

Энтропийно-информационный подход для оценивания качества текстов машинных переводов

Г. Н. Ахобадзе, Е. Ю. Русяева, А. В. Полтавский

Институт проблем управления им. В. А. Трапезникова РАН, Москва, Россия

Аннотация. В статье предлагается методика оценки качества текстов машинных переводов на основе энтропийно-информационного подхода. Дается анализ дисперсионного и энтропийного коэффициентов конкордации, используемых для оценки согласованности мнений экспертов при близких ранжировках различных объектов. Обосновывается перспективность применения энтропийного коэффициента конкордации, позволяющего зафиксировать факт разделения мнений на две противоположные группы. Это положение важно для проводимого исследования, поскольку в данной методике экспертной оценки переводов текстов важен учет разных мнений нескольких экспертов, привлеченных к экспертизе. Приводятся примеры расчета энтропийного коэффициента конкордации с изменяющейся ранговой системой, числом экспертов и ранжируемых объектов оценивания.

Ключевые слова: коэффициент конкордации, ранжировка текстов, мнение экспертов, информационная энтропия, ранг, матрица мнений экспертов, вероятностные параметры.

DOI 10.14357/20718632230402

EDN EOBUMC

Введение

Как известно, в нашем все более информатизирующемся мире больших данных разработано и продолжает мультипликативно увеличиваться число цифровых инновационных проектов и технологий, направленных как на создание самих цифровых объектов, так и на автоматизацию процедур их обработки. Естественно, что в нынешнем потоке разнородной информации только экспертной (человеческой) обработкой данных уже не обойтись. Поэтому растет интерес, даже среди представителей гуманитарного знания, к возможностям автоматизации определенных процедур анализа тестового и прочего мультимедийного контента. Важно учитывать, что в разработке новых адаптивных алгоритмов для сетевых информационных систем (ИС), при

создании информационно-аналитических систем (ИАС), необходимо учитывать фактор неопределенности.

Авторы в рамках интегративного подхода [1] используют метод энтропийного анализа не только для лингвистического анализа авторского (художественного) текста как объекта исследования, но и для расчетов информационной энтропии в экспертной оценке переводов такого рода текстов в создаваемой ИАС энтропийного оценивания данных. Именно на этом аспекте и будет заострено исследовательское внимание в данной работе. Акцент ставится на методике расчета информационной энтропии при анализе вариантов экспертных оценок разного количества текстовых объектов с учетом рангов оценивания. Считаем, что представленный метод энтропийного анализа в перспективе позволит

точнее и качественнее оценивать языковые характеристики как самих лингвистических (а далее и других, мультимедийных) объектов, так и качественнее идентифицировать экспертные мнения, тем самым лучше определять адекватность информации, как авторских, так и искусственных переводчиков текстов.

Проблема, которая требует решения, кроется, на наш взгляд, не только и не столько в растущей актуальности оперативного анализа информации и генерации текстов, чем сейчас заняты выложенные в открытом доступе программы типа ChatGPT [2], а в том, чтобы адекватно идентифицировать оригинальные авторские тексты. Актуальны методики, позволяющие адекватно оценить перевод текста оригиналу, методы анализа экспертных оценок для сличения подобия и соответствия, адекватности и полноте преобразованных данных. На поиск решения этих задач и направлено данное исследование. Причем в качестве гипотезы мы предполагаем, что интегративный подход и энтропийный анализ как знаковых систем и символической информации в сетевых структурах, так и энтропийный анализ экспертных оценок дают более адекватный расчет при формировании современных информационно-аналитических проектов.

1. Краткий обзор методик оценки энтропийного коэффициента конкордации в экспертных мнениях с ранговой ранжировкой

Авторы не ставили перед собой задачу сделать теоретическую обзорную статью, поэтому детальный обзор многочисленных и разноплановых подходов для оценки неопределенности и активно разрабатываемых к настоящему моменту энтропийных методов, подробно описанных, например, в работах Ю.С. Попкова [3], методов вербального анализа решений О.И. Ларичева [4] и методов, использующих математический аппарат нечеткой логики [5] не был проведен. Также в данном исследовании задача глубинной проработки методов группового принятия решений не ставилась, хотя считаем необходимым сослаться на ставшие уже классическими работы отечественных [6-9] и зарубежных авторов [10-12] по данной тематике. Подробнее об отличиях в указанных и наших подходах мы планируем рассказать в дальнейших работах.

Поясним нашу нынешнюю авторскую позицию. В отличие от указанных выше работ, разработанный авторами метод основывается на вычислении энтропийного коэффициента и является, по мысли авторов, более простым в реализации. В подтверждение этого положения приведем лишь некоторые доказательные ссылки из следующих научных работ.

