

4. *Friedman Milton*. The Optimum Quantity of Money in the Optimum Quantity of Money and Other Essays. Chicago: Aldine Publishing Company, 1969.

5. *Jorda O., Singh S.R., Taylor A.M.* Longer-Run Economic Consequences of Pandemics. Federal Reserve Bank of San Francisco. Working paper. 2020.

Literatura

1. *Vecherskaya S.E.* Postanovka I algoritm reshenia zadachi optimizatsii upravleniya. Uchebno-metodicheskoe posobie. M.: Redaktsionno-izdatelskii dom Rossiiskogo novogo universiteta, 2018.

2. *Gorborukov V.* Sovremennaya transportnaya logistika. Elektronnoe izdatelstvo Ridero, 2018.

3. *Enikolopov R.* Vvedenie // Ekonomicheskaya politika vo vremena COVID-19. M., 2020. S. 3–10. URL: [https://www.nes.ru/files/COVID19\(21apr2020\).pdf](https://www.nes.ru/files/COVID19(21apr2020).pdf) (data obrashcheniya: 03.05.2020).

4. *Friedman Milton*. The Optimum Quantity of Money in the Optimum Quantity of Money and Other Essays. Chicago: Aldine Publishing Company, 1969.

5. *Jorda O., Singh S.R., Taylor A.M.* Longer-Run Economic Consequences of Pandemics. Federal Reserve Bank of San Francisco. Working paper. 2020.

DOI: 10.25586/RNU.V9I187.20.02.P.066

УДК 330.46

И.С. Клименко, О.И. Даниленко, В.С. Зубарев

К ПРОБЛЕМЕ ВЫДЕЛЕНИЯ ИНФОРМАЦИИ ИЗ СООБЩЕНИЙ, СОДЕРЖАЩИХ ИНФОРМАЦИОННЫЙ ШУМ И НЕДОСТОВЕРНЫЕ СВЕДЕНИЯ

Рассматриваются содержательные аспекты декомпозиции принимаемых сообщений на основные составляющие: информацию, информационный шум и дезинформацию. Обсуждаются возможные подходы к их выявлению и разграничению применительно к принятию решений и управлению складываемым в них риском. Сформированы оригинальные критерии, ориентирующие лицо, принимающее решение, на достижение гарантированного результата, проявление максимального оптимизма и обеспечение оптимального решения в задаче формирования управляющего воздействия. *Ключевые слова:* неопределенность, декомпозиция, сведения, дезинформация, целевой эффект, критерий, ресурсы управления.

I.S. Klimenko, O.I. Danilenko, V.S. Zubarev

TO THE PROBLEM OF ISSUING INFORMATION FROM MESSAGES CONTAINING INFORMATION NOISE AND UNRELIABLE INFORMATION

Substantive aspects of the decomposition of received messages into the main components: information, information noise and misinformation are considered. Possible approaches to their identification and differentiation with respect to decision-making and risk management are discussed. Original criteria are formed to guide a decision maker to achieve a guaranteed result, to show maximum optimism and to ensure an optimal decision in the task of forming a managing influence.

Keywords: uncertainty, decomposition, information, misinformation, target effect, criterion, management resources.

Введение

Принятие решений в условиях неопределенности и порождаемого ею риска относится к числу наиболее сложных задач управления. Согласно фундаментальному критерию Харкевича [11], те или иные сообщения, отражающие состояние ситуации принятия решения и используемые лицом, принимающим решение (ЛПР) по назначению, в общем случае могут содержать три типа информационных объектов: информацию, информационный шум и дезинформацию [1; 3].

Различная ценность указанных составляющих сообщения неизбежно порождает задачу их адекватной интерпретации с целью выделения собственно информации, обеспечивающей повышение вероятности достижения цели. В такой постановке целесообразно опираться на прагматический аспект теории информации, согласно которому информация определяется как мера снятия неопределенности для достижения цели [12]. Это означает, что недостоверные, неактуальные, несвоевременные, неполные и неточные сведения автоматически не подпадают под понятие информации, поскольку их использование будет уменьшать вероятность достижения цели.

