

DIFFERENCIÁLT TÁJÉKOZTATÁSI RENDSZER SZERVEZÉSE IPARVÁLLALATOKNÁL

II. AZ INFORMÁCIÓFELHASZNÁLÓK CSOPORTOSÍTÁSÁNAK MÓDSZEREI

Schmél Ferencné

VBKM Transzvíll Gyára

Előző cikkünkben [1] áttekintést adtunk a differenciált tájékoztatás iparvállalati szükségességéről, és egy új módszert mutattunk be az igények felméréséhez. A rendelkezésre álló adatokat a hagyományos tájékoztatási rendszerek egyetlen halmazként kezelték, ezzel szemben a differenciált tájékoztatási rendszer arra a felismerésre épül, hogy az információkat hasznosítók viszonylag jól elkülöníthető csoportokat alkotnak. Ebben a tanulmányban a csoportképzés különböző módszereivel foglalkozunk.

A csoportosítás elvei és gyakorlata

A csoportosítás (osztályozás) lényege az, hogy egy adott alapsokaságból valamilyen közös jegyek alapján több kisebb – nem feltétlenül egyforma nagyságú – egységet képeznek. Ha erre a lebontásra csak egyszer kerül sor, akkor *egyszintű osztályozásról* beszélünk: ekkor az egyes csoportok egymás mellé rendelve. A kialakított csoportok valamelyikén (vagy mindegyikén) belül – most már az előbbiektől valamelyest eltérő jegyek szerint – megismételve ezt a műveletet, az ún. *hierarchikus osztályok* jönnek létre: itt mindegyik osztály valamelyik magasabb rendű csoport („gyűjtőfogalom”) részét képezi.

A tájékoztatási rendszer szervezésekor a vállalat szakemberei (a műszaki információkat felhasználó munkatársak) képezik az alaphalmazt. Mivel mindegyikük egyéni igényének maradéktalan és specializált kielégítése nem lenne gazdaságos és nem is indokolt, a tájékoztatás alanyait ésszerű módon csoportokba kell rendezni. Ezek a részhalmazok (szakembercsoportok) ezután homogén

egységként kezelendők: tagjaik azonos, de a másik csoporttól többé-kevésbé eltérő tájékoztatásban részesülnek. Belátható, hogy a differenciált tájékoztatási rendszer hatékonysága alapvetően a csoportosítás sikerességétől függ.

A hagyományos osztályozást másképpen *a priori* csoportosításnak is nevezhetjük, mivel az alaphalmaz felosztásakor, illetve az egyes részhalmazok kialakításakor az a meghatározó, hogy az egyedek (a halmaz elemei) melyik előre megadott ismérvek felelnek meg. A csoportosítás tehát valamilyen előre elhatározott kritérium-rendszer szerint történik. Az iparvállalati szakemberek esetében pl. a következő osztályozási lehetőségek kínálkoznak:

- a) iskolai végzettség (egyetem, főiskola, technikum, gimnázium stb.);
 - b) munkakör (tervező, fejlesztő, termelésirányító stb.);
 - c) tevékenységi szint (felsőszintű irányító, végrehajtó stb.);
 - d) szakterület (műszaki, közgazdasági, ügyviteli dolgozó stb.);
 - e) életkor vagy szakmai gyakorlat;
- illetve ezek különböző kombinációi.

Az ilyen módszerrel létrehozott – esetleg mégoly szűk – csoportok is általában elég heterogének a tájékoztatási igények, a tájékozódási szokások és az érdeklődési kör tekintetében. Az egyének szempontjából ez azt jelenti, hogy a kézhez kapott információk redundanciája („haszontalan” részének aránya) viszonylag magas, miközben körük sem teljes. Az egész csoport szemszögéből tehát a tájékoztatás határfoka nem kielégítő.

Felmerül a kérdés: nem lehetne-e úgy csoportosítani az iparvállalat szakembereit, hogy az azonos osztályba kerülők tájékoztatási igénye minél homogénebb legyen, mégpedig függetlenül attól, hogy egyébként mely területen tevékenykednek, milyen iskolai végzettséggel bírnak stb.

