Univerzita Karlova v Praze
Filozofická fakulta

Bakalářská práce

2009 **Jindřich Mynarz**

Univerzita Karlova v Praze
Filozofická fakulta
Ústav informačních studií a knihovnictví
Studijní program: informační studia a knihovnictví
Studijní obor: informační studia a knihovnictví

Bakalářská práce
Jindřich Mynarz

Metody automatizovaného filtrování informací v prostředí internetu

Methods for automated information filtering on the internet

Praha, 2009.

Vedoucí práce: PhDr. Petra Sluková

Vedoucí bakalářské práce: PhDr. Petra Sluková
Oponent bakalářské práce:
Datum obhajoby:
Hodnocení:

Prohlášení

Prohlašuji, že jsem bakalářskou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje.

V Praze, 20.5. 2009

podpis studenta

Identifikační záznam:

MYNARZ, Jindřich. *Metody automatizovaného filtrování informací v prostředí internetu [Methods for automated information filtering on the internet]*. Praha, 2009. vii, 57 s. Bakalářská práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví. Vedoucí bakalářské práce PhDr. Petra Sluková.

Abstrakt:

Bakalářská práce nabízí přehled metod filtrování informací, které je možné využít v prostředí internetu. Tyto metody slouží k automatické klasifikaci obsahu na žádoucí a nežádoucí. Představena jsou východiska, z nichž filtrování vychází. K nim patří specifika prostředí internetu a situace přehlcení informacemi. Popsány jsou typy dat a postupy, s nimiž metody filtrování informací pracují. Stěžejním obsahem je prezentace jednotlivých metod objasňující jejich funkci a účinek. Zváženy jsou také požadavky na aplikace těchto metod a jejich účinnost.

Abstract:

The aim of this work is to discuss the available methods for information filtering which can be used on the internet. These methods are used for automatic content classification to distinguish between desirable and undesirable items. This work introduces the specifics of internet and the problem of information overload which makes information filtering necessary. It describes the items and core processes that are involved in information filtering. The main attention is focused on the presentation of the particular methods. Dealing with their function and effect it also provides requierements and recommendations for the applications of these methods.

Klíčová slova:

filtrování informací, internetové filtry, kolaborativní filtrování, antispam

Obsah

Předmluva

Metody automatizovaného filtrování informací představují nástroje, které pomáhají při výběru z velkého množství informací. Jejich cílem je, aby uživatel získal přístup k žádoucím informacím, a aby nežádoucí informace nezískaly přístup k uživateli. Na rozdíl od vyhledávání informací, které představuje aktivní způsob získávání informací, je filtrování pasivním způsobem získávání informací. Prostředí internetu poskytuje velké množství informací a zároveň funguje na sdíleném technologickém základě, takže je pro aplikaci automatizovaných metod filtrování informací výbornou doménou.

Toto téma mne zaujalo, jelikož jde o technologii pomáhající lidem automatizací jejich činností. Filtrování informací je navíc zajímavé tím, že se kromě problému jak získat přístup k relevantním informacím, věnuje také tomu, jak se informací efektivně zbavovat.

Bakalářská práce obsahuje oproti svému zadání kapitolu *Základy*, která vznikla osamostatněním obecných postupů, jež využívají metody popsané v následujících kapitolách. Z důvodu změny použité terminologie dělení metod došlo také k přejmenování 3. kapitoly na *Filtrování založené na atributech*.

Tento text začíná úvodní kapitolou, která popisuje filtrování informací na obecné úrovni. Druhá část se věnuje základním prvkům a postupům, které metody filtrování informací využívají. Tím jsou stanovena východiska, na nichž staví následující tři kapitoly popisující hlavní metody, které se při filtrování informací uplatňují. Práce je zakončena závěrem, který navrhuje základní požadavky, jimž by systémy automatizovaného filtrování měly vyhovět.