Наличие неоднородных связей между различными компонентами сообщения требует рассматривать их как сложную систему, имманентно содержащую оптимизационную задачу, в рамках которой увеличение количества извлекаемой информации представляет собой целевую функцию, а наличие компонентов, снижающих вероятность достижения цели, играет роль ограничений.

Следует отметить следующее немаловажное обстоятельство. Очевидно, что использование дезинформации для принятия решения снижает вероятность достижения цели по определению. Однако и отражение в решении сведений, представляющих собой информационный шум (не изменяющий вероятность достижения цели при их реализации), неизбежно снижает эффективность информационных процессов в силу ухудшения показателей оперативности и ресурсоемкости (бесполезных затрат, как правило, лимитированных ресурсов и времени). Как следствие, может возникнуть ситуация, при которой необходимых ресурсов и/или времени может попросту не хватить для достижения цели.

Таким образом, задача корректной декомпозиции состава и структуры исходного сообщения на вышеуказанные компоненты приобретает высокую степень актуальности.

Декомпозиция сообщения

В общем случае указанные компоненты сообщения в виде определенных сведений могут быть рассредоточены в пределах сообщения и встречаться в них неоднократно. Поэтому представляется целесообразным рассматривать их как подмножества полного множества сведений, составляющих исходное сообщение $S_{\text{исх}}$. Тогда

$$S_{\text{и}}, S_{\text{ш}}, S_{\text{д}} \subset S_{\text{исх}} \text{ и } S_{\text{и}} \cup S_{\text{ш}} \cup S_{\text{д}}, \quad (1)$$

где $S_{\text{и}}, S_{\text{ш}}, S_{\text{д}}$ – подлежащие выявлению информационные объекты, а именно сведения, несущие информацию, информационный шум и дезинформацию. Подчеркнем, что априори распределение этих сведений по подмножествам неизвестно, что приводит к необходимости решения задачи их корректной декомпозиции.

Формально задача декомпозиции сообщения может быть представлена следующим образом:

$$R_A = K(F(S_{\text{исх}}) \rightarrow \{S_{\text{и}}, S_{\text{ш}}, S_A\}), \quad (2)$$

где R_A – результат декомпозиции; K – критерии или решающие правила, используемые ЛПР при идентификации сообщения; F – оператор преобразования.

Сложность задачи состоит в том, что априори сведения разного типа, в целом составляющие исходное сообщение, не имеют четких границ, поэтому структура сообщения, определяемая отношениями (неоднородными связями) между ними, однозначно не просматривается. Как следствие, принцип выбора (декомпозиции) не может быть строго формализован. Следовательно, ЛПР приходится руководствоваться собственной системой предпочтений при выборе критериев качества выявляемых им компонентов сообщения.

Под качеством сообщения будем понимать совокупность существенных свойств несомых им сведений, определяющих степень их пригодности для использования по назначению, в нашем случае – для формирования управляющего воздействия, адекватного реальной ситуации принятия решения.

Для интерпретации сообщения ЛПР, как известно, можно использовать три основных вида мыслительных процедур: расчетные, логические и эвристические, придерживаясь при этом критерия минимума эвристик [8]. Как правило, по предпочтительности применения указанные процедуры связаны отношением строгого порядка:

$$R_p \succ R_A \succ R_3, \quad (3)$$

где R_p , R_A и R_3 – результаты декомпозиции на основе применения расчетных, логических и эвристических процедур соответственно.

Что же касается подлежащих выявлению и разграничению подмножеств $S_{\text{и}}$, $S_{\text{ш}}$, S_A , то по степени исходной неопределенности и, соответственно, сложности интерпретации они не выглядят эквивалентными.

Действительно, рассмотрим существенные свойства подлежащих выявлению составляющих сообщения. Начнем с информационного шума, который, по нашему мнению, в наибольшей степени поддается распознаванию, поскольку его можно считать (по аналогии с теорией связи) аддитивной помехой, которая добавляется к полезным компонентам.

