BIULETYN INSTYTUTU SYSTEMÓW INFORMATYCZNYCH 12 1–7 (2013)

# Metoda podziału zbioru obiektów na wielokryterialne klastry jakościowe

A. AMELJAŃCZYK
aameljanczyk@wat.edu.pl

Instytut Systemów Informatycznych
Wydział Cybernetyki WAT
ul. S. Kaliskiego 2, 00-908 Warszawa

W pracy przedstawiono ogólną procedurę tworzenia rankingów jakościowych elementów ustalonego zbioru obiektów. Procedura polega na rekurencyjnym wyznaczaniu elementów ekstremalnych zbioru na podstawie przyjętej relacji preferencji. Efektem jej działania jest podział zbioru na klastry rankingowe (kategorie). Przedstawione metody podziału zbioru na klastry mogą być wykorzystywane w analizie jakościowej obiektów w wielu praktycznych zastosowaniach.

**Słowa kluczowe:** ranking, relacja preferencji, elementy ekstremalne, klastry rankingowe, kategorie, punkt idealny, funkcje rankingowe, ranking liniowy.

## 1. Wprowadzenie

Zagadnienie tworzenia rankingów zbioru obiektów ocenianych wielokryterialnie jest bardzo powszechne w wielu dziedzinach życia. Różni się od klasycznego zagadnienia optymalizacji tym, że określa „jakość” wszystkich obiektów od „najlepszego do najgorszego”, a nie wyłącznie element „najlepszy”, co jest domeną optymalizacji [1, 2, 7]. Typowe rankingi, konstruowane na potrzeby komercyjne czy też społeczne, tworzone są na ogół w oparciu o tzw. funkcje rankingowe lub też heurystycznie, a nawet intuicyjnie w oparciu o bardzo proste metody i narzędzia. Nie mają one zatem wystarczającego waloru obiektywizmu i często służą jedynie celom marketingowym, a nawet swego rodzaju manipulacjom. W wielu obszarach życia jest to jednak niedopuszczalne i bardzo szkodliwe. Takimi przykładami są procedury ustalania rankingu ofert w zamówieniach publicznych (przetargach) i wszelkich innych przedsięwzięciach, jak konkursy projektów, grantów itp., w których wynik rankingu ma konsekwencje finansowe, prestiżowe lub handlowe. Powszechnie stosowane metody polegające na tworzeniu tzw. „ocen ważonych” czy też „ważonych decyzji ekspertów” są obarczone subiektywizmem i możliwością kamuflowanej manipulacji. Łatwo można pokazać, że dobierając odpowiednie wartości współczynników wagowych lub wartości progów, można istotnie zmieniać ranking wynikowy.

Celem niniejszej pracy jest pokazanie metody tworzenia rankingów odpornych na możliwość manipulacji, obiektywnie wynikających z przyjętego modelu preferencji, zdefiniowanego w przestrzeni kryteriów oceny poszczególnych obiektów.

## 2. Ogólne zadanie tworzenia rankingu elementów zbioru

Do sformułowania zadania rankingu zbioru można podchodzić na wiele sposobów [3, 4, 5, 8]. Formalnie ranking elementów zbioru można rozumieć jako przekształcenie zbioru w pewien szczególny ciąg podzbiorów tego zbioru lub jego elementów. Jeśli liczność tego zbioru jest równa N, to maksymalnie możemy otrzymać N! różnych ciągów (N! różnych rankingów – uporządkowań).

Niech zatem Y – niepusty, skończony zbiór elementów (obiektów), który ma podlegać procedurze rankingowej. $R \subset Y \times Y$ – relacja poprzedzania rozumiana następująco: para (y, z) należy do relacji wtedy i tylko wtedy, gdy „element y jest przed elementem z”. Zdanie „y jest przed z” (albo: „y poprzedza z”) może być rozumiane bardzo szeroko. Najczęściej jednak jest rozumiane w kontekście jakościowym „y jest lepszy od z” [2, 3]. Relacja R bywa nazywana relacją preferencji poprzedzenia, albo relacją rankingowania. Parę (Y, R) nazywać będziemy zbiorem z relacją [2].

Symbolem $2^{Y\times Y}$ oznaczać będziemy zbiór wszystkich relacji generowanych przez zbiór Y. Oczywiście nie wszystkie relacje preferencji jakościowych prowadzą do rankingu jednoznacznego (liniowego) [5]. Szczególne znaczenie mają jednak takie relacje preferencji, które wprowadzają porządek, a szczególnie tak zwany porządek liniowy [6]. Podstawą tworzenia każdego algorytmu rankingowego są tak zwane elementy ekstremalne zbioru w przestrzeni ze zdefiniowaną relacją poprzedzania (jakości) [6]. Elementy ekstremalne zbioru wykorzystywane w rankingowaniu to elementy najmniejsze i minimalne lub największe i maksymalne.