Как отмечено в исследовании [13], «Сравнительная оценка дисперсионного и энтропийного коэффициентов конкордации показывает, что эти коэффициенты дают примерно одинаковую оценку согласованности экспертов при близких ранжировках. Однако если, например, вся группа экспертов разделилась в мнениях на две подгруппы, причем ранжировки в этих подгруппах противоположные (прямая и обратная), то дисперсионный коэффициент конкордации будет равен 0, а энтропийный коэффициент конкордации будет равен 0,7. Таким образом, энтропийный коэффициент конкордации позволяет зафиксировать факт разделения мнений на две противоположные группы». Это положение важно для нашего исследования, поскольку в нашей методике экспертной оценки переводов текстов важен учет разных мнений нескольких экспертов, привлеченных к экспертизе. Таким образом, при обработке результатов ранжирования могут возникнуть задачи определения зависимости между ранжировками двух экспертов, связи между достижением двух различных целей при решении одной и той же совокупности проблем или взаимосвязи между двумя признаками. В этих случаях мерой взаимосвязи может служить коэффициент ранговой корреляции.

Характеристикой взаимосвязи множества ранжировок или целей будет являться матрица коэффициентов ранговой корреляции. Давно известны, как подчеркнуто в [14], коэффициенты ранговой корреляции Спирмена и Кендалла, а также то, что метод экспертных оценок активно используется при принятии решений. Ведь, строго говоря, разного рода принятие решений, в том числе управленческих, далеко не всегда основывается только на принципе рациональности, научной доказательности, математически точных расчетах. Часто решения принимаются на основе обобщения своего/чужого опыта, субъективных экспертных оценок проблемы специалистами.

Когда речь идет о качественном, смысловом оценивании текстового объекта, метод экспертных оценок позволяет более или менее успешно проводить такое качественное сравнение объектов (по принципу «лучше-хуже»). Профессор Венедиктов статье «О показателе согласованности экспертных оценок» подвергает конструктивной критике коэффициент ранговой конкордации Кэнделла, доказывая его непригодность для оценки степени согласованности экспертных оценок и обосновывает перечень необходимых свойств, которыми должна обладать функция, адекватно вычисляющая такую согласованность [14]. В этом исследовании было доказано, что поведение функции Кэнделла на ряде наборов данных не является разумным: ее значения при весьма незначительном изменении аргумента могут отличаться существенным образом и, наоборот, обладающие качественными различиями исходные данные могут давать одинаковое значение коэффициента конкордации. Вместе с тем, «разумность» поведения искомой функции является неформализованным понятием и по этой причине не может быть положена в основу обоснования пригодности любого из предлагаемых подходов к вычислению коэффициента конкордации. В целях формализации соответствующих требований автор обосновал ряд правил, выполнение которых, по его мнению, полагается необходимым условием пригодности анализируемой функции к использованию ее для определения согласованности экспертных оценок.

Со своей стороны А. И. Буравлев в исследовании «О задачах многокритериального выбора» [15], анализируя методические подходы к решению многокритериальных задач, использующих различные свертки частных критериев и специальные процедуры построения парето-оптимальных решений, приходит к выводу, что необходимость решения многокритериальных задач во многом зависит от искусства выбора целевой функции (критерия), отражающей цель решения задачи и связанных с ней фазовых переменных, влияющих на достижение данной цели. Именно включение фазовых переменных в число критериев и приводит к возникновению многокритериальных задач. Пример, рассмотренный в [15], доказывает, что адекватный и корректный выбор целевого критерия и соответствующих фазовых

переменных исключает необходимость решения многокритериальной задачи.

Разумеется, авторы привели лишь очень краткий обзор современных работ по схожей тематике. Но даже этот небольшой перечень позволил нам убедиться в правильности выбора нами классических математических решений, подтверждение этого в [16]. Естественно, на сегодняшний день имеется много методов оценки информации и ее экспертизы, есть и довольно актуальный метод ранжирования, и метод идеальной точки и многие, многие другие. Эти методы применяются в теории принятия решений [17] к различным системам.

2. Формализация информационных процессов и модели обработки потоков данных

Предлагаемая авторами методика расчета информационной энтропии для экспертной оценки машинных переводов текстов основывается на анализе энтропийного коэффициента согласованности [18] мнений экспертов – переводчиков при ранжировании письменного текста машинного перевода.

