

Алгоритм улучшения согласованности структурной интероперабельности

An algorithm for improving the consistency of structural interoperability

Дулин С.К.^{1*}, Рябцев А.Б.²
Dulin S.K.^{1*}, Ryabtsev A.B.²

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук; Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте (АО НИИАС), Москва

²Московский физико-технический институт (Государственный университет), Москва

¹Federal Research Center Computer Science and Control of the Russian Academy of Sciences, Joint Stock Company Research and Design Institute for Information Technology, Signalling and Telecommunications in Railway Transportation (JSC NIAS), Moscow, Russian Federation

²Moscow Institute of Physics and Technology (Public university), Moscow

*skdulin@mail.ru



Дулин С.К.



Рябцев А.Б.

Резюме. Цель. Одной из наиболее сложных задач искусственного интеллекта является проблема согласованности совместно формируемой совокупности знаний. При этом существует весьма важный аспект процесса приобретения знаний – наличие средств сопровождения знаний. В общей схеме функционирования процедур сопровождения системы знаний особое место занимают процедуры, связанные с контролем согласованности взаимосвязанных компонентов. Среди направлений в области представления знаний выделилось исследование условий функционирования корпоративной системы знаний, в рамках которого одной из самых открытых проблем оказался поиск предпочтительной структуры интероперабельности компонентов системы знаний. Рассмотрим некоторое множество из n элементов, представляющих собой агенты, информационные системы или компоненты знаний, тогда можно поставить цель – оценить уровень интероперабельности и представить структуру интероперабельности на основании анализа предпочтительности установления отношений между элементами. **Методы.** Таким образом, необходимо учитывать влияние структуры взаимодействующих элементов на мотивацию к интероперабельности на основе определенных признаков или характеристик в структуре элементов, которые способствуют или препятствуют достижению интероперабельности. Было отмечено, что потенциальная возможность установления интероперабельности на основе структуры взаимосвязанных элементов, может быть определена как структурная интероперабельность. Соответственно, для изучения тенденции информационных систем или элементов к интероперабельности в зависимости от соотношения собственных признаков или характеристик была предложена методология структурного соответствия, позволяющая оценивать группы потенциально близких друг к другу элементов на основе аппарата структурной согласованности. Моделирование структурной интероперабельности на основе анализа структуры связей с помощью выбранного критерия согласованности позволяет найти ближайший к исходному множеству согласованный прообраз. Найденный прообраз своими подмножествами указывает на предпочтительную группировку элементов, при которой интероперабельность между ними устанавливается с наименьшей рассогласованностью. **Результаты.** В результате возникает разбиение множества потенциально взаимодействующих элементов на наборы мотивированных к взаимодействию элементов. В работе предлагается и обосновывается алгоритм нахождения для произвольной структуры элементов наиболее близкой согласованной структуры, на основе которой можно сделать вывод о выборе структуры интероперабельности. **Выводы.** Из анализа предлагаемого алгоритма и его альтернативы следует, что при сформулированных условиях существования структурной интероперабельности существует несколько вариантов выбора приемлемой структуры интероперабельности. Однако поиск оптимального варианта потребует перебора среди всех приемлемых вариантов, либо использование разумных эвристик, учитывающих специфику матрицы связности исходного знакового графа. Сравнивая представленные алгоритмы, необходимо заметить, что, алгоритм распространения меток носит оценочный характер, при этом, несмотря на сложность комбинаторных оценок и преобразований, алгоритм уменьшения рассогласованности на основе вектора поворшинных различий представляет собой инструмент перманентного анализа и управления согласованностью наборов мотивированных к интероперабельности элементов.

Abstract. Aim. The consistency of a collectively generated body of knowledge is one of the most difficult problems of artificial intelligence. Additionally, there is a very important aspect of

the knowledge acquisition process, i.e., the availability of knowledge support tools. Within the overall scheme of operation of knowledge system maintenance procedures, those associated with ensuring the consistency of interrelated components hold a special place. Out of the lines of knowledge representation research, the study of the operating procedures of corporate knowledge systems stood out. One of its most unconquered challenges is finding an optimal structure for the interoperability of knowledge system components. Let us consider a set of n elements that represent agents, information systems or knowledge components, then we can set the goal of assessing the level of interoperability and present the structure of interoperability based on an analysis of preferred relationships between elements. **Methods.** Thus, the effect of the structure of interacting elements on the motivation for interoperability is to be taken into account based on certain features or characteristics in the structure of elements that contribute or hinder the achievement of interoperability. It was noted that the potential for establishing interoperability based on a structure of interrelated elements can be defined as structural interoperability. Accordingly, for the purpose of studying the tendency of information systems or elements to be interoperable depending on the correlation of their own features or characteristics, a structural correspondence methodology was proposed that allows evaluating groups of potentially close elements using the structural consistency apparatus. Modelling structural interoperability based on the analysis of the connections structure using the selected consistency criterion allows finding the consistent preimage that is the closest to the original set. The subsets of the found preimage indicate the preferred grouping of elements that enables interoperability with the least inconsistency. **Results.** As a result, the set of potentially interacting elements is divided into sets of elements motivated to interact. This paper proposes and justifies an algorithm for finding – for a random structure of elements – the closest consistent structure that allows concluding on the choice of the interoperability structure. **Conclusions.** Out of the analysis of the proposed algorithm and its alternative follows that, under defined conditions for the existence of structural interoperability, there are several options for an acceptable interoperability structure. However, finding the optimal option will require searching through all acceptable options, or using reasonable heuristics that take into account the specifics of the connectivity matrix of the original signed graph. Comparing the presented algorithms, it must be noted that the label propagation algorithm is evaluative in its nature. Meanwhile, despite the complexity of combinatorial estimates and transformations, the inconsistency reduction algorithm based on the vertex-to-vertex difference vector is a tool for permanent analysis and managing the consistency of sets of elements motivated for interoperability.