Na użytek niniejszych rozważań korzystać będziemy z pojęć:

$Y^R_{inf}$ – zbiór elementów najmniejszych

$Y^R_{min}$ – zbiór elementów minimalnych,

które w ogólności można (bez żądania własności porządkowych [2] relacji) zdefiniować następująco [2]:

$$Y^R_{inf} = \{y \in Y | (y,z) \in R \text{ dla każdego } z \in Y \setminus \{y\}\} \tag{2.1}$$

$$Y^R_{min} = \{y \in Y | \text{ nie istnieje } z \in Y \setminus \{y\}, \text{ że } (z,y) \in R\} \tag{2.2}$$

Element najmniejszy zbioru Y to taki, który poprzedza wszystkie pozostałe elementy ze zbioru Y. Element minimalny zbioru Y to taki element, który nie jest poprzedzany przez żaden z pozostałych elementów zbioru Y. Do utworzenia rankingu elementów zbioru Y (w skrócie będziemy mówili „do utworzenia rankingu zbioru Y") możemy wykorzystać zbiory elementów ekstremalnych – w zależności od potrzeb: zbiór elementów najmniejszych (lub zbiór elementów minimalnych) w następujący sposób: ze zbioru Y wybieramy elementy, które poprzedzają wszystkie pozostałe elementy. Następnie ze zbioru pozostałych elementów wybieramy kolejne elementy, które poprzedzają wszystkie pozostałe, i tak aż zbiór pozostałych elementów stanie się pusty. Zapiszemy to następująco:

1. Wyznacz $Y^R_{inf}(1) = Y^R_{inf}$ (2.3)

2. Wyznacz $Y^R_{inf}(2) = (Y \setminus Y^R_{inf}(1))^R_{inf}$ itd.

3. Ogólnie zapiszemy:

$$Y^R_{inf}(k) = \left(Y \setminus \bigcup_{i=0}^{k-1} Y^R_{inf}(1)\right)^R_{inf}, \; k = 1,2,3,... \tag{2.4}$$

przyjmując, że $Y^R_{inf}(0) = \emptyset$

Algorytm kończy się w kroku k = K takim, że

$$\bigcup_{i=1}^{K} Y^R_{inf}(i) = Y \tag{2.5}$$

Podobną procedurę możemy zastosować, korzystając z zależności (2.2). Otrzymamy wtedy ( przyjmując, że $Y^R_{min}(0) = \emptyset$):

$$Y^R_{min}(k) = \left(Y \setminus \bigcup_{i=0}^{k-1} Y^R_{min}(i)\right)^R_{min}, \; k = 1,2,3,... \tag{2.6}$$

<u>Określenie 2.1</u>

Klastrem jakościowym (kategorią) *k*-tego stopnia nazywać będziemy zbiór $Y^R_{min}(k), \; k = 1,2,3,...$

Najszczęśliwszym przypadkiem jest oczywiście przypadek, gdy wszystkie zbiory $Y^R_{inf}(k)$ lub $Y^R_{min}(k)$ są jednoelementowe, czyli przypadek, gdy

$$|Y^R_{inf}(k)| = 1 \text{ lub } |Y^R_{inf}(k)| = 1 \text{ dla } k = 1, \dots, K \tag{2.7}$$

Mamy wtedy tzw. ranking jednoznaczny (liniowy). Taką sytuację mamy między innymi wtedy, gdy relacja jakości (rankingowania) R jest relacją liniowego porządku (zwrotna, antysymetryczna, przechodnia i spójna [6]). W przypadku gdy relacja R jest relacją liniowego porządku, symbolem $Y(R)_{inf}$ lub $Y(R)_{min}$ oznaczać będziemy ranking zbioru Y (ciąg podzbiorów (klastrów) jednoelementowych) uzyskany przy relacji preferencji R. Ciąg zbiorów $[Y^R_{inf}(k)]$, lub analogicznie $[Y^R_{min}(k)]$, dla $k = 1,2,...$ tworzy w ogólności ranking klas równoważności (klastrów) elementów najmniejszych lub minimalnych. W ogólnym przypadku każdemu klastrowi $Y^R_{min}(k)$ można przyporządkować liczbę rankingową jako pewną ogólną charakterystykę jakościową tego klastra.

