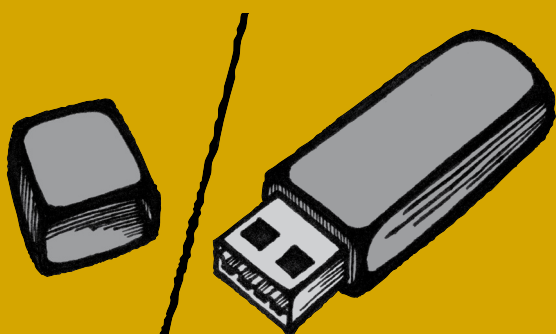


FAKULTA
APLIKOVANÝCH VĚD
ZÁPADOČESKÉ
UNIVERZITY
V PLZNI

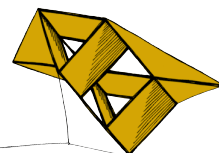
SBORNÍK

STUDENTSKÁ VĚDECKÁ KONFERENCE



BAKALÁŘSKÉ
STUDIJNÍ PROGRAMY

25.5.2017



Bakalářské studijní programy

Sborník rozšířených abstraktů

Editor: Vladimír Lukeš

Autor obálky: Petr Lobaz

Vydavatel: Západočeská univerzita v Plzni, Univerzitní 8, 306 14 Plzeň

Datum vydání: květen 2017

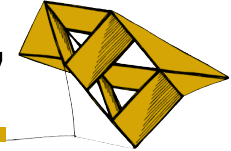
ISBN 978-80-261-0705-7

Obsah

<i>Sekce – Fyzika, matematika, mechanika</i>	5
Numerické řešení úlohy kontaktu elastických těles Jan Holeček	7
Růst a remodelace jaterního parenchymu Dominik Kasl	9
Numerický model pro stanovení ohybové tuhosti florbalové hole Dmytro Korch	11
Modely proudění na sítích pro popis tkáňové perfúze a saturace Martin Levý	13
Difúzní modely na grafech Jan Matas	15
<i>Sekce – Informatika, kybernetika</i>	17
Mobilní aplikace pro sběr medicínských dat David Bohmann	19
Návrh a implementace deskové hry pro kognitivní trénink Michal Horký	21
Poker Cards Recognition using Neural Networks Jan Hrdlička	23
Detekce přízvuků v ruštině s použitím klasifikátoru Anastasiia Chizhova	25
Analýza pohybu RC autíčka pro návrh autonomního řízení Jakub Matoušek	27
Systém pro správu publikací Ondřej Pittl	29
Explicitní sémantická analýza Michal Tušl	31

Sekce

Fyzika, matematika, mechanika



Numerické řešení úlohy kontaktu elastických těles

Jan Holeček¹

1 Úvod

Úloha kontaktu a její řešení se uplatňuje v mnoha technických aplikacích, například při modelování styku páru ozubených kol, návrhu konstrukcí, crash testech aj.

V tomto příspěvku bude stručně popsán vytvořený matematický model pro řešení úlohy jednostranného kontaktu elastického tělesa s dokonale tuhous překážkou a ukázány výsledky numerické simulace. Při formulaci vyjdeme z následujících zjednodušení. Předpokládáme malé deformace a posuvy (úloha elastostatiky) a linearizovanou překážku. Dále předpokládáme nulové tření v místě kontaktu. Úlohu řešíme ve 2D.

2 Model úlohy kontaktu bez uvažování tření

Pro vytvoření diskretizovaného matematického modelu byla použita metoda konečných prvků, odvození např. v Rohan E. (1997). Matematický model má podobu soustavy nehladkých algebraických rovnic (maximum je bráno po složkách):

$$\begin{cases} \mathbf{X}(\mathbf{R}\mathbf{K}\mathbf{u} - \mathbf{R}\mathbf{f}) = 0 \\ \max\{\mathbf{Y}(\mathbf{R}\mathbf{K}\mathbf{u} - \mathbf{f}), \mathbf{Y}\mathbf{R}\mathbf{u} - \mathbf{s}\} = 0. \end{cases} \quad (1)$$

\mathbf{K} je matice tuhosti, \mathbf{u} je vektor posuvů, \mathbf{f} je vektor sil působících v uzlech, \mathbf{s} je vektor vzdáleností uzlů na kontaktní hranici ve směru vnější normály na hranici. \mathbf{R} je matice rotace, která převádí globální souřadnicový systém na kontaktní hranici na lokální systém s osami ve směru tečny a normály v daném bodě kontaktní hranice. \mathbf{X} a \mathbf{Y} jsou tzv. restriční matice, které vybírají z vektorů jen některé stupně volnosti takto: vynásobením vektoru \mathbf{u} zleva maticí \mathbf{X} získáme stupně volnosti odpovídající posuvům uzlů mimo kontaktní hranici a posuvům uzlů na hranici ve směru tečny; vynásobením vektoru zleva maticí \mathbf{Y} získáme stupně volnosti odpovídající posuvům uzlů na hranici ve směru normály.

První rovnice v (1) je ekvivalentní s formulací úlohy elastostatiky, druhá rovnice popisuje kontaktní podmínku. Je zřejmé, že bude nulová síla v uzlech na kontaktní hranici (není kontakt), nebo bude v daném uzlu nulová vzdálenost mezi tělesem a překážkou (případně budou nulové obě hodnoty).

Takto získaná soustava nehladkých rovnic (1) byla numericky řešena pomocí tzv. Newtonovy metody s tlumením, jejíž tvar je:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha \mathbf{J}(\mathbf{x}_k)^{-1} \mathbf{F}(\mathbf{x}_k),$$

kde \mathbf{x}_k je vektor neznámých, $\mathbf{F}(\mathbf{x}_k)$ je vektor funkčních hodnot a $\mathbf{J}(\mathbf{x}_k)$ je Jacobiova matice a $\alpha \in (0, 1)$ je parametr tlumení.

¹ student bakalářského studijního programu Počítačové modelování v technice, obor Počítačové modelování, e-mail: holecekj@students.zcu.cz

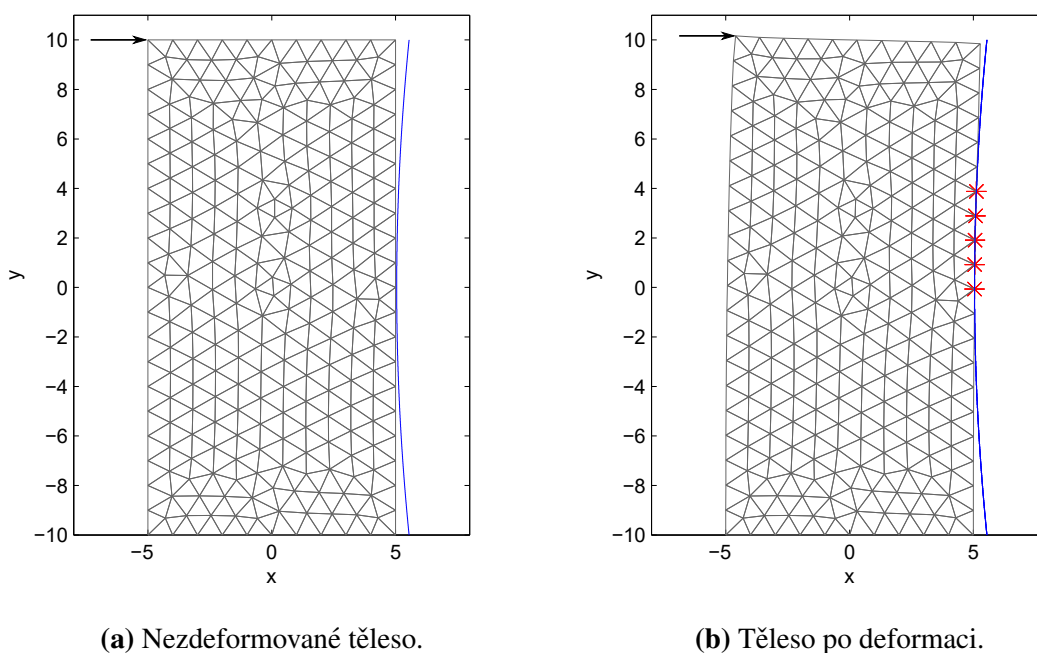
Jiný, též použitý způsob řešení je popsán v De Luca et al. (1996). Pro tento způsob je třeba přeformulovat druhou rovnici v (1) pomocí tzv. Fischerovy-Burmeisterovy funkce $\phi_{FB}(a, b) \equiv \sqrt{a^2 + b^2} - (a + b)$, viz. Pieracchini S. et al. (2003)), takže řešíme úlohu $\phi_{FB}(a, b) = 0$. Oba tyto způsoby vedou ke stejným výsledkům, neboť platí $\phi_{FB}(a, b) = 0 \Leftrightarrow \min(a, b) = 0$.

3 Výsledky

Matice \mathbf{K} a \mathbf{f} byly získány aplikací metody konečných prvků v software SfePy (Simple Finite Elements in Python), numerická simulace byla následně provedena v software MATLAB. Ukázka výsledků simulace je na obr. 1. Těleso je zatíženo osamělou silou působící v uzlu v levém horním rohu, na dolní hranici je těleso vetknuto. Uzly, ve kterých došlo ke kontaktu jsou označeny červeně, překážka vykreslena modře. Vyčíslením výrazu

$$\mathbf{Y}(\mathbf{R}\mathbf{K}\mathbf{u} - \mathbf{f})$$

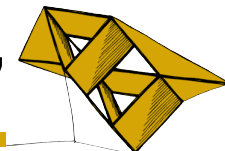
je možné určit uzlové kontaktní síly. Jejich znalost je nezbytná pro možné rozšíření modelu o nenulové tření na kontaktní hranici. V současné době je práce rozšiřována o model suchého tření.



Obr. 1: Výsledky numerické simulace v software MATLAB.

