SBORNÍK

prací účastníků vědecké konference doktorského studia
Fakulta informatiky a statistiky
Vysoké školy ekonomické v Praze

Vědecká konference se uskutečnila dne 12. února 2015
pod záštitou děkana FIS doc. RNDr. Luboše Marka, CSc.

Sestavení sborníku
prof. Ing. Petr Doucek, CSc.
proděkan pro vědu a výzkum

Nakladatelství Oeconomica – Praha 2015

ISBN 978-80-245-2081-0

OBSAH

Předmluva

V roce 2015 se konal další ročník semináře „Den doktorandů“, který byl letos dvacátým. Seminář se konal v polovině února tradičně pod gescí děkana Fakulty informatiky a statistiky. Seminář se stal tradiční platformou, kde doktorandi všech oborů doktorského studia fakulty presentují výsledky své vědecké a odborné práce. Pro mnohé z nich je to první vystoupení před odbornou veřejností, na němž získávají cenné zkušenosti a tříbí tak i formulace názorů a hypotéz. Kromě toho si vyzkouší presentaci závěrů výzkumné práce a argumentaci na jejich podporu. I v letošním roce byly příspěvky rozděleny do sekcí podle studovaných oborů. Do dvacátého ročníku „Dne doktorandů“ se celkem přihlásilo 14 účastníků. Z toho na oboru „Aplikovaná informatika“ to bylo 8 příspěvků, na oboru „Ekonometrie a operační výzkum“ 1 příspěvek a na oboru „Statistika“ 5 příspěvků. Semináře se zúčastnili studenti obou forem studia, jak presenční, tak i kombinované. Vzhledem k malému počtu přihlášených prací v programu „Ekonometrie a operační výzkum“ byla vytvořena jedna sekce pro obory „Ekonometrie a operační výzkum“ a „Statistika“.

Nedílnou součástí „Dne doktorandů“ je i práce hodnotících komisí, jejichž členové pečlivě sledují jednotlivá vystoupení a potom po oborech vybírají nejlepší práce k ocenění. Hlavními kritérii pro jejich rozhodování byly zejména kvalita a aktuálnost zpracovaného tématu, přístup k řešení vybraného problému, způsob použití metodiky, úroveň práce s reálnými daty a v neposlední řadě i schopnost prezentovat a argumentačně své výsledky obhájit v diskusi. Ti nejlepší z nich získávají prestižní „Cenu děkana FIS“, s níž je spojena i symbolická finanční odměna.

Za práci v komisích chci poděkovat všem jejím členům, kteří pracovali pod vedením předsedů - doc. Ing. Vojtěcha Svátka, Dr. (obor Aplikovaná Informatika) a prof. Ing. Hany Řezankové, CSc. (spojená komise pro obory „Ekonometrie a operační výzkum“ a obor „Statistika“). Že se jednalo o nelehkou práci, ukazuje i fakt, že ve všech vyhlášených kategoriích byla udělena dělená třetí místa. Komise se ovšem své úlohy zhostily na výbornou.

V letošním roce získali ceny za nejlepší příspěvky v jednotlivých kategoriích následující studentky a studenti:

Studijní program – Aplikovaná informatika

1. místo: Ing. et Ing. Stanislav Vojíř: Datamining klasifikačních busines rules prostřednictvím systému EasyMiner

2. místo: Ing. Martin Potančok: Sourcingové strategie IS/ICT ve zdravotnických zařízeních

3. místo: Ing. Marek Dudáš: Vizualizace schémat sémantického webu a data setů linked data

Ing. Tomáš Klíma: Projektové řízení penetračních testů IS

Studijní program – Kvantitativní metody v ekonomice

1. místo: Neuděleno

2. místo: Ing. Lenka Fiřtová: Aplikace metaheuristických algoritmů při vytváření jazykových testů

Ing. Zdeněk Šulc: Application of Goodall´s and Lin´s similarity measures in hierarchical clustering

3. místo: Mgr. Ing. Michal Rychnovský, MSc.: Mathematical Models for Loss Given Default Estimation

Ing. Lenka Vraná: Business cycle analysis with composite indicators

Oceněným studentům doktorského studia upřímně blahopřeji a doufám, že získané zkušenosti uplatní při své další práci, ať už vědecké nebo v praxi. Uznání také patří všem vědeckým a pedagogickým pracovníkům FIS – školitelům doktorandů, kteří se „Dne doktorandů“ zúčastnili.

Zvláštní poděkování pak patří studijní referentce doktorského studia paní Jitce Krajíčkové, díky níž byl seminář skvěle organizačně zajištěn, a Mgr. Lee Nedomové za obětavou práci při editaci a sestavení tohoto sborníku. Úplným závěrem bych chtěl poděkovat prof. RNDr. Janu Pelikánovi, CSc., který tuto tradici před dvaceti lety založil.

Prof. Ing. Petr Doucek, CSc,
Proděkan pro vědu a výzkum

Měření customer experience v bankovnictví

Jakub Albrecht

jakub.albrecht@vse.cz

Doktorand oboru Aplikovaná informatika

Školitel: prof. Ing. Jiří Voříšek, CSc., (vorisek@vse.cz)

Abstrakt: Cílem materiálu je představení možností a metod měření customer experience. Výstupem materiálu je desatero doporučení pro měření customer experience, které zároveň reflektuje relevantní specifika bankovního odvětví.

V úvodu článku jsou definovány základní pojmy a vysvětlen význam customer experience pro podniky. Následuje část věnovaná aktuálním metodám měření customer experience, existujícím metrikám a internímu hodnocení zralosti.

Další kapitola popisuje specifika customer experience v bankovnictví.

Výstupem článku je pak desatero doporučení pro měření customer experience v bankovnictví, které doporučuje následující body:

1. sledujte veškeré interakce
2. sledujte chování napříč kanály
3. přizpůsobte formu a obsah situaci
4. motivujte zákazníky
5. motivujte personál
6. vyberte si vhodnou metriku dle vlastního uvážení
7. konejte na základě získané zpětné vazby
8. provažte metriky customer experience na obchodní a finanční metriky
9. sledujte vývoj metrik v čase
10. definujte strategii rozvoje customer experience v podniku

Klíčová slova: customer experience, prožitek zákazníka, měření, metriky, bankovnictví

Úvod

Definice pojmů

Customer experience (CX) je chápána jako komplexní vyjádření celkového prožitku, který má konkrétní zákazník ve vazbě na podnik v průběhu celého jejich společného vztahu v nejširším slova smyslu. Uvažovány jsou přitom nejen objektivně-racionální složky tohoto prožitku, ale také složky subjektivní, emocionální a iracionální, které se významným způsobem na celkovém prožitku zákazníka podílejí. Disciplína, která se zabývá rozvojem a správou customer experience v podniku, se nazývá **customer experience management (CXM)**.

Definice pojmů customer experience a customer experience management se v různých zdrojích liší, pro účely této práce budou použity následující definice:

- *Customer experience* je vnímání a související pocity zákazníka, které vyplývají z jeho interakce se společností (volně přeloženo a upraveno z (ANON. nedatováno).
- *Customer experience* management je návrh a realizace interakcí společnosti se zákazníky takovým způsobem, aby byla naplňována nebo převyšována očekávání zákazníků a tím u nich bylo dosahováno vyšší spokojenosti, loajality a míry ztotožnění se se společností (volně přeloženo a upraveno z (ANON. nedatováno).

Základními společnými rysy většiny definic pojmu customer experience jsou:

- důraz na celkový prožitek klienta zahrnující jak vědomé a racionální aspekty, tak aspekty podvědomé a emocionální;
- zaměření na celý životní cyklus vztahu společnosti a zákazníkem, nikoliv na samostatné dílčí interakce a transakce.

Význam customer experience

Ekonomika a tržní prostředí doznaly zejména v uplynulém desetiletí poměrně revolučních změn, přičemž klíčovými faktory ovlivňujícími vztah společností a jejich zákazníků jsou:

- silná komoditizace produktů a služeb
- saturace trhů, zejména těch vyspělých
- globalizace a stírání geografických hranic trhů
- dostupnost a rychlost přístupu zákazníků k informacím
- rostoucí „síla“ zákazníků samotných zákazníků utvářet image společnosti, což souvisí jednak s faktory výše uvedenými, jednak s nástupem nových paradigmat komunikace a sdílení informací jako jsou sociální sítě a aplikace typu Web 2.0

Vývoj ekonomiky dle (Pine II a Gilmore 2011) ukazuje Obrázek 1. Poukazuje na fakt, že historicky prošla ekonomika různými etapami s různou úrovní nabízené ekonomické hodnoty, přičemž incentivem pro přechod na vyšší úroveň byla vždy komoditizace nabídky na aktuální úrovni a vidina zvýšení profitu cestou customizace.

Obrázek 1: Vývoj ekonomické hodnoty Zdroj: (ANON. nedatováno)

S ohledem na výše uvedené faktory hledají společnosti nové cesty jak diferenciovat svou nabídku. Cesta naslouchání skutečným potřebám zákazníků a budování jedinečné zkušenosti zákazníků během jejich interakce se společností se jeví jako win-win strategie, což dokládají i výsledky zákaznického průzkumu (Anonymous 2010), podle kterého je za jedinečnou zákaznickou zkušenost:

- 55% zákazníků ochotno zaplatit 10% nebo více nad rámec standardní ceny;
- 27% zákazníků ochotno zaplatit 15% nebo více nad rámec standardní ceny;
- 10% zákazníků ochotno zaplatit 25% nebo více nad rámec standardní ceny.

Customer experience přímo ovlivňuje zákaznické veličiny, jako jsou míra uspokojení potřeb, zákaznická loajalita (Mascarenhas et al. 2006) a ztotožnění se zákazníka se společností (anglicky advocacy) (Rabino et al. 2009).

Nejdůležitějšími koncepty disciplíny customer experience management jsou detailní znalost zákazníka a jeho potřeb, naplňování těchto potřeb a vnitřní konzistence společnosti zahrnující i autenticitu jejího projevu.

Měření customer experience

Metriky

Net Promoters Score (NPS)

Notoricky známý ukazatel, jehož konstrukce stojí na klasifikaci odpovědí na otázku, zda by zákazník doporučil podnik a jeho produkty a služby své rodině nebo známým. Čisté skóre se pak vypočítává jako

rozdíl procentuálního zastoupení odpovědí tzv. promotérů (respondentů s odpověďmi typicky 8 bodů a více na desetibodové škále) a tzv. detraktorů (respondentů s odpověďmi typicky 5 bodů a méně).

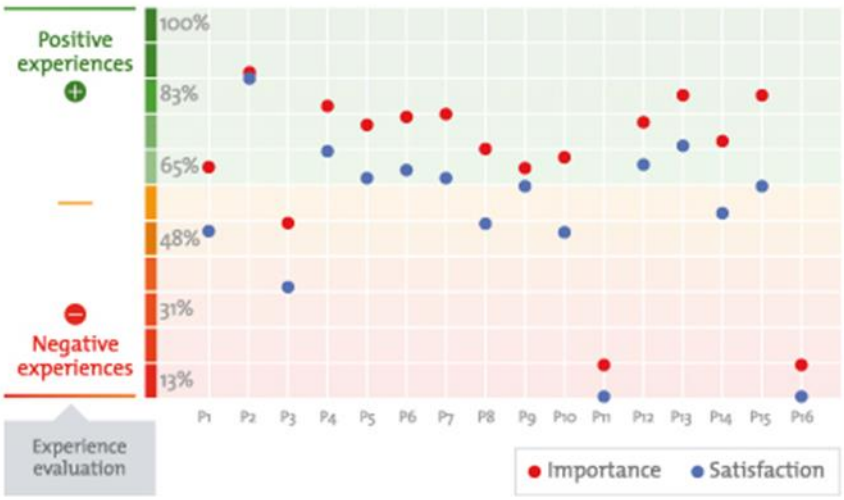
Další „jednootázkové“ metriky

Hodnocení na základě jediné otázky položené hodnotícímu zákazníkovi je celá řada. Lze uvažovat o metrikách vytvořených např. na základě následujících otázek:

- Kolik úsilí jste musel(a) vynaložit k provedení požadované interakce?
  (Customer Effort Score – CES) (Alfaro et al. nedatováno)
- Domníváte se, že podnik jedná ve Vašem nejlepším zájmu, nebo že spíše v zájmu vlastních zisků?
  (Customer Advocacy – CA) (Alfaro et al. nedatováno)
- Naplnila služba Vaše očekávání? (Beard 2013)
- Jak si stojí služba ve srovnání s Vaší představou o ideální službě? (Beard 2013)
- Jak jste celkově spokojen(a) s naším podnikem? (Beard 2013)
- Máte v plánu si znovu zakoupit službu, případně obnovit smlouvu, jakmile vyprší její platnost? (Beard 2013)

Experience Map

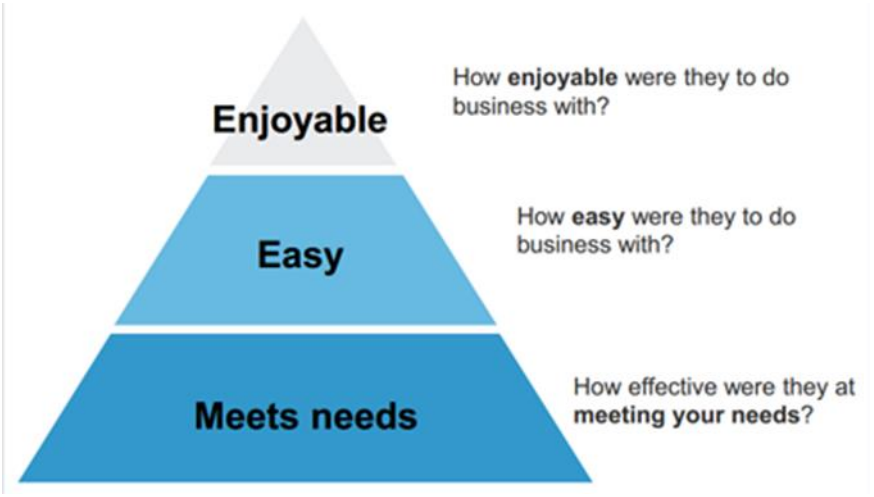
Jednoduchým přístupem, který poměřuje každou jednotlivou interakci zákazníka s podnikem v rovině důležitosti interakce a míry uspokojení zákazníka, je tzv. experience map. Na základě této mapy je možné clusterovat jednotlivé typy interakcí a identifikovat nejpalčivější problémy podniku v oblasti customer experience. Příklad experience map uvádí Obrázek 2.



Obrázek 2: Příklad experience map Zdroj: (Alfaro et al. nedatováno)

Forrester’s Customer Experience Index (CXi)

Forrester’s Customer Experience Index (CXi) je hodnocení úrovně customer experience podniků ve 14 různých odvětvích, které v USA a nověji i v Evropě a Číně pořádá nezávislá výzkumná společnost Forrester. Ta realizuje průzkum oslovením respondentů z řad skutečných zákazníků podniků, přičemž zákazníci se vyjadřují ke každému podniku jednotlivě. Strukturu průzkumu ukazuje Obrázek 3.



Obrázek 3: Skladba Forrester's Customer Experience Index Zdroj: (ANON. nedatováno)

Respondenti odpovídají pro každý jednotlivý podnik na následující 3 otázky, z nichž každá má vlastní škálu odpovědí ohodnocených od 1 do 5 bodů (Burns 2014):

- Pokud budeme uvažovat Vaše poslední interakce s touto společností, jak efektivní byli při uspokojování Vašich potřeb?

  Neuspokojili žádnou z mých potřeb (1) ↔ Uspokojili všechny mé potřeby (5)

- Pokud budeme uvažovat Vaše poslední interakce s touto společností, jak jednoduché bylo realizovat s nimi obchod?

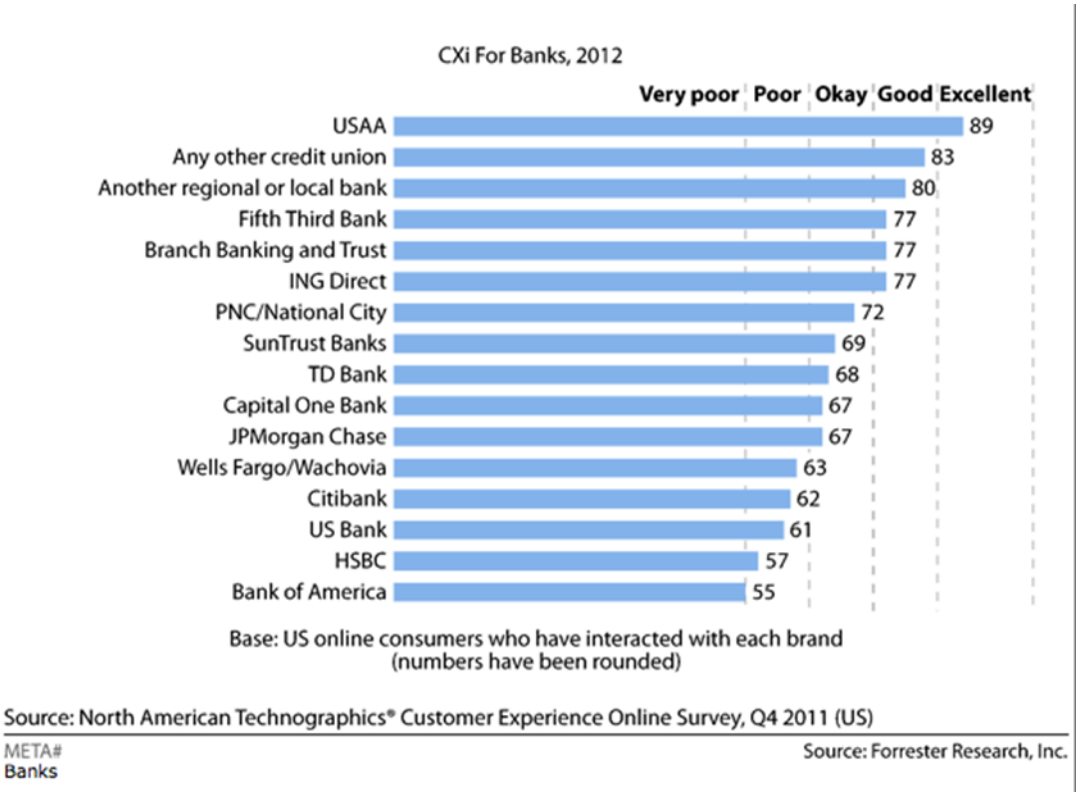
  Velmi obtížné (1) ↔ Velmi jednoduché (5)

- Pokud budeme uvažovat Vaše poslední interakce s touto společností, jak příjemné bylo realizovat s nimi obchod?

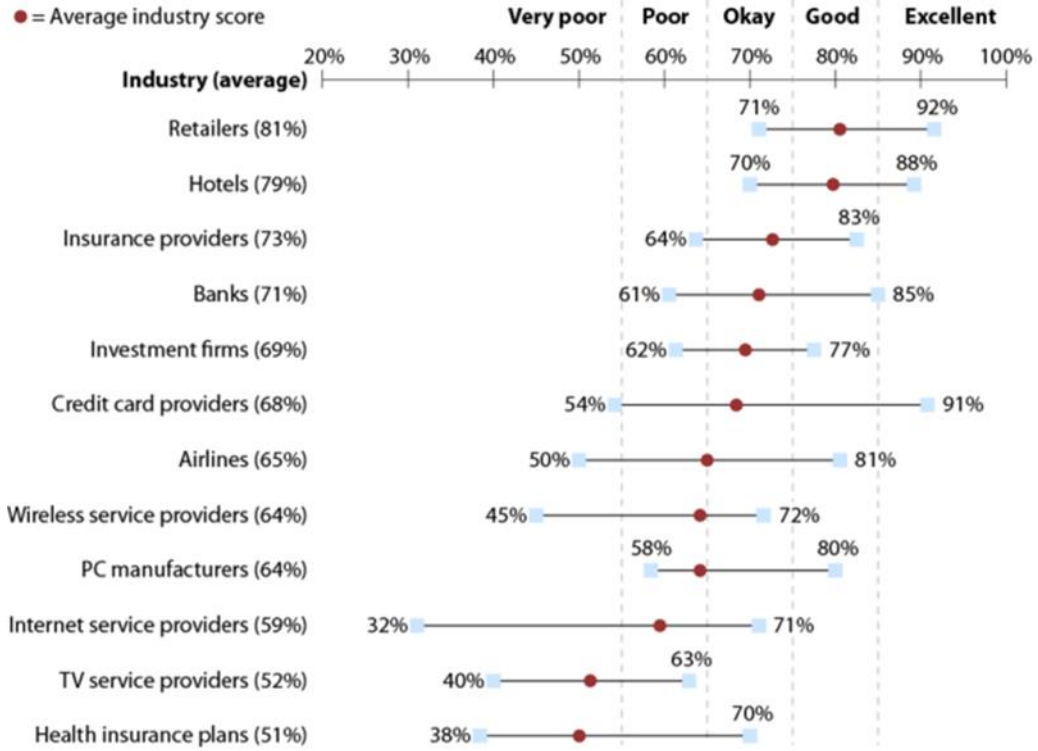
  Vůbec ne příjemné (1) ↔ Velmi příjemné (5)

Pro každou z výše uvedených tří oblastí se vyhodnotí skóre jako rozdíl procentuálního zastoupení odpovědí s bodovým hodnocení 4 nebo 5 a procentuálního zastoupení odpovědí s bodovým hodnocením 2 nebo 1. Výsledné CXi skóre se pak pro konkrétní podnik vypočte jako prostý průměr ze skóre podniku v každé ze tří oblastí. Mezi oficiálně hodnocené podniky se dostanou jen takové, ke kterým se vyjádřilo alespoň 100 respondentů.

Možnou výslednou podobu reportu CXi ukazuje Obrázek 4 a Obrázek 5.



Obrázek 4: Srovnání CXi pro jednotlivé podniky Zdroj: (ANON. nedatováno)



Obrázek 5: Srovnání CXi napříč odvětvími (2008) Zdroj: (ANON. nedatováno)

Interní hodnocení zralosti

Zatímco výše uvedené metriky měří reálnou úroveň customer experience konkrétních zákazníků a měření je tedy z pohledu podniku externího charakteru, je možné měřit také úroveň customer experience podniku interně. V tomto případě jde tedy spíše o zhodnocení zralosti capability customer experience management.

I tento typ hodnocení podporuje společnost Forrester. Na stránce (ANON. nedatováno) je možné vyplnit dotazník obsahující otázky na stav zavádění disciplíny customer experience management v podniku a využívání specifických procesů a artefaktů s disciplínou souvisejících. Výsledkem hodnocení je pak stanovení zralosti a doporučení na další rozvoj podniku, přičemž jsou definovány následující strategie (Burns 2014):

- opravení, tedy „nalezení nejbolavějších míst v oblasti customer experience a jejich náprava“
- povýšení, tedy „ustálení a normalizace přístupu k oblasti customer experience”
- optimalizace, tedy „používání sofistikovanějších metod customer experience
- diferenciace, tedy „přerámování zákaznických problémů, adresování nerealizovaných potřeb, přehodnocení celého ekosystému“

Specifika customer experience v bankovnictví

Specifiky customer experience v bankovním sektoru se zabývá např. (Rabino et al. 2009), (Philipp et al. 2013) a (Anonymous 2006), cennou agregací vývoje customer experience v oblasti retailového bankovnictví je každoroční World Retail Banking Report (ANON. nedatováno), (ANON. nedatováno), (ANON. nedatováno) a (ANON. nedatováno). Nejzásadnějšími odlišnostmi bankovnictví oproti ostatním odvětvím jsou:

- vysoká očekávání klientů v oblasti bezpečnosti, důvěry a respektování soukromí
- důležitost multikanálového přístupu a masivní rozvoj samoobslužných a mobilních kanálů
- vysoká míra regulovanosti odvětví ze strany národních i nadnárodních autorit

Desatero doporučení pro měření customer experience v bankovnictví

1. Sledujte veškeré interakce

Komplexní disciplína vyžaduje komplexní přístup i k měření a řízení. Jelikož definujeme customer experience jako celkový prožitek zákazníka ve vztahu k podniku v průběhu celého jejich společného životního cyklu, je nutné při měření customer experience uvažovat veškeré interakce, které zákazník s podnikem realizuje. Jde tedy o všechny následující typy interakcí:

a. Pre-sales, tedy interakce předcházející realizaci obchodního případu

Tento typ aktivit je v některých případech velmi obtížně měřitelný, neboť např. u zcela nových zákazníků podnik interaguje s protistranou, která je zcela mimo jeho sféru vlastnictví (např. ATL kampaně prezentované neklientům). Přesto je možné úroveň customer experience i v těch případech alespoň základním způsobem měřit, k čemuž mohou posloužit otázky typu „Jaké je Vaše současné povědomí o naší společnosti?“, „Jak byste hodnotil(a) image naší společnosti na základě Vašeho současného povědomí?“ nebo „Odkud jste se o nás dozvěděl(a)?“.

b. Sales, tedy aktivity, které přímo vedou k realizaci obchodního případu

c. After-sales, tedy aktivity následující po realizaci obchodního případu

V případě bankovnictví je tato skupina aktivit ve srovnání s ostatními odvětvími poměrně významná, neboť sem kromě občasných úprav parametrů obchodního případu řadíme i běžné používání zakoupeného produktu nebo služby, tedy standardní transakční chování.

2. Sledujte chování napříč kanály

Specifikem bankovnictví je široké spektrum distribučních kanálů, které banky nabízejí svým zákazníkům. S ohledem na předchozí bod desatera je pochopitelně nezbytné sledovat interakce zákazníka napříč všemi kanály. Multikanálový přístup k obsluze klienta a měření customer experience ale přináší podnikům další přidanou hodnotu:

a. Mezikanálové srovnání a řízení jednotlivých kanálů

Je možné získat parametry hodnocení z pohledu customer experience pro jednotlivé kanály, což podnikům umožní lépe řídit distribuční mix v jednotlivých kanálech i výkonnost jednotlivých kanálů.

b. Profil klienta a jeho preference

Je možné sledovat pro konkrétního zákazníka různé hodnoty customer experience při využívání různých kanálů a z těchto hodnot odvozovat osobní preference klienta a jeho subjektivní profil (např. klient preferující digitální samoobslužnou zónu vs. klient preferující asistenci).

3. Přizpůsobte formu a obsah situaci

S tím, jak v souladu s předchozími body roste poptávka podniků po zpětné vazbě a hodnocení od zákazníků, je nutné redefinovat formu i obsah zákaznických průzkumů. Oba parametry musí být v souladu s tím, kolik času a úsilí je od zákazníka vyžadováno k realizaci interakce. Můžeme rozlišit dva typy situací:

a. Okamžité situační průzkumy (short-loop)

Tento typ průzkumů navazuje na interakce, které jsou běžného operativního charakteru a reflektují primárně spokojenost klienta s průběhem a výsledkem prováděné operace. Hodí se tak převážně pro after-sales interakce, kdy zákazník přichází s jasným očekáváním následné operace a kritériem úspěchu je tak minimalizace času a úsilí, které klient během interakce vydává.

Jelikož interakce samotné jsou operativního charakteru, musí se i v případě průzkumu jednat o extrémně jednoduchý a časově nenáročný proces, který je spíše než průzkum chápán jako závěrečný krok realizované interakce. Optimálně se bude jednat o jedinou hodnotící otázku s hodnotící škálou odpovědí, případně doplněnou a možnost otevřené odpovědi v případě, kdy zákazník volí krajně pozitivní nebo krajně negativní hodnocení.

b. Komplexnější průzkumy (long-loop)

Komplexnější průzkumy jsou vhodné v případech, kdy je hodnocena customer experience:

- interakcí, které jsou zákazníkem chápány jako významné,
- za uplynulý časový horizont.

Komplexnější průzkumy je možné koncipovat jako delší dotazníky sestávající z více otázek, které mohou nabízet více otevřené možnosti odpovědí. Vždy ale důležité zachovat motivaci zákazníka k poskytnutí zpětné vazby prostřednictvím těchto průzkumů (viz následující bod desatera) a obsah i formu navrhovat tak, aby odpovídaly této motivaci.

4. Motivujte zákazníky

Měření customer experience není možné realizovat bez účasti skutečných zákazníků. Jelikož úroveň customer experience konkrétního zákazníka je vždy subjektivní veličinou, neobejdeme se při měření bez nutnosti aktivního kroku sdílení vlastní customer experience ze strany samotného zákazníka.

Abychom motivovali zákazníky k poskytování zpětné vazby, bez které není možné customer experience měřit, je možné využívat následující incentivy:

a. přímá odměna za poskytnutí zpětné vazby

Za poskytnutí zpětné vazby je možné zákazníka bezprostředně odměnit způsoby, které podniky běžně využívají v rámci svého standardního marketingu – slevou na vlastní produkty a služby, povýšením úrovně zákaznické péče, aj. V případě bank, které typicky disponují komplexními věrnostními programy, se nabízí propojení systému odměn za hodnocení customer experience právě s těmito věrnostními programy.

Výhodou tohoto způsobu odměňování je přímočarost a jednoduchost takovéhoto schémata. Nevýhodou může být „neupřímnost“ zákazníků, kdy zákazníci vyplňují hodnocení zběžně a bez skutečného zájmu o věc pouze proto, aby získali protihodnotu.

b. zapojení zákazníka do rozvoje společnosti

Pro specifický segment klientů může být výrazně vhodnější takový způsob motivace, kdy jim podnik prostřednictvím hodnocení customer experience dává možnost podílet se na směřování a řízení podniku. Kritickým faktorem úspěchu tohoto přístupu je schopnost dostatečně rychlé dodávky poptávaných změn a trasovatelnost mezi vstupními požadavky pocházejícími ze zákaznické zpětné vazby a konkrétními dodávanými změnami. Odměnou pro podnik, kterému se ve spolupráci s vlastními zákazníky podaří nastavit takovouto platformu vzájemné důvěry, je velmi silná a upřímná zpětná vazba od zákazníků a výrazné zvýšení jejich loajality a ztotožnění se s podnikem.

Velikost protihodnoty, kterou zákazník získává za poskytnutí vlastní zpětné vazby, by měla odpovídat časové náročnosti a množství úsilí, které k tomu zákazník potřeboval.

5. Motivujte personál

Implementace strategie maximalizující customer experience v podniku se neobejde bez fundamentální participace měkké lidské složky, což dokládá např. i (Maritz 2010), kde jsou „customer insight" týmy a manažeři spolu se zaměstnanci považováni za dva ze tří klíčových stakeholderů customer experience strategie, přičemž třetím klíčovým stakeholderem je zákazník samotný.

Pokud chce být podnik v oblasti customer experience úspěšný, musí dokázat správně motivovat a uspokojit následující skupiny lidí ve vlastních řadách, přičemž motivace a vedení lidí v jednotlivých kategorií a použitý způsob komunikace se bude mezi jednotlivými kategorie navzájem velmi lišit. Jde o následující kategorie zaměstnanců:

a. „Customer insight" týmy

Členové těchto týmů jsou přímo zodpovědní za definici a implementaci strategie maximalizace customer experience od obecné abstraktní úrovně až na úroveň znalosti a naplnění potřeb konkrétních segmentů zákazníků a nabídky podniku reflektující tyto potřeby. Je naprosto nezbytné, aby členy týmu byli odborníci vysoce specializovaní na oblast customer experience, ale zároveň lidé neodtržení od reality a reálné praxe, ve které se podnik i zákazník vyskytují. Nedílnou součástí práce těchto týmů je i inicializace změn motivovaných získanou zpětnou vazbou od zákazníků.

Jelikož úspěšná aplikace prozákaznického přístupu a strategie maximalizující customer experience vyžaduje fundamentální podporu celého podniku, neobejdou se „customer insight" týmy bez blízké spolupráce a podpory ze strany vrcholového managementu i HR útvarů.

„Customer insight" týmy definují „operativní" metriky úspěšnosti customer experience strategie, ale jejich vlastními KPI je dosahování požadované úrovně obchodních a finančních metrik, které jsou skutečným a hmatatelným benefitem podniku.

b. Front-line zaměstnanci a jejich liniový manažeři

V případě asistovaných kanálů je front-line personál přímým ztělesněním podniku v očích zákazníka při vzájemné interakci. Je proto bezpodmínečně nutné, aby všichni zaměstnanci znali, chápali, respektovali strategii podniku maximalizující customer experience a byli s ní ztotožněni, nebo alespoň motivovaní k jejímu dodržování a naplňování. Při motivaci, vedení a odměňování front-line personálu hraje klíčovou roli jejich liniový management, a proto musí být i tito manažeři shodným způsobem instruování a motivováni. Obdobným způsobem lze nahlížet na designéry přímých kanálů, kteří do značné míry suplují práci front-line personálu.

Front-line personál může vedle reprezentování podniku vykonávat ještě jednu důležitou úlohu. Může být prostředníkem sběru zpětné vazby od zákazníka a zároveň autorem cenných připomínek k podobě vztahu mezi podnikem a zákazníkem přímo „v terénu". Oba typy zpětné vazby jsou pro podnik nesmírně cenné a je proto zapotřebí vytvořit přirozenou a zdravou atmosféru, která zároveň front-line zaměstnance motivuje ke sběru a sdílení obou forem zpětné vazby.

c. Vrcholový management

Vrcholový management měří svět i podnik výrazně tvrdšími metrikami, než s jakými pracuje disciplína customer experience. Pro získání dostatečné podpory strategie maximalizující customer experience je

proto nutné poskytnou vrcholovému managementu důkazy o tom, že jde o správnou cestu poskytující podniku reálné a hmatatelné benefity. V tomto ohledu je klíčové prezentovat managementu skutečné přínosy ve formě zlepšení tvrdých finančně-ekonomických ukazatelů. Zlepšování customer experience je nekonečný proces, ale vrcholový management bude chtít alespoň dílčí výsledky už v krátkém časovém horizontu.

d. Všichni zaměstnanci podniku bez výjimky

Orientace na zákazníka a maximalizace jeho customer experience je pro podnik ve srovnání s dnes typickou produktově orientovanou a z pohledu podniku sebestřednou nabídkou fundamentální proměnou, při které je klíčovým faktorem úspěchu změna přístupu, myšlenkového nastavení i chování všech zaměstnanců promítající se do zcela nové podnikové kultury a nových požadavků na zaměstnance na všech úrovních. Tuto novou kulturu vyžadující dříve nebývalou zodpovědnost a angažovanost personálu je zapotřebí systematicky budovat a zaměstnance aktivním a pozitivním způsobem motivovat k tomu, aby prozákaznickou kulturu chápali, respektovali, aplikovali a sami rozvíjeli.

6. Vyberte si vhodnou metriku dle vlastního uvážení

Kapitola Metriky výše poskytuje ukázku metrik, které je možné využít k měření úrovně customer experience zákazníků. Ve skutečnosti není klíčové, jakou metriku si podnik zvolí, nýbrž to, jakým způsobem s hodnocením dále pracuje a využívá získané poznatky k vlastnímu rozvoji a recipročnímu zvyšování úrovně customer experience.

Při volbě metriky pro měření úrovně customer experience je vhodné, aby podnik respektoval následující principy:

- výběr standardizované metriky umožní podniku využít best practices trhu a srovnání s konkurenčními i nekonkurenčními podniky, které využívají shodnou metriku
- proprietární metrika může lépe reflektovat situaci a výzkumné potřeby konkrétní společnosti, přičemž je možné lépe zajistit kontinuitu výzkumu tak, že metrika customer experience může být kompatibilní s metrikami úrovně zákaznické péče, zákaznické spokojenosti nebo jinými dílčími metrikami používanými v podniku již dříve
- zvolená metrika by měla být dostatečně komplexní, aby reprezentovala tvrdé i měkké aspekty interakcí zákazníka s podnikem
- zvolená metrika by měla být konzistentně používána po co nejdelší časové období a při evoluci metriky je vhodné uvažovat alespoň o částečném zajištění zpětné kompatibility s původní podobou metriky

7. Konejte na základě získané zpětné vazby

Zpětná vazba získaná od samotných zákazníků je nesmírně cenným materiálem, ale zároveň i velkým závazkem podniku vůči jejímu autorovi. Pokud má být potenciál získané zpětné vazby bezezbytku využit a její autor aktivně motivován k další participaci, je nutné, aby po sběru zpětné vazby následovala „akční“ a „výkonná“ fáze procesu. Ne každá reakce zákazníků musí vyvolat realizaci změnového požadavku v podniku, ale v každém případě je vhodné požadavky analyzovat a k těm dostatečně konkrétním a specifickým sdělit autorovi vlastní reakci podniku. Pokud se podnik rozhodne realizovat požadavek získaný v rámci zpětné vazby, je vhodné respektovat následující pravidla:

- dodávka změny by měla být realizována v dostatečně krátkém a „odůvodnitelném“ časovém horizontu, přičemž rozhodující při posuzování délky implementace nejsou zvyklosti v podniku, nýbrž chápání autora požadavku
- pokud to charakter změny dovolí, měla by být autorovi požadavku umožněna participace na implementaci nebo alespoň validaci výsledků

8. Provažte metriky customer experience na obchodní a finanční metriky

Prozákaznický přístup se snahou o maximalizaci customer experience je nezbytností při budování dlouhodobě perspektivního a oboustranně vyváženého vztahu mezi zákazníkem a podnikem. V řízení podniku však nelze zapomínat ani na ekonomický rozměr podnikání a primární cíl podniku v podobě generování ekonomického zisku. Z toho důvodu je nutné definovat vazbu mezi strategií maximalizace

customer experience a obchodně-finanční situací podniku. Díky této vazbě je možné sledovat, jak naplňování strategie maximalizace customer experience přispívá k prosperitě podniku, čímž zároveň poskytuje vrcholovému managementu podniku důkazy o tom, že orientace podniku na maximalizaci customer experience byla správným rozhodnutím a investice související s tímto rozhodnutím jsou opodstatněné a rentabilní.

Kandidáty na vhodné obchodně-finanční metriky, které mají přímou provazbu na metriky customer experience, mohou být:

- profit podniku z konkrétních zákazníků a jejich segmentů
- share-of-wallet zákazníků
- počet nových zákazníků a náklady na tuto akvizici
- míra cross-sell a up-sell
- míra odchodovosti stávajících zákazníků a úspěšnost retence

9. Sledujte vývoj metrik v čase

Okamžité absolutní hodnoty zvolených metrik jsou sice zajímavé a důležité, neboť poskytují bezprostřední náhled na situaci podniku a v případě použití shodných metrik umožňují také srovnání s konkurenčními i nekonkurenčními podniky.

Klíčový je ale vždy vývoj metrik v čase, neboť jen ten ukazuje, jak se spolu s metrikami vyvíjí v čase i stav podniku. Pro zvolené specifické metriky customer experience i provázané metriky finanční a obchodní je proto nutné stanovit si pro následující časový horizont cílové hodnoty a průběžně sledovat, jak se aktuální hodnoty metrik k cílovým hodnotám přibližují či naopak jak se od nich vzdalují. Významné poznatky vedoucí k lepšímu řízení customer experience v podniku i podniku samotného, může přinést hlubší analýza trendů metrik a determinace faktorů, které mají na hodnoty metrik signifikantní vliv, neboť ovlivňováním těchto faktorů podnik může dosahovat lepších výsledků vyjádřených lepšími hodnotami metrik.

10. Definujte strategii rozvoje customer experience v podniku

Maximalizace customer experience zákazníků podniku nemůže být v rámci podniku jednorázovou, krátkodobou ani živelně neřízenou aktivitou. Jelikož bývá implementace strategie maximalizující customer experience změna celé filosofie podniku, neobejde se bez silné podpory nejvyššího managementu a dostatečného vlastního mandátu, rozpočtu a pravomocí na jedné straně, ale i zodpovědností a KPI na straně druhé. Je nutné zajistit, aby se strategie rozvoje customer experience stala nedílnou součástí celopodnikové strategie.

Nutnost existence a trvalého rozvoje strategie customer experience managementu dokládají i výsledky průzkumu (Burns 2014), v němž meziročně významně získávají podniky, které mají formální programy v oblasti customer experience a k tomu dedikované zdroje.

Závěr

Materiál přináší vzhled do problematiky customer experience a jejího měření, přičemž klade důraz na specifika této problematiky v oblasti bankovnictví.

Respektování doporučení, která jsou hlavní výstupem tohoto materiálu, může pomoci bankovním institucím lépe rozumět vlastním zákazníkům, naučit se rozumět jejich potřebám a zejména pak zvyšovat návratnost investic, které do oblasti customer experience realizují.

S příslušnými modifikacemi jsou uvedená doporučení platná nejen v bankovním sektoru, ale napříč odvětvími.

Literatura

ALFARO, Elena, Enrique BURGOS, Javier VELILLA, Fernando RIVERO, Hugo BRUNETTA, Santiago SOLANAS, Beatriz NAVARRO, Jaime CASTELLÓ, Carlos MOLINA, Jaime VALVERDE, Lluís MARTÍNEZ-RIBES, Borja MUÑOZ a José Ignacio RUIZ, nedatováno. Customer Experience: A

multidimensional vision of experience marketing [online]. [vid. 7. září 2014]. Dostupné z: http://www.thecustomerexperience.es/en/download/en/eBook_CustomerExperience.pdf

ANON., nedatováno. Customer Experience | Gartner IT Glossary [online] [vid. 12. červen 2014 a]. Dostupné z: http://www.gartner.com/it-glossary/customer-experience

ANON., nedatováno. Customer Experience Management (CEM) | Gartner IT Glossary [online] [vid. 12. červen 2014 b]. Dostupné z: http://www.gartner.com/it-glossary/customer-experience-management-cem

ANON., nedatováno. Customer Experience Maturity Assessment | Forrester Research [online] [vid. 5. říjen 2014 c]. Dostupné z: https://forrester.co1.qualtrics.com/SE/?SID=SV_3DAKQm2wyr0swCN

ANON., nedatováno. Forrester's 2008 Customer Experience Rankings. Customer Experience Matters [online]. [vid. 5. říjen 2014 d]. Dostupné z: http://experiencematters.wordpress.com/2008/12/15/forrester%e2%80%99s-2008-customer-experience-rankings/

ANON., nedatováno. Forrester's 2012 Customer Experience Index Reveals Enterprise Innovations [online] [vid. 5. říjen 2014 e]. Dostupné z: http://www.cmswire.com/cms/customer-experience/forresters-2012-customer-experience-index-reveals-enterprise-innovations-014495.php

ANON., nedatováno. The 2014 Customer Experience Index | Forrester Research [online] [vid. 13. červen 2014 f]. Dostupné z: http://solutions.forrester.com/customer-experience/prove-roi-49RA-161472.html

ANON., nedatováno. What is the Experience Economy? | SM2 Strategic [online]. [vid. 13. červen 2014 g]. Dostupné z: http://sm2strategic.com/what-is-the-experience-economy/

ANON., nedatováno. World Retail Banking Report 2011 [online]. [vid. 13. červen 2014 h]. Dostupné z: https://www.worldretailbankingreport.com/download/2/be600ae4

ANON., nedatováno. World Retail Banking Report 2012 [online]. [vid. 13. červen 2014 i]. Dostupné z: https://www.worldretailbankingreport.com/download/3/be610ae5

ANON., nedatováno. World Retail Banking Report 2013 [online]. [vid. 13. červen 2014 j]. Dostupné z: https://www.worldretailbankingreport.com/download/4/be620ae6

ANON., nedatováno. World Retail Banking Report 2014 [online]. [vid. 13. červen 2014 k]. Dostupné z: https://www.worldretailbankingreport.com/download/9/be630ae7

ANONYMOUS, 2006. Banks Differentiate Themselves When They Provide Superior Customer Experience. ABA Bank Marketing. 10., roč. 38, č. 8, s. 4. ISSN 15397890.

ANONYMOUS, 2010. RightNow Study Finds Customer Experience Impacts Revenue; 85% of Consumers Would Pay More for an Exceptional Customer Experience. Business Wire [online]. 13.10. [vid. 13. červen 2014]. Dostupné z: http://search.proquest.com.zdroje.vse.cz/docview/757495822/62CEE93F91E6413CPQ/1?accountid=17203

BAGDARE, Shilpa a Rajnish JAIN, 2013. Measuring retail customer experience. International Journal of Retail & Distribution Management. roč. 41, č. 10, s. 790–804. ISSN 09590552.

BEARD, Ross, 2013. Customer Satisfaction Metrics: 6 metrics you need to be tracking [online] [vid. 5. říjen 2014]. Dostupné z: http://blog.clientheartbeat.com/customer-satisfaction-metrics-6-metrics-you-need-to-be-tracking/

BURNS, Megan, 2014. The Customer Experience Index, 2014 [online] [vid. 5. říjen 2014]. Dostupné z: http://solutions.forrester.com/Global/FileLib/webinars/2014_Q1_NA_Webinar_CXi_1_deck.xlsx.pdf

MARITZ, 2010. The Three Dimensions of Customer Experience Measurement [online]. 2010. [vid. 7. září 2014]. Dostupné z: http://www.maritzresearch.com/~/media/Files/MaritzResearch/Case-Studies/CEEvolutionOverallPOVSummaryDoc.pdf

MASCARENHAS, Oswald A., Ram KESAVAN a Michael BERNACCHI, 2006. Lasting customer loyalty: a total customer experience approach. The Journal of Consumer Marketing [online]. roč. 23, č. 7, s. 397–405 [vid. 13. červen 2014]. ISSN 07363761. Dostupné z: doi:http://dx.doi.org.zdroje.vse.cz/10.1108/07363760610712939

PHILIPP, Klaus, Michele GORGOGLIONE, Daniela BUONAMASSA, Umberto PANNIELLO a Bang NGUYEN, 2013. Are you providing the „right" customer experience? The case of Banca Popolare di Bari. The International Journal of Bank Marketing [online]. roč. 31, č. 7, s. 506–528 [vid. 13. červen 2014]. ISSN 02652323. Dostupné z: doi:http://dx.doi.org.zdroje.vse.cz/10.1108/IJBM-02-2013-0019

PINE II, Joseph B. a James H. GILMORE, 2011. The Experience Economy, Updated Edition. Updated edition. Boston, Mass: Harvard Business Review Press. ISBN 9781422161975.

RABINO, Samuel, Stephen R. ONUFREY a Howard MOSKOWITZ, 2009. Examining the future of retail banking: Predicting the essentials of advocacy in customer experience. Journal of Direct, Data and Digital Marketing Practice [online]. 6., roč. 10, č. 4, s. 307–328 [vid. 13. červen 2014]. ISSN 17460166. Dostupné z: doi:http://dx.doi.org.zdroje.vse.cz/10.1057/dddmp.2009.12

JEL Classification: M19

Summary

Measuring of customer experience in banking

The purpose of this paper is the introduction of possibilities and methods for measuring the customer experience. The finding of the paper is recommendations “decalogue” for customer experience measuring which reflects the relevant specifics of banking industry as well.

In the beginning of the paper the terminology is defined and the significance of customer experience for enterprises is explained. Followed by the part devoted to current methods of customer experience measuring, current metrics and internal maturity evaluation.

The next chapter describes the specifics of customer experience in banking.

The result of the paper is the decalogue for the measuring the customer experience in banking which recommends the following points:

1. keep track of all interactions
2. keep track of the behavior across channels
3. adjust the form and the content to the situation
4. motivate customers
5. motivate stuff
6. choose the appropriate metric according to your discretion
7. execute action based on the collected feedback
8. tie the customer experience metrics to the business and finance metrics
9. keep track of the metrics progress during the time
10. define the strategy of the customer experience development in the enterprise

Key words: customer experience, measurement, metrics, banking

Sociální sítě mohou personalistům usnadnit nábor

Lucie Böhmová

lucie.bohmova@vse.cz

Doktorand oboru Aplikovaná informatika

Školitel: doc. Ing. Vlasta Střížová, CSc., (vlasta.strizova@vse.cz)

Abstrakt: Organizace, které chtějí být konkurence schopný, musí umět získat a udržet talentované zaměstnance. Takových je v populaci malý počet a personalisté musí využít k vyhledání vhodných kandidátů efektivní nástroje např. sociální sítě. Společnosti se zde propagují, nabízejí produkty, udržují kontakt s klienty a nejnověji také hledají zaměstnance. Sociální sítě se stávají významnou alternativou vůči doposud tradičnímu vyhledávání zaměstnanců přes pracovní portály (v ČR např. jobs.cz).

V USA je selekce zaměstnanců skrze sociální sítě běžnou praxí, neboť o jednotlivci poskytují v mnoha ohledech snadno a rychle dostupné informace. Existují online HR programy pro lustrování zaměstnanců právě na sociálních sítích, které poskytnou personalistovi ucelený přehled o dané osobě. Dokáží také vyhledat pasivní uchazeče na sociálních sítích dle daných kritérií. V České republice se v tomto ohledu jedná o výzvu relativně novou.

Práce je zaměřená na sociální sítě a jejich vliv při výběru zaměstnanců v organizacích. Cílem práce je zjistit, zda sociální sítě mohou být vhodných nástrojem pro nábor zaměstnanců v České republice, tak jako tomu je v zahraničí. K zjištění byla použita analýza výskytu klíčových slov na Facebooku a Twitteru. Dále průzkum veřejně dostupných informací z profilů uživatelů Facebooku.

Výzkum ukázal, že sociální sítě jsou vhodným prostředkem náboru i v České republice. Sice je tento nástroj na počátku, ale už teď je patrný velký potenciál.

Klíčová slova: HR, sociální sítě, nábor zaměstnanců, Facebook, Twitter.

Úvod

Sociální sítě začínají být využívány k jiným účelům než k těm, pro které byly původně vytvořeny. Typické je to pro Facebook a Twitter. Dle Qualmana (2011): „Sociální mediální platformy, jako jsou Facebook, YouTube a Twitter zásadně mění způsoby, jakým se chovají zákazníci a firmy. A to především tím, že spojují miliony lidí nástrojem okamžité komunikace.“ Pro firmy, které se na sociálních sítích etablovaly, představuje nelehký úkol propojit tak odlišné obory jako je IT, zákaznické služby, marketing, management do jedné smysluplné komunikace. Musí přijít s novými postupy, rolemi a zodpovědnostmi, metrikami, strategií, zodpovědět na výzvy a právní otázky, které mohou vyvstat (Wollan, Smith, 2011). Sociální sítě tedy pronikají do různých oborů a oblast Human resources není výjimkou.

V podmínkách rostoucích tlaků konkurence a globalizace, se pro organizace stále více stává klíčovým schopnost HR oddělení vyhledat, přijmout a udržet talentované zaměstnance. Důvodem je konkurenční výhoda, kterou tímto způsobem výběru zaměstnanců firmy získávají. „Organizace, které náboru zaměstnanců nepřikládá dostatečnou váhu, jsou odsouzeny k zániku. Získávání kvalitních zaměstnanců musí mít v kterémkoliv odvětví podnikání tu nejvyšší prioritu“ říká Steve Crabb, šéfredaktor časopisu Personnel Today. (Swain, 2012)

Osob na trhu práce je poměrně mnoho. Současná míra nezaměstnanosti v ČR se pohybuje okolo 5,2% (ČSÚ, 2014), což teoreticky znamená statisíce volných pracovníků. Bohužel lidí, kteří by byli dostatečně kvalifikováni, je nedostatek. Společnosti pochopili, že talentovaní zaměstnanci jsou z řad pasivních kandidátů[1]. Průzkum LinkedInu toto tvrzení dokazuje. 72% organizací v USA hledá nové zaměstnance pouze z těchto kandidátů. (LinkedIn, 2015) Proto běžná inzerce v některých případech nestačí a personalista musí sáhnout právě po moderních metodách, kterými jsou například sociální sítě.

HR má několik možností, jak objevit vysoce talentované, kvalifikované, motivované a oddané zaměstnance schopné přispět k podnikovým cílům. Existuje šest hlavních možností, jak na internetu

[1] Jsou takový kandidáti, kteří práci sami nevyhledávají. Pokud by však dostali lepší pracovní nabídku, tak by ji pravděpodobně přijali.

hledat kandidáty: sociální sítě, pracovní portály, agregátory pracovních nabídek, úřady práce, webové stránky konkrétních firem v záložce "volná místa" a specializovaná diskusní fóra. Další možností vyhledávání práce je inzerce v papírové podobě. Ovšem v současné době ji lze považovat takřka za archaismus. Oporou pro toto rozhodnutí je i statistika společnosti Factum Invenio, ze které vyplývá informace, že 80% nabídky a poptávky práce je v současnosti realizováno na internetu. (Jobs.cz, 2010)

Důležité při výběru kanálu, kterým bude organizace šířit poptávku po zaměstnancích a zároveň sama aktivně vyhledávat talentované jedince, jsou především ekonomické aspekty, efektivita a administrativní náročnost. Cílem firmy je sice uspořit finance a čas při výběru zaměstnanců, ale na druhé straně je jejím velkým zájmem, aby tyto úspory neměly vliv na kvalitu potencionálních zaměstnanců. Proto je třeba hledat inovativní řešení výběru, kterým mohou být právě sociální sítě, neboť o jedinci poskytují v mnoha ohledech více informací, než personalista může při pohovoru s kandidátem zjistit. Ve Spojených státech amerických je selekce zaměstnanců přes sociální sítě běžnou praxí. (Social Times, 2012) Ty největší sociální sítě, mezi které řadíme Facebook, LinkedIn a Twitter, vyvinuly nové nástroje pro efektivní propojení s trhem práce. Prvenství ve vztahu k HR drží LinkedIn, sociální síť pro profesionály, která je svou podstatou předurčená k využití pro personalistiku.

Ze studie, kterou uskutečnily pracovní portál CareerBuilder.com[2] a výzkumná agentura HarrisInteractive[3] vyplývá, že sociální sítě hýbou americkým pracovním trhem. Také ukazuje, v jaké míře již nyní američtí HR manažeři a zaměstnavatelé kontrolují uchazeče o zaměstnání pomocí sociálních sítí. Výzkumu se účastnilo 2667 HR manažerů, zaměstnavatelů a profesionálů s působností v USA. Ze studie vyplývá, že 45% čili 1200 zaměstnavatelů lustruje své potenciální zaměstnance přes sociální sítě a dalších 11% s tím plánuje v brzké době začít. Nejvíce se pro „lustraci“ využívá Facebook, který se umístil na prvním místě s 29%, dále také LinkedIn s 26%. Je zajímavé, že tato profesionální sociální síť, která slouží k setkávání profesionálů a diskutování o svých pracovních zájmech je až na druhém místě, i když rozdíl jsou pouze 3%. (Careerbuilder.com., 2011)

Důkazem, že sociální sítě jsou opravdu důležitým faktorem při přijímání zaměstnanců, je fakt, který není možné přehlédnout. 35% HR manažerů ve studii uvedlo, že na sociálních sítích našli takový obsah, který zapříčil nepřijetí uchazeče. (Careerbuilder.com., 2011)

Další významné zjištění přináší studie, kterou zveřejnila společnost Mashable. Ta potvrzuje důležitost osobní prezentace uchazeče na internetu, především na sociálních sítích. Dokazuje, že kromě kvalitního životopisu, který zašle uchazeč o práci zaměstnavateli, rozhoduje u přijímacího řízení profil na sociálních sítích a jeho tzv. internetová stopa, kterou zanechává při každodenním působení a komunikaci na webu. Díky tomu je v současnosti jednodušší vyhledat talentované zaměstnance. Můžeme takového člověka najít, pokud má profil na sociálních sítích, emailovou adresu, telefonní číslo, blog nebo osobní stránky nebo také přidal příspěvek na chatu, diskuzních fórech, fotografie (např. rajče.cz), zároveň se stal členem profesní asociace nebo o něm byla zmínka v novinách atd. V neposlední řadě také na pracovních portálech, kde lidé vkládají osobní informace (CV). Dále z výzkumu vyplývá, že 91% zaměstnavatelů lustruje na internetu, aby si ověřila identitu uchazeče na sociálních sítích. Personalisté využívají pro zjištění doplňujících informací o uchazeči především následující zdroje: Facebook (76%), Twitter (53%) a také LinkedIn (48%). (Mashable.com, 2011)

Klasické programy[4] pro personalistiku již ale nedostačující v dnešní době, kdy je vše online. Čas strávený užíváním sociálních sítí roste mnohonásobně vyšší měrou než čas strávený na internetu obecně. Personalisté pracují se sociálními sítěmi jako s doplňkovým nástrojem náboru. Hledají zde různé podrobnosti z profesního i soukromého života kandidátů, především pak reference a pracovní zkušenosti. Proto byly vytvořeny online HR programy, které využívají jako jeden ze zdrojů informace ze sociálních sítí.

---

[2] CareerBuilder je světovým lídrem v oblasti lidského kapitálu pomáhající společnostem upoutat jejich nejdůležitější aktiva - lidi. Stránka CareerBuilder.com je největší ve Spojených státech s více než 23 miliony návštěvníků, 1 milionem pracovních míst a 32 miliony životopisů.

[3] Mezinárodní společnost zabývající se strategickým výzkumem.

