператор Шоке

$$A_k = A_k^{\text{Ш}} = \sum_{i=1}^{i=m} [\mu_i^k(y_i) - \mu_{i+1}^k(y_{i+1})] g(Q_i), \quad (2)$$

где  $\mu_1^k(y_1) \geq \mu_2^k(y_2) \geq \dots \mu_m^k(y_m)$ ,  $Q_i = \{Y_1, \dots, Y_i\}$ ,  $i = 1, \dots, m$ ,  $\mu_{m+1}^k(y_{m+1}) = 0$ .

Нечеткий оператор Шоке обычно интерпретируется как обобщение понятия взвешенного среднего арифметического, а оператор Сугено —

как обобщение концепции взвешенной медианы (при агрегировании не менее трех источников информации).

Вследствие простоты наиболее распространены методы вычисления нечеткой меры, основанные на понятии  $g_\lambda$ -нечеткой меры, введенной Сугено [14]. Нечеткая мера называется  $g_\lambda$ -нечеткой мерой, если для нее справедливо условие: для всех  $Q, P \subset Y$  таких, что  $Q \cap P = \emptyset$ , имеет место  $g(Q \cup P) = g(Q) + g(P) + \lambda g(Q)g(P)$  для некоторого  $\lambda > -1$ .

Рассмотрим процедуру наиболее популярного метода вычисления  $g_\lambda$ -нечеткой меры [9, 14–16], обозначая ее по-прежнему просто  $g$ .

**Шаг 1.** Для каждого источника информации  $Y_i, i = 1, \dots, m$ , выбрать значение нечеткой меры  $g(Y_i) \in [0, 1]$  как степень важности этого источника  $Y_i$ . Значения  $g(Y_i)$  могут быть установлены экспертом, получены в результате наблюдений или каким-либо другим путем.

**Шаг 2.** Найти значение  $\lambda$ , используя уравнение

$$\lambda + 1 = \prod_{i=1}^m (1 + \lambda g(Y_i)). \quad (3)$$

**Шаг 3.** Для всех  $Q_i = \{Y_1, \dots, Y_i\}, i = 1, \dots, m$ , вычислить рекурсивно нечеткие меры  $g(Q_i)$ , используя следующие выражения:

$$g(Q_1) = g(Y_1), \quad g(Q_i) = g(Y_i) + g(Q_{i-1}) + \lambda g(Y_i)g(Q_{i-1}), \quad i = 2, \dots, m. \quad (4)$$

**Агрегирование на основе нечетких интегралов.** Прежде чем переходить к рассмотрению общей процедуры нечеткого агрегирования, рассмотрим процедуру формирования множества  $Y_i$  и процедуру распознавания отдельным алгоритмом  $i$  с помощью функции  $\mu(y_{ij_i})$ . В общем случае исходными для агрегирования являются  $i$  алгоритмов,  $i = 1, \dots, m$ , использующих источники скрытой информации. В данном случае источники скрытой информации и способы их агрегирования не рассматриваются, интерес представляют только результаты работы каждого из этих алгоритмов как источников новой отдельной информации (сигнала)  $Y_i, i = 0, \dots, m$ , и функции принадлежности  $\mu(y_{ij_i}), y_{ij_i} \in Y_i, i = 0, \dots, m, j_i = 0, \dots, n_i$ .

Основной задачей процедуры является агрегирование источников информации  $Y_i, i = 0, \dots, m$ . Для того чтобы сформировать множество  $Y_i$  и функции принадлежности  $\mu(y_{ij_i}), y_{ij_i} \in Y_i, i = 0, \dots, m, j_i = 0, \dots, n_i$ , каждый алгоритм предварительно обрабатывается в соответствии со следующей процедурой.

*Процедура 1.*

**Шаг 1.** Задаются совокупность пустых множеств  $Y_i^k = \emptyset, k = 1, \dots, K$ .

**Шаг 2.** Для каждого эталонного объекта  $k, k = 1, \dots, K$ , с использованием источников скрытой информации формируется своя эталонная модель  $G_i^k, k = 1, \dots, K$ .