Literatura

- Rohan E. (1997) Contact shape optimalization of elasto-plastic bodies. *Proceedings of the University of West Bohemia '97*. University of West Bohemia.
- De Luca T., Facchinei F., Kanzow C. (1996) A Semismooth Equation Approach To The Solution Of Nonlinear Complementarity Problems. *Mathematical Programming (1996)* 75.
- Pieracchini S., Gasparo M. G., Pasquali A. (2003) Global Newton-type methods and semismooth reformulations for NCP. *Applied Numerical Mathematics* 44 (2003).



Růst a remodelace jaterního parenchymu

Dominik Kasl¹

1 Úvod

V této práci se autor zabýval sestavením modelu popisující růst a remodelaci jaterního parenchymu a jeho následným numerickým řešením. Model byl navržen na bázi teorie směsí kombinované s konceptem objemových poměrů využívající klasických metod mechaniky kontinua a termodynamických principů. Model je především určen k simulaci remodelace jaterních cév a růstu, tedy změny objemového zastoupení, hepatocytů ve zkoumaném médiu. Dále byl diskutován vliv mechanického zatížení na samotný proces růstu a remodelace.

2 Remodelace sinusoid a vliv mechanických účinků na růst měkkých tkání

Samotný model byl simulován jako třífázové (složkové) porézní médium složené z 1 pevné složky reprezentující hepatocyty a ze 2 tekutých složek reprezentující krev, protékající jaterním parenchymem a nutrienty (živiny). Proces remodelace jaterních cév, tzv. sinusoid, byl v Ricken, T., (2010) realizován pomocí strukturního tenzoru:

$$M_S = a_S \otimes a_S = \mathbf{F}_s a_{0S} \otimes \mathbf{F}_s a_{0S}, \quad (1)$$

kde a_S je preferovaný směr toku, a_{0S} je počáteční směr toku, \mathbf{F}_s je deformační gradient tuhé složky a symbol \otimes značí tenzorový součin.

Samotný růst je dle Ambrosi, D., (2011) způsoben jednak biochemickou energií, viz Ambrosi, D., (2007), ale i mechanickým zatížením působící na médium. Např v Ricken, T., (2010) byl nárůst hmotnosti tuhé složky postulován následujícím způsobem:

$$\hat{\rho}^S = \hat{\rho}_{max}^S \hat{\rho}_{nN}^S \hat{\rho}_{J_s}^S \hat{\rho}_{\tau_v Mi}^S, \quad (2)$$

$$\hat{\rho}_{nN}^S = -e^{\kappa_{nN}(n^N)^2} + 1, \quad (3)$$

$$\hat{\rho}_{J_s}^S = -e^{\kappa_{J_s}(J_s-1)^2} + 1, \quad (4)$$

$$\hat{\rho}_{\tau_v Mi}^S = -2e^{-\log(2)\frac{\tau_v Mi}{\tau_v Mi0}} + 1, \quad (5)$$

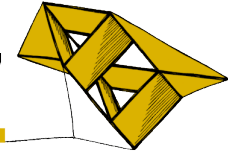
kde $\hat{\rho}_{max}^S$, κ_{nN} , κ_{J_s} jsou parametry určující vlastnosti materiálu, n^N je objemový poměr živin a $\tau_v Mi0$ je optimální efektivní napětí, u které se neočekávají žádné nárůsty hmotností.

Celý model je tvořen soustavou vzájemně propojených parciálně diferenciálních rovnic. Tento systém byl řešen metodou přímek, tedy metodou, ve které byly zdiskretizovány všechny proměnné až na jednu. Diskretizace byla provedena metodou konečných diferencí. Touto procedurou byl získán systém obyčejných diferenciálních rovnic.

¹ student bakalářského studijního programu Počítačové modelování v technice, obor Počítačové modelování, e-mail: dkasl@students.zcu.cz

Literatura

- Ambrosi, D.,Guillou, A. (1980) Growth and dissipation in biological tissues. *Continuum Mech. Thermodyn.* (2007) 19: 245–251.
- Ambrosi, D., Ateshian, G. A., Arruda, E. M., Cowin, S. C., Dumais, J., Goriely, A., Holzapfel, G. A., Hunphrey, J. D., Kemkemer, R., Kuhl, E., Olberding, J. E., Taber, L. A. Garikipati K. (2011) Perspectives on biological growth and remodeling. *Journal of the Mechanics and Physics of Solids.*
- Ricken, T., Dahmen, U., Dirsch, O. (2010) A biphasic model for sinusoidal liver perfusion remodeling after outflow obstruction. *Biomech Model Mechanobiol* (2010) 9:435–450.
- Ricken, T., Bluhm, J. (2010) Remodeling and growth of living tissue: a multiphase theory. *Arch Appl Mech* (2010) 80: 453–465.



Numerický model pro stanovení ohybové tuhosti florbalové hole

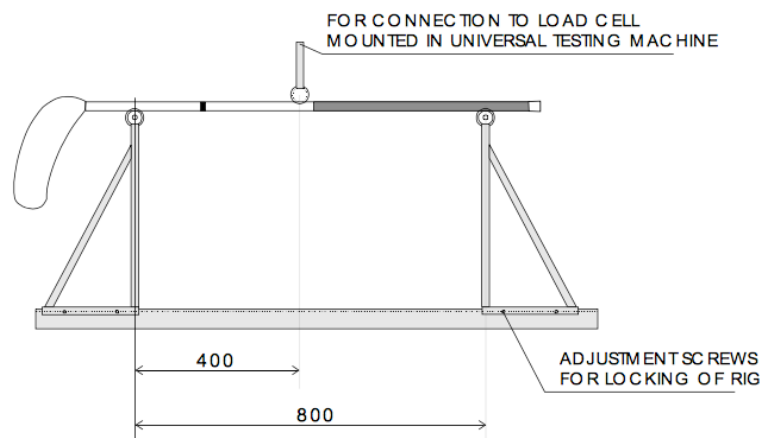
Dmytro Korch¹

Úvod

Florbal je mladý sport, ale v poslední době nabyt zvlášť u mládeže velké popularity. Dnes je informovanost lidí o tomto zábavném sportu mnohonásobně vyšší a u dětí je tento sport velmi oblíbený. Jelikož Česká republika je ve florbale jedna z nejlepších zemí světa jsou zde také podmínky a zázemí pro florbal jedny z nejlepších. Kvalita tohoto sportu se posunula velmi vpřed, je mnohem rychlejší, dynamičtější a taktičtější. Přesto, že je tento sport tak oblíbený je trochu překvapivé, že v České republice není žádná firma, která by se specializovala na florbal.

Tato práce se zabývá vytvořením numerického modelu pro stanovování ohybové tuhosti florbalové hole, verifikací tohoto modelu a analýzou vlivu skladby na výslednou tuhost.

Experimentální stanovení ohybové tuhosti



Obrázek 1: IFF certifikační zkouška na ohyb

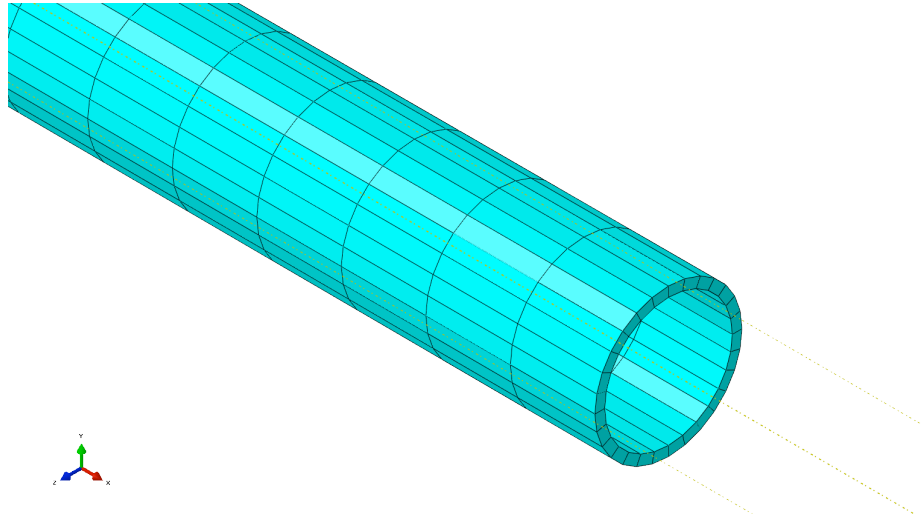
Jedním ze základních parametrů florbalové hole je její ohybová tuhost (tzv. flex). Experimentální stanovení této ohybové tuhosti (flexu) je definováno v IFF (International Floorball Federation) regulacích [1], které také definují geometrické parametry florbalové hole, a dalšího florbalového vybavení. Flex se určuje ze zkoušky tříbodovým ohybem, jak lze vidět na obr. 1. Hodnota flexu je průhyb hole, při zatěžující síle 300 N a vzdálenosti podpor 800 mm a obvykle se pohybuje v hodnotách 23 mm až 32 mm.

V rámci této práce byla uvedená metodika vyzkoušena na několika různých holích.

¹ student bakalářského studijního programu Počítačové modelování v technice, obor Mechanika, specializace Výpočty a design, e-mail: korkhd@students.zcu.cz

Experimentální stanovení ohybové tuhosti

Pro výpočtový systém Abaqus byl pomocí programovacího jazyka Python vytvořen parametrický model kompozitové trubky s různými průměry průřezu na obou koncích a konstantní tloušťce (viz obrázek 2).



Obrázek 2: Konečnoprvkový model florbalové hole

Tento model byl verifikován pomocí experimentální zkoušky tenkostěnné kompozitové trubky konstantního průřezu. Při této verifikaci byly testovány parametry sítě konečnoprvkového modelu, jako například velikost prvku v podélném směru trubky a počet prvků na průřezu. Hodnoty parametrů sítě, při kterých bylo dosaženo chyby mezi experimentem a numerickou simulací menší než 2%, byly pak použity v modelech florbalových holí.

