

ФОРМАЛИЗАЦИЯ ДАННЫХ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА НА ОСНОВЕ НЕЧЕТКИХ МНОЖЕСТВ РАЗНЫХ ТИПОВ

О.М. Полещук

ФГАОУ ВО «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)» (Мытищинский филиал), Россия, 141005, Московская обл., г. Мытищи, ул. 1-я Институтская, д. 1

poleshhukom@bmstu.ru

Разработаны модели формализации данных образовательного процесса в условиях различной исходной информации, для создания которых были использованы нечеткие множества первого типа, интервальные нечеткие множества второго типа и Z -числа. По всем построенным моделям даны анализ и рекомендации для использования их в практических целях. Нечеткие множества первого типа рекомендуется использовать для формализации статистических данных образовательного процесса, а также данных, полученных от единственного эксперта (экзаменатора). Интервальные нечеткие множества второго типа рекомендуется использовать для формализации статистических и экспертных данных образовательного процесса со случайными ошибками, а также данных, полученных от группы экспертов. Z -числа рекомендуется использовать для формализации данных образовательного процесса с учетом их достоверности. Приведенные числовые примеры в совокупности с теоретическими обоснованиями предоставляют возможность выбора модели для дальнейшего анализа данных в целях получения устойчивых конечных результатов и управляющих решений на их основе.

Ключевые слова: образовательный процесс, формализация данных, нечеткое множество, Z -числа

Ссылка для цитирования: Полещук О.М. Формализация данных образовательного процесса на основе нечетких множеств разных типов // Лесной вестник / Forestry Bulletin, 2026. Т. 30. № 1. С. 138–146. DOI: 10.18698/2542-1468-2026-1-138-146

Теория нечетких множеств показала свою эффективность при обработке информации образовательного процесса, поскольку позволяет принимать во внимание неопределенность в виде нечеткости [1]. Теория вероятностей учитывает случайность, но не неопределенность, которая возникает вследствие активного участия человеческого фактора в процессах обучения и управления и влияния на них. Теория нечетких множеств существенно обогащает аппарат обработки данных по образовательному процессу [2–4].

Лингвистические переменные, применяемые для формализации шкал, предназначенных для оценки качественных характеристик образовательного процесса, полностью ликвидировали некорректность оперирования элементами порядковых шкал, которые традиционно использовались ранее [4, 5]. Изучение свойств лингвистических переменных позволило выбрать такое их подмножество для моделирования реальных процессов, которое снизило риск ошибок в задачах принятия решений [5, 6] и дало импульс развитию методов нечеткого

анализа данных, в частности методов нечеткого кластерного анализа [7–12] и методов нечеткого регрессионного анализа [13–16].

Развитие теории нечетких множеств привело к появлению нечетких множеств второго типа [17]. Отличие нечетких множеств второго типа от нечетких множеств первого типа [1] состоит в том, что значениями функций принадлежности нечетких множеств первого типа являются числа, а значениями функций принадлежности нечетких множеств второго типа являются нечеткие множества первого типа.

С одной стороны, использование нечетких множеств второго типа позволяет сохранить больше исходной информации, поступающей из разных источников, чем использование нечетких чисел первого типа. Например, не усреднить экспертную информацию, а реально учесть мнения и опыт каждого эксперта в принятии группового решения. С другой стороны, оперировать нечеткими множествами второго типа достаточно сложно [18], поэтому для практических приложений стали использовать в основном интервальные нечеткие множества второго типа [18], значениями функции принадлежности которых являются интервалы. Применение интервальных нечетких множеств

второго типа для обработки информации образовательного процесса позволяет повысить достоверность поступающей информации и тем самым снизить риски ошибок при разработке управляющих воздействий по полученным результатам и выводам [19].

Важным этапом развития теории нечетких множеств стало появление Z -чисел [20]. В отличие от нечетких множеств второго типа, которые каждое значение универсального множества оценивают интервалом значений функции принадлежности, Z -числа оценивают все нечеткое множество целиком, причем в общем случае этой оценкой является нечеткое множество, а в частном — интервал.

Для того, чтобы использовать Z -числа, от эксперта требуется добавить к оценке свою уверенность в ней. Например, эксперт абсолютно уверен в том, что эффективность новой образовательной технологии высокая. Эксперту легче сформулировать подобное высказывание, чем дать числовые оценки, что всегда вызывает трудности и повышает риск ошибок. Получение числовых оценок обеспечивает формализация экспертной информации на основе Z -чисел, не привлекая для этого эксперта.

Разработка новых моделей формализации информации повлекла за собой разработку методов их анализа [21–29] в целях использования лучшей модели для решения конкретной практической задачи. Однако многие модели формализации данных образовательного процесса не нашли практического применения, возможно, по причине отсутствия публикаций с понятными примерами и объяснениями смысла самих формализаций.

Цель работы

Цель работы — разработка и сравнительный анализ моделей формализации данных образовательного процесса с примерами и пояснениями, которые дадут исследователям картину возможностей использования этих моделей для решения задач принятия решений в сфере образования.

Материалы и методы

Приведем основные понятия и определения.

Нечетким множеством первого типа \tilde{A} называется множество пар вида $\{(x, \mu_{\tilde{A}}(x)) : x \in X\}$; при этом значение функции принадлежности $\mu_{\tilde{A}}(x)$ называется степенью принадлежности x к \tilde{A} .