Согласно [1; 3] к существенным качественным характеристикам информационного шума, который в основном содержит достоверные сведения, относятся:

- неактуальность;
- несвоевременность;
- тривиальность;
- невозможность декодирования;
- неточность;
- неполнота;
- другие.

Очевидно, что тривиальные для себя сведения ЛПР способен выделить оперативно, не прибегая к сколько-нибудь сложным процедурам. *Неактуальные* и, как следствие, *несвоевременные* сведения также могут быть выявлены достаточно уверенно путем сопоставления с результатом целеполагания. *Не декодируемые*, т.е. недоступные для адекватной

Клименко И.С., Даниленко О.И., Зубарев В.С. К проблеме выделения информации...

интерпретации, компоненты сообщения целесообразно поначалу отнести к отдельной группе, в которой сосредотачиваются сведения, подлежащие дополнительному анализу. В частности, к этой группе ЛПР может отнести сведения, которые он посредством выполнения эвристических процедур оценивает, как *неполные*, т.е. не отвечающие признаку необходимости и достаточности для принятия обоснованного решения, адекватного реальной ситуации. Сюда же тяготеют сведения, которые представляются *неточными* в количественном отношении. Тем не менее некоторые сведения, имеющие признаки информационного шума, ЛПР может отнести к *категории неопределенных*, не исключая их потенциальной ценности для принятия решения.

После вычленения из сообщения информационного шума задача выявления информации упрощается. Речь идет о выделении сведений, обладающих очевидными свойствами достоверности, актуальности и своевременности, а также точности и полноты. Некоторые из них, однако, ЛПР может счесть сомнительными и также отнести к *категории неопределенных*.

При этом следует особо подчеркнуть, что неполные и неточно параметризованные данные могут не только нести информационный шум, но и оказаться умышленной дезинформацией. Естественно, для их адекватной декомпозиции и интерпретации потребуются дополнительные информационные и интеллектуальные (экспертные) ресурсы, а также дополнительный промежуток времени [2].

Поэтому в общем случае на начальном этапе анализа структуры и семантики сообщения ЛПР может, наряду с информационным шумом и полезной информацией, выделить определенные сведения, которые несут заведомые признаки недостоверности, т.е. искажают искомую оценку ситуации принятия решений. И снова часть недостоверных сведений может быть сохранена для *дополнительной проверки*, в том числе в силу их привлекательности для ЛПР.

Недостоверные сведения могут быть неоднородными в смысле природы и причин их появления в составе сообщения. В частности, они могут представлять собой:

- неумышленные ошибки и искажения действительности;
- умышленное искажение реальной действительности;
- целенаправленную дезинформацию;
- распространение слухов;
- разного рода мистификации.

Одним из признаков умышленной недостоверности сообщения является его *сенсационный характер*, который направлен на привлечение к нему внимания значительной группы лиц, которые могут счесть его потенциально полезным для себя. При этом источник такого сообщения будет стремиться к тому, чтобы содержание сообщения выглядело *актуальным* для многих лиц из этой группы. Следующая цель источника: признавая содержание сообщения актуальным для себя, ЛПР связывает с ним определенную *ценность* с точки зрения возрастания вероятности достижения своей конкретной цели.

Признание ЛПР наличия у сообщения признаков ценности логически приводит к выводу о том, что оно обладает определенной рыночной стоимостью, на что и рассчитывает источник недостоверного сообщения.

По некоторым оценкам, более чем в половине случаев сообщения сенсационного характера недостоверны, за исключением передачи их от априори надежного источника,

заинтересованного в успешном достижении цели (в частности, разрешения проблемной ситуации). В таком случае сообщение подобного рода действительно может позволить получить существенный выигрыш в ресурсах для достижения цели.

Таким образом, в процессе идентификации сообщения несомые им сведения ЛПР перераспределяет, формируя четыре подмножества:

$$R_A^* = \{S_n^*, S_c^*, S_{ш}^*, S_A^*\}, \quad (4)$$

где R_A^* – результат декомпозиции сообщения после выполнения идентификации; S_n^* – сведения, идентифицированные ЛПР как полезная информация; S_c^* – сведения, порождающие остаточную неопределенность, в том числе не декодированные и не идентифицированные сведения, среди которых ЛПР не исключает наличие полезной информации; $S_{ш}^*$ и S_A^* – сведения, уверенно отнесенные ЛПР к информационному шуму и дезинформации соответственно, которые он в дальнейшем полностью выводит из рассмотрения.