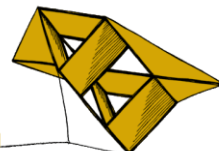
Takto vytvořeným numerickým modelem byla simulována ohybová zkouška pro stavení ohybové tuhosti (flexu) dle standardů IFF. Byla provedena analýza vlivu natočení vláken v jednotlivých vrstvách kompozitu a počet těchto vrstev na výsledný průhyb. Tato analýza byla uskutečněna pro kompozit se zvoleným typem uhlíkových vláken a tloušťkou jedné vrstvy 0,29 mm. Při této konfiguraci byla např. prokázána nevhodnost vrstvy s úhlem vláken 0° z důvodů příliš vysoké ohybové tuhosti.

Závěr

V systému Abaqus byl vytvořen parametrický model kompozitové trubky s možností nastavení různých průměrů průřezu na obou koncích. Tento numerický model byl verifikován pomocí experimentální zkoušky. Chyba mezi modelem a experimentem byla menší než 2%. Z této verifikace byly také určeny parametry sítě pro následný model ohybové zkoušky florbalové hole. Byla provedena analýza vlivu skladby vláken a jejich tloušťek na výsledný průhyb určující ohybovou tuhost – tzv. flex.

Literatura

[1] IFF MATERIAL REGULATION. IFF [online]. Helsinky [cit. 2017-05-21]. Dostupné z: <http://www.floorball.org/pages/EN/Material-Regulations>



Modely proudění na sítích pro popis tkáňové perfúze a saturace

Martin Levý¹

1 Úvod

Tato práce se zabývá modelováním 1D proudění a transportu na sítích pro popis tkáňové perfúze se zaměřením na tkáni jater, která jsou uvažována jako porézní prostředí, dále se zaměřuje na získání hodnot permeability sítě a na modelování saturace tkáně kontrastní látkou s uvažováním disperze. Cílem práce je možnost získání koeficientů permeabilit potřebné pro multi-kompartmentové modely jaterní perfúze viz Rohan et al. (2015) a usnadnit lékařům predikci saturace tkáně kontrastní látkou.

2 Modely proudění

Pro získání stavových hodnoty na uzlech a segmentech sítě byly popsány a implementované modely 1D proudění na síti následující:

- Model vazkého proudění zanedbávající setrvané účinky definovaný Poiseuilleovo rovnicí a rovnicí kontinuity viz Noskievič et al. (1987)
- Model všímající si setrvaných účinků proudění a zanedbávající vazkost proudění popsán Bernoulliho rovnicí a rovnicí kontinuity viz Noskievič et al. (1987)
- Kombinovaný model vycházející z Bernoulliho rovnice doplněné o ztrátový člen a rovnicí kontinuity viz Rohan et al. (2015)

Těmito modely byly následně získány hodnoty tlaků a rychlostí na síti.

3 Permeabilita sítě

V úloze řešení proudění pomocí Darcyho rovnice se vyskytuje tenzor Permeability \mathbf{K} viz Rohan et al. (2015), který lze získat z výsledků proudění na síti z předchozích modelů. Samotné hodnoty pro výpočet tenzoru Permeability \mathbf{K} jsou získány na pomyslném hranolku ohraničující v určitém bodě síť viz Debbaud et al. (2012).

4 Saturace sítě

Dále byla vyhodnocena saturace sítě v závislosti na čase s uvažováním disperze. Rovnice saturace na proudnici uvnitř segmentu lze v určitém bodě $[x,y,t]$ vyjádřit jako:

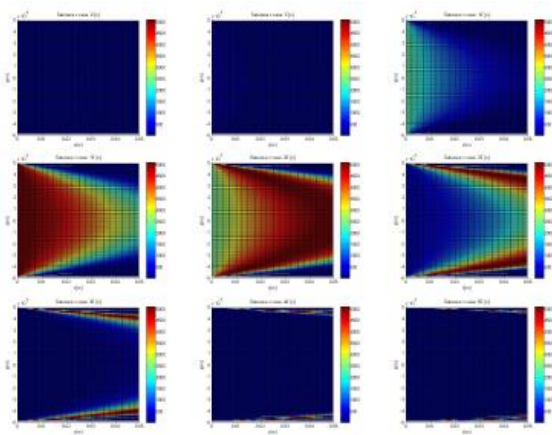
$$s(x, y, t) = s_0(t - \tau(x, y))$$

Kde s_0 je saturace na proudnici v prvním bodě segmentu a τ je transportní čas bolusu kontrastní látky.

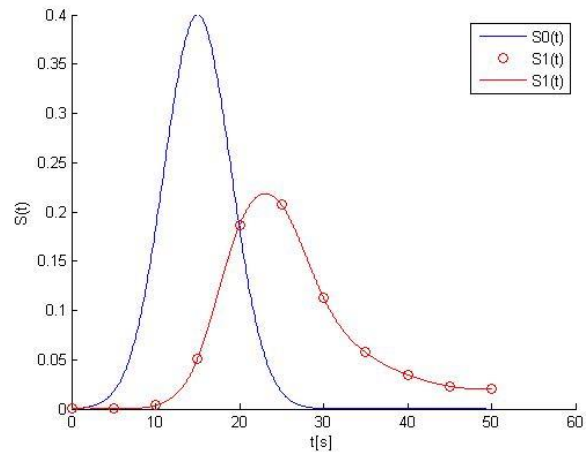
¹ student bakalářského studijního programu Počítačové modelování v technice, obor Počítačové modelování, e-mail: levym@students.zcu.cz

Pokud označíme saturaci v prvním uzlu $S_0=s_0/A$, kde A je průsvit segmentu, pak saturace v druhém uzlu je definována jako:

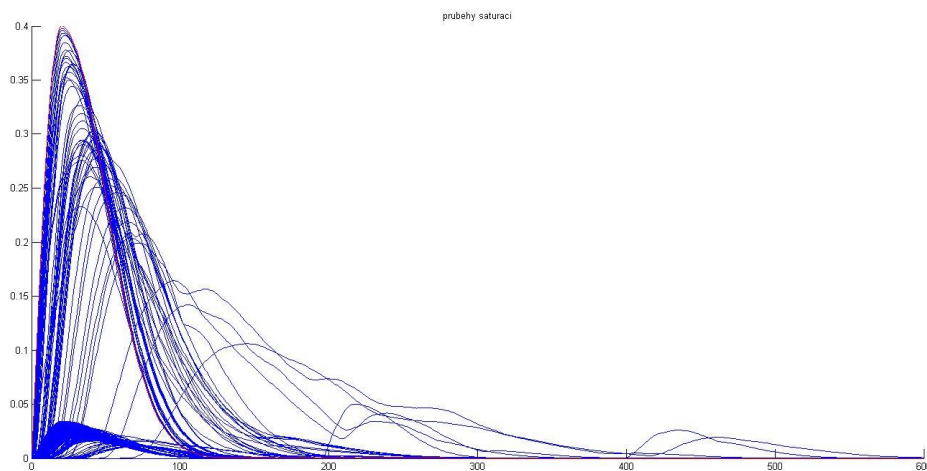
$$S_1(t) = \int_0^r s(l, y, t) y 2\pi dy = \int_0^r s_0(t - \tau(l, y)) y 2\pi dy = \int_0^r \frac{S_0(t - \tau(l, y))}{A} y 2\pi dy$$



Obrázek 1: Saturace segmentu



Obrázek 2: Saturace uzlů segmentu



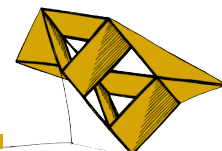
Obrázek 3: Saturace sítě

Literatura

Noskievič, J., 1987. Mechanika tekutin. SNTL Praha.

Debbaut, C., Vierendeels, J., Casteleyn, C., Cornillie, P., Van Loo, D., Simoens, P., Van Hoorebeke, L., Monbaliu, D., Segers, P. 2012 *Perfusion Characteristics of the Human Hepatic Microcirculation Based on Three-Dimensional Reconstructions and Computational Fluid Dynamic Analysis*.

Rohan, E., Lukeš, V., Brašnová, J., 2015. CT based identification problem for the multicom-
partment model of blood perfusion. In: Computational Vision and Medical Image
Processing V. London: CRC Press. s. 289-294. ISBN 978-1-138-02926-2.



Difúzní modely na grafech

Jan Matas¹

1 Úvod

Tato práce se zabývá matematickým modelováním difúze mezi oddělenými oblastmi, jejichž vzájemná propojení modelujeme grafy. S pomocí teorie grafů a obyčejných diferenciálních rovnic vytvoříme obecný difúzní model nad grafem se dvěma vrcholy, který následně zobecníme pro obecný souvislý neorientovaný graf.

Difúzi zde nechápeme pouze jako přesun z oblastí s vyšší koncentrací do oblastí s nižší koncentrací, ale jako obecnější proces přesunu daný difúzní funkcí. Je-li difúzní funkce lineární, modelujeme difúzi v již popsaném klasickém pojetí. Zvolíme-li ji ale nelineární, můžeme modelovat složitější procesy, např. shlukování a koexistence.

2 Difúzní modely pro graf se dvěma vrcholy

Mějme dvě místa (lokality), na kterých může přebývat nějaká populace, a která jsou navzájem propojena. Jedinci populace se tak mohou libovolně přemisťovat z jednoho místa do druhého. Tuto situaci můžeme znázornit grafem se dvěma vrcholy, x a y , které představují jednotlivé lokality, a neorientovanou hranou, reprezentující jejich propojení.

Dynamiku migrace jedinců populace mezi jednotlivými lokalitami můžeme popsat následující soustavou diferenciálních rovnic:

$$\begin{aligned}x'(t) &= \psi(x(t) - y(t)), \\y'(t) &= \psi(y(t) - x(t)),\end{aligned}\tag{1}$$

kde $x(t)$ a $y(t)$ reprezentují relativní četnosti jedinců celkové populace přebývajících v jednotlivých lokalitách a difúzní funkce $\psi(s)$ charakterizuje vnitřní tendence pohybu populace.

