

# PRÍSPEVOK KU GENEROVANIU TEXTOVÝCH POPISOV BITOVÝCH KALENDÁROV

## CONTRIBUTION TO GENERATING TEXT DESCRIPTION OF CALENDAR FROM ITS BIT MAP DESCRIPTION

Hynek Bachratý, Emil Kršák, Lubomír Sadloň, Marek Tavač<sup>1</sup>

---

*Anotace: Článok popisuje základnú schému práce algoritmu, ktorý z bitovej mapy kalendára vygeneruje jeho textový popis. Uvádzame ukážky práce algoritmu a pozitívne výsledky jeho testovania na kalendároch jazdy vlakov z databázy MERITS.*

*Klíčová slova: kalendár, bitová mapa kalendára, textový popis kalendára*

*Summary: The article describes the basic work scheme of an algorithm which generates a calendar text description from its bit map. We present samples of the algorithm work, and positive results of the algorithm being tested on the trains journey calendar from the MERITS database.*

*Key words: calendar, bit map of calendar, text description of calendar*

### 1. ÚVOD

Kalendárne údaje sú významnou súčasťou informácií o doprave a ich spracovanie a zobrazenie je dôležitou súčasťou mnohých informačných systémov. Najčastejšie popisujú dni jazdy dopravných prostriedkov alebo dni platnosti niektorých ďalších charakteristík dopravy. V ďalšom budeme uvažovať a hovoriť o kalendároch popisujúcich dni jazdy vlaku. Vytvorené a tu popisované algoritmy pre generovanie textu sú ale z veľkej časti univerzálne a zmenou používaných textov je možné výstupy prispôbiť aj pre iné situácie.

### 2. FORMULÁCIA PROBLÉMU

Kalendárne údaje sú najčastejšie zadávané v dvoch formách: v bitových mapách a pomocou textových popisov.

Termínom *bitová mapa* rozumieme súhrnnú informáciu, v ktorej je pre každý deň určitého obdobia zadané, či vlak v daný deň ide alebo nejde. Kalendár si teda možno predstaviť a v princípe tak býva aj dátovo reprezentovaný ako postupnosť jednotiek (ide)

---

<sup>1</sup> RNDr. Hynek Bachratý, Ph.D., Žilinská univerzita, Fakulta riadenia a informatiky, Katedra softvérových technológií, tel.: +421 415 134 127, e-mail: [Hynek.Bachraty@fri.utc.sk](mailto:Hynek.Bachraty@fri.utc.sk)  
Ing. Emil Kršák, Ph.D., Žilinská univerzita, Fakulta riadenia a informatiky, Katedra softvérových technológií, tel.: +421 415 134 100, e-mail: [Emil.Krsak@fri.utc.sk](mailto:Emil.Krsak@fri.utc.sk)  
Mgr. Ing. Lubomír Sadloň, Ph.D., Žilinská univerzita, Fakulta riadenia a informatiky, Katedra softvérových technológií, tel.: +421 415 134 100, e-mail: [Lubomir.Sadlon@fri.utc.sk](mailto:Lubomir.Sadlon@fri.utc.sk)  
Ing. Marek Tavač, Žilinská univerzita, Fakulta riadenia a informatiky, Katedra softvérových technológií, tel.: +421 415 134 130, e-mail: [Marek.Tavac@fri.utc.sk](mailto:Marek.Tavac@fri.utc.sk)

a núl (nejde) zodpovedajúcich jednotlivým dňom. Súčasťou informácie o kalendári (alebo o celej skupine kalendárov) je začiatok a koniec obdobia platnosti. Tým je najčastejšie zabezpečená aj väzba týchto dát na skutočný kalendár.

Bitovou mapou zadaný kalendár je optimálny z informatického hľadiska. Pomocou nej je možné ľahko editovať alebo zisťovať dni jazdy vlaku, porovnávať a kombinovať dáta viac kalendárov, zisťovať ich vlastnosti atď. Zobrazíť tieto údaje je možné v podstate len graficky, čo sa rozumne dá len na (dostatočne veľkom) monitore. Užívateľ musí mať prostriedky a aj čas takýto spôsob zobrazenia dátumovej informácie využívať. Táto forma zadania kalendára je naopak nevhodná pre (hromadné) tlačové výstupy, zobrazenie na malých obrazovkách (napr. PDA zariadenia), pre iný ako grafický prenos informácie o jazde vlakov atď.