**Шаг 3.** Для распознаваемого объекта по тем же принципам формируется модель  $G$ .

**Шаг 4.** Модель сравнивается с каждой моделью  $G_i^k, k = 1, \dots, K$ , в результате чего вычисляется множество отсчетов  $\{y_i^1, y_i^2, \dots, y_i^K\}$ , характеризующих близость модели  $G$  соответственно к моделям  $G_i^k, k = 1, \dots, K$ .

**Шаг 5.** Формируются множества  $Y_i^k \cup y_i^k, k = 1, \dots, K$ , которые принимаются за новые множества  $Y_i^k$ . Если множества  $Y_i^k$  перестают изменяться, то осуществляется переход к шагу 6 (могут использоваться и другие критерии перехода к шагу 6). В противном случае процедура начинается с шага 2.

**Шаг 6.** Множества  $Y_i^k$  объединяются, в результате чего получается множество  $Y_i = \bigcup_{k=1}^K Y_i^k$ , которое упорядочивается (если оно числовое, то упорядочивание осуществляется по возрастанию) и его элементы индексируются  $i = 1, \dots, m, j_i = 0, \dots, n_i$ , в результате получается множество  $Y_i = \{y_{ij_i} \in Y_i | i = 1, \dots, m, j_i = 0, \dots, n_i\}$ . На множестве  $Y_i$  задается функция принадлежности  $\mu(y_{ij_i}), y_{ij_i} \in Y_i, i = 1, \dots, m, j_i = 0, \dots, n_i$ .

Распознавание по любому отдельному алгоритму  $i$  с помощью функции  $\mu(y_{ij_i})$  может быть осуществлено в соответствии со следующей процедурой.

*Процедура 2.*

**Шаг 0.** С помощью процедуры 1 осуществляется формирование множества  $Y_i$  и функции принадлежности  $\mu(y_{ij_i}), y_{ij_i} \in Y_i, i = 0, \dots, m, j_i = 0, \dots, n_i$ .

**Шаг 1.** Для каждого эталонного объекта  $k, k = 1, \dots, K$ , с использованием источников скрытой информации формируется своя эталонная модель  $G_i^k, k = 1, \dots, K$ .

**Шаг 2.** Для распознаваемого объекта по тем же принципам формируется модель  $G$ .

**Шаг 3.** Модель  $G$  сравнивается с каждой моделью  $G_i^k, k = 1, \dots, K$ , и в результате вычисляется множество отсчетов  $\{y_i^1, y_i^2, \dots, y_i^K\} \subset Y_i$ , характеризующих близость модели  $G$  соответственно к моделям  $G_i^k, k = 1, \dots, K$ .

**Шаг 4.** Модель  $G$  считается совпадающей с той эталонной моделью  $G_i^k$  для которой значение  $\mu(y_i^k)$ , где  $y_i^k \in Y_i$  максимально.

Таким образом, функция принадлежности  $\mu(y_i^k)$ , где  $y_i^k \in Y_i$ , оценивает близость распознаваемой модели к соответствующей эталонной.

Основной задачей является агрегирование источников информации  $Y_i$ ,  $i = 1, \dots, m$ , для повышения надежности распознавания действий пользователя.

Таким образом, общая процедура агрегирования источников информации с помощью операторов Сугено или Шоке будет следующей.

*Процедура 3.*

**Шаг 1.** Для каждого источника информации (сигнала)  $Y_i$ ,  $i = 1, \dots, m$  выбрать значение  $g(Y_i) \in [0, 1]$  как степень важности источника информации  $Y_i$ .

**Шаг 2.** Найти значение  $\lambda$ , используя уравнение (3).

**Шаг 3.** Для распознаваемого объекта по каждому алгоритму  $i = 1, \dots, m$  и для каждого  $k = 1, \dots, K$  вычислить множество функций принадлежности  $\mu(y_i^k)$ ,  $y_i^k \in Y_i$ ,  $i = 1, \dots, m$ , с помощью процедуры 2.

**Шаг 4.** Для каждого  $k = 1, \dots, K$  упорядочить множество функций  $\mu(y_i^k)$  таким образом, чтобы  $\mu(y_{j_1}^k) \geq \mu(y_{j_2}^k) \geq \dots \geq \mu(y_{j_m}^k)$ ,  $j_n \in \{1, \dots, m\}$ .

