

Validace modelů na vybraných reprezentativních vzorcích dat – shrnutí za třetí monitorovací období

V rámci této aktivity byly ve třetím monitorovacím období realizovány zbývající experimenty tak, aby všechny vytvořené modely byly řádně otestovány a validovány. Ve třetím monitorovacím období probíhalo především testování dvou různých přístupů – proporcionálního a empirického – k odhadu počtu výhledového počtu zásahů a ověřování vlivu těchto přístupů na řešení úlohy o rozmístování středisek ZZS v daném kraji. Rovněž byl vytvořen a otestován fuzzy přístup ke zpracování rozdílů mezi koeficienty modelu získané z proporcionálního a empirického přístupu. Zde je však nutné konstatovat, že výsledky série experimentů provedených s tímto fuzzy přístupem ukázaly, že rozdíly v hodnotách účelové funkce jsou pouze minimální. Podrobnější informace o testování fuzzy přístupu jsou uvedeny dále v tomto textu.

Testovanie fuzzy prístupu k odhadu budúcich počtov volaní z jednotlivých komunít

V tejto kapitole sa budeme zaoberať fuzzy prístupom k odhadu budúceho dopytu po záchranej zdravotnej službe. Intenzita tiesňových volaní, t. j. počet volaní za jednotku času v jednotlivých obciach regiónu v budúcnosti je neznáma, jej skutočnú hodnotu nepoznáme, môžeme ju len odhadnúť. Odhad sa obvykle robí dvoma spôsobmi. Prvý z nich, nazvime ho proporcionálny prístup, vychádza z predpokladu, že počet výjazdov ZZS do obce je úmerný počtu jej obyvateľov. Druhý prístup je tzv. empirický prístup, ktorý vychádza z minulého vývoja a historických údajov o výjazdoch ZZS do obce. Navrhujeme nový prístup, kde neistý počet volaní popíšeme trojuholníkovým fuzzy číslom získaným zo vzťahu medzi proporcionálnym a empirickým odhadom. V tejto kapitole budeme skúmať rôzny tvar funkcie príslušnosti fuzzy čísiel spolu s príslušným návrhom siete staníc ZZS.

Budeme používať označenie zavedené v predchádzajúcej kapitole: nech I je množina kandidátov, kde môžu byť stanice a J množina miest, kde sú pacienti. Obidve množiny sú tvorené mestami a obcami v danom regióne, ktoré sú reprezentované uzlami na dopravnej sieti. Množiny I a J môžu byť totožné. Z množiny I chceme vybrať p prvkov, kde budú umiestnené stanice. Ďalej označíme symbolom t_{ij} najkratší čas, za ktorý môže prísť ambulancia z miesta $i \in I$ na miesto $j \in J$. Koeficienty f_i a u_i označujú minimálny potrebný a maximálny povolený počet ambulancií v mieste $i \in I$. Premenné y_i udávajúce počet ambulancií umiestnených v mieste $i \in I$ musia spĺňať nasledujúce podmienky: $f_i \leq y_i \leq u_i$ a $\sum_{i \in I} y_i = p$.

V ideálnom prípade ide na výjazd ambulancia z najbližšej stanice k miestu tiesňového volania. Môže však nastať situácia, že v okamihu, kedy príde tiesňové volanie, môžu byť všetky

ambulancie z najbližšej stanice na výjazde k iným pacientom. Dispečer rieši túto situáciu tak, že na výjazd pošle voľnú ambulanciu, ktoré je v druhej, prípadne tretej najbližšej stanici. Pravdepodobnosť, že k -ta najbližšia ambulancia je prvá voľná, označíme q_k pre $k = 1, \dots, r$. Platí $q_1 > q_2 > \dots > q_r$.

Situáciu, že v jednom mieste i môže byť viac ambulancií, riešime tak, že zavedieme označenie $T_{ij}(y_i)$ pre postupnosť y_i prvkov s rovnakou hodnotou t_{ij} . Zavedieme operátor $\min_k(T)$, ktorý vráti k -tu najmenšiu hodnotu z postupnosti T čísiel.