<u>Określenie 2.2</u>

Liczbą rankingową $L_k$ klastra stopnia k zbioru Y nazywać będziemy odległość tego klastra od kresu dolnego zbioru Y

$$L_k = \min_{y \in Y^R_{min}(k)} \min_{\overset{*}{y} \in \inf^Y_R} \left\| \overset{*}{y} - y \right\|, \; k = 1,2,... \tag{2.8}$$

Dodatkową charakterystyką klastrów są ich kresy dolne (górne) [2, 6]:

$$\left\{\overset{*}{y}{}^k\right\} = \inf_R Y^R_{min}(k), \; k = 1,2,... \tag{2.9}$$

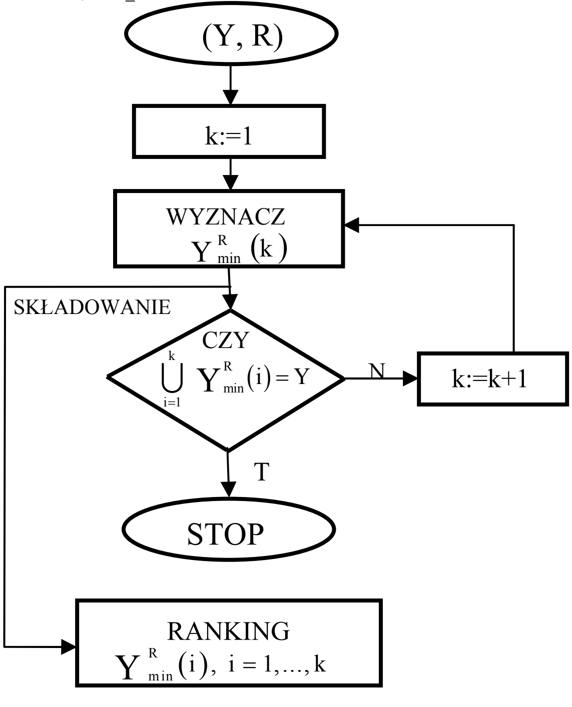
Liczby rankingowe klastrów, jak też ich kresy dolne, można wykorzystać do porządkowania klastrów oraz ich oceny jakościowej [3, 8].

Analogicznie, każdemu klastrowi $Y^{R}_{inf}(k)$ (elementów najmniejszych) można przyporządkować liczbę rankingową jako pewną ogólną charakterystykę jakościową tego klastra.

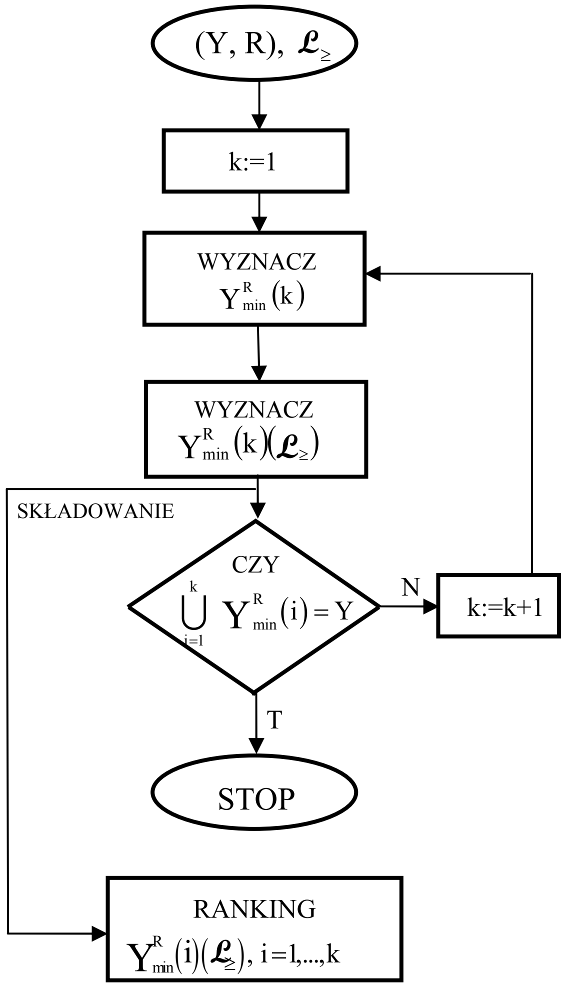
Zbiory (2.1) i (2.2) mają dużo bardzo interesujących „własności rankingowych". W praktycznych zastosowaniach rankingi elementów zbioru dotyczą głównie szeroko rozumianej jakości tych elementów i są tworzone w tak zwanej przestrzeni jakości.

Na rysunku 1 został przedstawiony ogólny schemat podziału zbioru Y na klastry rankingowe z wykorzystaniem elementów minimalnych zbioru.