Bitovou mapou zadané kalendáre sú teda v zásade vhodné pre prácu informačných systémov a špecializovaných užívateľov, ktorý dátumové údaje ako súčasť svojej práce zadávajú, menia, analyzujú atď. Nevhodné sú naopak pre hromadné využívanie, pre bežnú komunikáciu medzi ľuďmi a pod.

*Textový popis* kalendára je najčastejšie využívaný v dátumových poznámkach v rôznych typoch cestovných poriadkov a ďalších pomôcok. Vlastnosti tohto popisu dátumovej informácie sú v podstate opačné ako pri bitovej mape. Obsah kalendára je vyjadrený v relatívne krátkom texte. Hodí sa preto do tlačových výstupov alebo na malé zobrazovacie zariadenia. Platnosť kalendára je možné komunikovať aj ústne. Vzhľadom na našu skúsenosť s používaním tejto formy popisu dátumov vieme z textového popisu ľahko získať niektoré typy informácií, napríklad či vlak pôjde zajtra, budúci pondelok alebo na Vianoce. Problém by sme ale mali, ak by sme z textového popisu mali určiť počet dní jazdy vlaku v sledovanom období, zistiť ktorý z dvoch vlakov chodí častejšie alebo v ktoré dni chodia oba vlaky.

Problém, ktorý reálne existuje a ktorý sme sa pokúsili vyriešiť, spočíva v algoritmickej prepojení týchto dvoch spôsobov zadania dátumovej informácie. Pre spôsoby práce používané v oblasti dopravy zväčša platí, že najskôr sú vytvorené (alebo inak získané) bitové kalendáre, ktoré treba následne vyjadriť v textovej forme. V článku popíšeme pre tieto účely vytvorený algoritmus generovania textových popisov z bitovej mapy kalendára a prvé výsledky z jeho testovania.

### **3. ZÁKLADNÉ PREDPOKLADY A VLASTNOSTI RIEŠENIA**

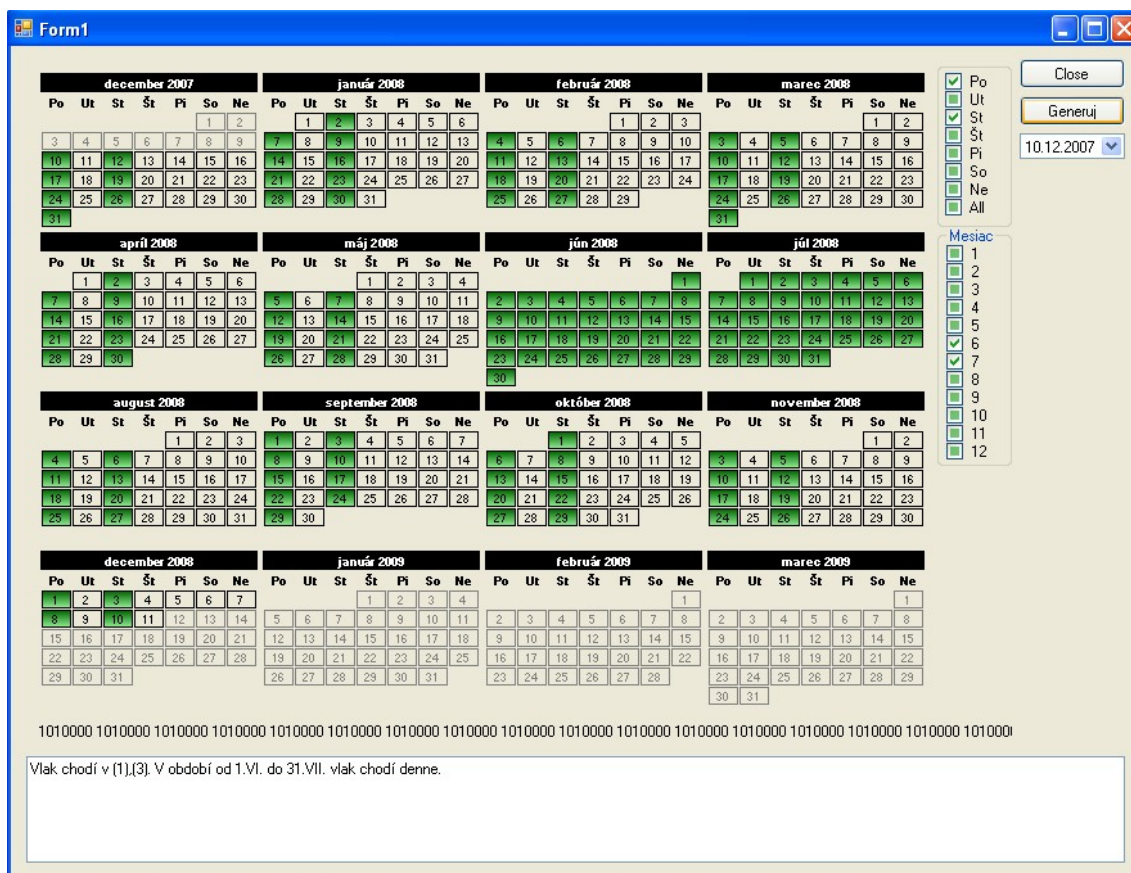
Základnou úlohou je vygenerovanie textu zodpovedajúceho danej bitovej mape kalendára. Pre bitovú mapu predpokladáme zadaný aj deň začiatku a deň konca obdobia platnosti kalendára. Tým je zároveň určená dĺžka obdobia platnosti. Pokiaľ uvažujeme o viac kalendároch, začiatok a koniec obdobia platnosti predpokladáme pre všetky za rovnaký. V prípade nedodržania tejto podmienky ju vieme algoritmicke formálne zabezpečiť predĺžením obdobia platnosti všetkých kalendárov na maximálny uvažovaný rozsah. Nami navrhnutý spôsob generovania textov zjednotenie obdobia platnosti jeho rozšírením zvládne bez väčších problémov.