Ключевые слова: интероперабельность, структурная согласованность, матрица связности
Keywords: interoperability, structural consistency, connectivity matrix

Для цитирования: Дулин С. К., Рябцев А.Б. Алгоритм улучшения согласованности структурной интероперабельности // Надежность. 2024. №2. С. 8-15. <https://doi.org/10.21683/1729-2646-2024-24-2-8-15>

For citation: Dulin S.K., Ryabtsev A.B. An algorithm for improving the consistency of structural interoperability. Dependability 2024;2:8-15. <https://doi.org/10.21683/1729-2646-2024-24-2-8-15>

Поступила: 21.03.2024 г. / **После доработки:** 19.04.2024 / **К печати:** 10.06.2024
Received on: 18.10.2023 / **Revised on:** 08.04.2024 / **For printing:** 10.06.2024

Введение

В работе [1] было рассмотрено влияние структуры взаимодействующих элементов на мотивацию к интероперабельности на основе определенных признаков или характеристик в структуре элементов, которые способствуют или препятствуют достижению интероперабельности [2]. Было отмечено, что потенциальная возможность установления интероперабельности на основе структуры взаимосвязанных элементов [3] может быть определена как структурная интероперабельность [4]. Соответственно, для изучения тенденции информационных систем или элементов к интероперабельности в зависимости от соотношения собственных признаков или характеристик была предложена модель структурного соответствия, позволяющая оценивать группы потенциально близких друг к другу элементов.

Используемая терминология и источники

Задача поиска предпочтительной структуры интероперабельности компонентов системы знаний определяет разбиение множества потенциально взаимодействующих элементов на наборы мотивированных к взаимодействию элементов. Эти элементы могут быть описаны вектором параметров (признаков), из сравнения которых можно сделать вывод о сходстве между элементами и, соответственно, потенциальной мотивации к интероперабельности.

В общем случае представим множество элементов $O = \{o_i\}$ ($i = 1, \dots, n$), где описание каждого элемента представлено в виде вектора из m признаков – атрибутов: $o_i = (p_1^i, \dots, p_m^i)$. Сравнивая любые два элемента этого множества на основе этих признаков, можно оценить их сходство.

Для двух элементов o_i и o_j , сходство которых устанавливается на основе k признаков $\{p_k\}$, функция сходства может выглядеть как [1]:

$$F(o_i, o_j) = 1 - \frac{1}{k} \sum_{i=1}^k w_{ml} \frac{|p_{ml}^i - p_{ml}^j|}{\max |p_{ml}^i - p_{ml}^j|},$$

где $0 \leq w_{ml} \leq 1$ – вес ml -того признака;

$\max |p_{ml}^i - p_{ml}^j|$ – диапазон значений ml -того признака.

Функция F принимает значения из $[0, 1]$, так что $F = 1$ – это абсолютное сходство элементов o_i и o_j , а $F = 0$ – абсолютное различие.

Далее предлагается перейти к знаковому графу, позволяющему на основе анализа согласованности его связей по выбранному критерию согласованности сгруппировать n элементов в P групп. Переход к знаковому графу осуществляется в результате выбора порогового значения α для функции F , так что когда $0 \leq F(o_i, o_j) \leq \alpha$ элементы o_i и o_j считаются несходными по k признакам, а в случае $\alpha < F(o_i, o_j) \leq 1$ – сходными. Сопоставляя знак «минус» связям с $0 \leq F(o_i, o_j) \leq \alpha$ и знак «плюс» – иначе, получим знаковую структуру, являющуюся дискретной знаковой моделью множества потенциально взаимодействующих элементов.

Знаковому графу может быть поставлена в соответствие матрица связности. Задача согласованности структурной интероперабельности решается на основе учета тернарных отношений, которые могут быть отнесены к согласованному или рассогласованному состоянию. В [1] эти состояния названы консонансными и диссонансными, соответственно. Если представить знаковый граф в виде совокупности тернарных отношений, то получится система анализа структурной согласованности потенциально взаимодействующих элементов. В качестве критерия согласованности предлагается выбрать тернарный критерий Хайдера [5], относящий тернарные отношения (треугольники) к консонансным, если положительная связь между любыми парами вершин задает тождественность связей этих вершин с третьей, в противном случае – к диссонансным.

При разбиении знакового графа на треугольники и исследовании согласованности множества по критерию Хайдера выясняется, что консонансное множество M_K (состоящее только из консонансных треугольников) представимо в виде двух подмножеств M_1 и M_2 : $M_K = M_1 \cup M_2$, так что любые два элемента o_i и o_j из одного подмножества связаны положительной связью, а принадлежащие разным подмножествам – отрицательной. Если M_1 или M_2 пустое, то такое множество называется тривиальным – оно содержит только положительные связи. Ассонансным множеством назовем множество, содержащее как консонансные, так и диссонансные треугольники.