[4] Tradičními programy jsou zde myšleny databáze mimo jiné pro ukládání detailů o náboru - například informace o kandidátech (CV, kontaktní údaje atd.), podrobnosti o pracovních nabídkách, zápisy z již provedených pohovorů.

Jedním z nich je "Social Intelligence", který poskytuje report aktivit jednotlivých uživatelů[5] na všech sociálních sítích. Součástí přehledu je i vyhodnocení dané osoby dle kritérií (oblast, znalosti, vzdělání atd.). Výhodou je, že HR takto může vyhledat i pasivní uchazeče. Podle magazínu Forbes (2014) Social Intelligence dělá pouze to, co dříve personalisté dělali ručně.

Článek hledá odpovědi, zda sociální sítě jsou vhodný nástroj pro výběr zaměstnanců. Porovnává dosavadní praxi v zahraničí, především USA, s českým prostředím. Zda jsou schopny obor HR posunout v nějakém směru kupředu nebo dokonce i nahradit velmi populární pracovní portály jako jobs.cz, prace.cz apod., které této oblasti v současnosti dominují. Případně do jaké míry mají sociální sítě šanci se prosadit i v oboru personalistiky a stát se plnohodnotným nástrojem náboru.

Výzkum

Byly provedeny dvě výzkumná šetření kvantitativní metodou, jejichž cílem bylo zjistit, zda jsou sociální sítě v České republice vhodným nástrojem pro nábor.

V první fázi výzkumu bylo zkoumáno 1400 uživatelských profilů z různých skupin na Facebooku v České republice. Skupiny byly vybrány z odlišných oblastí, s cílem udržet rozmanitost v datovém souboru. Sběr dat probíhal za pomoci studentů kurzů zaměřených na nová média a doktorandky na KSA Ing. Ludmily Malinové. Studenti vyhledávali veřejně dostupné informace (věk, vzdělání, bydliště, zaměstnání, dále zda mají viditelné fotografie, příspěvky, seznam přátel, skupiny) a ukládali je do strukturovaného soubor aplikace Excel. Termínem "veřejný" chápeme, že je k dispozici všem uživatelům Facebooku.

Následné výzkumné šetření bylo provedeno formou sběru dat pomocí webové aplikace, která vyhledávala vybrána klíčová slova, která uživatelé na sociálních sítích použijí pří hledání a nabídce zaměstnání. Aplikaci naprogramoval student KSA. Pro výzkum byl využit pouze Facebook a Twitter. Bohužel z technických důvodů nemohl být zařazen i LinkedIn. Sběr dat byl anonymní, proto nemáme k dispozici podrobnější demografické údaje (věk, pohlaví, vzdělání apod.) případně, zda klíčové slovo vložila soukromá osoba, pracovní portál nebo firma. Evidence klíčových slov probíhala po dobu tří měsíců v období od 25. 3. do 25. 6. Studie byla zaměřena pouze na Českou republiku, proto byla slova zvolena pouze v českém jazyce. Klíčová slova byla hledána ve všech pádech i osobách.

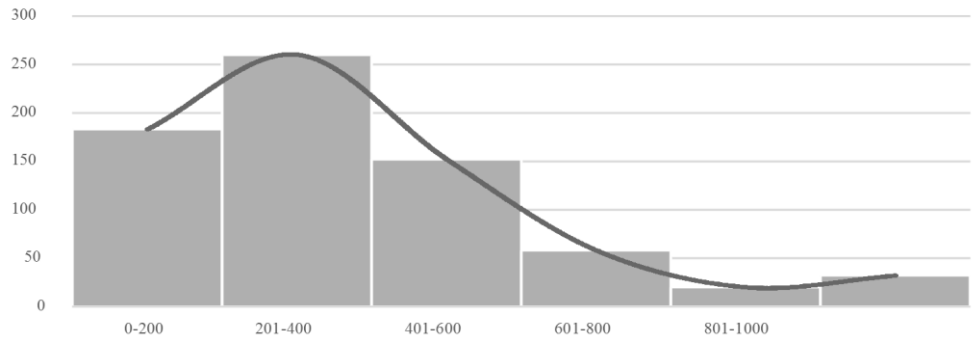
Klíčová slova byla rozdělena do tří skupin. První skupina obsahovala pouze slovo práce a zaměstnání. Druhá skupina klíčových slov byla zaměřena na hledání práce. Poslední okruh tvořila slova týkající se nabídky zaměstnání.

Diskuze

Z předchozího výzkumného šetření, které se zabývalo veřejnými profily na Facebooku bylo zjištěno, že dostupné informace má na svém profilu 62% uživatelů. Tito lidé mají veřejné především informace o jejich vzdělání (91,5 %), fotografie (62,5 %) a stránky (skupiny), které se jim líbí (61,3 %). Nejméně sdílí informace o vztahu (29,1 %). Pro HR jsou primárně zajímavé fotografie a skupiny, ze kterých lze vyčíst užitečné informace pro proces náboru. Případně usnadní rozhodnutí, zda pozvat kandidáta na osobní pohovor.

Z celkového vzorku mělo 47% uživatelů veřejnou informaci o počtu svých přátel na Facebooku. Průměrně má každý uživatel 399 přátel. Standardní odchylka je 324,4 a medián 327. V datovém souboru je normální rozdělení, jak je vidět v níže uvedeném histogramu počtu přátel. Vrchol normálního rozdělení leží v rozmezí 200 a 400 přátel, viz obrázek 1.

[5] Kteří mají veřejně dostupný profil. Social Intelligence dodržuje zákon the Fair Credit Reporting Act, který se zabývá ochranou osobních údajů.



Obrázek 1: Histogram počtu přátel s normálním rozdělením Zdroj: Böhmová, Malinová, 2013

Obrázek 1 ukazuje, že personalisté mohou vidět sociální prostředí přibližně u každého druhého kandidáta. Na základě vysokého počtu přátel lze snadněji získat reference vyhledáním kolegy ze současného nebo předchozího zaměstnání.

Z výsledků analýzy klíčových slov vyplývají následující fakta, viz tabulka 1. Průměrný denní výskyt klíčových slov z první skupiny (práce, zaměstnání apod.) je na sociálních sítích 1213 krát, konkrétně na Facebooku 950 krát a na Twitteru 262 krát. Průměrně v jeden den zadají uživatelé slovní spojení klíčových slov ve smyslu hledání práce 116 krát na obou sociálních sítích. Na druhou stranu nabídky práce se průměrně denně vyskytují 68 krát. Převládá poptávka po práci před nabídkou. Pokud bychom se zaměřili na konkrétní období, kdy jsou klíčová slova nejvíce použita, tak dominuje den čtvrtek a poté víkend (sobota i neděle). Statistiky výskytu slov mají lehce vzrůstající tendenci. Důvodem může být blížící se konec školního roku a tím i větší poptávka po práci.

Tabulka 1: Výskyt klíčových slov na Facebooku a Twitteru Zdroj: autor

	Twitter	**Celkem**	**Facebook**	**Celkem**
Zaměstnání	3576	23613	5160	85582
Práce	20037		80422	
Hledám zaměstnání	1725	4037	2829	6477
Hledám práci	2312		3648	
Nabízíme zaměstnání	904	2517	990	3563
Nabízíme práci	1613		2303	

Pracovní nabídky a poptávka po práci se začínají přesouvat na sociální sítě, tak jako tomu je v zahraničí, především v USA. Ve využití sociálních sítí v oblasti náboru dominuje Facebook před Twitterem. Důvodem může být široká škála (mladí/staří, základní/vysokoškolské vzdělání atd.) a vyšší počet uživatelů.

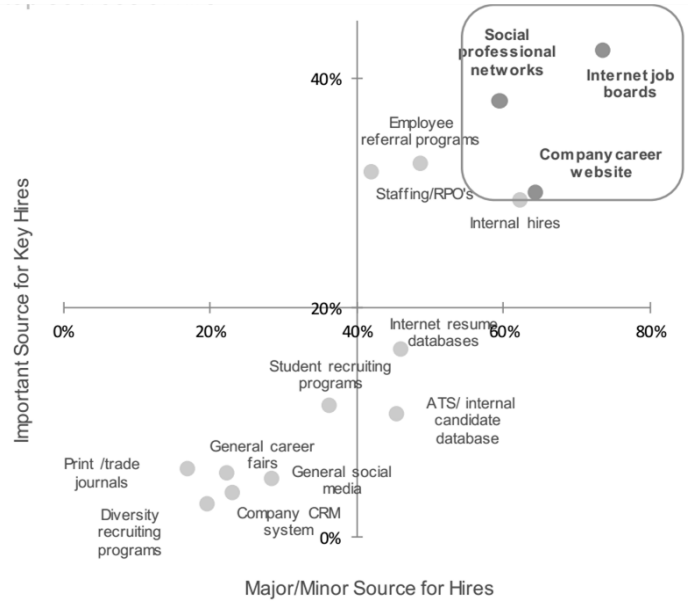
Z toho vyplývá, že pro personalisty má význam přesunout svoji pozornost i na sociální média a vkládat zde nabídky zaměstnání. Získat životopis od uchazečů o práci usnadní např. aplikace na Facebooku, díky, které mohou uživatelé vložit jejich CV přímo na sociální síti.

Z analýzy společnosti Effectix vyplývá, že počet uživatelů Facebooku i Twitteru v ČR neustále roste. Zatímco měl Facebook prudký nárůst počátkem roku 2010, dochází poté ke zpomalení, které je stále zřetelnější. V porovnání s těmito hodnotami roste Twitter stále rychleji. V březnu 2013 měl Facebook 3,8 mil. aktivních měsíčních uživatelů, z toho je cca 1,8 mil. uživatelů opravdu aktivních. Twitter měl 157 481 uživatelů na začátku roku 2013. (E15, 2013)

Sociální sítě jsou prostorem s velkým potenciálem pro HR a i čeští personalisté je začínají v poslední době objevovat. Dokládají to i výsledky studie, kterou realizovala společnost LMC na vzorku 234 českých personalistů, podle které sociální sítě typu Facebook nebo LinkedIn pracovně využívá 24 % dotázaných, 27 % z nich je užívá pouze k osobním účelům. 44% personalistů sociální sítě nevyužívá, protože by tím ztráceli čas. Zbylých 11 % respondentů zvažuje nad použitím komunitních sítí do budoucna. (Jobs.cz, 2010)

Přes nástup sociálních sítí zůstávají pracovní portály dlouhodobě jedničkou pro nabízející i poptávající na trhu práce. V krátkodobém horizontu se tato pozice nezdá být ohrožena. Tento fakt lze ilustrovat na výsledcích studie společnosti Factum Invenio, která říká, že 80 % internetové populace vyhledává pracovní příležitosti prostřednictvím internetu. Z toho téměř pětina uživatelů míří na sociální sítě a celých 75 % uchazečů hledá informace o nabídkách přímo na specializovaných pracovních portálech Jobs.cz a Práce.cz. (Jobs.cz, 2010) Problémem sociálních sítí je mnoho různých kandidátů, u kterých však není jisté, zda práci skutečně poptávají a zda stojí právě o pozici v dané firmě. Výhodou specializovaného on-line prostoru, kterými jsou pracovní portály je to, že se zde koncentrují ti, kdo skutečně práci hledají. Naopak pokud budeme hledat talentované zaměstnance, kteří jsou většinou pasivními uchazeči, sociální sítě jsou vhodným nástrojem.

LinkedIn vydal předpověď celosvětových trendů pro rok 2015 v náboru zaměstnanců. Mezi nimi je předpoklad, že sociální sítě budou ještě více využívány než doposud a dokonce předpokládají, že se stanou klíčovým zdrojem. Budou poskytovat dostatečné množství uchazečů, kteří budou i talentovaní, viz obrázek 2.



Obrázek 2: Zdroje potenciálních zaměstnanců Zdroj: LinkedIn, 2015

Závěr

Sociální sítě dávno neslouží pouze k zábavě, komerční využití nezadržitelně nastupuje. Nejedná se výhradě o reklamu, ale i o získávání informací a využití pro HR, především pro nábor zaměstnanců.

Personalisté pracují se sociálními sítěmi jako s doplňkovým nástrojem náboru. Hledají zde různé podrobnosti z profesního i soukromého života kandidátů, především reference a pracovní zkušenosti.

Sociální sítě tento trend evidují a začaly se mu obratně přizpůsobovat. Ty největší, mezi které řadíme Facebook, LinkedIn, i Twitter vyvinuly nové nástroje pro efektivní propojení s trhem práce. Facebook je využíván především ze strany personalistů, kteří se chtějí dozvědět o zájemci na pracovní místo více než jim sdělí LinkedIn, případně motivační dopis či životopis.

V zahraničí jsou sociální sítě využívány jako běžný prostředek pro hledání zaměstnanců, i když je samozřejmé, že touto cestou nelze hledat všechny typy profesí. LinkedIn vydal předpověď celosvětových trendů pro rok 2015 v náboru zaměstnanců s předpokladem, že sociální sítě budou ještě více využívány než doposud a dokonce se stanou klíčovým zdrojem. V České republice je využívání sociálních sítí v personalistice na počátku.

Na základě výzkumu a získaných sekundárních dat je patrné, že nárůst významu sociálních sítí je nesporný. Čeští uživatelé Facebooku vkládají na své profily mnoho informací. Více než 60% publikuje své fotografie veřejně pro všechny ostatní uživatele Facebooku. Necelá polovina uživatelů mají svou

zeď a seznam přátel snadno dostupné. Slova týkající se nabídky i poptávky po práci se na sociálních sítích Facebook a Twitter hojně vyskytují. Průměrně se slova jako práce, zaměstnání apod. vyskytne 1213 krát za den. Sociální média jsou vhodným prostředkem pro nábor i v České republice.

Literatura

Careerbuilder. Survey Reveals How Many Workers Commit Office Taboos. CareerBuilder.com. [Online] 2011. [Citace: 2015-01-18] Dostupné z: http://www.careerbuilder.com/share/aboutus/pressreleasesdetail.aspx?id=pr394 &sd=8%2F29%2F2007&ed=12%2F31%2F2007

ČSÚ. Statistika nezaměstnanosti vydaná 15. 05. 2014. Český statistický úřad. [Online] 2014. [Citace: 22. 01. 2015.] Dostupné z: http://www.czso.cz/x/krajedata.nsf/oblast2/zamestnanost-xc

E15. Sociální sítě v Česku: Facebook stále vede, firmy objevily YouTube. E15.cz [Online] 2013. [Citace: 2015-01-16] Dostupné z: http://zpravy.e15.cz/byznys/technologie-a-media/socialni-site-v-cesku-facebook-stale-vede-firmy-objevily-youtube-972707#utm_medium=selfpromo&utm_source=e15&utm_campaign=copylink

Forbes. Five Predictions For Social Media And Compliance In Financial Services In 2015. Forbes.com. [Online] 2015. [Citace: 2015-01-22] Dostupné z: http://www.forbes.com/sites/joannabelbey/2015/01/06/five-predictions-for-social-media-and-compliance-in-financial-services-in-2015/3/

Jobs. Pracovní portály vedou nad sociálními sítěmi. Jobs.cz. [Online] 2010. [Citace: 2015-01-18] Dostupné z: http://www.jobs.cz/poradna/novinky/aktualni-clanek/article/pracovni-portaly-vedou-nad-socialnimi-sitemi/

LinkedIn. 2015 Global Recruiting Trends. [online]. 2015 [cit. 2015-01-20]. Dostupné z: https://snap.licdn.com/microsites/content/dam/business/talent-solutions/global/en_US/c/pdfs/recruiting-trends-global-linkedin-2015.pdf

Malinova, L., Bohmova, L., 2013. Facebook User's Privacy in Recruitment Process. DOUCEK, P. -- CHROUST, G. -- OSKRDAL, V. (ed.). IDIMT 2013, Information Technology, Human Values, Innovation and Economy. Linz: Trauner Verlag universitat, 2013, ISBN 978-3-99033-083-8.

Mashable. Social networks study. Mashable.com. [Online] 2011. [Citace: 2015-01-16] Dostupné z: http://mashable.com/2011/10/23/how-recruiters-use-social-networks-to-screen-candidates-infographic/

QUALMAN, E., 2011. Digital Leader: 5 Simple Keys to Success and Influence. McGraw-Hill; 1 edition, 2011. ISBN 978-0071792424.

Social Times. 92% Of U.S. Companies Now Using Social Media For Recruitment. Leavenworthtimes.com [Online] 2012. [Citace: 2015-01-16] Dostupné z: http://socialtimes.com/social-media-recruitment-infographic_b104335

SWAIN A., BROWN J. N., 2012. Professional Recruiter's Handbook: Delivering Excellence in Recruitment Practice (2nd Edition). Kogan Page Ltd. 07/2012 ISBN: 9780749465414.

WOLLAN E., ZHOU C. a SMITH N., 2011. The Social Media Management Handbook: Everything You Need To Know To Get Social Media Working In Your Business. Wiley; 1 edition, 2011. ISBN 978-0470651247.

JEL Classification: M51.

Summary

Social networks could facilitate recruitment

Social networks don´t serve only for entertainment, but also are using for commercial purposes. This is not only about advertising, but also it is getting information and utilization for HR, especially for recruitment.

HR professionals working with social networks as a complementary recruitment tool. They are seeking various details of professional and private information of candidates, especially references and work experiences.

Social networks registered this trend and skillfully adapted. The largest, which belong Facebook, LinkedIn, and Twitter have developed new tools for effective connection with the labor market. Facebook is primarily used by HR professionals who want to know more information about the candidates because they can get there more than from LinkedIn, cover letter or resume.

Social networks are used as a common resource for finding employees abroad, although it is obvious that in this way we can not search for all types of professions. LinkedIn gone global trends forecast for 2015 in recruitment with the assumption that social networks will be more used than ever before, and even will become a key source. In the Czech Republic is using social networks for hiring process at the beginning, but already is evident their great potential.

Based on research and obtained secondary data is evident that the increase in the importance of social networks is indisputable. Facebook users carelessly publish a lot of personal information. More than 60% show their photos and around 50% of Facebook users, have their wall open for all Facebook users a similar percentage of users show the list of their friends publicly. Words related to the supply and demand for employment are very often on the social networks Facebook and Twitter. On average, the words such as labor, job, etc. occurs 1213 times per day. Social media are suitable tool for recruitment in the Czech Republic.

Key words: HR, social media, recruitment, Facebook, Twitter.

Komunikační protokoly NCIP, SIP 3.0 a budoucnost standardů RFID

Dušan Broďáni

brodanid@vse.cz

Doktorand oboru Aplikovaná informatika

Školitel: doc. Ing. Vilém Sklenák, CSc., (sklenak@vse.cz)

Abstrakt: RFID technologie implementována v knihovnách pokrývá RFID zařízení a komunikaci s knihovním informačním systémem. Aby byla zaručena kompatibilita mezi zařízeními od různých výrobců, zejména v případě pozdějšího rozšiřování systému, je nutné, aby komunikační protokoly splňovali přijaté standardy. Na poli komunikačních standardů s knihovním systémem již delší dobu vývoj stagnuje a situace dospěla do bodu, kdy dodavatelé používají proprietární řešení, protože současné protokoly jsou pro rozvoj a využití plného potenciálu RFID omezující. Výzkum současných implementací RFID a analýza a komparace dostupných systémů poukazuje na nebezpečí, kdy se knihovny mohou snadno stát závislými na jednom dodavateli. Existuje snaha tento stav napravit, ale zatím nebyl přijatý standard, který by se stal novou normou pro komunikaci s knihovním informačním systémem. Je proto důležité, aby knihovny trvaly na tom, že dodavatelé RFID, ILS nebo LMS budou rozvíjet jejich integrovaná řešení v rámci společného rámce dat.

Klíčová slova: RFID technologie, SIP 2.0, SIP 3.0, NCIP, NISO (National Information Standards Organization), LCF (BIC Library Communication Framework), ILS, knihovní informační systém, samoobslužní výpůjčky

Úvod

Implementace RFID technologie znamená pro knihovnu zásadní změnu, cílem které je optimalizace procesů, efektivnější využití lidských zdrojů a zejména zlepšení poskytovaných služeb pro uživatele. Nejedná se jen o nákup etiket a čtecích zařízení, neméně důležitým prvkem je knihovní informační systém, který data bude zpracovávat. Na trhu je již velký počet výrobců, kteří nabízejí ucelená řešení a pokrytí všech potřeb knihovny od zabezpečení knihovního fondu, samoobslužních výpůjček až po inventarizaci. V principu tyto systémy sledují stejný cíl, nicméně se od sebe odlišují technologii a dokonce i kompatibilitou se standardy, proto před samotným výběrem dodavatele je nutná podrobná analýza potřeb a požadavků knihovny. Problémem posledních let se ukazuje být stagnace v rozvoji komunikačního protokolu s knihovním informačním systémem, který se obecně používá, ale již nesplňuje požadavky na moderní protokol a kvůli svým omezením nedokáže plně využít potenciál RFID. Výzkum současných implementací RFID v knihovnách a analýza a komparace dostupných systémů poukazují na to, že se výrobci snaží tento problém obejít tvorbou vlastních řešení pomocí webových služeb a vytvářejí proprietární systémy. Knihovny často tomuto prvku nepřikládají dostatečnou důležitost, a to je činí do budoucna zranitelnými, protože se tak můžou stát závislými na jednom dodavateli. Lze očekávat, že v budoucnosti nastane potřeba systém rozšířit o nové funkce, se kterými se na začátku nepočítalo. Respektování standardů je proto důležité, i když se teď může zdát, že je na výběr buď řešení založené na starém nedokonalém komunikačním protokolu, nebo řešení, které toto omezení obchází, ale vzdaluje se od používaného standardu.

Zatím sice pomalu, ale přeci jen nastává posun ve vývoji nového protokolu, vydání nového standardu však bude vyžadovat ještě nějaký čas. Po stížnostech z řad dodavatelů byly vydány nové verze protokolů SIP (SIP 3.0 od společnosti 3M) a NCIP (NCIP 2.02 od National Information Standards Organization), které jsou ale jen vylepšením starších verzí. Na situaci zareagovala Londýnská Book Industry Communication, která vytvořila komunikační rámec pro data (LCF – Library Communication Framework), který má potenciál stát se součástí nového standardu komunikačního protokolu. Teď je za další vývoj protokolů SIP i NCIP odpovědná NISO, a všechno nasvědčuje tomu, že novým standardem se může stát nová verze protokolu NCIP, která již doznala značné vylepšení.

Princip a fungování RFID technologie v knihovnách

Systém RFID se skládá z dvou prvků:

- čtecí zařízení
- paměťové médium, které komunikuje s čtecím zařízením

Čtecí zařízení je zabudováno v pracovních stanicích výpůjčného protokolu, bezpečnostních branách, zařízeních pro samoobslužní výpůjčky, snímači pro inventarizaci položek a pod. (Obrázek 6: Library RFID Management system).

Obrázek 7: Library RFID Management system Zdroj: LibBest, 2014

Paměťové médium je malý čip s pamětí a anténou, zpravidla umístěn na papírové etiketě. Etiketa nemá zapotřebí žádný zdroj napájení, protože elektromagnetické pole generuje anténa připojena k RFID snímači. Když se v tomto poli ocitne RFID etiketa s transpondérem, v její anténě se generuje proud, který nabije kondenzátor a ten pak aktivuje čip. Důležité je, že etiketa samotná nevytváří vlastní elektromagnetické pole, ale jen mění pole vyvolané čtecím zařízením. To ve svém poli detekuje změny a převádí je na digitální data. Na to aby bylo možné využít všechny možnosti RFID technologie, je potřebné k uvedeným dvěma prvkům přidat také komunikaci RFID s knihovním systémem.

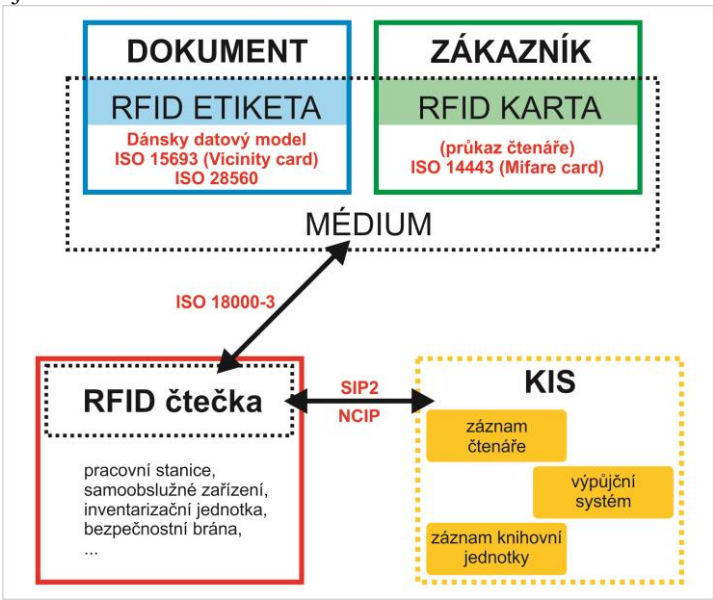
RFID etiketa spojuje identifikační i bezpečnostní funkci. Na čipu je uložen jednoznačný identifikátor publikace a další údaje, jako např. označení knihovny a země. Zaznamenává se zde informace, zda je publikace vypůjčena nebo zastřežena. Tento údaj je uložen v tzv. AFI byte a poskytuje více informací, než jen dva základní stavy. Je v něm uložena i informace o odvětví, ze kterého konkrétní etiketa pochází. Na základě toho bezpečnostní brána v knihovně reaguje jen na etikety v knihách ale k jinému zboží, které je zabezpečeno kompatibilním systémem RFID a prochází bránou, bude netečná.

Standardy a pravidla pro RFID

RFID systémy vyrábí mnoho různých firem po celém světě a pro knihovny ve světě jsou proto nasazované různé systémy. Od počátku se proto jevilo jako klíčové, aby výrobky různých firem byly navzájem kompatibilní a dokázali spolu komunikovat. Výrobci proto dodržují standardy ve formě norem. Pro koncového zákazníka to má výhodu v tom, že může kombinovat zařízení od více výrobců, které splňují danou normu. Standardy a pravidla určují:

- ***technické parametry RFID řešení*** – jako je vymezení frekvence bezdrátového přenosu, fyzikální vlastnosti etiket, přenos signálu, přenosový protokol – jsou důležité právě proto, aby zařízení od dvou různých výrobců byla navzájem kompatibilní,

- ***rozsah a způsob uložení informací na paměťovém médiu*** – toto je důležité, aby jakýkoliv systém uloženým datům „rozuměl“ a věděl, kde najít potřebný údaj,
- ***komunikaci s knihovním systémem*** – tj. způsob, jakým bude knihovní systém získávat informace uložené v paměťovém médiu a jak bude informace o dokumentech a čtenářích poskytovat nějakému zařízení RFID.



Obrázek 2: standardy RFID Zdroj: ITlib, 2013

Komunikace RFID – KIS SIP2

Původně proprietární přenosový protokol SIP společnosti 3M doznal značného rozšíření a především jeho druhá aktualizovaná verze SIP2 se tak stala standardem u většiny zařízení RFID, která komunikují s knihovním informačním systémem. Jeho funkcí je především předávání informací o jednotlivých položkách (knihách) a čtenářích. Často se však stává, že výrobci zařízení implementují ve svých zařízeních pouze podmnožinu tohoto protokolu, je tedy dobré před nasazením prověřit, zda jsou všechny potřebné funkce skutečně k dispozici (Zajíček, Rýzek, 2013).

NCIP (Z39.83 – NISO Circulation Interchange Protocol)

Standard definující přenosový protokol pro přenos informací a komunikaci zařízení RFID a informačního knihovního systému. Byl ustanoven organizací NISO a je založen na značkovacím jazyku XML (Zajíček, Rýzek, 2013).

Protokol SIP2 řídí komunikaci automatizovaného knihovního systému se systémem radiofrekvenční identifikace. Jeho nástupcem řešícím problémy vzniklé nedodržováním obecných standardů u všech výrobců knihovních systémů je protokol National Circulation Interchange Protocol (Národní cirkulační/oběhový protokol, dále NCIP) neboli Z39.83-2002, jenž definuje objekty a služby, které by se měly typicky vyskytovat, dále stanovuje jejich formální datovou strukturu a obsahový význam. NCIP se zabývá oblastí zajištění funkcí pro půjčování a vracení exemplářů, zajištění kontrolovaného přístupu k elektronickým zdrojům a usnadnění kooperativního řízení těchto funkcí jak pro tradiční dokumenty, tak pro digitální objekty. První verze NCIP protokolu se objevila v průběhu roku 2002. Tento standard však zatím nebyl přijat všemi výrobci radiofrekvenční identifikace pro knihovny.

NCIP, SIP 3.0 a LCF

Co přesně se děje s pokusy o zlepšení integrace ILS / LMS systémů s řešeními třetích stran? Primárním hráčem na jevišti je společnost 3M. Dlouho předtím, než RFID systém byl instalován v první knihovně, 3M podpořila integraci samoobslužních zařízení ILS / LMS přes SIP. Zpočátku se v 3M snažili prodat licenci do systémů dodavatelů, ale nakonec se přesvědčili, že výhody pro knihovny jsou větší, když další výrobce použije jejich protokol, než úzké obchodní zisky.

SIP byl uvolněn pro volné použití a v průběhu uplynulých desetiletí nasazován do provozu s cílem usnadnit širokou škálu služeb, které ani nebyly v plánu, kdy byl protokol poprvé navržen. Není překvapením, že se proto musel rozvíjet a růst na cestě ve snaze vyhovět novým nárokům.

Do roku 2006 SIP v jeho druhé zásadní revizi verze 2.0 značně vzrostl, co se týče jeho složitosti i rozsahu. Rozhodnutí uvolnit jej se pro vývojáře ukázalo být požehnáním i prokletím. Požehnáním byla jeho schopnost přizpůsobit se novým funkcím a prokletím, že neexistuje žádná na standardech založená kontrola nad použitím rozšíření. Následná snaha vytvořit národní datový model pro RFID vyústila do zveřejnění ISO 28560-2, ale nejběžnější stížnosti z řad poskytovatelů RFID řešení a ILS byly, že se SIP stal příliš omezený pro RFID aplikace. Takže po vydání ISO 28560 se pozornost přesunula k řešení problémů způsobených SIP.

Někteří dodavatelé ILS v USA zvolili do značné míry ignorovat SIP a místo toho vytvořili sadu webových služeb pro lepší integraci třetích stran včetně RFID. V reakci na tyto připomínky společnost 3M oznámila svůj záměr vytvořit vedle současného SIP2 i novou verzi 3.0, která byla vydána v roce 2012. Nakonec však 3M darovala autorská práva na protokol SIP pro NISO (National Information Standards Organization). Ani v roce 2014 se ale další rozvoj protokolu nedostal do plánovaného programu a NISO ponechává vedle sebe protokoly NCIP a SIP, i když se částečně překrývají. Bude mít NISO zájem přijmout a dále rozvíjet SIP když jejich protokol NCIP již nabízí RFID samoobsluhu? Nebo v rámci maximalizace zachování kompatibility musí knihovny používat oba protokoly? A vzhledem k současnému již dlouho trvajícímu stavu otázka ze všeho nejdůležitější – kdo by se zavázal k podpoře a rozvoji nového protokolu?

Jedná se o klíčové body, protože protokoly SIP i NCIP jsou obecně přijaty a používány a pro NISO možná bude i obtížné vysvětlit, proč vůbec práva k SIP přijala. Dodavatelům ILS tak poskytují nejlepší důvod, aby pracovali na proprietárních řešeních a v době, kdy NISO vydá novou verzi standardního protokolu již bude příliš pozdě, aby se dal vrátit džin zpět do láhve. V Londýně se problémem intenzivně zabývá Book Industry Communication (BIC), která přizvala k diskuzi 3M i NISO. Koncept je jednoduchý – vyvinout protokol k výměně dat v ILS. SIP jako protokol používá předepsané prostředky k výměně dat i skutečné hodnoty, které mají být použity. LCF (BIC Library Communication Framework) definuje datové prvky a hodnoty, ale nechal prostředky komunikace na jejich poskytovateli. Je to pragmatický přístup, který má nabídnout větší šanci na přijetí, protože dovoluje poskytovatelům určit prostředky, které se nejlépe hodí do jejich vývojového prostředí a zároveň umožňuje knihovnám zdokumentovat hodnoty používané k zajištění funkčnosti způsobem, který může být v případě potřeby předán jinému poskytovateli (Mick Fortune, 2013).

Harmonizace knihovních protokolů

I když se zdá, že vývoj v oblasti protokolů již delší dobu stagnuje, lidé pracující na protokolu NCIP se snaží o to, aby se stal novým standardem pro komunikaci s ILS / LMS. A nejedná se jen o snahu posunout dál protokol do nové verze, ale taky o kooperaci s lidmi kolem LCF.

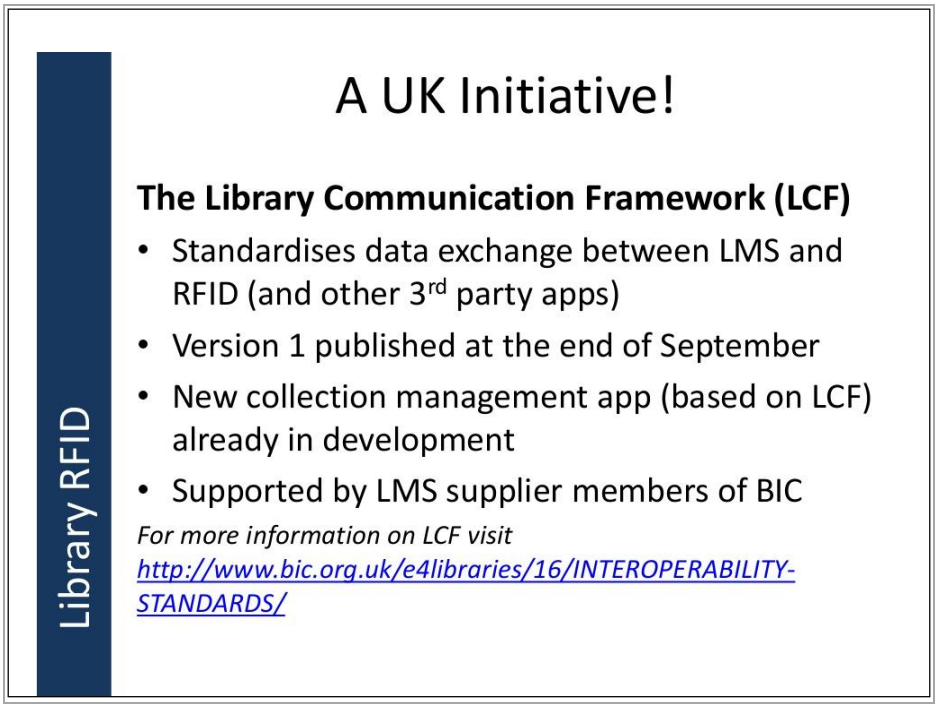
Lori Bowen Ayre popisuje její zkušenost ze schůzky výboru NCIP v Dublinu, Ohio:

Lidé z výboru pracují na rozvoji protokolu NCIP pro komunikaci s LMS. Mým cílem bylo zpochybnit tuto myšlenku, jelikož mým záměrem je zavést do knihovny komunikační rámec LCF, co je protokol vyvíjený v UK. Ale namísto odporu vůči LCF, jsou lidé z NCIP otevření naučit se více o problémech, které se snaží LCF řešit a jestli to nebude směr, kterým se možná bude muset vydat i NCIP. Vysvětlila jsem, že LCF byl vyvinut z velké části z důvodů tlaku dodavatelů, protože protokol SIP2 je příliš omezoval. Mám dojem, že lidi kolem LCF nevědí, že tvorba nové verze protokolu NCIP je nakloněná k spolupráci, ale může to být i kvůli přístupu k této otázce obecně ze strany NISO.

Proto věřím, že na NCIP bychom měli soustředit svou energii, pokud se jedná o standardní knihovní protokol. I když SIP2 je nejrozšířenější komunikační protokol pro knihovny, NCIP2 může umět všechno, co umí SIP2. Navíc to může dělat i se šifrováním. Také může umět všechno, co umí SIP3. Zatímco SIP 3.0 opravuje některé nejzávažnější problémy v SIP2, stále ještě není možné jej rozšiřovat jako NCIP. SIP je proto protokol, který musí být nahrazen.

NCIP si sebou nese špatnou pověst, protože jeho první verze velmi dobře nefungovala. V porovnání se SIP2 to bylo velké monstrum. Nová aktuální verze NCIP 2.02 je vysoce rozšiřitelná, bezpečná, flexibilní a velice dobře zvládnutá. I tak se ale často dodavatelé přiklánějí k SIP2, kromě případu, kdy se NCIP nelze vyhnout. SIP2 nepodporuje služby související se sdílením zdrojů. NCIP se používá mezi ILS a ILL software, ale i mezi ILS a ILS. SIP i NCIP lze použít i podpoře samoobslužního provozu a mezi samoobslužními zařízeními a systémy sdílení zdrojů je hodně překrývání. Například „Check In Item“, „Check Out Item“, „Lookup Item“, „Lookup User“ a „Renew Item“ jsou operace, které jsou u NCIP součástí základních služeb pro ILS-ILS i ILS-ILL, jakož i pro samoobslužní komunikaci. Jakmile dodavatel píše podporu pro těchto pět služeb, je na dobré cestě podpořit NCIP. Pokud přidá podporu pro jen čtyři další služby („Create User“, „Create User Fiscal Transaction“, „Undo Checkout Item“ a „Update User“), může zcela opustit SIP.

Hlavním bodem je, že musíme vyzvat všechny dodavatele, aby plně podporovali protokol NCIP. Na nějakou chvíli to může skončit u SIP2 / SIP3 ale nakonec to bude NCIP, ke kterému bude směrovat všechna komunikace. LCF nám dává vynikající podklad, o který musíme rozšířit NCIP (Obrázek 3: The Library Communication Framework). LCF definuje služby a datové prvky k využití technologie RFID. Dává smysl vzít LCF a stavět na něm nový protokol NCIP. Další výzvou je proto dostat ke společnému stolu lidi, kteří stojí za vývojem LCF s lidmi z vývoje NCIP a najít cestu ke kooperaci (Lori Bowen Ayre, 2013).

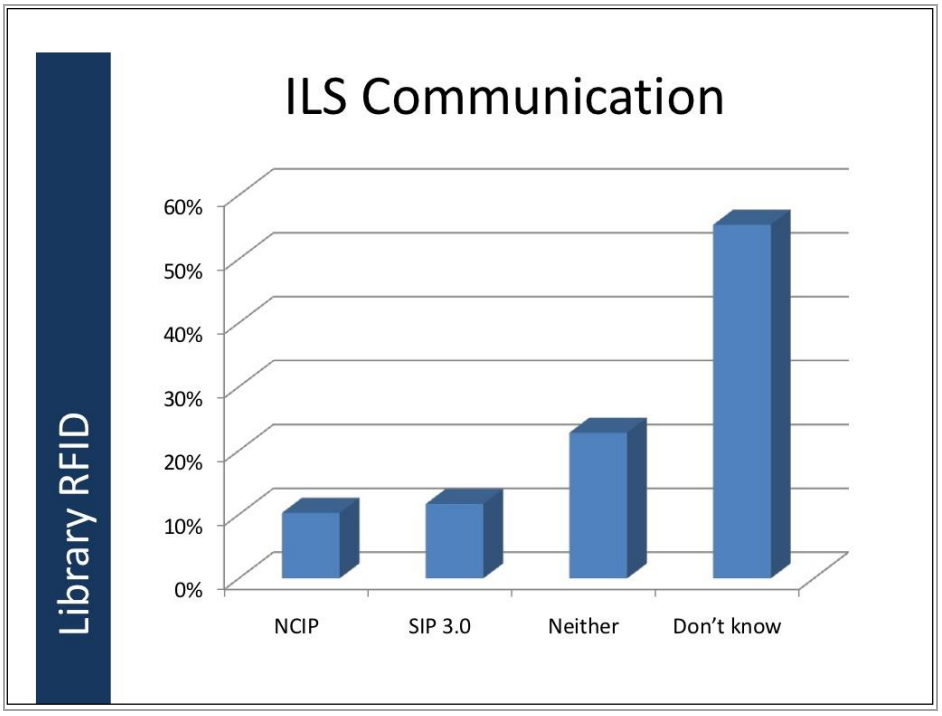


Obrázek 3: The Library Communication Framework Zdroj: Mick Fortune, 2013

Závěr

Stávající situace pro knihovny, které se rozhodují nasadit RFID systém proto není z pohledu budoucnosti jednoduchá. Dostávají se do zranitelné pozice, v které se jejich nynější zdánlivě dobré rozhodnutí může později ukázat jako neefektivní a v slepé uličce.

Dodavatelé ILS se můžou pokusit knihovny přesvědčit (do jisté míry oprávněně), že SIP je mrtev, nebo v nejlepším případě v zpoždění, a že nejlepší cestou je opustit staré myšlenky a přijmout svět API a webových služeb. Systém se tak stane zabezpečen před konkurencí, dodavatel je schopen velet při potřebách o jeho další rozšiřování, ale možná toto řešení bude atraktivnější pro méně bohaté knihovny. NISO jako nejvlivnější agentura na světě v oblasti norem se zdá být se současným stavem spokojená a v nejbližší době u ní nelze očekávat velké naděje na spasení.



Obrázek 4: ILS Communication Zdroj: Mick Fortune, 2013

Když se objeví požadavek knihovny na zavedení nebo upgrade systému RFID, často hledají funkční a ekonomicky nejvýhodnější řešení a nezajímají se o to, na jakých standardech je založeno řešení od dodavatele. Knihovna se tak lehce může dostat do pozice, kdy se stane závislá na svém dodavateli a v budoucnosti může mít problém zavedený systém rozšířit. Podle průzkumu (Obrázek 4: ILS Communication) je tento přístup v praxi běžný a poukazuje tak na nedostatečnou znalost pozadí nasazeného systému nebo tomuto bodu není přikládaná dostatečná důležitost.

Kam se tedy v této situaci obrátit? I když se od NISO objeví nová verze protokolu SIP 3.0 nebo NCIP, stále budou jen regulací pro řízení oběhu v ILS. Nové produkty, které využívají rozšířené možnosti RFID tagů po vydání normy ISO 28560 jsou již ve vývoji. Jedinou naději prevence zahlcení trhu pro knihovny od nepřeberného množství proprietárních řešení je trvat na tom, že dodavatelé RFID, ILS nebo LMS budou rozvíjet jejich integrovaná řešení v rámci společného rámce dat a jediný, který je k dispozici je LCF.

Literatura

LibBest, 2004-2014. *Library RFID System – System Architecture* [online]. Taipei, Taiwan: LibBest [vid. 2014-12-19]. Dostupný z: http://www.rfid-library.com/

The Galecia Group, 2000-2014. Lori Bowen Ayre: *Harmozation Of Library Protocols* [online]. Petaluma, CA, USA**:** Blog Post [vid. 2014-11-20]. Dostupný z: http://galecia.com/blogs/lori-ayre/harmonization-library-protocols

NISO, 2000-2014. *NCIP: NISO Circulation Interchange Protocol* [online]. Baltimore, MD, USA**:** NCIP workroom [vid. 2014-12-19]. Dostupný z: http://www.niso.org/workrooms/ncip

Book Industry Communication, 2014. *LCF: Library Communication Framework* [online]. London, UK: BIC [vid. 2014-12-19]. Dostupný z: http://www.bic.org.uk/114/LCF/

Changing Libraries, 2014. Mick Fortune: *NCIP, SIP3, BLCF and the future of RFID* [online]. Wordpress [vid. 2014-12-19]. Dostupný z: http://www.mickfortune.com/Wordpress/?p=864

Library Journal, 2012. Michael Kelley: *3M to Donate Copyright for SIP, a Key Library Communication Protocol, to NISO* [online]. New York, NY, USA: Library Journal [vid. 2014-11-19]. ISSN 0363-0277.

Dostupný z: http://lj.libraryjournal.com/2012/06/industry-news/3m-to-donate-copyright-for-key-library-communication-protocol-to-niso/

PAVEL ZAJÍČEK, JÁN RÝZEK, 2013. Standardy a pravidla pro technologii RFID. *ITlib*. 2/2013. ISSN 1336-0779.

JEL Classification: O2

Summary

Communication protocols NCIP, SIP 3.0 and future of RFID standards

RFID technology as implemented in libraries consists in the RFID device and communication with the library information system. To ensure compatibility between devices from different producers, especially in case of subsequent expansion of the system, it is necessary to ensure the compliance of communication protocols with recognized standards . In the area of communication standards with the library system there is already for a longer time a stagnation and the situation has reached a point where suppliers use proprietary solutions, because the current protocols represent a limitation to the development and use of the full potential of RFID. The research of current RFID implementations and analysis and comparison of available systems point to the danger where the libraries can easily become dependent on one supplier. There is an effort to remedy this situation, but no standard has been adopted yet, which would become the new standard for communication with the library information system. It is therefore important that libraries insist on RFID, ILS or LMS suppliers to develop their integrated solutions within the common data framework.

Key words: RFID technology, SIP 2.0, SIP 3.0, NCIP, NISO (National Information Standards Organization), LCF (BIC Library Communication Framework), ILS, library information system, self check

Visualization of semantic web vocabularies and linked datasets

Marek Dudáš

marek.dudas@vse.cz

Doktorand oboru Aplikovaná informatika

Školitel: doc. Ing. Vojtěch Svátek, Dr., (svatek@vse.cz)

Abstract: We analyzed several semantic web ontology and vocabulary visualization tools from the perspective of use case categories as well as low-level functional features and OWL expressiveness. A rule-based recommender was subsequently developed to help the user choose a suitable visualizer. Both the analysis results and the recommender were evaluated via a questionnaire. The questionnaire results rather confirm our hypotheses that different use cases need different visualization methods. We identified two specific use cases where the current visualization methods might be insufficient. The first one is ontology/vocabulary local coverage comparison. We propose using ontological background model visualization to make such comparison more easily comprehensible. The second is visualization of a vocabulary usage in a specific dataset. For this case, we propose a visualization method based on a dataset summarization using frequent paths and characteristic sets.

Keywords: semantic web, OWL, schema visualization, ontology visualization, dataset visualization, ontology local coverage comparison

Introduction

Numerous visualization methods have been proposed for OWL ontologies and many software tools implementing them appeared in the past decade, ranging from semantic-web-specific approaches to proposals of UML profiles (Brockmans et al., 2004, Kendall et al. 2009, Parreiras et al. 2010) leveraging on the similarity with software engineering models. However, so far no visualization method has been accepted by the majority of the semantic web community as de facto standard. One reason clearly is that different use cases require different ontology visualization methods. We defined a categorization of such possible use cases and analyzed available visualization tools with respect to the categorization. Based on the analysis, we created a recommender tool that suggests the most suitable visualization tool according to the given use case and ontology characteristics. We identified two use cases where the currently available visualization tools are insufficient: ontology local coverage comparison and visualization of ontology usage in a specific dataset. We propose original visualization methods for these two use cases.

Ontology Local Coverage Comparison

OWL allows to model the same real-world state of affairs using different combinations of language constructs, a kind of modeling styles. For example, in a lightweight ontology, data properties might be preferred, while in a more complex ontology, object properties might be used for describing the same relationships, allowing to state more facts about the property value. Current ontology visualization tools (as surveyed in (Katifori et al. 2007)) either display the OWL constructs directly or only offer very lightweight generalization, as in (Hayes et al. 2005). When the user wants to compare, in a visual manner, the local coverage of different ontologies, i.e. their capability to express a certain cluster of relationships, s/he has to first translate the OWL language constructs into his/her mental model to abstract from the modeling differences. A possible solution in such a situation is to express the above mentioned cluster of relationships in a modeling language that allows abstracting from the modeling style differences and thus making the mental model explicit. We believe that PURO ontological background models (OBMs) (Svátek et al. 2013) might fit this use case. We are developing PURO Modeler, a Javascript-based tool for graphical design of OBMs and ontology local coverage comparison. In the paper, we demonstrate its usage.

Visualizing Ontology Usage in a Dataset

In contrast to the RDBMS world, there is no such thing as a schema of an (RDF) dataset on the semantic web. There are of course RDFS/OWL ontologies and vocabularies, but those only define sets of concepts that can be used in a dataset rather than what combinations of concepts should be used to describe the data. The relationship between a vocabulary and a dataset is thus by far not as strict as the relationship between an SQL database and its schema. If the user encounters a dataset s/he is not familiar with, finding out what kind of data it contains is nontrivial (since up-to-date and complete documentation of the dataset is usually not present). The other way around, when the user encounters an ontology/vocabulary s/he is not familiar with, s/he can try to learn the proper usage of the ontology by reading the documentation and inspecting the axioms (such as domain/range), labels and comments in the RDFS/OWL source code. However, even the ontology documentation may be insufficient or missing, and the axioms (being mere inference rules) do not fully specify the usage, as said above. Then the only remaining option is to look at datasets where the ontology is used (provided they exist).

The user can obviously explore the dataset manually, either directly (by dereferencing its nodes in a Linked Data browser) or using exploratory SPARQL queries. However, manually browsing large datasets, even with the help of visualization tools such as RDF Gravity (Goyal and Westenthaler, 2004), is unwieldy, and even exploration via "clever" SPARQL queries, e.g., starting by listing all predicates used in the dataset and gradually refocusing, requires many non-trivial steps as well. There has been quite a lot of work done in the area of RDF dataset summarization in order to optimize SPARQL queries or provide query recommendation (Neumann and Moerkotte, 2011; Campinas, 2012) and optimize RDF storage (Pham, 2013). (Pham, 2013) is not only involved in optimization, but also mentions the possibility of creating an SQL schema of the RDF dataset, which should help the user understand the data structure. Helping the user see the structure of the dataset is also one of the goals of (Brunetti et al. 2013), which includes a method for showing an overview of classes used in the dataset.

We think that with a little effort it should be possible to reuse some of the existing methods to allow creating a visual overview of a dataset including both classes and properties used in it. In the paper we demonstrate a possible way of combining type-property paths (Presutti et al. 2011) and characteristic sets (Neumann and Moerkotte, 2011) into one graph showing the typical combinations of class instances and properties present in the dataset. Such visualization should allow the user to both (a) get an overview of the data and its structure in the dataset, and, (b) learn how the ontologies are used in it.

State of the Art

We are unaware of a recent analysis with as large coverage of visualization tools as in this paper, nor of an implemented system for tool recommendation. Our questionnaire was inspired by a survey (Ernst and Storey, 2003) done several years ago aimed at discovering which visualization tools are used by whom and for what tasks. (Katifori et al. 2006) describes the results of a comparative evaluation of (Protégé) Entity Browser, Jambalaya, TGViz and Ontoviz done with a group of test users. Time needed to perform a set of several predefined tasks in each tool by each user was measured and the users were also asked to fill in a questionnaire after using the tools. However, (Katifori et al. 2006) compares the tools using a single ontology and a specific use case, while our analysis is aimed at comparing the tools regarding different ontologies and use cases. A subsequent paper by the same group (Katifori et al. 2007) defines a categorization of visualization methods and tasks, and includes a thorough (but ageing) survey of ontology visualization tools. An approach similar to (Katifori et al. 2006) has been chosen in (Swaminathan and Sivakumar, 2012) for a comparative evaluation of Ontograf, OWL2Query and DLQuery: a group of users was asked to perform a set of predefined tasks aimed at finding information about instances and the time to complete each task was measured. A review of Protégé Entity Browser, Ontoviz, OntoSphere, Jambalaya and TGVizTab is available in (Sivakumar and Arivoli, 2011). Finally, (Da Silva et al. 2012) proposes rough guidelines for visualization tool design.

There is no prior research on visual comparison of ontology local coverage that we are aware of. Creating and visualizing mappings (Choi et al. 2006) between the compared ontologies might be useful but it only shows what the ontologies have in common. There is research on ontology comparison (Maedche and Staab, 2002) but that is rather focused on automatic computation of ontology similarity.

In our approach the actual analysis is left to the user; we however propose a method and means for supportive visualization. The PURO language, designed to be easily mappable to OWL, is used as interlingua for this purpose. Other modeling languages, e.g., entity-relationship diagrams (Chen, 1976), might be considered as alternatives to PURO. Of those, OntoUML (Albuquerque and Guizzardi, 2013) a version of UML for conceptual modeling, is probably the closest match; an existing tool, OLED, even allows to transform OntoUML models into OWL fragments. However, the purpose of OntoUML is conceptual modeling with easy validation and it is not designed for OWL ontology development or usage.

Our research of dataset summary visualization is mainly based on (Presutti et al. 2011), which presents the notion of type-property paths. They are used to obtain knowledge patterns (KP), which are quite similar to what we construct by typing the characteristic sets. The KPs are used to construct prototypical SPARQL queries that result into examples of typical data - instantiations of KPs. The main focus of (Presutti et al. 2011) is constructing such prototypical queries for a previously unknown dataset. We also plan to include the possibility of displaying similar typical instantiations of parts of the summarized graph, but at the moment we aim at summarization at the level of types rather than instances. Somewhat similar results are also offered in (Brunetti et al. 2013) which includes a method for displaying a summary of classes used in the dataset in a treemap (showing their hierarchy). (Pham, 2013) uses characteristic sets to summarize a dataset into a relational schema, offering an "SQL view of the RDF data." (Campinas et al. 2012) uses a summarization of the RDF graph in order to provide SPARQL query recommendations.

Ontology Visualization Tool Recommender

Visualization Tools Analysis

Based on ad hoc analyses of literature and own experience, we collected nine possible *use cases* of ontology visualization: making screenshots of selected parts of an ontology (*uc1*) or of its overall structure (*uc2*); structural error detection (*uc3*); checking the model adequacy (how well the ontology covers its domain) (*uc4*); building a new ontology (*uc5*); adapting an existing ontology, e.g., adding entities or transforming the style, to fit specific usage (*uc6*); analyzing an ontology in order to annotate data with (or create instances of) its entities (*uc7*); deciding about the ontology suitability for a specific use case (*uc8*); analyzing an ontology in view of mapping its entities to those from another ontology (*uc9*).

We then aggregated the use cases to four *categories*, named *Editing*, *Inspection*, *Learning* and *Sharing*, positioned in a 2-dimensional space. The first dimension distinguishes whether the user actively "develops" the ontology or just "uses" it. The second dimension is the required level of detail, i.e. whether an overview (e.g., class hierarchy) is enough or a detailed view (e.g., axioms related to a class) is needed. The categorization, incl. use cases is shown in Figure 1. *Editing* is a "development" category that needs both a general and detailed view, while *Inspection* focuses on a detailed view. Within visualization for "usage", *Learning* tends to use the general overview and *Sharing* tends to use the detailed view. The categories are not completely disjoint: *Inspection* overlaps with *Editing*, since the user usually needs to inspect the impact of her/his edits.

Inspired by categorization of tasks that should be supported by an information visualization application as defined in (Shneiderman, 1996) (and adapted in (Katifori et al. 2007)) and evaluation criteria classes described in (Freitas et al. 2002) we further identified seventeen relevant *features* (see Table 1) implemented in ontology visualization tools. Initially we identified 21 visualization tools. For the 11 we considered as "usable" (regarding their availability, stability and development state), we then characterized their supported features and language constructs. The eleven "usable" tools have been both evaluated in detail and included into the knowledge base of our recommender. The detailed analysis aimed to find out what features the tools implement (and how well), what language constructs they visualize and how they deal with larger ontologies.

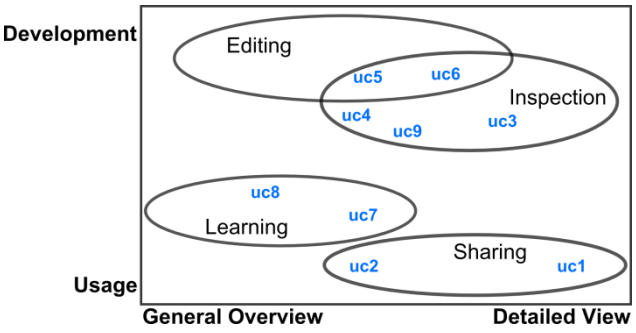


Figure 1- Ontology visualization use case categories

Table 1 shows the features we found as supported in each tool. The number indicates the support level as evaluated by us: "1" means "only implemented partially and/or in a rather user-unfriendly way"; "2" means "fully implemented"; empty cell means "not implemented".

Table 1 - Features implemented in each tool.

