

К. А. Осипенко, Е. А. Британов, Д. А. Волков,
А. Н. Осипенко, Н. Б. Осипенко

АЛГОРИТМ МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ РАЗБИЕНИЯ ГРУППЫ ЛЮДЕЙ НА ПОДГРУППЫ С РАЗНЫМИ СПЕЦИАЛИЗАЦИЯМИ

Работа посвящена описанию алгоритма многокритериальной оптимизации разбиения группы людей на подгруппы с разными специализациями. Для решения этой задачи был создан сайт, предоставляющий услугу, позволяющую повысить психологическую совместимость людей при выполнении ими общего дела. В основании используемых психологических тестов и оптимизации по совместимости лежат оригинальные авторские разработки. Для решения задачи был разработан эмпирический и интуитивно понятный алгоритм разбиения общей группы на подгруппы по критерию минимума расстояния субъектов до целевой ячейки деятельности.

Проблема эффективного разбиения некоторой совокупности людей на подгруппы с разными специализациями имеет немаловажное значение при работе с кадрами в любом коллективе. Для решения этой задачи был создан сайт, ориентированный на повышение психологической совместимости людей при выполнении ими общего дела: выстраивания семейной жизни, работы, учебы, досуга, спортивной игры, общения, оздоровления и т. д. Каждый человек обладает своим уникальным характером, темпераментом и способностями. У одного – какое-то качество (например, воля) сильно развито, у другого – его почти нет: у одного – позитив, а у другого – негатив (безволие) или псевдопозитив (пустая страсть). Причем, по отношению к разным видам деятельности и к разным людям у человека это качество может проявляться всеми тремя способами. Поэтому классическая диагностика: интроверт – экстраверт, холерик – флегматик или рождение в знаке Стрельца дает очень общее и статичное представление о человеке. Сайт помогает подобрать «нужный сплав» психологических элементов, развернуть их друг к другу позитивными сторонами, обеспечить устойчивость и эффективность сотрудничества, или же при формировании коллектива посоветует процедуру поиска новых людей с недостающими качествами.

В основании предлагаемых сайтом психологических тестов и последующей оптимизации по совместимости лежат оригинальные авторские разработки, в частности, таблица тем деятельности (и соответствующих качеств человека), аналогичная таблице химических элементов Д.И. Менделеева [1]. Для решения задачи был разработан эмпирический и интуитивно понятный алгоритм разбиения общей группы на подгруппы по критерию минимума расстояния субъектов до целевой ячейки деятельности. На начальном этапе работы алгоритма используется информация о паспортных данных человека [2].

Основные понятия и виды реализуемых расстояний. Пусть N – количество людей общей группы Ω , которые разбиваются на L команд: $\Omega_1, \Omega_2, \dots, \Omega_L$.

$$N = N_1 + N_2 + \dots + N_L,$$

где N_l – количество людей, которые должны остаться в l -ой команде ($l = \overline{1, L}$).

Будем считать, что $N_l \geq N/L - \Delta N_l$, где ΔN_l по умолчанию равно 1, иначе запрашивается у пользователя ($N/L > \Delta N_l \geq 1$).

Рассмотрим произвольную пару субъектов из общей группы Ω людей: S_f и S_r , ($1 \leq f \leq N, 1 \leq r \leq N$): $S_f \in \Omega$ и $S_r \in \Omega$. Пусть *корневая* ячейка для субъекта S_f : $R(S_f) = Z_{ij}$

и для субъекта S_r : $R(S_r) = Z_{pq}$ (корневая ячейка для субъекта определяется с помощью отдельной программы с использованием даты рождения, имени, а также результатов тестирования темперамента).

Целевая ячейка $Z_{i'j'}$ для l -ой группы задается изначально согласно смыслу совместной деятельности в l -ой группе ($l = \overline{1, L}$).

Итоговое расстояние между ячейками $\rho(R(S_n), Z_{i'j'})$ определяется как взвешенная сумма четырех расстояний:

$$\rho(R(S_n), Z_{i'j'}) = 0,125 \cdot \rho_T(R(S_n), Z_{i'j'}) + 0,125 \cdot \rho_H(R(S_n), Z_{i'j'}) + 0,25 \cdot \rho_{Re}(S_n, l) + 0,5 \cdot \rho_{Ps}(R(S_n), Z_{i'j'}) \quad (1)$$

Расстояния, используемые в (1), определены ниже.