Приведение ассонансного множества в консонанс по Хайдеру означает возможность представления его в виде двух классов эквивалентности элементов. Для поиска согласованного множества, состоящего более чем из двух подмножеств, необходимо расширить критерий Хайдера

до понятия поликонсонанса и консонанса степени P [1]. Поликонсонанс степени P соответствует согласованному состоянию множества, состоящего не более, чем из P подмножеств, так что элементы внутри каждого подмножества связаны только положительными связями, а из разных подмножеств – только отрицательными. Под консонансом степени P будем понимать поликонсонанс степени P , состоящий в точности из P классов.

Описание алгоритма

Алгоритм поиска структурного соответствия, позволяющий оценивать связность групп потенциально близких друг к другу элементов, заключается в оценке состояния ассонансного множества, моделирующего потенциальную возможность элементов к интероперабельности. Полученное в результате моделирования ассонансное множество сравнивается с некоторым консонансным множеством – прообразом и преобразуется, используя операцию повершинного переброса (изменение всех знаков элемента как вершины знакового графа на противоположные) так, чтобы найти ближайший по количеству различаемых связей консонансный прообраз.

Определим контур множества как совокупность множеств, образованную в результате всевозможных повершинных изменений, то есть изменений всех знаков связей у каждого элемента (вершины в представлении этого множества в виде знакового графа). Если ограничиться повершинным изменением каждой вершины не более одного раза, то получим цикл.

Отсюда следует, что экземпляры консонансных (диссонансных) множеств образует контур.

В контуре множества взаимно преобразуются не более чем за $\lfloor n/2 \rfloor$ повершинных изменений, и каждая пара множеств контура единственным образом задает содержащий их цикл.

Предлагаемый алгоритм нахождения ближайшего консонансного прообраза для заданного ассонансного множества заключается в его сопоставлении с неким консонансным прообразом для приближения этого прообраза на основе повершинного изменения связей так, чтобы различие консонансного прообраза от ассонансного множества стало меньше.

Возьмем два множества M_1 и M_2 из одних и тех же n объектов с различными связями между объектами. Для каждой пары объектов o_i и o_j из этих множеств определим число r_{ij} так, что:

а) $r_{ij} = 0$, если $i = j$;

б) $r_{ij} = 0$, если $i \neq j$ и связи между o_i и o_j в этих множествах совпадают;

в) $r_{ij} = 1$, если $i \neq j$ и связи между o_i и o_j в этих множествах различны.

Зададим набор из n чисел $(r_1, r_2, r_3, \dots, r_n)$, так что каждое из $r_i = \sum_{j=1}^n r_{ij}$ ($0 < r_i \leq n - 1$) равно числу различающихся знаков связей у объекта (вершины) o_i в том и другом множествах.

Можно рассматривать вектор $(r_1, r_2, r_3, \dots, r_n) = \vec{r}$ как критерий попершинного различия двух множеств из одних и тех же объектов, назвав его вектором попершинных различий (ВПП).

Сумма r_i равна двум суммам различающихся знаков связей.

Однако ВПП не позволяет однозначно задавать множество парой $(m, \vec{r})$ где m – экземпляр множества, $\vec{r}$ – ВПП m с заданным множеством [1].

Множеству, минимально удаленному по сумме изменяемых знаков связей от некоторого множества m , соответствует минимальная сумма компонентов вектора попершинных различий с m .

Обозначим через $M(\varphi)$ множество позиций, в которых располагаются φ различающихся знаков у двух строк матрицы связности, а через $K(M)$ – количество элементов множества M , то есть $K(M(\varphi)) = \varphi$.

Теорема 1. Пусть даны три строки с одинаковым количеством бинарных элементов: m_1, m_2, m_3 ; и числа несовпадающих значений элементов у каждой пары этих строк: r_{12}, r_{13}, r_{23} . Тогда справедливы следующие соотношения:

$$M(r_{13}) = (M(r_{12}) \cup M(r_{23})) / M(r_{12}) \cap M(r_{23});$$

$$r_{13} = r_{12} + r_{23} - 2K(M(r_{12}) \cap M(r_{23})).$$

Доказательство. Рассмотрим объединение позиций, по которым различаются строка m_1 от m_2 и строка m_2 от m_3 , то есть $M(r_{12}) \cup M(r_{23})$. Среди них есть позиции, в которых либо m_1 отличается от m_2 но m_2 не отличается от m_3 , либо m_1 не отличается от m_2 , а m_2 отличается от m_3 . Очевидно, что в этих позициях m_1 отличается от m_3 . Оставшиеся позиции являются позициями отличия m_1 от m_2 и m_2 от m_3 , и вследствие бинарности значений элементов строк значения элементов строк m_1 и m_3 в этих позициях не совпадают. Значит, если их исключить из числа позиций $M(r_{12}) \cup M(r_{23})$, то оставшееся множество и будет $M(r_{13})$, так как других различий нет. Множество исключаемых позиций состоит из $M(r_{12}) \cap M(r_{23})$. Учитывая вышесказанное, нетрудно заметить, что $K(M(r_{12}) \cap M(r_{23})) = r_{12} + r_{23} - K(M(r_{12}) \cup M(r_{23}))$. Отсюда $K(M(r_{13})) = r_{13} = K(M(r_{12}) \cup M(r_{23})) - K(M(r_{12}) \cap M(r_{23})) = r_{12} + r_{23} - 2K(M(r_{12}) \cap M(r_{23}))$.

Отметим, что теорема верна для любых трех последовательностей одинаковой длины с фиксированным положением бинарных элементов.