Ďalej vychádzame z predpokladu, že problém riešime pre kalendáre používané najmä pre verejnú osobnú dopravu. Slúžia teda veľkej skupine obyvateľstva ako dôležitá prístupová informácia k určitej službe. Táto služba predpokladá určitý stupeň pravidelnosti,

jednoduchosti a prehľadnosti týchto kalendárov. Angličania by povedali, že dni jazdy tvoria určitú vzorku, „pattern“. V opačnom prípade by neprehľadná dátumová informácia túto službu znehodnocovala. Hľadanie a nájdenie tejto pravidelnosti resp. vzorky tvorí dôležitú časť riešenia problému.

Algoritmus je samozrejme schopný spracovať ľubovoľný kalendár aj z inej oblasti, napr. kalendár nákladnej alebo špeciálnej dopravy, servisný kalendár atď. Tu je ale možné očakávať nižšiu efektivitu a horšiu úspešnosť algoritmu. Úzka skupina užívateľov takýchto špeciálnych kalendárov ale umožňuje predpokladať využitie špecifických možností pre zobrazenie kalendára, napr. grafické zobrazenie, rozsiahly text, možnosť dopytovania sa na jednotlivé dni jazdy atď.

Textové popisy kalendárov, ktoré budú výstupmi algoritmu, musia tiež spĺňať určité zásady. Text musí byť samozrejme jasný, jednoznačný, štruktúrovaný a typizovaný. Pokiaľ má popísať širšiu typovú paletu kalendárov, bude nutné použiť viac typov textu. Ich počet by ale nemal byť príliš veľký. Rozpoznanie typu kalendára a následné priradenie a spresnenie príslušného typu textu je ďalšou kľúčovou úlohou riešenia. Text musí mať určitú štruktúru, najlepšie spoločnú pre všetky typy a opäť nie nadmieru komplikovanú. Mal by obsahovať „vzorkovú“ časť, kde dôjde k najvýraznejšiemu komprimovaniu informácie popisom nájdenej pravidelnosti kalendára. Túto ale nebude možné vždy dosahovať absolútne presne a bude potrebné uviesť výnimky z tohto pravidla. Obe informácie je potrebné vyjadriť kompaktne a nadväzujúco, aby sme užívateľa neuviedli do omylu.