Taxonómiai alapfogalmak

Az automatikus osztályozás több tényezővel (paraméterrel, mutatóval) jellemzett dolgok csoportosítását teszi lehetővé úgy, hogy az együvé tartozás vagy elkülönülés megállapításakor mindegyik tényezőt objektív módon vesszük figyelembe. (Az automatikus osztályozás terjedt el leginkább a magyar szakirodalomban [2, 3] – emellett használatos még a *numerikus taxonómia*, a *klaszter analízis* és a *csoportosítási algoritmus*.) A cél olyan csoportok generálása az alapsokaságból, amelyekben belül a vizsgált tényezőket illetően a homogenitás biztosított.

Az alkalmazott módszer (matematikai apparátus) viszonylag egyszerű, így nincs akadálya annak, hogy bármelyik tájékoztatási szakember elsajátítsa. A következőkben előbb magát az eljárást ismertetjük, majd egy konkrét példán keresztül tesszük szemléletessé.

Elsőként a csoportosítás során figyelembe veendő tényezőket kell megválasztani (pl. az egyes szakemberek által a különböző tájékoztatási szolgáltatásokhoz rendelt preferenciasúlyokat).

A kiinduló adatokat egy olyan mátrixba rendezzük, amelynek sorai a válaszadókhoz (egyedekhez), oszlopai pedig a vizsgált tényezőkhöz tartoznak. A mátrix minden egyes eleme a sor szerinti megkérdezettnek az oszlop szerinti szolgáltatáshoz rendelt preferenciasúlya, tehát a paraméter aktuális értéke. Jelöljük ezt a mátrixot Z -vel: $Z = (z_{ij})$, ahol z_{ij} – az i -edik megkérdezett által a j -edik tényezőnek tulajdonított paraméter.

A következő lépésben meghatározzuk az egyes válaszadók közötti véleményeltérést kifejező ún. *taxonómiai távolságokat*:

$$d_{kr} = \sqrt{\sum_{j=1}^n (z_{kj} - z_{rj})^2} \quad (1)$$

ahol

z_{kj} – a k -edik válaszadó által a j -edik tényezőnek adott súlysám;

z_{rj} – az r -edik megkérdezett által ugyanannak (j -edik) tényezőnek tulajdonított súlysám;

n – a vizsgált tényezők száma.

A d_{kr} értékeket a D mátrixba rendezzük (könnyű belátni, hogy ez szimmetrikus).

Az automatikus osztályozáshoz meg kell adni egy olyan d_1^* *távolsági küszöbértékét*, amely az együvé

tartozás kritériuma lesz. Ennek megfelelően most képezzünk egy A mátrixot a következő szabály szerint:

$$\begin{aligned} \text{ha } d_{kr} \leq d_1^*, & \text{ akkor } y_{kr} = 1 \\ \text{ha } d_{kr} > d_1^*, & \text{ akkor } y_{kr} = 0 \end{aligned} \quad (2)$$

$$\text{és } d_{kk} = 1$$

A csoportképzés az A négyzetes mátrixon végrehajtható ún. *transzitiv lezárás*on alapszik [3]. Az A mátrixot addig szorozzuk önmagával, amíg olyan mátrixhoz jutunk, amely a további hatványozás során már nem változik, azaz

$$A^p = A^p \cdot A \quad (3)$$

ahol p – pozitív egész szám ($p \neq 0$). A mátrixalgebrában az ilyeneket *idempotens mátrixnak* nevezik [4]. (Megjegyezzük, hogy az A mátrix ilyen értelmű hatványozása mindig lehetséges, mivel A négyzetes. A szorzás során a részműveleteket a logikai szorzás és összeadás szabályának megfelelően kell elvégezni.)