Нечетким множеством второго типа называется нечеткое число, если $\mu_{\tilde{A}}(x) : R \rightarrow [0, 1]$, где R — подмножество действительной прямой [1].

Лингвистической переменной называется пятерка

$$\{X, T(X), U, V, S\},$$

где X — название переменной;

$T(X) = \{X_l, l = \overline{1, m}\}$ — терм-множество переменной X , то есть множество термов или названий лингвистических значений переменной X ;

V — синтаксическое правило, порождающее названия значений лингвистической переменной X ;

S — семантическое правило, которое ставит в соответствие каждому терму из $T(X)$ нечеткое подмножество универсального множества U [1].

При оценке любых параметров их точечными значениями риск ошибки всегда больше, чем при оценке интервальными значениями, поэтому для снижения этих рисков стали использовать интервальные нечеткие множества второго типа.

Интервальное нечеткое множество второго типа $\tilde{A} = (\tilde{A}, \tilde{A})$ традиционно определяется нижним \underline{A} и верхним \overline{A} нечеткими множествами соответственно с нижней $\mu_{\underline{A}}(x)$ и верхней $\mu_{\overline{A}}(x)$ функциями принадлежности [17].

Таким образом, нечеткие множества второго типа расширяют возможности нечетких множеств первого типа для формализованного представления экспертных оценочных процедур.

Z -числом называется упорядоченная пара нечетких чисел $Z = (\tilde{A}, \tilde{R})$, где \tilde{A} — нечеткое число с функцией принадлежности $\mu_{\tilde{A}}(x) : R \rightarrow [0, 1]$, которое является нечетким расширением значений действительной переменной X , а \tilde{R} — нечеткое число с функцией принадлежности $\mu_{\tilde{R}}(x) : [0, 1] \rightarrow [0, 1]$, которое является нечетким расширением значений меры надежности первой компоненты \tilde{A} [20].

Лингвистическим называется следующее Z -число: $Z = (\tilde{A}, \tilde{R})$, компоненты которого \tilde{A}, \tilde{R} являются значениями лингвистических переменных [19].

Преимущество представления данных на основе Z -чисел по сравнению с их представлениями на основе интервальных нечетких множеств второго типа заключается в том, что эксперт может оценить достоверность словами «Малодостоверно», «Не очень достоверно», «Достоверно», «Очень достоверно», «Полностью достоверно», которые легко формализуются с помощью нечетких множеств.

В работе [15] дано определение агрегирующего отрезка $[c_1, c_2]$ для нечеткого числа $\tilde{A} = (a_1, a_2, a_L, a_R)$:

$$c_1 = \int_0^1 \frac{2a_1 - (1-\alpha)a_L}{2} 2\alpha d\alpha = a_1 - \frac{1}{6}a_L,$$

$$c_2 = \int_0^1 \frac{2a_2 + (1-\alpha)a_R}{2} 2\alpha d\alpha = a_2 + \frac{1}{6}a_R.$$

Также в работе [15] определено расстояние между нечеткими числами

$$\tilde{A} = (a_1, a_2, a_L, a_R), \quad \tilde{B} = (b_1, b_2, b_L, b_R)$$

на основе их агрегирующих отрезков $[c_1, c_2]$ $[d_1, d_2]$

$$\rho(\tilde{A}, \tilde{B}) = \sqrt{(c_1 - d_1)^2 + (c_2 - d_2)^2}.$$

В работе [18] для интервального нечеткого множества второго типа $\tilde{A} = (\underline{\tilde{A}}, \overline{\tilde{A}})$ определена пара агрегирующих отрезков $[a'_1, a'_2]$, $[a''_1, a''_2]$, первый из отрезков определен для нечеткого множества $\underline{\tilde{A}}$, второй — для нечеткого множества $\overline{\tilde{A}}$. В этой же работе определено расстояние между интервальными нечеткими множествами второго типа \tilde{A}, \tilde{B} с агрегирующими отрезками соответственно $[a'_1, a'_2]$, $[a''_1, a''_2]$, $[b'_1, b'_2]$, $[b''_1, b''_2]$:

$$\rho(\tilde{A}, \tilde{B}) = \sqrt{(a'_1 - b'_1)^2 + (a'_2 - b'_2)^2 + (a''_1 - b''_1)^2 + (a''_2 - b''_2)^2}.$$

В работе [22] дано определение агрегирующего отрезка $[\delta_1, \delta_2]$ для Z-числа $Z = (\tilde{C}, \tilde{R})$, $\tilde{C} = (c_1, c_2, c_L, c_R)$, $\tilde{R} = (r_1, r_2, r_L, r_R)$:

$$\delta_1 = r_1 \left(c_1 - \frac{1}{6}c_L \right) - r_L \left(\frac{1}{6}c_1 - \frac{1}{12}c_L \right),$$

$$\delta_2 = r_2 \left(c_2 + \frac{1}{6}c_R \right) + r_R \left(\frac{1}{6}c_2 + \frac{1}{12}c_R \right).$$

На основе агрегирующих отрезков $[\delta_1^1, \delta_2^1]$, $[\delta_1^2, \delta_2^2]$ Z-чисел Z_1, Z_2 определено расстояние между ними

$$\rho(Z_1, Z_2) = \sqrt{(\delta_1^1 - \delta_1^2)^2 + (\delta_2^1 - \delta_2^2)^2}.$$

Все определенные расстояния используются для распознавания формализованной информации при ее анализе.