В общем случае при сопоставлении объектов, эксперты обычно в большей или меньшей степени расходятся во мнениях. В связи с этим возникает необходимость количественной оценки степени согласия экспертов. Для определения степени согласованности экспертов – переводчиков, ранжирующих рассматриваемый письменный текст машинного перевода, используется энтропийный коэффициент конкордации (коэффициент согласия) $Wэ$, выражаемый формулой $Wэ = 1 - H/H_{max}$, где H – энтропия, вычисляемая по формуле:

$$H = - \sum_{n=1}^N \sum_{i=1}^{Fn(L)} P_{ni} \log P_{ni}, \quad (1)$$

где N – число письменных текстов, L – число экспертов, $F_n(L)$ – число групп экспертов, присвоивших объекту O_n один и тот же ранг r_{ni} , P_{ni} – оценка вероятности того, что i -ая группа экспертов присвоит n -му объекту один и тот же ранг r_{ni} . Величина P_{ni} вычисляется в виде отношения количества экспертов L_{ni} , приславших объекту O_n один и тот же ранг r_{ni} к общему числу экспертов L : $P_{ni} = L_{ni} / L$.

Определение максимального значения энтропии $H_{\max}$ сводится к тому, что сначала необходимо оценить целое частное от деления числа экспертов на число письменных текстов Q , а затем остаток от деления числа экспертов на число объектов R . Следовательно, можно записать $L = N * Q + R$. После этого определяем число объектов, получивших по Q единиц и число объектов, получивших по $Q + 1$ единиц. Тогда в рассматриваемом случае для максимального значения энтропии можно принимать $H_{\max} = H_{\max}^Q + H_{\max}^{Q+1}$.

Здесь $H_{\max}^Q$ – максимальное значение информационной энтропии для письменных текстов, получивших по Q единиц, а $H_{\max}^{Q+1}$ – максимальное значение энтропии для письменных текстов, получивших по $Q + 1$ единиц. После вычисления $H_{\max}^Q$ и $H_{\max}^{Q+1}$ соответствующими формулами и с учетом допущения, например, $Q = 0$, то $L = R$ и $H_{\max}^Q = 0$ для $H_{\max}$ можно записать:

$$H_{\max} = N * \log L. \quad (2)$$

При $Q \geq 1$ $H_{\max}$ примет следующий вид:
 $H_{\max} = N * \log L$

$$[(N - R) * Q / L * \log L / Q + R * Q + 1 / L * \log L / Q + 1]. \quad (3)$$

В рассматриваемом случае для расчета коэффициента согласованности [19] мнений экспертов-переводчиков воспользуемся ранговой системой от 1-го ранга до 3-х рангов. При этом принимается, что первый ранг по значимости ниже, чем 2, а второй ранг – чем третий ранг. Кроме того, все эксперты, привлеченные к ранжированию машинного текста переводов, должны обладать высоким уровнем профессиональной компетентности.

3. Примеры расчета энтропийного коэффициента конкордации, информационной энтропии при отборе экспертных оценок перевода текстов

3.1. Первый вариант. Можно рассмотреть пример расчета, когда число экспертов равно 3, а число объектов равно 1, т.е., $L = 3$, $N = 1$. После этого составляется матрица, показывающая согласованность мнений экспертов в зависимости от рангов и объекта.

Матрица 1

	O_1
$r=1$	2
$r=2$	1
Σ	3

Согласно матрице 1, при наличии двух ранговой системы и одного текста машинного перевода, двое экспертов сошлись во мнениях по первому рангу. Поэтому для данного случая, соответственно, к расчету подлежит параметр r_{11} , т.е. первый ранг по объекту. Следовательно, для параметров из формулы (1) соответственно получаем, что $F_n(L) = 1$; $P_{ni} = L_{ni} / L = 2/3 = 0,6$. После подстановки этих цифровых значений в формулу (1) с учетом сумм, равным единицам, имеем: $H = -1 * 1 * 0,6 \log 0,6 = 0,13$.

Согласно выше приведенной методике, вычисление $H_{\max}$ предусматривает сначала определение параметров Q и R . Из матрицы 1 видно, что эти параметры соответственно равны 0 и 1. Тогда в этом случае для определения $H_{\max}$ воспользуемся формулой (2). Следовательно, получаем $H_{\max} = 0,47$. С учетом этого отношение $H/H_{\max} = 0,276$. В результате $W_s = 1 - H/H_{\max} = 1 - 0,276 = 0,724$.

Полученное значение коэффициента согласованности можно считать вполне приемлемым для оценки качества переведенного текста.

3.2. Второй вариант. Для данного случая принимаем, что число экспертов $L=3$, а объектов $N=2$, то есть, число экспертов больше числа объектов. $L = 3$, $N = 2$.