**Шаг 5.** Для каждого  $k = 1, \dots, K$  вычислить рекурсивно значения нечетких мер  $g(Q_i^k)$ , где  $Q_i^k = \{Y_{j_1}, \dots, Y_{j_i}\}$ ,  $i = 1, \dots, m$ , используя формулу (4).

**Шаг 6.** Вычислить для всех  $k = 1, \dots, K$  значения операторов  $A_k = A_k^C$  (или  $A_k = A_k^{III}$ ). Распознаваемый объект считается совпадающим с тем эталонным объектом, для которого значение  $A_k = A_k^C$  максимально.

**Метод нечеткого распознавания пользовательских профилей.** В общем случае каждая порция информации может содержать  $L$  объектов  $\theta_l$ ,  $l = 1, \dots, L$ , подлежащих распознаванию. Объекты различных множеств  $\Theta_l$  могут находиться в определенных в общем случае  $r$ -арных отношениях  $\Xi \in \Theta_{l_1} \times \Theta_{l_2} \times \dots \times \Theta_{l_r}$ ,  $\{l_1, l_2, \dots, l_r\} \subseteq \{1, \dots, L\}$ . Каждое такое отношение  $\Xi$  будем называть эталонным пользовательским профилем. По аналогии с распознаваемым объектом распознаваемый пользовательский профиль будем обозначать  $\xi = \langle \theta_{l_1}, \theta_{l_2}, \dots, \theta_{l_r} \rangle$ ,  $\{l_1, l_2, \dots, l_r\} \subseteq \{1, \dots, L\}$ , где  $\theta_{l_1}, \theta_{l_2}, \dots, \theta_{l_r}$  — распознаваемые объекты, а процесс установления ее сходства с каким-либо эталонным пользовательским профилем  $\Xi$  будем называть распознаванием профиля  $\xi$ . Сходство распознаваемого профиля  $\xi$  с эталонным профилем  $\Xi$ , характеризуемое ненулевым значением критерия сходства, будем записывать как  $\xi \approx \Xi$ . В случае если значение критерия схожести равно нулю, профиль  $\xi$  не схож с профилем  $\Xi$ . Это несходство записывается как  $\xi \neq \Xi$ .

Для вычисления сходства распознаваемого ( $\theta$ ) и эталонного ( $\Theta^j$ ) объектов используется оператор  $A^j[\mu_1(y_1), \mu_2(y_2), \dots, \mu_m(y_m)]$ . Аргументами этого оператора являются функции принадлежности  $\mu_1(y_1)$ ,

$\mu_2(y_2), \dots, \mu_m(y_m)$ , значения которых находятся в интервале  $[0, 1]$ . Значения оператора также находятся в интервале  $[0, 1]$ . Следовательно, оператор является функцией  $[0, 1]^m \rightarrow [0, 1]$ . Тогда если для каждого объекта распознаваемого профиля  $\xi = \langle \theta_{l_1}, \theta_{l_2}, \dots, \theta_{l_r} \rangle$  известно множество значений критериев  $A_{l_1}, A_{l_2}, \dots, A_{l_r}$  его сходства соответственно с объектами  $\Theta_{l_1}^{k_{l_1}}, \Theta_{l_2}^{k_{l_2}}, \dots, \Theta_{l_r}^{k_{l_r}}$  эталонного профиля, то, используя некоторый оператор агрегирования  $A$ , можно вычислить степень сходства распознаваемого профиля с эталонным как значение функции  $A[A_{l_1}, A_{l_2}, \dots, A_{l_r}]$ .

Процедура распознавания отдельного профиля  $\xi = \langle \theta_{l_1}, \theta_{l_2}, \dots, \theta_{l_r} \rangle$ ,  $\{l_1, l_2, \dots, l_r\} \subseteq \{1, \dots, L\}$ , использующая эту идею, будет выглядеть следующим образом.