Лемма. Пусть проведена последовательность попершинных изменений из k вершин. Если начальные значения попершинных различий у этих вершин с теми же вершинами другого, множества были $\{r_i\}$, $i = 1 \dots k$; то после изменений они станут равными $(n - k) - r_i + 2\beta_i$, где β_i – количество позиций в i -й строке, в которых находятся знаки взаимных связей k вершин, одновременно являющиеся знаками различения, учитываемыми в r_i .

Доказательство. Поскольку у i -й вершины последовательности имеется $k - 1$ знаков связей с другими вершинами этой последовательности, то среди $k - 1$ знаков α_i знаков не являются знаками различения и не входят в число r_i , а остальные β_i – являются. Тогда в

результате попершинного изменения i -й вершины (то есть одновременного изменения всех знаков ее связей на противоположные) ее попершинное различие станет равным $(n - 1) - r_i$, а α_i теперь будет количеством знаков различения, β_i – наоборот. Последующие $k - 1$ попершинных изменений остальных вершин последовательности приведут к тому, что $(n - 1) - r_i$ уменьшится на α_i и увеличится на β_i , став равным $(n - 1) - r_i - \alpha_i + \beta_i$. Но $\alpha_i + \beta_i = k - 1$, отсюда получим $(n - k) - r_i + 2\beta_i$.

Теорема 2. Пусть множества m_1 и m_2 состоят из одних и тех же объектов. Рассмотрим в цикле множества m_1 состояние m_1^k , полученное k попершинными изменениями, и в цикле множества m_2 состояние m_2^k , полученное k попершинными изменениями тех же k вершин. Тогда вектора попершинных различий множества m_1 от m_2^k и множества m_2 от m_1^k совпадают.

Доказательство. Согласно лемме, k строк m_1 и m_2^k будут различаться на $(n - k) - r_i + 2\beta_i^*$ и $n - k$ строк m_1 и m_2^k будут различаться на $k - r_i + 2\beta_i^*$, где r_i – компоненты вектора попершинных различий m_1 от m_2 . Аналогично для k строк m_2 и m_1^k : $(n - k) - r_i + 2\beta_i^{**}$ и $n - k$ строк m_2 и m_1^k : $k - r_i + 2\beta_i^{**}$. Но β_i^* и β_i^{**} входят в число знаков между k вершинами или $n - k$ вершинами, которые остались неизменными, значит, все $\beta_i^* = \beta_i^{**}$.

Отсюда следует, что рассматривая в качестве m_1 произвольное ассонансное множество, а в качестве m_2 – консонансное, можно искать минимально удаленный консонансный прообраз, находясь в цикле ассонансного множества m_1 , поставив затем в соответствие m_1 множество m_2^k , полученное той же последовательностью попершинных изменений, что и найденное нами минимально удаленное от m_2 состояние m_1^k .

Казалось бы, что достаточно брать произвольное консонансное множество, сравнивать с ним исходное ассонансное и производить последовательные попершинные изменения пока все попершинные различия не станут меньше или равными половине связей, получив в минимально удаленное состояние. Но этого оказывается недостаточно. Пример ассонансного множества на рис. 1 показывает, что различие этого множества с тривиальным консонансным множеством (0:9) характеризует вектор r_1 , у которого только первый компонент равен половине связей вершины, а остальные – меньше. Но (0:9) – не минимальный консонансный прообраз, так как вектор r_2 , различающий указанное консонансное множество и консонантное множество (5:4), имеет все

0:9					5:4					r_1		r_2	
01	+	+	+	+	-	-	-	-	-	4	0	4	0
+	02	+	+	+	-	+	-	+	+	2	2	2	2
+	+	03	+	+	+	-	-	+	-	2	2	2	2
+	+	+	04	+	-	+	-	+	+	2	2	2	2
+	+	+	+	05	+	-	+	-	-	2	2	2	2
-	-	+	-	+	06	+	+	+	+	3	2	3	2
-	+	-	+	-	+	07	+	+	+	3	2	3	2
-	-	+	-	+	+	+	08	+	+	3	2	3	2
-	+	-	+	-	+	+	+	+	09	3	2	3	2

Рис. 1. Матрица ассонансного множества и ВПП с консонансами 0:9 и 4:5

компоненты меньше половины связей и действительно определяет минимально удаленное консонансное множество (5:4).

Теорема 3. Перемещаясь по циклу множества можно обнаружить несколько векторов повершинных различий, у каждого из которых все компоненты меньше половины количества связей вершины. При этом соответствующие состояния отстоят друг от друга не менее чем на два повершинных изменения для четного n и не менее чем на три – для нечетного n .

Доказательство. Пусть мы находимся в состоянии с вектором $\vec{r}$: $r_i < \gamma$, где γ – половина количества связей вершины. Постараемся достичь еще одно состояние за k повершинных изменений, характеризуемое вектором $\vec{r}^k$: $r_i^k < \gamma$, $r_j^k < \gamma$, предположив, что оно существует. Для каждой из k вершин, согласно Лемме, компонента вектора $r_j^k = (n - k) - r_j + 2\beta_j$. Но у вектора $\vec{r}^k$ все компоненты также должны быть меньше γ , значит $(n - k) - r_j + 2\beta_j < \gamma$ и $k > n - \gamma - r_j + 2\beta_j$. Ясно, что k будет минимальным, когда $\beta_j = 0$ (следствие 1 из Леммы), а r_j принимает максимальное значение, равное $\gamma - 1$. Отсюда $k > n - 2\gamma + 1$.