1) Расстояние по разности темпераментов: $\rho_T(Z_{ij}, Z_{pq})$ – количество инверсий (максимум 5) в структуре темперамента при переходе от ячейки Z_{ij} к Z_{pq} : $T_{ij} = (t_{ij}^1, t_{ij}^2, t_{ij}^3, t_{ij}^4, t_{ij}^5)$, где $t_{ij}^m = 0$, если соответствующая составляющая темперамента преимущественно интенсивна, иначе $t_{ij}^m = 1$.

$$\rho_T(Z_{ij}, Z_{pq}) = \begin{cases} 1 - 0,2 \cdot [(t_{ij}^1 \wedge t_{pq}^1) + (t_{ij}^2 \wedge t_{pq}^2) + (t_{ij}^3 \wedge t_{pq}^3) + (t_{ij}^4 \wedge t_{pq}^4) + (t_{ij}^5 \wedge t_{pq}^5)] & \text{если } (j < 7 \vee (j = 7 \wedge i = (3 \vee 4))) \\ 0,1 \cdot \sum_{k=1}^5 |t_{ij}^k - t_{pq}^k| & \text{иначе для остальных случаев когда встречаются ячейки } Z_{17}, Z_{27}, Z_{57} \end{cases}$$

Здесь матрица $T_{ij}, i = \overline{1, 5}, j = \overline{1, 7}$ задается априори, исходя из теоретических соображений.

2) Расстояние по различию направлений деятельности сравниваемых ячеек: $\rho_H(Z_{ij}, Z_{pq}) = 0,0625 \cdot |H_{ij} - H_{pq}|$, где H_{ij} и H_{pq} – номера направлений описания деятельности человека $1 \leq H_{ij} \leq 18$ и $1 \leq H_{pq} \leq 18$ для ячеек Z_{ij} и Z_{pq} . Матрица $H_{ij}, i = \overline{1, 5}, j = \overline{1, 7}$ задается также априори, исходя из теоретических соображений.

3) Рейтинговое расстояние $\rho_{Re}(S_n, l) = \frac{Re[n, l] - 1}{L - 1}$, $n = \overline{1, N}$, $l = \overline{1, L}$, где $Re[n, l]$ – рейтинг (ранг) l -ой деятельности (на первом месте самая приоритетная деятельность). Рейтинги деятельностей задают субъекты путем заполнения таблицы «название деятельности – ее рейтинг (порядковый по приоритету номер)».

4) Расстояние по психологической несовместимости $\rho_{Ps}(R(S_n), Z_{i'j'})$.

Пусть $\alpha^1(K_{i'j'}^{n+}), \alpha^1(K_{i'j'}^{n0}), \alpha^1(K_{i'j'}^{n-})$ – степени выраженности по целевой ячейке $Z_{i'j'}$ псевдокачества, нормы качества и антикачества во мне в отношении от меня к другому возможному члену коллектива; $\alpha^2(K_{i'j'}^{n+}), \alpha^2(K_{i'j'}^{n0}), \alpha^2(K_{i'j'}^{n-})$ – степени выраженности этих трех составляющих качества, но уже в моем представлении от другого возможного члена коллектива по отношению ко мне, или другим людям. Пусть $\beta^1(K_{i'j'}^{n+}), \beta^1(K_{i'j'}^{n0}), \beta^1(K_{i'j'}^{n-}), \beta^2(K_{i'j'}^{n+}), \beta^2(K_{i'j'}^{n0}), \beta^2(K_{i'j'}^{n-})$ – соответствующие степени значимости (важности) трех разновидностей качества по ячейке Z_{ij} .

$$\rho'_{PS}(R(S_n), Z_{i'j'}) = 0,2(\alpha^1(K_{i'j'}^+) \beta^1(K_{i'j'}^+) + \alpha^2(K_{i'j'}^+) \beta^2(K_{i'j'}^+)) + \\ + 0,8(\alpha^1(K_{i'j'}^-) \beta^1(K_{i'j'}^-) + \alpha^2(K_{i'j'}^-) \beta^2(K_{i'j'}^-)) - (\alpha^1(K_{i'j'}^0) \beta^1(K_{i'j'}^0) + \alpha^2(K_{i'j'}^0) \beta^2(K_{i'j'}^0))).$$

Все перечисленные степени выраженности и значимости трех разновидностей качеств высчитываются с помощью отдельной программы по результатам специального психологического тестирования субъектов.

$$\rho_{PS}(R(S_n), Z_{i'j'}) = \frac{\rho'_{PS}(R(S_n), Z_{i'j'}) - \min_{i \in \{1, L\}} \rho'_{PS}(R(S_n), Z_{i'j'})}{\max_{i \in \{1, L\}} \rho'_{PS}(R(S_n), Z_{i'j'}) - \min_{i \in \{1, L\}} \rho'_{PS}(R(S_n), Z_{i'j'})}.$$

Начальное разбиение общей группы на подгруппы.

Первый критерий. Распределение субъектов по подгруппам по минимуму расстояния субъектов до целевой ячейки. Для этого организуется цикл по субъектам общей группы $S_n, n = \overline{1, N}: S_n \in \Omega_{l_n}^n$, если $l_n \rightarrow \min \rho(R(S_n), Z_{i'j'})$.