Bizonyított, hogy lezárás – azaz változatlan mátrixok bekövetkezése a hatványozás során – véges számú lépésben bekövetkezik, és az eredmény természetesen megint Boole-mátrix lesz (a Boole-mátrixok csupán egyeseket és zérusokat tartalmaznak). Ennek egy további igen fontos tulajdonsága az, hogy a sorok és oszlopok alkalmas cseréjével (permutálásukkal) ún. *blokkdiagonális* mátrixszá alakítható: az azonos blokkokhoz tartozó egyedeket célszerű a következőkben egy-egy csoportként kezelni.

A választott d_1^* küszöbérték nagyságától függően alakul majd a generált csoportok száma (a létrejött csoportok természetesen lehetnek egyeleműek is). Ha az előállított csoportok száma túl nagy, akkor egy $d_2^* > d_1^*$ küszöbértékkel célszerű a transzitiv lezárást újra elvégezni. Hasonlóképpen túl nagy d_1^* esetén igen kevés (de legalább két) csoport jön létre: ilyenkor a helyzet egy $d_2^* < d_1^*$ értékkel javítható.

Gyakorlati alkalmazás

A VBKM Transzvill Gyárában 1979-ben végeztünk felmérést a műszaki területeken dolgozó vezetők körében, a tájékoztatási igények feltárása céljából [5]. A vezetői tájékoztatási rendszer szervezése szempontjából tényezőnek tekintettük a különböző szolgáltatások iránti igényeket (pontosabban az azok mértékét kifejező súlysámokat), egyednek pedig a műszaki irányító apparátus tagjait. A csoportosítás célja az egyedek olyan részalmazainak létrehozása volt, amelyek tagjai hasonló jellegű szolgáltatásokat preferálnak, az egyes csoportok

pedig a differenciált tájékoztatás címzettjei. Az egyes csoportok munkakör, életkor (szakmai gyakorlat), iskolai végzettség stb. szerinti összetétele ily módon inhomogén lesz; az összetartozás egyedüli kritériuma a tájékoztatási igény.

A könnyebb áttekinthetőség kedvéért a komplex vizsgálatból most 8 vezetőt és 6-féle szolgáltatást ragadunk ki. Az egyes vezetők által a különböző szolgáltatásokhoz rendelt *preferenciasúlyokat* az 1. táblázat tartalmazza (a preferenciasúlyok meghatározásának módját lásd az [1]-ben). Ezután felhasználva az (1) képletet, kiszámítottuk az egyes vezetők közötti taxonómiai távolságokat, és a 2. táblázatban gyűjtöttük össze őket. Megfigyelhető, hogy a főatlóban zérusok állnak, ugyanis minden vezető véleménye „egybeesik a sajátjával” (azaz a távolság a saját véleményétől értelemszerűen éppen nulla). Látható továbbá az is, hogy a táblázat valójában egy szimmetrikus négyzetes mátrix, hiszen két szakértő véleményének eltérése „oda-vissza” azonos: $d_{ij} = d_{ji}$.

A kritikus távolsáértéket $d_1^* = 30$ -ra választottuk (a 2. táblázattal összevetve: $d_1^* > d_{\min} = d_{78} = 12,9$ és $d_1^* < d_{\max} = d_{57} = 108,9$). Felhasználva most a (2)-ben található relációkat, képeztük a 3. táblázatban feltüntetett \underline{A} mátrixot. Gyakorlatilag ez úgy történt, hogy ahol a 2. táblázatban 30-nál kisebb értéket (vagy éppen 30-at) találtunk, a 3. táblázat ugyanezen elemébe 1-et írtunk, ellenkező esetben pedig nullát. Ezután feltöltöttük a főátló alatti elemeket is úgy, hogy a meglévő értékeket a főátlóra tükröztük ($a_{ij} = a_{ji}$).