Для четного n получаем $k > 1$, а для нечетного – $k > 2$.

Пример множества из четного числа объектов, равноудаленного от трех консонансных множеств, отстоящих друг от друга на два повершинных изменения, приведен на рис. 2. При этом все $r_i < \gamma$.

0:8				2:6				4:4			
o ₁	+	+	+	-	-	-	+	3	3	1	
+	o ₂	+	+	+	-	-	-	3	3	1	
+	+	o ₃	+	-	+	+	+	1	3	3	
+	+	+	o ₄	-	+	+	+	1	3	3	
-	+	-	-	o ₅	+	+	+	3	3	1	
-	-	+	+	+	o ₆	+	+	2	0	2	
-	-	+	+	+	+	o ₇	+	2	0	2	
+	-	+	+	+	+	+	o ₈	1	1	3	

Рис. 2. Матрица ассонансного множества и ВПР с консонансами 0:8, 2:6 и 4:5

Другой пример на рис. 3 показывает существование двух векторов со значениями компонентов меньше γ , но задающих разное удаление приведенного на рис. 3 ассонансного множества от консонансных множеств (0:9) и (4:5). Таким образом, не только при равенстве повершинных различий числу γ нельзя считать, что это минимальное удаление, но и при достижении состояния,

0:9				4:5				0:9				4:5			
o ₁	+	+	+	+	-	+	-	3		3		2		2	
+	o ₂	+	+	-	+	-	-	3		3		2		2	
+	+	o ₃	+	+	-	+	-	3		3		2		2	
+	+	+	o ₄	-	+	-	+	3		3		2		2	
+	-	+	-	o ₅	-	+	+	3		3		3		3	
-	+	-	+	-	o ₆	+	+	3		3		3		3	
+	-	+	-	+	+	o ₇	+	2		2		2		2	
-	-	-	+	+	+	+	o ₈	3		3		1		1	
-	+	-	-	+	+	+	+	3		3		1		1	

Рис. 3. Два ВПР с неулучшаемыми компонентами, но с разным удалением по связям

где все $r_i < \gamma$, также нельзя быть уверенным, что именно оно и есть минимально удаленное.

Из вышесказанного следует, что алгоритм, основанный на уменьшении повершинных различий, должен обеспечивать нахождение всех потенциально минимальных состояний, а не останавливаться на первом найденном. В принципе, для множества m может не существовать вектора $\vec{r}$: $r_i < \gamma$, а количество минимальных состояний будет равно количеству типов консонансного множества, так что поиск состояния с вектором $\vec{r}$: $r_i < \gamma$ в этом случае ничего не даст.

Основная идея уменьшения рассогласованности множества за счет вершинных перебросов заключается в следующем. Возьмем некоторое консонансное множество и «испорчим» часть связей, переводя его в ассонанс. Представим теперь, что мы забыли этот ближайший консонансный прообраз полученного ассонанса, и будем искать его среди возможных консонансов в рамках выбранной степени. Если взять произвольное консонансное множество и сравнить с ним имеющееся ассонансное множество, то ВПР в общем случае будет содержать нулевые компоненты. Каждая компонента ВПР будет представлять собой число, учитывающее различие в связях ближайшего консонансного прообраза, взятого наугад консонансного множества, и различие, возникшее из-за «испорченных» связей. Итак, рассматриваемый ВПР учитывает два вида различий. Однако ближайший консонансный прообраз и взятое наугад консонансное множество принадлежат одному контуру как переводимые друг в друга некоторой последовательностью вершинных, перебросов, и тогда останется только минимальный набор диссонансных связей. Но простым перебором это сделать очень трудно, т.к. надо перебирать экземпляры консонансных множеств, а их даже для поликонсонанса степени 2 порядка 2^{n-1} .

Алгоритм уменьшения рассогласованности был предложен для поликонсонанса степени 2. Обобщим его применение для степени n .

Пусть имеется некоторое ассонансное множество. Возьмем произвольное консонансное множество $A_1 A_2 A_3 \dots A_N (n_1; n_2; n_3 \dots n_N)$ степени N , составленное из тех же объектов. Тогда некоторому объекту $o_i \in A_1 A_2 A_3 \dots A_N$ соответствует r_i – ВПР этих множеств. Рассмотрим знаки связей объекта o_i , например, в A_1 и A_2 , составленные из n_1 и n_2 объектов, соответственно. Если окажется, что у o_i среди связей с объектами подмножеств A_1 и A_2 более чем $1/2 (n_1 + n_2 - 1)$ отличаются от аналогичных связей того же o_i в ассонансном множестве, то o_i целесообразно перебросить из A_2 в A_1 , т.к. это уменьшит r_i , а следовательно, и общую сумму $\sum r_i$, равную, как уже отмечалось, удвоенной сумме диссонансных связей. Действуя аналогичным образом с другими объектами, мы придем к ситуации, когда ни один из объектов уже не имеет смысла перебрасывать в другое подмножество. Тогда, казалось бы, уничтожен один вид различия – межконсонансный, и остались только диссонансные связи в полученном минимально удаленном консонансном прообразе. Однако

это не всегда верно, т.е. из того, что каждый единичный вершинный переброс не выгоден, еще не следует, что нецелесообразно перебросить группу объектов.