Как видим, в общем случае подмножество S_c^* содержит в латентной форме все рассматриваемые типы информационных объектов, нуждающихся в дополнительной интерпретации с целью извлечения скрытой полезной информации:

$$S_c^* = \{s_n^*, s_{ш}^*, s_A^*\}, \quad (5)$$

где s_n^* , $s_{ш}^*$, s_A^* – однозначно не идентифицированные сведения, несущие соответственно информацию, шум и дезинформацию. Если ЛПР располагает некоторым запасом времени для доведения решения, то рациональным представляется проведение дальнейшей интерпретации сведений, составляющих подмножество S_c^* и выделение из него элемента s_n^* с целью пополнения подмножества S_n^* ценной информацией.

Формирование критериев выбора решения

Вернемся к множеству (1) и привлечем к рассмотрению задачи методологию взвешенного выбора решений в условиях статистической неопределенности («игры с природой») [5; 7; 10]. В практике выбора решений (стратегий) в условиях неопределенности центральным моментом является определение критерия, которым будет руководствоваться ЛПР в соответствии со своим отношением к риску. Условия ситуации принятия решения также оказывают существенное влияние на выбор ЛПР, поскольку в одних случаях допустимым является только гарантированный результат, а в других – возможен определенный риск. Под взвешенным выбором принято понимать такое решение, которое позволяет обеспечить оптимальное соотношение между оптимизмом и пессимизмом с точки зрения конкретного ЛПР.

Практика принятия решений в условиях неопределенности выработала достаточно широкий спектр критериев – от крайне осторожных до предельно оптимистичных. Взвешенный выбор при этом осуществляется за счет конструирования комбинированных критериев, среди которых наиболее распространенным является критерий Гурвица [9], либо на основе применения ранговых измерительных шкал [4; 6].

Рассматриваемую в настоящей работе задачу, по нашему мнению, целесообразно в первую очередь отнести к классу задач, ориентированных на достижение *гарантированного результата*. Это ограничение выглядит вполне обоснованным в ответственных ситуациях, когда ЛПР к тому же вынужден принимать решения в условиях дефицита времени.

Клименко И.С., Даниленко О.И., Зубарев В.С. К проблеме выделения информации...

Тогда естественным взвешенным решением является выбор на базе семантики максимального критерия Вальда [13].

В нашем случае такое решение будет сводиться к использованию по назначению исключительно информационной составляющей сообщения S_n^* при полном пренебрежении всеми остальными – с целью исключения любых элементов риска. Формализуем соответствующий такому решению критерий гарантированного результата:

$$K_{\text{гар}} : (\forall r_A^* \in S_n^*), \tag{6}$$

где r_A^* – элемент множества R_A^* . Этот критерий задает правило, согласно которому решение считается оптимальным с точки зрения обеспечения гарантированного возрастания вероятности достижения целевого эффекта, если оно основано исключительно на использовании детерминированной информационной компоненты исходного сообщения S_n^* .

Недостатком такого подхода является, естественно, риск потери благоприятных возможностей, которые могут быть связаны с наличием в составе сомнительных сведений S_c^* невыявленной информационной составляющей s_n^* . Однако для ее надежного выявления ЛПР должен, как отмечалось выше, располагать запасом времени и ресурсов для снятия неопределенности относительно состава и структуры подмножества S_c^* .

Наиболее оптимистичным решением, в духе семантики критерия максимакса, в нашем случае будет принятие эвристической гипотезы, согласно которой подмножество S_c^* не содержит шумовой и дезинформационной составляющих и может быть полностью включено в множество S_n^* :

$$K_{\text{max}} : (\forall r_A^* \in \{S_n^*, S_c^*\}). \tag{7}$$

Этот критерий задает правило, согласно которому решение считается оптимальным с точки зрения возможного максимального возрастания вероятности достижения целевого эффекта, если оно основано на использовании, наряду с детерминированной информационной компонентой исходного сообщения, всех неопределенных сведений, априори не отнесенных к шумовой и дезинформационным составляющим.