Перейдем к формализации данных образовательного процесса на основе нечетких множеств разных типов.

Рассмотрим модель прямого опроса эксперта для оценки некоторой характеристики образовательного процесса X в рамках шкалы с лингвистическими уровнями (значениями) $X_l, l = 1, m$. При опросе эксперта предлагается определить типичные для уровней $X_l, l = 1, m$ интервалы $(x_l^1, x_l^2), l = 1, m$, т. е. интервалы, для всех точек которых функции принадлежности этих уровней равны единице. Для некоторых уровней типичными могут быть точки (по одной для каждого уровня), а не интервалы. Универсальным множеством для построения функций принадлежности выбирается отрезок $[0, 1]$. Если характеристика количественная, то область ее значений путем несложных арифметических операций отображается в отрезке $[0, 1]$. Функции принадлежности нечетких чисел $\tilde{X}_l, l = 1, m$, формализующих лингвистические уровни шкалы, определяются следующим образом:

$$\tilde{X}_1 = \left(0, x_1^2, 0, \frac{x_2^1 - x_1^2}{2} \right),$$

$$\tilde{X}_l = \left(x_l^1, x_l^2, \frac{x_l^1 - x_{l-1}^2}{2}, \frac{x_{l+1}^1 - x_l^2}{2} \right),$$

$$\tilde{X}_m = \left(x_m^1, 1, \frac{x_m^1 - x_{m-1}^2}{2}, 0 \right), \quad l = \overline{2, m-1}.$$

Функции принадлежности лингвистических переменных могут быть разными. Наиболее часто используются методы построения кусочно-линейных функций принадлежности [4, 6]. Согласно этим методам, функции принадлежности соседних уровней лингвистических переменных пересекаются в точках, значения функций в которых равны 0,5, т. е. это точки наибольшей неопределенности для идентификации уровней. Для того, чтобы понять, линейны или нелинейны функции принадлежности, необходим анализ их значений ниже и выше точек пересечения.

Предположим, что, например, левое крыло линейной функции принадлежности имеет вид

$$y = \frac{5}{2}x - \frac{3}{4}$$

Найдем значения x , при которых значения функции равны 0,25 и 0,75.

Получаем

$$y(0,4) = 0,25; \quad y(0,6) = 0,75.$$

Если эксперт согласен с этими значениями, то левое крыло функции принадлежности действительно можно выбирать линейным.

Если эксперт считает, что значение функции принадлежности в точке $x = 0,4$, например, меньше, чем $0,25$, а значение функции принадлежности в точке $x = 0,6$ больше, чем $0,75$, то левое крыло функции принадлежности опишет нелинейная функция, для которой точка с координатами $(0,5; 0,5)$ будет являться точкой перегиба. На интервале $(0,3; 0,5)$ функция принадлежности будет выпуклой вниз, а на интервале $(0,5; 0,7)$ — выпуклой вверх.

Если эксперт считает, что значение функции принадлежности в точке $x = 0,4$, например, больше, чем $0,25$, а значение функции принадлежности в точке $x = 0,6$, меньше, чем $0,75$, то левое крыло функции принадлежности опишет нелинейная функция, для которой точка с координатами $(0,5; 0,5)$ будет являться точкой перегиба. На интервале $(0,3; 0,5)$ функция принадлежности будет выпуклой вверх, а на интервале $(0,5; 0,7)$ — выпуклой вниз. Аналогичный анализ необходим для правого крыла функции принадлежности.

Если левое крыло функции принадлежности некоторого уровня выбрать линейным (нелинейным), то правое крыло соседней функции принадлежности, с которой есть пересечение, также выбирают линейным (нелинейным). Если правое крыло функции принадлежности некоторого уровня выбрать линейным (нелинейным), то левое крыло соседней функции принадлежности, с которой есть пересечение, также выбирают линейным (нелинейным). Анализ функций принадлежности и их построение осуществляется последовательно от первого уровня до последнего.

Выбор функции принадлежности при формализации характеристик образовательного процесса имеет большое значение для дальнейшей обработки данных образовательного процесса и получения адекватных конечных результатов для принятия решений и управления.

Рассмотрим модель формализации статистических данных образовательного процесса, полученных в результате оценки качественной характеристики X в рамках шкалы с уровнями (значениями) $X_l, l = 1, m$.

Обозначим относительные числа объектов, оцененных уровнями $X_l, l = 1, m$, соответственно через $a_l, l = 1, m$. Построение функции принадлежности нечеткого числа $\tilde{X}_l, l = 1, m$, формализующего уровень $X_l, l = 1, m$ осуществляется таким образом, чтобы площадь фигуры, ограниченной графиком этой функции и осью абсцисс, была равна $a_l, l = 1, m$: например, при $a_1 = 0,1, a_2 = 0,5, a_3 = 0,4$ получим

$$\tilde{X}_1 = (0, 0,05, 0, 0,1), \quad \tilde{X}_2 = (0,15, 0,4, 0,1, 0,4), \\ \tilde{X}_3 = (0,8, 1,0, 0).$$

Рассмотрим модель формализации данных образовательного процесса на основе интервальных нечетких множеств второго типа [17].