В процессе выполнения этого цикла сохраним полученные расстояния в матрице RO: $RO[n, l] = \rho(R(S_n), Z_{i'j'})$. В итоге получаем вектор распределения субъектов по подгруппам: $LP[n] = l_n, n = \overline{1, N}$. Таким образом, формируется разбиение Ω .

Второй критерий. Проверка наполняемости групп субъектами. Если мощность множества $|\Omega_l| < N_l$, то к этому множеству добавляется субъект n_0^l ($n = n_0^l$ субъект S_n максимально удален от всех остальных групп и минимально близок к $Z_{i'j'}$, если ищем

$$\min(z), \text{ то это равносильно } \max(-z): n_0^l \rightarrow \max_{S_n \in (\Omega \setminus \Omega_l)} \{ \sum_{i \neq l}^L \rho(R(S_n), Z_{i'j'}) - \rho(R(S_n), Z_{i'j'}) \}.$$

Последняя операция осуществляется до тех пор пока $|\Omega_l| \geq N_l, l = \overline{1, L}$.

Корректировка начального разбиения. Корректировка начального разбиения осуществляется с использованием матрицы индивидуальных предпочтений и нерасположенностей к совместной деятельности, а также матрицы психологической несовместимости.

Пусть $Q[n_1, n_2], n_1 = \overline{1, N}, n_2 = \overline{1, N}$ – матрица индивидуальных предпочтений и нерасположенностей:

$$Q[n_1, n_2] = \begin{cases} 1, \text{ если } S_{n_1} \text{ предрасположен к } S_{n_2} \\ 4, \text{ если } S_{n_1} \text{ нерасположен к } S_{n_2} \\ 2, \text{ если } S_{n_1} \text{ не отметил } S_{n_2} \end{cases}$$

Матрица попарных психологических несовместимостей $C'[n_1, n_2], n_1 = \overline{1, N}, n_2 = \overline{1, N}$:

$$C'[n_1, n_2] = \alpha^1(K_{ij}^{n_1+}) \beta^1(K_{ij}^{n_1+}) + \alpha^2(K_{pq}^{n_1-}) \beta^2(K_{pq}^{n_1-}) + \\ + \alpha^1(K_{pq}^{n_2+}) \beta^1(K_{pq}^{n_2+}) + \alpha^2(K_{ij}^{n_1-}) \beta^2(K_{ij}^{n_1-}) - \\ - [\alpha^1(K_{ij}^{n_1^0}) \beta^1(K_{ij}^{n_1^0}) + \alpha^2(K_{ij}^{n_1^0}) \beta^2(K_{ij}^{n_1^0}) + \\ + \alpha^1(K_{pq}^{n_2^0}) \beta^1(K_{pq}^{n_2^0}) + \alpha^2(K_{pq}^{n_1^0}) \beta^2(K_{pq}^{n_1^0})]$$

$$C_{\max} = \max_{n_1, n_2 (n_1 < n_2)} C'[n_1, n_2], C_{\min} = \min_{n_1, n_2 (n_1 < n_2)} C'[n_1, n_2]$$

$$C[n_1, n_2] = \frac{C'[n_1, n_2] - C_{\min}}{C_{\max} - C_{\min}} \times \\ \times (1 - 0,5 \cdot (\rho_T(Z_{ij}, Z_{pq}) + \rho_H(Z_{ij}, Z_{pq})) \times (Q[n_1, n_2] + Q[n_2, n_1])).$$

В последней формуле использованы три критерия: критерий психологической несовместимости, принцип разнообразия внутри подгрупп; удаленность по матрице индивидуальных нерасположенностей.

Преобразуем двумерную треугольную матрицу $C[n_1, n_2]$, $n_1 < n_2$ в одномерный вектор $C^{sv}[m], m = \overline{1, N(N-1)/2}$ с помощью функции свертки индексов: $m = svertka(n_1, n_2) = N(n_1 - 1) + n_2$. При этом $C^{sv}[m] = C[n_1, n_2]$. Отсортируем вектор $C^{sv}[m], m = \overline{1, N(N-1)/2}$ по возрастанию. В результате получим вектор $C^{sort}[m'], m' = \overline{1, N(N-1)/2}$, а также вектор $M^{rank}[m'], m' = \overline{1, N(N-1)/2}$. $M^{rank}[m'] = m$ – номер элемента в векторе $C^{sv}[\cdot]$, который после сортировки в векторе $C^{sort}[\cdot]$ оказался на m' -ом месте $C^{sort}[m'] = C^{sv}[m]$. С помощью вектора $M^{rank}[m'], m' = \overline{1, N(N-1)/2}$ можно выделить наиболее совместимые и наименее совместимые пары субъектов. С помощью функции развертки: $m = M^{rank}[m']$: $rasvertka(m) = (n_1, n_2)$, где $n_1 = [m/N] + 1$, $n_2 = m - N \cdot [m/N]$ из первых элементов вектора $M^{rank}[\cdot]$ берутся близкие друг к другу пары субъектов S_{n_1} и S_{n_2} из последних элементов этого вектора – удаленные друг от друга пары субъектов.