Естественно, при этом существенно возрастает риск получения отрицательного эффекта, когда негативные последствия чрезмерно оптимистичного решения могут обернуться серьезной неудачей.

Поэтому в случае, когда ЛПР располагает запасом времени и соответствующими ресурсами, рациональным решением будет, как уже отмечалось выше, снятие неопределенности относительно содержания и структуры составляющей S_c^* для выявления скрытой в ней дополнительной информации s_n^* . Для этого могут быть привлечены данные научной литературы и патентного фонда, авторитетные эксперты соответствующей предметной области, использованы методы группового выбора решений.

Иными словами, ресурсы управления могут быть направлены на увеличение количества информации, извлекаемой из сообщения за счет проверки и экспертного оценивания неопределенных (сомнительных) сведений. Формализуем соответствующий критерий, который можно рассматривать как объективно оптимальный в рамках системы предпочтений ЛПР при условии, что промежуток времени на принятие окончательного решения $T_{\text{реш}}$ не превышает допустимого значения $T_{\text{доп}}$:

$$K_{\text{опт}} : (\forall r_A^* \in \{S_n^*, s_c^*\}) \text{ при } T_{\text{реш}} \leq T_{\text{доп}}. \tag{8}$$

Тем самым открывается дополнительная возможность управления риском принимаемых решений. Затраты ресурсов при этом можно рассматривать как плату за увеличение шансов на успешный результат (увеличение вероятности достижения цели).

Заключение

Проведенный в настоящей работе анализ содержания и структуры данных, используемых для принятия управленческих решений, позволил сформулировать критерии гарантированного результата и максимального оптимизма для ЛПР, стремящегося выделить ценную информацию из сообщений, содержащих смесь различных информационных объектов. Кроме того, развитый в работе подход открывает дополнительную возможность взвешенного выбора решений в пределах доступных временных ресурсов управления.

Литература

1. *Завгородний В.И.* Системное управление информационными рисками. Выбор механизмов защиты // Проблемы управления. 2009. № 1. С. 53–58.
2. *Клименко И.С., Белова Н.А., Шарипова Л.В.* К проблеме определения ценности информации в условиях информационного общества // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2018. Вып. 2. С. 54–62.
3. *Клименко И.С., Коровко П.Г., Шарипова Л.В.* К проблеме оценивания качества управленческих решений и эффективности управления // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2017. Вып. 1. С. 53–57.
4. *Клименко И.С., Плуталов М.А.* О ранжировании критериев выбора решений, сформированных на матрице риска // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2017. Вып. 2. С. 63–65.
5. *Клименко И.С., Плуталов М.А., Чеботарев Г.А.* К вопросу об оценивании оптимизма критериев выбора стратегий в «игре с природой» // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2015. Вып. 2. С. 19–23.
6. *Клименко И.С., Плуталов М.А., Чеботарев Г.А.* К формированию ранговой шкалы оптимизма критериев выбора решений в «игре с природой» // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2016. Вып. 3. С. 19–23.
7. *Клименко И.С., Плуталов М.А., Чеботарев Г.А.* Сравнительный анализ критериев выбора стратегий в «игре с природой» // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2015. Вып. 1. С. 57–61.
8. *Клименко И.С., Шарипова Л.В.* К проблеме системного анализа телекоммуникационных процессов // Вестник Российского нового университета. Серия «Сложные системы: модели, анализ и управление». 2016. Вып. 1–2. С. 82–86.
9. *Лабскер Л.Г.* К вопросу о проблеме сглаживания критерием Гурвица и экономическое приложении // Инновации и инвестиции. 2016. № 6. С. 134–145.
10. *Лабскер Л.Г.* Теория критериев оптимальности и экономические решения: монография. М.: КноРус, 2012. 744 с.
11. *Харкевич А.А.* О ценности информации // Проблемы кибернетики. 1960. Вып. 4.
12. *Floridi L.* The Philosophy of Information // *Metaphilosophy* / ed. by A.T. Marsoobian. Oxford, 2010. Vol. 41, № 3. P. 420–442.
13. *Wald A.* Contribution of the Theory of Statistical Estimation and Testing Hypothesis // *Annals Math. Statist.* 1939. Vol. 10. P. 299–326.