Schemat procedury rankingowej przedstawiony na rysunku 1 nie gwarantuje oczywiście rankingu liniowego. W sytuacji gdy niektóre zbiory $Y^{R}_{min}(k)$ nie są jednoelementowe, potrzebna jest dodatkowa „informacja preferencyjna" dotycząca poprzedzania w sensie jakości w ramach elementów tego samego klastra. W takim przypadku mogą być stosowane różne sposoby wewnętrznego rankingu (np. dodatkowe funkcje rankingowe). Bardzo często jednak decydent jest pytany o hierarchię ważności (leksykografię) poszczególnych kryteriów cząstkowych (cech jakościowych analizowanych obiektów). Praktyczne eksperymenty decyzyjne na ogół potwierdzają dużą podatność decydentów na takie „uzupełnienie preferencji" rankingowych. Tak ustalone preferencje leksykograficzne w ramach wewnętrznych rankingów oznaczać będziemy $\mathcal{L}_{\geq}$.




Rys. 1. Podział zbioru Y na klastry rankingowe




Rys. 1a. Ogólny schemat procedury rankingowej

Symbolem $Y^{R}_{min}(i)(\mathcal{L}_{\geq})$, $i = 1,\dots, k$ oznaczać będziemy ranking leksykograficzny klastra numer i. Idea ogólnego schematu procedury rankingowej zbioru Y przy relacji preferencji R wspomaganej leksykografią $\mathcal{L}_{\geq}$ przedstawiona jest na rysunku 2. W schemacie tym występuje obok głównej (zasadniczej) relacji rankingowania R, relacja pomocnicza $\mathcal{L}_{\geq}$, która „reaguje" w sytuacji, gdy zbiór $Y^{R}_{min}(k)$ ma liczność większą od 1.

## 3. Przestrzenie jakości – rankingi jakościowe

Procedury rankingowe w zastosowaniach praktycznych najczęściej oparte są na modelach jakościowych „rankingowanych obiektów". Stąd też pierwszym etapem jest ustalenie modelu jakościowego obiektów. Sprowadza się to oczywiście do zdefiniowania odpowiednich wskaźników jakości obiektów (cech jakościowych obiektów) [3]. Ich liczba i konkretna postać zależą oczywiście od potrzeb (celu) rankingu. W ten sposób obiektami porównań rankingowych stają się nie bezpośrednio obiekty, lecz ich modele jakościowe w przestrzeni jakości [1, 2]. Kolejnym etapem jest zazwyczaj ustalenie modelu preferencji jakościowych decydenta, najczęściej w postaci odpowiedniej relacji zdefiniowanej w przestrzeni jakości. Jest to podejście o wiele bardziej ogólne, bezpieczniejsze i „bardziej obiektywne” niż podejście polegające na przyjęciu skalarnej funkcji oceny (rankingu) np. w postaci sumy ważonej [5]. Definiowanie modelu preferencji w postaci relacji bywa najczęściej o wiele prostsze i nie wymagapodawania liczbowych wartości współczyn-ników, progów itp.

Praktyczne zadanie tworzenia rankingów jakościowych można zatem przedstawić następująco:
Niech:
$X \subset A$ – zbiór rozważanych obiektów (elementów) w pewnej przestrzeni obiektów A;
$\mathcal{N} = \{1,...,n,...,N\}$ – zbiór numerów wskaźników cząstkowych oceny obiektów;
B – przestrzeń ocen jakościowych obiektów (najczęściej $B = \mathcal{R}^N$);
$F : A \to \mathcal{R}^N$ – funkcja oceny obiektów;
$F(x) = y \in \mathcal{R}^N$ – ocena (model jakościowy) obiektu $x \in A$ mająca postać:
$F(x) = (F_1(x),..., F_n(x),...,F_N(x)) \in \mathcal{R}^N$
$F_n(x)$ – ocena obiektu $x \in X$ w sensie wskaźnika nr $n \in \mathcal{N}$ (3.1)
$F(X) = \{F(x) = y \in \mathcal{R}^N | x \in X\} = Y$ – obraz jakościowy zbioru obiektów X (3.2)

Model preferencji jakościowych (podstawa rankingu) możemy zdefiniować w postaci relacji preferencji R

$$R = \{(y,z) \in \mathcal{R}^N \times \mathcal{R}^N | y \text{ jest lepszy od } z\} \quad (3.3)$$

Jest to model preferencji „zbudowany” w przestrzeni B modeli jakościowych obiektów (elementów) ze zbioru X. Funkcję zdaniową $\varphi(y,z) \equiv$ „y jest lepsze od z” można definiować na bardzo wiele sposobów [2, 3].

Na etapie ustalania modelu jakościowego $F(x)$ obiektów $x \in X$ ważne jest, aby

$$\left|F^{-1}(\{F(x)\}) = 1\right| \text{ dla każdego } x \in X \quad (3.4)$$

Gwarantuje to jednoznaczność rankingu w zbiorze X, o ile występuje ona w zbiorze Y.