Необходимость такого подхода возникает в нескольких случаях:

- 1) оценку характеристики образовательного процесса осуществляет группа экспертов;
- 2) единственный эксперт не уверен в определении границы между терминами лингвистической переменной, формализующей шкалу для оценки этой характеристики;
- 3) степень уверенности эксперта в принадлежности значений универсального множества к термам выражается не числом, а интервалом;
- 4) статистические или экспертные данные содержат случайные ошибки.

Предположим, что эксперт определяет типичные интервалы $(x_l^1, x_l^2), l = 1, m$ (интервалы, на которых функции принадлежности равны единице) для уровней $X_l, l = 1, m$ шкалы, используемой для оценки характеристики X . При этом эксперт не уверен в точности их границ $(x_l^1, x_l^2), l = 1, m$, а поэтому предполагает, что x_l^1 принадлежит $(x_l^{11}, x_l^{12}), l = 1, m$, а x_l^2 принадлежит (x_l^{21}, x_l^{22}) . Тогда формализациями уровней $X_l, l = 1, m$ будут являться интервальные нечеткие множества второго типа $\tilde{X}_l = (\underline{\tilde{X}}_l, \overline{\tilde{X}}_l), l = 1, m$:

$$\underline{\tilde{X}}_1 = \left(0, x_1^{21}, 0, \frac{x_2^{11} - x_1^{22}}{2} \right),$$

$$\overline{\tilde{X}}_1 = \left(0, x_1^{22}, 0, \frac{x_2^{11} - x_1^{22}}{2} \right),$$

$$\underline{\tilde{X}}_l = \left(x_l^{12}, x_l^{21}, \frac{x_l^{11} - x_{l-1}^{22}}{2}, \frac{x_{l+1}^{11} - x_l^{22}}{2} \right),$$

$$\overline{\tilde{X}}_l = \left(x_l^{11}, x_l^{22}, \frac{x_l^{11} - x_{l-1}^{22}}{2}, \frac{x_{l+1}^{11} - x_l^{22}}{2} \right),$$

$$\underline{\tilde{X}}_m = \left(x_m^{12}, 1, \frac{x_m^{11} - x_{m-1}^{22}}{2}, 0 \right),$$

$$\overline{\tilde{X}}_m = \left(x_m^{11}, 1, \frac{x_m^{11} - x_{m-1}^{22}}{2}, 0 \right), l = \overline{2, m-1}.$$

Для формализации данных, полученных от единственного эксперта (экзаменатора), предлагается использовать нечеткие числа.

Если эксперт (экзаменатор) не уверен в принадлежности значений уровня используемой шкалы, то для формализации полученных от него данных предлагается использовать нечеткие множества второго типа.

Отдельно следует обратить внимание на случай формализации данных, полученной от разных экспертов (экзаменаторов) в целях выработки группового критерия оценки характеристики образовательного процесса. Если групповой критерий формализуется с помощью нечетких чисел, то существует большая вероятность потери индивидуальной информации, полученной от каждого эксперта (экзаменатора).

Формализация группового критерия на основе нечетких множеств второго типа позволяет учесть индивидуальные критерии экспертов (экзаменаторов) с помощью нижних и верхних функций принадлежности, которые соответствуют этим множествам.

Следовательно, можно сделать вывод о том, что при формализации индивидуальных критериев экспертов (экзаменаторов) в целях получения группового критерия оценки характеристики образовательного процесса предпочтительно использовать нечеткие множества второго типа.

Рассмотрим формализованный подход на основе Z -чисел. Отличие от формализованного подхода на основе интервальных нечетких множеств второго типа состоит в том, что оценкой уверенности эксперта (экзаменатора) в достоверности представленной информации является не интервал, а нечеткое множество. Преимущество такого представления состоит в том, что эксперт может оценить достоверность словами «Недостоверно», «Не очень достоверно», «Достоверно», «Очень достоверно», «Абсолютно достоверно», которые без труда будут формализованы с помощью нечетких множеств.

Используя полученные от эксперта данные, построим лингвистическую переменную с названием «Достоверность информации» и терминами: «Недостоверно», «Не очень достоверно», «Достоверно», «Очень достоверно», «Абсолютно достоверно», в соответствии которым поставим нечеткие числа \tilde{R}_p , $p = 1, 5$.

Информация, полученная в результате оценки характеристики X образовательного процесса, формализуется в виде множества лингвистических Z -чисел: $Z = (\tilde{X}_l, \tilde{R}_l)$, $l = 1, m$ а нечеткое число \tilde{R}_l равно одному из нечетких чисел \tilde{R}_p , $p = 1, 5$.

Подобная формализация открывает новые возможности представления, анализа и прогноза данных образовательного процесса с учетом ее достоверности, а также позволяет уменьшить субъективную составляющую и риски ошибок.

В зависимости от доступности и полноты поступающих данных образовательного процесса появляется возможность выбрать один (или несколько) типов нечетких множеств для их формализации и последующего анализа.