Ak w_j je intenzita požiadaviek v obci j , potom priemerný dojazdový čas $\underline{t}(y)$ môžeme vypočítať podľa vzťahu (1).

$$\underline{t}(y) = \sum_{j \in J} w_j \sum_{k=1}^r q_k \min_k \{ \cup_{i \in I} T_{ij}(y_i) \} / \sum_{j \in J} w_j \quad (1)$$

Problém návrhu siete staníc ZZS, ktorý minimalizuje priemerný dojazdový čas ku všetkým pacientom daného regiónu, je popísaný vzťahom (2).

$$\min \{ \underline{t}(y) : y_i \in Z^+, f_i \leq y_i \leq u_i \text{ for } i \in I \text{ and } \sum_{i \in I} y_i = p \} \quad (2)$$

Empirický prístup k odhadu intenzity požiadaviek je založený na dlhodobej štatistike počtu výjazdov. Empirický odhad počtu požiadaviek v obci $j \in J$ označíme e_j . Z celkového počtu výjazdov D v danom regióne v časovom období T môžeme vypočítať celkovú intenzitu ako D/T . Ak b_j predstavuje počet obyvateľov v obci j , potom koeficient $\lambda = (D/T)/(\sum_{j \in J} b_j)$ vyjadruje počet výjazdov pripadajúci na jedného obyvateľa za jednotku času. Proporcionálny odhad intenzity v obci j vyjadríme ako λb_j . Potom v účelovej funkcii (1) môžeme za váhu obce w_j dosadiť e_j v empirickom prístupe alebo b_j v proporcionálnom prístupe. V prípade fuzzy prístupu predpokladáme, že neznáma hodnota w_j sa môže meniť v rámci intervalu ohraničeného ľavou a pravou hodnotou. Fuzzy číslo je definované takzvanou funkciou príslušnosti, ktorá má hodnotu 0 mimo intervalu, v rámci intervalu od jeho ľavej hodnoty po tzv. hlavnú hodnotu funkcia lineárne rastie od 0 po 1 a v intervale medzi hlavnou a pravou hodnotou funkcia lineárne klesá od 1 po 0.

Fuzzy číslo w_j popíšeme tromi konštantami w_j^1, w_j^2 a w_j^3 ktoré predstavujú ľavú, hlavnú a pravú hodnotu. Funkcia príslušnosti μ_{w_j} definovaná pre každé reálne číslo x je popísaná vzťahom (3).

$$\begin{aligned} \mu_{w_j}(x) &= 0 \text{ pre } x \leq w_j^1 \text{ alebo } x \geq w_j^3 \\ \mu_{w_j}(x) &= \frac{x-w_j^1}{w_j^2-w_j^1} \text{ pre } w_j^1 \leq x \leq w_j^2 \\ \mu_{w_j}(x) &= (w_j^3-x)/(w_j^3-w_j^2) \text{ pre } w_j^2 \leq x \leq w_j^3 \end{aligned} \quad (3)$$

Z α -cut princípu [18, 21] vyplýva, že lineárna kombinácia ľubovoľnej množiny trojuholníkových čísiel je tiež trojuholníkové číslo. Teda $\sum_{j \in J} w_j$ je trojuholníkové číslo, ktorého ľavá, hlavná a pravá hodnota sú $\sum_{j \in J} w_j^1, \sum_{j \in J} w_j^2$ a $\sum_{j \in J} w_j^3$. Potom deliteľ vo vzťahu (1) je tiež trojuholníkové fuzzy číslo, ale celý výraz, ktorý je podielom dvoch fuzzy čísiel, nie je trojuholníkové fuzzy číslo, pretože jeho funkcia príslušnosti nie je po častiach lineárna.

Napriek tomu fuzzy číslo $\mathbf{w}_j / \sum_{j \in J} \mathbf{w}_j$ môžeme aproximovať trojuholníkovým fuzzy číslom $\underline{\mathbf{w}}_j$, ktorého ľavá, hlavná a pravá hodnota sú $w_j^1 / \sum_{j \in J} w_j^3$, $w_j^2 / \sum_{j \in J} w_j^2$ a $w_j^3 / \sum_{j \in J} w_j^1$. Potom účelovú funkciu (1) môžeme prepísať do tvaru (4).

$$\underline{t}(\mathbf{y}) = \sum_{j \in J} \underline{\mathbf{w}}_j \sum_{k=1}^r q_k \min_k \{ \cup_{i \in I} T_{ij}(y_i) \} \quad (4)$$

Výraz (4) je tiež trojuholníkové fuzzy číslo $\underline{t}(\mathbf{y})$, ktorého ľavá, hlavná a pravá hodnota sú dané vzťahom (5).