**Шаг 1.** Каждый объект  $\theta_{l_1}, \theta_{l_2}, \dots, \theta_{l_r}$  распознается отдельно сопоставлением соответственно с эталонными объектами  $\Theta_{l_1}^{k_{l_1}}, \Theta_{l_2}^{k_{l_2}}, \dots, \Theta_{l_r}^{k_{l_r}}$ ,  $k_{l_r} = 1, \dots, K_{l_r}$ ,  $\{l_1, l_2, \dots, l_r\} \subseteq \{1, \dots, L\}$  с помощью операторов агрегирования  $A_{l_1}, A_{l_2}, \dots, A_{l_r}$ . Если для всех распознаваемых объектов  $\theta_{l_1}, \theta_{l_2}, \dots, \theta_{l_r}$  найдены сходные с ними эталонные объекты  $\Theta_{l_1}^{\tilde{k}_{l_1}}, \Theta_{l_2}^{\tilde{k}_{l_2}}, \dots, \Theta_{l_r}^{\tilde{k}_{l_r}}$ ,  $\tilde{k}_{l_r} = 1, \dots, K_{l_r}$ , т.е. такие, что  $\theta_{l_1} \approx \Theta_{l_1}^{\tilde{k}_{l_1}}$ ,  $\theta_{l_2} \approx \Theta_{l_2}^{\tilde{k}_{l_2}}, \dots, \theta_{l_r} \approx \Theta_{l_r}^{\tilde{k}_{l_r}}$ , то перейти к шагу 2. Если хотя бы для одного распознаваемого объекта  $\theta_{l_1}, \theta_{l_2}, \dots, \theta_{l_r}$  не найдено ни одного сходного эталонного, то перейти к шагу 3.

**Шаг 2.** Профиль  $\xi = \langle \theta_{l_1}, \theta_{l_2}, \dots, \theta_{l_r} \rangle$  считается распознанным и схожим с профилем  $\Xi = \Theta_{l_1}^{\tilde{k}_{l_1}}, \Theta_{l_2}^{\tilde{k}_{l_2}}, \dots, \Theta_{l_r}^{\tilde{k}_{l_r}}$ , а значение критерия схожести профиля  $\xi$  с профилем  $\Xi$  равно  $A \left[ a_{l_1}^{\tilde{k}_{l_1}}, a_{l_2}^{\tilde{k}_{l_2}}, \dots, a_{l_r}^{\tilde{k}_{l_r}} \right]$ .

**Шаг 3.** Профиль  $\xi = \langle \theta_{l_1}, \theta_{l_2}, \dots, \theta_{l_r} \rangle$  не распознан.

Профили  $\Xi \in \Theta_{l_1} \times \Theta_{l_2} \times \dots \times \Theta_{l_r}$ ,  $\{l_1, l_2, \dots, l_r\} \subseteq \{1, \dots, L\}$ , будем называть профилями 1-го уровня и обозначать их  $\Xi_1$ . Профилями  $s$ -го уровня будем называть профили  $\Xi_s \in \Xi_{s-1}^1 \times \Xi_{s-1}^2 \times \dots \times \Xi_{s-1}^\nu$ , где  $\Xi_{s-1}^1, \Xi_{s-1}^2, \dots, \Xi_{s-1}^\nu$  — профили  $(s-1)$ -го уровня. Таким образом, профили первого уровня являются отношениями объектов, а профили  $s$ -го уровня, где  $s > 1$ , являются отношениями профилей  $(s-1)$ -го уровня. Для того чтобы распознать профили  $(s-j)$ -го уровня,  $j = 0, 1, \dots, s-2$ , необходимо распознать профили  $(s-j-1)$ -го уровня, отношением которых являются профили  $(s-j)$ -го уровня. Если при распознавании какого-либо профиля  $(s-j)$ -го уровня окажется, что хотя бы один профиль  $(s-j-1)$ -го уровня, входящий в отношение этого профиля  $(s-j)$ -го уровня, не может быть распознан, то процесс распознавания последнего прекращается.

В основу метода распознавания профилей  $s$ -го уровня  $\Xi_s \in \Xi_{s-1}^1 \times \Xi_{s-1}^2 \times \dots \times \Xi_{s-1}^\nu$  может быть положено развитие процедуры распознавания профилей первого уровня следующим образом.

**Шаг 1.** Каждый объект  $\theta_{l_1}, \theta_{l_2}, \dots, \theta_{l_r}$ , входящий хотя бы в один профиль первого уровня  $\Xi_1$ , распознается отдельно сопоставлением соответственно с эталонными объектами  $\Theta_{l_1}^{k_{l_1}}, \Theta_{l_2}^{k_{l_2}}, \dots, \Theta_{l_r}^{k_{l_r}}$ ,  $k_{l_r} = 1, \dots, K_{l_r}$ ,  $\{l_1, l_2, \dots, l_r\} \subseteq \{1, \dots, L\}$  с помощью операторов агрегирования  $A_{l_1}, A_{l_2}, \dots, A_{l_r}$ .