Матрица 2

	O_1	O_2
$r=1$	1	2
$r=2$	2	1
Σ	3	3

Как следует из матрицы 2, при наличии двух ранговой системы и двух текстов машинного перевода, двое экспертов сошлись во мнениях по первому и второму рангам, т.е. в данном случае к расчету подлежат параметры r_{12} и r_{22} . В этом случае, так как максимум согласованности мнений экспертов наблюдается в случаях первого ранга по второму объекту и второго ранга по первому объекту, то для вычисления H достаточен расчет либо r_{12} , либо r_{22} . В соответствии с этим (случай r_{12}), для

параметров из формулы (1) соответственно получаем, что $F_n(L) = 1$; $P_{ni} = L_{ni} / L = 2/3 = 0,6$. После подстановки этих цифровых значений в формулу (1) и с учетом сумм, равным единицам, имеем: $H = -1 * 1 * 0,6 \log 0,6 = 0,13$.

Аналогично первому варианту расчета для параметров Q и R по данным матрицы 2 можно принимать, что они равны соответственно 0 и 1. Следовательно, в формулу $H_{\max} = N * \log L$ после подстановки $N = 2$ и $L = 3$, имеем $H_{\max} = 2 * \log 3 = 0,95$. После вычисления отношения $H/H_{\max} = 0,13/0,95 = 0,14$ имеем, что $W_3 = 1 - H/H_{\max} = 1 - 0,14 = 0,86$. Это значение коэффициента согласованности можно считать приемлемым для оценки качества переведенных двух текстов в данном случае.

Как уже отмечали выше, из-за одинаковых максимальных значений согласованности мнений экспертов по параметрам r_{12} и r_{22} , по рассчитанному значению одного параметра можно судить о значении другого параметра. Иными словами, вычисленное значение коэффициента W_3 для случая r_{12} можно применить и для случая r_{22} .

Здесь, как показывают расчеты, имеет место равенство коэффициентов конкордации случаев первого ранга по второму объекту и второго ранга по первому объекту. В таком случае приоритет отдается второму тексту машинного перевода, так как второй ранг по значимости выше, чем первый ранг.

3.3. Третий вариант. Здесь принимается: $L = 2$, $N = 3$.

Матрица 3

	O ₁	O ₂	O ₃
r=1	2	0	2
r=2	0	2	0
Σ	2	2	2

Из матрицы 3 видно, что к расчету подлежат параметры r_{11} , r_{13} и r_{22} (случай максимумов согласованности во мнениях всех двух экспертов). В силу этого вычисление коэффициента конкордации для одного случая может быть применено и к другим двум случаям. Следовательно, для вычисления H по параметру, например, r_{11} , значения параметров $F_n(L)$ и P_{ni} могут быть представлены как $F_n(L) = 1$; $P_{ni} = L_{ni} / L = 2/2 = 1$. После подстановки этих цифровых значений в формулу (1) и с учетом сумм равным единицам, имеем $H = -1 * 1 * 1 \log 1 = 0$.

Согласно матрице 3 для параметров Q и R, можно записать $Q = 0$, а $R = 3$. Тогда в этом случае для определения $H_{\max}$ воспользуемся формулой (2). В итоге: $H_{\max} = N * \log L = 3 * \log 2 = 3 * 0,3 = 0,9$.

Следовательно, для коэффициента согласованности W_3 будем иметь $W_3 = 1 - H/H_{\max} = 1 - 0/0,9 = 1$. Отсюда можно сделать заключение о согласованности во мнениях всех экспертов в полном объеме. В соответствии с этим, по параметрам r_{13} и r_{22} также можно принимать значения коэффициентов в пределах единицы.

В этом случае ввиду равенства всех коэффициентов конкордации предпочтение можно отдать второму тексту машинного перевода, так как по рангу он выше, чем другие переводы.

3.4. Четвертый вариант, когда $L = 4$, $N = 4$.

Матрица 4

	O ₁	O ₂	O ₃	O ₄
r=1	3	2	0	1
r=2	0	1	0	1
r=3	1	1	4	2
Σ	4	4	4	4

Из матрицы 4 видно, что при 3-х ранговой системе ранжирования максимум согласованности мнений экспертов достигается в одном случае – третьего ранга по третьему объекту, т.е. к расчету подлежит r_{13} . Для вычисления H по параметру r_{13} , значения параметров $F_n(L)$ и P_{ni} могут быть представлены как $F_n(L) = 1$; $P_{ni} = L_{ni} / L = 4/4 = 1$. После подстановки этих цифровых значений в формулу (1) и с учетом сумм равным единицам имеем $H = -1 * 1 * 1 \log 1 = 0$.

Согласно матрице 4 для параметров Q и R, можно записать $Q = 1$, а $R = 0$. В этом случае, так как Q больше единицы, для вычисления $H_{\max}$ необходимо использовать формулу (3). После подстановки цифровых значений $N = 4$; $R = 0$; $Q = 1$; $L = 4$ в эту формулу для $H_{\max}$ можно записать $H_{\max} = 4 * [(4) * 0,25 * \log 4 * 0,5 * \log 2] = 0,36$.