$$\begin{aligned} \underline{t}^1(\mathbf{y}) &= \sum_{j \in J} \underline{\mathbf{w}}_j^1 \sum_{k=1}^r q_k \min_k \{ \cup_{i \in I} T_{ij}(y_i) \}, \\ \underline{t}^2(\mathbf{y}) &= \sum_{j \in J} \underline{\mathbf{w}}_j^2 \sum_{k=1}^r q_k \min_k \{ \cup_{i \in I} T_{ij}(y_i) \} \end{aligned} \quad (5)$$

$$\underline{t}^3(\mathbf{y}) = \sum_{j \in J} \underline{\mathbf{w}}_j^3 \sum_{k=1}^r q_k \min_k \{ \cup_{i \in I} T_{ij}(y_i) \}.$$

Štandardný prístup k optimalizačnej úlohe s fuzzy koeficientami spočíva v tom, že vytvoríme fuzzy množinu dobrých hodnôt účelovej funkcie a hľadáme také riešenie, ktorého fuzzy účelová funkcia patrí do fuzzy množiny dobrých hodnôt na maximálnej hladine dôveryhodnosti h . Fuzzy množinu dobrých hodnôt definujeme pomocou hodnôt účelových funkcií t_{min} a t_{max} optimistického a pesimistického riešenia úlohy (2). Optimistické riešenie získame pre najnižšie hodnoty $\underline{\mathbf{w}}_j$ a pesimistické pre najvyššie hodnoty $\underline{\mathbf{w}}_j$. Fuzzy množina dobrých hodnôt účelovej funkcie je definovaná funkciou príslušnosti (6).

$$\mu_{good}(x) = 0 \text{ for } x \geq t_{max}$$

$$\mu_{good}(x) = \frac{t_{max}-x}{t_{max}-t_{min}} \text{ for } t_{min} \leq x \leq t_{max} \quad (6)$$

$$\mu_{good}(x) = 1 \text{ for } x \leq t_{min}$$

Nasledujúca definícia hovorí, že trojuholníková fuzzy číslo t patrí do množiny dobrých hodnôt na hladine dôveryhodnosti h , ak existujú dve reálne čísla x a y také, že $x \leq y$ a $\mu_{good}(y) = \mu_t(x) = h$. Toto pravidlo vyjadruje podmienka (7).

$$(1-h)t^1(\mathbf{y}) + ht^2(\mathbf{y}) \leq x \leq (1-h)t_{max} + ht_{min} \quad (7)$$

Optimalizačný proces úlohy s neistými koeficientami $\underline{\mathbf{w}}_j$ pokračuje maximalizáciou hladiny dôveryhodnosti h za podmienok (2) a (7).

Keďže h je rozhodovacia premenná, podmienka (7) je nelineárna vzhľadom na premenné $y_i, i \in I$ a h .

Ak máme k dispozícii len nástroje diskretného lineárneho programovania, potom môžeme použiť Tanaka-Assaiov prístup. Tento prístup zafixuje h na nejakej hodnote a opakovane rieši vyššie definovaný problém minimalizáciou ľavej strany (7) pre rôzne hodnoty h . Výsledným riešením je prípustné riešenie pre maximálnu hodnotu h .

Experimentálne budeme vyšetrovať vzťah medzi odhadom intenzity požiadaviek a výsledným rozmiestnením ambulancií. Okrem empirického a proporcionálneho prístupu k

modelovaní dopytu použijeme aj fuzzy prístup. Výpočtové experimenty boli vykonané na testovacích príkladoch, ktoré vychádzajú z cestnej siete Žilinského samosprávneho kraja na Slovensku. V sieti je 315 miest a obcí. V každej z nich vznikajú požiadavky na obsluhu a zároveň tam môže byť sídlo ambulancie. Poznáme počet obyvateľov b_j v každej obci $j \in J$ a rovnako aj počet volaní e_j . Teda môžeme určiť váhy w_j pre proporcionálny aj empirický prístup. Na území kraja možno umiestniť $p = 36$ ambulancií. Z dôvodu veľkej heterogénosti vo vzťahu k váham bola vypočítaná celková váha W pripadajúca na jednu ambulanciu. Ak váha nejakej obce w_j prekročí hodnotu W , do vrcholu j sa priamo umiestni minimálny počet ambulancií $\lfloor w_j/W \rfloor$. Aby sme odvodili ďalšie testovacie úlohy, menili sme percentuálny podiel priamo priradených ambulancií pomocou zaokrúhľovacieho pravidla $\lfloor (w_j + \alpha)/W \rfloor$ [10, 16]. Percentuálny podiel začínal na hodnote 25%, čo zodpovedá podielu $\lfloor w_j/W \rfloor$. Nasledujúce hodnoty boli 30%, 35%, atď, až do 95%.