A következő lépés a *transzitiv lezárás*: az \underline{A} mátrixot megszorozzuk önmagával, a matematikai logika szabályait felhasználva, azaz:

$$\begin{array}{ll} 0 + 0 = 0 & 0 \cdot 0 = 0 \\ 0 + 1 = 1 & 0 \cdot 1 = 0 \\ 1 + 0 = 1 & 1 \cdot 0 = 0 \\ 1 + 1 = 1 & 1 \cdot 1 = 1 \end{array}$$

Ezután a kapott eredménymátrixot megint \underline{A} -val szorozzuk, és az eljárást addig folytatjuk, amíg újabb szorzással a mátrix változatlan marad. Esetünkben ez már az első szorzás után bekövetkezik, azaz

$$\underline{T} = \underline{A}^2 = \underline{A}^3 = \dots \quad (4)$$

A $\underline{T} = \underline{A}^2 = \underline{A} \cdot \underline{A}$ eredménymátrixot a 4. táblázat szemlélteti. Ha most ennek a 6. és 8. sorát, illetve ugyanezen két oszlopát felcseréljük, akkor az összes „egyes a főátló körül csoportosul” (azaz blokk-diagonális mátrixhoz jutunk), amint azt az 5. táblázat mutatja.

Példánkban mindössze egy többelemű blokk van: ez a 4, 5, 7 és 8 jelű sor és oszlop által meghatározott. A gyakorlatban ez azt jelenti, hogy a főmeős, a karbantartási és a termelési osztály vezetője, valamint a főművezető tájékoztatási igényei ebben a gyárban azonosnak

tekinthetők. A többi vezető tájékoztatási szükséglete ezekétől és egymásétól jobban eltér, ezért ők sajátos információellátásban részesülnek majd.

Összefoglalás

A csoportosítás lehetőséget nyújt arra, hogy a differenciált tájékoztatási rendszerben az egymástól különböző szolgáltatások számát az ésszerűség és a gazdaságosság határain belül csökkentjük. Az *automatikus osztályozás* több tényező egyidejű figyelembevételével objektív módon csoportba sorolja az azonos tájékoztatást igénylő szakembereket, függetlenül attól, hogy milyen más egyéni adottságúak.

Az automatikus osztályozás számítógépi programja BASIC nyelven mindössze 60 sor, futási ideje is mindössze néhány perc. Ez igen előnyössé és gazdaságossá teszi ezt az eljárást akár néhány száz szakember véleményének egyidejű érvényesítése esetén is.

1. táblázat

Különböző vezetők által meghatározott preferenciasúlyok mátrixa

	Szakkönyvek	Kézikönyvek	Szakfolyóiratok	Dokumentációs összeállítások	Kézikönyvtár	Fordítások
1. Főmérnök	42	45	55	84	72	21
2. Főtechnológus	82	63	75	44	40	75
3. Fejlesztési ov.	49	44	63	81	76	75
4. Főmeős	53	87	29	17	69	57
5. Karbantartási ov.	47	91	41	12	85	29
6. Műszaki tanácsadó	45	37	85	81	52	63
7. Termelési ov.	46	88	39	9	74	37
8. Főművezető	50	85	27	12	69	31

2. táblázat

Taxonómiai távolságok

	Fő- mérnök	Főtech- nológus	Fejlesztési ov.	Főmeős	Karban- tartási ov.	Műszaki tanács- adó	Termelési ov.	Főműve- zető
1. Főmérnök	0	88,7	54,7	99,0	88,1	56,1	89,5	88,0
2. Főtech- nológus		0	71,4	80,3	91,3	61,0	84,8	87,3
3. Fejlesztési ov.			0	94,5	105,5	54,0	103,4	105,8
4. Főmeős				0	34,4	96,2	25,3	26,8
5. Karban- tartási ov.					0	108,9	14,4	22,4
6. Műszaki tanácsadó						0	102,8	108,5
7. Termelési ov.							0	12,9
8. Főmű- vezető								0

3. táblázat

Összetartozási reláció
 $d^* = 30$

	1.	2.	3.	4.	5.	6.	7.	8.
1.	1	0	0	0	0	0	0	0
2.	0	1	0	0	0	0	0	0
3.	0	0	1	0	0	0	0	0
4.	0	0	0	1	0	0	1	1
5.	0	0	0	0	1	0	1	1
6.	0	0	0	0	0	1	0	0
7.	0	0	0	1	1	0	1	1
8.	0	0	0	1	1	0	1	1