После вычисления отношения $H/H_{\max}$ ($H = 0$ и $H_{\max} = 0,36$) для коэффициента имеем $W_3 = 1 - H/H_{\max} = 1 - 0/0,36 = 1$. Отсюда можно заключить, что наиболее качественным текстом машинного перевода по мнению экспертов можно считать третий текст.

3.5. Пятый вариант, когда $L = 3$, $N = 4$.

Матрица 5

	O ₁	O ₂	O ₃	O ₄
r=1	3	2	0	2
r=2	0	1	3	1
Σ	3	3	3	3

Из матрицы 5 видно, что при 2-х ранговой системе ранжирования максимумы согласованности мнений экспертов достигаются в двух случаях: первого ранга по первому объекту и второго ранга по третьему объекту, т.е. в данном случае к расчету подлежат r_{11} и r_{23} . Аналогично предыдущим случаям, для вычисления H достаточен расчет либо r_{11} , либо r_{23} .

Следовательно, для вычисления H по параметру, например, r_{11} , значения параметров $F_n(L)$ и P_{ni} могут быть представлены как $F_n(L) = 1$; $P_{ni} = L_{ni} / L = 3/3 = 1$. После подстановки этих цифровых значений в формулу (1) и с учетом сумм равным единицам имеем $H = -1 \cdot 1 \cdot \log 1 = 0$.

Согласно матрице 5 для параметров Q и R можно записать $Q = 0$, а $R = 3$. Тогда в этом случае для определения $H_{\max}$ воспользуемся формулой (2). В итоге $H_{\max} = N \cdot \log L = 4 \cdot \log 3 = 4 \cdot 0,47 = 1,8$. Следовательно, для коэффициента согласованности W_s будем иметь $W_s = 1 - H/H_{\max} = 1 - 0/1,8 = 1$. Отсюда можно сделать заключение о согласованности во мнениях всех экспертов в полном объеме при ранжировке текста 1. В соответствии с этим, по параметру r_{23} также можно принимать значение коэффициентов в пределах единицы. В этом случае ввиду равенства вычисленных двух коэффициентов конкордации предпочтение можно отдать третьему тексту машинного перевода, так как этот ранг по значимости выше, чем первый ранг.

4. Алгоритм работы устройства для анализа экспертных мнений в ИАС

Рассмотрим работу устройства системы индикации в цифровой среде ИАС.

Принцип работы данного устройства (Рис. 1), согласно, например, матрице 1, заключается в следующем.

Поясним нумерацию блоков: 1 – клавишный блок ввода, 2 – первый сумматор, 3 – второй сумматор, 4 – первый блок деления, 5 – блок

умножения и логарифмирования, 6 – первый блок вычисления коэффициента, 7 – второй блок вычисления коэффициента, 8 – первый комплексный формирователь, 9 – второй блок деления, 10 – первый блок вычитания, 11 – второй комплексный формирователь, 12 – третий блок деления, 13 – второй блок вычитания.

Для вычисления первой суммы из выражения (1), так как число переводов равно единице, с первого выхода клавишного блока ввода 1 сигнал, соответствующий единице, подается на вход первого сумматора 2. Для вычисления второй суммы из выражения (1) с учетом того, что число групп экспертов ($F_n(L)$) равно единице), сигнал, соответствующий единице, со второго выхода клавишного блока ввода, подается на вход второго сумматора 3. Для вычисления величины P_{ni} с третьего и одиннадцатого выходов клавишного блока ввода сигналы, соответствующие числам экспертов ($L = 3$) и экспертов, сошедшихся во мнениях ($L_{ni} = 2$), подают, соответственно, на первый и второй входы первого блока деления 4. Далее выходные сигналы первого, второго сумматоров и первого блока деления подают соответственно на первое, второе и третье плечо блока умножения и логарифмирования 5. После проведения соответствующих операций в этом блоке его выходной сигнал направляют на второе плечо второго блока деления 9. В устройстве для вычисления параметров Q и R сигналы с шестого, двенадцатого, седьмого и тринадцатого выходов клавишного блока ввода, соответствующие числам экспертов и письменных текстов, подают, соответственно, на первое и второе плечо первого блока вычисления коэффициента 6 и второго блока вычисления коэффициента 7.

Выходные сигналы первого и второго блоков вычисления коэффициентов, характеризующие величины коэффициентов Q и R , далее поступают на первое и второе плечо, соответственно, первого комплексного формирователя 8 и второго комплексного формирователя 11. Одновременно с этим, сигналы, соответствующие числам экспертов и письменных текстов, с четвертого, пятого, восьмого и девятого выходов клавишных блоков ввода, соответственно, подают на третье и четвертое плечо этих двух формирователей.