Pravdepodobnosti q_k pre $k = 1, \dots, r = 5$ boli na základe simulačných experimentov s reálnym systémom ZZS na Slovensku nastavené na hodnoty $q_1 = 0.798474$, $q_2 = 0.160913$, $q_3 = 0.0324282$, $q_4 = 0.00653513$ a $q_5 = 0.00164939$ [11]. Použili sme radiálnu formuláciu lokačnej úlohy (2) [1, 5, 8, 9, 15], ktorú sme riešili pomocou komerčného optimalizačného softvéru FICO Xpress 8.3 (64-bit, release 2017). Výpočtové experimenty sme vykonali na osobnom počítači Intel® Core™ i7 s 2.8 GHz procesorom a 16 GB RAM.

Výpočtová štúdia začala vyšetrením vzťahu medzi hodnotami b_j a e_j . Zistili sme, že korelačný koeficient je 0.99, čo indikuje veľmi silnú závislosť. Lineárna regresia odhalila nasledujúcu závislosť: $E_j^2 = ab_j + c^2$, kde $a = 0.0636$ a $c^2 = -17.0273$. Ďalej sme určili najužšiu oblasť v priestore E^2 pomocou dvoch paralelných ohraničujúcich priamok $E_j^1 = ab_j + c^1$ a $E_j^3 = ab_j + c^3$ takých, že $E_j^1 \leq e_j \leq E_j^3$ pre $j \in J$.

Prvým spôsobom ako popísať neistú hodnotu w_j je zostrojiť trojuholníkové fuzzy číslo w_j pre každé $j \in J$ také, že $w_j^1 = E_j^1$, $w_j^2 = E_j^2$ a $w_j^3 = E_j^3$. Druhý spôsob fuzzyfikácie spočíva v nahradení E_j^2 empirickou intenzitou e_j .

Vykonali sme experimenty pre každú zo štyroch možností nastavenia váh w_j a pre všetky testovacie príklady. Variant označený písmenom B považuje za váhu obce w_j počet jej obyvateľov b_j . Variant E považuje za váhy empirické hodnoty e_j . Variant F predstavuje fuzzy prístup, kde váha w_j je trojuholníkové fuzzy číslo dané trojicou $[E_j^1, E_j^2, E_j^3]$ a variant G tiež obsahuje fuzzy váhy, pričom w_j je dané trojicou $[E_j^1, e_j, E_j^3]$. Symbol $SolX$ označuje optimálne riešenie úlohy vo variante X , kde X má hodnotu B , E , F alebo G . Každé optimálne riešenie dané vektorom y rozhodovacích premenných bolo vyhodnotené pomocou účelovej funkcie pre všetky ostatné varianty. Teda napríklad $FceB_SolG$ znamená, že riešenie $SolG$ vypočítané s účelovou funkciou vo variante G bolo dosadené do účelovej funkcie s koeficientami pre variant B . Optimálne riešenie matematického modelu môžeme porovnať s aktuálnym rozmiestnením ambulancií, ktoré je označené ako $SolC$.

Hoci sme výpočtové experimenty vykonali pre všetkých 15 príkladov, kvôli ich podobnosti uvádzame v nasledujúcich tabuľkách len výsledky pre percentuálne podiely 25%, 35%, 45%, 55% a 65% počtu všetkých ambulancií, ktoré boli umiestnené priamo pred samotnou optimalizáciou.

Table 1 Výsledky pre pevné umiestnenie 25% ambulancií

CT [s]		FceB	FceE	FceF	FceG	avg	var
3.375	SolB	6.508	6.436	5.995	6.021	6.240	0.055
5.171	SolE	6.543	6.411	6.015	5.997	6.241	0.058
56.043	SolF	6.547	6.531	5.995	6.109	6.295	0.061
74.621	SolG	6.543	6.411	6.015	5.997	6.241	0.058
	avg	6.535	6.447	6.005	6.031		
	var	0.0002	0.0024	0.0001	0.0021		
	max	6.547	6.531	6.015	6.109		
	min	6.508	6.411	5.995	5.997		
	dif	0.039	0.120	0.020	0.112		
	SolC	6.975	6.822	6.397	6.379		