Po formálním zavedení modelu jsme se zabývali jejich asymptotickým chováním pro konkrétní volby difúzní funkce $\psi(s)$, například v případě lineární difúze je asymptoticky klidovým stavem vždy 0.5, viz průběh řešení na Obrázku 1.

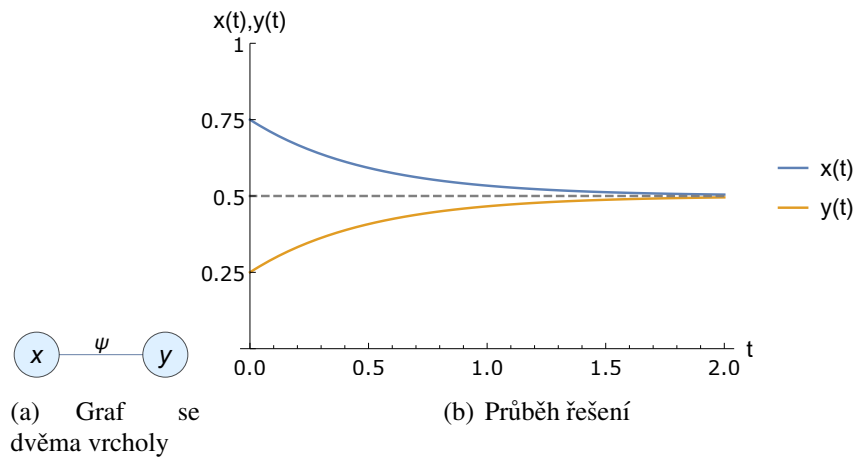
3 Difúzní modely pro obecný neorientovaný graf

Ve druhé části práce jsme se zabývali zobecněním výše popsaných modelů. Mějme obecný souvislý neorientovaný graf G s množinou vrcholů $V(G) = \{1, \dots, n\}$ popsany maticí sousednosti $\mathbf{A}_G = (a_{ij})_{n \times n}$. Pak můžeme soustavu rovnic popisující model difúze na grafu zapsat ve tvaru:

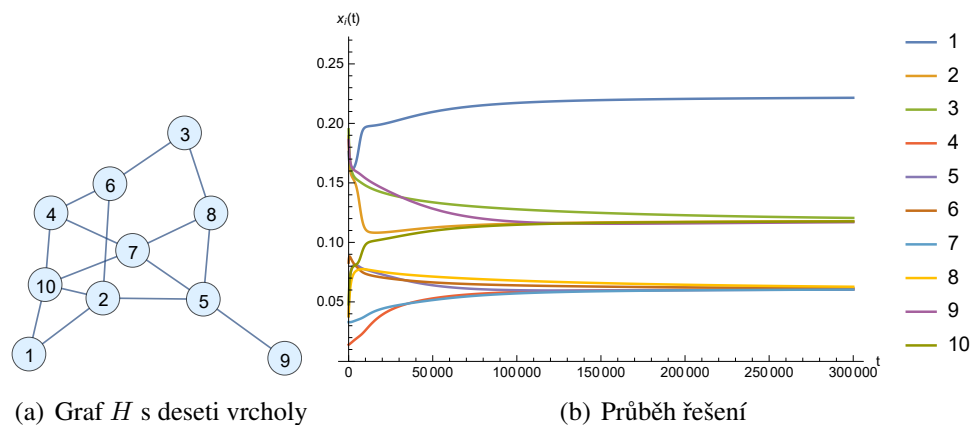
$$x'_i(t) = \sum_{j=1}^n a_{ij} \phi(x_i(t), x_j(t)), \quad i \in V(G), t \geq 0,\tag{2}$$

kde $\phi(u, v)$ je difúzní funkce.

¹ student bakalářského studijního programu Matematika, obor Matematika a finance, e-mail: janmatas@students.zcu.cz



Obrázek 1: Volbou difúzní funkce $\psi(s) := -rs, r > 0$ a dosazením do soustavy (1) získáme model lineární difúze. Na obrázku je zachycen průběh řešení počáteční úlohy pro takto vzniklou soustavu s počátečními hodnotami $x(0) = 0.75, y(0) = 0.25$ a $r = 1$.



Obrázek 2: Volbou difúzní funkce $\psi(s) := rrv(u - v)(v - (2u - u^2))(u - (2v - v^2)), r > 0$ a dosazením do soustavy (2) získáme model koexistence. Na obrázku je zachycen průběh řešení počáteční úlohy takto vzniklé soustavy nad grafem H pro počáteční podmínky $\mathbf{x}_0 = (0.176, 0.189, 0.195, 0.014, 0.038, 0.083, 0.033, 0.038, 0.186, 0.048)^T$ a $r = 1$.

Tyto modely nemusí popisovat pouze stěhování jedinců, ale i např. změny názorů, preferencí (volebních, spotřebitelských), strategií apod., které jsou reprezentovány vrcholy grafu a hrany pak představují možné přechody. Dalším zdokonalením modelů bychom mohli zkoumat chování voličů spotřebitelů, hráčů, atd.

4 Seznam literatury a citace

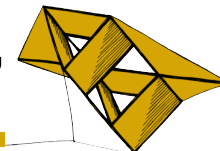
Literatura

de Vries, G., et al. (2006), *A Course in Mathematical Biology*, SIAM.

Kelley, W., Peterson, A. (2010), *The Theory of Differential Equations*, Springer.

Sekce

Informatika, kybernetika



Mobilní aplikace pro sběr medicínských dat

David Bohmann¹

1 Úvod

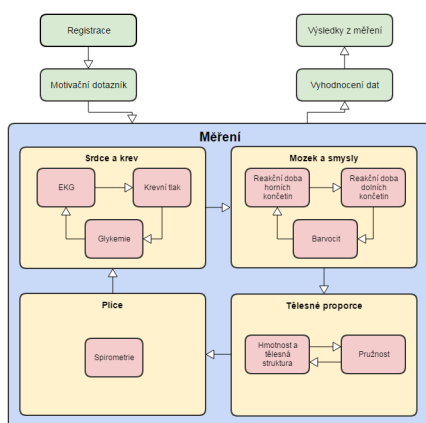
Každý den jsme zahlcováni informacemi o tom, co je zdraví prospěšné či škodlivé, co dělat, abychom byli hubení, zdraví, či abychom si udrželi mládí. V takovém množství informací se nemůže průměrný člověk vyznat. Většina reklam cílí na masy lidí bez ohledu na to, že organismus každé osoby funguje rozdílně a tudíž každý potřebuje k dosažení a udržení optimálního zdraví jiný přístup. Naším dlouhodobým cílem je poskytnout každé osobě komplexní informace o jeho zdravotním stavu a individuálně doporučit, v jakém ohledu je možné zdravotní stav a kondici zlepšit, ať se jedná o změnu stravování, doplnění vitamínů, či více pohybové aktivity nebo pravidelný spánek.

2 Exercise and Wellness Health Strategy Framework

Mobilní aplikace je vyvíjena jako součást projektu „Exercise and Wellness Health Strategy Framework [Brůha et. al (2017)“, který je vyvíjen na Katedře informatiky a výpočetní techniky Fakulty aplikovaných věd Západočeské univerzity.

V neuroinformatické laboratoři na KIV se nachází mnoho přístrojů, které je možno použít pro změření aspektů lidského zdraví a celkového tělesného stavu. V rámci projektu „Exercise and Wellness Health Strategy Framework“ měřená osoba projde několika stanovišti, přičemž každé je zaměřeno na konkrétní část komplexního testu a používá jiné měřicí zařízení. Průběh měření je znázorněn na Obr. 1.

Hlavními důvody pro tvorbu mobilní aplikace jsou urychlení procesu měření, možnosti měření v místech bez počítačové techniky a zvýšení uživatelské přívětivosti.



Obrázek 1: Průběh měření

¹ student bakalářského studijního programu Inženýrská informatika, obor Informatika, e-mail: bohmann@students.zcu.cz

3 Mobilní aplikace

Cílem bakalářské práce [Bohmann (2017)] bylo vytvořit aplikaci, která bude umožňovat zadávání naměřených medicínských dat, zobrazování již naměřených dat, registraci nového uživatele a vyplnění dotazníkové sekce.

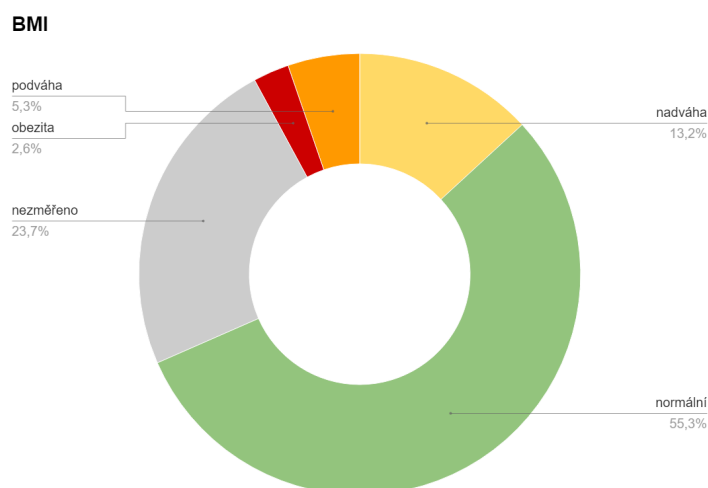
Kromě těchto splněných požadavků aplikace dokáže naskenovat QR kódy, což výrazně urychluje proces zadávání dat. Aplikace nabízí podporu více jazyků, zatím se jedná o jazyky angličtina, němčina a čeština.