	Zoom-out Overview	Radar View	Graphical Zoom	Focus on selected entity	History (undo/redo)	Pop-up Window / Tooltip	Incremental Exploration	Search	Hide selected entity	Filter specific entity type	Fisheye Distortion	Edge Bundles	Drag&Drop Navigation	Drag&Drop User Layout	Clustering	Integration with editing	Graphical editing
Jambalaya	1		1	1		2	1	2	2	2	1			2		2	
KC-Viz	2		2		2	2	2	1	2	2			2	2	2		
Ontograf			2	2		2	2	2		1			2	2		1	
Ontology Visualizer	2		2	2	2		2	2					0	2			
Ontoviz			2						2				2				
OWLGrEd	1	2	1		2	2						1	0	2		2	2
OWLViz	2		1	2		2	2	1	2	0			0	0		2	
Protégé Entity Browser	2					2	2	2								2	
SOVA	1		2	0				2		2			2	2			
TGVizTab	1		1	2			2	2	2	2	1		2				
TopBraid		2	1						2					2			2

The alignment between the categories and tool features is, mostly, intuitive: **Editing** ranges from developing a new ontology to merely changing a property value. Usually, the purpose of visualization is to find the entity to be changed or to become parent of a new entity. Important features are *Pop-Up Window* (to see detailed properties of a selected entity), *Search* (to easily find entities to be edited) and obviously *Integration with Editing* and *Graphical Editing*. **Inspection** needs a detailed view of the ontology to see errors or deficiencies. It is often implied by Editing, as the user needs to see the state of the ontology before and after an edit. *Pop-Up Window* and *Search* are thus important here as well (while *Integration with Editing* and *Graphical Editing* not so). Additionally, *Hide Selected Entity*, *Filter Specific Entity Type* and *Focus on Selected Entity* are useful, as it is easier to spot errors after hiding the previously checked elements and/or focusing on the unchecked ones. **Learning** means gaining knowledge about the domain the ontology covers or learning about the ontology itself so as to use it. The view need not be as detailed as for discovering errors (in Inspection). The user often only needs to see the available classes and properties, their hierarchy and their domain/range relationships; especially non-technical users who only want to learn the domain are unlikely to understand complex axioms anyway. Important features are thus *Zoom-Out Overview*, *Radar View* (to see an overview of the ontology) and *Incremental Exploration* (for user-friendly exploration of the ontology in more but not too much detail). **Sharing** is similar to Learning, except for one more actor to whom the ontology is to be explained and shared with. The visualization should thus support displaying a part of the ontology; e.g., a picture of it can be made for an article describing the ontology. *Hide Selected Entity* is important

for displaying the desired part only, and *Drag&Drop User Layout* is useful for achieving an appropriate layout, e.g., placing important entities into the center.

Table 2 shows the "suitability scores" (ss) of each tool for each of the use case categories. The scores are calculated from the values in Table 1 using the following simple formula:

$$ss = \sum ImportantFeatureScores \times \alpha + \sum OtherFeatureScores \times \beta$$

"Other features" are all features which are not specified as important for the given use case category, with the exception of Integration with Editing and Graphical Editing, which are only taken into account for Editing. The feature score is 0, 1 or 2 as shown in Table 1. The multiplication coeficients α and β were set to 3 and 0.5, respectively, since these values provided good discriminatory ability in our initial test with the recommender system. The suitability scores are normalized to interval <-3;3> in the recommender and used as weights for the appropriate rules.

Table 2 - Suitability scores of each tool in various aspects.

	Editing	Inspection	Learning	Sharing
Jambalaya	23,5	30,0	12,5	17,5
KC-Viz	18,0	28,0	20,5	20,5
Ontograf	20,5	25,0	12,5	12,5
Ontology Visualizer	12,0	17,0	17,0	12,0
Ontoviz	3,0	8,0	3,0	8,0
OWLGrEd	22,5	10,5	13,0	10,5
OWLViz	19,5	23,5	16,0	11,0
Protégé Entity Browser	20,0	14,0	14,0	4,0
SOVA	10,5	15,5	8,0	10,5
TGVizTab	12,5	27,5	15,0	12,5
TopBraid	9,5	8,5	8,5	13,5

The Recommender Implementation

The recommender (available as web application[6]) is built as a knowledge base (KB) for the NEST expert system shell (Berka, 2011). NEST covers the functionality of compositional rule-based expert systems (with uncertainty handling), non-compositional (Prolog-like) expert systems, and case-based reasoning systems. NEST employs a combination of backward and forward chaining and it processes uncertainty according to the algebraic theory of (Hajek, 1985). The KB consists of three layers. The top layer only contains one node, representing the recommendation of visualization tool. The middle layer contains two nodes, which aggregate the relevant answers from the user: Use case category (editing, inspection, learning and sharing) and OWL (the importance of particular OWL features). The bottom layer represents possible user answers to questions:

- Complex classes: importance of anonymous classes based on various OWL constructs such as union, complement, intersection etc.
- Intended usage: the nine use cases we identified above.
- Ontology size: fuzzy intervals for `small', `medium' and `large' ontologies.
- OWL features: importance of particular OWL features (object properties, interclass relationships, datatype properties and property characteristics), aiming to infer the overall importance of OWL support in the visualization.
- Favorite ontology editor: the user's preference for some (freely available) ontology editor: Protégé 3, Protégé 4 or Neon Toolkit.

[6] Available at http://owl.vse.cz:8080/OVTR/

Evaluation

As the evaluation of both the overall analysis and the recommender requires human expertise, we prepared a web-based anonymous questionnaire,[7] sent invitations to participate in it to several relevant mailing lists, and also asked several people from the area of ontology engineering, including authors of the surveyed visualization tools, directly. We gathered answers from 32 respondents. We discuss some of the results below.[8]

When introduced to a brief description of the use case categories, 13 respondents answered that the categorization makes "perfect sense", the same number of respondents stated that it makes "more or less sense", 5 replied that it "does not make much sense", and no one chose "does not make sense at all"; 1 respondent did not answer this question. To sum up, about 84% of respondents partly or fully agreed with the categorization. 29 respondents ran the consultation, of which 14 (48%) were satisfied and 7 (24%) partly satisfied with the resulting recommendation, i.e. 72% of respondents were partly or fully satisfied. In 23 (79%) cases, KC-Viz was recommended as the most suitable tool. Such a high percentage of one tool was mainly caused by the fact that most of the respondents (approx. 66%) entered an ontology size larger than 120 entities, which is considered by the recommender as "large", and KC-Viz is considered as the most suitable tool for large ontologies. These results suggest that we should reconsider the rules in the KB related to the size of the ontology (the risk of clutter when a larger ontology is visualized might be over-estimated for some tools).

Visualizing Ontology Local Coverage

PURO Modeler[9] is basically a graph editor. The user can create an OBM where the terms are represented by nodes and the relationships between them by edges. As the OBMs are models of specific real-world situations, such an example cluster of relationships has to be chosen. After constructing the model,[10] or loading an already existing one (possibly created by another user and shared), the user can select each node and annotate it with the ontologies that this particular instance of PURO term is covered by (i.e., the instance can be described with that ontology). The parts of the graph locally covered by each vocabulary are then highlighted so that the coverage could be easily compared in one diagram.

Example

Let us say we need to annotate data about books and their various issues (i.e. their manifestations). We can find several ontologies that might be used for it. **BIBFRAME**[11] contains class *Text* as a subclass of *Work*, class *Print* as a subclass of *Instance* and the *hasInstance* property for stating that an instance of *Print* is a manifestation of some *Text*. Figure 2b shows an example of possible usage of BIBFRAME to describe the book *The Semantic Web for the Working Ontologist* and its printed paperback manifestation. **FRBR**[12] ontology distinguishes between four different "levels of abstraction" represented by classes *Work*, *Expression*, *Manifestation* and *Item*, each with subclasses representing more specific types. There is a set of properties that can be used to link the instances of the classes, as shown in Figure 2c. **Schema.org** does not distinguish between the work and its manifestation through class membership, but contains properties *exampleOfWork* and *workExample* that serves that purpose. It contains the class *Book* and allows to specify the format of its instances by linking them to instances of *BookFormatType* with property *bookFormat* (Figure 2a).

[7] Available at http://owl.vse.cz:8080/OVQuestionnaire/
[8] The full evaluation analysis is available at http://bit.ly/1phLBdm
[9] Available at http://lod2-dev.vse.cz/puromodeler
[10] The methodology for the OBM modeling is yet to be formalized.
[11] http://bibframe.org/vocab
[12] http://purl.org/vocab/frbr/core

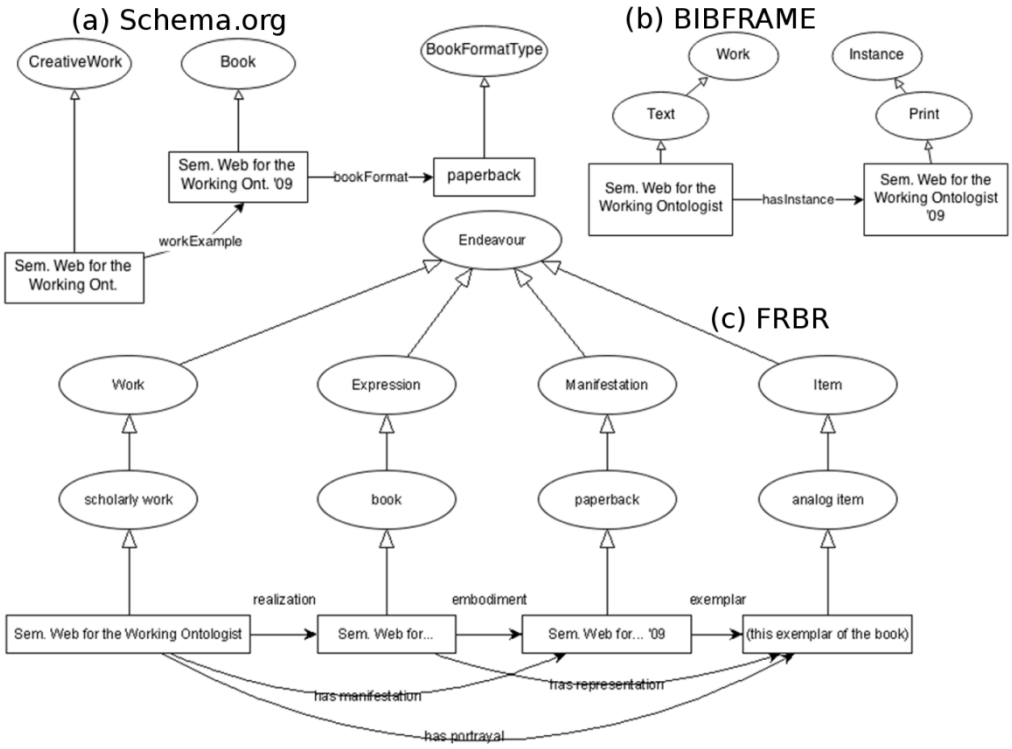


Figure 2 - Diagrams showing different OWL modeling styles of the relationship between a book and its issue.

For the purposes of the local coverage comparison, we chose the book *The Semantic Web for the Working Ontologist* (as an instance of the *B-type Text*, which is a subtype of *Work*) and its 2009 paperback issue (an instance of the *B-type Paperback*, a subtype of *Book*). An OBM of these two entities accompanied by a *B-object* representing an exemplar of the paperback issue (as a tangible item) and the relationships between them is in Figure 3 (exported as a screenshot from PURO Modeler). The color-coded highlighted parts of the model shows what of it can be expressed in each vocabulary. We can see that, e.g., the physical exemplar of the issue can only be described with FRBR (which "implements" the whole OBM), BIBFRAME allows to describe an issue of a book, but without specifying its format, and Schema.org does not allow to say that an instance of work is a *Text*.

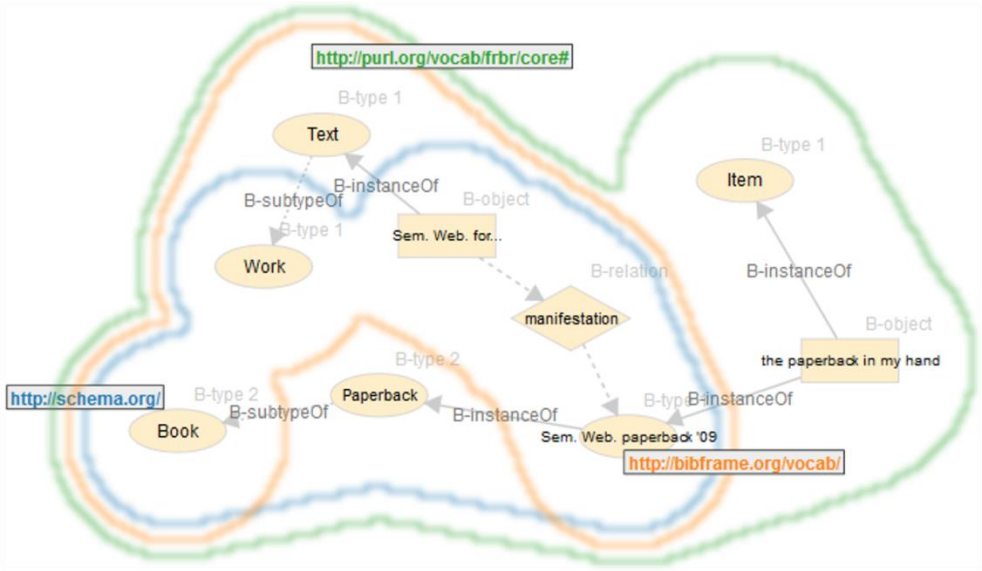


Figure 3 - The OBM of the relationship between a book and its issue. The amount of its coverage in various OWL representations (shown in Figure 2) is highlighted.

Visualizing Ontology Usage in a Dataset

We propose an algorithm that combines type-property paths and characteristic sets into one graph. Type-property path (of length 1) is a sequence "type1 - property - type2"; to stress its bipartite character, we will write it as "type1/property\type2". Type1 and type2 are the types of instances from the dataset that are connected by the property. A characteristic set is a set of predicates that are present in triples with the same subject. The algorithm for summarizing a dataset *D*, parametrized by a threshold *k*, is as follows:

1. Find the set of paths of length 1 with frequency in *D* higher or equal to *k*, i.e. a set of triplets $P = (t1, p, t2)$ such that there are at least *k* triples in *D* such that the predicate is *p*, the subject is instance of *t1* and the object is instance of *t2*.
2. Find the set of characteristic sets in *D*, i.e. a set $C = c_1,...,c_n$ such that each c_i is a characteristic set consisting of properties $p_{i1},...,p_{in}$.
3. The set of all subjects of a given characteristic set *c* will be denoted as $I(c)$.
4. Add possible types of subjects and objects to the predicates in characteristic sets: find the set of all triplets $S = (t1_i, p_k, t2_i)$ such that $c \in C, s \in I(c), p_k \in c$, and "*s* rdf:type $t1_i$.", "*o* rdf:type $t2_i$." and "*s* p_k *o*." are triples in *D*.
5. Create a subset S_p of all triplets $(t1, p, t2) \in S$ such that $(t1, p, x)$ or $(x, p, t1)$ is a triplet from P
6. Let $Sum = P \cup S_p$.
7. Create a graph $G = (V, E)$ where *V* is the set of vertices $\{x: (x, a, b) \in Sum \vee (a, b, x) \in Sum\}$ and *E* is the set of edges (x,y) with label *p* such that $(x, p, y) \in Sum$ (i.e. a graph from the summarized triplets as if they were RDF triples).

Simplified Example

Let us say we found a path BusinessEntity/offers\Offering and a characteristic set {gr:hasBusinessFunction, gr:hasPriceSpecification, gr:includesItem}. After typing the subjects/objects, the characteristic set can be represented by a graph as shown in Figure 4.

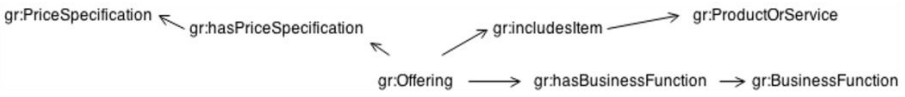


Figure 4 - An example of a typed characteristic set represented as a graph.

The typed characteristic set can then be merged with the path based on the fact that the class gr:Offering} is the subject type of the typed characteristic set and is present in the path. The resulting graph is shown in Figure 5.

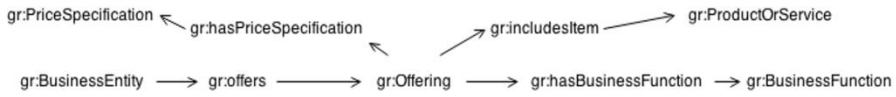


Figure 5 - An example of a path merged with a typed characteristic set.

We have experimentally implemented the summarization algorithm and summarized several datasets. Interactive visualization of the summarizations is available as a web application.[13]

Conclusions and Future Work

We overviewed and analyzed the current ontology visualization tools by taking into account newly formed (aggregated) use case categories, visualization tool features, language constructs, and scalability with respect to ontology size. For selecting a suitable visualizer we designed a simple recommender. We also performed a questionnaire-based evaluation of the recommender and of our analytical findings. The respondents generally agreed with the proposed use case categorization and and they were mostly satisfied with the recommender suggestions. During the analysis, we identified two use cases where

[13] http://lod2-dev.vse.cz/lodsight

none of the current visualization methods known to us seems sufficient and proposed possible more suitable visualization methods:

First, we argued that the task of ontology comparison in terms of local coverage is difficult if the ontologies use different modeling styles. We proposed that using a more general modeling language for visualizing the local coverage of several ontologies in one place, abstracting from the OWL modeling differences, might make the comparison easier, and that the PURO OBM language might be suitable for such a use case. We implemented PURO Modeler: a Javascript application for PURO OBM models design and visualization of the amount of their local coverage in various ontologies. The future work might include, besides the full evaluation, building a portal where the developers could publish visualizations of their ontology/vocabulary local coverage of typical situations.

Second, we suggested that analyzing the usage of a vocabulary in a specific dataset through common dataset browsing tools might be too time consuming. We demonstrated a possible way of achieving the dataset summary visualization through the use of already existing methods for finding type-property paths and characteristic sets. The result is a graph of classes (nodes) and properties (edges), where a connection of two classes by a property means that the dataset contains at least one pair of instances of the two classes connected by the property. We plan to add the possibility of displaying example instantiations of a part of the graph selected by the user. The future work will also of course include evaluation with users, once we overcome the limitations of our approach (mostly related to SPARQL query timeouts due to their complexity), and implementation of the algorithm into a user-friendly application.

References

1. ALBUQUERQUE, A., GUIZZARDI, G., 2013. An ontological foundation for conceptual modeling datatypes based on semantic reference spaces. In: *Research Challenges in Information Science (RCIS), 2013 IEEE Seventh International Conference*. ISBN 978-1-4673-2912-5.
2. BERKA, P., 2011. NEST: A Compositional Approach to Rule-Based and Case-Based Reasoning. In: *Advances in Artificial Intelligence*. ISSN 1687-7470.
3. BROCKMANS, S. et al. 2004. Visual modeling of OWL DL ontologies using UML. In: The Semantic Web - ISWC 2004. Springer Berlin Heidelberg.
4. BRUNETTI, J. M., GARCÍA, R., AUER, S., 2013. From Overview to Facets and Pivoting for Interactive Exploration of Semantic Web Data. In: *Int. J. Semantic Web Inf. Syst*. 9, 1–20. DOI 10.4018/jswis.2013010101
5. CAMPINAS, S. et al. 2012. Introducing RDF graph summary with application to assisted SPARQL formulation. In: *DEXA '12 Proceedings of the 2012 23rd International Workshop on Database and Expert Systems Applications*. IEEE. ISBN 978-0-7695-4801-2.
6. CHEN, P., 1976. The entity-relationship model -- toward a unified view of data. In: *ACM Transactions on Database Systems*. DOI 10.1145/320434.320440.
7. CHOI, N., SONG, I. Y., HAN, H., 2006. A survey on ontology mapping. In: *ACM Sigmod Record*. **35**(3), 34-41.
8. DA SILVA, I. C. S., FREITAS, C. M. D. S., SANTUCCI, G., 2012. An integrated approach for evaluating the visualization of intensional and extensional levels of ontologies. In: *Proceedings of the 2012 BELIV Workshop: Beyond Time and Errors-Novel Evaluation Methods for Visualization*. ACM. p. 2.
9. ERNST, N. A., STOREY, M.-A., 2003. A Preliminary Analysis of Visualization Requirements in Knowledge Engineering Tools. University of Victoria.
10. FREITAS, C. M. D. S. et al. 2002. On evaluating information visualization techniques. In: *Proceedings of the working conference on Advanced Visual Interfaces*. ACM. p. 373-374.
11. GOYAL, S., WESTENTHALER, R. 2004. Rdf gravity (rdf graph visualization tool). Salzburg Research.
12. HAJEK, P., 1985. Combining functions for certainty degrees in consulting systems. In: *International Journal of Man-Machine Studies*. **22**(1).

13. HAYES, P. et al. 2005. Collaborative knowledge capture in ontologies. In: *Proceedings of the 3rd international conference on Knowledge capture - K-CAP 2005.* ACM.
14. KATIFORI A. et al. 2006. A comparative study of four ontology visualization techniques in protege: Experiment setup and preliminary results. In: Information Visualization. IEEE. p. 417-423. ISBN 0-7695-2602-0.
15. KATIFORI, A. et al. 2007. Ontology visualization methods - a survey. In: ACM Computing Surveys (CSUR). ACM. 39(4). ISBN 978-1-4503-0548-8.
16. KENDALL, E. F et al. 2009. Towards a Graphical Notation for OWL 2. In: *Proceedings of the 6th International Workshop on OWL: Experiences and Directions*. CEUR.
17. MAEDCHE, A., STAAB, S., 2002. Measuring similarity between ontologies. In: *Knowledge engineering and knowledge management: Ontologies and the semantic web*. Springer Berlin Heidelberg. p. 251-263.
18. NEUMANN, T., MOERKOTTE, G., 2011. Characteristic sets: Accurate cardinality estimation for RDF queries with multiple joins. In: *Proceedings of the 27th International Conference on Data Engineering*. IEEE. p. 984-994. ISBN 978-1-4244-8958-9.
19. PARREIRAS, F. S., WALTER, T., GRONER, G., 2010. Visualizing ontologies with UML-like notation. In: Ontology-Driven Software Engineering. ACM. ISBN 978-1-4503-0548-8.
20. PHAM, M., 2013. Self-organizing structured RDF in MonetDB. In: *2013 IEEE 29th International Conference on Data Engineering Workshops (ICDEW)*. IEEE. p. 310-313. ISBN 978-1-4673-5303-8.
21. PRESUTTI, V. et al. 2011. Extracting core knowledge from Linked Data. In: *Proceedings of the Second Workshop on Consuming Linked Data*. CEUR.
22. SHNEIDERMAN, B., 1996. The eyes have it: A task by data type taxonomy for information visualizations. In: Proceedings of IEEE Symposium on Visual Languages. p. 336-343. ISBN 0-8186-7508-X.
23. SIVAKUMAR, R., ARIVOLI, P. V., 2011. Ontology Visualization Protege Tools: A Review. In: *International Journal of Advanced Information Technology*. **1**(4). Available from: http://airccse.org/journal/IJAIT/papers/0811ijait01.pdf
24. SVÁTEK, V. et al. 2013. Mapping Structural Design Patterns in OWL to Ontological Background Models. In: *K-CAP 2013*. ACM.
25. SWAMINATHAN, V., SIVAKUMAR, R. 2012. A Comparative Study of Recent Ontology Visualization Tools With a Case of Diabetes Data. In: *International Journal of Research in Computer Science*. **2**(3). p. 31.

JEL Classification: D83

Acknowledgement: This research was supported by VŠE IGA F4/34/2014.

Shrnutí

Vizualizace schémat sémantického webu a datasetů linked data

Analyzovali jsme několik nástrojů pro vizualizaci ontologií a slovníků sémantického webu z pohledu kategorií případů užití, podporovaných funkcí a jejich expresivity. Na základě analýzy byl vyvinut na pravidlech založený doporučovací systém, který uživateli pomáhá vybrat vhodný nástroj. Analýza i výsledný doporučovací systém byly evaluovány skrz dotazník. Výsledky dotazníkového šetření potvrzují hypotézu, že různé případy užití vyžadují různé vizualizační metody. Identifikovali jsme dva konkrétní případy užití, kde současné dostupné vizualizační metody mohou být nedostatečné. Prvním je srovnávání lokálního pokrytí slovníků a ontologií. Pro tento případ navrhujeme použití tzv. modelů ontologického pozadí, které může vést ke snáze pochopitelné vizualizaci. Druhým případem je vizualizace použití slovníků v konkrétním datasetu. Zde navrhujeme použití vizualizační metody založené na sumarizaci datasetu s použitím častých cest a charakteristických množin.

Klíčová slova: sémantický web, OWL, vizualizace schémat, vizualizace ontologií, vizualizace datasetu, srovnávání lokálního pokrytí ontologií

Projektové řízení penetračních testů IS

Tomáš Klíma

xklit10@vse.cz

Doktorand oboru Aplikovaná informatika

Školitel: Doc. Ing. Prokop Toman, CSc., (toman@vse.cz)

Abstrakt: V dnešní době dostupné metodiky penetračního testování bezpečnosti informačních systémů trpí ve většině případů neaktuálností a nekomplexností, což značně komplikuje jejich reálné použití. Autor od roku 2011 pracuje na vývoji metodiky PETA, jež se vyvinula z čistě technického popisu testování jednotlivých částí IT/IS infrastruktury ke komplexní metodice, která stojí zejména na principech vrstevnatosti a škálovatelnosti. Tvorba částí, které se zabývají problematikou řízení testů, jednoznačně ukázala, že optimálním přístupem je využití best practices ve formě určitého uznávaného rámce či metodiky. Tyto rámce či metodiky ale není možné slepě přijmout v plné šíři a původní podobě – je třeba je citlivě upravit tak, aby plně pokryly potřeby tak specifické oblasti, jakou je penetrační testování.

Při analýze problematiky projektového řízení penetračních testů se jako optimální ukázala metodika PRINCE2, jež je všeobecně srozumitelná a plní požadavek na snadnou upravitelnost pro specifické použití. V článku budou představeny principy, témata a procesy využitelné pro penetrační testy, které budou následně začleněny do metodiky PETA.

Klíčová slova: Projektové řízení, penetrační test, PRINCE2, PETA

Úvod

Poslední roky v oblasti bezpečnosti IS/IT nebyly pro její obránce nijak příznivé a dokonce se zdá, že pod vlivem úspěšnosti útočníků, ať už jsou to aktéři ze světa organizovaného zločinu, zpravodajských služeb, nebo jen pouzí hacktivisté, se pomalu začíná měnit i náhled na samotné základy tohoto oboru. Z populární zásady „prevent, detect, react“ se pomalu začíná stávat „complicate, detect, react“[14]. Pro ilustraci je také možné použít slova Roberta Muellera, ředitele FBI, jenž v roce 2012 pronesl: „There are only two types of companies: those that have been hacked, and those that will be.“[15] Díky tomu je celý segment IT bezpečnosti vystaven značnému zájmu a snaží se přicházet s novinkami s větším, či menším potenciálem při „komplikování“ jednotlivých útoků[16]. Ať už jde o biometrické formy autentizace, inteligentní monitoring síťových toků či pokročilé systémy pro sběr a analýzu logů typu SIEM. Všechna tato zařízení, či systémy, nás ale odvádějí od prověřeného konceptu, který je poměrně starý a klade důraz na ofenzivní stránku zabezpečení – tj. za účelem zvýšení zabezpečení své infrastruktury musí její ochránce nejprve začít přemýšlet jako útočník, získat jeho schopnosti a následně se i pokusit o reálný průnik – tedy sofistikovaně řečeno provést penetrační test (Mahmood, 2010).

Penetrační test využívá ty samé postupy a techniky jako útočníci za účelem odhalení maxima zranitelností, které by mohly při reálném útoku způsobit bezpečnostní incident. Jeho účelem je ověřit úroveň realizace bezpečnostních politik technickými a dalšími prostředky, a z pohledu útočníka kompromitovat cílový informační systém a identifikovat možné dopady reálného útoku. Přestože se nejedná o nový přístup, situace na poli metodik penetračního testování není nijak uspokojivá. Dostupné metodiky jsou buď zastaralé, nebo nejsou dostatečně komplexní, tudíž nepokrývají všechny oblasti potřebné pro reálný penetrační test. Tato situace přiměla autora k vytvoření metodiky vlastní. Při práci na této metodice bylo jedním z hlavních cílů pokrýt oblasti nedostatečně popsané, mezi něž patří jak jednotlivé speciální postupy (např. testy sociálním inženýrstvím, testy odolnosti vůči DoS[17] útokům), tak zejména propojení řízení (bezpečnosti) podnikového IS/IT s řízením penetračních testů.

[14] Nelze přeložit doslova, záměna prevent za complicate naznačuje, že již nelze útokům jednoznačně zabránit, spíše je jen komplikovat.
[15] „Existují jen dva typy společností: ty které již byly kompromitovány a ty, které to ještě čeká.“
[16] O aktuálnosti a významu problematiky jistě svědčí i zavedení Zákona o kybernetické bezpečnosti v ČR.
[17] Denial of service.

Projektové řízení penetračních testů a metodika PETA

Přestože je každý penetrační test bezpochyby unikátní a musí být proveden v určitém vymezeném období s danými zdroji, tudíž je možné jej považovat za projekt, není známa žádná metodika, která by se touto problematikou hlouběji zabývala. Pokud se podíváme do klasických metodik, pak je možné konstatovat, že např. v ISSAF (OISSG, 2006) lze nalézt základní elementy projektového řízení jako je řízení rizik, WBS (rozpad prací) a plán řešení problémů, či plán komunikace, ale úroveň detailu není na patřičné úrovni, tudíž i reálná použitelnost je nízká. Ostatní metodiky jako OSSTMM (Herzog, 2006) či (SANS, 2010), (NIST, 2008) se projektovým řízením nezabývají vůbec. V současnosti velmi ambiciózní projekt PTES (penetration testing execution standard)[18] se kupodivu projektovým řízením těchto testů také vůbec nezabývá.

O nápravu tohoto stavu se autor snaží ve své metodice PETA. Ta má své základy v roce 2011 kdy byla v práci (Klíma, 2011) představen metodika WIPE pro testování bezpečnosti bezdrátových sítí, která byla následně rozšířena pro použití při komplexních testech podnikového IS/IT. Přestože nejprve bylo cílem popsat zejména specifické postupy testování jednotlivých oblastí (např. testy sociálním inženýrstvím, testy odolnosti vůči DoS útokům) později došlo k přesunutí zájmu směrem k řízení samotných testů, což vyústilo v roce 2014 v analýzu propojení řízení (bezpečnosti) podnikového IS/IT s řízením penetračních testů, na jejímž základě byl sestaven model těžící z kontrolních cílů COBITu, který byl představen v publikaci (Klíma, 2014). Logickým krokem bylo následně se zaměřit na problematiku zvládání rozsáhlejších testů, na kterých již spolupracuje početnější tým jak na straně dodavatele, tak zadavatele. U těchto testů se v plném rozsahu projevuje důležitost řízení komunikace, rozsahu testu, zvládání výjimek a dalších oblastí, které zkušení projektoví manažeři dobře znají z projektů většího rozsahu. Vzhledem k specifičnosti penetračního testování ovšem není možné pouze přejmout best practices projektového řízení, ale je třeba je náležitě upravit. Tato problematika bude diskutována dále v článku, jelikož pro další postup je třeba si v krátkosti přiblížit metodiku PETA.

Metodika je rozdělena do následujících čtyř úrovní:

1. General procedure (Obecný postup)
2. Detailed procedure (Detailní postup)
3. Specific procedures (Speciální postupy)
4. Tools (Nástroje)

První úroveň je nazvána obecný postup a prezentuje hrubé rozdělení činností a projektování penetračních testů. Druhá, detailní, následně definuje podrobný rozpad činností až na úroveň jednotlivých kroků vedoucích od prvotního kontaktu s infrastrukturou až k jejímu, pokud možno úplnému, ovládnutí. Tyto kroky je možné vidět v tabulce 1. Jelikož jejich detailní popis je uveden v (Klíma, 2014), je z tohoto článku záměrně vynechán. Třetí úroveň obsahuje postupy testování konkrétních prvků IS a poslední, čtvrtá, pak popis nástrojů využitých při testu. Poslední dvě zmíněné úrovně jsou též vynechány, jelikož nemají vazbu na problematiku projektového řízení. Představené dělení je použito hlavně proto, že první dvě úrovně mohou být použity i v dlouhodobém horizontu bez ztráty aktuálnosti a pouze třetí a čtvrtou je třeba v budoucnosti aktualizovat.

[18] http://www.pentest-standard.org

Tabulka 1: Detailní postup Zdroj: autor

Planning	Testing	Reporting
1.1) Requirements identification	2.1) Information gathering	3.1) Cleanup
1.2) Stakeholder identification	2.2) Perimeter mapping	3.2) Document analysis
1.3) Project management team creation	2.3) Penetration	3.3) Report creation
1.4) Defining scope	2.4) Network scanning	3.4) Report presentation
1.5) Defining rules	2.5) Vulnerability scanning	
1.6) Testing team appointment	2.6) Penetration further	
1.7) Role description	2.7) Gaining access and escalation	
1.8) Kick off meeting	2.8) IS compromise	
	2.9) Maintaining access	
	2.10) Covering the tracks	

PRINCE2 - principy, témata a procesy

Vzhledem k tomu, že při analýze problematiky projektového řízení penetračních testů se jako optimální ukázalo využití metodiky PRINCE2, jež je všeobecně srozumitelná a plní požadavek na snadnou upravitelnost pro specifické použití, v následujících odstavcích se zaměříme na základní principy, témata a procesy, které jsou využitelná pro penetrační testy a jež budou následně začleněny do metodiky PETA.

Principy PRINCE2 a jejich úprava

Metodika PRINCE2 obsahuje sedm základních procesů, z nichž pět je použitelných pro oblast penetračních testů.

První z nich, Continued business justification[19], neboli zdůvodnění projektu, ovšem patří mezi ty, které nejsou považovány za klíčové. Vzhledem k tomu, že po zahájení penetračního testu nebývá zvykem přehodnocovat smysluplnost pokračování, nebo ho předčasně ukončovat (pokud nedojde k neočekáváným incidentům), je pro další postup tento princip vyřazen. Tomuto rozhodnutí také nahrává fakt, že penetrační testy nejsou projekty, které by zpravidla trvaly déle než týdny či měsíce, tudíž nedochází k revalidaci business casu (obchodního případu) po ukončení každé etapy.

Defined roles and responsibilities, neboli princip definovaných rolí a zodpovědností, je velmi důležitý zejména v rozsáhlejších projektech, při kterých je zapojen větší tým testerů. Jelikož je možné penetrační testy považovat za poměrně riskantní, musí existovat jasný písemný doklad o tom, kdo za co odpovídá (zejména v případě, že během testu dojde k narušení dostupnosti či spolehlivosti testovaného IS). Tento princip je v metodice PETA zakotven zejména v krocích 1.3, 1.6 a 1.7 detailního postupu a formalizován je v dokumentu s názvem *Team structure, rules, responsibilities.*

Focus on products, neboli zaměření na produkty, může na první pohled vypadat jako jeden ze základních principů, ale není tomu tak. Je třeba si uvědomit, že hlavním produktem, či artefaktem, je závěrečná zpráva, která zejména shrnuje nalezené zranitelnosti. To, co přináší zadavateli hodnotu, však není samotný dokument, ale proces jeho tvorby, tedy testování, na který by mělo být upřeno maximum sil. Analogií může být vysokoškolské studium, jehož přínosem není samotný diplom, ale spíše proces samotného získávání znalostí (resp. samotné znalosti).

Manage by stages, neboli řízení po etapách, patří mezi velmi důležité principy hned ze dvou důvodů. Jednak nám umožňuje sledovat metodicky správný postup rozdělený do logických částí, a také jasně

[19] Názvy jsou úmyslně (v celém článku) ponechány v původním znění.

odděluje fáze, ve kterých je tester v kontaktu s auditovaným IS, od těch, kdy se pouze připravuje, či se naopak věnuje tvorbě konečné dokumentace a na samotný IS „neútočí“. To je velmi důležité pro jasné odlišení problémů či výpadků, které má na svědomí tester, a těch zapříčiněných běžnými provozními okolnostmi. Čtenář má možnost se s principem dělení na jednotlivé fáze seznámit jak výše v tabulce 1, tak podrobněji níže, na obrázku 1.

Manage by exceptions, neboli řízení na základě výjimek, má v kontextu penetračních testů lehce odlišný význam od obecného chápání. Zatímco v klasických projektech bývá příčinou výjimky zpravidla přečerpání rozpočtu, či nedodržení termínu, u pentestů bývá touto příčinou již několikrát zmíněná degradace kvality služeb, či jiný bezpečnostní incident na straně zadavatele (kupř. při testech pomocí sociálního inženýrství může tímto incidentem být i zadržení samotného testera ostrahou). Pokud se tedy incident objeví, je třeba okamžité reakce, která by měla být iniciována na základě rozhodnutí managementu jak zadavatelské, tak dodavatelské strany, jelikož jedině tak je možné eliminovat potencionální škody na obou stranách. Problémy, vyskytnuvší se během projektu jsou zanášeny do dokumentu s názvem *Issue register* a měly by o nich být zmínka i v *Pentester diary*.

Learn from experience, neboli učení se na základě zkušenosti, je důležitý princip zejména v prostředí, kdy na testech spolupracuje několik týmů, které by měly mít přístup do určité „databanky moudrosti“, neboli formálně zaznamenaných zkušeností z předchozích projektů. Přestože každý informační systém a každá IT infrastruktura je bezesporu unikátní, společné prvky, ať už na úrovni používaného hardwaru, či operačních systémů, umožňují přenositelnost těchto zkušeností. V metodice PETA je tento princip zastoupen dokumentem s názvem *Lessons learned*, který je na konci projektu sestavován z průběžných poznámek.

Tailor to suit the project environment, neboli úprava pro projektové prostředí, je nejen jedním ze základních principů projektového řízení penetračních testů, ale na vyšší úrovni abstrakce i celé metodiky PETA, která nabádá její uživatele, aby vždy využili jen tu část, která je pro jejich projekt relevantní a ještě si tuto vybranou část upravili dle svých preferencí. Auditor by měl vždy vycházet z plného rozsahu metodiky a tu si dále zužovat a „ohýbat“, až mu vznikne metodika „upravená“, vhodná pro konkrétní projekt.

Tabulka 2: Relevantní principy dle PRINCE2 Zdroj:autor

PRINCE2	Relevantní
Continued business justification	NE
Defined roles and responsibilities	ANO
Focus on products	NE
Manage by stages	ANO
Manage by exceptions	ANO
Learn from experience	ANO
Tailor to suit the project environment	ANO

Témata PRINCE2 a jejich úprava

PRINCE2 popisuje sedm témat, z nichž je možné vybrat šest, která jsou uplatnitelná pro oblast penetračních testů.

Prvním z nich je téma Business case, neboli obchodní případ. Toto téma, považované za jeden z hlavních “driverů” projektového řízení, má též výsadní postavení v metodice PETA, přičemž je stanoveno, že pokud již proběhne zvážení všech rizik a přínosů na straně zadavatele a je přistoupeno k jeho zahájení, nemělo by již dojít k pochybám o smysluplnosti takového testu. K revalidaci obchodního případu během doby konání testu by tedy nemělo docházet. Obchodní případ je definován zejména v krocích 1.1 – 1.2 detailního postupu.

Téma Organization (organizace) má vazbu zejména k personálnímu obsazení jednotlivých pozic nutných ke zdárnému provedení testu. Součástí je též nutnost ustanovení jasného a formalizovaného plánu komunikace, který minimalizuje nejasnosti a chyby, zejména při řešení a eskalaci problémů a delegaci jednotlivých úkolů. V metodice PETA jsou relevantními body 1.3, 1.5, 1.6 a 1.7 a relevatními dokumenty pak *Team structure, roles, responsibilities* a *Communication strategy.*

Téma Quality (kvalita) je možné zmínit jen okrajově, jelikož jeho obsah je čtenáři zřejmý – pokrývá jak požadavky na kvalitu samotných testovacích postupů, tak formálních výstupů, zejména finální zprávy.

Plans, neboli plány, jsou obsahem dokumentu *Test plan and description* a jsou stanovovány zejména v kroku 1.4 a 1.5 detailního postupu. Úroveň jejich detailu je pak na dohodě zadavatelského a dodavatelského týmu, resp. jejich vedení.

Management rizik, popsaný v tématu Risk (riziko) je pro oblast penetračních testů zásadní, jelikož pouze jeho bezchybné zvládnutí umožní provést test kvalifikovaně a bez negativního dopadu na testovaný systém. Rizika by měla být identifikována a ohodnocena již před začátkem testu. Pokud nejsou rizika eliminována pomocí protiopatření, je nezbytná jejich formální akceptace ze strany vedení zadavatelského týmu. Není nutné pro jejich evidenci sestavovat samostatný dokument, stačí, když budou zahrnuta do *Test plan and description.*

Změna (Change) byla vyhodnocena jako nadbytečná, jelikož, jak již bylo zmíněno výše, po spuštění testu jsou jakékoliv zásadní změny nevítané a měly by být akceptovány jen v případě, že není jiného východiska.

Téma Progress (postup) má oporu v celém metodickém postupu dle PETA, přičemž základním dokumentem je *Stage progress report*. Detaily budou probrány v následujícím odstavci.

Tabulka 3: PRINCE2 – témata Zdroj:autor

PRINCE2	**Relevantní**
Business case	ANO
Organization	ANO
Quality	ANO
Plans	ANO
Risk	ANO
Change	NE
Progress	ANO

Procesy PRINCE2 a jejich úprava

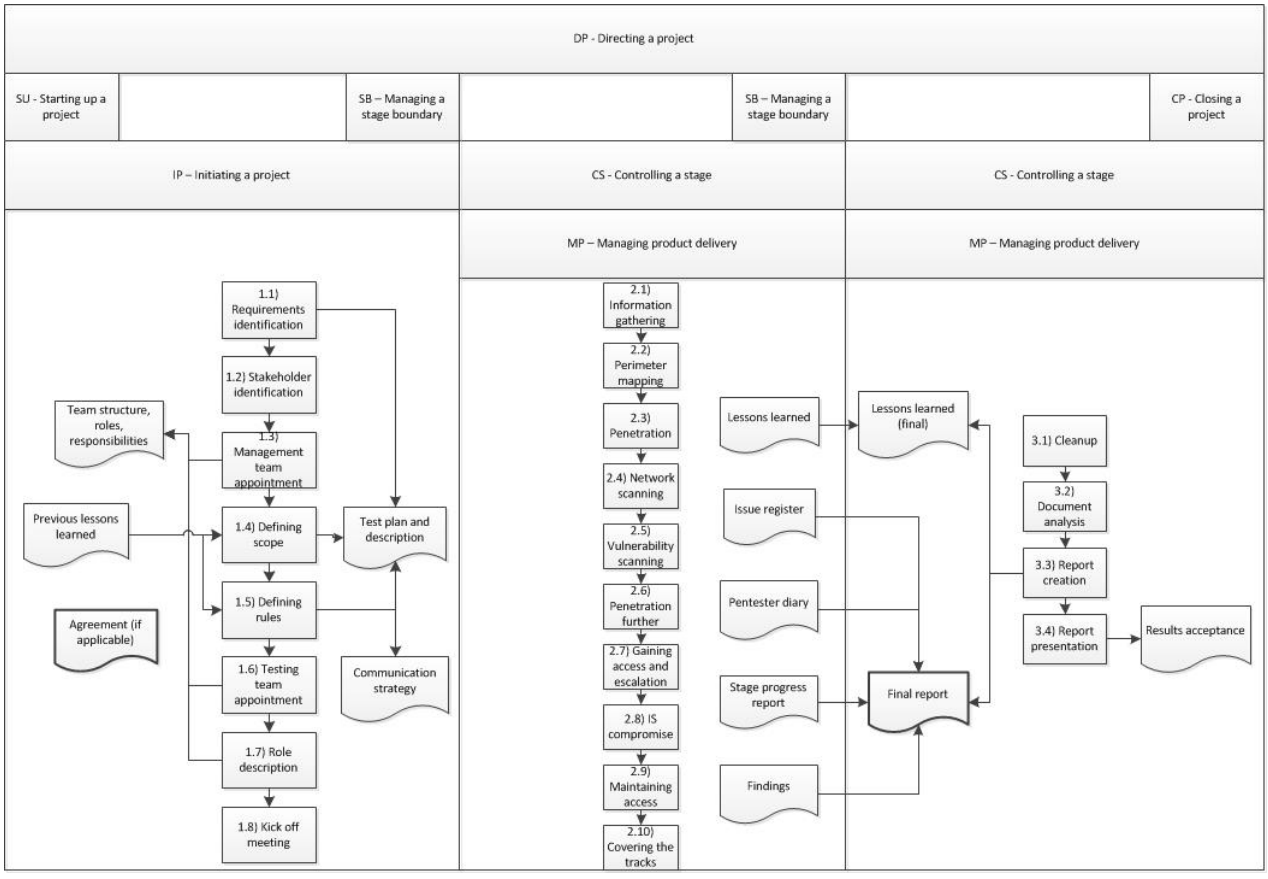
Všechny procesy uvedené v metodice PRINCE2 je možné považovat za relevantní pro penetrační testy. Jelikož správné pochopení úlohy těchto procesů při řízení penetračních testů je nezbytné pro další postup, bude v tomto odstavci představeno mapování na jednotlivé kroky obecného postupu PETA.

Starting up a project (nastartování projektu) a Initiating a project (inicializace projektu) lze hrubě namapovat na část plánování v obecném postupu metodiky PETA, přičemž během nich dochází k přípravným pracím a tester ještě není ve styku s cílovou infrastrukturou. Directing a project (řízení projektu) je průřezový proces, který běží po celou dobu trvání projektu a zaštiťuje ostatní procesy. Controlling a stage (kontrolní bod etapy) má vazbu na fázi testování a reportingu dle obecného postupu. Managing product delivery (řízení dodání produktu) není nijak významný proces, nicméně patří mezi relevantní, a v metodice PETA je zejména svázán s dokumentováním nálezů, problémů a postupu samotného testování a následné "destilace" výsledných shrnujících dokumentů, jež jsou předány zadavateli. Managing stage boundary (řízení rozhraní etap) má význam hlavně z důvodů vymezení kroků, při kterých je tester v kontaktu s cílovou infrastrukturou (jak již bylo poznamenáno výše). Closing a project (uzavření projektu) souvisí s řádným předáním výstupů a jejich akceptací ze strany zadavatele.

S využitím výše zmíněných souvislostí mezi metodikami PRINCE2 a PETA lze sestavit diagram (viz obrázek 1), který nejen graficky znázorňuje detailní postup uvedený v tabulce 1, ale zároveň i mapování procesů mezi oběma metodikami. Též jsou z diagramu patrné vazby jednotlivých doporučených dokumentů na relevantní kroky testování.

Tabulka 4: PRINCE2 – procesy Zdroj:autor

PRINCE2 process	Relevant
Starting up a project	ANO
Initiating a project	ANO
Directing a project	ANO
Controlling a stage	ANO
Managing product delivery	ANO
Managing stage boundary	ANO
Closing a project	ANO



Obrázek 8: Mapování PETA - PRINCE2 Zdroj:autor

Využití v praxi

Metodika PETA byla zatím použita při třech penetračních testech reálných informačních systémů.

Prvním z nich byl externí test infrastruktury, při kterém bylo třeba překonávat ochranu perimetru a tester nebyl vybaven žádnými předběžnými informacemi, jednalo se tedy o black – box přístup (administrátoři o plánovaném testu věděli). Zde byl ověřen zejména princip propojení řízení testů s řízením podnikového IT ve formě maximalizace přínosů zjištěných zranitelností, které byly dle kontrolních cílů COBITu distribuovány patřičným řešitelům.

Druhým byl test bezdrátových sítí, který byl celkově řízen dle metodiky PETA a bylo využito zejména speciálního postupu s názvem Wireless networks (Bezdrátové sítě) a relevantní části detailního postupu. Test byl proveden s částečnou znalostí testované části infrastruktury a za dohledu administrátorů. Výsledkem byl poměrně vysoký stupeň kompromitace.

Třetí test byl specializovaný na sociální inženýrství, při kterém byla simulována phishingová kampaň směřovaná na koncové zaměstnance. Byly využity zejména principy speciálního postupu Social engineering (Sociální inženýrství) a relevantní části detailního postupu. Jelikož se jedná o oblast, která běžně testována není, šlo o určitý krok do neznáma, který se ale ukázal jako velmi podstatný pro celkovou úroveň bezpečnostního povědomí a bezpečnostní odolnosti organizace.

Celkově je tedy možné říci, že v praxi byl prokázán přínos relevantních částí metodiky PETA – oproti předcházejícímu stavu, kdy bylo mnoho částí řešeno ad hoc dle různých nestrukturovaných popisů, např. (Selecký, 2012), (Scambray, 2009), (Dimkov, 2009), dochází k zlepšení komunikace, sofistikovanějšímu využití nálezů při zvyšování úrovně bezpečnosti IT/IS infrastruktury a obecně k přehlednějšímu řízení.

Závěr

Od roku 2011, kdy došlo k položení základních kamenů, se metodika PETA vyvinula z čistě technického popisu testování jednotlivých částí IT/IS infrastruktury ke komplexní metodice, která stojí zejména na principech vrstevnatosti a škálovatelnosti. Tvorba částí, které se zabývají problematikou řízení jednoznačně ukázala, že optimálním přístupem je využití best practices ve formě určitého uznávaného rámce či metodiky, která umožní snadnou pochopitelnost a přenositelnost. Tyto rámce či metodiky ale není možné slepě přijmout v plné šíři a původní podobě – je třeba je citlivě upravit tak, aby plně pokryly potřeby tak specifické oblasti, jakou je penetrační testování. Po využití rámce COBIT tedy byla pro potřeby projektového řízení zvolena metodika PRINCE2, která bezpochyby požadavek na všeobecnou uznávanost a pochopitelnost plní. Značnou výhodou této metodiky je též fakt, že její podstata je velmi jednoduchá a umožnuje dostatečnou volnost při úpravách.

Metodika PETA byla zatím úspěšně použita při třech penetračních testech reálných informačních systémů, přičemž při její validaci je dále plánováno využití simulace pro testování virtualizovaných informačních systémů postavených na základě technologií firem Oracle a VMWare. Též dojde ke komparativní analýze umožňující postihnout hlavní odlišnosti (a oblasti zlepšení) od současně používaných metodik.

Co se týká dalšího postupu, autor plánuje zejména doplnění mapování na COBIT 5 (v současné době je provedeno mapování na verzi 4.1) a dokončení jednotlivých speciálních testovacích postupů.

Literatura

DIMKOV, Trajce. Two methodologies for physical penetration testing using social engineering. Centre for Telematics and Information Technology University of Twente, Enschede, 2009, ISSN 1381-3625

HERZOG, Peter . Open-Source Security Testing Methodology Manual [online]. [s.l.]: ISECOM, 2006. Dostupné z WWW: <http://www.isecom.org/osstmm/>.

KLÍMA, Tomáš. HODNOCENÍ BEZPEČNOSTI BEZDRÁTOVÝCH SÍTÍ. Praha, 2011. Diplomová práce. KIT VŠE. Vedoucí práce L. Pavlíček.

KLÍMA Tomáš. Využití metodiky COBIT 4.1 při penetračním testování bezpečnosti IS. In *Sborník prací vědeckého semináře doktorského studia FIS VŠE.* Praha: Oeconomica, 2014, s. 43--49. ISBN 978-80-245-2010-0.

MAHMOOD, M. Adam. MOVING TOWARD BLACK HAT RESEARCH IN INFORMATION SYSTEMS SECURITY. MIS Quarterly. 2010, roč. 3, č. 34.

NIST. Technical Guide to Information Security Testing and Assessment: NIST Special Publication 800-115. 2008.

OFFICE OF GOVERNMENT COMMERCE. Managing successful projects with Prince2. 5th ed. London: TSO, 2009, xii, 327 s. ISBN 978-011-3310-593.

OPEN INFORMATION SYSTEMS SECURITY GROUP. Information systems security assessment framework [online]. [s.l.] : Open information systems security group, 2006. Dostupné z WWW: <http://www.oissg.org/downloads/issaf-0.2/index.php>.

SANS INSTITUTE. Penetration testing in the financial services industry. 2010.

SCAMBRAY, Joel; MCCLURE, Stuart: Hacking exposed, New York 2009, McGraw-Hill, ISBN: 978-0-07-161375-0.

SELECKÝ, Matúš. Penetrační testy a exploitace. 1. vyd. Brno: Computer Press, 2012, 303 s. ISBN 978-80-251-3752-9.

JEL Classification: L80, L86, M10, M15

Summary

Project management of IS penetration tests

Nowadays the methodologies of IS security penetration testing are in most cases out of date and incomplete, which greatly complicates their use in a real world. Since 2011 the author has been developing the methodology named PETA, which has evolved from a purely technical description of the testing steps to a comprehensive methodology. Projecting management guidelines of the methodology clearly showed that optimal approach is to use the best practices in form of certain recognized framework or methodology. These frameworks or methodologies can´t be blindly accepted in original form - it should be carefully tailored to fully meet the needs of a specific sector - penetration testing.

During the analysis of project management approaches to penetration tests PRINCE2 proved to be optimal methodology, which is universally understandable and fulfills the requirement for easy tailoring for specific applications. The article will introduce the principles, themes and processes that are useful for penetration testing, which will subsequently be incorporated into the PETA methodology.

Keywords: Project management, penetration test, PRINCE2, PETA

Analýza vlivů na sourcingové strategie IS/ICT ve zdravotnických zařízeních

Martin Potančok

xpotm03@vse.cz

Doktorand oboru Aplikovaná informatika

Školitel: prof. Ing. Jiří Voříšek, CSc., (vorisek@vse.cz)

Abstrakt: Příspěvek se zabývá problematikou sourcingových strategií informačních systémů a informačních a komunikačních technologií ve zdravotnických zařízení.

Cílem příspěvku je analyzovat faktory ovlivňující sourcingové strategie IS/ICT ve zdravotnictví. V souvislosti s tímto cílem jsou kladeny následující výzkumné otázky, jejichž odpovědi poslouží k dosažení hlavního cíle. Jaké faktory ovlivňují sourcingové strategie ve zdravotnických zařízeních? Jaká je klasifikace faktorů ovlivňujících sourcingové strategie ve zdravotnických zařízeních?

Výzkum v rámci tohoto příspěvku je založen na případové studii, kvalitativním výzkumu. Pro analýzu faktorů ovlivňujících sourcingové strategie ve zdravotnictví byl zvolen dvoufázový postup. Nejprve byla provedena analýza na obecné úrovni, na základě které byl definován prvotní rámec faktorů, který byl následně doplněn o problematiku zdravotnictví a využit při případové studii provedené v NEMOCNICI. Finální rámec obsahuje 18 faktorů a jejich klasifikaci. Pro každý z nich byl definován způsob zjištění vlivu a následného dopadu na strategie. Případová studie dále ukazuje, že přímí účastníci rozhodovacího procesu nepřikládají faktorům náležitou pozornost.

Závěry analýzy lze využít zdravotnickými zařízeními (především vrcholovým vedením a pracovníky IS/ICT oddělení) při rozhodovacím procesu o sourcingové strategii zejména při její tvorbě.

Klíčová slova: informační strategie, IS/ICT, řízení IS/ICT, sourcingová strategie, zdravotnictví

Úvod

Sourcingová strategie informačních systémů a informačních a komunikačních technologií (IS/ICT) ve zdravotnictví je velmi úzce svázaná s informační strategií daného zařízení [Voříšek et al., 2008]. Společná symbióza obou strategií je klíčová vzhledem k narůstající důležitosti IS/ICT používaných ve zdravotnických zařízení [Karpecki, 2010]. Tyto systémy a technologie lze již v současné době považovat za kritické prvky většiny léčebných a části ošetřovatelských procesů.

Sourcingovou strategii definuji jako **„dlouhodobý záměr činností, jehož cílem je navrhnout koncepci sourcingu IS/ICT tak, aby byly podpořeny cíle organizace".** Definice vychází ze základů informační strategie [Bruckner et al., 2012] a navazuje na sourcing, který je vnímán jako *„podnikový proces, jehož cílem je: rozhodnutí o tom, které služby, procesy a zdroje má podnik zajišťovat sám a které přenechat externím poskytovatelům, výběr nejvhodnějších poskytovatelů služeb, sepsání smluv s poskytovateli o obsahu a úrovni poskytovaných služeb, kontrola poskytovaných služeb a řízení vztahů s externími poskytovateli" [Voříšek et al., 2008]*. Pokud používám pojem sourcingová strategie bez uvedení specifikací, vždy je myšlena sourcingová strategie IS/ICT.