W dalszej części pracy rozważania związane z procedurami rankingowymi będą ilustrowane pewnym przykładem dotyczącym rankingu zbioru dwudziestu obiektów. Jakość tych obiektów jest oceniana wartościami dwóch cech. Poniższa tabela przedstawia zbiór tych obiektów (ponumerowanych indeksem k = 1, …, 20) oraz odpowiadające im wartości liczbowe cech jakościowych.

Tab. 1. Lista obiektów podlegających rankingowi

| k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $y_1^k$ | 2 | 4 | 5 | 6 | 6 | 6 | 5 | 3 | 2 | 1 |
| $y_2^k$ | 7 | 7 | 6 | 5 | 3 | 2 | 1 | 1 | 1 | 2 |

| k | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|
| $y_1^k$ | 0 | 1 | 2 | 4 | 5 | 4 | 3 | 2 | 3 | 3 |
| $y_2^k$ | 4 | 6 | 6 | 5 | 4 | 3 | 2 | 4 | 5 | 4 |

W przykładzie przyjęto założenie, że im większe są wartości rozpatrywanych cech, tym wyższa jest „ogólna jakość obiektu”. Modelem jakościowym obiektu nr k będzie zatem element:

$$y^k = (y_1^k, y_2^k) \in \mathcal{R}^2, k \in \mathcal{K} = \{1, ..., 20\} \quad (3.5)$$

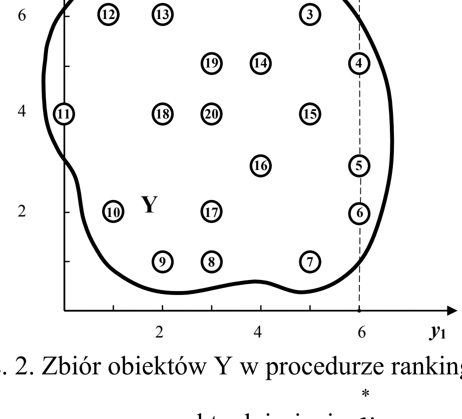
Zbiorem modeli jakościowych rozpatrywanych obiektów jest zbiór:

$$Y = \{y^k \in \mathcal{R}^2 | k \in \mathcal{K}\} \quad (3.6)$$

Na rysunku 2 dodatkowo został zaznaczony element idealny $\overset{*}{y}$, zdefiniowany następująco [2]:

$$\{\overset{*}{y}\} = \inf{}_R Y \quad (3.7)$$

Element idealny $\overset{*}{y}$ odgrywa bardzo ważną rolę w wielu procedurach rankingowych jako tak zwany „obiektywny punkt odniesienia jakościowego” [1, 2, 9].




Rys. 2. Zbiór obiektów Y w procedurze rankingowej oraz punkt odniesienia $\overset{*}{y}$

Na rysunku 2 przedstawiony został zbiór Y elementów podlegających procedurze rankingowania. Rankingi jakościowe zbioru Y mogą być tworzone w oparciu o zależności (2.1) lub (2.2). O jednoznaczności (liniowości) rankingów decydują głównie własności modelu preferencji R. Może się jednak zdarzyć, że przeciwobraz jednoznacznego rankingu zbioru Y nie będzie jednoznaczny w zbiorze X (patrz (3.4)). Oznaczać to będzie, że co najmniej dwa różne obiekty mają identyczny obraz w przestrzeni jakości (te same wartości wszystkich cech uwzględnianych w procesie modelowania [3].

## 4. Ranking leksykograficzny

W procedurach tworzenia rankingów bardzo ważną rolę odgrywają relacje porządku liniowego (zwrotne, antysymetryczne, przechodnie i spójne). Taką relacją preferencji jest relacja porządku leksykograficznego. Relacją preferencji leksykograficznego rankingowania nazywać będziemy relację $\mathcal{L}_{\geq}$ zdefiniowaną następująco:

$$\mathcal{L}_{\geq} = \{(y,z) \in Y \times Y \,|\, \text{istnieje } i \in \mathcal{N},\ \text{że } y_i > z_i \ \text{oraz } y_k = z_k \ \text{dla } k < i \ \text{lub } y = z\} \tag{4.1}$$

Ten model preferencji odpowiada sytuacji, gdy znana jest jedynie hierarchia ważności kryteriów cząstkowych (w tym przypadku zgodna z porządkiem naturalnym kryteriów).