Примеры. Рассмотрим результаты оценки знаний по математике у обучающихся. Формализуем их с помощью лингвистической переменной. Будем считать, что 10 % обучающихся получили оценку «2», 30 % — «3», 40 % — «4» и 20 % обучающихся получили оценку «5». Согласно работе [4] и полученной информации, термины «2», «3», «4», «5» лингвистической переменной «Знания по математике» имеют следующие функции принадлежности:

$$\tilde{2} = (0, 0,05, 0, 0,1);$$

$$\tilde{3} = (0,15, 0,25, 0,1, 0,3);$$

$$\tilde{4} = (0,55, 0,7, 0,3, 0,2);$$

$$\tilde{5} = (0,9, 1, 0,2, 0).$$

Все функции принадлежности являются T -числами, площади трапеций, ограниченных этими функциями соответственно для термов «2», «3», «4», «5» равны 0,1; 0,3; 0,4; 0,2.

Рассмотрим формализацию оценок знаний обучающихся на основе нечетких множеств второго типа [4].

Построение функций принадлежности осуществляется с учетом информации, на основе которой выше была построена лингвистическая переменная [5]. Будем считать, что при стремлении степени уверенности к единице интервал значений соответствующей функции принадлежности сужается и стремится к точечному значению. При стремлении степени уверенности к нулю интервал значений соответствующей функции принадлежности расширяется и имеет максимальную длину из всех возможных значений.

Функции принадлежности интервальных нечетких множеств второго типа, формализующих оценки «2», «3», «4», «5», получаем в виде

$$\tilde{2} = (\underline{\mu}_2(x), \overline{\mu}_2(x)), \underline{\mu}_2(x) = (0, 0,05, 0, 0,1),$$

$$\overline{\mu}_2(x) = (0, 0,05, 0, 0,15);$$

$$\tilde{3} = (\underline{\mu}_3(x), \overline{\mu}_3(x)), \underline{\mu}_3(x) = (0,15, 0,25, 0,05, 0,3),$$

$$\overline{\mu}_3(x) = (0,15, 0,25, 0,1, 0,35);$$

$$\tilde{4} = (\underline{\mu}_4(x), \overline{\mu}_4(x)), \underline{\mu}_4(x) = (0,55, 0,7, 0,25, 0,2),$$

$$\overline{\mu}_4(x) = (0,55, 0,7, 0,3, 0,25);$$

$$\tilde{5} = (\underline{\mu_5(x)}, \overline{\mu_5(x)}), \underline{\mu_5(x)} = (0,9, 1, 0,15, 0),$$

$$\overline{\mu_5(x)} = (0,9, 1, 0,2, 0).$$

Построим формализованное представление оценок знаний обучающихся на основе Z -чисел.

Функции принадлежности нечетких чисел, формализующих оценки «2», «3», «4», «5», не ограничивая общности, представим в виде

$$\tilde{2} = (0, 0,05, 0, 0,1);$$

$$\tilde{3} = (0,15, 0,25, 0,1, 0,3);$$

$$\tilde{4} = (0,55, 0,7, 0,3, 0,2);$$

$$\tilde{5} = (0,9, 1, 0,2, 0).$$

Построим лингвистическую переменную «Достоверность» с термами «Недостоверно», «Не очень достоверно», «Достоверно», «Очень достоверно», «Абсолютно достоверно» и соответствующими им нечеткими числами $\tilde{R}_p, p = \overline{1,5}$:

$$\tilde{R}_1 = (0, 0, 0,25), \tilde{R}_2 = (0,25, 0,25, 0,25),$$

$$\tilde{R}_3 = (0,5, 0,25, 0,25), \tilde{R}_4 = (0,75, 0,25, 0,25),$$

$$\tilde{R}_5 = (1, 0,25, 0).$$

Экзаменатору предлагается оценить достоверность информации, выбрав один из термов лингвистической переменной «Достоверность».

Таким образом, оценки знаний обучающихся на основе Z -чисел представляются следующим образом:

$$Z_2 = (\tilde{2}, \tilde{Q}_2), \quad Z_3 = (\tilde{3}, \tilde{Q}_3),$$

$$Z_4 = (\tilde{4}, \tilde{Q}_4), \quad Z_5 = (\tilde{5}, \tilde{Q}_5),$$

где $\tilde{Q}_2, \tilde{Q}_3, \tilde{Q}_4, \tilde{Q}_5$ равны соответственно одному из нечетких чисел $\tilde{R}_p, p = \overline{1,5}$.

Формализация данных образовательного процесса на основе Z -чисел позволяет представлять полученную информацию с определенным уровнем достоверности, тем самым повышая устойчивость конечных результатов и уменьшая риски ошибок в процедурах принятия решений.

В зависимости от типа поступающей информации выбирается соответствующий тип нечетких множеств для ее адекватной формализации и дальнейшего анализа.

Выводы

Разработанные и проанализированные модели формализации данных образовательного процесса показали, что теория нечетких множеств позволяет создавать модели, которые адекватно и полно отображают характеристики образовательного процесса, а также шкалы, используемые для их оценки, и оценочные действия экспертов.

Для каждой модели приведен числовой пример, который делает все теоретические разработки прозрачными и понятными при решении практических задач.

Совокупный теоретический и практический подход к построению моделей формализации данных образовательного процесса дает пользователям возможность выбора модели для снижения риска ошибок при выработке управляющих решений, а исследователям новые идеи для построения новых моделей и совершенствования известных.