Vzhledem k neustále rostoucím možnostem outsourcingu [Mechling, 2002]**, stále se zvětšujícímu trhu tohoto segmentu a předpokládané růstové tendenci do budoucna, je nutné správně nastavit sourcingovou strategii tak, aby mohlo dojít k efektivnímu využívání jednotlivých forem sourcingu.** Situaci na trhu outsourcingových služeb dokládají očekávání společnosti Gartner [Rivera et al., 2013]. V roce 2014[20] by, podle dosavadních odhadů, mělo dojít k nárůstu o 3,1 % na celkovou hodnotu 3,8 trilionu amerických dolarů [Rivera et al., 2014]. Je nutné poznamenat, že se jedná o data bez rozlišení oblasti podnikání a za celé IS/ICT odvětví. V tomto ohledu jsou také velmi pozitivní názory IS/ICT managementu, kde je očekáván do roku 2020 nárůst objemu outsourcingu v průměru o 15 % ve většině

[20] Skutečná data o trhu za rok 2014, v době psaní příspěvku (1. čtvrtletí 2015), nebyla zveřejněna. Stále zůstává platná perspektiva 2014-2017. [Rivera et al., 2013]

hlavních odvětví (finanční sektor, automobilový průmysl a finanční služby) [Eul et al., 2013]. Specifická data a odhady pro outsourcing IS/ICT ve zdravotnictví korelují s výše uvedenými daty. Do roku 2018 je očekáván růst o 7,6 %. Velmi podstatnou roli zde hrají zdravotnické informační systémy [MarketsandMarkets, 2013]. Zájem o outsourcing ve zdravotnictví dokládá také L. Willcocks, I. Oshri a J. Hindle zařazením zdravotnictví do skupiny s oblibou IS/ICT outsourcingu [Willcocks et al., 2009].

Cíle a výzkumné otázky

Hlavním cílem příspěvku je analyzovat faktory ovlivňující sourcingové strategie IS/ICT ve zdravotnictví. V souvislosti s tímto cílem jsou kladeny následující výzkumné otázky, jejichž odpovědi poslouží k dosažení hlavního cíle.

- Jaké faktory ovlivňují sourcingové strategie ve zdravotnických zařízeních?
- Jaká je klasifikace faktorů ovlivňujících sourcingové strategie ve zdravotnických zařízeních?

Výsledky studie [Diana, 2009] ukazují, že rozhodnutí ve zdravotnických zařízení jsou ovlivňována odlišnými faktory než rozhodnutí v jiných odvětvích. Z důvodu efektivní práce s faktory a možnosti jejich ovlivnění je pro přímé účastníky rozhodovacího procesu o sourcingové strategii podstatné vědět, které faktory mohou být ovlivněny přímo a které ne (klasifikace faktorů).

Struktura příspěvku je následující. Nejprve jsou analyzovány a popsány faktory ovlivňující sourcingové strategie na obecné úrovni, tedy bez rozlišení odvětví a následně je cíleno již pouze na oblast zdravotnictví. Po určení jednotlivých faktorů je určena jejich klasifikace.

Metody výzkumu

Výzkum v rámci tohoto příspěvku je založen na případové studii, kvalitativním výzkumu. Návrh a postup případové studie použité v tomto příspěvku vychází z [Lacity et al., 1996]. Pro případovou studii bylo vybráno zdravotnické zařízení střední velikosti [Feige et al., 2013] typu okresní nemocnice v České republice[21], které si nepřeje být jmenováno. Z tohoto důvodu všude, kde by se měl vyskytnout konkrétní název, použiji označení „NEMOCNICE“. Mezi zainteresované strany této studie patří pacienti, zástupci vrcholového vedení NEMOCNICE a pracovníci NEMOCNICE. Tyto tři strany zabezpečují dostatečně rozdílné pohledy na rozhodování o sourcingové strategii a tedy faktory ovlivňující sourcing.

V průběhu případové studie byly provedeny rozhovory jak s přímými účastníky rozhodovacího procesu o sourcingové strategii (vrcholové vedení a vedoucí pracovníci IS/ICT), tak nepřímými[22] (pacienti a pracovníci NEMOCNICE). Délka rozhovorů se pohybovala od 30 minut (pro nepřímé účastníky) do 3 hodin (pro přímé účastníky). Počátek každého rozhovoru byl nestrukturovaný tak, aby mohl být vyslechnut celý příběh a pohled účastníka. Následovala polostrukturovaná část s otázkami vyplývajícími ze základního okruhu faktorů ovlivňující sourcingové strategie. Dalším zdrojem informací bylo několik verzí sourcingové strategie, interní poznámky k její tvorbě, výroční zprávy a organizační struktura NEMOCNICE.

Případová studie byla provedena do hloubky v rámci středě velkého zdravotnického zařízení. Jak dokládají výsledky studie [Potančok, 2013], tato kategorie zdravotnického zařízení zahrnuje společné prvky IS/ICT malých i velkých zařízení. Předpokladem je tedy vysoká pravděpodobnost platnosti výsledků tohoto příspěvku pro ostatní kategorie ze státního sektoru.

Použití kvalitativního výzkumu v kontextu tohoto příspěvku je plně v souladu s [Yin, 2009] a [Myers, 2013].

---

[21] Jedná se o typického zástupce kategorie středně velkých lůžkových zařízení (vedle velikosti splňuje také další společné parametry, například spádová oblast, diverzifikace, specializace, vlastnictví).

[22] Nepřímými účastníky jsou označovány zainteresované strany, na které může mít změna sourcingové strategie IS/ICT dopad, tedy především pacienti, lékaři, zdravotní sestry a ostatní pracovníci IS/ICT oddělení.

Faktory ovlivňující sourcingové strategie IS/ICT ve zdravotnictví

Z definice problému a analýzy situace popsané v úvodu vyplývá jednoznačná nutnost zabývat se faktory ovlivňující sourcingové strategie IS/ICT. Aby bylo možné uceleně pokrýt celou oblast, byl zvolen následující dvoufázový postup.

Analýza specifických faktorů ovlivňující sourcingové strategie se nejprve zaměřila na faktory na obecné úrovni, tedy faktory platné bez určení odvětví a území, pomocí kterých byl definován počáteční okruh faktorů, který v příspěvku není uveden z důvodů popsaných cílů metod výzkumu a rozsahu. Počáteční okruh faktorů byl následně využit v případové studii (viz metody výzkumu) jako výchozí bod. Celý okruh byl validován a rozšířen pro podmínky zdravotnictví.

Následující část obsahuje výsledky případové studie a pro každý faktor uvádí jeho charakteristiku (důvody existence), způsoby zjištění jeho hodnot a určuje jeho vliv na strategii. Z důvodu přehlednosti jsou faktory řazeny abecedně.

Cíle, funkce a role IS/ICT
Faktor cíle, funkce a role IS/ICT shrnuje řadu prvků. Nejpodstatnější je vazba na globální strategii zdravotnického zařízení, která zahrnuje podnikové cíle a jejich priority, požadavky vlastníků procesů. [Myers et al., 1997]

Cíle, funkce a role IS/ICT lze nalézt v informační strategii, kde by měla být také zohledněna již zmiňovaná globální strategie. Pro ověření získaných informací je vhodné využít pozorování a polostrukturovaná interview s pacienty, zástupci vrcholového vedení a zaměstnanci zdravotnického zařízení.

Z důvodu shrnující funkce faktoru nelze na obecné úrovni definovat přesnou podobu jeho vlivu na sourcingovou strategii. Přesné požadavky na sourcing kladou až konkrétní cíle, funkce a role IS/ICT. Příkladem může být velmi důležitá role (postavení) systému PACS[23] v léčebném procesu, která vede k provozování vlastními silami. Na druhé straně méně podstatná role doplňkového systému Vitalmonitor[24] podporuje outsourcing.

Náklady a přínosy sourcingové strategie
Náklady a přínosy sourcingové strategie jsou zohledňovány v rámci analýzy nákladů a přínosů, která by měla být prováděna pro všechny varianty zvažovaného sourcingu. Jak uvádí model Management of Business Informatics (MBI): *„Aplikace a provoz aplikací jsou spojeny s určitými náklady. Ty by pochopitelně neměly přesáhnout přínosy, které projekt/aplikace přinesou. Proto je třeba u jednotlivých projektů/aplikací a služeb sledovat jejich náklady a přínosy.“ [Voříšek et al., 2012]*

Pro určení nákladů jednotlivých částí IS/ICT slouží odhady cen projektů a následné náklady na provoz. Výpočet musí zahrnovat také zvažovaný sourcingový přístup. Přínosy nemusí mít, a ve zdravotnictví zpravidla nemívají, finanční povahu. Z tohoto důvodu je nezbytné stanovit metriky, pomocí kterých budou měřeny, vyhodnocovány a číselně (finančně) vyjádřeny.

Náklady by nikdy neměly přesáhnout přínosy. Ve vztahu k sourcingu by interní ceny za služby neměly přesahovat ceny na trhu. Při příliš vysokých interních cenách by mělo docházet k outsourcingu.

Názory lékařů
Lékaři společně s pacienty tvoří většinu uživatelů IS/ICT. Závislost lékařů na IS/ICT a především jejich datech neustále narůstá, viz údaje v úvodní části tohoto příspěvku. Lékaři jsou tedy nepřímými účastníky rozhodovacího procesu a jejich názory mají vliv na sourcing systémů pracujících s pacientskými daty.

Pro zjišťování názoru lékařů jsou vhodná polostrukturovaná interview a průzkumy v podobě dotazníkových šetření. Získávání dat může probíhat kontinuálně nebo cíleně. Je ovšem nezbytné zvážit časovou náročnost těchto aktivit a připravovat je s ohledem na vytíženost lékařů. Velmi podstatný je

---

[23] Picture Archiving and Communication System (PACS) *„označuje komplexní řešení pro práci s obrazovými daty ve zdravotnictví. Obsahuje hardware i software“ [Potančok, 2010]*.

[24] Vitalmonitor je *„systém monitorující životní funkce pacienta pomocí bezkontaktní technologie zabudované přímo v lůžku, všechna data jsou v reálném čase zobrazována zdravotnickému personálu“ [Potančok, 2012]*.

správný výběr vzorku respondentů a případné sestavení dotazníku. Z tohoto důvodu je vhodné využít konzultace statistiků při sestavování požadavků na data, jejich sběru a vyhodnocování tak, aby nedošlo k zanesení nepřenosností nebo ovlivnění výsledků tohoto faktoru.

Negativní názory lékařů spíše omezují rozsah použitelných forem sourcingu. Pokud by měla být vyvolána ztráta důvěry lékařů v IS/ICT z důvodu přenosu služeb, procesů a zdrojů na externí partnery, dochází k omezení outsourcingu. Na druhé straně názory a požadavky pacientů mohou zvětšovat rozsah použitelných forem sourcingu. Příkladem mohou být požadavky na nové části IS/ICT, se kterými nemá zdravotnické zařízení dostatečné zkušenosti (viz faktor technická vyspělost, zkušenosti, znalosti a schopnosti).

Názory managementu na IS/ICT
Názory managementu (byznys managementu) zahrnují především vrcholové vedení zdravotnického zařízení, které je přímým účastníkem procesu rozhodování o sourcingové strategii. Faktor zahrnuje řadu prvků a dílčích postojů, které formují celkové názory managementu na IS/ICT. Jedná se především o podporu a vnímání odpovědnosti, očekávané přínosy a výstupy, záměry společnosti, preferovaný dodavatelský vztah, preferované formy sourcingu, postoj k transferu znalostí, požadavky na finanční efektivitu IS/ICT a systematické myšlení. Je důležité si uvědomit, že do celkového názoru se také promítají politické postoje vlastníků, které musejí být do určité míry reflektovány. Vzhledem k pozici managementu má tento faktor vliv na všechny části IS/ICT.

Zjišťování názorů managementu na IS/ICT probíhá pozorováním chování s doplněním potřebných interview. Obsáhlost problematiky neumožňuje použití průzkumů v podobě dotazníkových šetření.

Názory managementu působí na sourcingovou strategii v podobě preferencí formy sourcingu jednotlivých částí IS/ICT. Velmi podstatná je komunikace těchto názorů směrem k IS/ICT oddělení z důvodu synchronizace činností. Konzervativní přístup a malá informovanost umocňují předsudky a podporují insourcing. Naopak otevřenost a zájem o nové technologie často umožňují outsourcing.

Specializace a procesy
Specializace a procesy jsou interním faktorem, vztahuje se tedy plně k danému zdravotnickému zařízení. Specializace ovlivňuje především úzce zaměřené systémy (např. PACS systémy).

Protože specializace vychází ze zaměření zdravotnického zařízení, je přesně určená již z povahy zařízení. Procesy (léčebné, ošetřovatelské, IS/ICT apod.) jsou hodnoceny podle zralostní úrovně, např. CMMI (Capability Maturity Model) [Paulk et al., 1993].

Specializace zdravotnického zařízení tvoří požadavky na IS/ICT, pokud jsou příliš specifické a IS/ICT s nimi nemá dostatečné zkušenosti, vyhledává pomoc u externích partnerů (outsourcing) a naopak. Procesy se zralostní úrovní třetí (definovaná) [Paulk et al., 1993] a vyšší jsou vhodnější pro přenos na externího partnera, protože již existuje jejich definice. Nedefinované procesy nelze přenášet a musejí zůstat uvnitř zdravotnického zařízení. Důvodem nepřenositelnosti je neexistence definice jednotlivých činností procesu. Aby mohlo dojít k přenosu, musejí být činnosti popsány a proces definován. Tyto vlastnosti jsou následně zakotvené i ve smluvním vztahu mezi externím partnerem a zdravotnickým zařízením, jsou na ně také vázané odměny a sankce.

Standardizace a míra formalizace IS/ICT
Standardizaci a míru formalizace IS/ICT lze dělit na interní (dochází k vytváření interních standardů a norem) a obecně uznávanou (využívají se standardy a normy národně nebo mezinárodně uznávané). Ovlivnění lze pozorovat na celém IS/ICT (pro obecné standardy prolínající se napříč IS/ICT, např. infrastrukturní) nebo na dílčích částech (pro specifické standardy).

Aby bylo možné hovořit o standardizaci a míře formalizace IS/ICT ve vztahu k sourcingu, je nutné se zabývat národně nebo mezinárodně uznávanými standardy (např. datové standardy DASTA[25], HL7[26]). Při ověření dochází k porovnání implementace ve zdravotnickém zařízení s normami.

V případech, kdy zdravotnické zařízení využívá uznávané standardy (ne pouze interní) dochází ke zjednodušení procesu přenosu služeb, procesů a zdrojů na externího partnera (dochází např. k poklesu transakčních nákladů, viz níže). Obecně lze říci, že standardizace a míra formalizace zjednodušuje a podporuje outsourcing.

Stupeň integrace IS/ICT
Faktor stupeň integrace IS/ICT nezahrnuje pouze integraci v rámci IS/ICT, ale také integraci s byznysem. Zohlednění úrovně integrace a množství vazeb je podstatné v sourcingové strategii při přenechání zajišťování služeb, procesů a zdrojů externím partnerům.

Identifikaci stupně integrace IS/ICT lze provádět na základě modelu integrace IS/ICT podniku (integrace IS/ICT s byznysem a integrace IS/ICT) [Voříšek et al., 2008]. Při této identifikaci dochází k určení úrovně integrace (technologická integrace, integrace interních podnikových procesů, integrace podniku s okolím, integrace vizí, integrace hodnot, procesů a IS/ICT služeb, metodická integrace), na kterou zdravotnické zařízení dosahuje.

Vliv stupně integrace IS/ICT na sourcingovou strategii lze pozorovat v rovinách podle zvoleného modelu integrace IS/ICT podniku [Voříšek et al., 2008]. Obecně lze říci, že přílišná integrace na jakékoliv úrovni omezuje možnosti převodu pouze určité části IS/ICT na externího partnera. Určitá úroveň integrace ale naopak outsourcing může spíše podporovat. Příkladem je integrace zdravotnických zařízení na úrovni kraje, která umožňuje využití Business Intelligence (BI) nad daty dané skupiny zařízení a vede tedy k outsourcingu BI řešení.

Technická vyspělost, zkušenosti, znalosti a schopnosti
Faktor technická vyspělost, zkušenosti, znalosti a schopnosti zahrnuje kompletní IS/ICT oddělení. Protože hlavním zaměřením zdravotnického zařízení je léčba pacientů, není kladen příliš velký důraz na zkušenosti pracovníků IS/ICT oddělení. Podle výsledků případové studie 80 % provozních požadavků na IS/ICT ze strany uživatelů zvládne průměrně zkušený IS/ICT specialista. Faktor se tedy nejvíce projevuje u specifických částí, dále poté u vývoje, implementace apod.

Informace o technické vyspělosti, zkušenosti, znalosti a schopnosti lze získat z hodnocení jednotlivých zaměstnanců IS/ICT oddělení.

Získané výsledky faktoru musejí být při tvorbě sourcingové strategie porovnávány s požadavky. Na základě tohoto vyhodnocení dochází poté buď k upřednostňování přenosu služeb, procesů a zdrojů na externího partnera (pokud vyspělost, zkušenosti, znalosti a schopnosti neodpovídají požadavkům) nebo k zajištění služeb, procesů a zdrojů interně (pokud vyspělost, zkušenosti, znalosti a schopnosti odpovídají požadavkům).

Velikost zdravotnického zařízení a jeho struktura
Velikost zdravotnického zařízení a jeho struktura vytvářejí požadavky na podobu IS/ICT oddělení a důležitost jednotlivých IS/ICT procesů. Dané charakteristiky přímo formují sourcingovou strategii v oblasti přípustných možností.

Určení velikosti zdravotnického zařízení a jeho struktury lze provádět na základě počtu lůžek, zaměstnanců, pacientů apod. [Feige et al., 2013]. Údaje vycházející z interních informací zařízení jsou poté dosazeny do klasifikačního systému.

Důležitost IS/ICT procesů napříč kategoriemi zdravotnických zařízení ilustruje [Potančok, 2013]. Větší důležitost vyvolává potřebu vlastních IS/ICT oddělení a vytváří tedy větší možnosti insourcingu.

---

[25] Datový standard MZ ČR (DS, DASTA) je *„datový standard Ministerstva zdravotnictví ČR určenou pro vzájemnou komunikaci mezi zdravotnickými informačními systémy" [Potančok, 2012]*.

[26] Health Level Seven (HL7) je *„standard sloužící pro výměnu, správu a integraci dat sloužících péči o pacienta, administrativě, poskytování a hodnocení zdravotnických služeb" [HL7, 2010]*.

Naopak menší důležitost například u zdravotnických zařízení bez lůžkové části (ZZBLČ) nebo malých zdravotnických zařízení (MZZ) vede k větší míře outsourcingu.

Způsob vedení IS/ICT
Vedení IS/ICT společně s managementem zdravotnického zařízení tvoří přímé účastníky procesu tvorby sourcingové strategie. Vzhledem k této pozici má faktor vliv na všechny části IS/ICT.

Určení používaných způsobů vedení IS/ICT se nesmí spoléhat pouze na dokumentaci IS/ICT oddělení, která se může velmi odlišovat od skutečných přístupů. Z tohoto důvodu by určení mělo probíhat na základě interview se zaměstnanci oddělení a pozorování, ale také na základě analýzy procesů a struktury IS/ICT oddělení, používaných nástrojů apod.

Způsoby vedení IS/ICT vytvářejí preference na formy sourcingu, ale také formují jejich omezení. Je nezbytné, aby požadavky formy byly v souladu se způsoby vedení IS/ICT. Stejně jako u byznys managementu rozhoduje konzervativní nebo inovativní postoj vedoucích IS/ICT. Konzervatizmus upřednostňuje zabezpečení IS/ICT vlastními silami z důvodu obav přenosu IS/ICT na externího partnera a inovativnost společně s požadavkem na maximální efektivnost naopak outsourcing

Geografická poloha zdravotnického zařízení
Geografická poloha zdravotnického zařízení je transformovaná především do vzdálenosti mezi jím a poskytovatelem [Davis et al., 1974]. Kratší vzdálenosti umožňují pohotovou reakci na místě události. Faktor má největší dopad na provozní a servisní části IS/ICT.

Pro určení hodnot geografické polohy je nezbytné zjištění nejen vzdálenosti, ale především dojezdové doby. Tyto parametry musejí být ověřovány pro všechny poskytovatele a vyhodnocovány vzhledem k požadavkům IS/ICT.

Odlehlejší geografická poloha (větší vzdálenost mezi zařízením a poskytovateli) vytváří na sourcingovou strategii požadavek v podobě většího IS/ICT oddělení.

Konektivita
Zdravotnické zařízení přistupuje k datům a vyměňuje si je se svým okolím pomocí internetového nebo síťového připojení. Propustnost neboli tzv. konektiva je charakterizována především rychlostí a dobou odezvy. Vzhledem k důležitosti IS/ICT ve zdravotnictví a nepřetržitému provozu je také velmi podstatná stabilita a její garance. Nejvíce se tento faktor projevuje u extrémně kritických částí IS/ICT, které jsou náročné na datové přenosy (např. PACS systémy).

Hodnoty faktoru jsou zjišťovány pomocí měření (mělo by být prováděno opakovaně a v průběhu různých časových úseků). Dále je nutné posuzovat existenci a množství záložních spojení včetně jejich parametrů. Je vhodné se také zaměřit na vzdálenost mezi zdravotnickým zařízením a poskytovatelem, viz geografická poloha zdravotnického zařízení.

Parametry konektivity musejí být dodrženy podle jednotlivých požadavků IS/ICT v návaznosti na formu sourcingu. Dostatečná/nedostatečná konektivita umožňuje/znemožňuje především outsourcing provozu a využívání služeb datových center.

Dostupnost pracovníků
Faktor dostupnost pracovníků se zaměřuje především na pracovníky IS/ICT, kteří mohou být zaměstnáni ve zdravotnickém zařízení na požadované pozici sloužící k internímu zabezpečení dané služby, procesu nebo zdroje.

Vyhodnocení faktoru dostupnost pracovníků probíhá na základě počtu dostupných požadovaných pracovníků IS/ICT na trhu. Tyto údaje lze zjistit z vydávaných statistik nebo ze zkušeností (interních nebo externích) z výběrových řízení na dané pozice.

Pokud na trhu není dostupné dostatečné množství IS/ICT pracovníků, dochází k omezování možností sourcingu (především insourcingu).

Legislativa
Zdravotnické zařízení se musí striktně řídit platnou legislativou. Tato povinnost, jak již bylo uvedeno v úvodu příspěvku, je spojená především s citlivostí dat a prací s lidským životem. Příklady

legislativních požadavků na radiodiagnostické IS/ICT lze najít v [Potančok, 2010]. Tento faktor má vliv na celé zdravotnické zařízení a tedy na kompletní IS/ICT.

Právní oddělení nebo zástupce zdravotnického zařízení by měli identifikovat požadavky vytvářené legislativou. Tato identifikace by měla být prováděna ve spolupráci se zástupci oddělení IS/ICT.

Forma sourcingu musí plně respektovat platnou legislativu, podle své podoby umožňuje/znemožňuje jednotlivé formy. Tato skutečnost se vztahuje na všechny podoby sourcingu (insourcing, outsourcing, multisourcing atd.). Příkladem může být požadavek na ukládání dat na území České republiky. V takovémto případě nelze umístit datová uložiště do zahraničí a zvolit tedy takovouto formu outsourcingu.

Nabídka na trhu

Nabídka na trhu určuje aktuální možnosti IS/ICT včetně jejich forem vývoje, provozu, podpory atd. Tento faktor je podstatný především pro skupiny inovátorů[27] a early adopters[28], kde působí na nové části IS/ICT. [Kotler et al., 2007] Ostatní skupiny se již pohybují v možnostech IS/ICT s dostatečnou nabídkou na trhu.

K určení nabídky slouží analýza trhu. Lze čerpat z již hotových komerčních analýz, popřípadě je možné formovat požadavek na zcela novou analýzu, která může být prováděna interními nebo externími zdroji.

Aktuální nabídka na trhu vymezuje možné formy sourcingu jednotlivých částí IS/ICT a formuje tedy možnosti, předpoklady a omezení sourcingové strategie. V případě, že je nabídka dostatečná, dochází k podpoře outsourcingu a naopak.

Náklady na pracovní sílu

Náklady na pracovní sílu úzce souvisejí s dostupností pracovníků a následně s faktorem názory managementu na IS/ICT v podobě požadavků na ekonomickou efektivitu IS/ICT. Stejně jako dostupnost pracovníků ovlivňuje specifické části IS/ICT.

Náklady na pracovní sílu se dělí na již existující a plánované (potenciální/potřebné). Vycházejí jak z interních údajů zdravotnického zařízení na mzdové náklady, tak ze statistických údajů průměrné mzdy a požadavků trhu.

V případě, že náklady na pracovní sílu přesahují možnosti zdravotnického zařízení nebo příliš negativně ovlivňují ukazatel ekonomické efektivity IS/ICT, je nezbytné porovnání s nabídkou externích partnerů (viz výše faktor nabídka na trhu). Při příliš vysokých nákladech na pracovní sílu může docházet k omezení možností insourcingu a zabezpečování IS/ICT vlastními silami.

Názory pacientů

Pacienti ve své podstatě tvoří zákazníky zdravotnického zařízení a jsou nepřímými účastníky rozhodovacího procesu, stejně jako lékaři. Toto postavení způsobuje, že názory pacientů mají vztah především k sourcingu systémů pracujících s pacientskými daty. Nedochází k promítnutí do administrativních systémů, apod.

Pro zjišťování názoru pacientů je vhodný průzkum v podobě dotazníkového šetření, sběru názorů na sociálních sítích nebo vyhodnocování aktivity na internetových stránkách zařízení. Je možné volit kontinuální monitoring názorů na formy sourcingu nebo cílené ověření jednotlivých forem případně požadavků na služby, procesy a zdroje. Velmi podstatný je správný výběr vzorku respondentů a případné sestavení dotazníku. I v tomto případě, stejně jako u faktoru názory lékařů, je při sestavování požadavků na data, jejich sběru a vyhodnocování vhodná konzultace statistiků.

---

[27] *„Inovátoři jsou technologičtí, jsou odvážní a vychutnávají si zkoumání nových produktů a osvojování si jejich složitostí. Mohou-li inovaci získat za nízkou cenu, budou rádi provádět alfa nebo beta testování a reportovat „dětské nemoci" výrobku či služby." [Kotler et al., 2007]*

[28] *„Early adopters jsou názoroví lídři opatrně vyhledávající nové technologie, které jim mohou přinést zásadní konkurenční výhodu. Cenově jsou méně citliví a jsou ochotni přijmout nový produkt, pokud přináší personalizované řešení a dobrou servisní podporu (související služby)." [Kotler et al., 2007]*

Negativní názory pacientů spíše omezují rozsah použitelných forem sourcingu. Pokud by mohlo dojít ke ztrátě důvěry pacientů ve zdravotnické zařízení z důvodu přenosu služeb, procesů a zdrojů na externí partnery, dochází k omezení outsourcingu. Na druhé straně názory a požadavky pacientů mohou zvětšovat rozsah použitelných forem sourcingu. Příkladem může být požadavek na platbu poplatků pomocí SMS v NEMOCNICI. IS/ICT oddělení NEMOCNICE nemělo dostatečné zkušenosti a kapacity na zavedení této služby. Muselo tedy dojít k celkovému outsourcingu služby tak, aby byly naplněny požadavky pacientů.

Rozšířenost IS/ICT služeb na trhu
Rozšířenost IS/ICT služeb navazuje na faktor nabídka na trhu. Mohou nastat situace, ve kterých již služby na trhu sice existují, ale jejich rozšířenost je nedostatečná. Zohledněna by měla být především nabídka, ale následně také její akceptace ze strany zdravotnických zařízení.

Rozšířenost IS/ICT služeb na trhu lze identifikovat cenou, počtem poskytovatelů a implementací na daném trhu.

Sourcingová strategie by neměla připustit varianty, ve kterých stupeň rozšířenosti neodpovídá umístění zdravotnického zařízení v související kategorii schopnosti přijímat inovace a nové technologie. [Kotler et al., 2007] Pokud rozšířenost neodpovídá kategorii schopnosti přijímat inovace a technologie musí být upřednostněn insourcingu.

Transakční a nákupní náklady, frekvence transakcí
Faktor transakční a nákupní náklady zahrnuje veškeré náklady spojené se změnou sourcingu dané části IS/ICT. S tím souvisí také frekvence těchto změn neboli transakci. Vliv má na části, u kterých dochází ke změně.

Skutečné transakční náklady lze zjistit až po provedení daných transakcí z důvodu existence skupiny nepřímých nákladů. Při sestavování sourcingové strategie lze tedy vycházet buď z referenčních transakcí, nebo expertních odhadů. Vzhledem k náročnosti transakcí, se jejich počet pohybuje v řádu velmi malých jednotek.

Tento faktor vytváří na sourcingovou strategii požadavek v podobě správného rozhodnutí o sourcingu vzhledem k následujícím nákupním a transakčním nákladům. Nedochází tedy k přímému ovlivnění volby sourcingu.

Tabulka 2: Faktory ovlivňující sourcingové strategie IS/ICT ve zdravotnictví

Faktor	Klasifikace
Cíle, funkce a role IS/ICT	Interní
Názory lékařů	Interní
Názory managementu na IS/ICT	Interní
Specializace a procesy	Interní
Standardizace a míra formalizace	Interní
Stupeň integrace IS/ICT	Interní
Technická vyspělost, zkušenosti	Interní
Velikost a struktura zařízení	Interní
Způsob vedení IS/ICT	Interní
Geografická poloha zařízení	Interní/Externí

Faktor	Klasifikace
Konektivita	Interní/Externí
Náklady sourcingové strategie	Interní/Externí
Přínosy sourcingové strategie	Interní/Externí
Dostupnost pracovníků	Externí
Legislativa	Externí
Nabídka na trhu	Externí
Náklady na pracovní sílu	Externí
Názory pacientů	Externí
Rozšířenost IS/ICT služeb na trhu	Externí
Transakční a nákupní náklady	Externí

Závěr

Pro správné stanovení způsobů zajišťování služeb, procesů a zdrojů spojených s IS/ICT v návaznosti na hlavní cíle organizace slouží sourcingové strategie. Aby byly dodrženy všechny požadavky na tuto strategii, je nutné reflektovat faktory, které na ni více či méně působí.

Pro analýzu faktorů ovlivňujících sourcingové strategie ve zdravotnictví byl zvolen dvoufázový postup. Nejprve byla provedena analýza na obecné úrovni, na základě které byl definován prvotní rámec faktorů, který byl následně doplněn o problematiku zdravotnictví a využit při případové studii provedené

v NEMOCNICI. Finální rámec uvedený v tabulce 1 obsahuje 18 faktorů a jejich klasifikaci (interní, interní/externí a externí). Pro každý z nich byl definován způsob zjištění vlivu a následného dopadu na strategie. Případová studie dále ukazuje, že přímí účastníci rozhodovacího procesu nepřikládají faktorům náležitou pozornost. Je ale nezbytné, aby se snažili pokrýt celkové spektrum a nedocházelo k opomíjení, které by mohlo vést k chybám ve strategii.

Výše uvedené výsledky příspěvku jsou využitelné ve dvou rovinách, které odpovídají jednotlivých cílovým skupinám. První rovina využitelnosti je tvořena zdravotnickými zařízeními, především vrcholovým vedením, vedoucími a pracovníky IS/ICT oddělení. Pro tyto profese by příspěvek měl být východiskem při rozhodovacím procesu o sourcingové strategii, především při její tvorbě, jelikož poskytuje ucelený pohled na faktory. Druhá rovina se týká vědecké komunity. Příspěvek doplňuje poznání pro oblast zdravotnictví a tvoří základ pro další výzkum, viz níže.

Definice faktorů ovlivňujících sourcingové strategie IS/ICT ve zdravotnictví vznikla na základě případové studie provedené do hloubky. Následující výzkum by se měl zaměřit na **analýzu jak silně jednotlivé faktory sourcingové strategie ovlivňují.** Pro další rozvoj by bylo vhodné se zaměřit na **analýzu vlivu schopnosti přijímat inovace a nové technologie na faktor trendy a tendence sourcingu IS/ICT**, který nebyl v této studii zahrnut. Model sourcingové strategie musí být v souladu s požadavky, které jsou vytvářeny faktory, a je tedy následně nutné **definovat požadavky na model sourcingové strategie vyplývající z faktorů**.

Literatura

BRUCKNER, T., VOŘÍŠEK, J. & BUCHALCEVOVÁ, A. 2012. *Tvorba informačních systémů; Principy, metodiky, architektury*. Praha: Grada Publishing, a.s. ISBN 978-80-247-4153-6.

DAVIS, H. L., EPPEN, G. D. & MATTSSON, L.-G. 1974. Critical factors in worldwide purchasing. *Harvard Business Review*. roč. 52, č. 6, s. 81–90. ISSN 0017-8012.

DIANA, M. L. 2009. Exploring information systems outsourcing in US hospital-based health care delivery systems. *HEALTH CARE MANAGEMENT SCIENCE* [online]. roč. 12, č. 4, s. 434–450. ISSN 1386-9620. Dostupné z: doi:10.1007/s10729-009-9100-4

EUL, M., RÖDER, H., CHUA, S. G., HAGEN, C., MEYER, K. & CIOBO, M. 2013. IT 2020: Preparing for the Future - Technology Featured Article - A.T. Kearney. *A.T. Karney* [online]. [vid. 12. March 2014]. Dostupné z: http://www.atkearney.com/strategic-it/featured-article/-/asset_publisher/BqWAk3NLsZIU/content/it-2020-preparing-for-the-future/10192

FEIGE, T. & POTANČOK, M. 2013. Enterprise Social Networks as a Tool for Effective Collaboration in Health Care Facilities. In: *International Conference on Research and Practical Issues of Enterprise Information Systems, Confenis 2013*. S.l.: Trauner Verlag Universität, s. 364. ISBN 978-3-99033-081-4.

HL7. 2010. HL7 Česká republika :: O HL7 Česká republika [online]. [vid. 5. April 2010]. Dostupné z: http://www.hl7.cz/cz/hl7/about.html

KARPECKI, L. 2010. ICT ve velkých nemocnicích. *CIO Business World*. č. 4. ISSN 1803-7321.

KOTLER, P. & KELLER, K. L. 2007. *Marketing management*. Praha: Grada Publishing, a.s. ISBN 978-80-247-1359-5.

LACITY, M. C., WILLCOCKS, L. P. & FEENY, D. F. 1996. The Value of Selective IT Sourcing. *Sloan Management Review*.

MARKETSANDMARKETS. 2013. Healthcare IT Outsourcing Market by Application, Life Science & Infrastructure – 2018 | MarketsandMarkets [online]. [vid. 12. March 2014]. Dostupné z: http://www.marketsandmarkets.com/Market-Reports/healthcare-it-outsourcing-market-1219.html

MECHLING, J. 2002. *Information Age Government for Ontario* [online]. S.l. Panel on the Role of Government. Dostupné z: http://128.100.167.100/investing/reports/rp6.pdf

MYERS, M. D. 2013. *Qualitative Research in Business & Management*. 2nd. London: Sage. ISBN 978-0-85702-973-7.

MYERS, M. D. & YOUNG, L. W. 1997. Hidden agendas, power and managerial assumptions in information systems development: an ethnographic study. *Information Technology & People*. roč. 10, č. 3, s. 224–240. ISSN 0959-3845.

PAULK, M. C., CURTIS, B., CHRISSIS, M. B. & WEBER, C. V. 1993. Capability maturity model, version 1.1. *Software, IEEE*. roč. 10, č. 4, s. 18–27. ISSN 0740-7459.

POTANČOK, M. 2010. *Analýza užití digitalizace na radiodiagnostickém pracovišti*. Praha. Vysoká škola ekonomická v Praze.

POTANČOK, M. 2012. *Informační systémy ve zdravotnických zařízeních*. Praha. Vysoká škola ekonomická v Praze.

POTANČOK, M. 2013. IS/ICT processes importance and customization in different categories of health care facilities. In: *Interdisciplinární mezinárodní vědecká konference doktorandů a odborných asistentů QUAERE 2013*. Hradec Králové, The Czech Republic: s.n., . ISBN 978-80-905243-7-8.

RIVERA, J. & MEULEN, R. van der. 2013. Gartner Says Worldwide IT Outsourcing Market to Reach $288 Billion in 2013. *STAMFORD, Conn.* [online]. [vid. 27. September 2013]. Dostupné z: http://www.gartner.com/newsroom/id/2550615

RIVERA, J. & MEULEN, R. van der. 2014. Gartner Says Worldwide IT Spending on Pace to Reach $3.8 Trillion in 2014. *STAMFORD, Conn.* [online]. [vid. 12. March 2014]. Dostupné z: http://www.gartner.com/newsroom/id/2643919

VOŘÍŠEK, J., BASL, J., BUCHALCEVOVÁ, A., GÁLA, L., KUNSTOVÁ, R., NOVOTNÝ, O., POUR, J. & ŠIMKOVÁ, E. 2008. *Principy a modely řízení podnikové informatiky*. Praha: Vysoká škola ekonomická v Praze, Nakladatelství Oeconomica. ISBN 978-80-245-1440-6.

VOŘÍŠEK, J. & POUR, J. 2012. *Management podnikové informatiky*. Praha: Professional Publishing. ISBN 9788074311024.

WILLCOCKS, L. & OSHRI, I. 2009. *To bundle or not to bundle? Effective decision-making for business and IT services* [online]. S.l. LSE/Accenture. Dostupné z: http://www.cpwerx.de/SiteCollectionDocuments/PDF/240Accenture_Bundled_Outsourcing_To_Bundle_or_Not_to_Bundle.pdf

YIN, R. K. 2009. *Case study research: Design and methods*. 4th. Thousand Oaks: Sage publications. ISBN 1412960991.

JEL Classification: M15

Summary

Analysis of the effects upon the IS/ICT sourcing strategy in healthcare facilities

This paper deals with sourcing strategies for information systems and information and communication technologies in healthcare facilities.

The aim of this paper is to analyse factors affecting IS/ICT sourcing strategies in healthcare. In relation to this aim the following research questions were posed: What factors affect sourcing strategies in healthcare facilities? How can factors affecting sourcing strategies in healthcare facilities be classified?

The research described in this paper is based on a case study, qualitative research. A two-step procedure was chosen for the analysis of factors affecting sourcing strategies in healthcare. Firstly, an analysis was performed on a general level and an initial set of factors was defined. Secondly, the initial set of factors was supplemented by healthcare-specific issues and used in the case study conducted at a chosen hospital. The final set of factors contains 18 factors and their classification. For each of them, identification methods and impacts have been specified. The case study has also shown that direct participants in the decision-making process do not pay proper attention to these factors.

The outcomes of the analysis can be used by healthcare facilities (particularly senior management and IS/ICT department staff) in sourcing strategy decision making process.

Key words: healthcare, information strategy, IS/ICT, IS/ICT management, sourcing strategy

Strategie sourcingu IT služeb

Michal Šebesta

michal.sebesta@vse.cz

Doktorand oboru Aplikovaná informatika

Školitel: prof. Ing. Jiří Voříšek, CSc., (vorisek@vse.cz)

Abstrakt: Článek se zabývá zkoumáním existujících oblastí outsourcingu, dostupných modelů sourcingu v literatuře, a jejich následnou analýzou a kategorizací. Prezentuje původní integrovaný pohled formou jednotné taxonomie strategií sourcingu IT služeb, která může být potenciálně využívána zákaznickými organizacemi. Byly identifikovány tři hlavní přístupy v současné době využívané v teorii a praxi řízení IT a ty byly dále podrobeny analýze. Problémem je vzájemná odlišnost těchto přístupů, s čímž souvisí obtížná porovnatelnost výzkumu zahrnujícího dané sourcingové modely. Nejednotnost pak činí výběr vhodné strategie vzhledem ke komplexnosti problematiky značně obtížnou. Samotný návrh nové univerzální taxonomie strategií sourcingu IT služeb se zaměřuje na identifikaci specifik jednotlivých existujících přístupů, jejich vymezení v současné praxi, a zdůraznění rizik spojených s jednotlivými přístupy. Zvláštní důraz je pak kladen na jejich provázání, identifikaci odlišností a nesouladů, a na vazbu na poslední trendy v poskytování IT.

Klíčová slova: Outsourcing, strategie, IT služby, řízení IT, kategorizace, organizace.

Úvod

Současná doba je z hlediska řízení podnikové informatiky specifická především masivním rozvojem infrastruktur, které umožnily nástup Cloud Computingu a souvisejícího konceptu Software-as-a-Service[29]. Tyto technologie či lépe řečeno způsoby poskytování umožnily využívání donedávna nedostupných technologií řadě společností, především pak ze segmentu malých a středních podniků. Rozšiřující se trh těchto IT služeb poskytovaných formou cloudu, rostoucí počet poskytovatelů, a zvyšující se komplexita podnikové informatiky, vyžadují jasně definované metody a inovované či zcela nové přístupy k rozhodování o případném využití těchto služeb organizacemi formou outsourcingu.

V ideálním případě by tato rozhodnutí měla vycházet z podnikové strategie. Nicméně existující IT strategie v mnoha podnicích nepokrývají jasně aspekt sourcingových strategií. Proto jsou tato rozhodnutí v řadě případů činěna bez konkrétního strategického pohledu či plánu. Jedním z důvodů, který k tomu přispívá, může být i neexistence jednotné taxonomie těchto strategií.

Mluvíme-li o sourcingových strategiích je vhodné zúžit námi zkoumanou oblast na pohled zákaznické organizace[30]. Především z toho důvodu, že se samotné podnikové strategie a cíle jak zákaznických tak i poskytovatelských organizací v praxi značně liší. Jak vyplývá už z jejich charakteristiky, jejich sourcingové strategie jsou založené na jiných principech a tedy samy o sobě odlišné.

V následujícím textu rozebírám nejdříve existující formy sourcingu, přičemž se zaměřuji na v současnosti aktuální outsourcing informačních technologií, někdy nazývaný IT outsourcing. V návaznosti na to analyzuji existující sourcingové modely a uvádím původní kategorizaci sourcingových strategií. Tato taxonomie je také hlavním přínosem tohoto příspěvku.

Formy sourcingu

Předtím, než se budu věnovat samotným sourcingovým strategiím, je vhodné zmínit také klasifikaci samotné oblasti sourcingu. Oproti sourcingovým strategiím, kde je vhodné používat obecnější pojem sourcing, je v případě forem sourcingu vzhledem k zvyklostem v teorii a praxi lepší využít pojem outsourcingu, který tvoří podskupinu sourcingu. Zde uvedené pojetí bylo formulováno na základě analýzy existujících oblastí v praxi i v odborné literatuře IT managementu. V tomto směru můžeme identifikovat následující oblasti aplikace sourcingu:

[29] Software poskytovaný formou služby, někdy je též uváděna zkratka SaaS.
[30] Označovaná jako customer organisation či end-user organisation.

- Outsourcing podnikových procesů.[31]
- Outsourcing informačních technologií.[32]
- Outsourcing vývoje softwaru.[33]

Tato klasifikace staví především na principu vzájemné konceptuální oddělenosti informačních technologií od podnikových procesů. K podpoře využití tohoto principu přispívá i fakt, že je využíván v rámci řady existujících přístupů k podnikovým architekturám[34].

Krom toho může být ilustrativním příkladem i model SPSPR, zmiňovaný například Voříškem a Jandošem (2010), který jasně definuje roli informačních technologií resp. IT služeb jako podpůrnou k daným podnikovým procesům.

Zatímco outsourcing podnikových procesů (BPO) lze z pohledu své charakteristiky považovat za oblast využívanou především velkými organizacemi, oblast outsourcingu informačních technologií (ITO) lze považovat za oblast využívanou napříč segmenty. Současný rozvoj cloudových technologií a souvisejícího Software as a Service konceptu navíc činí tuto oblast outsourcingu ještě více dostupnou zejména pro malé a střední podniky. Outsourcing vývoje softwaru (SDO) je spíše separátní kategorií, kterou lze potenciálně zkombinovat s oběma výše uvedenými oblastmi. Tato oblast SDO cílí na vývoj IT a nikoli na jejich provoz či zajišťování, je však vhodné ji zde zařadit z důvodu jejího vlivu na kalkulace a návazné rozhodování o interně poskytovaných IT službách.

Rozdělení oblastí outsouringu na BPO a ITO je klíčové pro korektní nástin možných strategií. Každá z oblastí má svůj specifický kontext a tedy i rozhodovací mechanismy včetně strategie se mohou mírně lišit.

Jak vyplývá z výše uvedeného, IT služby řadíme do oblasti outsouringu informačních technologií, a další text se tedy zabývá výhradně touto oblastí.

Outsourcing informačních technologií

Jak zmiňují Loh a Venkatraman (1992), IT outsourcing lze definovat jako významný příspěvek externích dodavatelů hmotných a/nebo lidských zdrojů spojený s dodáním celé nebo zvláštní části IT infrastruktury uživatelské organizaci. Jak je patrné z této definice, outsourcing informačních technologií nesestává pouze z IT služeb, ale může být sám o sobě kategorizován do několika forem. Například velice vhodné rozdělení uvádí v tomto směru Kern et al. (2002a), který navrhuje následující kategorizaci IT outsourcingu založenou na různých druzích využívaných zdrojů:

- *Insourcing* – využití interních zdrojů a jejich interního řízení.
- *Buy-in* – získání vlastnictví původně externích zdrojů pro interní řízení a provoz.
- *Tradiční outsourcing* – převod vlastnictví původně interních zdrojů na poskytovatele a následné řízení těchto zdrojů tímto poskytovatelem jménem zákazníka.
- *ASP* – pronájem zdrojů vlastněných poskytovatelem, dodávka služeb se uskutečňuje přes Internet.

Dané rozdělení by pak mohlo být z našeho pohledu doplněno o kategorii SaaS. Ačkoli je v hrubém pohledu navrhovaná kategorie velice podobná již uvedené ASP, je narozdíl od ní založená na multitenantnosti, což uvádějí jako jednu z charakteristik například Mell a Grance (2011). Tedy se v případě kategorie SaaS jedná o značně standardizované aplikace s omezenými možnostmi přizpůsobování[35] specifickým potřebám zákaznických organizací. Tato dodatečná kategorie by tedy mohla být definována jako:

- *Software-as-a-Service* – poskytovatel zajišťuje organizaci přístup k instanci standardizované multitenantní aplikaci využívající zdroje tohoto poskytovatele či jeho subdodavatelů, dodávka služeb se uskutečňuje přes Internet.

[31] Business Process Outsourcing (BPO).
[32] Information Technology Outsourcing (ITO), někdy označován též jako IT Outsourcing.
[33] Software Development Outsourcing (SDO).
[34] Podniková architektura ve smyslu Enterprise Architecture (EA).
[35] V angl. customisation.

Podíváme-li se na outsourcing informačních technologií z pohledu možných vzájemných vztahů poskytovatele a zákazníka, nalezneme řadu dalších taxonomií. V návaznosti na zkoumání těchto přístupů byly vytipovány některé kategorizace, které podrobíme další analýze. Dle mého názoru lze hodnotit jako velice relevantní i v současné době například přístup, který uvádí Da Rold a Berg (2003) z Gartner Research, klasifikaci uváděnou v ITILu (2007a), či taxonomii jež používají Currie a Willcocks (1998). Tyto přístupy se však vzájemně odlišují.

Da Rold a Berg (2003) rozlišují deset variant sourcingových modelů: multisourcing, brand service company, prime contractor, mix joint venture, outsourcing joint venture, best-of-breed consortium, client organization consortium, full outsourcing, insourcing, a internal delivery.

Na druhou stranu ITIL V3 Service Strategy (ITIL, 2007a) uvádí šest možných sourcingových modelů rozdělených do tří kategorií: multivendor sourcing (prime, consortium, a selective outsourcing), traditional sourcing (full service outsourcing), a internal sourcing (shared services, a internal services).

Currie a Willcocks (1998) pak narozdíl od předchozích identifikují celkem čtyři typy sourcingových rozhodnutí: total outsourcing, multiple-supplier sourcing, joint venture / strategic alliance sourcing, a insourcing.

Výběr jednotlivých sourcingových modelů závisí na širším kontextu rozhodování, a toto rozhodování má převážně strategický charakter. Vztahové modely sourcingu od Gartner Research (Da Rold & Berg, 2003), tzv. sourcingové struktury v ITIL (2007a), a čtyři typy outsourcingových rozhodnutí jak jsou uvedeny dle Currie a Willcocks (1998) se navzájem do určité míry překrývají. Proto lze pokládat za žádoucí sestatvit jednotnou klasifikaci, která by mohla sloužit jako kategorizace možných strategií outsourcingu IT služeb.

Strategie sourcingu IT služeb

V návaznosti na analýzu výše uvedených zdrojů si v rámci tohoto článku dovolím přednést vlastní kategorizaci strategií sourcingu IT služeb, která reflektuje výše uvedené přístupy. Integruje poznatky v nich uvedené a v návaznosti na ně tyto poznatky aplikuje na současnou situaci řízení IT s ohledem na výše diskutované trendy.

Navrhovaná kategorizace strategií sourcingu má následující strukturu:

- Singesourcing,
- Multisourcing,
- Consortium Sourcing,
- Prime Contractor Sourcing,
- Internal Services,
- Joint Venture.

Jednotlivým kategoriím se podrobněji věnuji v následujícím textu.

Singlesourcing

U této strategie organizace vytěsňuje (outsourcuje) určitou část či více částí systému pouze jednomu poskytovateli. Rozsah vytěsňovaných částí je ovlivněn především funkcionalitou nabízenou konkrétním poskytovatelem a jeho možnostmi, resp. výpočetní kapacitou. Je nutné si uvědomit, že s rostoucím počtem vytěsněných služeb[36] se také zvyšuje závislost organizace na tomto poskytovateli. V takovém případě může potenciálně nastat nepříjemná situace, kdy může být v budoucnu pro organizaci obtížné přejít k jiným poskytovatelům těchto aplikačních služeb. Naopak v případě, že je takto outsourcována nepatrná část funkcionality systému, je riziko závislosti na poskytovateli přijatelné. Výhodou strategie singlesourcingu tkví v tom, že není tak rozsáhlá provázanost mezi jednotlivými systémy jako u využití více poskytovatelů.

Currie a Willcocks (1998) také zmiňují možnost takzvaného celkového outsourcingu (total outsourcing), který definují jako outsourcing více než 70–80% IT služeb organizace externímu poskytovateli.

[36] A především také s rostoucím poměrem oproti interně zajišťovaným částem systému.

Outsourcing jednomu poskytovateli se též někdy označuje jako full outsourcing, tento název však lze pokládat jako příhodnější formu komplexního outsourcingu ve smyslu jak jej uvádí například Voříšek (2008). V souvislosti s výše uvedeným přístupem se pak dá komplexní outsourcing považovat za extrémní formu celkového outsourcingu.

V řadě případů singlesourcingu však podnik vytěsňuje jen určité IT služby, přičemž ostatní zajišťuje interně.

Průzkum, který zmiňuje Kern et al. (2002b) uvádí, že úspěchem skončilo v minulosti pouze 38% případů velkých kontraktů IT outsourcingu, zatímco u selektivního sourcingu (v našem případě tomuto termínu odpovídá multisourcing) byla úspěšnost více než 77%. I přesto může být tato singlesourcingová strategie vhodná pro podniky, které chtějí získat externí know-how formou IT služby v jedné specifické oblasti.

Multisourcing

Další možnou strategií je multisourcing, někdy také označovaný jako multivendor sourcing (ITIL, 2007b) selective outsourcing (Da Rold & Berg, 2003), či multiple-supplier sourcing (Currie & Willcocks, 1998). Oproti předchozímu typu organizace v tomto případě odebírá různé služby od více poskytovatelů. Realizace tohoto typu sourcingové strategie může přinést komplikace především se zmiňovaným narůstajícím počtem poskytovatelů v současné době. Řízení dodávek jednotlivých služeb od různých poskytovatelů a jejich případná integrace do funkčního celku tak hraje klíčovou roli v oblasti kapabilit dané organizace. Z hlediska zajištění funkcionality lze však tuto strategii považovat za jednu z nejvíce přínosných, jelikož neobsahuje omezení na jednoho poskytovatele[37] a umožňuje výběr menších služeb zaměřených na konkrétní specifické činnosti. Za předpokladu správného výběru poskytovatelů resp. služeb tak může organizace potenciálně dosáhnout nejlepšího pokrytí funkcionality vyžadované podnikovými procesy.

Na druhou stranu, vyšší počet poskytovatelů s sebou také přináší pravěpodobnost problémů spojených se vzájemnou komunikací, se sledováním dodržování podmínek ujednaných v rámci SLA, a s dalšími souvisejícími oblastmi. To mohou být například již zmíněné zvýšené nároky na integraci, náklady na tvorbu a udržování kontraktu, a tak podobně.

Zvláštní pozornost by pak měla být věnována due diligence, kterou by organizace měla provádět v rámci selekce poskytovatelů vhodných pro spolupráci. Ačkoli by toto zkoumání důvěryhodnosti a schopností poskytovatele mělo probíhat v rámci všech strategií zahrnujících externí poskytovatele, při vyšším počtu externích subjektů je důležité zavést systematický postup hodnocení včetně podrobné specifikace kritérií. Důvodem nutnosti existence takového postupu, který by usnadňoval zde kritickou selekci jednotlivých služeb, je fakt, že se se zvyšujícím počtem externích poskytovatelů stává tato činnost velice komplexní a nepřehlednou. Jedním z takových přístupů řešících tuto oblast může být SOURCER framework (Sebesta, 2010).

Jak uvádím v rámci tomto přístupu, ve většině případů existuje na trhu více poskytovatelů daného typu služby, a tedy je možné provést podrobné srovnání a kalkulace těchto variant. Z tohoto pohledu je multisourcingová strategie ideální pro použití v rámci trhu SaaS služeb. Poskytovatelé Software-as-a-Service zpravidla publikují pro potenciální i současné zákazníky SLA či pravidla kontraktu, které obsahují informace o parametrech jejich služeb[38]. Tento fakt umožňuje na základě definovaných požadavků selekci tzv. best of breed služeb v rámci tvorby celkového portfolia IT služeb organizace. Nezanedbatelnou výhodou je při vhodném výběru poskytovatelů také rozsáhlejší možnosti využití moderních technologií a přístupů v daném odvětví. Je nutné zmínit, že u této strategie není nutné pokrýt veškeré potřeby prostřednictvým externích poskytovatelů. Kombinace interně a externě zajišťovaných služeb je u této strategie možná a dokonce můžeme říci, že i prospěšná z důvodu existence služeb, které není vhodné vytěsňovat. Těmito službami rozumíme například služby zahrnující určitou unikátní znalost

[37] Což je typické u singlesourcingu.

[38] V řadě případů není uvedeno specificky SLA. Podmínky jsou ale specifikovány na webových stránkách poskytovatele v rámci zásad a podmínek použití služby, které uživatel musí elektronicky odsouhlasit v rámci registrace. Ačkoli jsou tyto podmínky kontraktů většinou pevně stanovené, v některých případech je možné pomocí renegociací získat určité slevy či dočasné zlepšení podmínek/úrovní služby.

či poskytují dané organizaci konkurenční výhodu. Konkurenční výhodu zde chápeme jako tzv. competitive advantage, jak uvádí například Barney (1991), Peteraf (1993), či Porter (2004).

Přestože tato strategie může dlouhodobě zvýšit celkové náklady na IT, využití best of breed služeb může organizaci vypomoci získat a využívat inovativní technologie ve více různých oblastech IT služeb bez nutnosti vysokých investic do infrastruktury, vývoje, a souvisejících oblastí.

Consortium Sourcing

Tato strategie kombinuje principy singlesourcingu a multisourcingu. Organizace v tomto případě využívá službu poskytovanou konsorciem poskytovatelů. Toto konsorcium je zpravidla založeno k účelu uspokojení poptávky po službě většího rozsahu, resp. po specifické kombinaci integrovaných služeb. Poskytovatelé prezentují svá řešení dohromady, přičemž musí být schopni své IT služby do určité míry vzájemně integrovat a garantovat vzájemné propojení v dohodnuté kvalitě.

Ve zde prezentovaném rozdělení této kategorii strategie nejvíce odpovídá pojetí ITIL (2007a), kdy v případě modelu konsorcia organizace sama provede předběžnou selekci pro ni zajímavých poskytovatelů. Tito potenciální členové konsorcia jsou pak zpravidla přizváni k jednání, kde společně prezentují své nabídky včetně jednotného rozhraní pro správu a řízení svých IT služeb.

Model outsourcingu, který uvádí Rold a Berg (2003) jako best of breed consortium z pohledu zde uvedené kategorizace do strategie consortium sourcingu neřadíme, jelikož je bližší strategii prime contractor, které se věnujeme dále.

Prime Contractor Sourcing

Další strategií sourcingu může být využití prime contractora, tedy hlavního poskytovatele. Podobně jako předchozí přístup, i tato strategie kombinuje některé principy singlesourcingu a multisourcingu. Narozdíl od konsorcia jí však lze považovat za víc orientovanou směrem k singlesourcingu. Důvodem tohoto rozdílu je, že narozdíl od konsorcia je strategie hlavního dodavatele založena na dohodě mezi zákaznickou organizací a pouze jedním poskytovatelem[39]. Tato dohoda má formu jednotného kontraktu na všechny části outsourcovaného systému dané skupině poskytovatelů, přičemž hlavní poskytovatel je zodpovědný za kontrakt i za související SLA na jednotlivé služby.

Hlavní poskytovatel pak komunikuje s ostatními poskytovateli ve skupině a zajišťuje, aby byly splněny sjednané podmínky, dodržovány povinnosti vyplývající z kontraktu, a aby byly služby poskytovány v souladu s ujednanými úrovněmi kvality dle SLA.