Zdroj: Autori

Obr.1 - Ukážka vývojového prostredia pre testovanie algoritmu

Dôležitými kritériami kvality vytvoreného algoritmu bude čo najmenšia alebo akceptovateľná dĺžka generovaných textov a čo najvyšší relatívny počet kalendárov, pre ktoré bude algoritmus schopný text vygenerovať.

#### 4. HĽADANIE A POPIS PRAVIDELNOSTI KALENDÁRA

Základným princípom využitým pre vyhľadávanie vzorky platnosti kalendára je pravidelnosť založená na týždňovom rytme nášho „dátumového“ života. Základné vzorky preto zodpovedajú opakovaniu jednotlivých dní týždňa (pondelok až nedeľa). Zadaná bitová mapa sa porovnáva so všetkými 128 kombináciami týchto siedmich základných kalendárov. Hovoríme o tzv. vzorových kalendároch. Sedemdňový cyklus je v prípade potreby možné modifikovať na iný počet dní, aj keď iný ako sedemdňový kalendár sa nepoužíva. Táto možnosť môže byť využitá pri spracovávaní kalendárov viazucich sa na iné ako týždňové cykly údržby rôznych zariadení a podobne.

Pre určenie najvhodnejšieho vzorového kalendára sa ako kritérium používa počet výnimiek, teda dní zadaného kalendára ktorých platnosť nezodpovedá testovanému vzorovému kalendáru. Môže ísť o tzv. pozitívne výnimky (vlak ide aj v dni nad rámec testovaného vzoru) a negatívne výnimky (vlak nejde v dňoch testovaného vzoru). V optimálnom prípade sa testovaný vzorový kalendár presne zhoduje so zadaným kalendárom a počet výnimiek je nula. V tomto prípade je možné testovanie ukončiť. Naopak jedným z parametrov algoritmu je hodnota maximálneho počtu akceptovateľných výnimiek. Pokiaľ je prekročená, testovaný vzorový kalendár nepovažujeme za vyhovujúci. Je možné, že za vhodný nebudeme považovať žiadny vzorový kalendár a hľadanie pravidelnosti nebude úspešné.

Pre testovanie kalendárov viazaných na určitý štát je vhodné doplniť ďalšie dva vzorové kalendáre. Ide o tzv. sviatky (nedele a štátne sviatky) a pracovné dni (pondelok až piatok bez sviatkov), ktoré sa zvyknú označovať + a X. Získavame tak miesto 7 až 9 základných kalendárov, ktoré navzájom kombinujeme. Kalendár + pritom ale nemá zmysel kombinovať s kalendármi obsahujúcimi nedeľu a kalendár X s kalendármi obsahujúcimi len dni pondelok až piatok.

Popis nájdenej vzorky predstavuje hlavnú zložku generovaného textu. Formu označenia dní týždňa je možné zvoliť, pri konkrétnej implementácii algoritmu sa využívajú špeciálne fonty. Štandardne má text tvar typu

„**Vlak chodí v Po, Št, Pi**“ alebo

„**Vlak chodí v (1), (4), (5)**“.

Pre niektoré špeciálne prípady sa používajú zjednodušené a zaužívané texty, ako

„**Vlak chodí denne**“,

„**Vlak nechodí**“,

„**Vlak chodí v Po až Pi**“ alebo

„**Vlak chodí v So a Ne**“.

Voliteľne je možné využívať aj tzv. negatívne poznámky. Pokiaľ je vymenovanie dní týždňa príliš dlhé, je možné použiť modifikovaný text

„**Vlak chodí denne okrem (3),(4)**“ alebo

„**Vlak nechodí v (3),(4)**“

namiesto textu

„**Vlak chodí v (1),(2),(5),(6),(7)**“

Pri druhom uvedenom texte je informácia, že vlak chodí vo všetky dni okrem stredy a štvrtka len implicitná, text je preto síce kratší, ale nepresnejší. Malá predbežná anketa ukázala, že vymenovanie troch dní v negatívnej forme už nie je užívateľmi akceptované.