Table 2 Výsledky pre pevné umiestnenie 35% ambulancií

CT [s]		FceB	FceE	FceF	FceG	avg	var
4.104	SolB	6.533	6.443	6.002	6.027	6.251	0.057
5.469	SolE	6.583	6.423	6.063	6.008	6.269	0.058
55.389	SolF	6.567	6.540	6.006	6.117	6.308	0.062
80.320	SolG	6.563	6.420	6.026	6.005	6.254	0.059
	avg	6.562	6.457	6.024	6.040		
	var	0.0003	0.0024	0.0006	0.0021		
	max	6.583	6.540	6.063	6.117		
	min	6.533	6.420	6.002	6.005		
	dif	0.050	0.120	0.061	0.112		
	SolC	6.975	6.822	6.397	6.379		

Table 3 Výsledky pre pevné umiestnenie 45% ambulancií

CT [s]		FceB	FceE	FceF	FceG	avg	var
2.625	SolB	6.570	6.492	6.021	6.072	6.289	0.060
2.797	SolE	6.610	6.428	6.064	6.012	6.278	0.062
37.685	SolF	6.595	6.542	6.014	6.118	6.317	0.065
49.278	SolG	6.610	6.428	6.064	6.012	6.278	0.062
	avg	6.596	6.472	6.041	6.053		
	var	0.0003	0.0023	0.0005	0.0020		
	max	6.610	6.542	6.064	6.118		
	min	6.570	6.428	6.014	6.012		
	dif	0.040	0.114	0.050	0.106		
	SolC	6.975	6.822	6.397	6.379		

Table 4 Výsledky pre pevné umiestnenie 55% ambulancií

CT [s]		FceB	FceE	FceF	FceG	avg	var
3.069	SolB	6.693	6.606	6.137	6.178	6.404	0.062
3.357	SolE	6.732	6.542	6.180	6.118	6.393	0.065
28.935	SolF	6.703	6.657	6.120	6.226	6.426	0.066

38.518	SolG	6.732	6.542	6.180	6.118	6.393	0.065
	avg	6.715	6.587	6.154	6.160		
	var	0.0003	0.0023	0.0007	0.0020		
	max	6.732	6.657	6.180	6.226		
	min	6.693	6.542	6.120	6.118		
	dif	0.040	0.115	0.060	0.107		
	SolC	6.975	6.822	6.397	6.379		

Table 5 Výsledky pre pevné umiestnenie 65% ambulancií

CT [s]		FceB	FceE	FceF	FceG	avg	var
2.672	SolB	6.767	6.594	6.215	6.167	6.436	0.064
2.672	SolE	6.778	6.590	6.230	6.163	6.440	0.064
27.342	SolF	6.781	6.686	6.197	6.253	6.479	0.066
31.983	SolG	6.778	6.590	6.230	6.163	6.440	0.064
	avg	6.776	6.615	6.218	6.187		
	var	0.0000	0.0017	0.0002	0.0015		
	max	6.781	6.686	6.230	6.253		
	min	6.767	6.590	6.197	6.163		
	dif	0.014	0.096	0.033	0.089		
	SolC	6.975	6.822	6.397	6.379		

Prvé štyri riadky každej tabuľky udávajú hodnoty účelových funkcií pre optimálne riešenia *SolB*, *SolE*, *SolF* a *SolG*. Stĺpec označený *CT[s]* udáva výpočtový čas optimalizácie v sekundách. Stĺpce *avg* a *var* udávajú priemer a smerodajnú odchýlku účelových funkcií pre riešenie v danom riadku. Môžeme vidieť, že priemery sú takmer rovnaké a smerodajná odchýlka je veľmi malá. Z tohto pohľadu teda nemožno určiť najlepší variant.

Na rozdiel od riadkových priemerov a smerodajných odchýlok, priemery a odchýlky v stĺpcoch naznačujú rozdiely medzi variantami. Keď do rovnakej účelovej funkcie dosadíme všetky optimálne riešenia, potom sú ich hodnoty veľmi podobné, smerodajná odchýlka je veľmi malá a rádovo menšia než odchýlky v riadkoch. Ak vezmeme do úvahy priemery jednotlivých stĺpcov, fuzzy varianty F a G dominujú nad variantami s proporcionálnymi (B), resp. empirickými (E) váhami.