Список литературы

- [1] Zadeh L.A. Fuzzy logic and approximate reasoning // *Synthese*, 1975, v. 80, pp. 407–428.
- [2] Poleshchuk O.M., Komarov E.G., Chernova T.V. Influence of research and development activities on professional performance of aerospace students // *AIP Conference Proceedings*, 2019, no. 2171(1), p. 140003. DOI: 10.1063/1.5133293
- [3] Poleshchuk O.M., Tumor S.V. The Analysis of Student Performance During Face-to-Face and Distance Learning under Z-Information // *Lecture Notes in Electrical Engineering*, 2022, v. 857, pp. 393–402. DOI: 10.1007/978-3-030-94202-1_37
- [4] Poleshchuk O., Komarov E. Expert Fuzzy Information Processing // *Studies in Fuzziness and Soft Computing*, 2011, v. 268, pp. 1–239.
- [5] Ryjov A. Fuzzy Linguistic Scales: Definition, Properties and Applications. *Soft Computing in Measurement and Information Acquisition. Studies in Fuzziness and Soft Computing* // Reznik L., Kreinovich V. (eds). 2003, v. 127. DOI: 10.1007/978-3-540-36216-6_3
- [6] Poleshchuk O.M. Creation of linguistic scales for expert evaluation of parameters of complex objects based on semantic scopes // *International Russian Automation Conf. (RusAutoCon–2018)*, 2018, pp. 1–6.
- [7] Runkler T.A., Katz C. Fuzzy clustering by particle swarm optimization // *Proceedings of the IEEE Int. Conf. on Fuzzy Systems*, 2006, pp. 34–41.
- [8] Liu H.C., Yih J.M., Wu D.B., Liu S.W. Fuzzy C-mean clustering algorithms based on Picard iteration and particle swarm optimization // *Proceedings of the Int. Workshop on Geoscience and Remote Sensing (ETT and GRS–2008)*, 2008, pp. 75–84.
- [9] Poleshchuk O., Komarov E. The determination of rating points of objects with qualitative characteristics and their usage in decision making problems // *Int. J. of Computational and Mathematical Sciences*, 2009, v. 3, no. 7, pp. 360–364.

- [10] Chen M., Ludwig A. Particle swarm optimization based fuzzy clustering approach to identify optimal number of clusters // *J. of Artificial Intelligence and Soft Computing Research*, 2014, v. 4, no. 1, pp. 43–56.
- [11] Darwish A., Poleshchuk O. New models for monitoring and clustering of the state of plant species based on semantic spaces // *J. of Intelligent and Fuzzy Systems*, 2014, v. 3, no. 26, pp. 1089–1094.
- [12] Phyo O., Chaw E. Comparative Study of Fuzzy PSO (FPSO) Clustering Algorithm and Fuzzy C-Means (FCM) Clustering Algorithm // *National J. of Parallel and Soft Computing*, 2019, v. 1, no. 1, pp. 62–67.
- [13] Tanaka H, Ishibuchi H. Identification of possibilistic linear models // *Fuzzy Sets and Systems*, 1991, v. 41, pp. 145–160.
- [14] Chang Y.-H. Hybrid fuzzy least- squares regression analysis and its reliability measures // *Fuzzy Sets and Systems*, 2001, v. 119, pp. 225–246.
DOI:10.1016/S0165-0114(99)00092-5
- [15] Domrachev V.G., Poleshchuk O.M. On the construction of a regression model under fuzzy source data // *Avtomatika i Telemekhanika*, 2003, v. 11, pp. 74–83.
- [16] Arefi M. Quantative fuzzy regression based on fuzzy outputs and fuzzy parameters. *Soft Computing-A Fusion of Foundations, Methodologies and Applications*, 2020, v. 24(1), pp. 311–320.
DOI:10.1007/s00500-019-04424-2
- [17] Liu F., Mendel J.M. Encoding words into interval Type-2 fuzzy sets using an interval approach // *IEEE Trans. Fuzzy Systems*. 2008, v. 16(6), pp. 1503–1521.
- [18] Poleshchuk O., Komarov E. A fuzzy linear regression model for interval type-2 fuzzy sets // *Annual Conf. of the North American Fuzzy Information Processing Society – NAFIPS’2012*, 2012, p. 6290970.
DOI: 10.1109/NAFIPS.2012.6290970
- [19] Полещук О.М. Кластерный анализ Z-информации на основе эталонной системы нечетких определенных принадлежности // *Лесной вестник / Forestry Bulletin*, 2024. Т. 28. № 2. С. 150–155.
DOI: 10.18698/2542-1468-2024-2-150-155
- [20] Zadeh L.A. A Note on Z-numbers // *Information Sciences*, 2011, v. 14, no. 181, pp. 2923–2932.
- [21] Sadikoglu F, Huseynov O, Memmedova K. Z-Regression analysis in psychological and educational researches // *Procedia Comput. Sci.*, 2016, v. 102, pp. 385–389. DOI: 10.1016/J.PROCS.2016.09.416
- [22] Poleshchuk O. Fuzzy regression model with input and output Z-numbers // *IOP Conf. Series: Materials Science and Engineering*, 2020, v. 919(5), p. 052041. DOI: 10.1088/1757-899X/919/5/052041
- [23] Poleshchuk O.M. Multiple Z-Regression with Fuzzy Coefficients // *Advances in Intelligent Systems and Computing*, 2021, v. 1306, pp. 63–70.
DOI: 10.1007/978-3-030-64058-3_8
- [24] Poleshchuk O. Quintile multiple regression with fuzzy coefficients and initial Z-information // *E3S Web of Conferences*, 2023, v. 431, p. 05015.
DOI: 10.1051/e3sconf/202343105015
- [25] Jamal M., Khalif K., Mohamad S. The implementation of Z-numbers in fuzzy clustering algorithm for wellness of chronic kidney disease patients // *J. of Physics: Conf. Series*, 2018, v. 1366, p. 012058.
- [26] Aliev R., Guirimov B. Z-number clustering based on general Type-II fuzzy sets // *Advances in Intelligent Systems and Computing*, 2018, v. 896, pp. 270–278.
- [27] Aliev R.A., Pedrycz W., Guirimov B.G., Huseynov O.H. Clustering method for production of Z-numbers based if-then rules // *Information Sciences*, 2020, v. 520, pp. 155–176.
- [28] Полещук О.М. Кластерный анализ экспертной информации на основе Z-чисел // *Лесной вестник / Forestry Bulletin*, 2022. Т. 26. № 1. С. 143–148.
DOI: 10.18698/2542-1468-2022-1-143-148
- [29] Poleshchuk O. Clustering Z-information based on a system of fuzzy reference requirements // *E3S Web of Conferences*, 2023, v. 420, p. 06022.