4. táblázat

A tranzitív lezárás utáni \underline{A}^2 mátrix

	1.	2.	3.	4.	5.	6.	7.	8.
1.	1	0	0	0	0	0	0	0
2.	0	1	0	0	0	0	0	0
3.	0	0	1	0	0	0	0	0
4.	0	0	0	1	1	0	1	1
5.	0	0	0	1	1	0	1	1
6.	0	0	0	0	0	1	0	0
7.	0	0	0	1	1	0	1	1
8.	0	0	0	1	1	0	1	1

5. táblázat

Permutálás utáni eredménymátrix

	1.	2.	3.	4.	5.	6.	7.	8.
1.	1	0	0	0	0	0	0	0
2.	0	1	0	0	0	0	0	0
3.	0	0	1	0	0	0	0	0
4.	0	0	0	1	1	1	1	0
5.	0	0	0	1	1	1	1	0
6.	0	0	0	1	1	1	1	0
7.	0	0	0	1	1	1	1	0
8.	0	0	0	0	0	0	0	1



SCHMÉL Ferencné: Differenciált tájékoztatási rendszer szervezése iparvállalatoknál. 2. rész

Az ötrészes cikksorozat egy iparvállalati differenciált tájékoztatási rendszert feltételezve, matematikai számítási eljárásokat mutat be e rendszer tökéletesítési lehetőségeként. Az e számban közölt második közlemény az információfelhasználók csoportosításának egyik módszerét mutatja be, amely az automatikus osztályozáson alapul, s mint ilyen, lehetőséget nyújt a szolgáltatások számának az ésszerűség és a gazdaságosság határain belüli csökkentésére.

* * *

Mrs. SCHMÉL, S.: Organization of selective information services at industrial enterprises. Part 2

Assuming a selective information system at an industrial enterprise, the five-part series presents mathematical methods as a possible means of bringing the system to perfection. The second part, published in this issue, presents a method for forming user groups. The method is based on the theory of automatic classification and it facilitates a reasonable and economical reduction in the number of services.

* * *

Irodalom

1. SCHMÉL F.-né: Differenciált tájékoztatási rendszer szervezése iparvállalatoknál. 1. A tájékoztatási igények felmérésének korszerű módszerei = Tudományos és Műszaki Tájékoztatás, 28. köt. 7. sz. p. 286-291.
2. KINDLER J.-PAPP O.: Komplex rendszerek vizsgálata. Bp. Műszaki Könyvkiadó, 1977. p. 135-141.
3. PÁRNICZKY G.: A statisztikai informatika alapjai. Bp. Statisztikai Kiadó, 1976. p. 135-166.
4. SCHARNITZKY V.: Mátrixszámítás. Bp. Műszaki Könyvkiadó, 1970. p. 105-109.
5. SCHMÉL F.-né: Differenciált tájékoztatási rendszer szervezése iparvállalat műszaki vezetői részére. Szakdolgozat, 1980. p. 12-25.

ШМЕЛ, Ж.: Организация дифференцированной информационной системы на промышленных предприятиях. 2 часть

Серия статей, состоящая из 5 частей, рассматривает математические вычислительные процессы как возможность усовершенствования системы для случая дифференцированной информационной системы промышленного предприятия. Эта, вторая часть рассматривает один способ группировки потребителей информации, который основан на автоматической классификации и, как таковой, дает возможность уменьшить в разумных пределах и, соблюдая экономичность, число услуг.

* * *

Frau SCHMÉL, S.: Organisierung eines differenzierten Informationssystems bei Industriebetrieben, II

In der aus fünf Teilen bestehenden Artikelreihe werden für den Fall eines differenzierten Informationssystems eines Industriebetriebes mathematische Berechnungsverfahren als Vervollkommnungsmittel dieses Systems dargestellt. In diesem zweiten Teil der Mitteilung wird eine Methode der Gruppierung der Informationsbenutzer beschrieben, welche auf der automatischen Klassifikation beruht und als solche Möglichkeiten zur Reduzierung der Zahl der Dienstleistungen innerhalb der Grenzen der Rationalität und Wirtschaftlichkeit bietet.

* * *