

# AUTOMATICKÁ TVORBA TISKOVÝCH VÝSTUPŮ OBSLUŽNÝCH TRAS

## AUTOMATIC CREATION OF PRINT OUTPUT OF SERVICE ROUTES

Marek Pomp<sup>1</sup>

---

*Anotace: V příspěvku je popsán postup vytváření itinerářů na závěr procesu optimalizace obslužných tras. Jsou zde podrobně rozebrány použité metody a programové prostředky pro automatizovanou tvorbu kvalitních tiskových.*

*Klíčová slova: dobře dokumentované programy, značkovací jazyk, knitr, Markdown, itinerář,*

*Summary: The paper describes the process of creation itineraries as an endpoint of the process of optimization of service routes. There are described methods and software tools used for an automated creating high-quality print outputs.*

*Key words: well-documented programs, markup language, knitr, Markdown, itinerary,.*

### ÚVOD

V rámci projektu *Efektivní přístup k úsporným a adaptabilním systémům údržby a obsluhy dopravních sítí* je řešena problematika návrhu optimálních obslužných tras. Úlohy jsou řešeny za pomoci počítačových programů, jejichž výstupem jsou databázové tabulky popisující optimální trasu. V tomto příspěvku bude řešen problém vhodného způsobu pro přepis databázových souborů s optimální trasou do formy čitelné pro člověka, do itinerářů v tiskové kvalitě.

## 1. POUŽITÉ PROGRAMOVÉ PROSTŘEDKY

### 1.1 Program R

Program R je software určený pro zpracování dat a jejich grafickou prezenci (R Development Core Team 2010). V průběhu vytváření optimálních obslužných tras je program R používán jednak jako výpočtový systém pro optimalizaci trasy, což v tomto příspěvku nebudeme rozebírat, ale také jako prostředek pro automatickou editaci rozsáhlých datových souborů a konverzi mezi různými formáty dat.

Editace a konverze datových souborů se vyskytuje jednak na samém začátku celého procesu. Vstupní data, která poskytují zadavatelé úloh, jsou vždy ve formátu a struktuře, která není bez rozsáhlých úprav použitelná při dalším zpracování. Tady slouží program R jako vstupní preprocesor.

---

<sup>1</sup> RNDr., Marek Pomp, Ph.D., VŠB - Technická univerzita Ostrava, Ekonomická fakulta, Katedra matematických metod v ekonomii, Sokolská třída 2416/33, 702 00 Ostrava, Tel.: +420 597 322 505, E-mail: [marek.pomp@vsb.cz](mailto:marek.pomp@vsb.cz)

Dále je průběhu některých úloh zapotřebí převod výstupního formátu dat z jednoho programu na vstupní formát v programu dalším, tady je program R nasazen jako prostředek pro konverzi formátů a případnou editaci a úpravu dat.

Na závěr celého procesu je program R použit pro vytváření tiskových výstupů optimalizované obslužné trasy, na vytváření itinerářů.

Vstupem pro vytváření itinerářů jsou databázové soubory s charakteristikou obsluhovaných uzlů a soubor popisující optimální trasu průchodu těmito uzly. V programu R jsou tyto informace propojeny a pomocí balíku *knitr* převedeny do podoby vhodné k tisku.

## 1.2 Knitr

Konverzi výstupů a výsledků z počítačových programů do kvalitní tiskové podoby začal jako jeden z prvních řešit kolem roku 1981 Donald E. Knuth se svou ideou literárního programování založenou na jazyku WEB a systému TeX (Knuth 1984). Na podobných principech v posledních letech vzniklo velké množství nástrojů.