Za popisom pravidelnosti nasleduje vymenovanie výnimiek v nasledujúcej forme:

„**Vlak chodí v (5),(6) okrem 20.XII., 14.III. a 18.IV. Vlak ide 15.I. a 14.V.**“

alebo

„**Vlak chodí denne okrem (5) a nejde 20.XII., 14.III. a 18.IV. Vlak ide 15.I. a 14.V.**“

Zvolili sme princíp, že negatívne výnimky sú uvádzané skôr ako pozitívne hneď za popisom pravidelnosti. Vychádzame pritom zo skúsenosti, že užívateľ je náchylný po získaní pozitívnej informácie skončiť analyzovanie textového popisu kalendára. Mohol by tak mylne predpokladať, že vlak v daný deň chodí, aj keď platí opak. Preto sú negatívne výnimky uvedené čo najskôr a v jednom súvetí s popisom pravidelnosti. Naopak v prípade neúspechu má naša psychika tendenciu pokračovať v skúmaní, pozitívne výnimky preto môžu nasledovať neskôr v ďalšej vete.

Pri vymenovaní výnimiek sa výpis jednotlivých dní riadi viacerými pravidlami. Súvislé úseky za sebou nasledujúcich dní sú uvádzané len prvým a posledným dňom. Označenie mesiaca sa používa len pri jeho zmene a označenie roku len vtedy, ak sa daný deň v období platnosti nachádza viac krát.

„**Vlak ide 10.XII.2007-12.XII.,19.-23.II., 13.,14.V., 29.VI.-3.VII.,15.,16.X.,  
9. 11.XII.2008.**“

Uvedený príklad ilustruje zároveň jednu z voliteľných možností vygenerovania textu v prípade nenájdenia vhodnej pravidelnosti a teda neúspechu algoritmu. Ide tu o vymenovanie všetkých platných dní kalendára. Takýto text ale býva často dlhý a nie vždy využiteľný. Ďalšou voliteľnou možnosťou je preto aj generovanie prázdneho reťazca alebo upozorňujúceho textu, napr. „**Textový popis kalendára nie je k dispozícii**“.

## 5. HĽADANIE A POPIS TYPU KALENDÁRA

Vyššie popísaný spôsob vyhľadávania pravidelností by prinášal dobré výsledky len v prípade, že vlak chodí s určitou pravidelnosťou v celom období platnosti kalendárov. Presne také sú totiž aj testované vzorové kalendáre. Pokiaľ by ale vlak napríklad v pondelky chodil (alebo nechodil) len v období prázdnin, vysoký počet výnimiek by túto pravidelnosť v celoročnom horizonte neobjavil resp. neakceptoval. Táto skutočnosť nás viedla k myšlienke vytipovať a následne vyhľadávať v zadanej bitovej mape určité typy kalendárov. Typy by mali hlavne identifikovať obdobie, v ktorom vlak jazdí a len v ňom sa bude testovať a

vyhľadávať pravidelnosť. Pre každý typ kalendára bude následne generovaný aj príslušný typový text, dopĺňajúci informáciu o nájdenej pravidelnosti.

Počet typov sme sa rozhodli udržať pre prehľadnosť nízky, zároveň ale musíme byť schopní popísať s ich pomocou čo najväčšie množstvo bitových máp. Z dôvodu prehľadnosti sme sa rozhodli aj odolať možnosti využitia rekurzívneho vyhľadávania typov, ktoré sa dá ľahko programovo realizovať. V súčasnosti algoritmus využíva 5 nižšie popísaných typov. Zadaná bitová mapa je testovaná na príslušnosť k jednotlivým typom, pričom môže spadať aj pod viaceré z nich. Ak sa tak stane, vyberie sa typ s najmenším počtom výnimiek od nájdenej pravidelnosti. V prípade rovnakého počtu sa využíva priorita typov, ktorá zodpovedá poradiu v akom ich uvádzame.

Štandardným typom je *celoročný kalendár*. V tomto prípade sa generujú texty uvádzané v predchádzajúcej časti.