**Шаг 2.** Каждый профиль 1-го уровня  $\Xi_1$ , для всех объектов которого найдены сходные с ними эталонные объекты, считается распознанным и для него вычисляется критерий схожести (значение оператора агрегирования)  $A_1$ . После этого осуществляется переход к шагу 3. Если такие профили не найдены, то распознанных профилей 1-го уровня и более не существует и выполнение процедуры прекращается.

**Шаг 3.** Задается значение уровня  $s = 2$  и осуществляется переход к шагу 4.

**Шаг 4.** Если найдены профили  $s$ -го уровня  $\Xi_s$  для всех профилей  $(s - 1)$ -го уровня, для которых найдены ненулевые значения критериев схожести, то эти профили  $\Xi_s$  считаются распознанными, для них вычисляются критерии схожести (значение оператора агрегирования)  $A_s$ . Если существуют профили уровня  $s + 1$ , то шаг 4 снова выполняется со значением  $s = s + 1$ . В противном случае выполнение прекращается.

Если нет профилей  $\Xi_s$ -го уровня для всех профилей  $(s - 1)$ -уровня, для которых найдены ненулевые значения критериев схожести, то распознанных профилей  $s$ -го уровня не существует и выполнение процедуры прекращается.

Эксперименты по нечеткому агрегированию запросов и по нечеткому распознаванию профилей пользователей показали, что релевантность результатов поиска улучшилась на 15 % по сравнению с существующим методом на статистически значимой выборке из 1500 поисковых запросов от более чем 30 пользователей.

**Заключение.** Рассмотрен метод персонализации в гипертекстовых сетях на основе нечеткого агрегирования и распознавания с помощью нечетких мер и интегралов. Применение предлагаемого метода может быть полезно в случае, если нет возможности применить другой способ персонализации ресурсов, например развернуть полную онтологию или семантическую сеть организации, но создано множество инструментов для электронной коммуникации и взаимодействия. Основные отличия и преимущества этого метода от аналогов заключаются в следующем.

Во-первых, появляется возможность принимать во внимание как меру значимости каждого источника информации в отдельности, так и меру значимости каждого подмножества источников информации в процессе иерархического распознавания за счет использования для агрегирования нечетких интегралов по нечеткой мере.



Во-вторых, повышается точность распознавания профилей пользователей за счет охвата всех стадий распознавания, начиная от агрегирования запросов одного пользователя и заканчивая агрегированием профилей отдельных пользователей в профиль группы пользователей.

В-третьих, появляется перспектива развития интеллектуального и интуитивного взаимодействия пользователя с компьютером, поскольку при этом задействуются скрытые источники информации о конкретном пользователе.