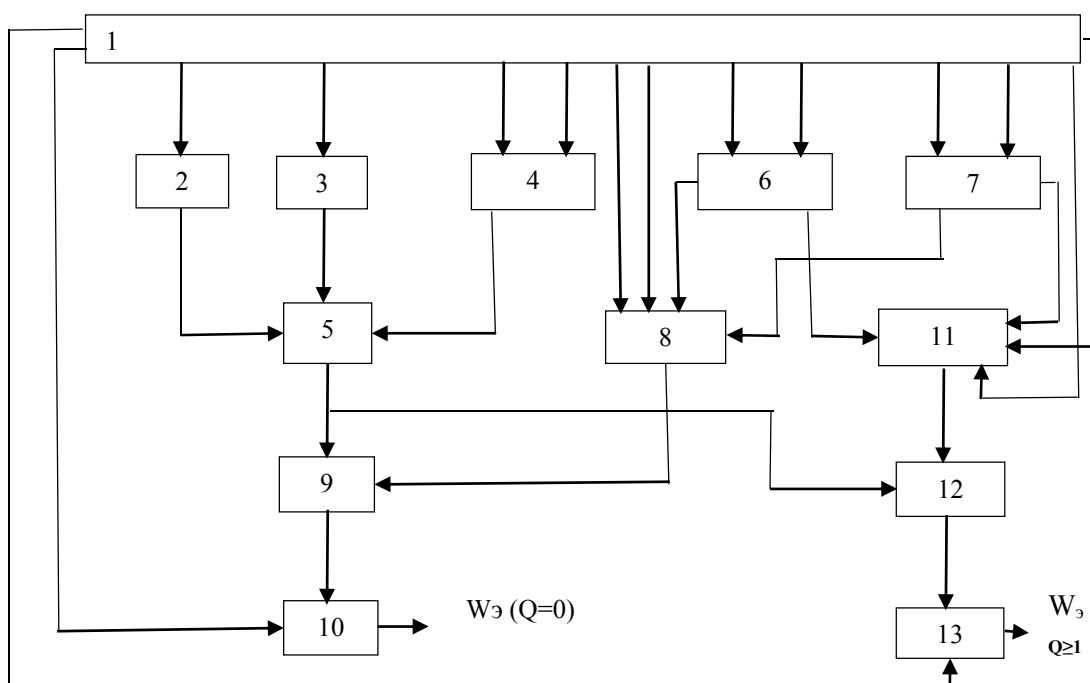


Рис. 1. Устройство системы индикации в цифровой среде ИАС

После этого выходной сигнал первого комплексного формирователя, обусловленный равенством $Q = 0$ и вычисленный по формуле (2), поступает на первое плечо второго блока деления. Сигнал с выхода этого блока, осуществляющий операцию $H/H_{\max}$, далее поступает на первый вход первого блока вычитания 10. Так как на второй вход этого блока поступает сигнал с десятого выхода клавишного блока ввода, соответствующий единице, то благодаря этому в этом блоке производится операция выражения $W_z = 1 - H/H_{\max}$. Другими словами, по выходному сигналу первого блока вычитания вычисляется коэффициент согласованности мнений экспертов при $Q = 0$.

В случае $Q \geq 1$, определение коэффициента W_z сводится к тому, что благодаря поступлениям выходных сигналов первого и второго блоков вычисления коэффициентов и клавишного блока ввода (восьмой и девятый выходы) на соответствующие плечи второго комплексного формирователя 11, на его выходе образуется сигнал, рассчитанный по формуле (3). Далее этот сигнал поступает на первый вход третьего блока деления 12. Согласно устройству на второй вход этого блока поступает сигнал с выхода блока умножения и логарифмирования, характеризующий

величину H . Благодаря этому на выходе третьего блока деления образуется сигнал, обусловленный делением $H/H_{\max}$. После этого для вычисления коэффициента W_z , выходной сигнал с выхода третьего блока деления подается на первый вход второго блока вычитания 13.

В силу того, что сигнал с двенадцатого выхода клавишного блока ввода, соответствующий единице, поступает на второй вход второго блока вычитания, все это вкуче дает возможность определить коэффициент конкордации согласно формуле $W_z = 1 - H/H_{\max}$.

По принципу действия предлагаемого устройства, цифровые значения клавишного блока ввода выбираются по данным матрицы, характеризующей число экспертов и ранжируемых объектов, а также значений параметров $F_n(L)$ и L_{ni} .

Заключение

По полученным результатам можно сделать следующие выводы:

- на основе энтропийно-информационного подхода показана возможность оценивания качества текстов машинных переводов при их ранжировании экспертами-переводчиками;

— для автоматического вычисления коэффициента конкордации разработана и предложена структурная схема устройства, реализующего рассматриваемый подход оценивания качества переведенных текстов.