Сведения об авторе

Полещук Ольга Митрофановна — д-р техн. наук, профессор, зав. кафедрой «Высшая математика и физика», ФГАОУ ВО «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)» (Мытищинский филиал), poleshhukom@bmstu.ru

Поступила в редакцию 27.08.2025.

Одобрено после рецензирования 23.09.2025.

Принята к публикации 04.12.2025.

EDUCATIONAL PROCESS DATA FORMALIZATION ON BASIS OF VARIOUS FUZZY SETS

O.M. Poleshchuk

BMSTU (Mytishchi branch), 1, 1st Institutskaya st., 141005, Mytishchi, Moscow reg., Russia

poleshchukom@bmstu.ru

The paper develops models for formalizing the data of the educational process in the conditions of various initial information. For modeling, the paper uses first type fuzzy sets, interval second type fuzzy sets, and Z -numbers. All the constructed models have been analyzed and recommended for use in solving various practical problems. First type fuzzy sets are recommended to be used to formalize statistical data of the educational process, as well as data obtained from a single expert (examiner). Interval second type fuzzy sets are recommended to be used to formalize statistical and expert data of the educational process with random errors, as well as data obtained from a group of experts. Z -numbers are recommended to be used to formalize the data of the educational process, taking into account their reliability. The numerical examples given in the article, together with theoretical justifications, provide an opportunity to choose a model for further data analysis in order to obtain sustainable final results and control decisions based on them.

Keywords: educational process, data formalization, fuzzy set, Z -number

Suggested citation: Poleshchuk O.M. *Formalizatsiya dannykh obrazovatel'nogo protsessa na osnove nechetkikh mnozhestv raznykh tipov* [Educational process data formalization on basis of various fuzzy sets]. *Lesnoy vestnik / Forestry Bulletin*, 2026, vol. 30, no. 1, pp. 138–146.
DOI: 10.18698/2542-1468-2026-1-138-146