Na rysunku 3 przedstawiono wyniki działania algorytmu rankingowego (2.4) oraz stożek $\Lambda(\mathcal{L}_{\geq})$, który generuje relację (4.1) w $\mathcal{R}^N$ dla $N = 2$ [2]. Zgodnie z zależnościami (2.3) i (2.4) otrzymujemy:

$$Y_{\inf}^{\mathcal{L}_{\geq}}(1) = Y_{\inf}^{\mathcal{L}_{\geq}} = \{4\}$$

$$Y_{\inf}^{\mathcal{L}_{\geq}}(2) = \left(Y \setminus Y_{\inf}^{\mathcal{L}_{\geq}}(1)\right)_{\inf}^{\mathcal{L}_{\geq}} = \{5\} \text{ itd.}$$

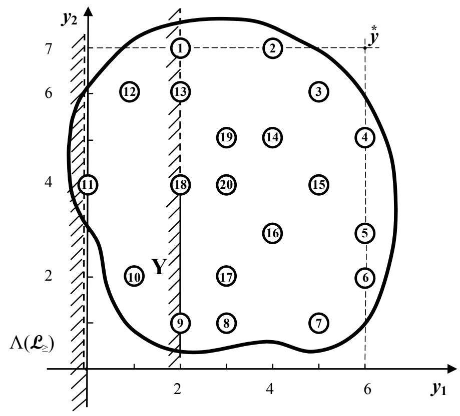
$$Y_{\inf}^{\mathcal{L}_{\geq}}(20) = \{11\}$$

Relacja $\mathcal{L}_{\geq}$ jest relacją liniowego porządku, stąd ranking jest jednoznaczny.

Otrzymany ranking $Y(\mathcal{L}_{\geq})_{\inf}$ ma postać:

$$Y(\mathcal{L}_{\geq})_{\inf} = (4, 5, 6, 3, 15, 7, 2, 14, 16, 19, 20, 17, 8, 1, 13, 18, 9, 12, 10, 11).$$

Przykładowo, element nr 18 poprzedza elementy o numerach: 9, 10, 11, 12, 18. Oznacza to, że element ten jest lepszy od tych elementów (formalnie jest też lepszy od samego siebie, gdyż relacja jest zwrotna). Najlepszym elementem jest zatem element o numerze 4, potem 5 itd., a najgorszym elementem jest element o numerze 11. Zaznaczono to na rysunku 3.




Rys. 3. Ranking leksykograficzny zbioru Y

Ranking $Y(\mathcal{L}_{\geq})_{\min}$ jest identyczny z rankingiem $Y(\mathcal{L}_{\geq})_{\inf}$ ponieważ $Y_{\inf}^{\mathcal{L}_{\geq}} = Y_{\min}^{\mathcal{L}_{\geq}}$ dla zadań z relacją $\mathcal{L}_{\geq}$ [2].

## 5. Ranking paretowski

Paretowska relacja jakości, będąca podstawą rankingu paretowskiego, jest najczęściej stosowaną w praktyce relacją preferencji [10]. Relację paretowską oznaczać będziemy symbolem „$\geq$”. Relacja ta prowadzi do uzyskania rozwiązań optymalnych w zadaniu wielokryterialnej maksymalizacji.

$$\geq = \{(y,z) \in Y \times Y \,|\, y_n \geq z_n \ \text{dla każdego } n \in \mathcal{N}\} \tag{5.1}$$

Analizy jakościowe w przypadku relacji co najmniej zwrotnych i przechodnich w procedurach rankingowych wygodnie jest prowadzić, korzystając z pojęcia stożka $\Lambda(R)$ związanego z rozpatrywaną relacją R. Związek $\Lambda(R)$ i R jest następujący [10]:

$$R = \{(y,z) \in Y \times Y \,|\, z \in \{y\} + \Lambda(R)\} \tag{5.2}$$

Wygodnym pojęciem jest również tzw. funkcja poprzedzania $d : Y \to 2^Y$, zdefiniowana następująco:

$$d(y) = \{z \in Y \,|\, (y,z) \in R\},\ y \in Y \tag{5.3}$$

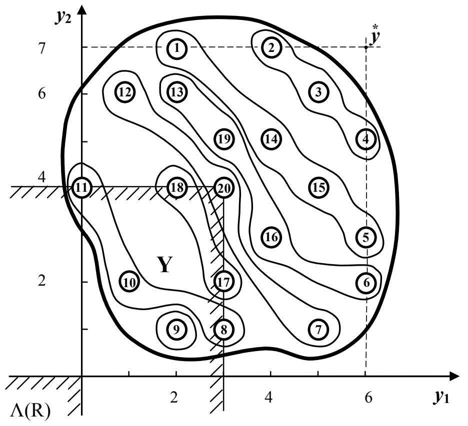
Zależność (5.3) możemy też zapisać następująco:

$$d(y) = \{y\} + \Lambda(R) \tag{5.4}$$

Na rysunku 4 zaznaczono między innymi

$$d(20) = \{20\} + \Lambda(R) = \{20, 17, 18, 8, 10, 11, 9\}$$

Na rysunku 4 przedstawiony został ranking paretowski (podział zbioru na klastry Pareto) wynikający z paretowskiej relacji preferencji jakości „$\geq$”.