Aplikace je vyvíjena v Apache Cordova s využitím frameworku Ionic postaveném na základě AngularJS. Tvorba tedy zahrnuje použití JavaScriptu, HTML a CSS.

Zásadní výhodou tohoto přístupu je možnost multiplatformního nasazení jednou napsané aplikace, a to jak na Android a iOS, tak i na Windows Phone. Aplikace může být nasazena na více než 99% mobilních zařízení.

4 Dosažené výsledky

Mobilní aplikace byla v průběhu vývoje v roce 2017 testována na Dni otevřených dveří FAV a několika projektových dnech. Následující graf (Obr. 2) znázorňuje rozložení BMI u měřených osob na DOD FAV. Alarmující je, že i přes nízký průměrný věk osob (méně než 20 let) více než pětina nemá normální BMI. Téměř čtvrtina účastníků odmítla toto měření absolvovat.

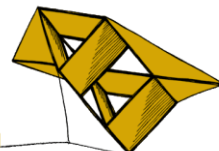


Obrázek 2: BMI osob změřených v rámci DOD FAV 2017

Literatura

Bohmann, David (2017) *Mobilní aplikace pro sběr medicínských dat*. Fakulta aplikovaných věd, Katedra informatiky a výpočetní techniky, Západočeská univerzita v Plzni, Plzeň.

Brůha, P., Mouček, R., Šnejdar, P., Bohmann, D., Kraft, V., Rehor, P. (2017) Exercise and Wellness Health Strategy Framework - Software Prototype for Rapid Collection and Storage of Heterogeneous Health Related Data. *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies HEALTHINF, (BIOSTEC 2017)*, Volume 5, pp. 477-483, SciTePress, Setúbal



Návrh a implementace deskové hry pro kognitivní trénink

Michal Horký¹

1 Úvod

Bakalářská práce popisuje výběr, návrh a implementaci deskové hry pro kognitivní trénink. Kognitivním tréninkem je myšleno cvičení kognitivních funkcí, mezi které patří pozornost, paměť, logické myšlení, řeč a prostorová orientace. Mého zadání se ovšem týká hlavně schopnost soustředění a meditace (zklidnění mysli). Pro získání těchto údajů byla použita EEG čelenka NeuroSky MindWave Mobile.

Čelenka obsahuje NeuroSky ThinkGear technologii, která umožňuje zařízení měřit mozkové vlny uživatele. Zahrnuje senzor s EEG elektrodami, kontaktní bod v ušním klípu (uzemnění) a čip, který zpracovává naměřená data. Senzor s elektrodami se přikládá na čelo uživatele, kde se zachycují elektrické potenciály způsobené mozkovou aktivitou² ve frontální části mozku. V ní probíhá většina našich vědomých myšlenek a rozhodnutí. Napětí naměřené na lidské pokožce je velmi malé a rušené. Čip má tedy za úkol zesílit signál a odstranit z něj rušení v podobě svalových pohybů. Na takto upravený signál se aplikuje eSense algoritmus od stejné firmy, který zjišťuje rozsah aktivity a vypočítává výsledné hodnoty soustředění a meditace. Některé části eSense algoritmu se dynamicky učí a přizpůsobují se přirozenému kolísání a vývoji aktivity uživatele.

2 Výběr a úprava hry

Rozhodl jsem se pro implementaci hry Lodě s vlastní modifikací. Hra je upravená tak, aby hráče co nejvíce bavila a aby byla schopna využít zmíněnou čelenku. Pomocí určité úrovně soustředění hráč může ovlivnit průběh hry. Proto do hry byly přidány herní akce bombardování a oprava lodě. Dále byla přidána chráněná políčka, jejichž zasažením vznikne mezi hráči tzv. duel. Byly provedeny i jiné úpravy, ale ty nevyužívají eSense hodnoty.

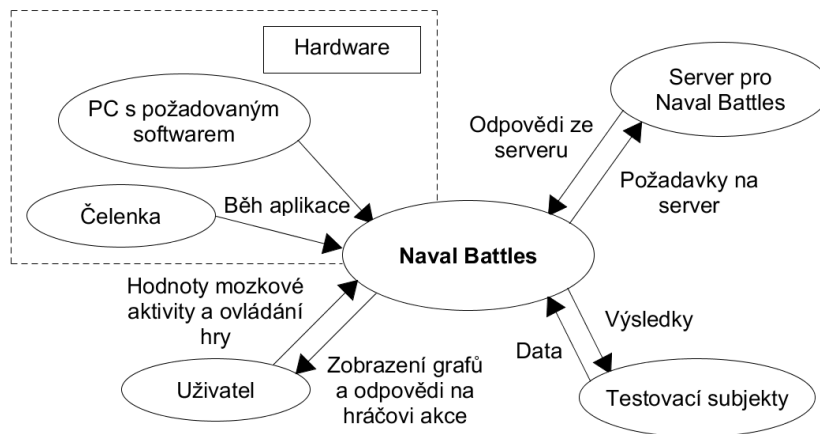
Bombardování je herní akce, kterou může hráč využít pouze při dosažení určité úrovně soustředění, kterou udrží na vybrané nebo vyšší úrovni po určitou dobu. Hráč pak určí, do které oblasti chce provést bombardování. Zvolená oblast je kompletně zničena včetně všech chráněných políček. Opravu může naopak hráč využít při určité úrovni meditace. Hráč si, kromě opravy lodě, může upravit pozice nezasažených lodí a pozice chráněných políček.

Duel je stav hry, při kterém se určitou dobu snímá buď soustředění, nebo meditace obou hráčů. Po uplynutí určité doby je určen lepší z nich. Lepším hráčem je ten, který byl lepší po delší dobu. Pokud byl lepší útočník, pole, na které střílel, se zničí. Jestliže byl lepší obránce, pole si zachránil a navíc zůstává nadále chráněné. Pokud se jednalo o vyrovnaný souboj, určí se průměrná hodnota a lehce se zvýhodní útočník. Pokud totiž i průměrná hodnota bude stejná, vyhrává útočící hráč.

¹ student bakalářského studijního programu Inženýrská informatika, obor Informatika, e-mail: horkmi@students.zcu.cz

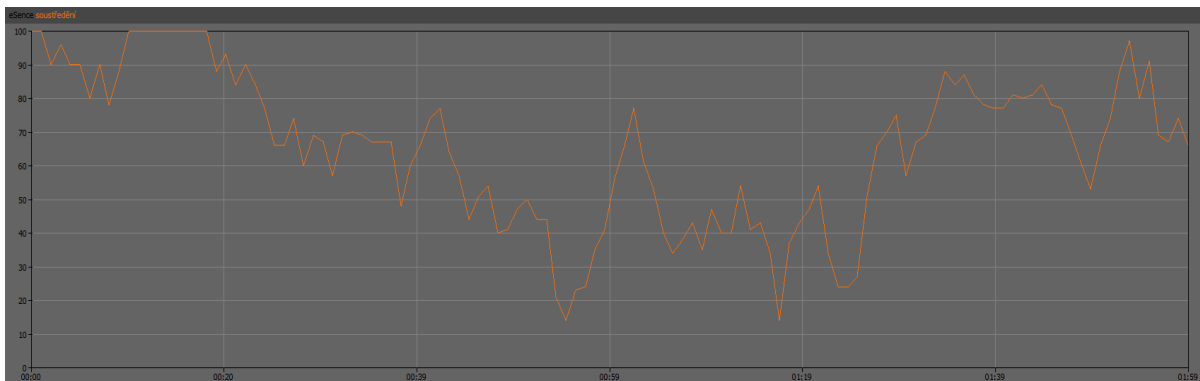
² přenos elektrických signálů mezi neurony

3 Výsledky



Obrázek 1: Kontextový diagram práce

Celá práce se skládá ze dvou aplikací. Jedná se o aplikaci klienta a serveru. Server zajišťuje komunikaci mezi klienty, uchovává nastavení her a řídí jejich průběh. Do aplikace klienta byl, během vývoje, přidělán modul pro průzkum mozkové aktivity. Pomocí něho bylo provedeno měření u šestnácti subjektů. Měření každého subjektu trvalo celkem čtyři minuty. Níže je uveden graf, na kterém lze pozorovat změny v okolí, ve kterém probíhalo měření. Průzkum mozkové aktivity byl do aplikace přidán kvůli hře proti počítači, kterou plánují dodělat v budoucnu. Hráč si vybere obtížnost v podobě věkové kategorie a počítači se z naměřených dat určí průměrná křivka soustředění a meditace.

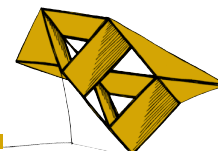


Obrázek 2: Ukázka soustředění jednoho měřeného subjektu

Prvních dvacet sekund probíhalo měření v tiché místnosti. Poté do místnosti přišli další lidé, jejichž přítomnost způsobila zhoršení soustředění měřeného subjektu. Přibližně po minutě odešli a soustředění subjektu začalo opět narůstat.

Zdroje

Developer Toolkit | Learn | Exchange | Publish | NeuroSky Developer Program [online]. Copyright © 2017 NeuroSky [cit. 19.05.2017]. Dostupné z: http://developer.neurosky.com/docs/doku.php?id=thinkgear_communications_protocol.



Poker Cards Recognition using Neural Networks

Jan Hrdlička¹, Miroslav Hlaváč²

1 Introduction

This problem of designing an artificial intelligence for playing poker consists of several different tasks. One of the tasks is automatic cards recognition and identification based on image captured by a video camera. In modern tournaments, the cards are usually put on a transparent material face down, where only the dealer can see the actual front side. The purpose of this work was to choose and create methods for automatic segmentation and identification of the cards color and values.