Nasledujúcim typom je *kalendár s obdobím „chodí“*. Zodpovedajú mu kalendáre, kde pred a(lebo) za obdobím dní jazdy vlaku, teda na začiatku a(lebo) konci obdobia platnosti vlak nechodí vôbec. Minimálny rozsah obdobia, v ktorom vlak jazdí (nie denne, ale s určitou následne testovanou pravidelnosťou) sa zadáva ako parameter algoritmu. Pre tento typ kalendára sa generujú nasledovné texty:

**„Vlak chodí len v období od 1.XII. do 31.VIII. v Po, St, Št, Pi.“**, **„Vlak chodí len v období od 1.XII. do 31.VIII. denne okrem (2), (6).“** alebo s výnimkami **„Vlak chodí len v období od 1.I. do 31.VIII. denne okrem Ut, So a nejde 9.IV., 14.V. a 27.V.. Vlak ide 14.IV., 14.VII. a 15.VIII.“**

Rozvinutím tohto typu je *kalendár s dvomi obdobiami „chodí“*. V takýchto kalendároch sa miesto jedného nachádzajú dve obdobia kedy vlak chodí, oddelené obdobím, kedy vlak nechodí vôbec. Generované texty vyzerajú takto:

**„Vlak chodí v období od 4.IV. do 31.V. a od 4.VII. do 31.VIII. v (1),(5),(6),(7).“**, **„Vlak chodí v období od 2.V. do 30.VI. a od 1.IX. do 31.X. denne okrem (4).“**, **„Vlak chodí v období od 30.I. do 30.IV. a od 1.VII. do 30.IX. denne okrem 31.I. a 15.VII.“**

Ďalší typ je *kalendár s obdobím „nechodí“*. V tomto prípade sa v bitovej mape kalendára hľadá vo vnútri obdobia platnosti jedno obdobie, v ktorom vlak nechodí vôbec. Minimálny rozsah obdobia „nechodí“ je opäť daný voliteľným parametrom. Generujú sa nasledovné texty:

**„Vlak mimo obdobia od 1.IV. do 2.VIII. chodí v Po až Pi.“**, **„Vlak mimo obdobia od 1.V. do 31.VII. chodí denne okrem Po, Ne.“** alebo **„Vlak mimo obdobia od 1.IV. do 2.VIII. chodí denne okrem So, Ne a nejde 19.II., 11.III. a 19.VIII.. Vlak ide 26.X., 8.XI. a 22.XI.“**

Posledný typ je *kalendár s obdobím „chodí denne“*. Na rozdiel od predchádzajúceho typu bitová mapa musí obsahovať dostatočne dlhé obdobie, v ktorom vlak chodí denne. Toto obdobie je uvedené v samostatnej vete na konci vygenerovaného textu:

**„Vlak chodí v (1),(3). V období od 1.VI. do 31.VII. vlak chodí denne.“**

**„Vlak chodí v (2),(3),(6) okrem 23.I. a 16.IX. Vlak ide 30.X. a 13.XI. V období od 1.VII. do 31.VIII. vlak chodí denne.“**

Pri testovaní hľadania typu kalendára sme narazili na problém špecifických mimoriadnych dní jazdy vlaku, ktoré narúšali súvislé obdobia v ktorých vlak nejazdí a znemožňovali tým jeho priradenie ku inak jasnému typu. (Ide napríklad o trasy „prázdninových“ vlakov, ktoré sú ako posilové využité pred vianočnými alebo veľkonočnými sviatkami.) Tieto dni sme nazvali *izolované dni jazdy* a identifikujeme ich ako dni, pred a za ktorými existuje dostatočne dlhé (parametrami algoritmu určené) obdobie v ktorom vlak nejazdí. Tieto dni sú identifikované a dočasne odstránené z bitovej mapy ešte pred začiatkom určovania typu kalendára. Ich maximálny počet je jedným z parametrov algoritmu. Izolované dni sú vymenované v špeciálnej doplnkovej vete „**Vlak ide tiež...**“ na konci generovaného textu:

„**Vlak chodí v období od 4.IV. do 28.VII. v (1),(5),(6),(7). Vlak ide tiež 6.II. a 18.IX.**“

## 6. JAZYKOVÉ OKIENKO

Obsah generovaných textov závisí samozrejme od kvality algoritmov, ktoré vyhľadajú príslušný typ kalendára, jeho pravidelnosť a výnimky z nej. Dôležitá je ale aj ich jazyková kvalita. Niektoré zásady sme už uviedli vyššie, teraz spomenieme niektoré jazykové špeciality týkajúce sa textového popisu dní jazdy vlaku.