Rys. 4. Podział zbioru Y na klastry Pareto

Z uwagi na to, że zbiór elementów najmniejszych dla relacji „$\geq$” najczęściej jest pusty, tworząc ranking zbioru Y skorzystamy z zależności (2.2).
Otrzymamy więc:

$$Y^{\geq}_{\min}(1) = Y^{\geq}_{\min} = \{2, 3, 4\}$$

$$Y^{\geq}_{\min}(2) = \left(Y \backslash Y^{\geq}_{\min}(1)\right)^{\geq}_{\min} = \{1,\ 5,\ 14,\ 15\}$$

$$Y^{\geq}_{\min}(3) = \left(Y \backslash \left(Y^{\geq}_{\min}(1) \cup Y^{\geq}_{\min}(2)\right)\right)^{\geq}_{\min} = \{6,\ 13,\ 16,\ 19\}$$

$$Y^{\geq}_{\min}(4) = \left(Y \backslash \left(Y^{\geq}_{\min}(1) \cup Y^{\geq}_{\min}(2) \cup Y^{\geq}_{\min}(3)\right)\right)^{\geq}_{\min} =$$

$$= \{7,\ 12,\ 20\},\ \text{itd.}$$

$$Y^{\geq}_{\min}(7) = \left(Y \backslash \bigcup_{i=1}^{6} Y^{\geq}_{\min}(i)\right)^{\geq}_{\min} = \{9\}.$$

Jak widać, zastosowanie paretowskiej relacji preferencji nie prowadzi w tym przypadku do rankingu jednoznacznego. Otrzymaliśmy klasy (klastry) równoważności $Y^{\geq}_{\min}(k)$, $k = 1, \ldots, 7$ w większości o liczności większej od 1. „Najwyższy” w rankingu jest klaster $Y^{\geq}_{\min}(1)$, potem $Y^{\geq}_{\min}(2)$, aż do $Y^{\geq}_{\min}(7)$. Do oceny jakościowej poszczególnych klastrów można wykorzystać dodatkowe charakterystyki w postaci liczb rankingowych klastrów oraz ich kresów dolnych, które łatwo wyznaczyć (2.8), (2.9). Zachodzi przy tym:

$$L_1 \leq L_2 \leq \ldots \leq L_k \leq \ldots \tag{5.5}$$

oraz

$$\overset{*}{y}{}^{1} \geq \overset{*}{y}{}^{2} \geq \ldots \geq \overset{*}{y}{}^{k} \geq \ldots \tag{5.6}$$

Gdyby przyjęte preferencje jakościowe „uzupełnić” o preferencje leksykograficzne (w przypadku zbiorów $Y^{\geq}_{\min}(k)$ o liczności większej od 1), można byłoby dokonać ich „cząstkowych rankingów wewnętrznych” (patrz schemat na rysunku 1). Dla leksykografii naturalnej otrzymalibyśmy następujące rankingi częściowe:

$$Y^{\geq}_{\min}(1) \to Y^{\geq}_{\min}(1)(\mathcal{L}_{\geq}) = (4, 3, 2)$$

$$Y^{\geq}_{\min}(2) \to Y^{\geq}_{\min}(2)(\mathcal{L}_{\geq}) = (5, 15, 14, 1)$$

$$Y^{\geq}_{\min}(3) \to Y^{\geq}_{\min}(3)(\mathcal{L}_{\geq}) = (6, 16, 19, 13)$$

$$Y^{\geq}_{\min}(4) \to Y^{\geq}_{\min}(4)(\mathcal{L}_{\geq}) = (7, 20, 12)\ \text{itd.}$$

oraz w konsekwencji jednoznaczny, jakościowy ranking końcowy (od obiektu najlepszego, którym jest obiekt o numerze 4 do obiektu najgorszego, którym jest obiekt numer 9):

$$Y(\geq, \mathcal{L}_{\geq})_{\min} = (4, 3, 2, 5, 15, 14, 1, 6, 16, 19,$$
$$13, 7, 20, 12, 17, 18, 8, 10, 11, 9).$$

Uzyskane klastry o numerach od 1 do 7 tworzą natomiast podział zbioru Y na pewne „warstwy” (kategorie) elementów równoważnych w sensie jakościowym. Najlepszym klastrem jest klaster o numerze 1, a najgorszym klaster numer 7.