## СПИСОК ЛИТЕРАТУРЫ

1. B u s h V., As we may think // The Atlantic Monthly. – 1945. – Vol 1. – P. 101–108.
2. S u l l i v a n D. Google now personalizes everyone’s search results // <http://searchengineland.com/google-now-personalizes-everyones-search-results>. 2009. 15 p. (дата обращения 13.11.2011).
3. C h i u W. Web site personalization // <http://www.ibm.com/developerworks/Web-sphere/library/techarticles/hipods/personalize.html> 2001. 23 p. (дата обращения 13.11.2011).
4. Y e e G. O., K o r b a L. Security Personalization for Internet and Web Services // International Journal of Web Services Research (IJWSR). – 2008. – No. 5(1). – P. 1–23.
5. C a s a s R., B l a s c o M a r I. R., R o b i n e t A., D e l g a d o A., Y a r z a A., M c g i n n J., P i c k i n g R. User modelling in ambient intelligence for elderly and disabled people // <http://www.springerlink.com/index/5085668312150127.pdf>. – 2008. – 8 p. (дата обращения 13.01.2012).
6. М а р м а н и с X., Б а б е н к о Д. Алгоритмы интеллектуального интернета. – СПб.-М.: Символ, 2011. – 466 с.
7. L o i a V., N i k r a v e s h M., Z a d e h L. A. Fuzzy logic and the Internet. – Berlin.: Springer-Verlag, 2010. – 337 p.
8. M a r i c h a l J. -L. An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria // <http://ikojadin.perso.univ-pau.fr/kappalab/pub/MarITFS2000.pdf>. – 2000. – P. 800–807 (дата обращения 13.01.2012).
9. G r a b i s c h M., R o u b e n s M. Application of the Choquet Integral in multicriteria decision making // Fuzzy Measures and Integrals – Theory and Applications. – Physica Verlag, 2000. – P. 415–434.
10. G r a b i s c h M. The application of fuzzy integrals in multicriteria decision making // European Journal of Operational Research. – 1996. – Vol. 89 (3). – P. 445–456.
11. T a k a h a g i E. On Identification methods of  $\lambda$ -fuzzy measures using weights and  $\lambda$  // Japanese Journal of Fuzzy Sets and Systems. – 2000. – Vol. 12 (5). – P. 665–676.
12. С а к у л и н С. А. Визуализация оператора агрегирования на основе интеграла Шоке по нечеткой мере 2-го порядка // Вестник ИРГТУ. – 2007. – Т. 2, № 2 (30). – С. 45–50.
13. C u i S., F e n g B. A. Fuzzy integral method to merge search engine results on Web: Lecture notes in computer science. Computational intelligence and security // [http://dx.doi.org/10.1007/11596981\\_107](http://dx.doi.org/10.1007/11596981_107). – 2005. – P. 731–736. (дата обращения 13.01.2012).
14. А в е р к и н А. Н., Б а т ы р ш и н И. З., Б л и ш у н А. Ф. Нечеткие множества в моделях управления и искусственного интеллекта. – М.: Наука, 1986. – 312 с.

15. Алфимцев А. Н. Нечеткое агрегирование мультимодальной информации в интеллектуальном интерфейсе // Программные продукты и системы. – 2011. – № 3. – С. 44–48.
16. Marichal J. On Choquet and Sugeno integrals as aggregation functions // Fuzzy Measures and Integrals. – 2000. – Vol. 40. – P. 247–272.

Статья поступила в редакцию 10.04.2012

Александр Николаевич Алфимцев родился в 1983 г., окончил МГТУ им. Н.Э. Баумана в 2005 г. Канд. техн. наук, доцент кафедры “Информационные системы и телекоммуникации” МГТУ им. Н.Э. Баумана. Автор 40 научных работ, в том числе трех патентов на изобретения, в области методов искусственного интеллекта, распознавания образов, компьютерного зрения.

A.N. Alfimtsev graduated from the Bauman Moscow State Technical University in 2005. Ph. D. (Eng.), assoc. professor of “Information Systems and Telecommunications” department of the Bauman Moscow State Technical University. Author of 40 publications, including three patents for inventions in the field of artificial intelligence methods, pattern recognition, computer vision.

Сергей Александрович Сакулин родился в 1976 г. Окончил МГТУ им. Н.Э. Баумана в 2001 г. Канд. техн. наук, инженер кафедры “Информационные системы и телекоммуникации” МГТУ им. Н.Э. Баумана. Автор восьми научных работ в области методов искусственного интеллекта, распознавания образов, формализации и визуализации экспертных знаний.

S.A. Sakulin graduated from the Bauman Moscow State Technical University in 2001. Ph. D. (Eng.), engineer of “Information Systems and Telecommunications” department of the Bauman Moscow State Technical University. Author of eight publications in the field of artificial intelligence methods, pattern recognition, expert knowledge formalization and visualization.

Владимир Валентинович Девятков окончил Ленинградский Государственный институт точной механики и оптики в 1963 г. Д-р техн. наук, профессор, заведующий кафедрой “Информационные системы и телекоммуникации” МГТУ им. Н.Э. Баумана. Автор более 120 научных работ в области систем искусственного интеллекта, мультиагентных систем, распознавания образов.

V.V. Devyatkov graduated from the Leningrad Institute of Precise Mechanics and Optics in 1963. D. Sc. (Eng.), professor, head of “Information Systems and Telecommunications” department of the Bauman Moscow State Technical University. Author of more than 120 publications in the field of artificial intelligence systems, multiagent systems, pattern recognition.