Пусть $\Omega_1, \Omega_2, \dots, \Omega_L$ – подгруппы субъектов, полученные в результате начального разбиения исходной группы.

Общая схема алгоритма корректировки начального разбиения представлена следующими этапами.

Этап 1. Расчет показателей качества объединения субъектов в подгруппы.

Этап 2. Проведение цикла «переброски» субъектов из одной группы в другую с целью оптимизации критериев средней попарной совместимости и разнообразия в подгруппах при относительном сохранении средней удаленности субъектов S_{n_1} от своей целевой ячейки.

Пусть $iter = 1$ (первая итерация по «переброскам»).

2.1. Блок объединения в одну подгруппу двух близких по критерию совместимости субъектов

2.2. Блок разнесения по разным подгруппам двух далеких по критерию совместимости субъектов.

Положим $iter = iter + 1$. Если $N - 2 \cdot iter \leq 2$, то идти на вывод результатов, иначе идти на блок 2.1 (объединение в одну подгруппу двух близких субъектов).

Этап 3. Вывод результатов. Вывод вектора $LP[n], n = \overline{1, N}$ и векторов по подгруппам $Nom^{ln}[n_i], n_i = \overline{1, NL[l_n]}, l_n = \overline{1, L}$.

На базе приведенного выше алгоритма написана на языке C# с использованием фреймворка ASP.NET MVC программная реализация многокритериальной оптимизации разбиения группы людей на подгруппы с разными специализациями, которая в свою очередь является ядром созданного сайта «Совместимость» в среде Microsoft Visual Studio. Задача сайта помочь пользователям выработать высоковероятные альтернативы по

подбору людей, однако окончательный выбор и ответственность всегда остаются за лицом, принимающим решения.

Литература

1 Осипенко Н. Б. Пример «выращивания» регрессионной модели социального явления на базе критерия правдоподобности ее интерпретации / Н. Б. Осипенко, А. Н. Осипенко, К. А. Осипенко // Проблемы физики, математики и техники. – 2013. – № 4(17). – 85–88 с.

2 Осипенко, К. А. Метод регрессионного моделирования продолжительности жизни по дате рождения / К. А. Осипенко, Н. Б. Осипенко // Творчество молодых 2012: сборник научных работ студентов и аспирантов УО «ГГУ им. Ф. Скорины»: в 2 ч. / Гомельский гос. ун-т им. Ф.Скорины; отв. ред. О. М. Демиденко. – Гомель, 2012. – Ч. 1. – С. 194–197.

УДК 519.25

К. А. Осипенко

ОБ ИСПОЛЬЗОВАНИИ АПРИОРНЫХ ЗНАНИЙ ПРИ РАСПОЗНАВАНИИ ГРУПП РИСКА СМЕРТНОСТИ НА ОСНОВЕ ПАСПОРТНЫХ ДАННЫХ

В статье развивается предложенная в предыдущих работах идея массовой экспресс-диагностики групп риска смертности жителей в рамках трехэтапного скрининга здоровья населения Беларуси. В качестве доказательства работоспособности этой идеи было показано, что используя только паспортные данные уже для значительной части людей можно распознавать обобщенную группу риска. С практической точки зрения знание своей основной группы риска поможет человеку выстроить оптимальную стратегию здорового образа жизни.

Введение. В статье [1] авторами был описан вариант трехэтапного скрининга здоровья населения Беларуси, где на первом этапе предлагается осуществить превентивную экспресс-диагностику групп риска смертности обследуемого человека с помощью стандартной информации из поликлинической базы данных. В качестве доказательства работоспособности этой идеи было показано, что используя только паспортные данные (дата рождения, фамилия, имя, отчество, семейное положение, количество детей), уже для значительной части людей можно распознавать обобщенную группу риска. В первую группу входят люди с высокой вероятностью смерти с быстрой потерей жизнеспособности организма (инфаркты, инсульты, несчастные случаи), во вторую – с медленной потерей жизнеспособности (онкологическая или иная хроническая болезнь). С практической точки зрения знание своей основной группы риска поможет человеку выстроить правильную стратегию профилактического поддержания своего здоровья. В связи с этим в настоящей статье развивается тема прогноза соответствующих групп риска для живущих людей.

Описание исходных данных для статистического анализа. В качестве исходного материала для пробного исследования возможностей использования паспортных сведений на первом этапе скрининга послужила выборка из 88 знаменитых людей преимущественно из артистической среды. По каждому человеку из интернета брались данные в виде строки: имя; фамилия; страна, в которой жил человек; тип