Рассмотрим последовательность повершинных перебросов из k объектов. Любой r_i может быть записан как $r_i = \beta_i + \psi_i + v_i$, где β_i – количество различающихся знаков связей между k объектами переброса, ψ_i – количество различающихся знаков с остальными $n_1 + n_2 - k$ объектами обоих подмножеств и v_i – различающиеся знаки с $n - (n_1 + n_2)$ объектами, не принадлежащими рассматриваемым подмножествам. Если начальные значения повершинных различий у этих объектов с теми же объектами другого множества были r_i , то после произведенных перебросов они станут равными $n_1 + n_2 - k - r_i + 2\beta_i + 2v_i$. Доказательство этого легко проводится по аналогии с Леммой. Пусть имеется все то же консонансное множество, в котором нет ни одного объекта $o_i \in A_2$, такого, что соответствующие связи с n_1 и n_2 объектами A_1 и A_2 имеют более чем $1/2 (n_1 + n_2 - 1)$ различий со связями тех же объектов ассонансного множества. Найдем условие, при котором целесообразно перебросить целую группу объектов из A_2 в A_1 . Обозначим через γ число $1/2 (n_1 + n_2 - 1)$. Различающиеся знаки v_i можно опустить из рассмотрения, т.к. они не имеют отношения к проводимым перебросам. Итак, до проводимых перебросов каждый объект из группы перебрасываемых объектов имел $\beta_i + \psi_i$ различающихся знаков со всеми остальными $n = n_1 + n_2$ объектами обоих подмножеств. После проведенных перебросов число различий стало $(n_1 + n_2 - k) - \psi_i + \beta_i$. По начальному условию ни один объект из группы перебрасывать нецелесообразно, т.е. все $\beta_i + \psi_i \leq \gamma$, аналогично после перебросов мы не должны выйти из локального минимально удаленного состояния, т.е. все $(n_1 + n_2 - k) - \psi_i + \beta_i \leq \gamma$. Отсюда возникают ограничения на β_i : $\beta_i \leq \frac{k-1}{2}$ и на ψ_i : $\beta_i + \gamma - (k-1) \leq \psi_i \leq \gamma - \beta_i$. Условие улучшения минимальной удаленности по знакам

$$\sum_{i=1}^k \beta_i + \sum_{i=1}^k \psi_i > \sum_{i=1}^k (n-k) - \sum_{i=1}^k \psi_i + \sum_{i=1}^k \beta_i \text{ дает}$$

$$\sum_{i=1}^k \psi_i > 1/2(n-k) \cdot k.$$

$$\text{Кроме того, } \sum_{i=1}^k \psi_i + \sum_{i=1}^k \beta_i \leq \gamma \cdot k, \text{ значит}$$

$$\sum_{i=1}^k \beta_i < \frac{k(k-1)}{2}.$$

Предположим, что последовательность из $k-1$ перебросов еще не обеспечивает улучшение локального минимума, т.е. $\sum_{i=1}^{k-1} \psi_i \leq \frac{(n+k+1)(k-1)}{2}$. Тогда переброс группы из k объектов будет целесообразен, если существует o_k : $\sum_{i=1}^{k-1} \psi_i + \psi_k > \frac{(n-k) \cdot k}{2}$, т.е. $\psi_k > \gamma - (k-1)$.

Интерпретация условия $\sum_{i=1}^k \psi_i > \frac{(n-k) \cdot k}{2}$ почти оче-

видна: если среди связей некоторой группы из k объектов с остальными $n-k$ объектами более половины отличается от связей тестируемого ассонансного множества, значит, эта группа имеет тенденцию к перебросу в другое подмножество, что тем самым поставит в соответствие ассонансному множеству менее удаленное по сумме различающихся знаков связей консонансное множество [6]. Поиск такой группы объектов даже для поликонсонанса степени 2 является очень сложной задачей, т.к. заранее невозможно определить даже потенциальную возможность вхождения каждого объекта в конструируемую для предполагаемого переброса группу.

Альтернативный подход

Другим подходом к поиску близкого консонансного прообраза, не требующим решения тяжелых комбинаторных задач, является алгоритм распространения меток [7].

В основе алгоритма лежит идея о том, что у вершины графа сообщество будет таким же, как и у большинства ее соседей.

Сам алгоритм работает следующим образом:

Изначально каждой вершине графа присваивается своя уникальная метка.

В цикле, пока решение не сойдется, или не будет достигнуто максимальное количество итераций, у каждой вершины обновляется метка на ту, которая чаще встречается у ее соседей. Если среди соседей самых часто встречаемых меток несколько, выбирается случайная.

Для перехода к формальному описанию алгоритма, нужно ввести ряд понятий:

Матрица связности $\mathbf{A}$ – симметричная квадратная матрица, показывающая наличие или отсутствие связей между объектами.

Диагональная матрица степеней $\mathbf{D}$ – диагональные элементы равны степеням вершин графа, а степень вершины графа – число связей этой вершины с другими вершинами:

$$\mathbf{D} = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}.$$

Матрица перехода $P = \mathbf{D}^{-1} \mathbf{A}$, где

$$\mathbf{D}^{-1} = \begin{pmatrix} \frac{1}{d_1} & 0 & \dots & 0 \\ 0 & \frac{1}{d_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{d_n} \end{pmatrix}.$$

Матрица «цветов» Y имеет размерность $N \times M$, где N – число объектов (вершин графа), а M – число уникальных цветов на нулевой итерации алгоритма (максимальное число кластеров). В нашем случае удобно считать, что $N = M$, то есть изначально каждая вершина графа имеет свой уникальный цвет. Тогда без потери общности

$$Y(0) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \quad Y(t+1) = PY(t).$$

При этом новые цвета вершин определяются как $\arg \max_{color} Y$.