Většinou tento hlavní poskytovatel chrání své vlastní zájmy a výdělky plynoucí z kontraktu použitím podkladových kontraktů, které uzavírá s ostatními členy skupiny, a které mají za úkol mu zajistit dostatečnou kvalitu služeb poskytovaných těmito poskytovateli. Da Rold a Berg (2003) uvádí formu best-of-breed consortium, kde poskytovatelé tvoří dočasná uskupení, aby se společně účastnili konkrétních soutěží a výběrových řízení ve veřejném či soukromém sektoru[40]. V našem přístupu budeme tyto formy řadit do strategie prime contractor sourcingu.

V rámci strategie hlavního poskytovatele je nejdůležitější vytvořit funkční síť dohod, resp. kontraktů, které skupině umožní kontrakt zajistit v plném rozsahu a plné kvalitě. V případě volnější varianty vztahů mezi hlavním poskytovatelem a ostatními poskytovateli je kritické zajištění spolehlivé komunikace a dlouhodobých vztahů mezi jednotlivými subjekty, a to v rámci celé skupiny.

Internal Services

Samostatnou kategorií strategií jsou takzvané strategie interních služeb. Ty mohou mít několik podob, které probíráme v této části jednotlivě. Dle požadované úrovně nezávislosti poskytovatele na zákaznické organizaci můžeme identifikovat tři strategie interních služeb: Genuine Internal Services (ryze interní služby), Shared Services Center (sdílené centrum služeb), a Independent Service Organisation (nezávislá servisní organizace).

[39] V praxi se lze setkat s pojmy prime contractor, main contractor, nebo také main provider.
[40] Například vládní zakázky, či velké komerční zakázky.

Genuine Internal Services

Tato strategie je nejčistší formou strategie interních služeb. Je založena na využívání výhradně interních služeb, které jsou poskytovány vlastními pracovnímky či skupinami pracovníků v rámci dané organizační jednotky. Tyto jednotky nemusí být pouze jednotlivá oddělení, ale může se jednat také o regionální pobočky organizace či v některých případech i specifické skupiny v rámci projektů napříč odděleními.

Základní složkou tohoto konceptu je fakt, že organizace zajišťuje služby s využitím svých vlastních zdrojů. Velice často se toto týká i samotných organizačních jednotek – každá jednotka může mít svého vlastního poskytovatele uvnitř této jednotky. Opět, náhled, který je zde uplatňován je v tomto podobnější přístupu ITIL (2007a), než ostatním, které byly analyzovány.

Zdroje jako servery, software, infrastruktura, zaměstnanci, nebo databáze jsou obvykle dedikovány určité organizační jednotce a nikoli celé organizaci[41].

Shared Services Centre

Tento typ strategie interních služeb sestává ze založení separátní organizační jednotky, která posléze poskytuje ICT služby ostatním jednotkám v organizaci. ITIL (2007a) uvádí, že tato jednotka typicky řídí a spravuje své vlastní náklady a výnosy. Vyúčtovává pak poskytované služby ostatním jednotkám, jakoby se jednalo o klasického „externího" zákazníka. Z našeho pohledu by toto centrum sdílených služeb mohlo být definováno jako separátní nákladové středisko v rámci organizační struktury. Jak zmiňují Da Rold a Berg (2003), toto centrum může také poskytovat služby externím subjektům na jednotlivých trzích IT služeb. Tak může provozovateli přinášet dodatečný zisk. V tomto případě je však vždy vhodné zvážit sdílení know-how s externími subjekty s ohledem na celkovou podnikovou a IT strategii podniku.

Narozdíl od výše zmíněných konceptů (Da Rold & Berg, 2003; ITIL, 2007a) zde zařazujeme centrum sdílených služeb nikoli jako separátní kategorii, ale jako variantu strategie interních služeb. Důvodem je především obdobná vysoká důvěryhodnost tohoto poskytovatele, stejně jako zde definovaná charakteristika tohoto přístupu, která indikuje jeho náležitost do sféry vlivu organizace.

Použití strategie sdíleného centra služeb zjednodušuje implementaci forem buy-in a nebo insourcingu, stejně tak jako případný budoucí přechod k zajištění služeb externě (a tedy využití jiné strategie). To je možné zejména díky existenci již zavedených struktur pro interní zajišťování služeb. Dalším důvodem je pak transparentnější vykazování nákladů a výnosů ve srovnání s využitím strategie ryze interních služeb.

Independent Service Organisation

Tato strategie je založena na využití oddělené servisní organizace zcela vlastněné zákaznickou organizací. Jelikož se v tomto případě jedná o separátní entitu, může být tato strategie využita v kombinaci s některými dalšími strategiemi, jako například Prime Contractor, Joint Venture, či typicky Shared Service Centre. Možnost fungování na různých bázích ostatně zmiňují i Da Rold a Berg (2003).

Více než u předchozí strategie, nezávislá servisní organizace klade důraz na oddělení zákaznické organizace od poskytovatelské složky. V tomto případě je tímto poskytovatelem dceřiná či přidružená společnost poskytující dané IT služby. To také znamená, že tyto služby jsou poskytovány ve stejném režimu, jako by byly poskytovány na volném trhu. V případě komplexnějšího systému služeb, tato strategie může potenciálně ovlivnit zvýšení kvality IT služeb[42].

Joint Venture

Tato sourcingová strategie je založená na obchodní dohodě joint venture. Da Rold a Berg (2003) definují joint venture jako vytvoření nové byznys entity dvěma nebo více partnery. Tito partneři založí společně vlastněnou organizaci, která může později figurovat jako poskytovatel vybraných IT služeb. Míra

[41] Vytvoření separátní organizační jednotky zahrnuje až strategie Shared Services Centre.
[42] Ačkoli toto zvýšení je podmíněno volbou vhodné právní formy a geografické lokace této dceřinné/přidružené servisní organizace.

kontroly nad touto entitou poskytovatele je dána obchodním podílem partnera v této poskytovatelské organizaci. S mírou kontroly se na druhou stranu pojí také odpovědnost daného partnera z hlediska kvality poskytované služby.

Míra rizika při využití této strategie závisí zejména na rozdělení odpovědnosti mezi jednotlivé partnery v joint venture. Ve většině případů by tato strategie měla být využívána pouze k dočasnému překlenutí právních, kulturních, či procesních překážek pojících se k poskytování dané služby či skupiny služeb. Jak vyplývá z případových studií, které prezentují Currie a Willcocks (1998), klíčovou výhodou přístupu joint venture je potenciální redukce některých rizik jedno-poskytovatelský a více-poskytovatelských outsourcingových kontraktů. V dlouhodobém horizontu se pak joint venture většinou přemění v běžnou outsourcingovou dohodu s tradičními rolemi zákazníka a poskytovatele služby. To implikuje převod určitých aktiv, odpovědností, ale také pravomocí a kontrolních mechanismů na jednu ze zúčastněných stran, tj. na poskytovatele, zákazníka, či na organizační složku jednoho z nich. Souběžně s tím dochází ke změně strategie organizace na některou z uvedených výše.

Joint venture s účastí pouze dvou partnerů je pravděpodobně nejčastějším případem aplikace této strategie. Nicméně existují alternativní možnosti. Například pro malé a střední podniky může být atraktivní především varianta joint venture s účastí většího počtu partnerů. Skupina nezávislých podniků, která sdílí potřebu určité specifické služby, v tomto případě založí joint venture. Toto uskupení pak poskytuje tyto služby zakládajícím členům a může také poskytovat služby na volném trhu IT služeb. Konkrétně tato varianta joint venture strategie umožňuje získat značné úspory z rozsahu ve srovnání s interními formami zajišťování služeb v jednotlivých organizacích odděleně[43].

Závěr

Rostoucí počet dostupných IT služeb a poskytovatelů na trhu tvoří převážně neprozkoumanou oblast vyžadující nové přístupy k rozhodování o případném outsourcingu těchto služeb organizacemi. Tato rozhodnutí by měla v ideálním případě vycházet z podnikové strategie. Nicméně, v mnoha případech existující IT strategie nepokrývají jasně aspekt outsourcingové strategie. Proto jsou tato rozhodnutí mnohdy činěna bez konkrétního strategického pohledu či plánu. Pro to, aby organizace mohly plně využít potenciálu Cloud Computingu a souvisejícího konceptu Software-as-a-Service, je nutné, aby v návaznosti na existující strategie měly zvolenou jasnou sourcingovou strategii pro IT služby. Na její realizaci by se měly zaměřit v rámci taktického a operativního řízení IT. Příkladem může být výběr poskytovatelů služby, výběr a zavádění konkrétního řešení, volba vhodných forem IT sourcingu.

Aby organizace mohly tuto strategii správně zvolit, chybí v současné době konsensus na taxonomii těchto sourcingových strategií. Řada existujících přístupů pracuje s rozdílným názvoslovím a zahrnuje různé varianty strategií. Tím vzniká potenciální problém nekonzistence výzkumu v rámci organizací odkazujících na různé přístupy. To je markantní především v případě oblasti IT služeb, kde se očekává v následujících nárůst podílu outsourcingu. Zmíněné problémy může ulehčit vhodná definice možných sourcingových strategií, na které by se organizace mohly odkazovat, a které by jim ulehčily volbu určité strategie v jejich konkrétních podmínkách.

V tomto článku jsem se věnoval shrnutí existující oblasti outsourcingu, a dále jsem analyzoval existující sourcingové modely. V návaznosti na to byla navržena univerzální taxonomie sourcingových strategií IT služeb.

Samotné sourcingové strategie IT služeb cílí především na oblast outsourcingu informačních technologií (IT outsourcingu). Zde je nutné chápat rozdíl oproti ostatním oblastem, jež jsou outsourcing

[43] Otázkou však je, jak je tento model kooperace uplatnitelný na vysoce konkurenčním typu trhů. Za hlavní problém lze považovat to, že využívání služeb poskytovaných v rámci této varianty strategie s sebou nese také sdílení citlivých dat spojených s IT službou mezi jednotlivými partnery v celém joint venture. Například data o typech využívaných služeb, o jejich nastavení, o dalších souvisejících vlastnostech, a o odebíraném množství služeb včetně možných sezónních výkyvů, mohou po podrobné analýze sloužit jako informační základ pro konkurenci v rámci dané joint venture. Proto je před zvažováním této strategie vhodné provést podrobnou analýzu citlivých dat v organizaci, studii trhu, a nástin možných scénářů a dopadů v případě jejich úniku ke konkurenci. Je tedy možné, že v některých případech nebude využití strategie joint venture žádoucí.

podnikových procesů a outsourcing vývoje softwaru. Jak bylo uvedeno, do této oblasti spadají též poslední trendy Cloud Computingu, resp. Software-as-a-Service, ke kterým bylo v rámci analýzy strategií přihlíženo.

Outsourcing informačních technologií byl dále diskutován ze dvou pohledů – z hlediska různých druhů využívaných zdrojů, a z hlediska modelů sourcingu. V pohledu využívaných zdrojů byl existující pohled v literatuře (insourcing, buy-in, tradiční outsourcing, a ASP) obohacen právě o novou kategorii Software-as-a-Service, přičemž byly zdůrazněny odlišnosti od konceptu ASP.

Hlavním přínosem práce je analýza přístupů k modelům sourcingu a v návaznosti na to prezentace původní kategorizace sourcingových strategií IT služeb. V rámci této taxonomie bylo představeno šest základních strategií: singesourcing, multisourcing, consortium sourcing, prime contractor sourcing, internal services, a joint venture. U těchto strategií byly definovány základní charakteristiky, byly identifikovány podstatné výhody či nevýhody, a v případě odchylek od existujících přístupů či existence nesouladu pak byla zdůrazněna vazba na tyto přístupy.

V budoucnu tato taxonomie sourcingových strategií ulehčí mapování jednotlivých přístupů k sourcingu IT služeb, a umožní podnikům lépe definovat očekávání a rizika spojená s jednotlivými strategiemi.

Z hlediska budoucího výzkumu je plánováno zkoumání reálného využití těchto strategií v podnicích, přičemž je v plánu se zaměřit především na průzkum úspěšnosti těchto strategií v závislosti na odvětví a velikosti daného podniku.

Literatura

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), 99–120.

Currie, W. L., & Willcocks, L. P. (1998). Analysing four types of IT sourcing decisions in the context of scale, client/supplier interdependency and risk mitigation. *Information Systems Journal*, 8(2), 119–143.

Da Rold, C., & Berg, T. (2003). Sourcing Strategies: Relationship Models and Case Studies. Research Report. Gartner.com. ID:R-18-9925. Retrieved from http://www.gartner.com/id=385260.

ITIL (2007a). ITIL V3 – Service Strategy. Norfolk, England : Office of Government Commerce.

Kern, T., Willcocks, L. P., & Lacity, M. C. (2002a). Application Service Provision: Risk Assessment and Mitigation. MIS Quarterly Executive, 2(1).

Kern, T., Willcocks, L. P., & Lacity, M. C. (2002b). Netsourcing: Renting Business Applications and Services Over a Network (1st ed.). Pearson Education.

Loh, L., & Venkatraman, N. (1992). Diffusion of Information Technology Outsourcing: Influence Sources and the Kodak Effect. *Information Systems Research*, 3(4), 334–358.

Mell, P. & Grance, T. (2011). The NIST Definition of Cloud Computing. Gaithersburg, MD : National Institute of Standards and Technology. NIST Special Publication 800-145.

Porter, M. E. (2004). *Competitive Advantage*. New York; London: Free Press.

Peteraf, M. A. (1993). The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal*, 14(3), 179–191.

Sebesta, M. (2010a). Towards a framework for effective outsourcing practice within the new application service provisioning trends. *Proceedings of the 3rd IEEE International Conference on Computer Science and Information Technology* . Chengdu, China : IEEE Press.

Vorisek, J., et al. (2008). Principy a modely rizeni podnikove informatiky. Prague, Czech Republic: Oeconomica.

Vorisek, J., & Jandos, J. (2010). ICT Service Architecture - Tool for Better Enterprise Computing Management. *IDIMT-2010 Information Technology - Human Values, Innovation and Economy*. Universitatsverlag Rudolf Trauner, Linz. 2010.

JEL Classification: M15, M10.

Summary

Explaining IT Service Sourcing Strategies

The paper examines existing outsourcing areas and available sourcing models in the literature. Subsequently it analyzes these approaches, and focuses on their categorization. The main contribution is then a proposal for universal IT service sourcing strategies taxonomy that can be utilized within customer organizations. Three major approaches mentioned in academic outlets and frequently used in the current business practice were identified and analyzed. Main problem with these approaches is their diversity. This in the end increases difficulties associated with research of sourcing models, especially with the use comparative studies of potential business cases. Given the complexity of the whole information systems context, the differences in these approaches then make selection of appropriate sourcing strategy much more difficult than it could be in ideal situation. The proposed universal taxonomy focuses on identification of existing approaches and their positioning within contemporary business practice. It also highlights risks associated with these approaches. Emphasis is then placed on the interconnections between these approaches, on identification of differences, and on reflection of latest trends in IT outsourcing.

Key words: Outsourcing, strategy, IT services, IT management, taxonomy, organization.

Datamining klasifikačních business rules prostřednictvím systému EasyMiner

Stanislav Vojíř

stanislav.vojir@vse.cz

Doktorand oboru Aplikovaná informatika

Školitel: doc. RNDr. Petr Strossa, CSc. (kizips@vse.cz)

Abstrakt: Business rules jsou efektivním konceptem pro zaznamenávání dílčích podmínek pro rozhodování, která ovlivňují chod podnikových procesů. Jednotlivá pravidla jsou do systémů zadávána obvykle ručně, prostřednictvím doménových expertů. V současné době jsou však zkoumány možnosti jejich (polo)automatického získávání prostřednictvím dataminingových metod. Cílem tohoto příspěvku je shrnout základy myšlenkového konceptu business rules, jeho praktického nasazení a představit možnost získávání sad business rules prostřednictvím dolování asociačních pravidel.

Po základní charakteristice specifik business rules jakožto myšlenkového konceptu jsou v příspěvku zhodnoceny možnosti dostupných specifikací pro záznam business rules z hlediska jejich využitelnosti, včetně podpory v podobě softwarových komponent pro integraci s podnikovými informačními systémy a se systémy pro podporu rozhodování. Následně je popsána možnost generování klasifikačních business rules z asociačních pravidel, popsán proces tvorby odpovídající znalostní báze a uvedeny možnosti realizace samotné klasifikace prostřednictvím různých typů pravidel – včetně modelového příkladu kombinace pravidel z většího množství zdrojů pro postupné řešení klasifikační úlohy.

Navržené přístupy byly experimentálně implementovány v rámci data miningového systému EasyMiner. Business rules získaná transformací z asociačních pravidel jsou vyhodnocována systémem JBoss Drools.

Klíčová slova: data mining, asociační pravidla, business rules, EasyMiner, klasifikace, Drools

Úvod

V současné době lze registrovat stále rostoucí poptávku po vhodných nástrojích a datových modelech pro podporu rozhodování a řízení podnikových procesů, stejně tak jako poptávku po získávání zajímavých informací z historických dat. Komerční i nekomerční organizace shromažďují řadu informací o svých klientech, potažmo zákaznících. Jedním z cílů vytváření komunit prostřednictvím „klientských karet“, či prostým sledováním činnosti uživatelů internetových obchodů je zjišťování informací o jejich chování – například za účelem zlepšení marketingových aktivit.

Jedním z vhodných způsobů analýzy dat je dolování asociačních pravidel. Tato pravidla je možné využívat nejen v podobě popisného modelu. Lze je též transformovat do podoby klasifikačních business rules, která mohou být aktivně využívána pro řešení klasifikačních úloh a lepší řízení podnikových procesů – například lepší zacílení posílání reklamy, či detekci odchylek v datech.

Cílem tohoto textu je představit možnost dolování klasifikačních business rules za využití systému EasyMiner a shrnout možnosti jejich praktického využití. V jednotlivých kapitolách je nejprve upřesněna problematika oblasti business rules a jejich praktického využití, představena možnost dolování asociačních pravidel a jejich následné transformace do podoby business rules prostřednictvím systému EasyMiner a popsány možnosti využití vytvořených sad business rules pro řešení klasifikačních úloh.

(Polo)automatické získávání sad business rules je jedním z aktuálních témat výzkumu. Z existujících softwarových nástrojů lze jmenovat komponentu *Rule Learner*, která je součástí systému *Open Rules* (OpenRules, Inc.), snahy o automatické získávání pravidel lze však zaregistrovat také u některých dalších nástrojů. Oproti stávajícím návrhům je v přístupu implementovaném v aplikaci LISp-Miner využíváno výsledku dataminingu asociačních pravidel, přičemž lze volit mezi pravidly získanými aplikací R, či aplikací LISp-Miner. Byly též testována výhodnost použití „bohatých“ asociačních pravidel získaných metodou GUHA.

Business rules a jejich použití v praxi

Business rules jsou jedním z možných přístupů k modelování podniku a popisu rozhodování. Nejedná se o konkrétní technologii či metodiku, ale spíše o způsob modelování, na kterém je postavena řada dílčích přístupů. Termín „business rules“ je z hlediska překladu do češtiny poněkud problematický, neboť není k dispozici větší množství odborných materiálů v českém jazyce a překlad „obchodní pravidla“ je v češtině využíván spíše ve svém širším kontextu – pro označení pravidel pro realizaci obchodů atp. V české literatuře jsou také využívány další částečně počeštěné varianty, jako například „business pravidla“ či „byznys pravidla“. Pro zabránění nejasnostem bude v tomto textu pro označení business rules jakožto přístupu k modelování a skupině technologií využíván původní anglický termín.

Principy tvorby business rules

Základním principem business rules přístupu k modelování je: „*Rules build on facts, and facts build on terms.*“ (Ross, 2003 str. 165) Tento princip vystihuje celou základní podstatu a postupnou tvorbu sad pravidel. Základním stavebním kamenem jsou termíny (terms). Termínem by měla být vždy označena konkrétní entita reálného světa (například věc, stav či funkce). Na základě termínů a jejich vztahů jsou následně definována fakta. Tyto definice mohou mít podobu jednoduchých vět, které popisují vztah mezi dvěma termíny – například: „Zákazník zadává objednávku.“ Fakta tedy neoznačují konkrétní entitu, ale jednoduchou základní znalost. Termíny (a v některých přístupech také fakta) je vhodné definovat v rámci terminologického slovníku, jenž je následně využíván pro definici samotných pravidel. Je nutné, aby byl vytvářený slovník sdílen všemi pravidly, která patří do dané báze (optimálně všemi pravidly daného podniku vůbec), neboť jen tak lze zajistit jejich konzistenci a koherentnost. Pro tvorbu business rules je dále vhodné dodržovat celou řadu dalších principů. Pravidla by měla být vyjádřena explicitně, v prostém jazyce, měla by být řízena a měla by být motivována identifikovatelnými a zároveň důležitými faktory. V zásadě se tedy jedná o jednoduchá pravidla, která jsou shromažďována do sad („rule sets“).

Z hlediska logiky jsou business rules vyjádřeními predikátového kalkulu ve tvaru „antecedent → konsekvent“. Predikáty jsou z pohledu business rules označovány jako fakta. Pravidla mohou obsahovat logické spojky (AND, OR, NOT) a kvantifikátory (existenční i univerzální). V rámci business rules přístupu mohou jednotlivé logické formule obsahovat nejen proměnné (obecné specifikace termínů), ale také konkrétní hodnoty. (Ross, 2003) Právě konkrétní hodnoty a jejich případná následná abstrakce je často vhodnou cestou tvorby business rules modelu na základě informací od podnikových pracovníků.

V reprezentaci řady standardů jsou samotná pravidla zaznamenávána v IF-THEN tvaru. Antecedent daného pravidla je označován jako podmínka, konsekvent je jeho tělem. Tělo pravidel v IF-THEN tvaru může obsahovat nejen konkrétní logickou formuli, ale v řadě přístupů také programový kód, který má být vykonán.

Využívání business rules v praxi

Business rules lze rozdělit do typů dle jejich účelu. Lze tedy definovat pravidla deklarativní (definující další fakta), represivní (zajišťující konzistenci odvozovací báze), odvozovací a také pravidla akční/reakční (vyvolávající konkrétní další aktivitu).

Nejde o jednu konkrétní, obecně využívanou technologii, ale o skupinu specifikací, technologií a nástrojů, které je vhodné vzájemně kombinovat a využít je k posunu analytických i programátorských přístupů k řešení softwarové podpory podnikových procesů směrem k oddělení aplikační a business logiky. Business rules jsou díky své jednoduchosti srozumitelná nejen softwarovým odborníkům, ale také například managementu organizací. Ve většině nástrojů je zároveň možné sadu business rules oddělit od samotného výkonného jádra aplikací a využívat jejich úprav například pro změnu kritérií rozhodování. Pro srozumitelnost uveďme jeden příklad: Mějme zdravotní pojišťovnu, která je povinna kontrolovat data získávaná od lékařů za účelem provádění revizních kontrol v případě abnormálně vysokých nároků na úhrady či příliš často vykazovaným provedením obdobných vyšetření. Prostřednictvím specifikace rozhodovacích (klasifikačních) business rules lze testovat všechna příchozí data a upozorňovat na sledované abnormality. Podmínkami jednotlivých business rules jsou sledované

podmínky (například „nárůst vykazovaných úkonů u pacientů jednoho lékaře o X procent“). V případě naplnění podmínky je spuštěn obchodní proces zajišťující provedení kontroly (například přidáním do seznamu úkolů pro revizního pracovníka).

Sady business rules jsou do systémů obvykle zadávány ručně. Ačkoliv pro zadávání pravidel může být k dispozici jednoduchý a srozumitelný softwarový nástroj, přesto tento postup vyžaduje velmi zkušené pracovníky, kteří zadávají do systému své vlastní znalosti. Rizikem tohoto zadávání pravidel je také velká pravděpodobnost neúplnosti vytvořené sady pravidel z důvodu opomenutí některých kritérií. Z těchto důvodů jsou v posledních letech zkoumány možnost poloautomatické či automatické tvorby sad business rules. Lze nalézt snahy o tvorbu business rules na základě analýz nestrukturovaných textů zpracovávaných v rámci organizace, zapojení dataminingových modelů do rozhodování atp. (OpenRules, Inc.) Jedním z vhodných přístupů je také využití dataminingu asociačních pravidel – viz následující kapitola.

Systémy pro práci s business rules a využívané specifikace

Jelikož nejsou business rules jednotnou technologií, ale spíše myšlenkovým konceptem, na kterém je postavena řada specifikací a nástrojů, je pro jejich praktické využívání nutné vybrat nejvhodnější kombinaci formátů a systémů pro zpracování sad pravidel.

Business rules je možné zaznamenávat v celé řadě formátů, které se liší nejen svým zaměřením a podobou záznamu, ale také svými vyjadřovacími schopnostmi. Z hlediska typu záznamu lze jednotlivé formáty rozdělit na formáty postavené na zjednodušení a formalizaci přirozeného jazyka (které jsou vhodné hlavně na sdílení informací v podobě datových modelů), grafické (například záznamy v podobě UML diagramů) a záznamy v podobě pseudokódu či XML struktury. Mezi nejznámější a nejrozšířenější specifikace business rules lze zařadit *SBVR* (specifikace zahrnující několik typů záznamů, nejčastěji užívanou variantou jsou textové záznamy pomocí jednoduchých pravidel formulovaných ve větách – *Structured English* či *RuleSpeak*), *RuleML* (formát navržený jako cross-industry standart pro záznam nejen business rules, ale také dalších logických pravidel, umožňuje provázání s dalšími XML specifikacemi), *Drools DRL* a *Jess Rules* (formáty svázané s konkrétními odvozovacími enginy, vhodné pro praktické využívání pravidel v komponentnách informačních systémů) či *Jena Rules* (pravidla zachycená v podobě tripletů, vhodný formát pro kombinaci s ontologiemi). Srovnání jednotlivých variant je k dispozici v tabulce Tabulka 3. Kromě těchto obecně definovaných formátů je nutné uvést ještě formát business rules definovaný Microsoftem, který je používán v rámci *BizTalk serveru*.

Tabulka 3: Srovnání jazyků pro záznam business rules

	Zaměření	**Typy pravidel**
SBVR	business komunikace, popis	všechny typy pravidel
RuleML	technická komunikace, popis, odvozování	všechny typy pravidel
R2ML	technická komunikace, popis, odvozování	deklarativní, restrikční
SWRL	popis, odvozování, rozšíření ontologií	všechny typy pravidel
RIF	technická komunikace, popis, odvozování, rozšíření ontologií	všechny typy pravidel
DRL	popis, odvozování	restrikční, akční
Jess	popis, odvozování	všechny typy pravidel
Jena Rules	popis, odvozování	deklarativní, restrikční
JRules	popis, odvozování	restrikční, akční
Open Rules	popis rozhodování, odvozování	rozhodovací

	Formát pravidel	**Srozumitelnost pro management**
SBVR	textový, grafický, XML	velká
RuleML	XML	střední
R2ML	XML	malá

	Formát pravidel	**Srozumitelnost pro management**
SWRL	textový, strukturovaný, XML, RDF	velká
RIF	strukturovaný, XML	malá
DRL	IF-THEN	střední
Jess	strukturovaný, IF-THEN	malá
Jena Rules	strukturovaný, RDF	střední
JRules	IF-THEN, textový	střední
Open Rules	tabulky	velká

	Srozumitelnost pro implementátory	**Software pro správu pravidel**	**Technologie**
SBVR	střední	ne	-
RuleML	střední	ano	Java
R2ML	malá	ne	Java
SWRL	střední	ano	Java, C++
RIF	střední	ne	Java, Python
DRL	velká	ano	Java, .NET
Jess	velká	ne	Java
Jena Rules	střední	ne	Java
JRules	velká	ano	Java, .NET
Open Rules	velká	ano	XLS, Java

Pro praktické využívání jsou však nezbytné nejen vhodné specifikace formátů pravidel, ale také konkrétní aplikace, které umí s danými formáty pracovat. Pro praktické nasazení je nutné mít k dispozici dostatečně výkonný *business rules engine* (odvozovací komponenta, která v praxi zpřístupňuje sadu pravidel jako celek a umožňuje její vyhodnocování) a *business rules management system* (obvykle označovaný zkratkou BRMS, jedná se o komponentu umožňující správu sad pravidel) a podpůrné nástroje – editor pravidel, automatizaci umožňující načasování vyhodnocování atp.

V současné době jsou business rules využívána nejčastěji jako komponenty expertních a doporučovacích systémů, či v rámci systémové integrace na úrovni *Enterprise Service Bus*. (Chappell, 2004)

Pro výběr konkrétní specifikace formátu business rules a odpovídajících softwarových komponent je vhodné vycházet z definice cíle využití business rules, zhodnotit možnost oddělení obchodní a aplikační logiky při implementaci informačního systému a též zhodnotit vazby na již užívané softwarové komponenty. Na základě testování lze říci, že v zásadě všechny nejrozšířenější komponenty jsou kvalitativně na podobné úrovni – výběr by tedy měl být ovlivněn také „platformou“, kterou využívá zbytek informačního systému. Většina business rules nástrojů je implementována v jazyce Java či .NET.

Pro další práci byl na základě ohodnocení variant pro zapojení do systému EasyMiner zvolen systém JBoss Drools, respektive jeho odvozovací engine *Drools Expert*. Důvodem pro tuto volbu byly široké možnosti obsahu jednotlivých business rules, otevřenost celého systému pod open-source licencí a také možnost kombinovat daný odvozovací engine s vlastním systémem pro správu a editaci pravidel.

Získávání business rules z asociačních pravidel

Jak již bylo zmíněno v předchozí části textu, jedním z vhodných způsobů získávání sad business rules je datamining asociačních pravidel a jejich následná transformace. Byly zkoumány možnosti využití nejen jednoduchých asociačních pravidel získávaných algoritmem APRIORI (tj. pravidla často používaná pro analýzy „nákupních košíků“), ale také komplexní pravidla získávaná metodou GUHA.

GUHA asociační pravidla

Metoda GUHA je původní česká metoda na hledání zajímavých vztahů (hypotéz) v datech, jejíž principy byly formulovány již v roce 1966. Její vývoj však pokračuje dodnes. Byly zkoumány možnosti tranfromace asociačních pravidel získávané GUHA procedurou ASSOC, implementovanou v systému LISp-Miner (respektive jeho součástí 4ft-Miner). (Šimůnek, 2010)

Obecně lze 4ft-asociační pravidlo definovat ve tvaru

$$\varphi \approx \psi$$

kde φ, ψ jsou logické kombinace atributů a jejich vztah je definován pomocí zobecněného kvantifikátoru ($\approx$). Antecedent (φ) pravidla i jeho konsekvent (ψ) mohou obsahovat nejen konjunkci jednotlivých atributů, podporovány jsou také logické spojky disjunkce a negace a vyjádření preference prostřednictvím závorek. Při hledání pravidel prostřednictvím systému LISp-Miner mohou být pro jednotlivé atributy v rámci pravidla definovány nejen konkrétní hodnoty, ale také spojení více hodnot z původního datasetu (množina, interval). Kvantifikátor, který vyjadřuje vztah mezi antecedentem a konsekventem, je booleovská funkce definovaná nad frekvencemi čtyřpolní tabulky:

Tabulka 4: Čtyřpolní tabulka **Zdroj: (Šimůnek, 2010)**

	ψ	$\neg\psi$	$\sum$
φ	a	b	r
$\neg\varphi$	c	d	s
$\sum$	k	l	n

Kvantifikátorů ($\approx$) existuje celá řada, mezi nejznámější z nich patří *fundovaná implikace* a *above average dependance*. Fundovaná implikace je kombinací běžně používaných měr zajímavosti konfidence *(confidence)* a podpora *(support)*, above average dependance je převoditelná na míru zajímavosti *lift* a konfidenci. Fundovanou implikaci lze vyjádřit vzorcem:

$$\frac{a}{a+b} \geq p \ \wedge \ a \geq BASE$$

Systém LISp-Miner je využíván nejen v podobě klasické desktopové aplikace pro platformu Microsoft Windows. Jeho části jsou k dispozici také prostřednictvím nadstavbové webové služby *LM Connect*, které je využitelná buď samostatně, či v rámci webového systému EasyMiner.

Systém EasyMiner

Systém EasyMiner[44] je výzkumným projektem Katedry informačního a znalostního inženýrství, jehož cílem je zpřístupnění dolování asociačních pravidel v prostředí běžného internetového prohlížeče, bez nutnosti instalace jakýchkoliv specializovaných aplikací či doplňků. EasyMiner je koncipován jako komplexní systém podporující všechny kroky procesu dolování z dat dle metodologie CRISP-DM. Uživatelé nahrají do systém vlastní data ve formátu CSV a definuje jejich předzpracování – buď zadáním definice vlastní, nebo znovupoužitím již existující definice předzpracování. Samotný proces dolování asociačních pravidel je zprostředkován v jednoduchém webovém rozhraní, ve kterém uživatel definuje vzor hledaných pravidel a míry zajímavosti (obvykle konfidenci, podporu či lift). Výsledky se zobrazují opět ve webovém rozhraní, kde má uživatel možnost s nimi zároveň dále pracovat – vybrat zajímavá pravidla, využít je k tvorbě analytických zpráv či také k tvorbě klasifikačního modelu v podobě sady business rules. (Vojíř, 2013)

Systém EasyMiner v současné době podporuje již dvě volitelné alternativy dolovacích komponent. První z nich, která umožňuje nalézat asociační pravidla prostřednictvím metody GUHA, je webová služba *LM Connect* zprostředkovávající funkcionalitu aplikace *LISp-Miner*.[45] Tato varianta umožňuje hledat komplexní asociační pravidla, která obsahují v antecedentu či konsekventu nejen sadu atributů s konkrétními hodnotami, ale také složitější výrazy – konjunkce, disjunkce, negace, vyjádření preferencí prostřednictvím závorek či dynamické seskupení hodnot atributu. Jelikož však nad některými typy dat

[44] http://www.easyminer.eu
[45] http://lispminer.vse.cz

vyžaduje dolování prostřednictvím aplikace LISp-Miner (a tím i nad ní postavenými komponentami) velké množství času, byla vytvořena druhá dolovací komponenta, která využívá balíčku *arules* statistického systému R. Pravidla získávaná tímto způsobem nedosahují komplexnosti GUHA asociačních pravidel, ale jejich dolování je v případě velkých datasetů časově úspornější. Mezi oběma komponentami se navíc mohou uživatelé přepínat – lze tedy využít pro část dataminingové analýzy dolování prostřednictvím komponenty *EasyMiner arules* a pro další část analýzu prostřednictvím komponenty *LM Connect*.

Kromě dolování asociačních pravidel byla v systému EasyMiner implementována také podpora pro tvorbu sad business rules použitelných pro řešení klasifikačních úloh. Uživatel má možnost exportovat vybraná asociační pravidla do podoby business rules a následně je kombinovat s pravidly zadanými ručně prostřednictvím *BusinessRules Editoru*. Kompletní sadu pravidel je možné přímo ve webovém prostředí otestovat z hlediska schopnosti klasifikace dat a následně exportovat pravidla ve formátu DRL.[46]

Slovník pro transformaci asociačních pravidel do podoby business rules

Pokud mají být nalézaná asociační pravidla transformovaná do podoby business rules, která budou nadále obecně použitelná, pak prvním, nezbytným krokem celého procesu transformace definice terminologického slovníku. Pokud je uvažována práce pouze s jedním konkrétním datovým zdrojem (například jedna datová matice s informacemi o zákaznících podniku) a vytvořená business rules budou využívána k odvozování nad příchozími údaji se stejnými vlastnostmi, pak lze jako terminologický slovník použít obecný popis dané datové matice (databázové tabulky). V průběhu dolování asociačních pravidel jsou však data obvykle rozličnými způsoby předzpracována. Z jednotlivých datových sloupců vznikají atributy, které jsou využívány k popisu nalezených asociačních pravidel. Jejich názvy a hodnoty jsou však poněkud odlišné od původních dat. Při předzpracování dochází například ke slučování některých hodnot do množin či intervalů.

Bude-li uvažována situace, kdy mají být vytvořená pravidla využívána k automatické klasifikaci hodnocených objektů (například doporučení vhodného zboží pro konkrétního zákazníka), je nutné v rámci transformace asociačních pravidel do podoby business rules provést opačné transformace, než které byly provedeny při předzpracování – tak, aby definice konkrétního business rule odpovídala původním datům. Pro lepší srozumitelnost uveďme konkrétní příklad. V datové matici existuje sloupec „*věk*“ s celočíselnými hodnotami od 10 do 80. V průběhu předzpracování byl definován atribut „*věkové kategorie*“ s kategoriemi „*děti*“ s hodnotami *<0;15)*, „*mladiství“* s hodnotami v intervalu *<15;18)* a dospělí *<18;∞)*. Pokud je v nalezeném pravidle uveden atribut „*věkové kategorie (mladiství)“*, je nutné jej převést na dílčí podmínky ve tvaru „*(věk >= 15) AND věk < (věk < 18)*“.

Tato situace může být zároveň ještě více komplikovaná v případě využití GUHA asociačních pravidel získaných prostřednictvím systému LISp-Miner, který podporuje dynamické seskupování hodnot v průběhu dolování pravidel. Jednotlivé vícehodnoty (například „*věk(děti, mladiství)*“) je nutné rozložit na dílčí podmínky spojené disjunkcí. Takto rozepsaná pravidla jsou dobře zpracovatelná odvozovacími enginy pro zpracování business rules, ale již nejsou tak dobře srozumitelná pro uživatele.

Z tohoto důvodu jsou pro práci v systému business rules pravidla uživatelům prezentována v podobě, která je podobná zápisu asociačních pravidel v kombinaci s prvky ze specifikace Structured English formátu SBVR, pro praktické odvozování je však využito formátu DRL, který využívá rozepsané podmínky uvedené v předchozích odstavcích. Pro propojení těchto dvou přístupů byl navržena a implementována znalostní báze pro sdílení báze znalostí o předzpracování dat, kdy jsou v systému uchovávány informace o jednotlivých atributech a jejich vazbách na původní data. Tyto vazby je možné uchovávat v podobě instancí ontologických tříd (či v podobě sady provázaných tabulek v relační databázi). Koncept této znalostní báze byl podrobněji představen v (Vojíř, 2014).

[46] DRL – formát pravidel využívaný business rules systémem JBoss Drools

Využívání business rules pro řešení klasifikačních úloh

Jedním z vhodných způsobů využívání business rules na základě výsledků úloh dolování asociačních pravidel je jejich využívání pro podporu rozhodování (klasifikace entit) v rámci expertních a doporučovacích systémů. Výhodou využití modelu získaného na základě asociačních pravidel oproti běžně využívaným klasifikačním dataminingovým modelům (rozhodovacím pravidlům, stromům atd.) je výrazně nižší pravděpodobnost přeučení vytvořeného modelu a zároveň jeho větší flexibilita. Pravidla získaná dolováním je možné velmi jednoduše kombinovat s pravidly zadanými doménovými experty či například s aktuálním zadáním z managementu podniku.[47]

V rámci provedeného testování bylo ověřeno, že klasifikace prostřednictvím modelu získaného transformací sady asociačních pravidel umožňuje dosahovat lepších výsledků, než konkurenční metody. Testování bylo provedeno za použití veřejně dostupných UCI datasetů, podrobné výsledky byly zveřejněny v rámci konference RuleML (Kliegr, 2014), včetně představení algoritmu brCBA, který se od ostatních CBA algoritmů odlišuje využíváním komplexních GUHA pravidel.

Transformace asociačních pravidel do podoby business rules není zcela bezproblémová. Pokud mají být vzniklá pravidla využívána ke klasifikaci, je vhodné stanovit omezení konsekventu pravidla na přítomnost pouze jednoho atributu (optimálně s jednou konkrétní, nepředzpracovanou hodnotou). V průběhu klasifikace je nejprve vyhodnocena shoda datových atributů hodnocené entity s antecedenty jednotlivých pravidel. Druhým krokem je vyřešení možné „aktivace“ většího množství pravidel a v případě rozličných hodnot konsekventů vyřešení tohoto konfliktu. Vhodným způsobem řešení konfliktů pravidel je určení jejich preference dle jejich konfidence, podpory a délky antecedentu. Antecedent je zde podmínkou klasifikačních pravidel, preferována jsou pravidla s kratšími (tj. obecnějšími) podmínkami. (Thabtah, 2006)

V některých případech je možné podmínku omezení na jeden atribut v konsekventu pravidla odstranit, avšak u výsledku klasifikace je nutné počítat s větší mírou nejistoty. Tuto nejistotu může řešit například uživatel expertního systému, či může být výběr z výsledných atributů učiněn obslužným systémem. Možnou implementací, u které není na škodu doporučení většího množství atributů, může být systém doporučování vhodného zboží v rámci elektronického obchodu. Do zasílaného obchodního sdělení může být zahrnuto zboží z více kategorií a potencionální výskyt většího množství atributů či hodnot ve výsledku klasifikace může vést k lepším marketingovým výsledkům.

Vytvořenou sadu pravidel (*ruleset*) je možné využívat buď v podobě samostatné komponenty či webové služby, nebo jej zapojit do komplexnější sady business rules či definice podnikových procesů (například v jazyce BPEL). Varianta zapojení klasifikačních business rules do komplexnější sady pravidel je nyní implementována v rámci TAČR grantu „Automatizovaná extrakce byznys pravidel se zpětnou vazbou“, který je řešen od poloviny roku 2014.

Kombinování business rules z více zdrojů, postupné odvozování

Jak již bylo zmíněno u popisu významu slovníku pro práci s business rules, byl navržen koncept znalostní báze pro zajištění nezbytné míry abstrakce pro využívání pravidel vybraných z výsledků většího množství dataminingových úloh. Tento koncept (Vojíř, 2014) navazuje na předchozí výzkum v rámci projektu SEWEBAR, v rámci kterého byla navržena báze doménových znalostí pro sdílení informací o předzpracování. (Balhar, 2010) Tato myšlenka byl dále rozvinuta do podoby znalostní sítě, ve které je možné uchovávat nejen informace o meta-atributech, ale také o jejich vazbách na atributy, které tvoří uložená pravidla. Přímo ve znalostní bázi je možné seskupovat pravidla do skupin, které jsou následně použitelné pro datamining. Pokud vycházíme z předpokladu, že všechna data jsou namapována na meta-atributy a jejich formáty (které tak tvoří základní terminologický slovník), je možné kombinovat pravidla získaná z rozdílných datasetů. Z hlediska procesu vyhodnocování sady pravidel dojde k postupné aktivaci většího množství pravidel, přičemž pokud konsekventem aktivovaného pravidla není cílový atribut, je obsah daného konsekventu vložen jakožto další fakty do dalšího kola odvozování. Pro lepší srozumitelnost vyjděme z modelového příkladu:

[47] Takovýmto zadáním může být například dočasná preference konkrétní skupiny zákazníků v rámci akční marketingové kampaně.

Mějme systém pro doporučování vhodných nabídek balíčků služeb pro zákazníky telefonního operátora. Cílem společnosti je maximalizovat zisk - zákazníkům žádající změnu služeb by tedy měla být nabídnuta taková nabídka, kterou je pro zákazníka dostatečně zajímavá, ale zároveň co nejvýhodnější pro společnost. Zároveň společnost nestanovuje ceny individuálně – pracovníci obchodního oddělení mají k dispozici sadu připravených balíčků služeb s konkrétní cenovou nabídkou. Ve znalostním bázi máme dílčí sady pravidel:

- pravidla klasifikující zákazníky do skupin dle nejčastěji používaných služeb (získaná na základě dataminingové analýzy statistik provozu sítě v minulém období)
- pravidla definující nejpravděpodobněji akceptovanou nabídku (získaná dolováním asociačních pravidel nad záznamy od ostatních zákazníků z minulého období)

Obě sady pravidel se sice vztahují k zákazníkům a jim poskytovaným službám, ale každá sada pravidel byla získána jiným způsobem a v jiném časovém období. Jelikož jsou všechny záznamy pravidel ve formátu business rules, není nutné k nim přistupovat jako k jedné sadě rozhodovacích (klasifikačních) pravidel, která by všechna nabízela pouze jeden cílový „atribut“ s konkrétní hodnotou. Pro výběr nabídky pro konkrétního zákazníka jsou do systému nejprve načteny v podobě faktů jeho osobní atribut s příslušnými hodnotami (informace o využívání služeb v předchozích 2 měsících, věk, pohlaví, bydliště, průměrná cena vyúčtování). Na základě postupného vyhodnocování dojde postupně:

1. k určení typu předpokládaných využívaných služeb – aktivace pravidla:
   data(<2000MB;∞)) AND hovory((0;30)) → služba(3G) OR služba (LTE)
2. do atributů zákazníka jsou přidány položky *služba(3G)* a *služba(LTE)*
3. výběr vhodného balíčku – aktivace pravidel:
   služba(3G) AND vyúčtování(<300;600)) AND region(Praha) → balíček(dataStandard1)
   služba(LTE) AND vyúčtování(<500;1000)) AND region(Praha) → balíček(dataStandard2)
   služba(3G) AND vyúčtování(<300;600)) AND pohlaví(muž) → balíček(dataIndividual2)
   …
4. na základě výběru dle nejvyšší konfidence vybráno pravidlo:
   služba(3G) AND vyúčtování(<300;600)) AND pohlaví(muž) → balíček(dataIndividual2)

Tento postup postupného vyhodnocování lze využít nejen v rámci sad pravidel získaných na základě dataminingových úloh, ale také v kombinaci s jednoduchým ovlivněním výsledků v podobě ručně vloženého pravidla (například z důvodu dočasné slevové akce).

Z hlediska řešení konfliktů je možné buď v každém kroku odvozování provést vyřešení konfliktu a aktivovat pouze nejvhodnější pravidlo[48]. Alternativním přístupem je vyhodnocení většího množství variant – buď prostřednictvím zahrnutí míry „vhodnosti“ pravidla, které daný fakt (atribut s konkrétní hodnotou) přidalo od odvozovací báze, nebo prostřednictvím vracení se v rámci odvozovacího procesu. Vyhodnocení konfliktů v každém kroku odvozování je rozhodně výhodnější z hlediska časových a paměťových nároků.

Práce s business rules v systému EasyMiner

Prostřednictvím systému EasyMiner je možné nejen dolovat asociační pravidla a vytvářet příslušné analytické zprávy, ale též vytvářet klasifikační modely v podobě sad business rules. uživatelé mají možnost vybírat zvolená asociační pravidla, která následně exportují do znalostní báze, mají možnost doplňovat pravidla zadaná ručně (prostřednictví editoru business rules) a kombinovat je s pravidly získanými z většího množství dataminingových úloh. Proces tvorby sady klasifikačních business rules lze rozdělit do těchto kroků:

1. definice dataminingové úlohy, dolování asociačních pravidel, výběr vhodných výsledků
2. export vybraných pravidel do báze znalostí
3. výběr pravidel z jiných úloh, editace business rules
4. otestování klasifikačního modelu
5. export výsledného klasifikačního modelu v podobě DRL

[48] „Vhodnost“ pravidla je hodnocena dle nastavené strategie pro řešení konfliktu. Nemusí jít o konfidenci či podporu, ale například také o cenu, která je uvedena v daném pravidle.

Pro testování klasifikačního modelu je k dispozici komponenta *ModelTester*, která umožňuje vyhodnotit úspěšnost klasifikace nejen nad trénovacími daty, ale také nad volitelně nahranými daty testovacími. Uživateli vrací nejen informace o přesnosti (*accuracy*) a úplnosti (*recall*) klasifikace. Pro sestavení vhodné sady pravidel je vhodná také informace četnosti využívání jednotlivých pravidel. K řešení konfliktních výsledků pravidel je využito řazení pravidel dle konfidence, podpory a délky podmínky.

V rámci implementovaného editoru business rules máš uživatel možnost ovlivňovat obsah znalostní báze i jednotlivá pravidla. Pro umožnění práce s pravidly nejen získanými přímou transformací výsledků dataminingových úloh, ale též s pravidly zadanými uživatelem je využíváno abstrakce do podoby meta-atributů (dle popisu v předchozí podkapitole). Uživatel před zahájením dolování asociačních pravidel namapuje jednotlivé sloupce z datové matice na meta-atributy. Dané namapování je posléze využíváno také v editoru business rules, ve kterém má tak uživatel možnost konzistentně pracovat se všemi pravidly, která jsou tvořena z dané skupiny meta-atributů (bez ohledu na předzpracování v procesu dolování).

Závěr

V rámci tohoto příspěvku byla popsána možnost využívání transformace asociačních pravidel do podoby business rules pro řešení klasifikačních úloh. Byly porovnány nejčastěji využívané formáty pro záznam business rules a otestována možnost jejich praktického využití. V textu byly popsány dva možné přístupy k řešení klasifikace. První z těchto přístupů již byl v minulosti otestován na UCI datasetech. Jeho použití je efektivní a uživatelsky velmi dobře srozumitelné, avšak plně nevyužívá možností logického odvozování v rámci business rules systémů. Druhý představený přístup k řešení klasifikace za použití sady business rules umožňuje kombinovat pravidla získaná na základě rozdílných dílčích datasetů a více využívá možností business rules enginů.

Cílem dalšího výzkumu je provedení kvalitativního i rychlostního testování druhého navrženého způsobu řešení klasifikačních úloh. Dalším dílčím cílem je realizace uživatelského testování vytváření klasifikačních modelů v rámci aplikace EasyMiner.

Literatura

BALHAR, J., T. KLIEGR, D. ŠŤASTNÝ a S. VOJÍŘ, 2010. Elicitation of background knowledge for data mining. In: *Proceedings of Znalosti 2010.* Jindrichuv Hradec, s. 283-286.

CHAPPELL, David A., 2004. *Enterprise Service Bus*. O'Reilly Media, 2004. ISBN 0-596-00675-6.

KLIEGR, T., J. KUCHAŘ, D. SOTTARA a S. VOJÍŘ, 2014. Learning Business Rules with Association Rule Classifiers. In: *Rules on the Web. From Theory to Applications*. Springer International Publishing. s. 236-250. ISBN 978-3-319-09869-2.

OPENRULES, INC. Rule Learner. *Open Rules.* [online] OpenRules, Inc. [cit. 2015-01-10]. Dostupné z: http://openrules.com/rulelearner.htm.

ROSS, Ronald G. 2003. *Principles of the Business Rule Approach*. Addison-Wesley Professional. ISBN 978-0-201-78893-8.

ŠIMŮNEK, Milan, 2010. *Systém LISp-Miner*. Praha : Oeconomica, 2010. ISBN: 978-80-245-1699-8.

THABTAH, Fadi. 2006. Pruning techniques in associative classification: Survey and comparison. *Journal of Digital Information Management*, č. 3, sv. 4, s. 197–202.

VOJÍŘ, Stanislav. 2014. Concept of semantic knowledge base for data mining of business rules. In: *Znalosti 2014.* Praha : Oeconomica, 2014. ISBN 978-80-245-2054-4.

VOJÍŘ, Stanislav, R. Škrabal, A. HAZUCHA, M. Šimůnek a T. KLIEGR, 2013. SEWEBAR - výuka dolování asociačních pravidel prostřednictvím webové aplikace. In: *Znalosti 2013.* Ostrava: Vysoká škola báňská-Technická univerzita, 2013. ISBN 978-80-248-3189-3.

JEL Classification: C88, D83

Dedikace: Prezentovaný výzkum byl podpořen z IGA 20/2013

Summary

Data mining of classification business rules using EasyMiner system

Business rules are an effective mind-concept for saving and sharing of partial criteria for decision making in business processes. Individual rules are usually collected from domain experts. These experts input the rules into the knowledge base manually. Currently, a target of research in the area of business rules is (semi)automatic exploration of business rules using data mining methods. The aim of this paper is to summarize the basics of the concept of business rules, their practical application and the possibilities of obtaining sets of business rules via data mining of association rules.

After the basic characteristics of business rules as an intellectual concept, the paper contains a comparison of mostly used specifications of business rules, including the evaluation of support in software component for integration with information and decision support systems. Consequently, the main part is description of business rules exploration using the data mining of association rules and building of appropriate knowledge base. Sets of business rules can be used for solving of classification tasks. The rules can also be combined from different sources. The paper contains a model use case of consecutive solving of classification task.

The proposed approach has been experimentally implemented within the data mining system EasyMiner. Business rules gained from transformation of association rules are evaluated using JBoss Drools system.

Key words: data mining, association rules, business rules, EasyMiner, classification, Drools

Aplikace metaheuristických algoritmů při vytváření jazykových testů

Lenka Fiřtová

xfirl02@vse.cz

Doktorandka oboru Ekonometrie a operační výzkum

Školitel: doc. Ing. Tomáš Cahlík, CSc. (tomas.cahlik@vse.cz)

Abstrakt: Úkolem řady znalostních či psychologických testů je rozdělit respondenty do předem definovaných kategorií. Skóre, které odděluje dvě různé kategorie, se nazývá hraniční skóre. Článek se zabývá problémem výběru takových testových úloh z disponibilní množiny, aby bylo při předem stanovené hodnotě hraničního skóre dosaženo optimálních hodnot pravděpodobnosti správné klasifikace respondentů do kategorií. Problém lze formulovat jako úlohu nelineárního programování s binárními proměnnými. Kvůli charakteru proměnných a účelové funkce jde však o výpočetně složitou úlohu, která by byla jen obtížně řešitelná běžnými optimalizačními algoritmy. K řešení úlohy budou proto využity metaheuristické algoritmy, konkrétně simulované žíhání a genetický algoritmus, které budou uzpůsobeny charakteru úlohy.

Klíčová slova: simulované žíhání, genetický algoritmus, Measurement Decision Theory, Bayesova věta, optimalizace, výpočetní složitost, vývoj jazykových testů, nelineární účelová funkce, binární proměnné

Úvod

Následující text zabývá aplikací metaheuristických algoritmů při sestavování testů, které mají za úkol rozřadit respondenty do předem definovaných kategorií v závislosti na úrovni jejich dovedností. Jedním z praktických příkladů je rozřazování testovaných do úrovní stanovených Společným evropským referenčním rámcem (Common European Framework of Reference for Languages, dále CEFR), který definuje 6 kategorií znalosti jazyka: A1, A2, B1, B2, C1, C2.

Článek je zaměřen na situaci, kdy je úkolem testu rozřadit respondenty do právě dvou skupin (dosahujících či nedosahujících alespoň úrovně B2), proto je možné dílčí skupiny agregovat do dvou větších celků. Budeme tedy pracovat se skupinou B1*, která bude zahrnovat skupiny A1, A2, B1, a se skupinou B2*, která bude zahrnovat skupiny B2, C1, C2. Níže uvedený postup je však možné zobecnit i pro více než dvě kategorie. Budeme uvažovat testy, které se skládají z úloh s právě jednou správnou odpovědí, kde za každou správnou odpověď získá respondent 1 bod a za každou špatnou odpověď 0 bodů.

V tomto kontextu se často pracuje s principy tzv. Measurement Decision Theory (dále MDT), které budou podrobněji vyloženy dále. Klíčový je však fakt, že pokud je test sestaven z úloh, u nichž jsou z předchozího testování k dispozici odhady jejich parametrů úspěšnosti v jednotlivých skupinách, je možné na základě principů MDT určit, jakého skóre musí respondent nejméně dosáhnout, aby mohl být zařazen do té které kategorie. Takové skóre nazýváme hraničním skóre. Spočítat hodnotu hraničního skóre po sestavení testu není nijak problematické. Může však nastat situace, kdy je hraniční skóre zadáno předem, a test musí být sestaven tak, aby zadaná hodnota skóre byla skutečně hraniční. To v praxi nastává například tehdy, když je potřeba udržet konstantní hraniční skóre napříč jednotlivými termíny zkoušky, či dokonce napříč jednotlivými roky. Navíc na testy bývají obvykle kladeny mnohé další požadavky. Musí například splňovat požadavek na minimální průměrnou hodnotu diskriminace, neboli schopnost rozlišit mezi lepšími a horšími respondenty, a musí obsahovat předem daný počet úloh ze zadaných subtémat či okruhů. Často se také sestavuje více variant testu paralelně (například pro různé termíny téže zkoušky). Sestavit takovýto test, či dokonce několik variant testu, při respektování všech uvedených požadavků je bez pomoci vhodného algoritmu velice obtížné. Cílem bude tedy vytvořit algoritmus, který z disponibilních úloh se známými odhady parametrů úspěšnosti pro jednotlivé skupiny (B1*, B2*) a se známými odhady parametru diskriminace umožní vybrat takové úlohy, aby pro všechny výsledné varianty testu platilo, že

- každá úloha je nejvýše v jedné variantě testu;

- průměrná diskriminační schopnost úloh v každé variantě testu je rovna minimální stanovené hranici;
- počet úloh v každé variantě testu odpovídá specifikaci, obsahuje-li test více subtémat, pak počet úloh v každém subtématu odpovídá specifikaci;
- hraniční skóre v každé variantě testu je co možná nejblíže předem stanovené hodnotě.