2 Data Recording and Cards Segmentation

The images of the cards were captured by several different devices (camera, mobile phone, etc.) in different conditions and lighting. Two different sets of cards were used for the recordings. To simulate the conditions of the tournament placement, the cards were randomly put on black background without overlapping. This allowed us to use some of the common computer vision techniques to segment the white cards from the black background based on adaptive thresholding and region detection. The results can be seen in Figure 1. The cards were then rotated and reshaped, and a dataset was created for the purpose of training a neural network. The cards were labeled by manually selecting a template for each type of a card and then using Template Matching to find cards similar to the templates.

3 Convolutional Neural network

The network architecture we constructed for the purpose of this work was based on experience with networks used for ImageNet competition. These networks are specifically designed for the task of classification of images. Since our network only classifies cards and no other images, the architecture is less complex than networks used for ImageNet like Alexnet (Krizhevsky (2012)) or Segnet (Badrinarayanan (2015)). The details of the tested architecture are shown in the results Table 1. The convolutional layers used had kernel size 32×32 with stride one and maxpooling layers had kernel size 2×2 with stride two. The fully connected layer at the end of each network had 52 neurons (same as the number of different cards). The *fc1* in the second network had 4096 neurons.

¹ student of Bachelors degree program Computer Science and Engineering, field Computer Control of Systems and Processes, e-mail: janhrdli@students.zcu.cz

² student of the PhD program Applied Sciences and Informatics, Department of Cybernetics, Computer Vision specialization & PhD student at ITMO, University, Saint Petersburg, e-mail: mhlavac@kky.zcu.cz

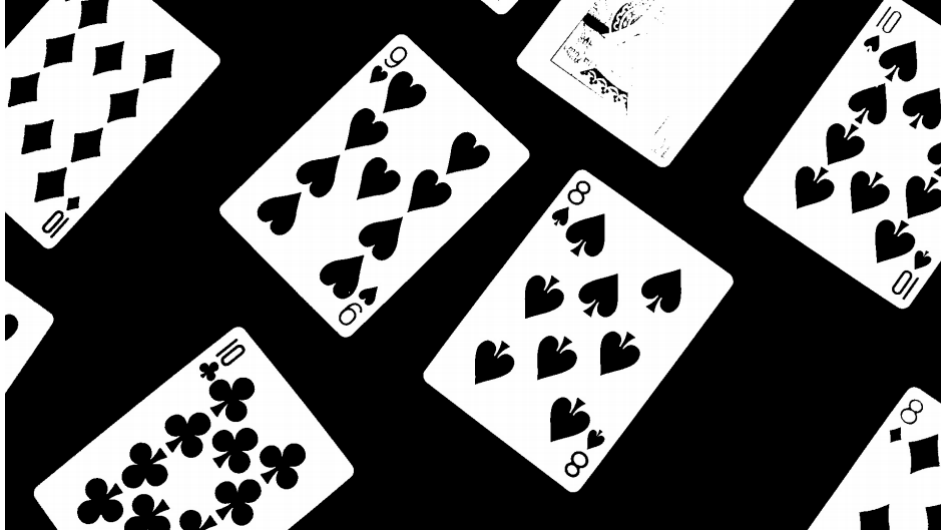


Figure 1: Cards segmentation by thresholding.

4 Results

We developed an automatic method for poker cards segmentation from video file and created several network architectures to classify the cards. The two best architectures are presented here in Table 1. For comparison, the accuracy of the Template Matching method was measured as 81.8 % on the same data. The networks proved to be faster and more accurate than Template matching and thus viable for the development of the next step of automatic poker playing machine.

Table 1: Two developed network architectures with corresponding results.

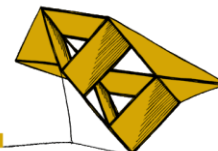
Network architectures	Classification accuracy
conv1,pool,conv2,pool,conv3,pool,fc1	92.21%
conv1,pool,conv2,pool,fc1,fc2	92.67%

Acknowledgement

This work is supported by grant of the University of West Bohemia, project No. SGS-2016-039. Computational resources were supplied by the Ministry of Education, Youth and Sports of the Czech Republic under the Projects CESNET (Project No. LM2015042) and CERIT-Scientific Cloud (Project No. LM2015085) provided within the program Projects of Large Research, Development and Innovations Infrastructures.

References

- Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *In Advances in neural information processing systems (pp. 1097-1105)*.
- Badrinarayanan, V., Kendall, A., Cipolla, R. (2015). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *arXiv preprint arXiv:1511.00561*.



Detekce přízvuků v ruštině s použitím klasifikátoru

Anastasiia Chizhova¹

Úvod

V současné době se naše katedra zabývá zpracováním řeči, jejím rozpoznáváním a syntézou. Při převodu textu do syntetizované řeči se setkáme s obrovským počtem komplikací, které přímo souvisí s daným jazykem, který právě používáme. Já se zabývám ruštinou, konkrétně slovy, které mají nejednoznačný přízvuk, tzv. homografy. Přízvuk takových slov je závislý na kontextu a je ho třeba různými způsoby syntetizovat. Jako příklad uveďme 2 různé věty se stejnými slovy «году»:

- «в следующем году» (*čes.* *V příštím roce*)
- «к 2010 - му году» (*čes.* *Do roku 2010*).

Pro většinu slov lze používat slovník, ale pro homografy existuje několik případů, protože umístění přízvuku ovlivňuje význam slova. Cílem mého experimentu je predikovat přízvuk pouze na základě textu.

Při zpracování daného problému používám metodu “Učení s učitelem” (Psutka (2016)).

Příprava trénovacích dat

Z velkého počtu novinových textů byl sestaven list nejčastěji používaných slov, z nichž jsem vybrala několik nejednoznačných a pro každé z nich našla 100 vět s různými případy kontextu a přízvuku. Ve větách jsem ručně označila v daném slově přízvuk pomocí apostrofu před přízvučnou samohláskou. Dále jsem to musela klasifikovat do dvou tříd:

- je přízvuk (1)
- není přízvuk (0)

Nejprve jsem pro každou samohlásku daného slova vytvořila řetězec s L předchozími a P následujícími znaky a k němu přiřadila třídu 0 nebo 1 (viz. Obr. 1).

```
vшем_году--- 0
ем_году----- 1
ющем_году--- 0
ем_году----- 1
-_му_году--- 1
му_году----- 0
2017_году--- 1
17_году----- 0
2008_году--- 1
08_году----- 0
```

Obrázek 1 Výstupní soubor s označením klasifikace pro velikost kontextu zleva L-6 a zprava P-5

¹ studentka bakalářského studijního programu Inženýrská informatika, obor Inteligentní komunikace člověk - stroj, e-mail: chizhova@students.zcu.cz

Klasifikátory a klasifikace

Řetězce znaků jsem převedla na číselné vektory pomocí metody DictVectorizer z balíčku scikit-learn (Pedregosa *et al.* (2011)). Každá položka vektoru tak odpovídala výskytu konkrétního znaku na konkrétní pozici v řetězci. Zkoušela jsem 2 klasifikátory: Logistic Regression (LogReg) a Support Vector Machine (SVM), kde pro SVM jsem zkoušela 2 různá nastavení: rbf a linear. Proto, abych získala co nejpřesněji úspěšnost klasifikátoru, použila jsem křížovou validaci pomocí Leave-One-Out, což funguje tak, že ze všech dat vždy vynechá jednu položku na testování a ze zbytku se klasifikátor natrénuje, přičemž se použijí všechny možné kombinace trénovacích a testovacích dat. Testovala jsem různě dlouhé levé a pravé kontexty. Výsledky klasifikace jsem porovnávala se správnými odpověďmi pomocí funkce `f1_score` (viz. Tab. 1).

Typ klasif.	LogReg			SVM		
Kontext	L-20 P-5	L-10 P-10	L-5 P-8	L-20 P-5	L-10 P-10	L-5 P-8
Году	86,57%	86,57%	84,73%	88,89%	89,00%	80,38%
Слова	88,00%	87,44%	87,50%	82,83%	85,02%	87,62%
Стороны	86,73%	88,66%	83,72%	86,70%	89,23%	82,57%

Tabulka 1 Výsledky klasifikátoru LogReg a SVM(linear) (`f1_score`)

Závěr

Z výsledků experimentu je zřejmé, že je možné pouze z textového okolí slova s nejednoznačným přízvukem s relativně vysokou úspěšností určit pozici přízvuků ve slově. Lepším z používaných klasifikátorů byl LogisticRegression, který pro všechna testovaná slova udává úspěšnost od 63,03% až do 88,66%. Maximální úspěšnost klasifikátoru SVM(rbf) je 80,73% a SVM(linear) je 89,23%. A pro většinu slov platí, že důležitější je levý kontext. Dále jsem už zkusila zvětšit počet trénovacích dat, což ve většině případů zvýší úspěšnost klasifikátoru, a mám v plánu přepočítat úspěšnost klasifikátoru pro celá slova.

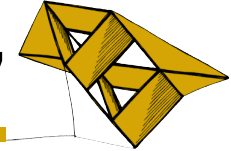
Poděkování

Příspěvek byl podpořen projektem Ministerstva školství, mládeže a tělovýchovy číslo LO1506.

Literatura

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., a Duchesnay, E. (2011) Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research*, 12, pp. 2825-2830.

Psutka, J. (2016) Učební texty z předmětu Základy strojového učení a rozpoznávání, ZČU v Plzni.



Analýza pohybu RC autíčka pro návrh autonomního řízení

Jakub Matoušek¹

1 Úvod

Autonomní řízení osobních i nákladních automobilů v současné době přitahuje čím dál větší pozornost. Velké společnosti jako Google nebo Tesla investují nemalé prostředky na vývoj těchto autonomních řídicích systémů. Za zjednodušenou úlohu zabývající se tímto problémem může být uvažováno řízení modelu RC autíčka bez lidského zásahu. Pro návrh takového řízení je nejprve nutné ověřit model pohybu autíčka.

Cílem této práce bylo tedy ověření pohybového modelu RC autíčka a vytvoření základního grafického rozhraní s možností ovládní pohybu autíčka. Práce byla navíc rozšířena o úlohu detekce kolize autíčka s okrajem dráhy při znalosti jejího tvaru.