## 6. Podsumowanie

W pracy przedstawiono ogólne podejście do problematyki rankingu elementów analizowanego zbioru Y. Jest to podejście znacznie bardziej ogólne niż podejście oparte wyłącznie na funkcjach rankingowych. Podstawą rankingu może być dowolny model preferencji. Najczęściej jest on przedstawiony w postaci relacji $R \subset Y \times Y$. „Skuteczność procedury rankingowej” zależy oczywiście od własności relacji preferencji. Do definiowania relacji preferencji można użyć również dowolnej funkcji rankingowej. Definiowanie preferencji rankingowych w postaci relacji jest jednak postępowaniem o wiele bardziej obiektywnym i praktycznym niż definiowanie funkcji rankingowej. W pracy rozważono relacje preferencji typu Pareto [2, 10] oraz preferencje hierarchiczne.

Ogólna procedura rankingowa oparta na relacji preferencji określa podział zbioru na „klastry jakościowe” (kategorie), tworzące pewien ciąg rankingowy klastrów. Klastry te można opatrzyć odpowiadającą im „liczbą rankingową” świadczącą o ich jakości. Liczność klastrów zależy oczywiście od własności relacji preferencji i „natury” zbioru Y.

W pracy przedstawiono uniwersalny algorytm rankingowy pozwalający otrzymać ranking liniowy niezależnie od własności podstawowej relacji preferencji. W algorytmie tym wykorzystywana jest pomocniczo relacja leksykograficzna w zakresie klastrów o liczności większej niż 1. Praktyka decyzyjna pokazuje, że decydenci bardzo łatwo akceptują wykorzystywanie pomocniczej relacji leksykograficznej w sytuacji pojawiania się klastrów o liczności większej niż 1. Podejście takie eliminuje konieczność określania jakichkolwiek wag, czy też progów. Większość stosowanych w praktyce relacji rankingowych to relacje quasi-porządku lub porządku. Relacje te pozwalają tworzyć bardzo interesujące rankingi klastrów jakościowych. Znalazły one między innymi zastosowanie w analizach scientometrycznych dotyczących wpływu poszczególnych naukowców, źródeł bibliograficznych, instytucji naukowych czy też państw na rozwój nauki w różnych dyscyplinach. Mogą być też z powodzeniem wykorzystywane w wielu rankingach dotyczących jakości „ważnych obiektów", w tym jakości wyrobów na rynku.

## 7. Bibliografia


[1] A. Ameljańczyk, *Optymalizacja wielokryterialna*, WAT, Warszawa, 1986.

[2] A. Ameljańczyk, *Optymalizacja wielokryterialna w problemach sterowania i zarządzania*, Ossolineum, Wrocław, 1984.

[3] A. Ameljańczyk, „Matematyczne aspekty modelowania pajęczynowego obiektów", *Biuletyn Instytutu Systemów Informatycznych*, Nr 4, 9–16 (2009).

[4] D. Bouyssou, T. Marchant, „An axiomatic approach to noncompensatory sorting methods in MCDM, I: The case of two categories", *EJOR* (2007).

[5] J.P. Brans, Ph. Vincke, "A preference ranking organization method: The PROMETHEE method", *Management Science*, Vol. 31, No. 6 (1985).

[6] H. Rasiowa, *Wstęp do matematyki współczesnej*, PWN, Warszawa, 2005.

[7] T.L. Saaty, "Rank from comparisons and from ratings in the analytic hierarchy/networkprocesses", *EJOR* (2006).

[8] F. Seo, M. Sakawa, *Multiple Criteria Decision Analysis in Regional Planning: Concept*, *Methods and Applications*, D. Reidel Publishing Company, Dordrecht, Boston, Lancaster, Tokyo, 1988.

[9] P.L. Yu, G. Leitmann, "Compromise solutions, domination structures and Salukwadze's solution", *JOTA*, Vol. 13 (1974).

[10] P.L. Yu, G. Leitmann, "Nondominated decision and cone convexity in dynamic multicriteria decision problems", *JOTA*, Vol. 14 (1974).


# The method of distribution of a set of objects into multi-criteria quality clusters

A. AMELJAŃCZYK

The paper presents a general procedure for creating quality rankings of objects. Ranking procedure fixed set of elements by recurrent determining the extreme elements of the set on the basis of its preference relation. The result of the procedure is to divide the set into ranking clusters (categories). The method splits the set of the clusters can be used in qualitative analysis of objects in many practical applications.

**Keywords:** ranking, preference relation, extreme elements, ranking clusters, categories, ideal point, ranking functions, linear ranking.