Zo školy si ešte pamätáme, že vety obsahujú prísudok, teda pre obsah vety určujúce sloveso. V slovenčine sa nám ako optimálne javí používanie tvarov slovesa *chodiť* a slovesa *ísť*. Prvé používame pre popis pravidelnosti čo súvisí s tým, že v slovenčine sa toto slovo používa pre popis pravidelne sa opakujúcich dejov. Naopak slovo *ísť* sa odporúča pre vyjadrenie deja jednorazového, používame ho preto pri popise výnimiek. Pokiaľ kalendár nebude popisovať dni jazdy vlaku, odporúčame zvoliť inú dvojicu príbuzných, ale minimálne dokonavým a nedokonavým vidom odlišných slovies. Druhou dôležitou časťou vety je podmet. V tu uvedených ukážkach používame slovo *vlak*. Je ho ľahké nahradiť univerzálnejším slovom *spoj* a vytvorili sme aj verziu textov so zamlčaným podmetom. Takéto texty je možné univerzálne použiť napríklad pre popis kalendárov jazdy vlakov a zároveň priamych vozňov.

Upozorním ešte na vedomé používanie výrazných spojok a častíc *okrem*, *mimo*, *tiež*, ktorých význam si užívateľ má a musí dobre uvedomiť. V prípade použitia negatívnej poznámky sa vyhýbame dvojitému záporu. Negatívne výnimky preto neuvádzame druhým použitím slova „okrem“, ale používame text: „**Vlak chodí denne okrem (5) a nejde 19.V.**“. Všeobecne sa snažíme generovať len jednoduché vety alebo jasne štruktúrované súvetia. V príspevku uvádzaných príkladoch sa využíva (mierne) dlhší variant formulácií, je možné zvoliť aj opačnú, čo najstručnejšiu a kratšiu formu textov.

Treba si uvedomiť, že jazyková formulácia je nadstavbou nad spomenutými algoritmami. Mala by prezentovať (a nie znehodnotiť) ich výsledky, čo je podľa nášho presvedčenia nie jednoduchá, ale zvládnuteľná úloha. Umožní nám využiť algoritmické jadro aj pre textové popisy bitových kalendárov z inej ako dopravnej oblasti a prevedenie algoritmu do iných jazykov.

## 7. TESTY A VYUŽITIE

Algoritmus sme pilotne overovali na sade kalendárov medzinárodnej databázy osobných vlakov MERTIS. Okrem celej databázy sme zvlášť testovali sady kalendárov vlakov ŽSSK a ČD.

Generovanie textov je riadené viacerými už spomenutými parametrami. Z dôležitých spomeniem, že pre testovanie bol maximálny počet výnimiek stanovený na 15 a maximálny počet izolovaných dní na 5. Minimálna dĺžka obdobia pri určovaní typu kalendára bola 14 dní. V prípade neúspechu algoritmu bol generovaný text, ktorý vymenoval všetky platné dni kalendára vo forme pozitívnych výnimiek.

Tab.1 - Kalendáre vlakov ŽSSK. Spolu 791 kalendárov, 10 (1,26%) s izolovanými dňami

Typ kalendára	Počet/% kal.	Priem. počet výnimiek	Priem. počet znakov textu	Min./max. počet znakov
Neúspech	45 / 5,69 %	171	245	42 / 538
Celoročný	396 / 50,06 %	5	61	13 / 144
Obd. „chodí“	219 / 27,69 %	0	61	45 / 121
Obd. „chodí+ chodí“	26 / 3,29 %	2	92	66 / 131
Obd. „nechodí“	83 / 10,49 %	4	87	48 / 154
Obd. „chodí denne“	22 / 2,78 %	5	101	67 / 164

Tab.2 - Kalendáre vlakov ČD. Spolu 1352 kalendárov, 14 (1,04%) s izolovanými dňami

Typ kalendára	Počet/% kal.	Priem. počet výnimiek	Priem. počet znakov textu	Min./max. počet znakov
Neúspech	155 / 11,46 %	183	295	45 / 585
Celoročný	631 / 46,67 %	5	59	13 / 146
Obd. „chodí“	341 / 25,22 %	1	68	45 / 155
Obd. „chodí+ chodí“	65 / 4,81 %	2	92	65 / 159
Obd. „nechodí“	96 / 7,10 %	3	85	48 / 142
Obd. „chodí denne“	64 / 4,73 %	6	116	63 / 174