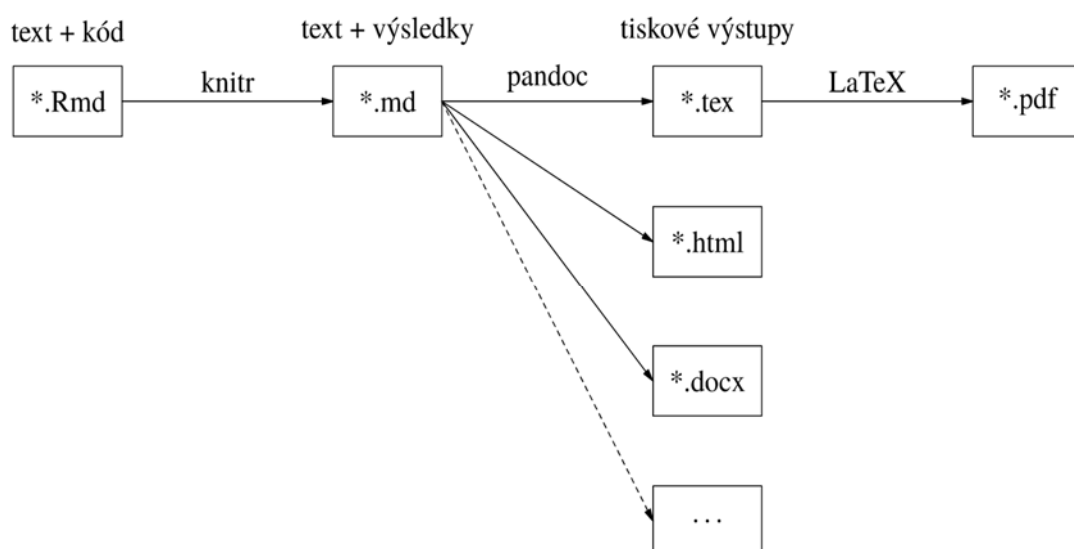
V roce 2002 Friedrich Leisch vytvořil systém *Sweave*, který umožňuje kombinovat text ve formátu LaTeX a zdrojový kód, který je vyhodnocen programem R (Leisch 2002). Tímto systémem se inspiroval v roce 2012 Yihui Xie a vytvořil balík *knitr* (Xie 2015), ve kterém obohatil základní možnosti *Sweave* a integroval výstupy z výpočtů programu R nejen do LaTeXu, ale také do dalších značkovacích jazyků (markup languages), jako je např. HTML nebo Markdown (Gandrud 2013). V roce 2012 Yihui Xie využil nástroj pro konverzi mezi značkovacími jazyky *pandoc* (Pandoc Development Team 2016) a značkovací jazyk *Markdown* k vytvoření nástrojů *R Markdown* respektive *Bookdown* (Xie et al. 2018), (Xie 2016), které již nejsou na vstupu omezeny jazykem R, ale je možné je použít pro generování výstupů z mnoha jiných programovacích jazyků (z více cca 50 jazyků lze namátkou vybrat Fortran, Haskell, C, Perl, Python nebo Julia).

V dokumentu, obvykle s příponou *Rmd*, je směs textu a zdrojového kódu. Text je od kódu oddělen vizuální značkou, třemi zpětnými apostrofy. Za horním oddělovačem bývá nastaven jazyk, ve kterém se má kód vyhodnotit a jsou zde také další přepínače pro vyhodnocení a zobrazení kódu (viz Obr. 1). Knitr nechá v dokumentu *Rmd* nejprve vyhodnotit kód, a posléze je text i s výsledky výpočtů konvertován do formátu, který lze zobrazit na obrazovce počítače, nebo vytisknout (viz Obr. 2).

```
Text
```{jazyk přepínače}
kód, který bude vyhodnocen a jeho výstupy vloženy do výsledného textu
```
text pokračuje
```

Zdroj: Autor

Obr. 1 - Příklad vložení výsledku výpočtů do textu, za pomoci *knitru*



Zdroj: Autor

Obr. 2 - Schéma práce knitru

### 1.3 Pandoc

*Pandoc* je utilita pro konverzi mezi značkovacími jazyky. Tato utilita je portována pro všechny hlavní operační systémy (a jsou k dispozici i její zdrojové soubory, psané v jazyce Haskell). V současnosti zvládá konverzi mezi více než 30 různými jazyky (např. MS-Word, LaTeX, různé dialekty HTML, Markdown, E-Pub apod.). Tuto utilitu je možné spouštět v příkazovém řádku. Např. převod mezi jazykem Markdown a formátem MS-Word bychom zařídili řádkem

```
pandoc -i soubor.md -o soubor.docx
```

Často je použití utility *pandoc* před uživatelem skryto a *pandoc* je použit jako filtr spouštěný na pozadí jinými programy. Takto program *pandoc* využíván *knitrem*.