Algoritmus bude naprogramován v prostředí *R*.

Measurement Decision Theory

Jelikož další postup bude využívat poznatků MDT, bude zde tato teorie stručně objasněna. MDT uvedl Rudner (2009), avšak v oblasti testování se jí zatím nedostalo takové pozornosti jako například tzv. Item Response Theory, která však vyžaduje práci s velkými počty respondentů, což nemusí být praxi vždy dosažitelné. MDT vychází z Bayesovy věty, kterou lze formulovat následovně:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)},$$

kde $P(A)$ resp. $P(B)$ jsou pravděpodobnosti jevů A resp. B, $P(A|B)$ je pravděpodobnost jevu A podmíněná výskytem jevu B a $P(B|A)$ je pravděpodobnost jevu B podmíněná výskytem jevu A. V našem případě se jevem A rozumí příslušnost do určité kategorie, jevem B pak získání určitého skóre.

Chceme-li tedy spočítat pravděpodobnost, že testovaný patří do skupiny B1* resp. B2*, za podmínky, že získal určité skóre, potřebujeme následující vstupní informace:

- pravděpodobnost, že náhodně vybraný testovaný patří do kategorie B1*, resp. B2*;
- pravděpodobnost, že testovaný získá určité skóre za podmínky, že patří do skupiny B1* resp. B2*.

K získání první informace stačí mít k dispozici údaje o rozložení kategorií v cílové populaci. U druhé požadované informace je situace složitější, protože pravděpodobnost získání určitého skóre v testu závisí na parametrech úloh, které daný test obsahuje.

Předpokládejme nyní, že test obsahuje celkem k úloh. Označme r_i průměrnou úspěšnost skupiny B1* v i-té úloze (tedy podíl osob ze skupiny B1*, které úlohu vyřeší), s_i pak průměrnou úspěšnost skupiny B2* v i-té úloze. Protože za každou úlohu může uchazeč získat právě 1 bod, je průměrná úspěšnost v každé úloze zároveň i střední hodnotou skóre v této úloze. Střední hodnota skóre v celém testu je pak součtem středních hodnot skóre v jednotlivých úlohách.

K určení rozptylu skóre v testu je potřeba znát kovarianční matici úloh tvořících test: kovarianční matici pro skupinu B1* resp. B2* označme $\mathbf{\Sigma}_{\mathrm{B1*}}$ resp. $\mathbf{\Sigma}_{\mathrm{B2*}}$. Na diagonále kovarianční matice jsou rozptyly skóre v jednotlivých úlohách a mimo diagonálu jsou kovariance skóre mezi jednotlivými dvojicemi úloh. Rozptyl v i-té úloze je roven $r_i(1 - r_i)$ pro B1* resp. $s_i(1 - s_i)$ pro B2*, neboť odpověď na úlohu je náhodnou veličinou s alternativním rozdělením (1 = správná odpověď, 0 = špatná odpověď). Kovarianci úloh lze pro každou dvojici úloh spočítat v případě, že byly všechny úlohy obsaženy současně v jednom pilotním testu. V opačném případě je možným řešením spočítat kovariance pro úlohy, které tento požadavek splňují, a na jejich základě kovariance pro zbylé dvojice úloh odhadnout. Rozptyl skóre v celém testu je pak roven $\mathbf{e^T \Sigma_{B1*} e}$ pro B1* resp. $\mathbf{e^T \Sigma_{B2*} e}$ pro B2*, kde $\mathbf{e}$ značí k-složkový vektor samých jednotek.

Označme $\mu_{\mathrm{B1*}}$ resp. $\mu_{\mathrm{B2*}}$ střední hodnotu skóre skupiny B1* resp. B2* v celém testu. Dále označme $\sigma_{\mathrm{B1*}}$ resp. $\sigma_{\mathrm{B2*}}$ rozptyl skóre skupiny B1* resp. B2* v celém testu. Budeme-li předpokládat, že skóre v každé skupině má normální rozdělení, případně že lze díky velkému počtu testových úloh využít důsledky centrální limitní věty, pak pro skóre skupiny B1* resp. B2* v celém testu platí:

skóre skupiny B1* ~ $N(\mu_{\mathrm{B1*}}, \sigma_{\mathrm{B1*}})$, kde $\mu_{\mathrm{B1*}} = \sum_{i=1}^{k} r_i$ a $\sigma_{\mathrm{B1*}} = \mathbf{e^T \Sigma_{B1*} e}$, resp.
skóre skupiny B2* ~ $N(\mu_{\mathrm{B2*}}, \sigma_{\mathrm{B2*}})$, kde $\mu_{\mathrm{B2*}} = \sum_{i=1}^{k} s_i$ a $\sigma_{\mathrm{B2*}} = \mathbf{e^T \Sigma_{B2*} e}$.

Pravděpodobnost získání určitého skóre skupinou B1* je tedy při platnosti výše uvedených předpokladů hustotou pravděpodobnosti normálního rozdělení se střední hodnotou μ_{B1*} a rozptylem σ_{B1*}, obdobně pro skupinu B2*. Aby bylo určité skóre hraniční, musí mít obě skupiny stejnou pravděpodobnost, že jej získají. Pro dvě skupiny platí, že pravděpodobnost získání hraničního skóre je v každé skupině rovna 0,5. Označme *h* hodnotu hraničního skóre. Pro hraniční skóre tedy platí, že

$$P(\mathrm{B1}^*|h) = \frac{d(h|\mathrm{B1}^*)\cdot P(\mathrm{B1}^*)}{d(h|\mathrm{B1}^*)\cdot P(\mathrm{B1}^*)+d(h|\mathrm{B2}^*)\cdot P(\mathrm{B2}^*)} = 0{,}5,$$

kde

$P(\mathrm{B1}^*|h)$ značí pravděpodobnost, že řešitel patří do skupiny B1*, pokud získal skóre *h*;

$d(h|\mathrm{B1}^*)$ značí hustotu pravděpodobnosti hraničního skóre za předpokladu, že testovaný patří do skupiny B1*, v níž má skóre normální rozdělení se střední hodnotou μ_{B1*} a rozptylem σ_{B1*};

$d(h|\mathrm{B2}^*)$ značí hustotu pravděpodobnosti hraničního skóre za předpokladu, že testovaný patří do skupiny B2*, v níž má skóre normální rozdělení se střední hodnotou μ_{B2*} a rozptylem σ_{B2*};

$P(\mathrm{B1}^*)$ značí pravděpodobnost, že náhodně vybraný testovaný patří do skupiny B1*;

$P(\mathrm{B2}^*)$ značí pravděpodobnost, že náhodně vybraný testovaný patří do skupiny B2*.

Obdobný vztah platí pro $P(\mathrm{B2}^*|h)$.

Matematický model

Nyní bude uveden obecný matematický model pro výše uvedený problém. Předpokládejme, že sestavujeme dva testy, přičemž máme na výběr *n* úloh a v každém testu je *k* subtémat. Subtémata se myslí v kontextu jazykových testů například úlohy na poslech, úlohy na porozumění textu apod. Zároveň budeme předpokládat, že sestavitel testu může některé úlohy do testu vložit pevně, například kvůli jejich obsahovým vlastnostem. Dále budeme předpokládat, že specifikace obou testů je shodná, tedy že každý test obsahuje stejná subtémata a stejný počet úloh v každém z nich, a tedy i celkově. Jde tedy de facto o dvě různé varianty téhož testu. Účelovou funkci lze definovat různými způsoby tak, abychom minimalizovali odchylku spočítaného a zadaného hraničního skóre v obou testech zároveň. Zde se bude pracovat se součtem čtvercových odchylek. V modelu budou obsažena tři vlastní omezení.

(1) Každá úloha musí být nejvýše v jednom testu.
(2) Počet úloh v každém testu musí odpovídat specifikaci, tedy od každého subtématu v něm musí být příslušný počet úloh.
(3) Každý test musí splňovat požadavek na minimální průměrnou diskriminaci.

Proměnné x_i, y_i budou binární, kde $x_i = 1$ při zařazení *i*-té úlohy do prvního testu, $y_i = 1$ při zařazení *i*-té úlohy do druhého testu.

Parametry

r_i = průměrná úspěšnost skupiny B1* v *i*-té úloze;
s_i = průměrná úspěšnost skupiny B2* v *i*-té úloze;
$\mathbf{\Sigma}_{\mathrm{B1}*}$ = kovarianční matice parametrů úspěšnosti úloh ve skupině B1*
$\mathbf{\Sigma}_{\mathrm{B2}*}$ = kovarianční matice parametrů úspěšnosti úloh ve skupině B2*
d_i = diskriminační schopnost *i*-té úlohy;
$t_{ij} = 1$, pokud patří *i*-tá úloha do *k*-tého subtématu, jinak 0, kde $j = 1 \ldots k$;
h_1 = cílové hraniční skóre v prvním testu;
h_2 = cílové hraniční skóre ve druhém testu;
D = minimální požadovaná diskriminace každého testu;
Tj = počet úloh, které mají být v každém testu z *j*-tého subtématu.

Účelová funkce

minimalizovat

$z = [P(\mathrm{B1}^*|h_1) - 0{,}5]^2 + [P(\mathrm{B1}^*|h_2) - 0{,}5]^2$, čili

$$z = \left[\frac{d(h_1|\mathrm{B1}^*)\cdot P(\mathrm{B1}^*)}{d(h_1|\mathrm{B1}^*)\cdot P(\mathrm{B1}^*) + d(h_1|\mathrm{B2}^*)\cdot P(\mathrm{B2}^*)} - 0{,}5\right]^2 + \left[\frac{d(h_2|\mathrm{B1}^*)\cdot P(\mathrm{B1}^*)}{d(h_2|\mathrm{B1}^*)\cdot P(\mathrm{B1}^*) + d(h_2|\mathrm{B2}^*)\cdot P(\mathrm{B2}^*)} - 0{,}5\right]^2$$

kde

$d(h_1|\mathrm{B1}^*)$ značí hustotu pravděpodobnosti normálního rozdělení se střední hodnotou $\sum_{i=1}^{n} x_i r_i$ a rozptylem $\mathbf{x^T \Sigma_{B1*} x}$;

$d(h_2|\mathrm{B1}^*)$značí hustotu pravděpodobnosti normálního rozdělení se střední hodnotou $\sum_{i=1}^{n} y_i r_i$ a rozptylem $\mathbf{y^T \Sigma_{B1*} y}$

$d(h_1|\mathrm{B2})$značí hustotu pravděpodobnosti normálního rozdělení se střední hodnotou $\sum_{i=1}^{n} x_i s_i$ a rozptylem $\mathbf{x^T \Sigma_{B2*} x}$;

$d(h_2|\mathrm{B2}^*)$ značí hustotu pravděpodobnosti normálního rozdělení se střední hodnotou $\sum_{i=1}^{n} y_i s_i$ a rozptylem $\mathbf{y^T \Sigma_{B2*} y}$.

Za podmínek

(1) $x_i + y_i \leq 1 \qquad \forall i = 1 \ldots n;$

(2) $\sum_{i=1}^{n} t_{ij} x_i = T_j \quad \forall j = 1 \ldots k;$

$\sum_{i=1}^{n} t_{ij} y_i = T_j \quad \forall j = 1 \ldots k;$

(3) $\sum_{i=1}^{n} d_i x_i \geq D;$

$\sum_{i=1}^{n} d_i y_i \geq D;$

$x_i \in \{0,1\}$	je-li *i*-tá úloha zařazena do prvního testu, pak x_i = 1, jinak 0	$\forall i = 1 \ldots n;$
$x_i = 1$	pro úlohy, které jsou do prvního testu vloženy pevně	$\forall i = 1 \ldots n;$
$y_i \in \{0,1\}$	je-li *i*-tá úloha zařazena do druhého testu, pak y_i = 1, jinak 0	$\forall i = 1 \ldots n;$
$y_i = 1$	pro úlohy, které jsou do druhého testu vloženy pevně	$\forall i = 1 \ldots n;$

Metody řešení

Optimalizace

Bylo by možné postupovat standardními optimalizačními metodami, například některou z modifikací metody větvení a mezí (Chinneck, 2004). To však skýtá dva potenciální problémy. Zaprvé, pomocí optimalizace najdeme jeden „nejlepší“ test z hlediska parametrů, bez ohledu na obsahové vlastnosti úloh. Bylo by vhodnější vygenerovat několik testů s uspokojivými parametry a následně z nich sestavitele nechat vybrat ten nejvhodnější po obsahové stránce.

Zadruhé se jedná o výpočetně náročnou úlohu, neboť zde pracujeme s nelineární účelovou funkcí v kombinaci s binárními proměnnými. Jak uvádí Murray a Kien-Ming (2008), nelineární optimalizační problémy s celočíselnými proměnnými jsou NP-obtížné i v nejjednodušším případě, kdy je účelová funkce kvadratická a omezující podmínky lineární.

Simulované žíhání

Tato metaheuristika se inspirovala procesem žíhání oceli, odtud její název. Podrobně algoritmus popsali například Bertismas a Tsitsiklis, (1993). Definujeme teplotu $T > 0$, parametr ochlazování *s* z intervalu (0, 1), dobu žíhání *h* a počet iterací *N*. Algoritmus upravený pro potřeby našeho problému funguje na následovně:

(1) Vygenerujeme náhodné řešení *x*: náhodně vybereme úlohy do každého z testů tak, aby byly splněny veškeré omezující podmínky uvedené výše u formulace matematického modelu. Účelová funkci tohoto řešení má hodnotu *f*(*x*). Označíme toto řešení jako dosud

nejlepší nalezené řešení x^*. Každé řešení má podobu matice s dvěma sloupci. Každý sloupec tvoří binární vektor indikující zařazení úlohy do daného testu, kde jednička na pozici i, j znamená, že i-tá úloha je součástí j-tého testu.

(2) Opakujeme h-krát: Zvolíme náhodně jiné řešení z okolí bodu x, a to tak, že jednu úlohu z jednoho z testů odebereme a vložíme do něj jinou. Pokud je tato nově vkládaná úloha zároveň i ve druhém testu, pak ji z druhého testu odebereme a vložíme do něj úlohu odebranou z prvního testu. Pokud tato nově vkládaná úloha ve druhém testu není, pak zůstane druhý test beze změny. Jestliže nové řešení splňuje veškeré omezující podmínky, pak:
   i. Pokud $f(x') < f(x)$, pak $x = x'$, tedy pokud je účelová funkce řešení x' nižší, přijmeme jej jako nové řešení a budeme prohledávat jeho okolí. Zároveň označíme toto řešení jako nové nejlepší řešení x^*.
   ii. Pokud $f(x') \geq f(x)$, pak $x = x'$ s pravděpodobností $e^{-\Delta/T}$, kde $\Delta = f(x') - f(x)$. Tedy i pokud je účelová funkce řešení x' vyšší, přijmeme jej jako nové řešení, jehož okolí budeme prohledávat, avšak pouze s určitou pravděpodobností, která závisí na tom, o kolik vyšší tato účelová funkce je. Čím horší je hodnota účelové řešení x' horší proti hodnotě účelové funkce původního řešení, tím nižší je pravděpodobnost, že x' přijmeme jako nové řešení. Prakticky to lze udělat tak, že vygenerujeme hodnotu z rovnoměrného rozdělení (0, 1), a je-li menší než $e^{-\Delta/T}$, přijmeme x' jako nové řešení.

(3) Snížíme teplotu ($T = T \cdot s$) a vracíme se do bodu 2. Opakujeme N-krát. Výsledným řešením je řešení x^* z posledního kroku výpočtu, tedy nejlepší řešení nalezené v průběhu algoritmu.

Genetický algoritmus

Tato metaheuristika vychází z Darwinovy evoluční teorie. Algoritmus popsal například Whitely (2003). Definujeme počet jedinců v generaci r, pravděpodobnost mutace z intervalu (0,1) a počet generací g. Algoritmus upravený pro potřeby našeho problému funguje na následovně:

(1) Vygenerujeme r náhodných řešení: náhodně vybereme úlohy do každého z testů tak, aby platilo, že každá úloha se vyskytuje maximálně v jednom testu a že každý test odpovídá specifikaci z hlediska počtu úloh v jednotlivých subtématech. To představuje výchozí generaci, v níž je r jedinců. Každé řešení má tudíž podobu matice s r řádky. Každý řádek tvoří binární vektor. První polovina vektoru indikuje zařazení úloh do prvního testu, druhá polovina vektoru indikuje zařazení úloh do druhého testu. Spočítáme hodnoty účelových funkcí každého jedince, který splňuje podmínku minimální stanovené diskriminace, a jedince s nejlepší hodnotou účelové funkce označíme x^*, její hodnotu pak $f(x^*)$.

(2) Opakujeme g-krát:
   i. Vybereme nejvhodnější rodiče pro reprodukci, a to tak, aby byli s větší pravděpodobností vybráni rodiče s lepší hodnotou účelové funkce.
   ii. Křížením rodičů vzniknou děti. Standardně se vezmou dva binární řetězce, které se na náhodném místě přetrhnou. Zde je však potřeba mít na paměti, že výsledný test musí odpovídat specifikaci (musí obsahovat daný počet úloh od daných subtémat), a nelze tudíž jen tak prohodit části řetězce, protože pak by bylo potřeba provádět korekce nově vzniklých jedinců, čímž by mohla být porušena vhodnost daného řešení. Proto postupujeme následovně. Uvažujme dva rodiče, kteří představují dvě dvojice testů. Náhodně vybereme jedno či více subtémat. Všechny úlohy spadající do daného subtématu mezi těmito dvěma rodiči prohodíme, a to v obou testech. Tím bude zajištěno, že takto vzniklí jedinci budou odpovídat specifikaci a zároveň bude stále každá úloha nejvýše v jednom testu.

iii. U nově vzniklých jedinců provedeme s určitou pravděpodobností náhodné mutace. Pravděpodobnost mutace je jedním ze vstupních parametrů algoritmu. Opět je potřeba si uvědomit, že testy po mutaci musí odpovídat zadané specifikaci. Zvolíme jeden ze dvou typů mutace: v prvním případě vybereme po jedné úloze z každého testu z téhož subtématu a prohodíme je mezi testy navzájem. Ve druhém případě z jednoho testu náhodně odebereme některou úlohu a nahradíme ji úlohou, která dosud v žádném testu není. V obou případech bude platit, že výsledné testy budou odpovídat specifikaci.

(3) Totéž provedeme *g*-krát. Výsledné řešení představuje nejlepší jedinec z poslední generace.

Výsledky experimentů na simulovaných datech

Celý postup bude ověřen na konkrétním příkladu na simulovaných datech. Cílem bude sestavit dvě varianty testu po 60 úlohách. Testy mají určit, kteří respondenti dosahují alespoň úrovně B2 dle CEFR. Budeme předpokládat, že osob ze skupiny B1* je v populaci 75 %, osob ze skupiny B2* pak 25 %. Hraniční skóre bylo pevně stanoveno na *x* bodů, kde za *x* budou dosazovány postupně všechny celočíselné hodnoty skóre z intervalu (28, 42).

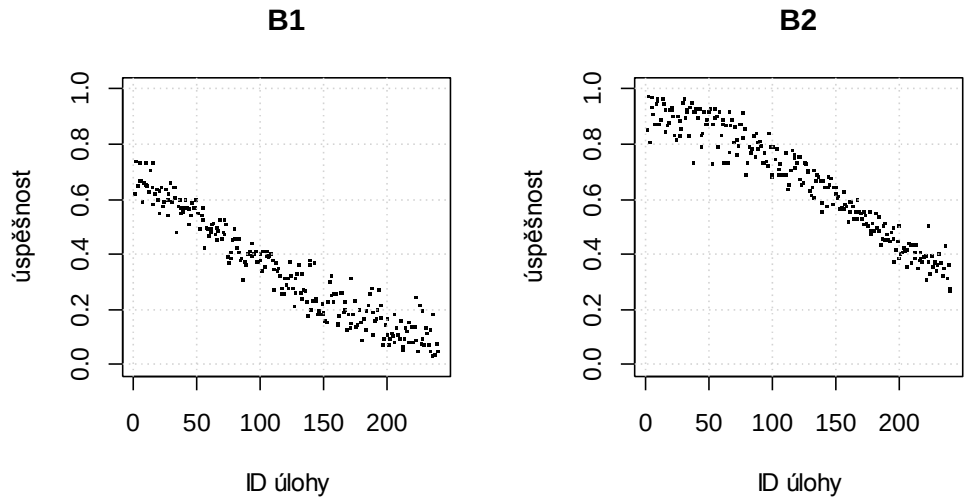
Kromě respektování předem daného hraničního skóre musí testy odpovídat zadané specifikaci: každý z nich musí obsahovat 20 úloh na porozumění textu, které sestavitel testu vybere předem, a dále 10 úloh gramatiku, 10 úloh na konverzaci, 10 poslechových úloh a 10 úloh na slovní zásobu, které budou vybírány algoritmem, tedy celkem 5 subtémat. Důvodem je, aby byla do algoritmu zabudována možnost stanovit pevně úlohy, které v testu být musí. Od každého subtématu, z něhož budou úlohy vybírány algoritmem, bude přitom k dispozici 50 úloh.

Úkolem tedy bude k 20 pevně stanoveným úlohám v každém testu vybrat z dvou set použitelných úloh 40 dalších tak, aby:

- každá úloha byla nejvýše v jednom testu;
- průměrná diskriminační schopnost úloh v každém testu byla minimálně 0,45;
- počet úloh v každém testu odpovídal specifikaci;
- hraniční skóre v každém testu bylo co možná nejblíže stanovené hodnotě *x* bodů.

U každé úlohy jsou z pilotního testování k dispozici informace o tom, s jakou úspěšností ji řešily kategorie B1* a B2*. Dále je k dispozici informace o diskriminační schopnosti každé úlohy, a to napříč celou populací. V neposlední řadě je k dispozici kovarianční matice parametrů úspěšnosti úloh pro kategorie B1* a B2*.

Tyto údaje získáme prostřednictvím simulace. Vygenerujeme si matici odpovědí pro 200 osob úrovně B1* a 200 osob úrovně B2* pro úlohy různé obtížnosti a diskriminace a spočítáme výběrovou úspěšnost v jednotlivých skupinách a výběrovou diskriminaci. Dále spočítáme výběrové kovarianční matice v každé z obou skupin. Úlohy náhodně rozřadíme do subtémat, tak aby počty odpovídaly výše uvedeným. Výběrovou úspěšnost zachycuje obrázek níže. Průměrná hodnota výběrové diskriminace je 0,57, obrázek není uveden. V realitě se stává, že některá úloha je pro horší uchazeče jednodušší než pro lepší a že diskriminace úlohy je záporná. Reálně bývají hodnoty diskriminace v průměru nižší než 0,57, avšak požadavky na průměrnou diskriminaci celého testu bývají také nižší. Cílem následujícího textu je především vyzkoušet popsané algoritmy, proto budeme pracovat s takto odhadnutými parametry úloh, byť spíše odrážejí, jak by se úlohy chovat měly, než jak se v realitě mnohdy chovají.

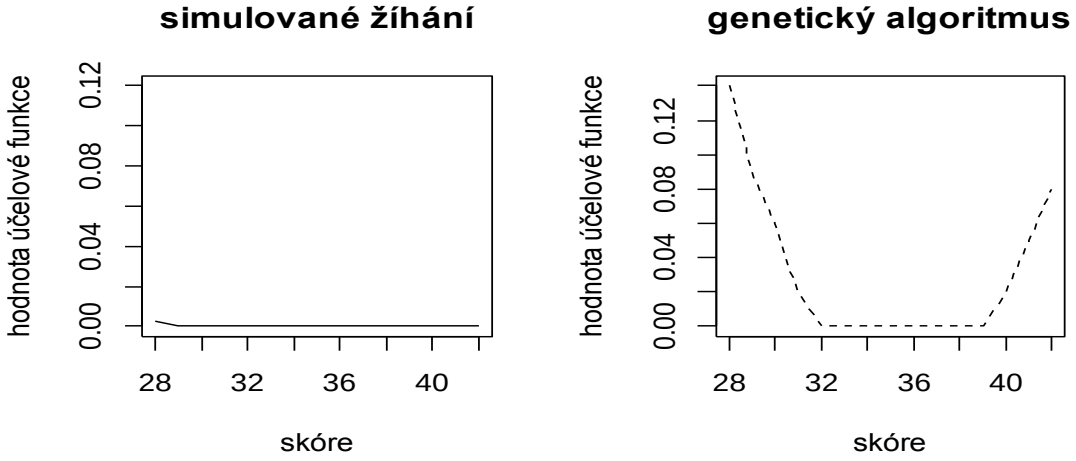


Obrázek 1: Odhad parametrů úspěšnosti úloh pro simulaci Zdroj: vlastní zpracování

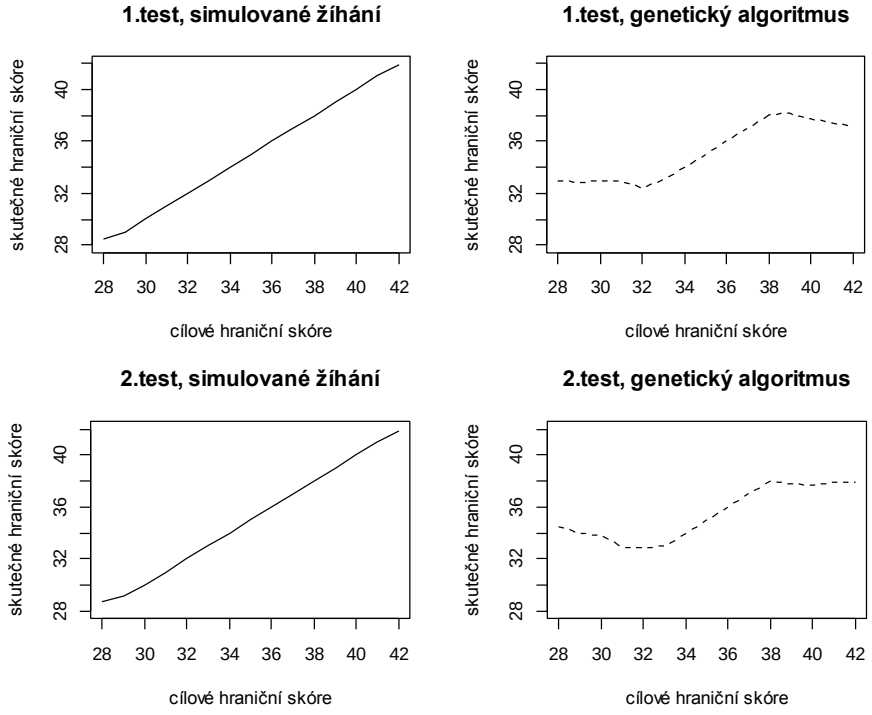
Schopnost simulovaného žíhání i genetického algoritmu najít vhodné řešení závisí mimo jiné na nastavení jejich parametrů. V obou případech platí, že pokud algoritmus necháme běžet déle, může nalézt lepší řešení, avšak za cenu delší doby prohledávání. V průběhu testování algoritmů budeme pracovat se vstupními hodnotami, které jsou určitým kompromisem mezi snahou nalézt co nejlepší řešení a snahou nalézt řešení za co nejkratší dobu.

V případě simulovaného žíhání to bude výchozí teplota 100, parametr ochlazování 0,8, doba žíhání (počet prohledávaných řešení v každé iteraci) 10, počet iterací 100. V případě genetického algoritmu to bude 50 jedinců v každé generaci, pravděpodobnost mutace 0,99 a počet generací 20. V obou případech se tudíž při jednom běhu algoritmu prohledá 1000 různých řešení.

Pro každou hodnotu hraničního skóre z intervalu (28, 42) provedeme 10 replikací pro každý z obou algoritmů. Výsledek simulace zachycují obrázky 2 a 3 níže. Obrázek 2 zachycuje průměrné hodnoty účelové funkce při použití simulovaného žíhání a genetického algoritmu pro jednotlivé hodnoty hraničního skóre napříč replikacemi, obrázek 3 pak průměrnou hodnotu skutečného hraničního skóre napříč replikacemi v jednotlivých testech, které představují finální řešení, pro jednotlivé hodnoty cílového hraničního skóre.



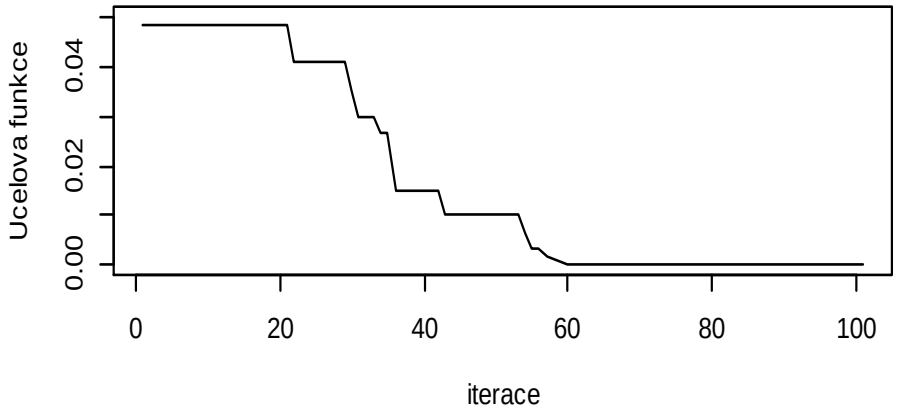
Obrázek 2: Hodnoty účelové funkce Zdroj: vlastní zpracování



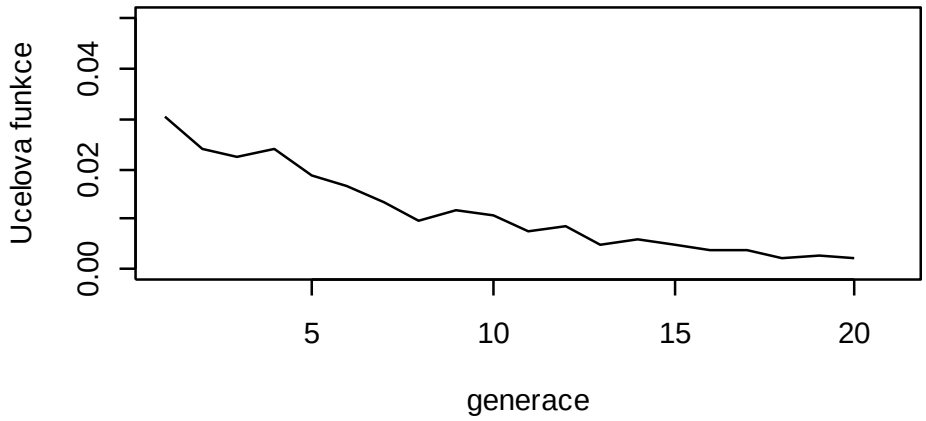
Obrázek 3: Hodnoty hraničního skóre výsledných řešení Zdroj: vlastní zpracování

Jak je z obrázků patrné, simulované žíhání se ukázalo jako vhodnější algoritmus pro řešení tohoto problému. Genetický algoritmus byl schopen nalézt vhodné řešení, čímž rozumějme řešení, kdy se skutečná hodnota hraničního skóre neodchýlila v žádném z testů od cílové hraniční hodnoty o více než 0,1 bodu, pouze v intervalu od 33 do 38 bodů. Oproti tomu simulovaným žíháním jsme našli řešení splňující tyto podmínky i pro hodnoty 30, 31, 32, 40 a 41 bodů. Je otázka, zda pro cílové hodnoty hraničního skóre pod 30 bodů či nad 41 bodů takové řešení existuje a zda by jej bylo možné úpravou vstupních hodnot algoritmu, samotného algoritmu nebo opakovaným spuštěním algoritmu nalézt.

Obrázky 4 a 5 zachycují typický průběh konvergence účelové funkce pro každý z algoritmů. Je patrné, že díky vygenerování většího počtu výchozích řešení je hodnota účelové funkce při použití genetického algoritmu zpočátku nižší ve srovnání se simulovaným žíháním. U simulovaného žíhání lze však pozorovat rychlou konvergenci k optimu, a výsledné řešení získané simulovaným žíháním bývá lepší ve srovnání s genetickým algoritmem. Simulované žíhání lze tedy pro řešení dané úlohy doporučit bez nutnosti dalších modifikací. Důvodem, proč genetický algoritmus v tomto případě nekonverguje k optimu rychleji, může být jedna z omezujících podmínek modelu, totiž nutnost dodržet specifikaci testu. Standardně se při použití genetického algoritmu náhodně kříží části řetězce, což zde nelze, neboť by tím mohlo dojít k porušení přípustnosti řešení. Výsledný řetězec by totiž mohl obsahovat jiný než žádoucí počet úloh v jednotlivých subtématech i v testu jako celku, případně by se jedna úloha mohla vyskytnout v obou testech, neboť jeden rodič představuje jednu dvojici testů. Proto byl algoritmus modifikován tak, že se prohazují části řetězce odpovídající celým subtématům a následně se provádějí náhodné mutace, což má však pravděpodobně za následek snížení schopnosti algoritmu nalézt co nejlepší řešení ve srovnání s jeho standardní formulací.



Obrázek 4: Konvergence při použití simulovaného žíhání Zdroj: vlastní zpracování



Obrázek 5: Konvergence při použití genetického algoritmu Zdroj: vlastní zpracování

Závěr

Článek se zabýval sestavováním testů, jejichž úkolem je rozřadit respondenty do předem definovaných kategorií. Konkrétně se zaměřil na otázku, jakým způsobem lze sestavit z disponibilních úloh takový test, aby hodnota hraničního skóre, které odděluje dvě kategorie, byla rovna předem zadané hodnotě. Problém byl formulován jako úloha nelineárního programování s binárními proměnnými a pro její řešení byly využity dva metaheuristické algoritmy modifikované pro potřeby úlohy: simulované žíhání a genetický algoritmus. Simulované žíhání se ukázalo jako vhodnější algoritmus a umožnilo najít přijatelné řešení pro různě zadané cílové hodnoty hraničního skóre. Genetický algoritmus poskytoval horší výsledky, a to pravděpodobně kvůli nutnosti zakomponovat do modelu omezující podmínky, které narušovaly standardní průběh algoritmu. V další práci by bylo vhodné zaměřit se na výběr nejlepších vstupních hodnot pro algoritmus simulovaného žíhání (počet iterací, teplota, rychlost ochlazování, počet prohledaných řešení) a ověřit jeho schopnost najít přijatelné řešení i pro jiné než v článku zmíněné situace. Dále by bylo přínosné pokusit se zakomponovat omezující podmínky úlohy do genetického algoritmu jiným než stávajícím způsobem, tak aby poskytoval lepší výsledky.

Literatura

BERTSIMAS, D., TSITSIKLIS, J.: Simulated Annealing. *Statistical Science* [online]. Volume 8, Number 1, 10-15, 1993 [citováno 18.1.2015]. URL http://www.mit.edu/~dbertsim/papers/Optimization/Simulated%20annealing.pdf

CHINNECK, J.: *Practical optimization, Chapter 13: Binary and Mixed-Integer Programming*. Carleton University, Canada, 2004 [citováno 28.1.2015]. URL http://www.sce.carleton.ca/faculty/chinneck/po/Chapter13.pdf

MURRAY, W., KIEN-MING, N.: An algorithm for nonlinear optimization problems with binary variables [online]. *Springer*, 2008 [citováno 18.1.2015]. URL http://web.stanford.edu/class/cme334/docs/2012-10-22-murray_ng_binary.pdf

R CORE TEAM: *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

RUDNER, L. M.: Scoring and Classifying Examinees Using Measurement Decision Theory. *Practical Assessment & Research Evaluation* [online]. Volume 14, Number 8, April 2009 [citováno 18.1.2015]. ISSN 1531-7714. URL http://pareonline.net/getvn.asp?v=14&n=8

WHITLEY, D.: *A Genetic Algorithm Tutorial*. Computer Science Department, Colorado, 2003 [citováno 18.1.2015]. URL http://www.cs.colostate.edu/~genitor/MiscPubs/tutorial.pdf

JEL Classification: C11, C61.

Dedikace: Příspěvek je zpracován jako součást výzkumné činnosti společnosti Scio.

Summary

Application of Metaheurstic Algorithms in the Context of Language Test Development

This paper focused on the development of tests whose objective is to classify examinees into several predefined categories. In this context, it is possible to use the so-called Measurement Decision Theory, based on the Bayes' theorem, so the principles of this theory were briefly explained. The paper dealt with the question of how to develop a test when the value of the cut-off score, i.e. the score which separates examinees into different categories, is set in advance. This turned out to be a computationally complex problem with a nonlinear objective function and binary variables. Such problems are difficult to solve by common optimization methods, so metaheuristic algorithms may prove a better option here. In this paper, two of them were used, namely simulated annealing and genetic algorithm, both modified with respect to the nature of the problem. Simulated annealing gave satisfactory results, which were significantly better compared to those of the genetic algorithm. This might have been caused by specific constrains incorporated into the genetic algorithm to ensure the feasibility of the solution. More in-depth study is required though to evaluate the best ways to solve the above mentioned problem.

Key words: simulated annealing, genetic algorithm, Measurement Desicion Theory, Bayes' Theorem, optimization, computational complexity, language test development, nonlinear objective function, binary variables

Modelování úrokových sazeb při proměnlivé volatilitě

Dana Cíchová Králová

xkrad900@vse.cz

Doktorand oboru statistika

Školitel: prof. Ing. Josef Arlt, CSc., (josef.arlt@vse.cz)

Abstrakt: Článek se zabývá ohodnocením platnosti a funkčnosti úrokového modelu BGM v oblasti risk managementu po finanční krizi a po následné dluhové krizi v EU. V současné době lze pozorovat různé deformace trhu, které jsou způsobeny regulací trhu, vysokou úrovní likvidity v systému nebo rostoucí rolí centrálních bank na trzích. Změna charakteru trhu je v článku demonstrována na vývoji volatility úrokových sazeb zkonstruované pomocí modelu GARCH a na vývoji kotovaných volatilit swapcí v období před finanční krizí, v období finanční krize a po odeznění finanční krize. Vzhledem k tomu, že většina běžně používaných úrokových modelů (včetně BGM modelu) byly vyvinuty za odlišných podmínek na trhu před rokem 2007, je validace těchto modelů velmi důležitá. Zatímco v krizi model BGM selhal, po odeznění finanční a dluhové krize se model jeví jako robustní i vůči změnám v charakteru trhu, které nastaly během finanční krize a po ní.

Klíčová slova: volatilita, GARCH, úrokové riziko, BGM model, euro

Úvod

Cílem tohoto článku je posoudit, zda po odeznění finanční krize a po uklidnění dluhové krize v EU je i nadále možné dále využívat modely úrokových měr pro řízení rizik portfolia obsahující instrumenty navázané na úrokovou míru a držené do splatnosti. Charakteristickým rysem trhu v období počínajícím finanční krizí je zvýšená regulace trhu, nárůst likviditního a kreditního rizika a odlišné vztahy mezi jednotlivými finančními instrumenty a účastníky trhu. Vzhledem k tomu, že většina v současnosti používaných modelů úrokových měr vznikla v době před finanční krizí, kdy na trhu platily odlišné podmínky, je validace modelů úrokových měr důležitá. Dalším cílem tohoto článku je popis chování modelu Brace-Gatarek-Musiela (dále jen „BGM“) v obdobích, která jsou charakteristická odlišným chováním trhu. V tomto článku se budu zabývat obdobím před začátkem finanční krize, obdobím finanční krize a obdobím po finanční a dluhové krizi.

Článek je organizován následovně: V první části je popsána volatilita eurových úrokových sazeb pomocí modelu GARCH ve sledovaných obdobích. V této části je taktéž popsán vývoj kotovaných volatilit swapcí[49], když tyto volatility závisí především na aktuálních informacích obchodníků na rozdíl od volatilit získaných pomocí modelu GARCH, které závisí především na historickém vývoji úrokových sazeb. Ve druhé části článku je potom aplikován komplexní úrokový model celé výnosové křivky BGM, na základě kterého jsou provedeny simulace možného vývoje eurových úrokových sazeb v jednotlivých obdobích a tyto simulace jsou srovnány se skutečným vývojem sazeb.

Veškeré výpočty v tomto článku byly učiněny v softwaru Matlab a EViews.

Analýza vývoje volatility

V této části se zaměříme na volatilitu eurových sazeb peněžního trhu, protože spolu s nárůstem likviditního a kreditního rizika lze pravděpodobně pozorovat změny ve vývoji volatility jednotlivých úrokových sazeb tvořících výnosovou křivku peněžního trhu. Pro modelování volatility použijeme model GARCH. Pro úplnost uvádíme, že popisem vývoje volatility

[49] Možnost vstoupit do úrokového swapu. Výměnou za opční prémii získává kupující právo, ale ne povinnost, vstoupit do daného swapu s emitentem ke stanovenému budoucímu datu.

konkrétní sazby získáme popis volatility jednoho konkrétního bodu na příslušné výnosové křivce.

Model ARCH (autoregressive conditional heteroscedasticity), aplikovaný Englem (1982) na modelování inflace v UK, položil základní kámen rozsáhlé třídě modelů, které charakterizují podmíněnou volatilitu. Jejich význam spočívá v tom, že umožňují zachytit měnící se podmínky nejistoty na trhu. Do této třídy modelů patří i model GARCH (generalized ARCH), který odstraňuje některé nedostatky základního modelu ARCH. Model GARCH(m,s) má tvar

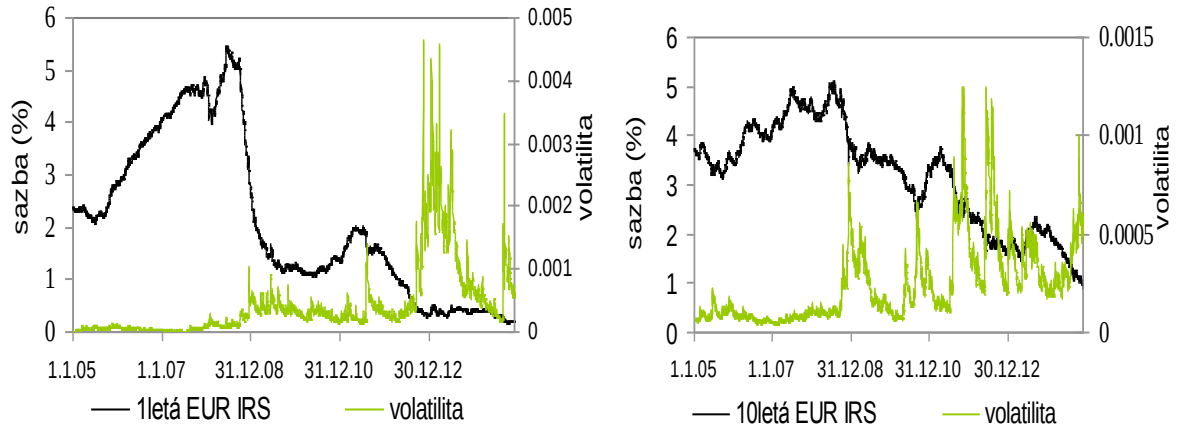
$$y_t = \mu_t + e_t, \qquad e_t = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i e_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2, \tag{1}$$

kde y_t je časová řada, μ_t označuje podmíněnou střední hodnotu a σ_t^2 představuje podmíněný rozptyl. Podmíněná střední hodnota μ_t je modelována pomocí vhodné rovnice střední hodnoty, která je obvykle lineární (často se jedná o lineární regresní model nebo ARMA proces). Dále v rovnici figurují odchylky od podmíněné střední hodnoty e_t, které jsou nekorelované stejně rozdělené náhodné veličiny s nulovou střední hodnotou, a ε_t značí nezávislé stejně rozdělené náhodné veličiny s nulovou střední hodnotou a jednotkovým rozptylem. Parametry modelu α_i , β_j splňují následující omezení

$$\alpha_0 > 0, \qquad \alpha_i \geq 0, \qquad \beta_j \geq 0, \qquad \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1, \tag{2}$$

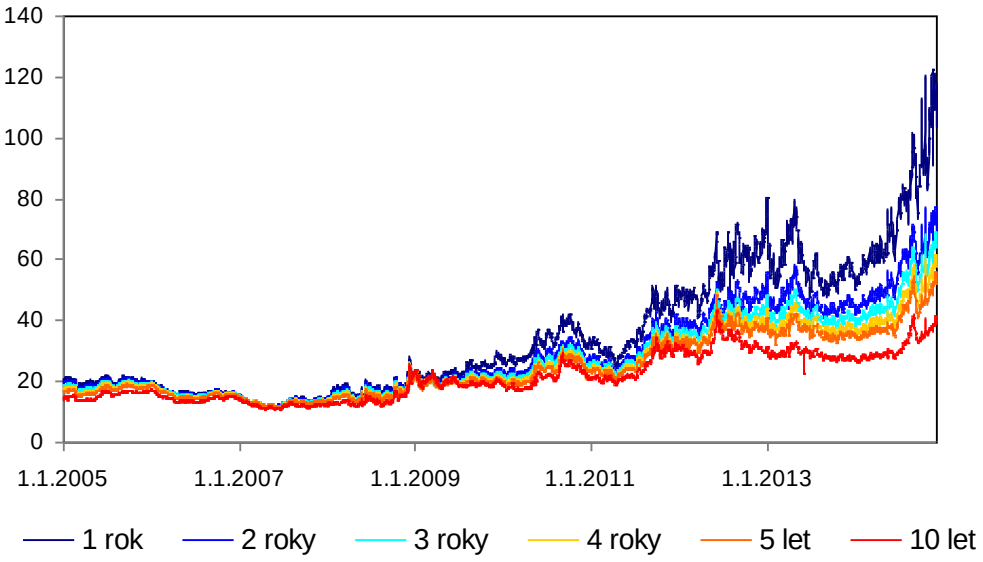
kde poslední nerovnost je postačující podmínkou pro existenci rozptylu $\mathrm{var}(e_t)$. Informace o modelování volatility jsme čerpali z knihy (Arlt, Arltová-2009).

Na obrázku 1 jsou zobrazeny volatility 1leté a 10leté EUR IRS sazby modelované pomocí modelu GARCH, ze kterých je patrné, že volatilita úrokových sazeb se před začátkem finanční krize vyvíjela odlišně oproti období po krizi. Volatilita sazeb před finanční krizí je ve srovnání s obdobím po krizi poměrně nízká a pohybovala se v relativně úzkém koridoru. V průběhu roku 2008 lze pozorovat výrazný, několikanásobný nárůst volatility oproti období před rokem 2008. Vysokou nervozitu trhů dle vývoje volatility lze pozorovat u 1leté i 10leté EUR IRS sazby po krachu Lehman Brothers v září 2008. Zajímavé je však, že volatilita úrokových sazeb se po odeznění finanční krize nesnížila, ale naopak vzrostla a v současné době se i po znatelném uklidnění nervozity trhů pohybuje zejména u 10leté EUR IRS sazby v průměru na svých maximálních hodnotách, což je o cca 300 % více oproti úrovním před finanční krizí. Volatilita 1leté EUR IRS sazby se po odeznění finanční krize během první poloviny roku 2013 prudce zvýšila na přibližně čtyřnásobek svých úrovní z období finanční krize, poté rychle klesla pod své úrovně z období finanční krize a na podzim 2014 opět skokově narostla na úrovně vyšší než v období finanční krize avšak nižší než v roce 2013. Pomocí modelu GARCH jsme dále popsali volatility ostatních splatností EUR IRS s podobnými závěry jako u vývoje volatility 10leté EUR IRS sazby.



Obrázek 1: Vývoj volatility 1letého a 10letého EUR IRS Zdroj: Bloomberg a vlastní výpočty

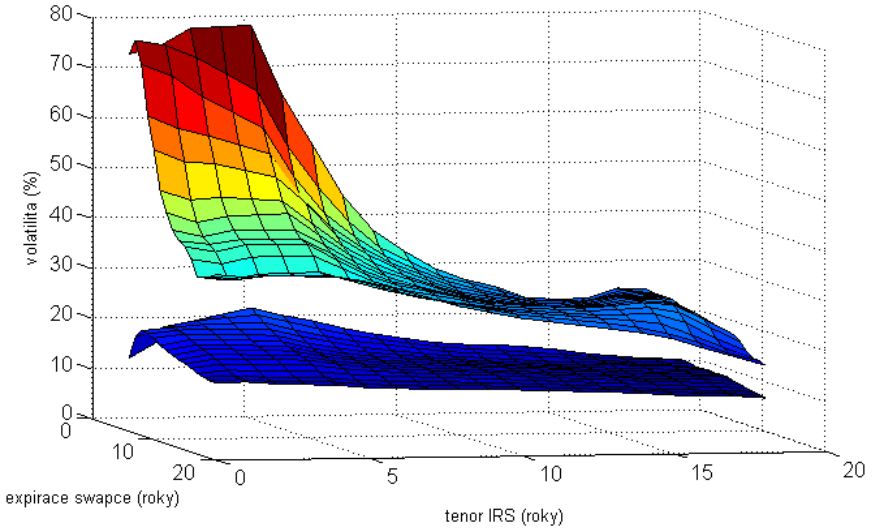
Výše popsaný vývoj volatilit EUR IRS sazeb získaný pomocí modelu GARCH je částečně konzistentní s vývojem kotací volatilit swapcí. Obrázek 2 zobrazuje kotace Blackových volatilit evropských at-the-money (dále jen ATM) swapcí s datem expirace 3 roky a délkou podkladového EUR IRS od 1roku do 10let. I z těchto tržních kotací je patrné, že volatility swapcí spojenými s příslušnými podkladovými IRS všech zobrazených délek se nacházejí po odeznění finanční krize na několikanásobně vyšších úrovních oproti jejich úrovním před finanční krizí. V roce 2014 se nacházejí kotované volatility na svých maximálních úrovních. Dále je patrné, že s delší dobou do expirace podkladového IRS kótovaná volatilita klesá. Je nutné připomenout, že zatímco volatility modelované pomocí GARCH jsou založené na historických pozorováních, kotované volatility v sobě obsahují pouze aktuální očekávání obchodníků se swapcemi (volatility swapcí kotuje např. mezibankovní broker ICAP). Uvedené vlastnosti volatilit dokumentuje i obrázek 3, který zobrazuje celou matici kotací Blackových volatilit ATM evropských swapcí k datu 1. říjnu 2005 a k 1. říjnu 2012. Z obrázku 3 je jasně patrný posun celé plochy volatilit směrem k vyšším úrovním, matice kotací volatilit k pozdějšímu datu dosahuje vyšších hodnot. Obecně lze říci, že při relativně normálních situacích na trhu klesá volatilita swapcí spolu s prodlužující se délkou podkladového IRS stejně jako s prodlužující se dobou do expirace swapce.



Obrázek 2: Vývoj kotací volatilit evropských ATM EUR swapcí (%) Zdroj: Bloomberg

Pozn.: Kotované Blackovy volatility evropských ATM swapcí s podkladovým EUR IRS s datem expirace 3 roky a délkou IRS 1, 2 , 3, 4, 5, a 10let.

Vysoké úrovně kotací volatilit swapcí po ukončení finanční krize částečně souvisí dle našeho názoru s historickou zkušeností obchodníků, kteří se do svých kotací volatilit reflektují zkušenosti z finanční krize, umocněnou vysokou nejistotou pramenící ze zásahů centrálních bank a snahy autorit o regulace finančních trhů, neboť pro obchodníky je značně rizikové předpovědět chování politiků a centrálních bankéřů. V neposlední řadě je vysoká volatilita swapcí pravděpodobně způsobena i samotnými úrokovými sazbami nacházejícími se na historicky nejnižších úrovních, kdy i relativně malá změna absolutní úrovně sazby v rámci minimálního možného přírůstku kotace znamená relativně vysokou procentuální změnu oproti situaci, kdy se absolutní úrovně úrokových sazeb nacházely na několikanásobně vyšších úrovních.



Obrázek 3: Kotace volatilit evropských ATM EUR swapcí k 1.10.2005 a k 1.10.2012 Zdroj: Bloomberg

Pozn.: Kotované Blackovy volatility evropských ATM swapcí s podkladovým EUR IRS. Expirace swapcí je zobrazena na ose x, délka podkladového IRS je zobrazena na ose y. Spodní plocha zobrazuje matici volatilit kotovaných ke dni 1. 10.2005, horní plocha zobrazuje matici volatilit kotovaných ke dni 1.10.2012.

BGM model

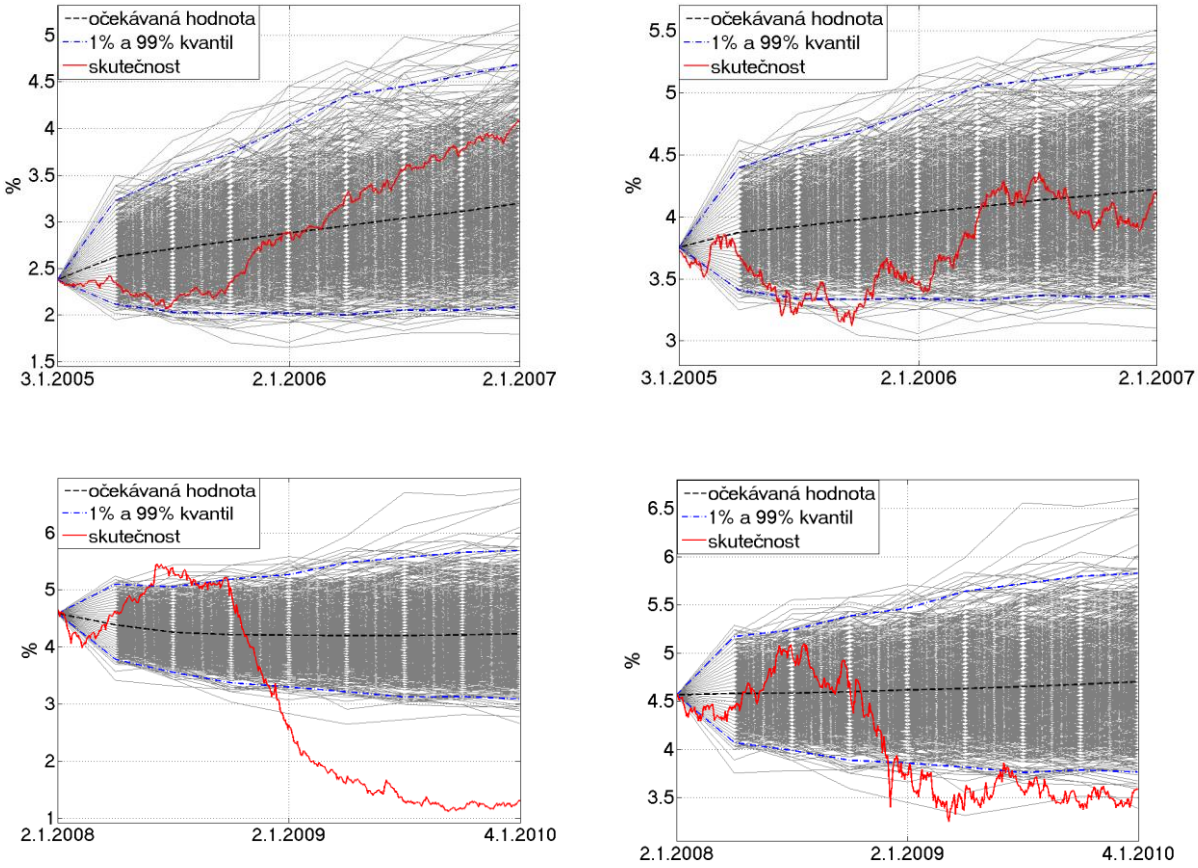
V této části popíšeme dynamiku výnosové křivky pomocí BGM modelu, který je mezi kvantitativními analytiky velmi populární a často užívaný. Jak již bylo zmíněno, hlavní využití modelu spočívá v oceňování derivátů navázaných na úrokovou míru. Tento model vytvořili v roce 1994 Miltersen, Sandmann a Sondermann (1997), poté v roce 1995 model vylepšili do podoby aplikovatelné v praxi Brace, Gatarek a Musiela (1997). Do současné podoby model upravil Jamshidian (1997), přičemž využil abstraktní formulaci z článku (Musiela, Rutkowski-1997).