References

- [1] Zadeh L.A. Fuzzy logic and approximate reasoning. *Synthese*, 1975, v. 80, pp. 407–428.
- [2] Poleshchuk O.M., Komarov E.G., Chernova T.V. Influence of research and development activities on professional performance of aerospace students. *AIP Conference Proceedings*, 2019, no. 2171(1), p. 140003. DOI: 10.1063/1.5133293
- [3] Poleshchuk O.M., Tumor S.V. The Analysis of Student Performance During Face-to-Face and Distance Learning under Z -Information. *Lecture Notes in Electrical Engineering*, 2022, v. 857, pp. 393–402. DOI: 10.1007/978-3-030-94202-1_37
- [4] Poleshchuk O., Komarov E. Expert Fuzzy Information Processing. *Studies in Fuzziness and Soft Computing*, 2011, v. 268, pp. 1–239.
- [5] Ryjov A. Fuzzy Linguistic Scales: Definition, Properties and Applications. *Soft Computing in Measurement and Information Acquisition. Studies in Fuzziness and Soft Computing // Reznik L., Kreinovich V. (eds), 2003, v. 127. DOI: 10.1007/978-3-540-36216-6_3*
- [6] Poleshchuk O.M. Creation of linguistic scales for expert evaluation of parameters of complex objects based on semantic scopes. *International Russian Automation Conf. (RusAutoCon–2018)*, 2018, pp. 1–6.
- [7] Runkler T.A., Katz C. Fuzzy clustering by particle swarm optimization // *Proceedings of the IEEE Int. Conf. on Fuzzy Systems*, 2006, pp. 34–41.
- [8] Liu H.C., Yih J.M., Wu D.B., Liu S.W. Fuzzy C -mean clustering algorithms based on Picard iteration and particle swarm optimization. *Proceedings of the Int. Workshop on Geoscience and Remote Sensing (ETT and GRS–2008)*, 2008, pp. 75–84.
- [9] Poleshchuk O., Komarov E. The determination of rating points of objects with qualitative characteristics and their usage in decision making problems. *Int. J. of Computational and Mathematical Sciences*, 2009, v. 3, no. 7, pp. 360–364.
- [10] Chen M., Ludwig A. Particle swarm optimization based fuzzy clustering approach to identify optimal number of clusters. *J. of Artificial Intelligence and Soft Computing Research*, 2014, v. 4, no. 1, pp. 43–56.
- [11] Darwish A., Poleshchuk O. New models for monitoring and clustering of the state of plant species based on semantic spaces. *J. of Intelligent and Fuzzy Systems*, 2014, v. 3, no. 26, pp. 1089–1094.
- [12] Phyo O., Chaw E. Comparative Study of Fuzzy PSO (FPSO) Clustering Algorithm and Fuzzy C -Means (FCM) Clustering Algorithm. *National J. of Parallel and Soft Computing*, 2019, v. 1, no. 1, pp. 62–67.
- [13] Tanaka H, Ishibuchi H. Identification of possibilistic linear models. *Fuzzy Sets and Systems*, 1991, v. 41, pp. 145–160.
- [14] Chang Y.-H. Hybrid fuzzy least-squares regression analysis and its reliability measures. *Fuzzy Sets and Systems*, 2001, v. 119, pp. 225–246. DOI:10.1016/S0165-0114(99)00092-5
- [15] Domrachev V.G., Poleshchuk O.M. On the construction of a regression model under fuzzy source data. *Avtomatika I Telemekhanika*, 2003, v. 11, pp. 74–83.
- [16] Arefi M. Quantitative fuzzy regression based on fuzzy outputs and fuzzy parameters. *Soft Computing-A Fusion of Foundations, Methodologies and Applications*, 2020, v. 24(1), pp. 311–320. DOI:10.1007/s00500-019-04424-2
- [17] Liu F., Mendel J.M. Encoding words into interval Type-2 fuzzy sets using an interval approach. *IEEE Trans. Fuzzy Systems*, 2008, v. 16(6), pp. 1503–1521.
- [18] Poleshchuk O., Komarov E. A fuzzy linear regression model for interval type-2 fuzzy sets. *Annual Conf. of the North American Fuzzy Information Processing Society – NAFIPS'2012*, 2012, p. 6290970. DOI: 10.1109/NAFIPS.2012.6290970

- [19] Poleshchuk O.M. *Klasternyy analiz Z-informatsii na osnove etalonnoy sistemy nechetkikh opredeleniy prinadlezhnosti* [Cluster analysis of Z-information based on a reference system of fuzzy identification]. *Lesnoy vestnik / Forestry Bulletin*, 2024, vol. 28, no. 2, pp. 150–155. DOI: 10.18698/2542-1468-2024-2-150-155
- [20] Zadeh L.A. A Note on Z-numbers. *Information Sciences*, 2011, v. 14, no. 181, pp. 2923–2932.
- [21] Sadikoglu F, Huseynov O, Memmedova K. Z-Regression analysis in psychological and educational researches. *Procedia Comput. Sci.*, 2016, v. 102, pp. 385–389. DOI: 10.1016/J.PROCS.2016.09.416
- [22] Poleshchuk O. Fuzzy regression model with input and output Z-numbers. *IOP Conf. Series: Materials Science and Engineering*, 2020, v. 919(5), p. 052041. DOI: 10.1088/1757-899X/919/5/052041
- [23] Poleshchuk O.M. Multiple Z-Regression with Fuzzy Coefficients. *Advances in Intelligent Systems and Computing*, 2021, v. 1306, pp. 63–70. DOI: 10.1007/978-3-030-64058-3_8
- [24] Poleshchuk O. Quintile multiple regression with fuzzy coefficients and initial Z-information. *E3S Web of Conferences*, 2023, v. 431, p. 05015. DOI: 10.1051/e3sconf/202343105015
- [25] Jamal M., Khalif K., Mohamad S. The implementation of Z-numbers in fuzzy clustering algorithm for wellness of chronic kidney disease patients. *J. of Physics: Conf. Series*, 2018, v. 1366, p. 012058.
- [26] Aliev R., Guirimov B. Z-number clustering based on general Type-II fuzzy sets. *Advances in Intelligent Systems and Computing*, 2018, v. 896, pp. 270–278.
- [27] Aliev R.A., Pedrycz W., Guirimov B.G., Huseynov O.H. Clustering method for production of Z-numbers based if-then rules. *Information Sciences*, 2020, v. 520, pp. 155–176.
- [28] Poleshchuk O.M. *Klasternyy analiz ekspertnoy informatsii na osnove Z-chisel* [Cluster analysis of expert information based on Z-numbers]. *Lesnoy vestnik / Forestry Bulletin*, 2022, vol. 26, no. 1, pp. 143–148. DOI: 10.18698/2542-1468-2022-1-143-148
- [29] Poleshchuk O. Clustering Z-information based on a system of fuzzy reference requirements. *E3S Web of Conferences*, 2023, v. 420, p. 06022.

Author's information

Poleshchuk Ol'ga Mitrofanovna — Dr. Sci. (Tech.), Professor, Head of Higher Mathematics and Physics Department, of the BMSTU (Mytishchi branch), poleshhukom@bmstu.ru

Received 27.08.2025.

Approved after review 23.09.2025.

Accepted for publication 04.12.2025.