Na výstupu z utility *pandoc* může být i soubor typu PDF. Konverzi do formátu PDF neprovádí samotný *pandoc*, ale použije jako mezistupeň konverzi do formátu TeX ze kterého je PDF vyroben nějakým klonem pdfTeXu. Samozřejmě musí být program TeX v počítači nainstalován, což by pro nepoučeného uživatele mohl být problém. V programu R proto je možné nainstalovat minimalizovanou verzi TeXu pomocí balíku *TinyTeX*. Balík *TinyTeX* dokáže nainstalovat pdfTeX na všech běžně používaných operačních systémech.

V rámci projektu *pandoc* byl upraven jednoduchý značkovací jazyk *Markdown*, který v tomto systému tvoří jakýsi páteřní jazyk.

### 1.4 Markdown (RMarkdown)

Značkovací jazyk *Markdown* (Gruber 2004) byl vytvořen v roce 2004 jako jazyk pro převod jednoduchých textů do HTML. Brzy se rozšířil ve sféře blogů a v poslední době slouží v různých verzích a mutacích také k publikaci knih a dalších komplikovanějších dokumentů.

Dokument v jazyce Markdown je obyčejný neformátovaný textový soubor, tj. elektronický soubor obsahující jen lidmi čitelné znaky a mezery.

Po formální stránce může být dokument rozdělen na dvě části, hlavičku a vlastní dokument. Hlavička ve formátu *YAML* (Ben-Kiki et al. 2005) specifikuje vlastnosti dokumentu a je od vlastního textu oddělena trojicí pomlček (viz Obr 3).

```
---  
title: Titul publikace  
output: pdf_document  
geometry: a4paper  
fontsize: 12pt  
lang: cs-CZ  
---
```

Zdroj: Autor

Obr. 3 - Hlavička jednoduchého dokumentu

Tělo dokumentu je formátováno pomocí strukturních značek. Formát *Markdown* je velmi jednoduchý, nadpis kapitoly je označen znakem hash # na začátku řádku, položky v nečíslovaných výčtech označuje znak hvězdička \* na začátku řádku apod. Přesto je v tomto jazyce možné zapsat i strukturně poměrně komplikované texty. Např. číslované výčty jsou značeny automaticky zvolenou posloupností, je podporováno automatické generování citací z databází (podobně jako v BibTeXu) a v součinnosti s LaTeXem je umožněna automatická tvorba obsahu a rejstříku apod. (viz Obr 4 a Obr. 5).

K jazyku Markdown bylo vytvořeno mnoho speciální rozšíření a dialektů. Např. tento formát dokumentu byl upraven pro potřeby jazyka R a utility *knitr* na dialekt RMarkdown.

```
# Nadpis kapitoly
```

```
text
```

1. první položka výčtu
1. automaticky bude očíslována druhá položka výčtu

Odkaz na literaturu [[@xie2016bookdown](#)] bude programem pandoc vyhodnocen automaticky a příslušná publikace bude vybrána z databáze a přidána do seznamu použité literatury.

```
![Tady vkládáme obrázek ze souboru](obrazky/logo.png){width=20%}
```

Zdroj: Autor

Obr. 4 - Tělo dokumentu v jazyce Markdown

```
NADPIS KAPITOLY
```

```
text
```

1. první položka výčtu
2. automaticky bude očíslována druhá položka výčtu

Odkaz na literaturu (Xie 2016) bude programem *pandoc* vyhodnocen automaticky a příslušná publikace bude vybrána z databáze a přidána do seznamu použité literatury.



Obrázek: Tady vkládáme obrázek ze souboru.

Zdroj: Autor

Obr. 5 - Výsledný vzhled dokumentu z Obr. 4

## 2. KONVERZE TABULEK V PROGRAMU R

Při vyváření dokumentů balíkem *knitr* je vždy zapotřebí přihlídnout ke specifickým možnostem výsledného výstupního formátu. Napsat v jazyce R Markdown univerzální dokument, který by bylo možné převádět na ekvivalentní dokumenty ve větší množině formátů, je možné jen ve velmi úzkých mantinelech, které vymezují speciální typografické a jiné požadavky autora. Ovšem v případě komplikovaných požadavků na vzhled tabulek je velmi často nutné ustoupit od univerzálnosti a vyrovnat se s různými speciálními vlastnostmi výstupních systémů. Autor si tedy musí způsob pro konverzi tabulek vybrat v závislosti na výstupních formátech, které by rád pro svůj dokument využil.

Při vytváření tabulek má autor na výběr z následujících možností.