2 Řešení

Zjednodušený model pohybu autíčka lze najít například v práci od Rajesh Rajamani (2012). Tento model popisuje polohu autíčka, úhel natočení v rovině a úhel skluzu autíčka. Použitím Eulerovy dopředné metody diskretizace lze pohyb autíčka v každém časovém okamžiku $k = 0, 1, 2, \dots$ popsat diskrétním stavovým modelem

$$\begin{aligned}
 x_{1,k+1} &= u_{2,k} \cdot \cos(x_{3,k} + \beta_k) \cdot T_s + x_{1,k}, \\
 x_{2,k+1} &= u_{2,k} \cdot \sin(x_{3,k} + \beta_k) \cdot T_s + x_{2,k}, \\
 x_{3,k+1} &= \frac{u_{2,k} \cdot \cos(\beta_k)}{2l} \tan(u_{1,k}) \cdot T_s + x_{3,k}, \\
 \beta_k &= \tan^{-1} \left(\frac{l \tan(u_{1,k})}{2l} \right),
 \end{aligned} \tag{1}$$

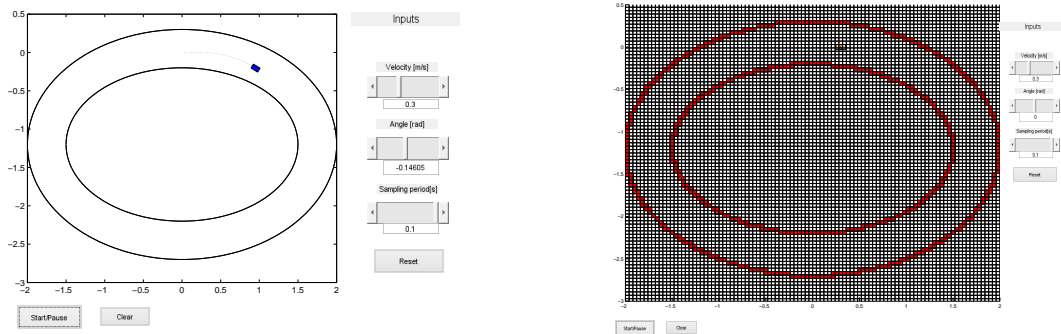
kde $x_{1,k+1}[m]$ je poloha těžiště autíčka v ose x, $x_{2,k+1}[m]$ je poloha těžiště autíčka v ose y, $x_{3,k+1}[rad]$ je úhel natočení autíčka vzhledem k inerciální vztažné soustavě, $\beta_k[rad]$ je úhel skluzu, $l[m]$ je vzdálenost předního a zadního kola, $u_{1,k}[rad]$ je úhel natočení předních kol, $u_{2,k}[m/s]$ je rychlost a $T_s[s]$ je perioda vzorkování. Stav systému je $\mathbf{x}_k = [x_{1,k}, x_{2,k}, x_{3,k}]^T$ a vstup systému je $\mathbf{u}_k = [u_{1,k}, u_{2,k}]^T$. Předpokládá se, že počáteční stav je známý. Pro zpřehlednění dalšího textu je vhodné zavést redukovaný stav $\mathbf{x}_k^p = [x_{1,k}, x_{2,k}]^T$.

Obrázek 1a představuje grafické rozhraní zobrazující pohyb autíčka po dráze ve tvaru elipsy. Pomocí příslušných tlačítek lze simulaci spustit, zastavit, pozastavit, nebo tlačítkem clear resetovat. Další součástí grafického rozhraní jsou posuvníky, pomocí kterých lze za běhu simulace nastavit rychlost autíčka, úhel natočení předních kol, či změnit velikost periody vzorkování. Posuvníky lze tlačítkem reset vrátit do původních poloh.

¹ student bakalářského studijního programu Aplikované vědy a informatika, obor Kybernetika a řídicí technika, e-mail: matoujak@students.zcu.cz

Pro další účely řešení úlohy návrhu autonomního řízení může být užitečné, použitím vhodné agregační funkce, vytvořit konečnou diskrétní množinu stavu polohy \mathbf{x}_k^p . Taková množina odpovídá $\mathcal{X}^p = \{\bar{x}_1^p, \bar{x}_2^p, \dots, \bar{x}_N^p\}$, kde \bar{x}_i^p je diskrétní stav polohy autíčka a N je počet všech možných diskrétních stavů polohy. Ilustrace množiny \mathcal{X}^p v podobě jednotné mřížky je zobrazena na obrázku 1b.

Detekce kolize slouží k odhalení kontaktu autíčka s okrajem dráhy. Za předpokladu vhodné matematické reprezentace tvaru dráhy a znalosti stavového modelu pohybu autíčka (1) lze různými způsoby ověřit, zda autíčko vyjelo z dráhy. V práci byla tato problematika řešena různými způsoby jak pro spojité, tak diskrétní stav polohy autíčka.



(a) Gr. rozhraní pro spojitou polohu autíčka

(b) Gr. rozhraní pro diskrétní stav pohybu autíčka

Obrázek 1: Grafická rozhraní

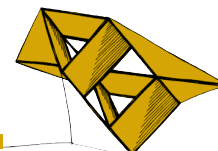
3 Závěr

V práci bylo analyzováno chování a pohyb RC autíčka při změně rychlosti, úhlu natočení přených kol a periody vzorkování. K vizualizaci pohybu po dráze posloužilo grafické rozhraní.

Další vývoj práce by měl směřovat k návrhu řízení pro tento model autíčka bez apriorní znalosti tvaru dráhy, tedy bez znalosti okolí. Toto řízení by mělo minimalizovat vhodně zvolené návrhové kritérium, ohodnocující pohyb autíčka na dráze. Budoucí cíle zároveň předpokládají implementaci algoritmů na reálný model RC autíčka.

Literatura

Rajesh Rajamani (2012) Vehicle Dynamics and Control *Mechanical Engineering Series*. Springer, Heidelberg, Germany chapter 2.



System pro správu publikací

Ondřej Pittl¹

1 Úvod

Související bakalářská práce si kladla za cíl vytvoření webové aplikace sdružující a spravující výsledky publikační činnosti vědců zvolených vědeckých pracovišť². Informace, se kterými aplikace pracuje, jsou získávány externě z bibliografických služeb a jsou aplikací uchovávány v databázi. Společně s vypočtenými statistikami jsou prezentovány uživateli formou seznamu autorů a seznamu jejich publikací. Publikace jsou navíc formátovány dle citační normy ČSN ISO 690:2011 (Firstová 2011) nebo BiBTeX. Obsah obou seznamů je filtrován, řazen a umožněn je i jeho export do souboru ve formátu CSV či TXT.

2 Volba dat a jejich získání

Poskytované informace nebývají napříč bibliografickými službami konzistentní, neboť každá z nich poskytuje jinou kolekci dat. Ta bývá obvykle doprovázena statistickými údaji, zejména počtem citací publikace a metrikami ohodnocujícími míru vědeckého přínosu autora, například h-index, viz Hirsh (2005), g-index, viz Egghe (2006), a i-index, viz Connor (2011).

Data, s nimiž tento systém pro správu publikací pracuje, jsou získávána dolováním webových stránek zvolených bibliografických služeb. Výběr služeb je výsledkem komplexní rešerše jejich sedmi významných zástupců³. Sledovaným faktorem rešerše byla nejen kvalita a kvantita poskytovaných informací, ale také jejich dostupnost a způsob, jakým jsou data službou indexována. Časté jsou dva druhy indexace a to manuální, kterou provádí člověk, a automatická, kterou zajišťuje autonomní algoritmus. Automaticky indexující služby obvykle disponují objemnou databází informací, zatímco služby využívající manuální indexaci se pyšní spíše důvěryhodností a správností poskytovaných informací. Pro další zpracování byly zvoleny služby obojího typu, konkrétně se jedná o automaticky indexující službu GoogleScholar⁴ a manuálně indexující služby ResearchGate⁵ a DBLP⁶.

Všechny zahrnuté služby jsou dolovány vyvinutým obecným parserem, který se řídí popisy jednotlivých služeb uvedených v konfiguračním souboru ve formátu XML. Získaná data z webových stránek daných služeb jsou uchovávána v MySQL databázi, jejíž struktura byla navržena podle informací poskytovaných službami. Tato aplikace indexuje především bibliografické údaje vztahující se k publikaci (například jméno autora či autorů publikace, její název, rok a místo vydání, ISSN a/nebo ISBN atd.) a také afiliace autorů. Nad vydolovanými daty jsou například počítány hodnoty základních metrik ohodnocujících autory, konkrétně h-index,

¹ Student bakalářského studijního programu Inženýrská informatika, obor Informatika, e-mail: pittl@students.zcu.cz. Vedoucí práce: Ing. Michal Nykl, Ph.D.

² Příkladem takového pracoviště může být Katedra informatiky a výpočetní techniky ZČU.

³ BibSonomy, CiteSeerX, DBLP, Google Scholar, Mendeley, ResearchGate a Web of Science.

⁴ Oficiální stránky: <https://scholar.google.com>

⁵ Oficiální stránky: <https://www.researchgate.net>

⁶ Oficiální stránky: <http://dblp.uni-trier.de>

g-index, i5-index a i10-index.

Jak jsem již uvedl výše, informace poskytované službami nebývají konzistentní, například i z důvodu nepřesností způsobených chybami indexace či uvedením odlišných informací (např. užitím zkratk, vynecháním slova apod.). Tento systém s nepatrnými odchylkami v názvech publikací počítá a na základě podobnosti (tj. procentuální shody) názvů publikací, množiny jejich autorů a roků vydání sdružuje příslušné záznamy publikací v databázi. Spoléhá se však na korektnost uvedených jmen autorů.