Literatura

1. *Zavgorodniy V.I.* Sistemnoe upravlenie informatsionnymi riskami. Vybora mekhanizmov zashchity // Problemy upravleniya. 2009. № 1. S. 53–58.
2. *Klimenko I.S., Belova N.A., Sharapova L.V.* K probleme opredeleniya tsennosti informatsii v usloviyakh informatsionnogo obshchestva // Vestnik Rossiyskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2018. Vyp. 2. S. 54–62.
3. *Klimenko I.S., Korovko P.G., Sharapova L.V.* K probleme otsenivaniya kachestva upravlencheskikh resheniy i effektivnosti upravleniya // Vestnik Rossiyskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2017. Vyp. 1. S. 53–57.
4. *Klimenko I.S., Plutalov M.A.* O ranzhirovaniy kriteriev vybora resheniy, sformirovannykh na matritse riska // Vestnik Rossiyskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2017. Vyp. 2. S. 63–65.
5. *Klimenko I.S., Plutalov M.A., Chebotarev G.A.* K voprosu ob otsenivaniy optimizma kriteriev vybora strategiy v "igre s prirodoy" // Vestnik Rossiyskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2015. Vyp. 2. S. 19–23.
6. *Klimenko I.S., Plutalov M.A., Chebotarev G.A.* K formirovaniyu rangovoy shkaly optimizma kriteriev vybora resheniy v "igre s prirodoy" // Vestnik Rossiyskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2016. Vyp. 3. S. 19–23.
7. *Klimenko I.S., Plutalov M.A., Chebotarev G.A.* Sravnitel'nyy analiz kriteriev vybora strategiy v "igre s prirodoy" // Vestnik Rossiyskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2015. Vyp. 1. S. 57–61.
8. *Klimenko I.S., Sharapova L.V.* K probleme sistemnogo analiza telekommunikatsionnykh protsessov // Vestnik Rossiyskogo novogo universiteta. Seriya "Slozhnye sistemy: modeli, analiz i upravlenie". 2016. Vyp. 1–2. S. 82–86.
9. *Labsker L.G.* K voprosu o probleme sglazhivaniya kriteriem Gurvitsa i ekonomicheskoe prilozhenie // Innovatsii i investitsii. 2016. № 6. S. 134–145.
10. *Labsker L.G.* Teoriya kriteriev optimal'nosti i ekonomicheskie resheniya: monografiya. M.: KnoRus, 2012. 744 s.
11. *Kharkevich A.A.* O tsennosti informatsii // Problemy kibernetiki. 1960. Vyp. 4.
12. *Floridi L.* The Philosophy of Information // Metaphilosophy / ed. by A.T. Marssoobian. Oxford, 2010. Vol. 41, № 3. P. 420–442.
13. *Wald A.* Contribution of the Theory of Statistical Estimation and Testing Hypothesis // Annuals Math. Statist. 1939. Vol. 10. P. 299–326.

DOI: 10.25586/RNUV9187.20.02.P.073

УДК 003.26

К.А. Ахрамеева, Е.Ю. Герлинг, Д.Ю. Мицковский, С.В. Прудников

ИСПОЛЬЗОВАНИЕ МЕТРИКИ BLEU ДЛЯ ОЦЕНКИ
ЕСТЕСТВЕННОСТИ ТЕКСТА ЛИНГВИСТИЧЕСКИХ СТЕГОСИСТЕМ

Представлены результаты исследования использования метрики BLEU для оценки естественности текста лингвистических стегосистем. Рассмотрены два метода вложения дополнительной информации: метод замены синонимов и метод перестановки слов в предложении. В рамках исследования производится сравнение оценок текста с помощью метрики BLEU при различных стегопреобразованиях текста, а также оценивается влияние изменений, внесенных стегосистемой, на