### 2.1 Převod jednoduchých tabulek příkazem *kable*

V balíku *knitr* je k dispozici funkce *kable()*. Touto funkcí je možné univerzálně vytvářet jednoduché tabulky pro všechny výstupní formáty *knitru*. Speciální formátování těchto tabulek však není možné.

### 2.2 *KableExtra* převod do html a pdf

Balík *KableExtra* vytváří nadstavbu nad funkcí *kable* a umožňuje i velmi speciální formátování tabulek, jejich záhlaví a zápatí, jednotlivých řádků, sloupců, nebo přímo buněk (Pomp 2018). Nevýhodou tohoto balíku je, že umožňuje jen export do formátů pdf a html.

### 2.3 Balík *xtable*, komplexní systém pro převod do pdf a html

Balík *xtable* je jeden z nejstarších a nejkompexnějších balíčků pro formátování tabulek v programu R. Vznikl před rokem 2000 pro formátování tabulek do formátu LaTeX a byl později rozšířen i na formát html. Jeho opravdu silné možnosti může trochu brzdit to, že autor dokumentu by měl znát způsob formátování tabulek v jazyce LaTeX.

### 2.4 Systém *pander*, převod do Markdown

Balík *pander* je knihovna jazyka R, která slouží ke konverzi tabulek a dalších objektů do jazyka Markdown. Protože je jazyk Markdown jakýmsi středobodem při konverzích programem *pandoc*, je *pander* poměrně univerzální systém. Pomocí tohoto balíku lze konvertovat tabulky z programu R nejen do formátů pdf a html ale také docx a dalších.

Navzdory své univerzálnosti poskytuje balík *Pander* velmi silné možnosti a umožňuje komplikované formátovací zásahy do struktury tabulky.

## 2.5 Další možnosti převodu do MS-Word

Protože převod tabulek z programu R do dokumentů ve formátu MS-Word z předchozích způsobů obstojně zvládá jen balík *pander*, je dobré zmínit, že existují další balíčky, které umí řešit tuto problematiku. Je to např. dvojice balíčků *huxtable* a *flextable*.

## 3. ŘEŠENÍ PROBLÉMU

V rámci projektu *Efektivní přístupu k úsporným a adaptabilním systémům údržby a obsluhy dopravních sítí* jsou vytvářeny itineráře pro svoz komunálního odpadu. Pro koncového uživatele jsou výstupy připraveny ve formátech PDF, docx a ve formátu xlsx. Ovšem je možné výstupy připravit také ve formátu html a dalších.

Jako vstupní data slouží

- seznam obsluhovaných adres, který poskytuje klient obvykle ve formátu xlsx
- výstupy z procesu optimalizace trasy, pořadí obsluhovaných uzlů ve formě řady čísel,
- speciální parametry úlohy v textovém souboru s předem danou strukturou.

Speciální parametry umožňují snadnou a flexibilní modifikaci úlohy. Speciálními parametry úlohy jsou

- administrativní údaje, jako je datum, údaje o koncovém uživateli apod., které slouží k vytváření hlaviček a titulních listů
- technické údaje, např. kapacita svozových vozů pro případné rozdělení svozu na části atd.
- pokyny pro skript, výstupní formát pro itinerář, typ údajů které budou do itineráře zahrnuty apod.

Vlastní řešení problému vypadá následovně. V jazyce *R Markdown* byl vytvořen dokument *itinerar.Rmd*. Tento soubor je pak příkazem v příkazovém řádku (nebo skriptu) v operačním systému Linux

```
Rscript -e "rmarkdown::render('itinerar.Rmd')"
```

převeden do požadovaných formátů (viz Obr. 6). Dokument ve formátu R Markdown je navržen tak, aby generování itinerářů mohlo probíhat automaticky, pouze na základě obsahu souboru se speciálními parametry. Všechny potřebné funkce a operace jsou uvedeny v tomto dokumentu, v rámci vloženého kódu v jazyce R. Dokument tedy v podstatě sám o sobě

- načítá data z příložených souborů a upravuje jejich strukturu,
- řadí obsluhované adresy v pořadí optimální obslužné trasy,
- podle zadaných parametrů vybírá údaje o obsluhovaném místě,
- vytváří tabulky s popisem obsluhovaných míst,
- podle zadaných parametrů vytváří titulní strany a případně generuje další speciální pokyny v itinerářích.