Tab.3 - Kalendáre [plnej databázy MERITS. Spolu 21 389, z toho 642 kalendárov (3,00%) s izolovanými dňami

Typ kalendára	Počet/% kal.	Priem. počet výnimiek	Priem. počet znakov textu	Min./max. počet znakov
Neúspech	2282 / 10,67 %	161	223	36 / 770



Celoročný	7157 / 33,46 %	6	61	13 / 158
Obd. „chodí“	7240 / 33,85 %	3	80	45 / 176
Obd. „chodí+ chodí“	3606 / 16,86 %	2	92	62 / 196
Obd. „nechodí“	794 / 3,71 %	5	95	46 / 174
Obd. „chodí denne“	310 / 1,45 %	7	113	62 / 204

Pilotné testovanie prinieslo dobré výsledky. Pre kalendáre ŽSSK bol v 44 prípadoch použitý aj kalendár + a v 89 prípadoch kalendár X. Vzhľadom na príbuznosť našich štátov boli približne v rovnakých počtoch využité aj pri kalendároch ČD. Percento kalendárov, v ktorých sa využili izolované dni v zásade priamo zvyšuje úspešnosť algoritmu.

Rôznymi nastaveniami parametrov je možné ďalej zvyšovať úspešnosť algoritmu a kvalitu generovaných textov. Podrobné testovanie týchto zákonitostí pripravujeme. Jednoduchou cestou je zvyšovanie počtu výnimiek, čím sa ale spolu so zvýšením úspešnosti zvýši aj priemerná dĺžka generovaných textov. Ďalšou možnosťou je zníženie počtu dní potrebných pre akceptáciu obdobia pri určovaní typu kalendára. Dôležitá bude aj podrobná analýza bitových máp, pri ktorých algoritmus nebol úspešný. Tá môže ukázať potrebu zavedenia ďalších typov kalendárov.

Zjednodušená verzia algoritmu je momentálne využívaná na generovanie textov IS pre vyhľadávanie spojení ŽSSK, ktorý sme tiež vytvorili na našom pracovisku. Výsledky algoritmu sú uspokojivé a z jeho nasadenia získavame aj cenné pripomienky a podnety. V ďalšom predpokladáme nasadenie plnohodnotného algoritmu, a to aj v oblasti generovania textových popisov bitových máp pri tvorbe ZCP a KCP (SJŘ a KJŘ) v systémoch ZONA a SENA.

*Príspevok vznikol ako súčasť riešenia VZ MSM 0021267505 Teorie dopravních systémů.*

## POUŽITÁ LITERATURA

- [1] ŠOTEK K., BACHRATÝ H. *Nové možnosti simulačních modelů reálného prostředí v železniční dopravě*, příspěvek do zborníka II konferencie Teorie dopravních systémů, 17. Január 2007, Pardubice, vydala DPJF Univerzita Pardubice, str. 9-18, ISBN 978-80-7194-927-5
- [2] AMCHA, R., BACHRATÝ, H., KRÝŽE, P., VESELÝ, P. *Příspěvek k problematice dodatkových vlaků v prostředí IS SENA*. Příspěvek do zborníka konferencie DOPRAVNÍ SYSTÉMY 2005, 29. november 2005, Pardubice, vydala DPJF Univerzita Pardubice, str. 331-338, ISBN 80-7194-805-5
- [3] BLAHO, P.; GAŠPARÍK, J. *Some aspects of enterprise access to infrastructure of Slovak Republic railways (ŽSR)*. In: Transport XXI wieku, medzinárodná vedecká konferencia, Stare Jablonki, 18.-21.9.2007, Politechnika Warszawska, 2007, s. 91-96, ISBN 978-83-7204-618-5
- [4] ZÁHOROVÁ V. *Určení nástupištní koleje pro zpožděný příjezdějící vlak pomocí fuzzy regulátoru*, In sborník příspěvků konference Inteligentní systémy pro praxi: XI ročník

Ročník 3., Číslo 5., 2008

konference, Bohdaneč, 30.-31. ledna 2008, Ostrava:AD&M konferenční servis, 2008, str.  
103-108, ISBN 978-80-7399-354-2

Recenzent: RNDr. Miroslav Benediktovič  
Univerzita Pardubice, FEI, Katedra softwarových technologií