Литература

1. Полтавский А.В., Русыева Е.Ю. Entropy Evaluation of Data Flows Using an Integrative Approach / Proceedings of the 15th International Conference Management of Large-Scale System Development (MLSD). Moscow: IEEE, 2022. p.1-4. doi: 10.1109/MLSD55143.2022.9934166
2. Introducing ChatGPT and Whisper APIs. Developers can now integrate ChatGPT and Whisper models into their apps and products through our API. Электронный ресурс. <https://openai.com/blog/introducing-chatgpt-and-whisper-apis>. Дата обращения 27.04.2023.
3. Попков Ю.С. Математическая демоэкономика. Макросистемный подход. М.: ЛЕНАНД. 2021. 560 с.
4. Ларичев О. И. Вербальный анализ решений. М.: Наука, 2006. 181 с.
5. Нечеткие множества в моделях управления и искусственного интеллекта / А. Н. Аверкин, И. З. Батыршин, А. Ф. Блишун и др.; Под ред. Д. А. Поспелова. М.: Наука, 1986. 312 с.
6. Алескеров Ф. Т., Хабина Э. Л., Шварц Д. А. Бинарные отношения, графы и коллективные решения. М.: ГУ ВШЭ, 2006. 298 с.
7. Вольский В. И., Лезина З. М. Голосование в малых группах: процедуры и методы сравнительного анализа. М.: Наука, 1991. 192 с.
8. Китаев Н. Н. Групповые экспертные оценки. М.: Знание, 1975. 64 с.
9. Миркин Б. Г. Проблема группового выбора. М.: Наука, 1974. 256 с.
10. Multi-objective Group Decision Making: Methods, Software and Applications with Fuzzy Set Techniques / J. Lu, G. Zhang, D. Ruan, F. Wu. London: Imperial College Press, 2007. 400 pp.
11. Эрроу К. Дж. Коллективный выбор и индивидуальные ценности. М.: ГУ ВШЭ, 2004. 204 с.
12. Льюис Р. Д., Райфа Г. Игры и решения. М.: Иностранная литература, 1961. 642 с.
13. Жиликова Е.В., Ларин С.П. Методы и приемы проведения независимой экспертизы // Вестник Воронежского государственного университета. Серия: Экономика и управление, 2009, №2, с.108-116.
14. Венедиктов А.А. О показателе согласованности экспертных оценок // Вооружение и экономика, 2018, №3 (45), с. 52-66.
15. Буравлев А.И. О задачах многокритериального выбора // Вооружение и экономика, 2021, №1 (55), с. 131-138.
16. Клевцев Н.И. Сравнительный анализ методов многокритериального ранжирования альтернатив // Научный вестник: Финансы, банки, инвестиции, 2018, №2, с.153-163.
17. Ларичев О. И. Теория и методы принятия решений. М.: Логос, 2002. 296 с.
18. Оценка согласованности мнений экспертов. Электронный ресурс. Дата обращения 05.03.2023. <https://studfile.net/preview/9186952/page:22/>
19. Бурков В. Н., Новиков Д. А. Элементы теории графов. Электронный ресурс. Дата обращения 20.03.2023. http://www.mtas.ru/start/t_garf.pdf

Русыева Елена Юрьевна. Федеральное государственное бюджетное учреждение науки Институт проблем управления им. В. А. Трапезникова Российской академии наук. Старший научный сотрудник, кандидат философских наук. Область научных интересов: философия управления, информационные технологии лингвистического анализа текстов, когнитивное моделирование, концептуализация вычислительных моделей ИАС. E-mail: rusyaeva@ipu.ru.

Ахобадзе Гурами Николаевич. Федеральное государственное бюджетное учреждение науки Институт проблем управления им. В. А. Трапезникова Российской академии наук. Ведущий научный сотрудник, доктор технических наук. Область научных интересов: вычислительная техника, информационные технологии ИАС. E-mail: ahogur@yandex.ru

Полтавский Александр Васильевич. Федеральное государственное бюджетное учреждение науки Институт проблем управления им. В. А. Трапезникова Российской академии наук. Ведущий научный сотрудник, доктор технических наук, профессор. Область научных интересов: вычислительная техника, информационные технологии ИАС. E-mail: avp57avp@yandex.ru

Entropy-Information Approach for Evaluating the Quality of Machine Translation Texts

G. N. Akhobdze, E. Yu. Rusyaeva, A. V. Poltavsky

V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia

Abstract. Based on the entropy-information approach, a method for assessing the quality of machine translation texts is proposed. An analysis is given of the dispersion and entropy coefficients of concordance used to assess the consistency of expert opinions with close rankings of various objects. The prospects of using the entropy coefficient of concordance, which makes it possible to fix the fact of division of opinions into two opposing groups, are substantiated. This provision is important for the study, since in this method of

expert evaluation of text translations, it is important to take into account the different opinions of several experts involved in the examination. Examples of calculating the entropy coefficient of concordance with a changing rank system, the number of experts and ranked evaluation objects are given.