На рис. 4 приведен пример работы данного алгоритма. Сначала каждому узлу присваивается своя уникальная метка, и далее запускается их распространение по графу. По ходу распространения меток соседи внутри одного сообщества достаточно быстро приходят к общему значению метки. Далее такие сообщества начинают расширяться,

пока это возможно. На примере видно, что вершины 3 и 5 хоть и соединены ребром, имеют разные цвета. Так получилось, потому что основная часть соседей каждой из двух этих вершин (вершины 2 и 4 – для вершины 3, и вершины 6, 7 и 8 – для вершины 5) имеют общий цвет метки.

Является ли решение, найденное данным алгоритмом, оптимальным с точки зрения достижения минимального по сумме компонент ВПР, неизвестно. Однако алгоритм имеет эффективную реализацию для распределенных вычислительных систем, что делает его удобным для использования на практике. Говоря про оптимальность, стоит отметить, что алгоритм можно как минимум использовать как быстрый метод для поиска первоначального консонансного прообраза для алгоритма из предыдущего раздела.

Заключение

Из приведенного описания предлагаемого алгоритма и его альтернативы следует, что при сформулированных условиях существования структурной интероперабельности существует несколько вариантов

выбора приемлемой структуры интероперабельности. Однако поиск оптимального варианта потребует перебора среди всех приемлемых вариантов, либо использование разумных эвристик, учитывающих специфику матрицы связности исходного знакового графа. Сравнивая представленные алгоритмы, необходимо заметить, что алгоритм распространения меток носит оценочный характер, при этом, несмотря на сложность комбинаторных оценок и преобразований, алгоритм уменьшения рассогласованности на основе вектора попершинных различий представляет собой инструмент перманентного анализа и управления согласованностью наборов мотивированных к интероперабельности элементов.

Библиографический список

1. Розенберг И.Н., Дулин С.К., Дулина Н.Г. Моделирование структуры интероперабельности средствами структурной согласованности // Информатика и ее применения. 2023. Т. 17. Вып. 1. С. 57-65. DOI 10.14357/19922264230108

Algorithm: Label propagation, Zhu et al. 2002

Input: Graph G , labels Y_k

Output: labels $\hat{Y}$

Compute adjacency matrix A

Compute $d_i = \sum_j A_{ij}$

Compute $D = \text{diag}(d)$

Compute $P = D^{-1}A$

Initialize $Y(0) = (Y_k, 0), t = 0$

repeat

$Y(t+1) = PY(t)$

$Y_k(t+1) = Y_k(t)$

$t = t + 1$

until $Y(t)$ converges;

$\hat{Y} = Y(t)$

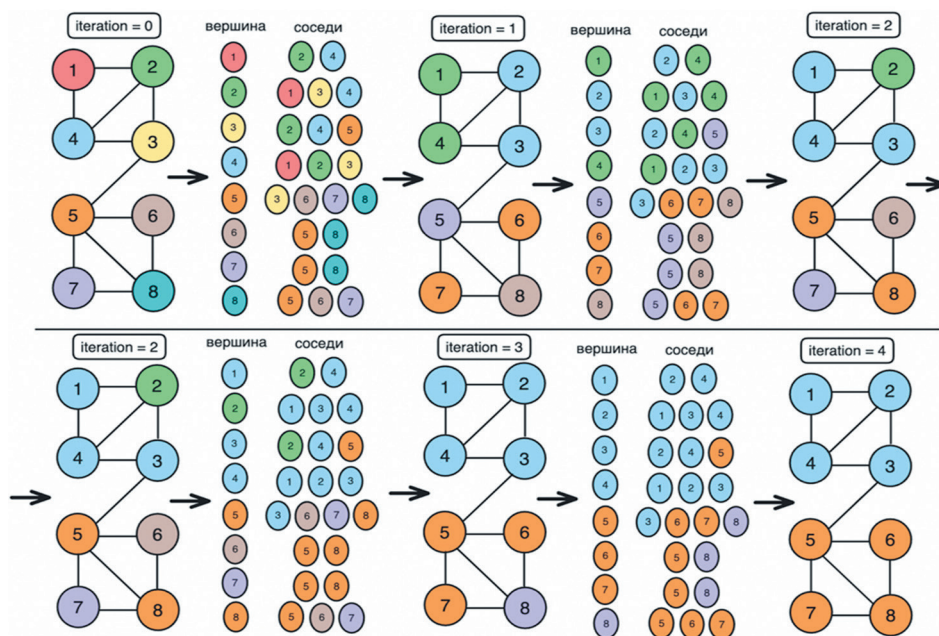


Рис. 4. Схема работы алгоритма распространения меток. Цвет вершины соответствует метке / кластеру

2. ГОСТ Р 55062-2012 Информационные технологии (ИТ). Системы промышленной автоматизации и их интеграция. Интероперабельность. Основные положения. М.: Стандартинформ, 2014. IV, 8 с.