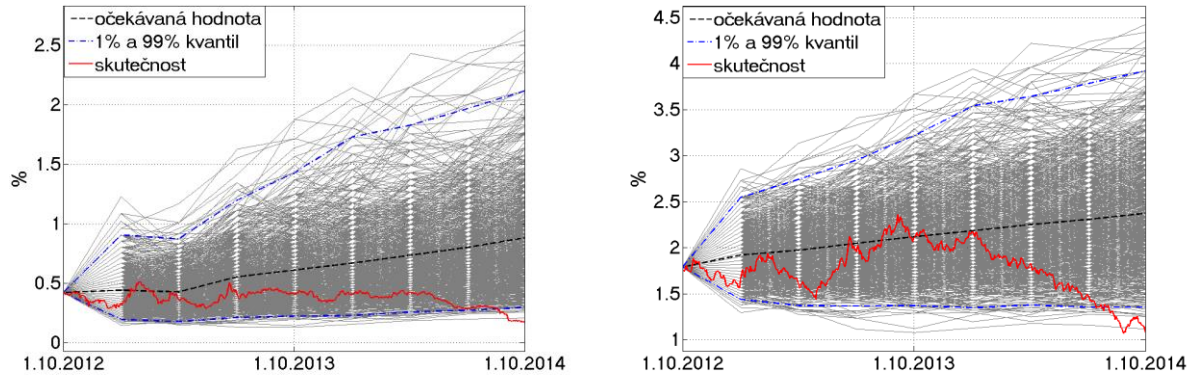
Uvažujme posloupnost časových okamžiků $0 = T_0 < T_1 < \cdots < T_m$, kde vzdálenost mezi sousedními okamžiky je konstantní. Forwardovou úrokovou míru v čase t pro období $[T_{n-1}, T_n]$ označme $L_n(t)$. V modelu BGM se forwardová úroková míra v rizikově neutrálním prostředí řídí procesem:

$$dL_n(t) = L_n(t)\gamma_n(t) \cdot dW_n(t), \qquad (3)$$

kde deterministická funkce $\gamma_n(t)$ je volatilita příslušné forwardové míry a $W_n(t)$ označuje Wienerův proces pod danou forwardovou mírou. Důkladný popis teorie a odvození modelu BGM lze nalézt například v (Brace, Gatarek, Musiela -1997). V tomto článku jsme se zaměřili především na aplikaci modelu, při níž jsme poznatky čerpali z knihy (Gatarek, Bachert, Maksymiuk-2007).

Pro modelování forwardových sazeb podle BGM modelu je třeba nejprve odhadnout volatility forwardových sazeb. Ke kalibraci se obvykle využívají jako vstupní data volatility capů či swapcí. Algoritmů kalibrace je celá řada viz například (Rebonato-2002) či (Brigo, Mercurio, Morini-2005). My jsme zvolili lokálně jednofaktorový přístup, který využívá aktuální volatility swapcí a předpokládá, že volatility jsou v daných časových intervalech konstantní. Touto metodou získáme odhady volatilit forwardových sazeb $\gamma_n(t)$. Vstupními daty jsou volatility swapcí a spotové sazby peněžního trhu. Pro aproximaci chybějících hodnot vstupních dat jsme použili nelineární regresi, detailnější popis této regrese se nachází v (Bertocchi, Dupačová, Moriggia-2000). Simulace jsme generovali na základě binomického stromu, kde jsme pro každý uzel určili hodnoty uvažovaných forwardových sazeb pomocí lognormální aproximace sazeb. Forwardové sazby pro čas nula se získají z aktuální známé spotové křivky peněžního trhu. Poté ze zkonstruovaného binomického stromu vygenerujeme simulace forwardových sazeb. Vytvořili jsme 1000 simulací pro 2letý horizont v každém sledovaném období. Na závěr převedeme simulace forwardových sazeb na spotové sazby, podrobnější vysvětlení lze nalézt v (Hull-2005).





Obrázek 4: Simulace a skutečná realizace 1letého (1.sloupec) a 10letého (2.sloupec) EUR IRS

Zdroj: Bloomberg a vlastní výpočty

Obrázek 4 zobrazuje dvouleté simulace 1leté a 10leté EUR IRS sazby provedené pomocí BGM modelu v období před krizí, tj. ke 3. lednu 2005 (první řádka), v době finanční krize, tj. ke 2. lednu 2008 (druhá řádka) a po ukončení finanční krize a po uklidnění dluhové krize v EU tj. k 1. říjnu 2012 (třetí řádka). V prvním sledovaném období před krizí setrvává skutečný vývoj 1leté i 10leté sazby kromě několika dní v intervalu daným 1% až 99% kvantily simulací. 10letá EUR IRS sazba se sice v průběhu 2. poloviny 2005 dostala pod 1% kvantil simulací, avšak neopustila prostor simulací a od druhého čtvrtletí 2006 osciluje okolo své očekávané hodnoty.

Obrázek 4 dále zobrazuje ve druhé řádce následující sledované období v době finanční krize. Při srovnání simulací se skutečným vývojem sazeb lze pozorovat jednoznačné selhání modelu BGM hlavně v případě 1leté sazby, když model nedokázal popsat prudký růst sazeb v první polovině roku 2008 stejně jako následující prudký propad sazeb v druhé polovině téhož roku. Skutečné sazby všech splatností se po celý rok 2009 pohybují mimo prostor daný 1% a 99% kvantilem simulací, avšak u splatností 10let a více se skutečné sazby pohybují alespoň v prostoru simulací.

Třetí řádka obrázku 4 zobrazuje dvouleté simulace 1leté a 10leté EUR IRS sazby po ukončení finanční krize a uklidnění dluhové krize v EU. Při srovnání simulací se skutečným vývojem sazeb lze pozorovat, že se skutečný vývoj 1leté i 10leté EUR IRS sazby pohybuje v prostoru daným 1% a 99% kvantilem simulací kromě přibližně posledního měsíce, kdy se pravděpodobně z důvodu dalšího snížení depozitních sazeb od ECB na ještě nižší záporné hodnoty dostaly tyto sazby pod 1% kvantil. Z vývoje simulací je zřejmé, že model BGM očekával v říjnu 2012 rostoucí trend sazeb všech modelovaných splatností, zatímco skutečný vývoj sazeb signalizuje spíše klesající trend hlavně v průběhu roku 2014. Fakt, že se skutečný vývoj sazeb dostal pod hranici 1% kvantilu, ještě neznamená, že model BGM nedokázal vysvětlit skutečnou volatilitu sazeb. Vzhledem k tomu, že v realitě může nastat pouze jedna z možných cest stochastického procesu (3) a zásahy ECB v současném rozsahu se v okamžiku tvorby simulací jevily jako velice nepravděpodobné, tj. byly zahrnuty obchodníky do kotací volatilit swapcí s velice nízkou váhou, sleduje skutečný vývoj sazeb pravděpodobně extrémní cestu procesu (3). Dalším zajímavým poznatkem je fakt, že v případě sazeb blízkých 0 model BGM rozpoznal asymetrické riziko mezi dalším omezeným poklesem sazeb a možností výraznějšího růstu sazeb, když rozdíl mezi 99% kvantilem simulací a očekávanou hodnotou je vyšší než rozdíl mezi očekávanou hodnotou simulací a 1% kvantilem.

Podobný vývoj jako vykazují zde zobrazené 1letá a 10letá EUR IRS sazba lze pozorovat i u ostatních splatností EUR IRS, které však nejsou v tomto článku zobrazeny z důvodu nedostatku prostoru. Obecně však lze říci, že sazby o splatnosti delší než jeden rok ale kratší než 10 let

vykazují nižší volatilitu simulací i skutečné realizace oproti volatilitě 1leté sazby avšak vyšší volatilitu oproti volatilitě 10leté sazby. Simulace sazeb i skutečné realizace sazeb o splatnosti více než 10 let pak vykazují volatilitu nižší než 10letá sazba.

Závěr

V tomto článku jsme prozkoumali několik málo aspektů využití modelů úrokových měr pro účely risk managementu portfolia finančních aktiv navázaných na úrokovou míru v období před finanční krizí, v jejím průběhu a po jejím skončení.

V první části jsme díky modelování volatility pomocí modelu GARCH učinili závěr, že volatilita na trzích výrazně vzrostla oproti období před krizí a v současné době se stabilizovala na úrovních vyšších, než činila její výše v průběhu finanční krize a o málo nižších než při dluhové krizi. Z kotací volatilit swapcí, které se nacházejí po ukončení finanční krize na svých maximálních úrovních, lze usuzovat, že obchodníci do těchto kotací reflektují své zkušenosti z období finanční krize umocněné vysokou nejistotou pramenící ze zásahů centrálních bank a snahou autorit o regulace finančních trhů v období po finanční krizi.

Ve druhé části jsme provedli srovnání skutečného vývoje úrokových sazeb a sazeb simulovaných za použití modelu BGM, z čehož jsme učinily závěr, že zatímco v krizi model BGM selhal, po odeznění finanční a dluhové krize se model jeví jako robustní i vůči změnám v charakteru trhu, které nastaly během finanční krize a po ní. Je však důležité si uvědomit, že intervence bank a zavedení různých typů regulací trhu vnáší do trhu předefinování vztahů, které modely úrokových měr vysvětlit neumějí a nemohou. Na druhé straně je však pravda, že pokud budoucí dopady těchto regulací již do svých úvah zanesou obchodníci kotující volatility swapcí, může model BGM promítnout tato očekávání do simulací sazeb. Jeho velkou výhodou je, že odhad parametrů modelu probíhá na základě aktuálních kotací, které v sobě nesou aktuální informace, a tudíž mohou hodnoty parametrů modelu zareagovat okamžitě na nové skutečnosti, zatímco odhad parametrů z historických dat toto neumožňuje. Další výhodou modelu BGM je skutečnost, že modeluje celou výnosovou křivku najednou, a tudíž v simulacích nedochází k potlačení volatility delšího konce výnosové křivky a k dalším nekonzistentnostem jako v případě simulací provedených na základě modelů krátkodobé úrokové míry, viz například (Cícha, Zimmermann-2006) a (James, Weber-2002).

Dále je nutné upozornit, že hlavní využití modelů úrokových měr je bezarbitrážní oceňování finančních aktiv, kterému se tento článek nevěnuje. Ze závěrů tohoto článku není možné učinit závěry v této oblasti, jelikož bezarbitrážní oceňování je podstatně více citlivé na porušení vazeb mezi finančními veličinami, než použití modelů úrokových měr pro účely řízení úrokového rizika portfolia instrumentů navázaných na úrokovou míru.

Je zřejmé, že rozsah tohoto článku neumožňuje provedení hlubších analýz všech aspektů vývoje finančních trhů ve sledovaných obdobích ve vztahu k aplikaci modelů úrokových měr, a proto tento článek spíše otevírá diskuzi a ponechává prostor pro další výzkum v této oblasti.

Literatura

ARLT, J. a ARLTOVÁ, M., 2009. *Ekonomické časové řady*. 1. vyd. Praha: Professional Publishing, s. 127-162. ISBN 978-80-86946-85-6.

BERTOCCHI, M., DUPAČOVÁ, J. and MORIGGIA,V., 2000. *Sensitivity of Bond Portfolio's Behavior with Respect to Random Movements in Yield Curve: A Simulation Study.* Annals of Operation Research 99. Netherlands: Kluwer Academic Publishers, p. 267-286.

BRACE, A., GATAREK, D. and MUSIELA, M., 1997. *The market model of interest rate dynamics.* Mathematical Finance, Vol. 7, No. 2, p. 127-156.

BRIGO, D., MERCURIO, F. and MORINI, M., 2005. *The LIBOR model dynamics: Approximations, calibration and diagnostics* . European journal of operational resarch, Vol. 163, No. 1, p. 30-41.

CÍCHA, M. a ZIMMERMANN, P., 2006. *Nové možnosti využití modelů úrokových měr v podmínkách ČR*. In FISCHER, J. (ed.). Sborník prací účastníků vědeckého semináře doktorského studia FIS VŠE. Praha: VŠE FIS, s. 150-157. ISBN 80-245-1037-5.

ENGLE, R. F., 1982. *Autoregressive conditional heteroscedasticity with the estimates of the variance of United Kingdom inflations.* Econometrica, Vol. 50, No. 4, p. 987-1007.

GATAREK, D., BACHERT, P. and MAKSYMIUK, R., 2007. *The LIBOR market model in practice*. Chichester, England: John Wiley & Sons, 2007. ISBN-13: 978-0-470-01443-1.

HULL, C. J., 2005. *Options, futures & other derivatives*. 6. vyd. New Jersey: Prentice Hall, ISBN-13: 978-0131499089.

JAMES, J. and WEBER, N., 2002. *Interest Rate Modeling*. Chichester: John Wiley & Sons, Ltd., ISBN-13: 978-0471975236.

JAMSHIDIAN, F., 1997. *LIBOR and swap market models and measures*. Finance and Stochastics, No. 1, p. 293-330.

MILTERSEN, K., SANDMANN, K. and SONDERMANN, D. 1997. *Closed form solutions for term structure derivatives with log-normal interest rates*. J. Finance, Vol. 52, No. 1, p. 409–30.

MUSIELA, M. and RUTKOWSKI, M., 1997. *Continuous-time term structure models: Forward measure approach.* Finance and Stochastics, Vol. 4, No. 1, p. 261-91.

REBONATO, R., 2002. *Modern Pricing of Interest-Rate Derivatives the LIBOR Market Model and Beyond. Princeton*, New Jersey, USA: Princeton University Press, ISBN: 9780691089737.

JEL Classification: G170, G180, C580.

Dedikace: Příspěvek je zpracován jako součást výzkumného projektu IGA 128/2014.

Summary

Interest Rate Modeling under Changing Volatility

The aim of this paper is a validation of popular BGM interest rate model with respect to the aftermath of the last global financial crisis and of following debt crisis in the EU and to answer question whether it is still possible to use this model for managing interest rate risk of portfolio. Nowadays, various market distortions can be observed. They are caused by market regulations, high liquidity level in the system or increasing role of central banks in markets. Change of the market character is shown on a development of interest rates volatility using GARCH model and on a development of quoted swaptions volatilities in the period before, during and after the Financial crisis. Given the fact that the most of the currently used interest rate models (including BGM model) were developed under different market conditions before 2007 is a validation of these models very important. While the BGM model failed during the Financial crisis it seems to be robust to changes in the nature of the market that occurred since September 2007.

Key words: volatility, GARCH, interest rate risk, BGM model, euro

Hodnocení výsledků fuzzy shlukování

Elena Makhalova

elena.makhalova@vse.cz

doktorand oboru statistika

Školitel: doc. Ing. Iva Pecáková, CSc. (iva.pecakova@vse.cz)

Abstrakt: Jedním z nejobtížnějších úkolů ve shlukové analýze je nalezení vhodného počtu shluků. Pro zjištění vhodného (správného) počtů shluků již existuje řada koeficientů. Otázka správného počtu shluků je však stále aktuální, protože mnohdy s rostoucím počtem shluků se snižuje úspěšnost stanovení jejich správného počtu. Tento článek se především věnuje problematice hodnocení výsledků fuzzy shlukové analýzy u datových souborů s velkým počtem shluků.

Klíčová slova: fuzzy shlukování, hodnocení výsledků, vhodný počet shluků

Úvod

Shluková analýza je vícerozměrná statistická metoda, jejímž základním cílem je zařadit objekty do skupin, a to především tak, aby dva objekty stejného shluku si byly více podobné, než dva objekty z různých shluků. Pro splnění tohoto úkolu je potřeba vyřešit řadu dálších otázek. Základní otázky ve shlukování jsou dvě: stanovení podobnosti dvou objektů a hodnocení výsledků shlukování. V tomto příspěvku se budeme věnovat tématu hodnocení výsledků fuzzy shlukování.

Fuzzy shluková analýza je založena na teorii fuzzy množin, kterou v roce 1965 navrhnul L.Zadeh (odkaz). Idea fuzzy shlukování spočitává v následujícím: každý objekt patří ke každému shluku s určitou mírou příslušnosti. Tato míra příslušnosti může nabývat hodnot v intervalu <0,1> , součet měr příslušnoti jednoho objektu ke všem shlukům se rovná 1. Otázkou však je, jak stanovit správný počet shluků. Pro stanovení vhodného (správného) počtu shluků se využívají různá kritéria. K takovým kritériím patří Dunnův koeficient a jeho modifikovaný tvar, obrysový (silhouette) koefficient, Xieho – Beniho index a další. V tomto článku se budeme věnovat hodnocení výsledků fuzzy shlukování pomocí uvedených indexů u datových souborů s velkým počtem shluků.

Fuzzy shlukování

Ve fuzzy shlukové analýze je pro každý *i*-tý objekt a *j*-tý shluk určena míra příslušnosti u_{ih},se kterou je daný objekt *i* zařazen do *j*-tého shluku. Jedná se o postup, který patří k nehierarchickým metodám shlukování. Jeho výhodou je, že může indikovat zařazení jednoho objektu do více shluků. Výstupem fuzzy shlukové analýzy je tedy matice příslušností jednotlivých objektů do shluků, kde pro míry příslušnosti musí platit:

$u_{ij} \geq 0$, *pro i=1,2,...,n a h = 1, 2, ...,k,* číslování a odsazení vzorců!!

h nebo *j*?

$$\sum_{h=1}^{k} u_{ij} \equiv 1, \text{ pro } i = 1, 2, \ldots, n,$$

kde
u_{ij} – míra příslušnosti pro variantu i do shluku j,
k – počet variant.

Existuje mnoho metod, které mohou být aplikovány pro fuzzy shlukování. Mezi ně lze zařadit například algoritmus fuzzy k-průměrů, který byl použit v tomto výzkumu. Algoritmus fuzzy *k*-průměrů je určen pro kvantitativní data a založen na minimalizaci účelové funkce J_m, která je vyjádřena vzorcem:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{c} u_{ij}^{m}\left\|x_i - c_j\right\|^2, 1 \le m \le \infty .$$

Zde

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{\left\|x_i - c_j\right\|}{\left\|x_i - c_k\right\|}\right)^{\frac{2}{m-1}}}$$

a

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \times x_i}{\sum_{i=1}^{N} u_{ij}^{m}},$$

u_{ij} je míra příslušnosti pro variantu i do shluku j,
$||x_i - c_j||$ je vzdálenost mezi objektem x_i a centrem shluka c_j,
m je fuzzifikátor (*fuzzy index*).

Tímto fuzzifikátorem m je reálné číslo větší než 1, které určuje „mlhavost“ (*fuzziness*) výsledných shluků. Na základě praktických aplikací se nejčastěji volí hodnota dva. Fuzzy shlukování je iterativní algoritmus, který hledá lokální optimální řešeni pro funkci J_m. Iterace končí, je-li dosaženo $\max_{ij}\{\left|u_{ij}^{(g+1)} - u_{ij}^{(g)}\right|\} < \varepsilon$, kde ε je kritérium mezi 0 a 1, g je počet iteraci.

Hodnocení fuzzy shlukování

K hodnocení výsledků fuzzy shlukování lze využít kritéria založená na variabilitě nebo na podobnostech (Řezanková, H. 2009) . V tomto článku se budeme věnovat především koeficientům založeným na variabilitě.
Ke kritériím založeným na variabilitě patří Dunnův koeficient rozkladu, modifikovaný Dunnův koeficient, obrysový (silhouette) koeficient a Xieho-Beniho index. Základním koeficientem pro hodnocení fuzzy shlukování je Dunnův koeficient rozkladu.

Dunnův koeficient vyjadřuje míru překrytí mezi shluky a lze ho využít ke stanovení optimálního počtu shluků. Dunnův koeficient se počítá podle vzorce

$$PC = \sum_{j=1}^{k}\sum_{i=1}^{n}\frac{u_{ij}^2}{n},$$

kde

u_{ij} – míra příslušnosti pro variantu i do j–tého shluku, jeho hodnoty leží v intervalu od $1/k$ do 1.

Optimální počet shluků k_{opt} pomocí Dunnova koeficienta lze získat nalezením maximální hodnoty tohoto koeficientu pro rozmezí od 2 do $(n - 1)$ shluků (Löster, 2011):

$$k_{opt(PC)} = \max_{2 \le k \le n-1} PC$$

Interval možných hodnot, kterých nabývá Dunnův koeficient je závísly na počtu shluků, a proto pro hodnocení shlukování většinou se používá jeho modifikovaný tvar:

$$PC_{\mathrm{mod}} = 1 - \frac{k}{k-1} \times (1 - PC)$$

Hodnoty tohoto koeficientu se nachází v intervalu od 0 (úplné fuzzy shlukování) do 1 (pevné shlukování). Optimální počet shluků k_{opt} stejně jako v případě Dunnova koeficientu, stanovíme nalezením maximální hodnoty tohoto koeficientu pro všechny hodnoty z předem zvoleného intervalu počtu shluků (Löster, 2011):

$$k_{opt(PC_{\mathrm{mod}})} = \max_{2 \le k \le n-1} PC_{\mathrm{mod}}$$

Jiné kritérium podle kterého můžeme hodnotit vysledné shluky – kompaktnost a separace. K tomu slouží Xieho-Beniho index, který využívá vstupní data:

$$XB = \frac{\pi}{s} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{k} u_{ij}^2 \times D^2(x_i, \bar{x}_j)}{n' \times \min_{1 \le j \le j' \le k} D^2(\bar{x}_j, \bar{x}_{j'})} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{k} u_{ij}^2 \times \left\| x_i - \bar{x}_j \right\|^2}{n' \times \min_{1 \le j \le j' \le k} \left\| \bar{x}_j - \bar{x}_{j'} \right\|^2}$$

Xieho-Beniho index je založen na tzv. Funkci platnosti a na jednotlivých příslušnostech.
Funkce platnosti je závislá na vstupních datech, a jak na vzdálenosti mezi centroidy shluků tak i na geometrické míře vzdálenosti. (Löster, 2011):Vážená Euklidova vzdálenost dij mezi i-tým objektem *j*-tého shluku a centroidem *j*-tého shluku (vahami jsou tady míry příslušnosti *i*-tého objektu do *j*-tého shluku). Pro dobré rozdělení bz měla hodnota tohoto indexu nabývát minimální z možných hodnot:

$$d_{ij} = u_{ij} \times D(x_i, \bar{x}_j)$$

Tehdy fuzzy variabilita *j*-tého shluku je definována podle vzorce:

$$\sigma_j = \sum_{i=1}^{n} d_{ij}^2$$

Celková variabilita σ původních dat je definována jako součet shlukových variabilit pro všechny shluky podle vzorce:

$$\sigma = \sum_{j=1}^{k} \sigma_j = \sum_{j=1}^{k}\sum_{i=1}^{n} d_{ij}^2$$

Fuzzy počet objektů, které patří do j-tého shluku, lze stanoví:

$$n'_j = \sum_{i=1}^{n} u_{ij}$$

Celkový počet objektů se stanoví jako součet fuzzy počtu objektů přes všechny shluky:

$$n' = \sum_{j=1}^{k} n'_j$$

Ukazatel celkové kompaktnosti shluků π je definován jako podíl celkové variability a celkového počtu objektů:

$$\pi_j = \frac{\sigma}{n'}$$

měří, jak jsou jednotlivé shluky kompaktní. Čím nižší hodnoty ukazatel kompaktnosti nabývá, tím více kompaktní jsou jednotlivé shluky a jedná se o lepší rozdělení.
S rostoucím počtem shluků dochází k poklesu hodnoty tohoto indexu a jeho interpretace nemá, zejména při vysokých hodnotách k (blízkých n), význam. Optimální počet shluků k_{opt} lze stanovit nalezením minima z jednotlivých indexů počítaných pro všechny počty shluků z určitého intervalu (Löster, 2011):

$$k_{opt(XB)} = \min_{2 \le k \le n-1} XB$$

SC je nejsložitější index pro hodnocení shlukování. Tento index je schopen odhadnout kompaktnost a míru separaci v datové struktuře, a ne jenom pro každý shluk zvlášť a hodnotí jak každý objekt ve shluku si nejvíc podobný a jak se liší vůči objektům z ostatních shluků.
Pro jednotlivý objekt i obrysový koeficient je založen na dvou mírách: průměrná vzdálenost mezi i-objektem a každém objektem v tomto shluku A_{iCi} a na minimální průměrné vzdálenosti mezi i-objektem a objekty v jiných shlucích. Výsledný obrysový koeficient má tvar:

$$SC = \frac{1}{n}\sum_{i=1}^{n} \frac{\min(A_{ij,}\, j \in C_{-1}) - A_{iC_i}}{\max(\min(A_{ij,}\, j \in C_{-1}), A_{iC_i})}.$$

Kde
C_{-i} centry shluků,co nezahrnují oobjekt i, C_i je centrem shluku, co zahrnuje v sobě objekt i.
SC je mírou toho, jak těsně jsou seskupeny všechny objekty ve shluku. Vysoká hodnota koeficientu (čím víc se blíže k jedničce) odpovídá nejlepšímu shlukování. *SC* hodnotí vhodný počet shluků správně, pokud variabilita ve shlucích je malá.

Provedené experimenty

Fuzzy k - means shlukování bylo aplikovano na 10 různých datových souborech s různým, ale předem známem, počtem shluků. Každý datový soubor obsahuje různý počet shluků: od 3 do 12. Každý shluk obsahuje 70 jednotek, tehdy celkový počet jednotek v každém datovém souboru je závislý na počtu shluků. Počet proměnných pro každý datový soubor je stejný – pět (numerické, prediktivní). Shluky jsou se překrývající.

Výsledky tohoto shlukování byli ohodnocenény pomocí koeficientů.

Úkolem nášého výzkumu bylo najít správný počet shluků v datovéch souborech pomocí různých koeficientů a ohodnotit, jak fungují uvedené koeficienty pro různý počet shluků. Pro shlukování byla použitá metoda fuzzy k - means shluková analýza s Euklidouskou vzdáleností. Výzkum byl proveden na datových souborech, které obsahují 3 až 12 shluků. Koeficienty PC_{mod} a XB . Většinou indikují počet shluků správně (v případech 3 až 7 shluků, tabulky 1-5). S rostoucím počtem shluků této koeficienty neuvádí počet shluků vždy správně. Tak v tabulce 6 (správný počet shluků je 8). *XB* uvádí vhodný počet shluků 9, a PC_{mod} indikuje počet správně. Avšak pro případ 9 shluků (Tabulka 9)oba dva koeficienty uvádějí počet shluků spatně – 10. Podobná situace nastává i v případě pro 10 a 11 shluků (Tabulky 8 a 9), kde oba dva koeficienty uvádějí vhodný počet shluků o 1 víc, než ve skutečnosti je. Zajímavé je, že pro případ 12 shluků (Tabulka 10) PC_{mod} na rozdíl od *XB* indexu, uvádí vhodný počet shluků správně. Tehdy můžeme říci, že v 70% případech PC_{mod} index hodnotí počet shluků správně oproti 50% úspěšnosti indexu *XB*.

Situace s indexy *SC* a *PC* není až tak moc optimistická. *SC* index správně odhadnul počet shluků ve třech případech z deseti (úspěšnost pro daný výzkum je 30%) a *PC* index ve dvou ze deseti (úspěšnost pro daný výzkum je 20%). Oba dva této indexy mají tendence ukazovat menší počet shluků, než ve skutečnosti je. *SC* index správně uvedl počet shluků jen pro případ 3 (Tabulka 1), 5 (Tabulka3) a 7(Tabulka5) shluků. V ostatních případech (pro 4,6,8,9,10,11 a 12 shluků)koeficient nefinguje. Trochu horší je situace s koeficientem *PC,* který správně odhadnul počet shluků jen ve dvou případech, a to pro 3 a 7 shluků. Pro ostatní shluky má tento koeficient tendence ukazovat významné nižší počet shluků, což svědčí o tom, že není schopen posoudit o struktuře ve shlucích, protože bere v úvahu jenom pouze míry příslušnosti objektů a nehodnotí shluky z hlediska jejich kompaktnosti a vzdálenosti mezi centriody.

Tabulka 1: Data s 3 shluky Zdroj: vlastní výpočet

Název koeficientu / Počet shluků	SC	PC	PC_{mod}	XB
2	0,5468	0,7878	0,5756	0,5321
3	**0,6953**	**0,8469**	**0,7704**	**0,2185**
4	0,5173	0,6839	0,5786	0,2227
5	0,5178	0,5967	0,4958	0,2423
6	0,5073	0,4485	0,3382	0,3123
7	0,3385	0,3607	0,2541	0,4562
8	0,1723	0,3756	0,2864	0,5055
9	0,1622	0,3379	0,2551	0,5384
10	0,1891	0,3206	0,2451	0,5531

Tabulka 2: Data s 4 shluky Zdroj: vlastní výpočet

Název koeficientu / Počet shluků	SC	PC	PC_{mod}	XB
2	0,2630	**0,7182**	0,4364	0,8766
3	**0,4697**	0,6514	0,4772	0,5008
4	0,3637	0,6888	**0,5850**	**0,4876**
5	0,4902	0,5672	0,4590	0,5051
6	0,4824	0,4872	0,3846	0,4988
7	0,3511	0,4062	0,3072	0,5087
8	0,3385	0,3679	0,2776	0,5326
9	0,2825	0,3306	0,2469	0,5586
10	0,2495	0,2956	0,2173	0,5908

Tabulka 3: Data s 5 shluky Zdroj: vlastní výpočet

Název koeficientu / Počet shluků	SC	PC	PC_{mod}	XB
2	0,3910	**0,6872**	0,3745	0,6555
3	0,4660	0,6582	0,4873	0,5643
4	0,5703	0,6712	0,5616	0,5601
5	**0,5767**	0,6619	**0,5773**	**0,3010**
6	0,5030	0,5702	0,4843	0,4445
7	0,4915	0,0381	0,1223	0,5667
8	0,4017	0,4391	0,3589	0,6322
9	0,3840	0,3837	0,3067	0,6660
10	0,2399	0,3436	0,2707	0,7654

Tabulka 4: Data s 6 shluky Zdroj: vlastní výpočet

Název koeficientu / Počet shluků	SC	PC	PC_{mod}	XB
2	0,3456	**0,7261**	0,4523	0,7655
3	0,3799	0,6022	0,4033	0,5988
4	0,3611	0,5994	0,4659	0,5999
5	**0,4605**	0,5163	0,3954	0,5067
6	0,3867	0,4976	**0,4971**	**0,3877**
7	0,3496	0,4289	0,3338	0,4001
8	0,3321	0,3794	0,2907	0,4045
9	0,3234	0,3419	0,2596	0,5599
10	0,3234	0,3027	0,2253	0,5877

Tabulka 5: Data s 7 shluky Zdroj: vlastní výpočet

Název koeficientu / Počet shluků	SC	PC	PC_{mod}	XB
2	0,2541	0,5328	0,0656	0,8871
3	0,3616	0,4742	0,2113	0,7117
4	0,4638	0,5012	0,3349	0,7117
5	0,4088	0,5517	0,4397	0,6540
6	0,5504	0,5648	0,4778	0,6432
7	**0,5814**	**0,5691**	**0,4973**	**0,3221**
8	0,4308	0,5019	0,4307	0,5444
9	0,4480	0,4477	0,3786	0,4577
10	0,4490	0,4105	0,3450	0,4954

Tabulka 6: Data s 8 shluky Zdroj: vlastní výpočet

Název koeficientu / Počet shluků	SC	PC	PC_{mod}	XB
2	0,4146	**0,6749**	0,3498	0,5661
3	0,3952	0,6056	0,4084	0,5511
4	0,4495	0,5571	0,4095	0,5872
5	0,2928	0,5776	0,4721	0,5023
6	0,4381	0,6142	0,5271	0,4871
7	**0,5761**	0,5972	0,5300	0,4776
8	0,5103	0,5411	**0,5355**	0,4444
9	0,4272	0,4891	0,4252	**0,4120**
10	0,4506	0,4378	0,3754	0,4599

Tabulka 7: Data s 9 shluky Zdroj: vlastní výpočet

Název koeficientu / Počet shluků	SC	PC	PC_{mod}	XB
2	0,2341	0,5333	0,1233	0,7554
3	0,3446	0,4741	0,2334	0,6350
4	0,5038	0,5312	0,3311	0,5411
5	0,5118	0,5407	0,4444	0,4883
6	0,5422	**0,5800**	0,4442	0,4880
7	0,5814	0,5690	0,4567	0,3809
8	**0,6611**	0,5419	0,4600	0,4801
9	0,4480	0,5467	0,4989	0,4592
10	0,4490	0,5607	**0,5777**	**0, 3334**
11	0,5112	0,4778	0,4780	0,3546
12	0,5331	0,4723	0,4391	0,3444
13	0,5551	0,4659	0,3861	0,4819
14	0,5556	0,4129	0,3831	0,5111

Tabulka 8: Data s 10 shluky Zdroj: vlastní výpočet

Název koeficientu / Počet shluků	SC	PC	PC_{mod}	XB
2	0,4146	0,6001	0,4418	0,6369
3	0,3952	0,6056	0,4166	0,5510
4	0,4495	0,5001	0,4111	0,4875
5	0,2928	0,5006	0,4561	0,5068
6	0,4381	**0,6142**	0,5201	0,4002
7	0,5761	0,5877	0,5551	0,3527
8	0,5103	0,5444	0,5371	0,3444
9	**0,8123**	0,4881	0,5221	0,3212
10	0,4506	0,4328	0,6987	0,3112
11	0,4194	0,3980	**0,7651**	**0,2310**
12	0,4213	0,3812	0,6439	0,4566
13	0,4111	0,3716	0,6198	0,5561
14	0,3640	0,3511	0,5410	0,5678

Tabulka 9: Data s 11 shluky Zdroj: vlastní výpočet

Název koeficientu / Počet shluků	SC	PC	PC_{mod}	XB
2	0,2001	0,5328	0,1656	0,6011
3	0,3634	0,4612	0,2993	0,5699
4	0,4911	0,5009	0,3376	0,5661
5	0,5122	0,5107	0,4367	0,5011
6	0,5522	0,5223	0,4678	0,4989
7	0,5634	0,5230	0,4999	0,4934
8	0,5699	0,5421	0,5219	0,4917
9	**0,6455**	0,5564	0,6421	0,4888
10	0,5590	**0,7105**	0,6929	0,4876
11	0,5321	0,6555	0,7322	0,4677
12	0,5321	0,6512	**0,7450**	**0,4302**
13	0,5122	0,6499	0,6875	0,4313
14	0,4877	0,6121	0,5551	0,4558

Tabulka 10: Data s 12 shluky Zdroj: vlastní výpočet

Název koeficientu / Počet shluků	SC	PC	PC_{mod}	XB
2	0,5046	0,5432	0,4443	0,8997
3	0,5111	0,5551	0,4879	0,8661
4	0,5232	0,5567	0,5112	0,8441
5	0,5622	0,5987	0,5231	0,8123
6	0,5741	0,6019	0,5449	0,7435
7	**0,6761**	0,6321	0,5677	0,7122
8	0,6623	0,6544	0,6129	0,6542
9	0,6544	0,6971	0,6321	0,5999
10	0,6310	**0,7778**	0,6476	0,5349
11	0,6119	0,7120	0,6976	0,4323
12	0,5987	0,7001	**0,7364**	0,3331
13	0,5876	0,6441	0,7009	**0,2311**
14	0,5666	0,5889	0,6554	0,2567

Závěr

O hodnocení výsledků shlukováni pomocí koeficientu můžeme říct následující:

Koeficienty PC_{mod} a *XB* většinou ukazují správný počet shluků, ale jen u datových souborech s malým počtem shluků. Pomocí funkce platnosti, která je závislá na vstupních datech, a jak na vzdálenosti mezi centroidy shluků tak i na geometrické míře vzdálenosti, *XB* index je schopen poznat strukturu shluku a dobře identifkovat vhodný počet. Dunnův koeficient vyjadřuje míru překrytí mezi shluky, ale není schopen (stejně jako SC) posoudit strukturu ve shlucích. Víceméně PC_{mod} a *XB* můžeme používat pro soubory dat s malým počtem shluků (shluky jsou překrývajíce se). S koeficientami *PC* a *SC* musíme být více opatřené při hodnocení výsledků, protože *PC* vyjadřuje jenom míru překrytí mezi shluky a SC totiž nefunguje u dat s velkou variabilitou.

Literatura

GOJUN, G.− CHAOGUN, M., JIANHONG, W. (2007): *Data Clustering theory, algorithms, and applications*. : Siam, Society for Industrial and Applied Mathematics

HÖPPNER, F., KLWONN, F., RUNKLER, T. (1999): *Fuzzy Cluster Analysis*. London: Wiley

KRUSE, R., DÖRING, C., LESOT, M., J. (2007): *Fundamentals of fuzzy clustering, advances in fuzzy clustering and its applications*. London: Wiley.

VALENTIE DE OLIVIERA, J. (2007): *Advances in Fuzzy clustering and its applications.* London: Wiley

BRODOWSKI., S. (2011): A Validity Criterion for Fuzzy Clustering. *Computational collective intelligence: technologies and applications.* Berlin: Springer-Verlag Berlin

LÖSTER, T., PAVELKA, T. (2013): Evaluating of the Results of Clustering in Practical Economic Tasks. *International Days of Statistics and Economics*. Slaný:

Melandrium, 804−818 p.

LÖSTER, T. (2011): Hodnocení výsledků metod shlukové analýzy, disertační práce, Vysoká škola ekonomická v Praze

LÖSTER, T. (2012): Kritéria pro hodnocení výsledků shlukování se známým zařazením do skupin založená na konfuzní matici. *Forum Statisticum Slovacum*, 8/7, 85−89 p.

REZAEE, B., (2010): A cluster validity index for fuzzy clustering. *Fuzzy sets and systems.* Amsterdam: Elsevier Science BV

ŘEZANKOVÁ, H. − HÚSEK, D., SNÁŠEL., V. (2009): Shluková analýza dat. Praha: Kamil Mařík – Professional Publishing

ŘEZANKOVÁ, H. − HÚSEK, D. (2012): Fuzzy Clustering: determining the number of clusters. *Computational Aspects of Social Networks* (CASoN), 277–28p.

XIE, NN. (2011): A Classification of Cluster Validity Indexes Based on Membership Degree and Applications. *Web information systems and mining.* Berlin: Springer-Verlag Berlin

ZADEH, L.A. (1965). Fuzzy Sets. *Information and control 8,* 338-353p.

ZALIK, K.R. (2011). Validity index for clusters of different sizes and densities. *Computer science, artificial intelligence.* Amsterdam: ELSEVIER SCIENCE BV

JEL Classification: C18, C38, C69.

Dedikace: Příspěvek je zpracován jako součást výzkumného projektu IGA No. 6/2013 „Hodnocení výsledků metod shlukové analýzy v ekonomických úlohách."

Summary

Evaluation results of fuzzy clustering

Let's sum up the results of current research:

Coefficients PC_{mod} and *XB* usually show the correct number of clusters, but only for data sets with a small number of clusters. Validation function is dependent on input data, on the distance between the centroids of the clusters and geometric extent distance. With the validation function help *XB* index is able to discover the structure of the cluster and define the suitable number of clusters. *PC* coefficient express the degree of overlapping between clusters, but can't (like *SC*) discover the structure in clusters. More or less PC_{mod} and *XB* can be used for data sets with a small number of clusters (clusters are overlapping). Using coefficients PC and SC must be more fitted in evaluating the results fuzzy clustering, because PC only expresses the degree of overlap between the clusters, not the cluster's structure.

Key words: fuzzy clustering, estimating the results, appropriate number of clusters.

Mathematical Models for Loss Given Default Estimation

Michal Rychnovský

michal@rychnovsky.com

Doktorand oboru statistika

Školitel: Prof. Ing. Josef Arlt, CSc. (arlt@vse.cz)

Abstract: The aim of the present work is to find new methods for predicting the Loss Given Default in the financial and banking data. The standard approaches to predicting recoveries based on linear and logistic regression are capable of working only with the vintages of data that are already observable for a given period of time. This research aims to incorporate the most recent partial recoveries, where the full observation period has not passed yet, in order to increase the predictive power of the modeling. Therefore, due to its capability of working with censored data, the survival analysis methodology is proposed to be tested. In this paper two different applications of the Cox model are introduced, and both models are then applied to real financial data to be compared to the common methods. It is shown that in the terms of the weighted residual sum of squares, the standard models based on linear and logistic regression perform better on the old sample where the recoveries are fully observable; whereas, the survival analysis based models better predict the LGD for the future.

Key words: Survival analysis, Cox model, Loss Given Default.

Introduction

Before starting with the models, let's briefly explain about the general situation of credit risk. In the banking and financial sector the most risk is connected to the event called default. Default is usually defined as a violation of debt contract conditions, such as a lack of will or a disability to pay a loan back. In the case of default, the creditor (e.g. a bank or other financial institution) suffers a loss.

In the terms of estimating these losses we often talk about the expected loss (EL), which can be calculated as

$$EL = PD \cdot LGD \cdot EAD;$$

where PD is the probability of default, LGD is the rate of loss given the default and EAD is the exposure at default (i.e. the unpaid principal at the moment of default). All of these measures can then be estimated for all potential clients based on their characteristics.

In this paper we deal with the loss given default (LGD) modeling and propose some new techniques for its estimation. The main challenge for the methods is the fact that to be able to estimate the recoveries in a given period, we need to use some very old data (i.e. vintages observable for the whole time). This means that we usually use old data samples and completely omit the recent history (or do only some expert adjustments of the model).

And this was the main motivation for incorporating this censored data to develop a model that could outperform the standard models in the terms of predicting future recoveries. For this purpose we apply the survival analysis theory[50] – and the Cox model in particular – and introduce several possible directions, which are then tested and implemented to the real business data.

Recovery Rate Models

First let's introduce a little about the recovery process. In reality when a default occurs[51] there will be a collection process in place. During this period the clients according to their capacities either pay nothing,

[50] Survival analysis is a common approach to model the time to death of biological organisms or failure in mechanical systems – generally a time to some defined event for observed subjects. In this theory censored data is a natural part incorporated in the models. For more information about the survival analysis modeling and the Cox model one can refer to Kalbfleisch and Prentice (1980), Cox (1972, 1975) and Peto (1972).

[51] There are various definitions of default, but for our purposes we can assume e.g. 90 days default after any payment – i.e. the client didn't pay the payment in time and neither within 90 days after the scheduled due date.

or repay a part or the whole owed exposure – in one or more payments. These payments are usually referred to as recoveries.

If we now denote the recovered part of the whole exposure as recovery rate (RR), we can express the loss given default as

$$LGD = 1 - RR.$$

Then the problem transforms into modeling the expected clients' recoveries in time, denote it $RR(t)$ conditional to clients' characteristics.

Before starting the explanation about the models, let's briefly explain a little bit more about the structure of the available data. For the LGD modeling we can only use the accounts which have actually defaulted and we can observe the recovery process for some time. And as already mentioned above, this is the biggest problem of the LGD or RR modeling – the longer we want to observe the recovery process, the older data we have to use. Therefore, if we want to model for example the recovery after 3 years, we can only use the 3 years old defaults, since the fresh data is not fully observable (we can call it censored). Later on we can see how this data can be used by the survival models.

Linear Regression Model

Standard approaches usually consider a univariate target variable such as the recovery rate after a fixed time interval t, and model it conditional to the characteristics of the client. The most straightforward way one can think of is using the linear regression. Besides the already mentioned disadvantage of the old data sample, this approach also struggles with the limitations for the target variable (since the RR or LGD are rates between 0 and 1, or the recovery amount between 0 and EAD) and usually very few positive recoveries (target more than 0).

Logistic Regression Model

Alternatively, one can transform the target variable into a binary target (e.g. 1 for full recovery and 0 otherwise) and use the logistic regression to estimate the probability of recovery. For this model the definition of the target variable can vary on the percentage of exposure that we consider as recovery. Later on we can compare 11 models using different bounds in the target definition and always get the expected recovery as

$$RR = P_1 RR_1 + (1 - P_1) RR_0;$$

where P_1 is the probability of recovery from the model and RR_1 and RR_0 are the average recovery rates of the recovered and unrecovered parts of the portfolio (based on the target definition).

The weakness of this approach is the fact that it doesn't use the full information about the partial recoveries (as all or some of them are assigned to zero). However, due to the above mentioned fact of only few positive recoveries, this model is often used in finance. More information about using of these standard methods in LGD modeling can be found in Rychnovský (2009).

Full Repayment Survival Model

The first approach based on the survival analysis theory considers observing the clients and measuring their time until they fully repay or achieve some given threshold. Same as for the logistic regression model we consider a binary target with various definitions of what to consider as recovery. Therefore, using the terminology of survival analysis we can say that the subjects are the clients and exit is defined as a repayment of the exposure. It is assumed that every client will repay eventually and every observation without a recovery is considered as censored in time.

Then using the Cox model for every t we get the survivor function $S(t)$ which stands for the probability that the client will "survive" time t (i.e. will not repay until time t). Therefore, using the definition of the full repayment and the expected recovery rate $RR(t)$, we can use the probability of repayment and get

$$RR(t) = 1 - S(t),$$

and for $LGD(t)$ directly

$$LGD(t) = S(t).$$

Similarly as for the case of the logistic regression introduced above, this model does not take into account the fact that for some observations the exposure was partially recovered. Same as above this can be partially fixed by considering the weighted average of the recoveries for the case of both target options. Then $RR(t)$ can be computed as

$$RR(t) = \big(1 - S(t)\big)RR_1(t) + S(t)RR_0(t);$$

Where $RR_1(t)$ is the average recovery rate of the recovered accounts in time t and $RR_0(t)$ is the average recovery rate of the unrecovered accounts. The partial recoveries are also the motivation of the second approach.

Currency Unit Survival Model

The second survival-based model understands every currency unit (or alternatively percentage) of the owed exposure at default as a subject and its repayment as the exit. Therefore, every client is represented by a set of units and their repayment times fully describing their payment history. Then, for the repaid proportion of units we observe an exit and the rest are censored to the maximal time of observation.

Again, it is assumed that every unit will be once repaid and using the theory we can understand the survivor function $S(t)$ as the probability that a currency unit with some client's characteristics will not be paid until time t. Then the expected recovered proportion of the client's exposure can be expressed as

$$RR(t) = 1 - S(t),$$

and thus

$$LGD(t) = S(t).$$

For this model, we can use either the individual currency units (such as 1 EUR) or the proportions of the owed exposures (such as 1%) as subjects. Whereas the second approach is more balanced on the client level, the first currency unit approach can put more weight to high exposure cases which can be preferred by financial companies in practice.

Real Data Application

All the models are applied to the real business data in order to compare the results and the predictive power. The data transformation and modeling is performed in SAS.

Data Overview

The data for this research was kindly provided by Česká spořitelna, a.s. It is a database of 4,000 defaulted accounts with the following properties:

- It is a homogenous product from a non-secured retail business.
- There is some unspecified definition of default identical for the whole portfolio.
- There was a unified collection process for all the accounts.
- For each account there is a history of net recoveries – all discounted by time and collection costs.
- For each account there is a set of characteristics Z_1 to Z_8, which are assumed to have predictive power. Here Z_4 is categorical and the rest are numerical; by Z_0 we denote the intercept of the models.[52]

[52] It is not the aim of this research to further categorize or optimize the set of predictors or interpret the results (which even couldn't be responsibly done since no information about the meaning of the predictors is provided). Therefore, all of the predictors are used for the models without any transformation.

For the accounts where the full collection history up to 36 months can be observed, the cumulative recoveries are summarized in Figure 1. As we can see from the graph, the recoveries of some accounts are greater than 1 which means that more than the owed amount was actually recovered in the collection process (due to collection fees etc.).

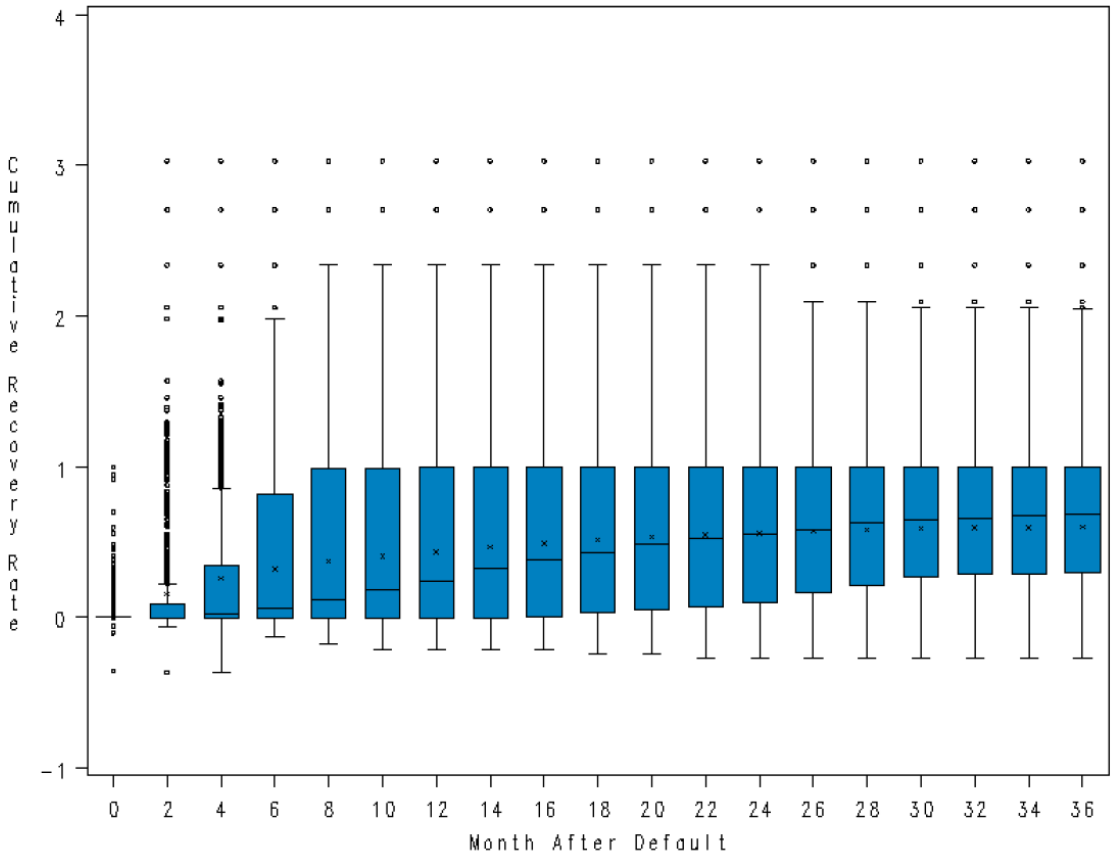


Figure 9: Provided data overview.

Data Structure for Modeling and Comparison

First let's have a look at the structure of data provided in the given sample. There is a time interval of months coded as 162–220 (without any real months reference), with a classical triangle structure – for the first month 162 we can observe the full history of 58 months, whereas for the last month we have no time to observe.

Due to the sample size of our data the 26 months recovery have been chosen for comparison. Therefore, we can only observe the full recovery for the vintages 162–194. See Figure 2 for reference. As we want to compare the predictive power of the models, we have to shorten the period for modeling and leave some data vintages for out-of-sample ex ante prediction performance testing. Therefore, for the models construction we use only the observations before month 194. There we have the full history of vintages 162–168 (we denote it as area D1) and censored data of vintages 169–194 (we denote it as area D2). The rest of the vintages 169–194 with observations after month 194 (denoted as area D3) are left for comparison of the models. For more illustration about the data structure again refer to Figure 2.

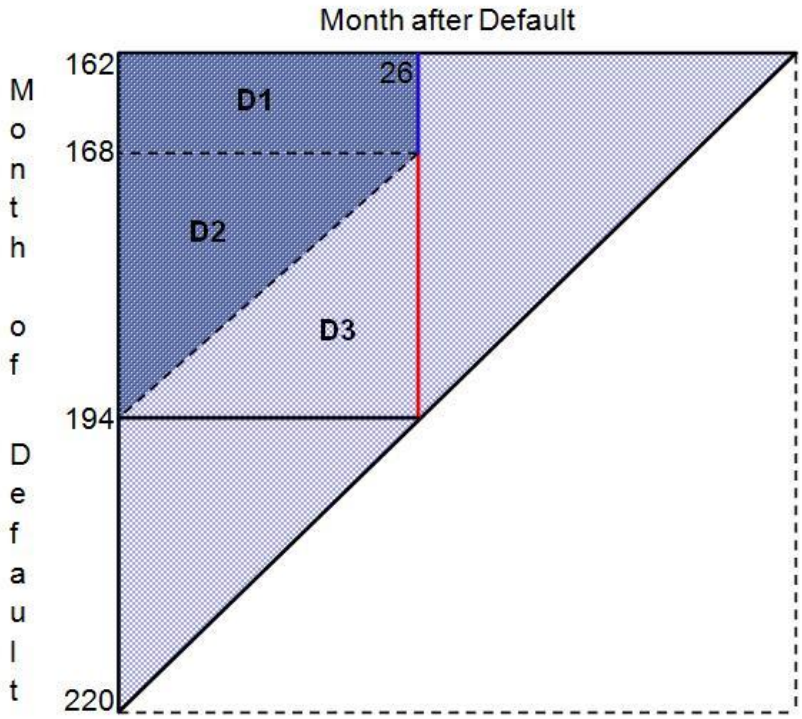


Figure 2: Triangle data and its structure for model development.

Goodness of Fit Definition

Each of the introduced models is then developed on the data sample of D1 and D2 (D1 for the linear and logistic regression and D1+D2 for the survival models), and the recovery after 26 months is estimated for each account in vintages 162–194 (D1+D3). Then we can compare the estimates with the real values using the weighted R^2defined as

$$R^2 = 1 - \frac{\sum_{i=1}^{n} w_i\left(RR(26)_i - \widehat{RR(26)_i}\right)^2}{\sum_{i=1}^{n} w_i\left(RR(26)_i - \overline{RR(26)_{pool}}\right)^2},$$

where

$$\overline{RR(26)_{pool}} = \sum_{i=1}^{n} w_i RR(26)_i$$

is the weighted average of $RR(26)_i$ and

$$w_i = \frac{EAD_i}{\sum_{k=1}^{i} EAD_k}$$

are the weights corresponding to the exposure at defaults.

Then for each model we compute the weighted R^2 separately on the area of known (development) recoveries D1, on the area of future (out-of-sample) recoveries D3, and the whole available data together D1+D3.

Results and Comparison

Linear Regression Model

For the linear regression model (weighted by the exposures) only the data of D1 is used for model development. Then the model is used to estimate the recoveries in the groups D1 and D3. The values of the modified R^2 are summarized in Table 1 and the coefficients' estimates later in Table 6. We can see that he R^2 is highest for the development sample and then rapidly decreases for the prediction.

Table 5: Results of the linear regression model

N(D1)	R^2(D1)	N(D3)	R^2(D3)	N(D1+D3)	R^2(D1+D3)
600	0.1621	1735	0.0064	2335	0.0558

Logistic Regression Model

The next tested approach for the recovery modeling is the weighted logistic regression model. The accounts from the development area D1 are divided into two groups according to their recovery rate (e.g. less than 0.1 and more than 0.1), and the probability $P(Z)$ that an account with characteristics Z will belong to the latter category is modeled using the weighted logistic regression. Then for all accounts from D1 and D3 we compute the estimated recovery as

$$RR(Z) = P(Z)RR_1 + \big(1 - P(Z)\big)RR_0;$$

where RR_0 is the weighted average recovery in the first group and RR_1 is the weighted average recovery in the second group (both computed on the D1 sample).

Altogether, 11 models are computed with different values of bounds dividing the accounts to the recovered and non-recovered categories. The results of all models can be found in Table 2. Again, the prediction for the D1 area is very weak – sometimes with negative R^2. We can see that the bound 0.1 has the best performance on D1+D3. The parameters estimates for this model can be found later in Table 6.

Table 2: Results of the logistic regression model

Bound	N(D1) N(D1)	R^2(D1)	N(D3)	R^2(D3)	N(D1+D3)	R^2(D1+D3)
0	600	0.1249	1735	0.0205	2335	0.0562
0.1	600	0.1423	1735	0.0356	2335	0.0718
0.2	600	0.1510	1735	0.0116	2335	0.0566
0.3	600	0.1510	1735	−0.0253	2335	0.0297
0.4	600	0.1591	1735	−0.0055	2335	0.0463
0.5	600	0.1529	1735	0.0038	2335	0.0514
0.6	600	0.1512	1735	0.0117	2335	0.0567
0.7	600	0.1490	1735	0.0040	2335	0.0505
0.8	600	0.1431	1735	0.0216	2335	0.0618
0.9	600	0.1261	1735	−0.0008	2335	0.0410
1	600	0.0344	1735	0.0113	2335	0.0256

Full Repayment Survival Model

For the full repayment survival model we use all the accounts from D1 and D2 (i.e. censored by the observation time of 194) for development. Again the accounts are divided into two groups according to their recovery rate (e.g. less than 0.1 and more then 0.1), and the probability that an account will belong to the latter category is understood as $1 - S(Z; 26)$, where $S(Z; 26)$ is the value of the survivor function of an account with characteristics Z in time $t = 26$. Then for all accounts from D1 and D3 we compute the estimated recovery as

$$RR(Z;26) = \big(1 - S(Z;26)\big)RR_1(26) + S(Z;26)RR_0(26);$$

where $RR_0(26)$ is the weighted average recovery after 26 months in the first group and $RR_1(26)$ is the weighted average recovery after 26 months in the second group (both computed on the D1 sample) – i.e. the same values as for the logistic regression model.

Again, 11 models are computed with different values of bounds dividing the accounts to the recovered and non-recovered categories. The results of all models can be found in Table 3. Same as for the logistic model, the bound 0.1 has the best performance on D1+D3, and the corresponding parameters estimates can be found in Table 6.