Keywords: concordation coefficient, text ranking, expert opinion, information entropy, rank, expert opinion matrix, probabilistic parameters.

DOI 10.14357/20718632230402

EDN EOBUMC

References

1. Poltavsky A.V., Rusyaeva E.Yu. Entropy Evaluation of Data Flows Using an Integrative Approach. In: Proceedings of the 15th International Conference Management of Large-Scale System Development MLSD 2022. 26–28 September 2022, Moscow, Russia: IEEE; 2022. P.1-4. doi: 10.1109/MLSD55143.2022.9934166
2. Introducing ChatGPT and Whisper APIs. Developers can now integrate ChatGPT and Whisper models into their apps and products through our API. Available from: <https://openai.com/blog/introducing-chatgpt-and-whisper-apis>. [Accessed 27 April 2023]
3. Popkov Yu.S. Mathematical Demoeconomics. macrosystem approach. M.: LENAND. 2021. 560 p. (In Russ).
4. O. I. Larichev, Verbal decision analysis. M.: Nauka, 2006. 181 p. (In Russ).
5. Averkin A.N., Batyrshin I.Z., Blishun A.F. et al. Fuzzy sets in control and artificial intelligence models; Ed. D. A. Pospelova. M.: Nauka, 1986. 312 p. (In Russ).
6. Aleskerov F. T., Khabina E. L., Shvarts D. A. Binary relations, graphs and collective solutions. M.: GU VSHE, 2006. - 298 p. (In Russ).
7. Volsky V. I., Lezina Z. M. Voting in small groups: procedures and methods of comparative analysis. M.: Nauka, 1991. 192 p. (In Russ).
8. Kitaev N. N. Group expert assessments. M.: Knowledge, 1975. 64 p. (In Russ).
9. Mirkin BG The problem of group choice. M.: Nauka, 1974. 256 p. (In Russ).
10. Multi-objective Group Decision Making: Methods, Software and Applications With Fuzzy Set Techniques; J. Lu, G. Zhang, D. Ruan, F. Wu. London: Imperial College Press, 2007. 400 pp.
11. Arrow K. J. Collective choice and individual values. M.: GU VSHE, 2004. 204 p. (In Russ).
12. Lewis R. D., Raifa G. Games and solutions. M.: Foreign Literature, 1961. 642 p. (In Russ).
13. Zhilyakova E.V., Larin S.P. Methods and techniques for conducting an independent examination. Bulletin of the Voronezh State University. Series: Economics and Management. 2009; 2:108-116. (In Russ).
14. Venediktov A.A. On the indicator of consistency of expert assessments. Armament and Economics. 2018;3 (45):52-66 (In Russ). Available from: <http://www.viek.ru/45/52-66.pdf>. [Accessed 17 April 2023]
15. Buravlev A.I. On the tasks of multi-criteria choice. Armament and Economics. 2021;1 (55): 131-138. (In Russ). Available from: <http://www.viek.ru/55/131-138.pdf>. [Accessed 07 April 2023]
16. Klevets N.I. Comparative analysis of methods for multi-criteria ranking of alternatives // Scientific Bulletin: Finance, banks, investments. 2018;2:153-163 (In Russ).
17. Larichev O. I. Theory and methods of making decisions. Moscow: Logos; 2002. 296 p. (In Russ)
18. Evaluation of the consistency of expert opinions. Available from: <https://studfile.net/preview/9186952/page:22/> [Accessed 05 April 2023]. (In Russ)
19. V. N. Burkov and D. A. Novikov, Elements of Graph Theory. Available from: http://www.mtas.ru/start/t_garf.pdf [Accessed 03 May 2023]. (In Russ)

Rusyaeva Elena Y. PhD, V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences 65 Profsoyuznaya street, Moscow 117997, Russia. Research interests: philosophy of management, information technologies of linguistic analysis of texts, cognitive modeling, conceptualization of computational models of IAS. E-mail: rusyaeva@ipu.ru.

Akhobadze Gurami N. Professor, Doctor of technical sciences, V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences 65 Profsoyuznaya street, Moscow 117997, Russia. Area of scientific interests: computer engineering, information technologies IAS. E-mail: ahogur@yandex.ru

Poltavsky Alexander V. Professor, Doctor of technical sciences, V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences 65 Profsoyuznaya street, Moscow 117997, Russia. Area of scientific interests: computer engineering, information technologies IAS. E-mail: avp57avp@yandex.ru