3. Ullberg Johan, Johnson Pontus, Buschle Markus. A Modeling Language for Interoperability Assessments // *Lecture Notes in Business Information Processing*. 2011. V. 76. Pp. 61-74. DOI: 10.1007/978-3-642-19680-5_7

4. Макаренко С.И., Соловьева О.С. Основные положения концепции семантической интероперабельности сетевых систем // *Журнал радиоэлектроники*. 2021. № 4. 24 с. DOI:10.30898/1684-1719.2021.4.10

5. Heider F. The psychology of interpersonal relation. N.Y., 1982. 336 p. DOI: 10.30898/1684-1719.2021.6.3

6. Дулин С.К., Дулина Н.Г., Кожунова О.С. Синтез геоанализа в пространственных инфраструктурах на основе связанных данных // *Информатика и ее применения*. 2019. Т. 13. № 1. С. 82-90. DOI: 10.14357/19922264190112

7. Xiaojin Zhu, Ghahramani Zoubin. Learning from labeled and unlabeled data with label propagation // *Tech. Rep., Technical Report CMU-CALD-02-107*, Carnegie Mellon University. 2002. URL: <https://mlg.eng.cam.ac.uk/zoubin/papers/CMU-CALD-02-107.pdf> (дата обращения: 01.04.2024)

References

1. Rozenberg I.N., Dulin S.K., Dulina N.G. Modeling the structure of interoperability by means of structural consistency. *Informatics and Applications* 2023;17(1):57-65. DOI 10.14357/19922264230108. (in Russ.)

2. GOST R 55062-2012. Information technologies. Industrial automation systems and integration. Interoperability. Basic principles. Moscow: Standartinform; 2014. (in Russ.)

3. Ullberg J., Johnson P., Buschle M. A Modeling language for interoperability assessments. *Lecture Notes in Business Information Processing* 2011;76:61-74. DOI: 10.1007/978-3-642-19680-5_7.

4. Makarenko S.I., Solovieva O.S. [Key provisions of the concept of semantic interoperability of network-centric systems]. *Journal of radio electronics* 2021;4. DOI:10.30898/1684-1719.2021.4.10. (in Russ.)

5. Heider F. The psychology of interpersonal relation. New York; 1982. DOI: 10.30898/1684-1719.2021.6.3.

6. Dulin S.K., Dulina N.G., Kozhunova O.S. Synthesis of geodata in spatial infrastructures based on related data. *Informatics and Applications* 2019;13(1):82-90. DOI: 10.14357/19922264190112. (in Russ.)

7. Xiaojin Z., Ghahramani Z. Learning from labeled and unlabeled data with label propagation. Technical Report CMU-CALD-02-107, Carnegie Mellon University; 2002. (accessed 01.04.2024). Available at: <https://mlg.eng.cam.ac.uk/zoubin/papers/CMU-CALD-02-107.pdf>.

Сведения об авторах

Дулин Сергей Константинович (Dulin Sergey Konstantinovich) – 109029, Россия, Москва, Нижегородская ул., д.27 стр.1, АО «Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте» (АО «НИИАС»), главный научный сотрудник, доктор технических наук, профессор; 119333, Россия, Москва, ул. Вавилова, д.44 к.2, Федеральный исследовательский центр «Информатика и управление» Российской академии наук, ведущий научный сотрудник, e-mail: skdulin@mail.ru; *Надежность* № 3, 2018; *Системы и средства информатики* Т. 28, № 2. 2018; Т. 29 № 3. 2019; Т. 30, № 2. 2020; Т. 31, № 3. 2021; Т. 32, № 1. 2022; Т. 33, № 2, 2023; *Железнодорожный транспорт* № 8, 2018; *Информатика и ее применения* Т. 13. № 1. 2019; Т. 14, Вып. 1, 2020; Т. 17, Вып. 1, 2023; *Автоматика, связь, информатика*. №11, 2021.

Рябцев Антон Борисович (Ryabtsev Anton Borisovich) – 117303, Россия, Москва, Керченская ул., д.1а к.1, Московский физико-технический институт (Государственный университет), аспирант, ryabtsev.ab@phystech.edu; *Образовательные ресурсы и технологии* № 3(44), 2023.

About the authors

Sergey K. Dulin, Doctor of Engineering, Professor, Chief Researcher, JSC Research and Design Institute for Information Technology, Signalling and Telecommunications in Railway Transportation (JSC NIIS), 27, bldg 1 Nizhegorodskaya St., Moscow, 109029, Russia; Lead Researcher, Federal Research Center Computer Science and Control of the Russian Academy of Sciences, 44, korp. 2 Vavilova St., Moscow, 119333, Russia, e-mail: skdulin@mail.ru; *Dependability* 2018;3; *Systems and Means of Informatics* 2018;28(2); 2019;29(3); 2020;30(2); 2021;31(3); 2022;32(1); 2023;33(2); *Zheleznodorozhny transport* 2018;8; *Informatics and Applications* 2019;13(1); 2020;14(1); 2023;17(1); *Automation, Communications, Informatics* 2021;11.

Anton B. Ryabtsev, post-graduate student, Moscow Institute of Physics and Technology (Public university), 1a, korp. 1 Kerchenskaya St., Moscow, 117303, Russia, ryabtsev.ab@phystech.edu; *Educational Resources and Technologies* 2023;3(44).

Вклад авторов в статью

Дулин С.К. разработал алгоритм улучшения согласованности структурной интероперабельности. Рябцев А.Б. исследовал и развил алгоритм распространения меток, сравнил его возможности с разработанным Дулиным С.К. алгоритмом улучшения согласованности структурной интероперабельности.

Конфликт интересов

Автор заявляет об отсутствии конфликта интересов.