Table 3: Results of the full repayment survival model

Bound	$N(\mathbf{D1})$ $N(\mathbf{D1})$	$R^2(\mathbf{D1})$	$N(\mathbf{D3})$	$R^2(\mathbf{D3})$	$N(\mathbf{D1+D3})$	$R^2(\mathbf{D1+D3})$
0	600	0.0747	1735	0.0606	2335	0.0721
0.1	600	0.1012	1735	0.1355	2335	0.1336
0.2	600	0.0972	1735	0.1334	2335	0.1310
0.3	600	0.0925	1735	0.1091	2335	0.1121
0.4	600	0.0983	1735	0.0988	2335	0.1062
0.5	600	0.0922	1735	0.1024	2335	0.1071
0.6	600	0.0872	1735	0.1021	2335	0.1056
0.7	600	0.0874	1735	0.0994	2335	0.1037
0.8	600	0.0733	1735	0.0885	2335	0.0921
0.9	600	0.0668	1735	0.0792	2335	0.0836
1	600	0.0324	1735	0.0248	2335	0.0349

Currency Unit Survival Model

The last tested model is the currency unit survival model. Again, this model is developed on the full data of D1 and D2. In this model we observe the life cycle of every 100 CZK of the exposure as our subject. Then the survivor function $S(Z;26)$ is interpreted as the probability that a particular amount of money (100 CZK) of an account with characteristics Z will survive 26 months. Therefore, $1 - S(Z;26)$ is the expected recovery rate of the account.

Again, the modified R^2 and the parameters estimates can be found in Tables 4 and 6.

Table 4: Results of the currency unit survival model

$N(\mathbf{D1})$	$R^2(\mathbf{D1})$	$N(\mathbf{D3})$	$R^2(\mathbf{D3})$	$N(\mathbf{D1+D3})$	$R^2(\mathbf{D1+D3})$
600	0.0987	1735	0.1202	2335	0.1218

Comparison of Results

In Table 5 we can find the comparison of the four tested models. We can see that for the estimation on the development data D1 the standard approaches of the linear and logistic regression models reach better results than the survival models. Here the linear regression performs a little better than the logistic regression (even when we take into account the best model on D1 with $R^2 = 0.1591$).

However, when we compare the predictive power of the models in the ex-ante sample D3, we see that the additional information contained in the censored area D2 brought substantially better results to the survival models. Particularly the full repayment survival model with the bound of 0.1 seems very useful for this type of data.

Table 5: Comparison of results

Model	R^2(D1)	R^2(D3)	R^2(D1+D3)
Linear model	0.1621	0.0064	0.0558
Logistic model	0.1423	0.0356	0.0718
Full repayment model	0.1012	0.1355	0.1336
Currency unit model	0.0987	0.1202	0.1218

Table 6: Estimated parameters for all models[53]

Parameter	Linear model	Logistic model	Full repayment	Currency unit
β_0	1.0216	-1.6197	–	–
β_1	-0.0038	0.0071	0.0088	0.0056
β_2	$6.02 \cdot 10^{-7}$	$-7.02 \cdot 10^{-6}$	$1.92 \cdot 10^{-6}$	$1.07 \cdot 10^{-6}$
β_3	-0.0601	0.3168	0.0251	0.0014
β_{41}	-0.0582	-1.7130	-0.4934	-0.6112
β_{42}	0	0	-1.0318	-0.8827
β_{43}	-0.2607	0.4806	-0.5239	-0.4763
β_{44}	0	0	-0.0571	-0.9107
β_{45}	0	0	-0.6475	-0.5947
β_{46}	$-6.51 \cdot 10^{-4}$	-0.1235	-0.2684	-0.2033
β_{47}	0.1412	-1.6065	-0.2703	-0.3872
β_{48}	-0.0554	-0.5790	-0.1336	-0.3650
β_{49}	0.0461	-0.2139	0.0373	-0.1347
β_5	$1.88 \cdot 10^{-6}$	-0.00003	$1.58 \cdot 10^{-6}$	$1.95 \cdot 10^{-6}$
β_6	-0.5307	0.1777	-1.0498	-1.2301
β_7	0.3634	-2.8989	0.6739	0.5237
β_8	0.0792	-1.6675	0.4454	0.2779

Conclusions

The aim of this paper was to propose new methods for the loss given default modeling, where the most emphasis is put on using the fresh and unfinished observations for model development. As censoring is an integral part of survival analysis models, it is quite natural to apply their substance to improve the predictive power.

Both introduced models give us formulas to compute the expected recovery for any time within the observed period and thus complete information about the whole payment perspective.

Finally the survival models were applied to the real business data and their predictive power was measured and compared with the standard linear and logistic regression models. From this comparison it shows that the new approach outperforms the standard models in the terms of prediction and the suggested goodness-of-fit. Therefore, we can believe that there is a good potential for further application.

[53] As for the Cox model, the baseline level is given by the baseline hazard function; there is no intercept in the two survival models. Moreover, parameters β_{42}, β_{44} and β_{45} were set to 0 due to collinearity in the data used for the linear and logistic models.

For this purpose the nonparametric Cox model was chosen from the variety of widely used survival models;[54] however, a parametric alternative – such as the Accelerated Failure Time (AFT) model – can be used instead.[55]

Finally, it is fair to say that for the survival models we used the assumption that every debt would be repaid eventually, which is not true in the financial practice. Therefore, such assumption should prevent us from extending the predictions over the maximal observed time t (which could be tempting especially for the parametric models).

Bibliography

Cox, D.R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, volume 34, no. 2, pages 187–220. ISSN 1467-9868.

Cox, D.R. (1975). Partial likelihood. *Biometrika*, volume 62, no. 2, pages 269–276. ISSN 0006-3444.

Kalbfleisch, J. and Prentice, R. (1980). *The statistical analysis of failure time data*. Wiley New York. ISBN 978-04-710-5519-8.

Peto, R. (1972). Contribution to the discussion of a paper by D.R. Cox. *Journal of the Royal Statistical Society*, volume 34, page 205–207.

Rychnovský, M. (2009). *Matematické modely LGD (Mathematical models for LGD)*. Master Thesis, Faculty of Mathematics and Physics, Charles University, Prague.

JEL Classification: C51, G32.

Shrnutí

Matematické modely pro odhad LGD

Cílem této práce je najít nové metody predikce LGD (Loss Given Default, tj. ztráta při defaultu) ve finančních a bankovních datech. Standardní přístupy k odhadu budoucích výtěžností založené na lineární a logistické regresi jsou schopny pracovat pouze s daty, kdy již můžeme pozorovat výtěžnosti po uplynutí stanovené doby. Cílem tohoto výzkumu je zahrnout do vývoje také nejčerstvější částečné výtěžnosti, kde celá pozorovaná doba ještě neuplynula, a tím zvýšit predikční schopnost modelu. Právě kvůli schopnosti práce s cenzorovanými daty bylo navrženo vyzkoušet metodologii analýzy přežití. V této práci jsou představeny dvě odlišné aplikace Coxova modelu a oba modely jsou aplikovány na reálná finanční data, kde jsou srovnány s obvyklými metodami. Ukazuje se, že ve smyslu váženého součtu čtverců odchylek vyhovují standardní modely lineární a logistické regrese lépe na starším vzorku, kde jsou výtěžnosti plně pozorovatelné, zatímco pro predikci budoucího LGD jsou navržené modely založené na analýze přežití přesnější.

Klíčová slova: analýza přežití, Coxův model, LGD.

[54]The Cox model was chosen mainly due to the fact that the collection process is often differentiated into different activities in various times, and the hazard function can thus have multiple peaks. However, a parametric alternative can be tested as well.

[55] For more information about parametric models one can refer to Kalbfleisch and Prentice (1980).

Application of Goodall's and Lin's similarity measures in hierarchical clustering

Zdeněk Šulc

zdenek.sulc@vse.cz

Doktorand oboru statistika

Školitel: prof. Ing. Hana Řezanková, CSc., (hana.rezankova@vse.cz)

Abstract: The paper evaluates clustering performance of five similarity measures in nominal data clustering, which were introduced by Boriah in 2008, and compares them with the overlap measure, which is well known, and thus, it can serve as a reference similarity measure. The examined measures were derived from already known measures. There are four measures which were derived from the original Goodall's measure introduced in 1966, and one measure derived from the Lin measure, which was introduced in 1998. Originally, these measures were developed for outlier detection in categorical data, but their use can be much broader, such as clustering or classification. Unfortunately, they have never been examined in these fields of use. The aim of this paper is to analyze their quality from a point of view of their clustering performance in hierarchical cluster analysis with the complete linkage, which usually provides the best results. The various software was used for the analysis: Matlab, IBM SPSS and MS Excel. The resulting clusters are evaluated by means of the WCM coefficient based on the Gini coefficient, which measures within-cluster variability, and the pseudo F coefficient, which determines the optimal number of clusters. The similarity measures are evaluated on four different datasets, which come from the UCI Machine Learning Repository. That ensures the comparability of results with other researchers. The results show that some of examined similarity measures perform very well; thus, they can be recommended for the use in hierarchical cluster analysis with nominal variables.

Key words: similarity measures, hierarchical clustering, nominal variables, Goodall measure, Lin measure

Introduction

Goodall (1966) introduced the similarity measure for nominal data, which took into account relative frequencies of categories when comparing two observations by a given variable. On the contrary to previous approaches, represented e.g. by Sokal (1963), which treated similarity by means of transformation into binary variables, Goodall used an approach which enabled to evaluate similarity of nominal variables directly. He described the disadvantages of binary transformation by these words: "*The various methods of coding used have the disadvantage either of giving different weight to the attributes according to the number of different values they may take, or of losing information derivable from the ordering of the different values*." Since that time there have been introduced many similarity measures which do not need binary transformation, e.g. (Lin, 1998), (Eskin, 2002), (Le and Ho, 2005), etc. In such papers, the introduced similarity measures performed outstandingly well; however, most of them have not been examined outside the papers where they were introduced. There are only a few articles, where similarity measures from different sources have been evaluated, e.g. (Boriah et al., 2008), (Desai et al., 2011) or (Šulc, 2014). Furthermore, Boriah et al., (2008) introduced a few modifications of similarity measures she examined. Particularly, she introduced four modifications of the original Goodall's measure and one modification of the Lin's measure. These similarity measures have been evaluated only from a point of view of an outlier detection, which was the main topic of Boriah's article. Apart from that, they have been mentioned only in (Aggarwal, 2013) as a part of summary of similarity measures. Thus, their clustering performance is practically unknown.

The aim of this paper is to evaluate the similarity measures, which have been derived by Boriah et al., (2008), from the aspect of their clustering performance in hierarchical cluster analysis, because this area of interest presents the most typical way of their application. The similarity measures are going to be evaluated at the same datasets and by the same criteria as in the previous research, see (Šulc, 2014), to make the outcomes comparable in a broader way.

This paper has the following structure. Section 2 presents all examined similarity measures, Section 3 introduces their evaluation criteria. Section 4 contains an experimental application of these measures to four different datasets and the last section summarizes the results.

Derived similarity measures

All the similarity measures, which are examined in this paper, were firstly introduced in (Boriah et al., 2008). This article has been focused on outlier detection, thus, the similarity measures has been examined from that aspect. Since their introduction, they have never been evaluated.

Boriah et al., (2008) introduced five new similarity measures. Four measures are variations of the original Goodall's measure and one of the Lin's measure. According to Goodall's measures, the authors justified their creation by much lower computational expensiveness than the original one. The measures take values which do not exceed the interval from zero to one; the higher value indicates higher similarity. When comparing two categories, all measures change the value in case of match, whereas the value in case of mismatch stays always zero. The adjusted Lin's measure is more complex to the original measure, especially in a weight system. It takes values from zero; its upper boundary is not set.

The computation of every dissimilarity measure consists of three steps. In the first one, similarity is computed between two objects for a given variable; in the second one, a similarity measure across all variables is determined, and in the last one, a similarity measure is transformed into a dissimilarity measure. In this paper, all formulas are based on data matrix $\mathbf{X} = [x_{ic}]$, where $i = 1, 2, ..., n$ (n is the total number of objects); $c = 1, 2, ..., m$ (m is the total number of variables). All the similarity measures use the following notation: $p_c(x)$ is a relative frequency of value x in the variable c.

The Goodall 1 measure is the most similar one to the original Goodall's measure; computation of similarity on the variable level (first step of computation) is exactly the same, i.e.

$$S_c(x_{ic}, x_{jc}) = \begin{cases} 1 - \sum_{q \in Q} p_c^2(q) \text{ if } x_{ic} = x_{jc} \\ 0 \text{ otherwise} \end{cases}, \; Q \subseteq X_c : \forall q \in Q, p_c(q) \leq p_c(x_{ic}), \tag{1}$$

where q represents a subset of categories in the c-th variable whose relative frequencies are lower or equal to $p(x_{ic})$. The Goodall 1 similarity measure takes values from 0 to $1 - 2n(n-1)$ on a variable level.

In the original Goodall's measure, dependencies among variables were analyzed. Goodall 1 represents much simpler approach, where similarity between objects $\mathbf{x}_i$ and $\mathbf{x}_j$ across all variables is computed as an arithmetic mean:

$$S(\mathbf{x}_i, \mathbf{x}_j) = \frac{\sum_{c=1}^{m} S_c(x_{ic}, x_{jc})}{m}. \tag{2}$$

Since the range of possible values does not exceed the interval from zero to one, the transformation of a similarity measure into a dissimilarity measure can be performed as

$$D(\mathbf{x}_i, \mathbf{x}_j) = 1 - S(\mathbf{x}_i, \mathbf{x}_j). \tag{3}$$

The Goodall 2 measure assigns higher similarity to the infrequent matches under condition that there are also other categories, which are even less frequent than the examined one. In other words, this measure assigns higher similarity if the relative frequencies of categories are approximately equally distributed, see (Boriah et al., 2008). On a variable level, similarity between two objects is expressed as

$$S_c(x_{ic}, x_{jc}) = \begin{cases} 1 - \sum_{q \in Q} p_c^2(q) \text{ if } x_{ic} = x_{jc} \\ 0 \text{ otherwise} \end{cases}, \; Q \subseteq X_c : \forall q \in Q, p_c(q) \geq p_c(x_{ic}). \tag{4}$$

In comparison of this equation to Equation (1), the only difference is in definition of q. By Goodall 2, q represents a subset of categories in the c-th variable whose relative frequencies are higher or equal to $p(x_{ic})$. The range of this measure is same as by the Goodall 1 measure. It is also possible to use the Equations (2) and (3) for computation of similarity and dissimilarity measure across all variables.

The Goodall 3 measure was proposed to assign higher similarity if the infrequent categories match regardless on frequencies of other categories. Similarity between two objects by the *c*-th variable is expressed by the formula

$$S_c(x_{ic}, x_{jc}) = \begin{cases} 1 - p_c^2(x_{ic}) \text{ if } x_{ic} = x_{jc} \\ 0 \text{ otherwise} \end{cases}. \tag{5}$$

The range of this measure is equal to the one in Goodall 1 and Goodall 2. Also, the second and third step of computation can be completed by using Equations (2) and (3).

The Goodall 4 assigns higher similarity if the frequent categories match. Actually, when measuring similarity between two objects, this measure provides complement results of Goodall 3 to one, i.e.

$$S_c(x_{ic}, x_{jc}) = \begin{cases} p_c^2(x_{ic}) \text{ if } x_{ic} = x_{jc} \\ 0 \text{ otherwise} \end{cases}. \tag{6}$$

The range of possible values is opposite to Goodall 1 thru Goodall 3 measures, i.e. from $2/(n(n-1))$ to 1. The other steps of computation follow Equations (2) and (3).

The Lin 1 measure has a more complex system of weights. Boriah et al., (2008) describes it the following way. *It gives lower weight to mismatches if either of the mismatching values are very frequent, or if there are several values that have frequency in between those of mismatching values. Higher weight is given when there are mismatches on infrequent values and there are a few other infrequent values. For matches, lower weight is given for matches on frequent categories or matches on values that have many other values of the same frequency. Higher weight is given to matches on rare values.*

Similarity between two objects by the *c*-th variable is treated according to the expression

$$S_c(x_{ic}, x_{jc}) = \begin{cases} \sum_{q \in Q} \ln p_c(q) \text{ if } x_{ic} = x_{jc} \\ 2 \ln \sum_{q \in Q} p_c(q) \text{ otherwise} \end{cases}, \tag{7}$$

$$Q \subseteq X_c : \forall q \in Q, p_c(x_{ic}) \le p_c(q) \le p_c(x_{jc}) \text{ assuming that } p_c(x_{ic}) \le p_c(x_{jc}),$$

where q represents a subset of categories in the c-th variable whose relative frequencies are in between relative frequency of $p(x_{ic})$ and $p(x_{jc})$. The range of possible values of the Lin 1 measure is from $-\ln(n/2)$ to 0.

The similarity measure across all variables is computed as

$$S(\mathbf{x}_i, \mathbf{x}_j) = \frac{\sum_{c=1}^{m} S_c(x_{ic}, x_{jc})}{\sum_{c=1}^{m} (\ln p(x_{ic}) + \ln p(x_{jc}))}. \tag{8}$$

Since the range of this similarity measure exceeds interval from zero to one, a different formula has to be used for a transformation into a dissimilarity measure:

$$D(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{S(\mathbf{x}_i, \mathbf{x}_j)} - 1. \tag{9}$$

All the above introduced similarity measures are further compared to a reference measure, the overlap measure (simple matching coefficient), which does not take any advantage from frequencies distribution in a dataset and simply assigns value 1 in case of match and 0 otherwise.

Criteria for cluster evaluation

Similarity measures can be used in variety of statistical tasks, such as clustering, classification or outlier detection. Each of these areas requires a different type of evaluation. The quality of similarity measures has to be measured indirectly, by means of quality of produced clusters. However, the quality can be measured by many ways. Some authors evaluate it from a point of view of their classification abilities, see (Morlini and Zani, 2012), or the way, how measures behave if a dataset contains outliers, see (Boriah et al., 2008). In this paper, the quality is evaluated from a point of view of their within-cluster variability. This approach corresponds the most to the original purpose of cluster analysis, which is not to classify objects into already known classes, but to reveal the original structure of the dataset by dividing it into several homogenous groups.

Measuring the within-cluster variability is based on a principle that with an increasing number of clusters provided by hierarchical cluster analysis, the within-cluster variability decreases. In this paper, it is measured by means of within-cluster mutability coefficient, which was introduced in (Řezanková et al., 2011), and which is based on the Gini coefficient:

$$WCM\,(k) = \sum_{g=1}^{k} \frac{n_g}{n \cdot m} \sum_{c=1}^{m} \frac{K_c}{K_c - 1} \left(1 - \sum_{u=1}^{K_c} \left(\frac{n_{gcu}}{n_g} \right)^2 \right), \tag{10}$$

where n is a number of objects, n_g is a number of objects in the g-th cluster ($g = 1, ..., k$), n_{gcu} is a number of variables in the g-th cluster by the c-th variable with the u-th category ($u = 1, ..., K_c$). This coefficient takes values from 0 to 1; the higher values indicate higher variability. The within-cluster mutability can be also measured by measures based on the entropy, see for instance (Šulc and Řezanková, 2014). Because these results are very similar to those provided by measures based on the Gini coefficient, evaluation criteria based on entropy are not used in this paper.

Since with an increasing number of clusters the WCM coefficient continually decreases, it is necessary to choose a statistics, which determines the optimal number of clusters. In (Řezanková et al., 2011) was introduced the pseudo F coefficient with the formula:

$$PSFG(k) = \frac{(n-k)(WCM\,(1) - WCM\,(k))}{(k-1)WCM\,(k)}, \tag{11}$$

where $WCM(1)$ is variability in the whole dataset, and $WCM(k)$ variability in the k-cluster solution. The highest value of this coefficient across all cluster solutions indicates the optimal one.

Experimental application

To evaluate the quality of similarity measures, which were introduced in Section 2, a particular method of hierarchical clustering must be set. With regards to the previous research, see (Šulc and Řezanková, 2014), the complete linkage method was chosen. To enable the direct comparison with the previous research, see (Šulc, 2014), the examined similarity measures are applied on the same datasets. Their overview is presented in Table 1.

Table 6: Description of analyzed datasets

	Car	Post-operative	Hayes Roth	Breast Cancer
number of objects (n)	1728	90	132	683
number of variables (m)	6	7	5	9
min. number of categories	3	2	3	9
max. number of categories	4	3	4	10

All datasets come from the UCI Machine Learning Repository, see (Bache and Lichman, 2013), where one can find their detailed description. The Car datasets was made artificially, for its objects contain all possible combinations of properties, each of them occurs exactly one time. This structure may be challenging for some similarity measures. The datasets Post-operative, Hayes Roth and Breast Cancer

represents three types of data structure complexity, from the simple one (Post-operative) to the complex one (Breast Cancer).

For each dataset, a set of proximity matrices was computed according to similarity measures introduced in Section 2. Further, hierarchical cluster analysis with the complete linkage for two to six cluster solution was performed on each proximity matrix. Each of these solutions was evaluated by the criteria introduced in Section 3. The computed values of the WCM coefficient are in Table 2. The bold values represent the optimal solutions for a given similarity measure and dataset, which were computed by the PSFG coefficient.

Table 2: Values of the WCM coefficient

Dataset	Measure	WCM(1)	WCM(2)	WCM(3)	WCM(4)	WCM(5)	WCM(6)
Car	Goodall 1	1.0000	0.9444	0.8889	**0.8333**	0.8194	0.8056
	Goodall 2	1.0000	0.9444	0.8889	**0.8333**	0.8194	0.8056
	Goodall 3	1.0000	0.9444	0.8889	**0.8333**	0.8194	0.8056
	Goodall 4	1.0000	0.9167	**0.8333**	0.8056	0.7778	0.7500
	Lin 1	1.0000	0.9444	0.8889	**0.8333**	0.8194	0.8056
	overlap	1.0000	**0.9405**	0.9005	0.8777	0.8454	0.8054
Post-operative	Goodall 1	0.7463	0.6986	0.6445	**0.5814**	0.5492	0.5090
	Goodall 2	0.7463	0.6973	0.6363	0.6022	0.5369	**0.4981**
	Goodall 3	0.7463	**0.6810**	0.5938	0.5431	0.4955	0.4739
	Goodall 4	0.7463	0.7106	0.6919	0.6803	0.6165	**0.5572**
	Lin 1	0.7463	**0.6768**	0.6225	0.5810	0.5456	0.5224
	overlap	0.7463	0.6962	0.6622	**0.6079**	0.5790	0.5503
Hayes Roth	Goodall 1	0.9404	0.8267	**0.7164**	0.6753	0.6145	0.5794
	Goodall 2	0.9404	**0.8376**	0.7616	0.7254	0.6999	0.6279
	Goodall 3	0.9404	0.8312	**0.7221**	0.6570	0.6218	0.5713
	Goodall 4	0.9404	0.8089	**0.6929**	0.6744	0.6611	0.6472
	Lin 1	0.9404	0.8659	0.8232	**0.7319**	0.6991	0.6837
	overlap	0.9404	0.8675	0.7629	0.7117	**0.6095**	0.5900
Breast Cancer	Goodall 1	0.7267	0.7203	0.6948	**0.5739**	0.5705	0.5628
	Goodall 2	0.7267	0.7126	**0.6800**	0.6735	0.6654	0.6550
	Goodall 3	0.7267	0.7203	0.7085	**0.5771**	0.5733	0.5652
	Goodall 4	0.7267	**0.6640**	0.6608	0.6545	0.6520	0.6500
	Lin 1	0.7267	**0.5874**	0.5796	0.5735	0.5689	0.5632
	overlap	0.7267	0.7170	0.6972	**0.5720**	0.5674	0.5616

The Car dataset contains the maximal possible variability of objects, which is expressed by value 1 of the WCM coefficient in the whole dataset (one-cluster solution). All the similarity measures, except for Goodall 4, are evaluated by the completely same scores of WCM. The Goodall 4 measure provides clusters with the highest variability decrease not even among measures presented in this dataset, but also among measures presented in (Šulc, 2014). When comparing the performance of presented similarity measures to the basic one, the overlap measure, they provide slightly better clusters, however, the difference is not very big. Most of measures prefer a four-cluster solution, Goodall 4 a three-cluster solution and overlap a two-cluster solution.

In the Post-operative dataset the clusters provided by Goodall 1, Goodall 2 and Goodall 3 are evaluated similarly from a point of view of WCM. When one take into account that low-cluster solution is usually preferred, the Goodall 3 results are the best. Clusters of Lin 1 measure have a very good score in the two-cluster solution. The worst results are provided by Goodall 4. For measures Goodall 3 and Lin 1 was determined a two-cluster solution, for Goodall 1 and overlap a four-cluster solution and for Goodall 2 and Goodall 4 a six-cluster solution.

In the Hayes Roth dataset, all the Goodall measures have good values of WCM, the best ones are provided by the Goodall 4 measure. The Lin 1 measure provides the worst clusters among all the measures. Whereas for Goodall 2 a two-cluster solution was determined, for Goodall 1, Goodall 3 and Lin 1 it was a three-cluster solution.

In the Breast Cancer dataset, Lin 1 has the best outcomes of WCM across most cluster solutions. Similarly as the original Lin measure, see (Lin, 1998), it provides good clusters if applied to a dataset with a more complex structure. From Goodall's measures, the best clusters were provided by Goodall 4, which is followed by Goodall 1, Goodall 3 and Goodall 2.

In comparison to other examined similarity measures in this paper and in (Šulc, 2014), the Goodall's adjusted measures provide much smaller decrease of within-cluster variability in the low-cluster solutions. That is the reason, why their optimal solution is usually with more than two clusters. The measures Goodall 1, Goodall 2 and Goodall 3 have very similar results, the Goodall 3 offers best performance of them. The Goodall 4 measure has a slightly different behavior; its low-cluster solutions have higher decrease of variability, but with increasing cluster solutions the decrease starts to slow down ominously. According to above mentioned facts, the best Goodall's measures are Goodall 3 and Goodall 4. Thus, it is obvious that the complex system of weights of Goodall 1 and Goodall 2 does not have any significant influence on cluster quality. Except for the Hayes Roth dataset, Lin 1 measure performs very well, especially in datasets with more complex structure. The values of WCM are very similar to the original Lin measure, see (Šulc, 2014); sometimes are slightly better and vice versa. Thus, there also arises the question if a complex system of weights of this measure has any significant importance. The overlap measure usually provides substandard results. It prefers cluster solutions with a higher number of clusters.

Conclusion

In this paper, five similarity measures were compared and evaluated from a point of view of their clustering performance in hierarchical cluster analysis with the complete linkage. Further, their results were compared with clustering based on a basic similarity measure, the overlap measure. The evaluation criterion was based on within-cluster variability represented by the WCM coefficient and determining the optimal cluster solutions by the PSFG coefficient. The examined similarity measures were applied on four well known datasets; it enables the possibility for other researchers to compare additional measures to those, which were presented in this paper.

There were introduced four different types of the original Goodall's measure. Although each of them has a different weight system, it is apparent that they all belong to one family of measures. The Goodall 3 and Goodall 4 measures have the best results among them. Moreover, these measures are not computationally expensive. Therefore, they are going to be included for additional examinations in future research. The Lin 1 measure demonstrated similar properties as its original measure; it provided good results, especially on datasets with more complex structure. This measure is going to be examined in further research as well.

References

AGGARWAL, C. (2013). *Data Clustering*. Hoboken: CRC Press.

BORIAH, S., CHANDOLA and V., KUMAR, V. (2008). Similarity measures for categorical data: A comparative evaluation. In: *Proceedings of the 8th SIAM International Conference on Data Mining, SIAM*, p. 243-254. Available at: http://www-users.cs.umn.edu/~sboriah/PDFs/BoriahBCK2008.pdf.

BACHE, K. and LICHMAN, M. (2013). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

ESKIN, E., ARNOLD, A., PRERAU, M., PORTNOY, L. and STOLFO, S., V. (2002). A geometric framework for unsupervised anomaly detection. In: *Applications of Data Mining in Computer Security*, p. 78–100.

DESAI, A., HIMANSHU, S. and VIKRAM, P. (2011). DISC: Data-intensive similarity measure for categorical data. In: *Advances in Knowledge Discovery and Data Mining*, p. 469-481.

GOODALL, V.D. (1966). A new similarity index based on probability. *Biometrics*. Vol. 22, No.4, p. 882.

LE, S. and HO, T. (2005). An association-based dissimilarity measure for categorical data. *Pattern Recognition Letters*, Vol 26, No 16, p. 2549-2557.

LIN, D. (1998). An information-theoretic definition of similarity. In: *ICML '98: Proceedings of the 15th International Conference on Machine Learning*. San Francisco. Morgan Kaufmann Publishers Inc., p. 296-304.

MORLINI, I. and ZANI, S. (2012). A new class of weighted similarity indices using polytomous variables. In: *Journal of Classification*, Vol 29, No 2, p. 199-226.

ŘEZANKOVÁ, H., LÖSTER and T., HÚSEK, D. (2011). Evaluation of categorical data clustering. In: *Advances in Intelligent Web Mastering – 3*. Berlin. Springer Verlag, p. 173-182.

SOKAL, R. (1963). The Principles and Practice of Numerical Taxonomy. *Taxon*, Vol 12, No 5, p. 190.

ŠULC, Z. (2014). Comparison of the new approaches in nominal data clustering (in Czech). In: *Sborník prací vědeckého semináře doktorského studia FIS VŠE*. Praha: Oeconomica, p. 214-223. Available at:

http://fis.vse.cz/wp-content/uploads/2014/05/DD_FIS_2014_CLANKY_CELY-SBORNIK.pdf.

ŠULC, Z. and ŘEZANKOVÁ, H. (2014). Evaluation of recent similarity measures for categorical data. In: *AMSE*. Wroclaw: Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, p. 249-258. Available at: http://www.amse.ue.wroc.pl/papers/Sulc,Rezankova.pdf.

JEL Classification: C38

Shrnutí

Použití Goodallových a Linových měr podobnosti v hierarchickém shlukování

Článek hodnotí kvalitu shlukování pěti měr podobnosti určených pro nominální proměnné, které byly poprvé představeny Boriah v roce 2008, a srovnává je s mírou overlap, která je dobře známá, a proto může sloužit jako referenční míra. Zkoumané míry byly odvozeny z již známých měr. Čtyři míry byly odvozeny z původní Goodallovy míry představené v roce 1966 a jedna z Linovy míry, která byla představena v roce 1998. Tyto míry byly původně vytvořeny pro účely detekce odlehlých pozorování v kategoriálních datech, ale jejich použití může být mnohem širší, např. shlukování nebo klasifikace. Bohužel, v těchto oblastech však nebyly nikdy prozkoumány. Cílem tohoto článku je prozkoumat kvalitu těchto měr z hlediska kvality shlukování objektů při použití shlukové analýzy s metodou nejvzdálenějšího souseda, která obvykle poskytuje nejlepší výsledky. Při analýze byl použit následující software: Matlab, IBM SPSS a MS Excel. Výsledné shluky jsou hodnoceny prostřednictvím koeficientu WCM založeného na Giniho koeficientu, který měří vnitroskupinovou variabilitu, a pseudo F koeficientu, který slouží k určení optimálního počtu shluků. Míry podobnosti jsou hodnoceny na čtyřech různých datových souborech, které pochází z UCI Machine Learning Repository. Tím je zajištěna možnost porovnání výsledků s dalšími výzkumníky. Výsledky výzkumu poukazují na některé míry podobnosti, které poskytují velmi dobré výsledky, a mohou být tak doporučeny k používání v hierarchické shlukové analýze s nominálními proměnnými.

Klíčová slova: míry podobnosti, hierarchické shlukování, nominální proměnné, Goodallova míra, Linova míra

Composite Indicators of Czech Business Cycle

Lenka Vraná

lenka.vrana@vse.cz

Doktorand oboru statistika

Školitel: doc. Ing. Markéta Arltová, Ph.D. (marketa.arltova@vse.cz)

Abstract: Composite indicators are recognized to be an eligible tool of the business cycle analysis, especially because they can be easily interpreted although they summarize multidimensional relationships between individual economic indicators.

These indicators can be divided into groups of leading, coincident and lagging ones with regard to the reference time series (usually GDP or industrial production index). The time series have to be seasonally adjusted and the trend component has to be removed. We use Hodrick-Prescott filter to find the trend and to smooth the cycle component.

There have been several attempts of constructing the composite indicators of the Czech business cycle. These analyses were based mostly on the comparison of cross correlations between the reference time series and the individual economic indexes. However, cross correlations should be used only in combination with other criteria such as the average lead/lag time between the turning points or number of extra and missing cycles.

This paper describes the history and methodology of composite indicators construction. It examines possibilities of Bry-Boschan non-parametric algorithm of the turning points detection on dating the Czech business cycle indicators. The results are compared with the simplified method that uses cross correlations only.

Key words: dating business cycle, turning points, composite indicators, cycle measurement

Introduction

In recent years due to the economic slowdown, the public attention was focused on the possibilities of forecasting the business cycle movements. One of the methods used for the business cycle analysis is based on the study of composite indicators which combine several individual economic indexes. The indexes can be divided into groups of leading, coincident and lagging ones with regard to the reference time series (usually gross domestic product (GDP) or industrial production index).

The leading composite indicator should be able to predict future states of economic activity – when will the economy switch from the expansion phase into the contraction phase or vice versa. The coincident indicator serves mainly to confirm the hypothesis about the state the economy is currently in. It also may replace GDP or industrial production index as the reference series for the evaluation of the individual indicators (this approach is used in the USA by the Conference Board). The lagging indicator should certify the cycle behavior and the dating of the turning points.

The construction of composite indicators usually follows methodology created by the Organization for Economic Co-operation and Development (OECD) or by the Conference Board. Most of the analyses of the Czech business cycle use OECD methodology which will be employed in this paper as well. For more information about the Conference Board methodology see its Business Cycle Indicators Handbook (2001) or Ozyildirim et al. (2010).

Czech economists who study composite indicators usually simplify the evaluation phase of OECD's algorithm. The goal of this paper is to describe OECD methodology and evaluate impacts of these simplifications on the forecasting ability of the composite indicators. Several papers focused on these topics have already been published; see Vraná (2014a, 2014b).

Composite business cycle indicators – history and methodology

In 1930's two American economists, Arthur Frank Burns and Wesley Clair Mitchell, worked on new methods how to measure business cycles and determine recessions. In 1946 they published work on

business cycle indicators which contained one of the first lists of leading, coincident and lagging indicators as well as set of instructions how to track the cycle.

Burns and Mitchell's methodology spread worldwide in the following years. It was executed manually and it required lots of personal judgment and therefore it wasn't quite objective.

In 1971 Gerhard Bry and Charlotte Boschan introduced their algorithm to automate the turning points detection. It was one of the first programmed approaches that were published and, with the fast development of information technologies, was than widely implemented.

OECD and other organizations still use Bry-Boschan algorithm with only slight changes. In the first proposal, Bry and Boschan used 12-month moving average, Spencer curve and short-term moving average of 3 to 6 month to detect the turning points. Nowadays none of these are necessary because some other techniques (like Hodrick-Prescott filter) are used to smooth the time series without shifting the turning points.

OECD methodology (Gyomai and Guidetti, 2012) consists of five steps: 1. pre-selection phase, which is passed only by long time series of indicators that have justified economic relationship with the reference series, broad coverage of economic activity and high frequency of observations, 2. filtering phase, when the time series are seasonally adjusted and de-trended, 3. evaluation phase, when only the best individual indicators with the strongest relationship with the reference series are selected to be included in the composite indicators, 4. aggregation phase, when the composite indicators are created and 5. presentation of the results.

Pre-selection

When constructing the composite indicators the eligible individual series has to be selected first. According to Gyomai and Guidetti (2012) only long time series of indicators that have justified economic relationship with the reference series, broad coverage of economic activity, high frequency of observations (preferable monthly), that were not subject to any significant revisions and that are published soon should pass the pre-selection phase.

The selection of individual indicators is highly dependent on the reference time series. Usually GDP or index of industrial production is used as reference series. GDP should respond to the cyclical movements better but it is quarterly statistics and it is necessary to convert it to the monthly estimates. OECD had used the industrial production index until March 2012 and then switched to the adjusted monthly GDP. We will use the index of industrial production because it shows strong co-movements with GDP series in the Czech Republic and it is available monthly.

Filtering

The second phase of the composite indicator construction is called the filtering. The main task of this stage is to decompose the individual time series. The series have to be seasonally adjusted and the trend component has to be removed.

Zarnowitz and Ozyildirim (2006) and Nilsson and Gyomai (2011) compare several methods how to estimate the cyclical component: phase average trend, Hodrick-Prescott filter, Christiano-Fitzgerald and Baxter-King band pass filters, etc. OECD has used Hodrick-Prescott filter to detrend the series since December 2008.

Hodrick-Prescott filter divides the series into two parts (τ_t – trend component and c_t – the cyclical component) and optimizes expression

$$\min_{\tau_t}\left[\sum_t\left(y_t-\tau_t\right)^2+\lambda\sum_t\left(\tau_{t+1}-2\tau_t+\tau_{t-1}\right)^2\right]. \tag{1}$$

It minimizes the difference between the trend and the original series and smooths the trend as much as possible at the same time. The λ parameter prioritize the latter from the two contradictory goals – the higher the λ, the smoother the trend.

Hodrick-Prescott filter deals with the series as with the system of sinusoid and it keeps in the trend only those with low frequency (high wave length). According to OECD the business cycles last 10 years at maximum therefore the fluctuations with lower wave length should be kept in the cycle component. Ravn and Uhlig (2002) recommend to set the λ parameter equal to 129 600 for monthly series. Nilsson and Gyomai (2011) use the Hodrick-Prescott filter twice: first with high λ to find the trend and then with low λ to smooth the cycle component. They also confirm that Hodrick-Prescott filter gives clear and steady turning point signals.

After the trend component is estimated, it is subtracted from the original data set (this is called the deviation cycle[56] then).

Evaluation

After the cycle components of all the individual indicators are found, turning points are detected. Not every peak or trough of the cycle is considered as the turning point though. Bry-Boschan algorithm (Bry and Boschan, 1971) is used to determine the turning point.

The cycle components of all the individual indicators are compared to the reference series. OECD uses several methods how to evaluate their relationship: the average lead (lag) times between the turning points, cross correlations and number of extra and missing cycles. Then the selected individual indicators are divided into groups of leading, coincident and lagging ones and the composite indicators are created.

OECD use Bry-Boschan algorithm to detect the turning points of the time series. Goodwin (1993), Diebold and Rudebusch (1996) and Levanon (2010) also mention regime switching models as alternative to estimate the phases of the cycle. Bruno and Otranto (2004) provides comprehensive overview of several parametric and non-parametric methods for turning points detection.

Aggregation and presentation of the results

The last two phases of OECD methodology are not described in this paper because they don't relate closely with the topic. For more information about the rest of the process see Gyomai and Guidetti (2012).

Czech composite business cycle indicators

Besides OECD there are several other authors who study composite indicators of the Czech economy. These authors usually simplify the evaluation phase of the computational algorithm. It has become common practice to use only cross correlations to test the relationship between individual indicators and the reference series as correlations are easy to compute; see Czesaný and Jeřábková (2009), Pošta and Valenta (2011), Svatoň (2011) or Tkáčová (2012). However, it is highly recommended to use the cross correlations only in the combination with other criteria such as analysis of the turning points (Gyomai and Guidetti, 2012).

The cross correlations measure the linear dependency between the reference series and individual indicator with applied time-lag. Then the maximum of absolute value of cross correlations is found and the individual indicator is included into one of the composite indicators (leading, coincident or lagging) according to the time-lag where the maximal value appeared.

The advantage of cross correlations is that they are easy to compute and that their results can be quickly evaluated. However, cross correlations measure only general fit between individual indicators and the reference time series. The goal of the composite indicators is nevertheless to predict the turning points of the reference time series (when will the economy switch from the expansion phase into the contraction phase or vice versa) not to forecast the level of economy. This can cause problems, because the indicators can show strong linear dependency on the reference time series and still miss some of the cycles, contain extra cycles or the lead (lag) times of their turning points can significantly fluctuate.

[56] Cycles in smoothed growth rates can be used instead of deviation cycles; for more information see Mintz (1969).

Pošta and Valenta (2011) explain the choice of the cross correlations by the lack of available data. As they analyze only the data from business surveys, which start in January 2003, their decision is understandable. The other authors mentioned above do not provide any reasons why not to analyze also the timing of the turning points.

It can be assumed that ignoring the turning points of the time series means loss of information which is essential for the predictive capabilities of composite indicators. We will show in this paper that cross correlations give different results than the turning points analysis and also provides less details about the relationship between the reference series and individual indicators.

Application of Bry-Boschan algorithm

Data

Bry-Boschan algorithm for turning points detection will be illustrated with the cycle analysis of selected economic indicators. We analyze dataset of 83 individual economic indicators which are available from January 2002 to August 2012. Index of industrial production is set as the reference time series.

Cross correlations were used as a pre-selection method: only individual indexes with maximal absolute value of cross correlations higher than 0.6, which occurred when the individual index was shifted forward in time, would be dated by Bry-Boschan algorithm. Only 10 out of 83 individual indexes passed the criteria (tab. 1); for more details see Vraná (2013a). As Czech economy is small and open, several selected indexes are indicators of German economic activity.

Cycle Chronology

We demonstrate the Bry-Boschan algorithm in detail on three selected time series: index of industrial production (reference series), German business expectations indicator and industrial producer price index – installation of industrial machinery and equipment (fig. 1).

Tab. 1: Cross correlations and average lead (-) or lag (+) time of 10 selected individual indicators.

Economic indicator	Cross correlations		Average lead/lag time of turning points
	Max. value	Lead time	
New orders in industry	0,70	-1	1,83
New orders in industry - Manufacture of electrical equipment	0,73	-1	0,00
Industrial producer price index - Electrical equipment	0,79	-2	-2,50
Industrial producer price index - Installation of industrial machinery and equipment	-0,77	-5	-4,00
Composite confidence indicator	0,88	-1	-0,43
Business confidence indicator	0,90	-1	0,00
IFO (DE) - R 1 : Business climate	0,86	-3	-4,00
IFO (DE) - R 2 : Business situation	0,89	-1	-2,50
IFO (DE) - R 3 : Business expectations	0,71	-5	-4,67
Stock market index PX	0,83	-3	4,50

Over the period from January 2002 to August 2012 the Bry-Boschan algorithm found three cycles measured from peak to peak in IIP time series. The average length of the cycle phase is 17.3 months. When working with growth cycles, the phase of expansion is usually longer than the phase of the contraction (Goodwin, 1993). This is obvious also in our data: the average duration of the expansion phase is 22.7 months and the average duration of the contraction phase is 12.0 months.

The series of German business expectations includes double turn: peaks in April 2006 and in May 2007. However, the minimum cycle duration is 15 months, so the lower of these two peaks (April 2006) had to be eliminated. The trough in October 2006 couldn't be accepted either, because the turns have to alternate. IFO Business expectation is indicator with highest average lead time when compared to the reference series (lead of 4.67 months).

The industrial producer price index – installation of industrial machinery and equipment shows strong countercyclical behavior. It tends to increase when the economy is slowing down and vice versa. This is indicated also by the negative value of cross correlation. Countercyclical indexes can be included in

the composite indicators, but they have to be inverted first. However, this price index skips some of the cycles of IIP.

The other individual indicators don't miss any cycles indicated by IIP, but index of new orders in industry - manufacture of electrical equipment contains one extra cycle.

The turning point with the lowest spread among all the 10 individual indicators was the trough in May 2009 (measured in IIP) – this means that the switch between the cycle phases happened almost together in all the series. The turning point with the highest spread was the peak in May 2004 (in IIP).

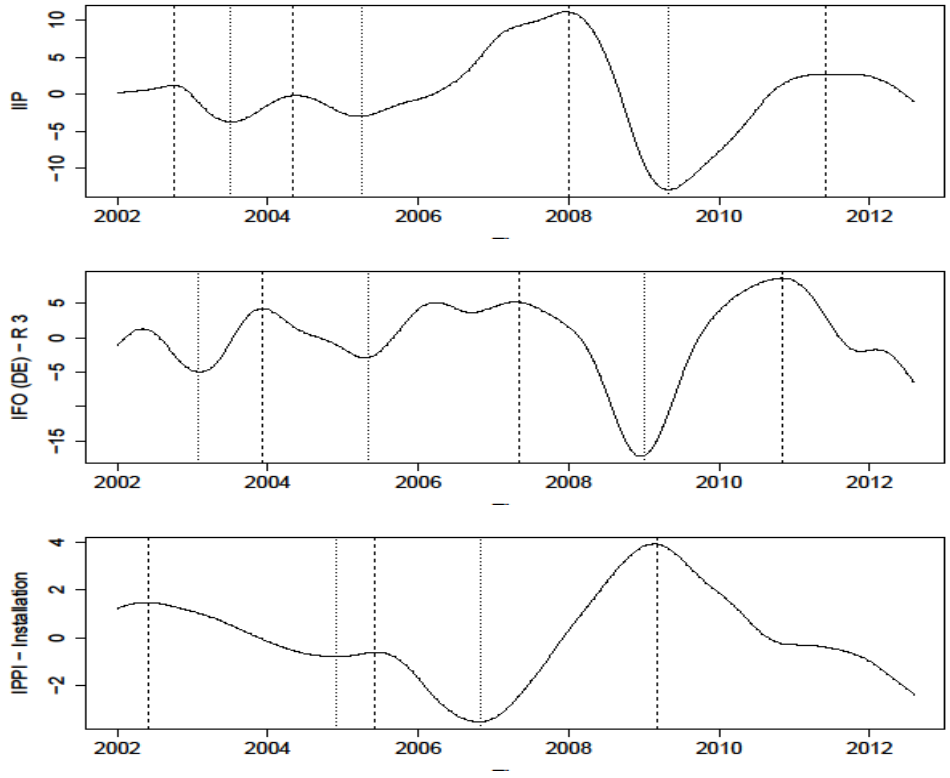


Fig. 1: Cycle components of index of industrial production, German business expectations indicator and industrial producer price index – installation of industrial machinery and equipment with emphasized peaks and troughs, which were detected by Bry-Boschan algorithm.

The average lead/lag times of all analyzed indicators (tab. 1) prove, that the cross correlations alone cannot determine, which time series should be included in leading, coincident or lagging composite indicators. Cross correlations measure only general fit between individual indicators and the reference time series not the relationship between the turning points, which is essential for the right selection.

Only 6 out of 10 indicators (that were indicated as leading ones by cross correlations) show the lead in turning points: both industrial producer price indexes, composite confidence indicator and the three IFO indicators. The other indicators behave more like coincident or lagging ones when their turning points are considered.

Composite leading indicator

We proved that not all the individual economic indicators with strong correlation with reference time series show also the corresponding shifts in turning points. We can also compare the performance of composite leading indicator (CLI) that was constructed based on cross correlations only with CLI that uses also turning points analysis. We can assume that the latter one will predict the turning points of the whole economy (measured here by IIP) better than the first one, although it was the primary goal of both of them.

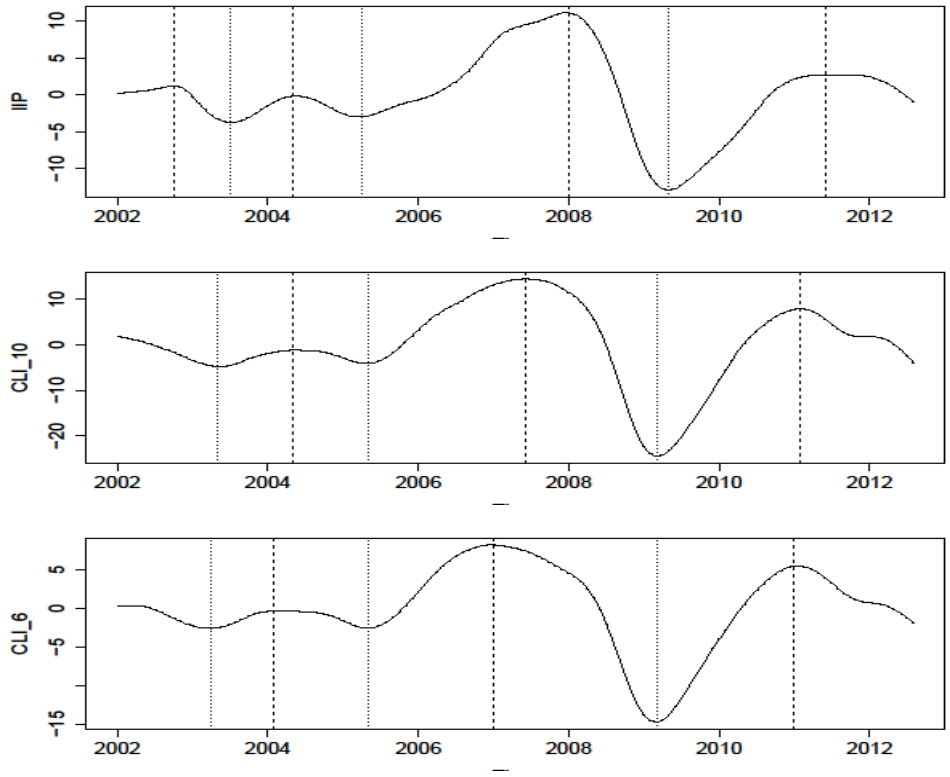


Fig. 2: Cycle component of index of industrial production and composite leading indicators with emphasized peaks and troughs, which were detected by Bry-Boschan algorithm.

The first composite indicator (CLI_10) includes all 10 economic indicators, which showed maximal absolute value of cross correlations higher than 0.6 when shifted forward in time. The second one (CLI_6) contains only a subselection of the discussed indicators – only those with positive average lead time of the turning points.

Both constructed composite indicators shows significant lead when compared with IIP cycle. The average lead time (tab. 2) is equal to 2.3 and 4.0 months for CLI_10 and CLI_6, respectively.

Tab. 2: Dates of the turning points of index of industrial production cycle and composite leading indicators and leads (-), coincidences (0) and lags (+) in months relative to IIP cycle turns.

Peak/Trough	Turning point dates			Leads, coincidences and lags	
	IIP	CLI_10	CLI_6	CLI_10	CLI_6
P	Oct-02	-	-	-	-
T	Jul-03	May-03	Apr-03	-2	-3
P	May-04	May-04	Feb-04	0	-3
T	Apr-05	May-05	May-05	+1	+1
P	Jan-08	Jun-07	Jan-07	-7	-12
T	May-09	Mar-09	Mar-09	-2	-2
P	Jun-11	Feb-11	Jan-11	-4	-5

The indicators don't miss any of the cycles (except the first peak, which is too close to the beginning of the time series). They manage to indicate all the cycle turning points in advance, except the trough in April 2005 – the troughs of both CLIs were lagged by one month. The lead/lag times of the turning points are rather unstable: from -7 to +1 month for CLI_10 and even more erratic for CLI_6 (from -12 to +1 month). Both indicators were able to predict the peak in January 2008 and the following slowdown of the economy, which can be linked to the global financial crisis.

We can use also cross correlations to assess the linear relationship between the IIP cycle and the composite indicators. The maximal cross correlation (0.97) of CLI_10 appears when the indicator is

shifted 3 month forward in time. The maximal cross correlation of CLI_6 is slightly lower (0.95), but it occurs with 4-month lead time. This means that there is strong general fit between the values of IIP cycle and both constructed composite indicators. However, the ability to predict the turning points between expansion and contraction phases of economy remains the key asset of leading composite indicators.

Conclusions

This paper focused mainly on the evaluation phase of the OECD methodology of composite indicators construction. In the evaluation phase cyclical components of all evaluated individual indicators are compared to the reference series. Their relationship can be described by several methods: the average lead (lag) time between the turning points, cross correlations and number of extra and missing cycles. The selected individual indicators are then divided into groups of leading, coincident and lagging ones and the composite indicators are created.

In the Czech Republic it has become common practice to use only cross correlations to test the relationship between individual indicators and the reference series. This paper compared results of the cross correlation analysis with the Bry-Boschan algorithm for tracking the turning points of the cycle component. It has proved that cross correlations gave different results than the turning points analysis and also provided less information about the relationship between the reference series and the individual indicators.

References

BRUNO, G., & OTRANTO, E. *Dating the Italian business cycle: a comparison of procedures.* Istituto di Studi e Analisi Economica, 2004. Working Paper, 41, 25.

BRY, G., & BOSCHAN, C. *Cyclical analysis of time series: selected procedures and computer programs*. National Bureau of Economic Research, 1971. 216.

BURNS, A. F., & MITCHELL, W. C. *Measuring Business Cycles*. NBER, 1946. Retrieved from http://papers.nber.org/books/burn46-1

CZECH STATISTICAL OFFICE. *Business Cycle Surveys - Methodology*. Retrieved 5 25, 2013, from http://www.czso.cz/eng/redakce.nsf/i/business_cycle_surveys

CZESANÝ, S., & JEŘÁBKOVÁ, Z. Kompozitní indikátory hospodářského cyklu české ekonomiky. Statistika(3), 2009. 256-274.

DIEBOLD, F. X., & RUDEBUSCH, G. D. *Measuring Business Cycles: A Modern Perspective*. Review of Economics and Statistics, 1996. 78, pp. 67-77.

GOODWIN, T. H. *Business-Cycle Analysis with a Markov-Switching Model*. Journal of Business & Economic Statistics, 1993. 11(3), pp. 331-339.

GYOMAI, G., & GUIDETTI, E. *OECD System of Composite Leading Indicators*. OECD Publishing, 2012. 18.

LEVANON, G. *Evaluating and comparing leading and coincident economic indicators*. Business Economics, 2010. 45(1), 16-27.

MINTZ, I. *Dating Postwar Business Cycles: Methods and Their Application to Western Germany, 1950-1967*. New York: National Bureau of Economic Research, 1969.

NILSSON, R., & GYOMAI, G. (2011). *Cycle extraction: A comparison of the Phase-Average Trend method, the Hodrick-Prescott and Christiano-Fitzgerald filters*. OECD Publishing, 23.

OZYILDIRIM, A., SCHAITKIN, B., & ZARNOWITZ, V. (2010). *Business cycles in the euro area defined with coincident economic indicators and predicted with leading economic indicators*. Journal of Forecasting, 29(1-2), 6-28.

POŠTA, V., & VALENTA, V. *Composite Leading Indicators Based on Business Surveys: Case of the Czech Economy*. Statistika, 2011. 48(1), pp. 12-18.

RAVN, M., & UHLIG, H. *On adjusting the Hodrick-Prescott filter for the frequency of observations*. Review of Economics and Statistics, 2002. 84(2), 371-376.

SVATOŇ, P. *Composite Leading Indicators: A Contribution to the Analysis of the Czech Business Cycle*. Prague: Ministry of Finance of the Czech Republic, 2011.

THE CONFERENCE BOARD. *Business Cycle Indicators Handbook*. New York: 2001. Retrieved from https://www.conference-board.org/publications/publicationdetail.cfm?publicationid=852

TKÁČOVÁ, A. *Kompozitný predstihový indikátor hospodárskeho cyklu českej ekonomiky*. Politická ekonomie(5), 2012. 590-611.

VRANÁ, L. *Alternative to the Construction of Czech Composite Indicators*. International Days of Statistics and Economics (pp. 1502-1511). Slaný: Melandrium, 2013.

VRANÁ, L. Business Cycle Analysis: *Comparison of Approaches to Cycle Measurement*. Mathematical Methods in Economics 2013 (pp. 1016-1021). Jihlava: College of Polytechnics Jihlava, 2013.

VRANÁ, L. Business Cycle Analysis: *Tracking Turning Points*. AMSE (pp. 277-283). Wroclaw: Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, 2014.

VRANÁ, L. *Czech Business Cycle Chronology*. The 8th International Days of Statistics and Economics (pp. 1623-1632). Slaný: Melandrium, 2014.

ZARNOWITZ, V., & OZYILDIRIM, A. *Time series decomposition and measurement of business cycles, trends and growth cycles*. Journal of Monetary Economics, 2006. 53(7), 1717-1739.

JEL Classification: C32, E32

Shrnutí

Kompozitní ukazatele českého hospodářského cyklu

Tento článek se zabývá využitím kompozitních ukazatelů pro analýzu cyklu, popisuje historii jejich vzniku a metodiku konstrukce.

O sestavení kompozitních ukazatelů českého hospodářského cyklu se již v minulosti pokoušelo několik autorů. Jejich analýzy využívaly pro zhodnocení vztahu mezi referenční řadou a jednotlivými ekonomickými indexy zejména křížové korelace, které jsou výpočetně nenáročné. Křížové korelace by však nikdy neměly být používány samostatně, ale pouze v kombinaci s dalšími metodami, např. průměrným časem předstihu/zpoždění mezi body obratu nebo porovnáním počtu cyklů.

Tento článek zkoumá možnosti aplikace neparametrického algoritmu Brye a Boschanové, který vyhledává body obratu, na česká data. Článek porovnává výkonost kompozitních ukazatelů českého hospodářského cyklu, pro jejichž konstrukci byly využity jen křížové korelace, s těmi, které využívají také analýzu bodů obratu.

Klíčová slova: datování hospodářského cyklu, body obratu, kompozitní ukazatele, analýza cyklu