3 Vizualizace výsledků

Uchovávané informace jsou strukturovaně prezentovány uživateli formou seznamu autorů a seznamu publikací, které jsou doplněny o hodnoty vypočítaných metrik. Výpis publikací je navíc formátován v souladu s citační normou ČSN ISO 690:2011 nebo BiBTeX. Oba seznamy lze řadit a také filtrovat nastavením kritérií filtrů. Obsah filtrovaných a seřazených seznamů je možné exportovat do souboru ve formátu CSV či TXT. Do exportovaného výstupu je umožněno zahrnout rozšiřující informace (spočtené metriky, odkaz na publikaci atd.).

4 Závěr

Implementovaný systém pro správu publikací byl navržen s ohledem na snadnou rozšiřitelnost o další bibliografické služby dodržením třívrstvé architektury, využitím obecného parseru, který využívá přehledný konfigurační soubor s definicemi služeb, a správou závislostí nástrojem Composer⁷. Tato aplikace je silně závislá na zvolených bibliografických službách a jakákoliv změna z jejich strany může znemožnit proces dolování informací a tak i použitelnost aplikace.

Frontend aplikace je plně responzivní⁸, přičemž důraz byl kladen na přehlednost aplikace a jednoduchost designu. Proces dolování a následného zpracování informací je uživateli dynamicky vizualizován grafy a textovým logováním v reálném čase. Přístup na administrační stránku, umožňující ovládání tohoto systému, je podmíněn přihlášením do systému.

Vytvořená aplikace je využitelná například pro získávání informací o publikační činnosti vědeckých pracovníků na Západočeské univerzitě, ale může být využita i jinou vědeckou institucí, přičemž je otevřená dalšímu vývoji.

Literatura

Connor, J. (2011) *Google Scholar Citations Open To All* [online]. Google Inc. [cit. 2017/05/13]. Dostupné z: <https://scholar.googleblog.com/2011/11/google-scholar-citations-open-to-all.html>.

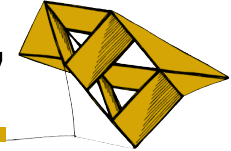
Egghe, L (2006) Theory and practise of the g-index. *Scientometrics*. 69, 1, s. 131–152. ISSN 0138-9130. doi: 10.1007/s11192-006-0144-7. Dostupné z: <https://doi.org/10.1007%2Fs11192-006-0144-7>.

Firstová, Z. (2011) *Nová citační norma ČSN ISO 690:2011* [online]. [cit. 2017-05-13]. Univerzitní knihovna Západočeské univerzity v Plzni. Dostupné z: <https://sites.google.com/site/novaiso690/home>.

Hirsh, J.E. (2005) An Index to Quantify An Individual's Scientific Output. *Proceedings of the National Academy of Sciences*. 102(46), s. 16569–16572. ISSN 1091-6490.

⁷ Oficiální stránky: <https://getcomposer.org>

⁸ Responzivita zajišťuje čitelnost na široké škále zařízení o různých rozlišeních jejich displejů.



Explicitní sémantická analýza

Michal Tušl¹

1 Úvod

Zpracování přirozeného jazyka – NLP (*Natural Language Processing*) je rozvíjejícím se oborem, který je čím dál víc využíván. Používají ho velké společnosti jako jsou Google nebo Facebook ve svých vyhledávačích, zobrazování příspěvků a dalších funkcích. Jádrem tohoto zpracování je sémantika, která se zabývá významem výrazu z různých úrovní jazyka. V tomto případě se tedy jedná hlavně o zpracování a částečné porozumění textu počítačem. V současné době existuje již mnoho metod, které dokážou určit význam textu. Metody trénování bez učitele se učí samy na velkém korpusu dat. Tyto metody jsou založené na distribuční hypotéze, předpokladu, že význam slova lze odvodit z jeho použití (distribuce v textu).

2 Popis metod

Jednou z těchto metod je i metoda ESA (*Explicit Semantic Analysis*) (Gabrilovich a Markovitch, 2009). ESA vytváří vícerozměrné vektory, kde každý prvek vektoru je jeden kontext. Tyto vektory reprezentují sémantiku jednotlivých slov. Na základě vypočtených vektorů pak lze porovnávat podobnost významu slov i souvislých textů. K učení významu potřebuje metoda velké množství dat. K tomu je použita Wikipedie, která obsahuje velké množství článků z různých oborů (kontexty). Výsledek této reprezentace lze využít pro další aplikace jako je například vyhledávání informací, strojový překlad, opravy pravopisu a gramatiky, jazykové modelování pro rozpoznávání řeči, dialogové systémy a další. Metoda ESA používá TF-IDF pro výpočet spojitosti mezi slovem t a kontextem d . Tyto hodnoty ukládá do matice M , kde řádky i představují slova a sloupce j články (kontexty), v kterých se slova vyskytla:

$$M_{i,j} = \text{tf}(t_i, d_j) \cdot \log \frac{n}{df_i}. \quad (1)$$

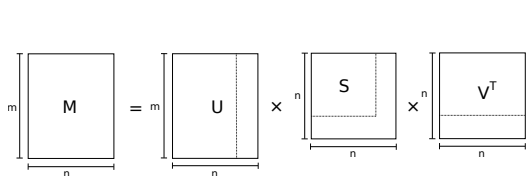
Hodnota tf je definována jako: $1 + \log N(t_i, d_j)$, kde $N(t_i, d_j)$ je počet výskytů slova t_i v dokumentu d_j . Hodnota $df_i = |\{d_k : t_i \in d_k\}|$ představuje počet dokumentů (článků), ve kterých se vyskytuje výraz (slovo) t_i . Sémantická interpretace slova t_i je řádek i z matice M . Význam slova je dán vektorem kontextů spojeným s jejich TF-IDF hodnotami, které odráží spojitost mezi každým tématem a vybraným slovem.

Tuto metodu jsem následně rozšířil o použití kategorií ve Wikipedii jako kontextů. Výpočet míry asociace mezi slovem a kategorií se vypočte pomocí PMI (*Pointwise mutual information*). PMI určuje do jaké míry jsou dva jevy na sobě závislé. V tomto případě tedy do jaké míry je slovo t_i závislé na určité kategorii c_j :

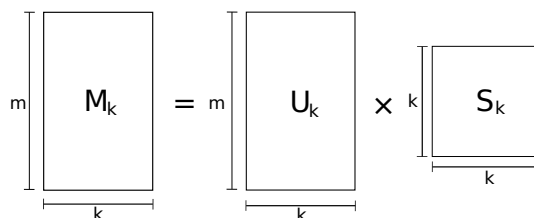
$$\text{pmi}(t, c) = \log \frac{p(t, c)}{p(t)p(c)} = \log \frac{p(t|c)}{p(c)}. \quad (2)$$

¹ student bakalářského studijního programu Inženýrská informatika, obor Informační systémy, e-mail: tuslm@students.zcu.cz

Na výslednou matici této metody byl dále aplikován singulární rozklad (SVD), který redukuje dimezi matice, čímž snižuje paměťové nároky a zároveň zlepšuje výsledky metody. SVD je rozklad matice na tři matice U , S , V , viz Obrázek 1. Pro redukci dimenze je potřeba nejdříve redukovat matice U a S na požadovanou dimenzi k . Aproximace původní matice M s redukovanou dimenzí na k se získá vynásobením matic U_k s maticí S_k , součinem matic pak vznikne matice M_k , viz Obrázek 2. Tato metoda nebyla ještě nikde představena, jedná se tedy o zcela novou metodu nazvanou LSC (*Latent Semantic Categories*).



Obrázek 1: SVD – rozklad matice



Obrázek 2: SVD – redukce dimenze

Singulární rozklad byl též aplikován na metodu ESA, z které takto vznikla metoda LSA (*Latent Semantic Analysis*) (Landauer et al. , 1998). U metod LSC a LSA lze měnit velikost cílové dimenze, proto jsem u nich provedl měření pro různé velikosti dimenze. Nejlépe vyšly výsledky pro dimenzi 2500, které jsou uvedeny v tabulce 1.

3 Výsledky

Všechny prezentované metody byly testovány na datasetech *RG-65* a *WS-353* obsahující lidmi definovanou významovou podobnost párů slov. Shodnost výsledků implementace s dataseť byla určována pomocí Spearmanovy (SC) a Pearsonovy korelace (PC). Výsledky metod jsou uvedeny v tabulce 1. Všechny metody dosáhly dobrých výsledků pro český i anglický jazyk.

	angličtina				čeština				čeština – stemming			
	RG-65		WS353		RG-65		WS353		RG-65		WS353	
Model	PC	SC	PC	SC	PC	SC	PC	SC	PC	SC	PC	SC
ESA	0,57	0,76	0,41	0,51	0,70	0,75	0,32	0,51	0,38	0,83	0,38	0,49
PMI	0,51	0,58	0,49	0,48	0,58	0,62	0,41	0,45	0,47	0,70	0,43	0,44
LSA	0,68	0,72	0,52	0,54	0,65	0,67	0,46	0,44	0,66	0,70	0,49	0,51
LSC	0,63	0,62	0,55	0,56	0,57	0,62	0,43	0,42	0,59	0,59	0,48	0,48

Tabulka 1: Výsledky metod pro anglický a český jazyk.

Literatura

- Gabrilovich, E., Markovitch, S. (2009) Wikipedia-based Semantic Interpretation for Natural Language Processing. *J. Artif. Int. Res.*, pp. 443—498.
- Landauer, T. K., Foltz, P. W., Laham, D. (1998) An Introduction to Latent Semantic Analysis. *Discourse Processes*, pp. 259—284.

Název: SVK FAV 2017– bakalářské studijní programy, sborník rozšířených abstraktů

Editor: Vladimír Lukeš

Autor obálky: Petr Lobaz

Vydavatel: Západočeská univerzita v Plzni, Univerzitní 8, 306 14 Plzeň

Datum vydání: květen 2017

ISBN 978-80-261-0705-7