| Č. | Ulice   | Č.p. | GPS              | Typ | ks | Celkem |
|----|---------|------|------------------|-----|----|--------|
| 32 | Písečná | 54   | 49.5689, 18.7877 | 110 | 1  | 464    |
| 33 | Písečná | 64   | 49.5690, 18.7872 | 110 | 1  | 478    |
| 34 | Písečná | 62   | 49.5692, 18.7872 | 110 | 1  | 493    |
| 35 | Písečná | 60   | 49.5696, 18.7870 | 110 | 1  | 508    |
| 36 | Písečná | 61   | 49.5693, 18.7869 | 110 | 1  | 522    |
| 37 | Písečná | 176  | 49.5692, 18.7868 | 110 | 1  | 536    |
| 38 | Písečná | 99   | 49.5691, 18.7862 | 110 | 1  | 551    |
| 39 | Písečná | 117  | 49.5687, 18.7864 | 110 | 1  | 566    |
| 40 | Písečná | 94   | 49.5700, 18.7867 | 110 | 1  | 580    |

Zdroj: Autor

Obr. 6 - Ukázka z itineráře

## ZÁVĚR

Dokumenty v jazyce *Markdown* a skripty v programu R umožňují snadné a jednoduché automatické vytváření itinerářů, které lze flexibilně exportovat do různých formátů, podle požadavků koncového uživatele. Skripty pro vytváření itinerářů lze snadno připojit k jiným výpočetním úlohám, a proto je možné zautomatizovat celý proces od vstupní úpravy dat, která poskytuje zadavatel úlohy, přes výpočet optimální trasy až po tisk výsledných itinerářů, které pak mohou používat řidiči projíždějící obslužné trasy.

Vytvoření takové fronty úloh, jejíž řešení probíhá automaticky v návaznosti na výsledky předchozích kroků, minimalizuje vliv lidského faktoru, který je náchylný k chybám a v procesu řešení úlohy by mohl působit rušivě.

*Príspevek byl vytvořen v rámci Programu na podporu aplikovaného výzkumu a experimentálního vývoje EPSILON, Technologické agentury České republiky: „Efektivní přístupy k úsporným a adaptabilním systémům údržby a obsluhy dopravních sítí“ TH02010930.*

## LITERATURA

- (1) BEN-KIKI, Oren, Clark EVANS a Brian INGERSON. *Yaml ain't markup language (yaml™) version 1.1. yaml.org, Tech. Rep. 23. 2005.*
- (2) GANDRUD, Christopher. *Reproducible Research with R and R Studio*. Boca Raton: Chapman; Hall/CRC. 2013. ISBN 978-1498716963.

- (3) GRUBER, John. *Daring Fireball: Markdown* [online]. 2004 [cit. 2019-08-01]. Dostupné z: <http://daringfireball.net/projects/markdown>
- (4) KNUTH, D. E. Literate Programming. *The Computer Journal* [online]. **27**(2), 97–111. 1984. ISSN 0010-4620.
- (5) LEISCH, Friedrich. Sweave: Dynamic Generation of Statistical Reports Using Literate Data Analysis. In: Wolfgang HÄRDLE a Bernd RÖNZ, ed. *Compstat 2002 — Proceedings in Computational Statistics*. B.m.: Physica Verlag, Heidelberg, s. 575–580. 2002. ISBN 3-7908-1517-9.
- (6) PANDOC DEVELOPMENT TEAM. *Pandoc: the swiss-army knife for converting files from one markup format into another*. [online]. 2016 [cit. 2019-08-01]. Dostupné z: <http://pandoc.org>
- (7) POMP, Marek. Tabulky v dobře dokumentovaných statistických výpočtech. *Zpravodaj Československého sdružení uživatelů TeXu* [online]. (1–4), 22–37. 2018. ISSN 1211-6661.
- (8) R DEVELOPMENT CORE TEAM. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. 2010. ISBN 3-900051-07-0.
- (9) XIE, Yihui. *Bookdown: Authoring Books and Technical Documents with R Markdown*. New York: Chapman; Hall/CRC. 2016. ISBN 978-1138700109.
- (10) XIE, Yihui. *Dynamic documents with R and knitr*. Boca Raton: Chapman; Hall/CRC. 2015. ISBN 978-1498716963.
- (11) XIE, Yihui, Joseph J ALLAIRE a Garrett GROLEMUND. *R markdown: The definitive guide*. B.m.: Chapman; Hall/CRC. 2018. ISBN 978-1138359338.