Oblast filtrování informací je nesmírně rozsáhlá a aktivní. Velký podíl této problematiky patří do informatických a matematických oborů. Tato práce je však orientována na obecné znaky a postupy filtrování informací, a proto je pro první dvě kapitoly zabývající se obecnými aspekty a funkčními principy určeno nejvíce místa.

Metody filtrování informací představují aplikaci informačních a komunikačních technologií, díky čemuž je značná část používané terminologie zavedena pouze v angličtině. Proto tato práce respektuje zavedené anglické termíny, které jsou mezinárodně používány.

Jelikož téma této práce zahrnuje internet, čerpá významně z on-line informačních zdrojů. Velkou část citovaných materiálů tvoří příspěvky z konferencí a odborných periodik. Neocenitelnými zdroji pro práci byly databáze a digitální úložiště jako *ACM Digital Library* nebo *SpringerLink*. Naprostá většina použitých zdrojů je napsána anglicky, přestože angličtina nebývá vždy mateřským jazykem autorů.

Bibliografické záznamy uvedené v této práci jsou zpracovány v souladu s citačními normami *ČSN ISO 690 Dokumentace – Bibliografické citace – Obsah, forma a struktura*[1] a *ČSN ISO 690-2 Informace a dokumentace – Bibliografické citace – Elektronické dokumenty nebo jejich části*.[2] Uvnitř textu jsou citace uváděny v kulatých závorkách pomocí prvního údaje a roku vydání.

Děkuji vedoucí práce PhDr. Petře Slukové za konzultování obsahu a struktury.

1ČSN ISO 690 (01 0197). *Dokumentace – Bibliografické citace – Obsah, forma a struktura.* Praha : Český normalizační institut, 1996. 31 s.

2ČSN ISO 690-2 (01 0197). *Informace a dokumentace – Bibliografické citace – Část 2: Elektronické dokumenty nebo jejich části.* Praha : Český normalizační institut, 2000. 22 s.

1 Úvod

1.1 Vymezení

Filtrování informací spočívá v jednoduché klasifikaci na informace žádoucí a informace nežádoucí. Tato rozhodnutí dokáží nejlépe provádět lidé, avšak kvůli neustále narůstajícímu množství informací je zapotřebí je automatizovat.

Rozhodnutí o tom, zdali je informace žádoucí či nežádoucí, záleží na více předpokladech. Může být založeno na objektivně stanovitelném obsahu informace, uživateli informace a jeho vkusu (resp. informačním potřebám), a kontextu, v němž je informace komunikována.
Filtrování informací předchází získání přístupu k informačním zdrojům. Poté je následováno zobrazením informací.

1. Prvním krokem, který tvoří nutný předpoklad pro filtrování informací, je *sběr informačních zdrojů*. Informační zdroje jsou entity obsahující informace ve formě, kterou je uživatel schopen interpretovat. Vlastní filtrování informací se nezabývá tím, jak k informačním zdrojům získat přístup. To spadá do oblasti vyhledávání informací. Filtrování informací proto začíná až tehdy, když je dosaženo přístupu k informacím.
2. Filtrování informací slouží k *výběru (selekci) z informačních zdrojů*. Buď jsou vybrány informace žádoucí nebo jsou odstraněny informace nežádoucí.
3. Po provedení výběru informací je na řadě jejich *zobrazení*. Informace, které prošly krokem filtrování, jsou prezentovány uživateli.

Filtrování informací spadá mezi procesy zpracování informací. Ty lze obecně charakterizovat pomocí informační potřeby, kterou se snaží uspokojit, a zdrojů informací, z nichž pro tento úkol čerpají.

Tabulka 1.1: Srovnání procesů zpracování informací (převzato z Oard, 1996)

Proces	Informační potřeba	Zdroje informací
Filtrování informací	stálá a specifická	dynamické a nestrukturované
Vyhledávání informací	dynamická a specifická	stálé a nestrukturované
Přístup k databázím	dynamická a specifická	stálé a strukturované
Extrakce informací	specifická	nestrukturované
Upozorňování	stálá a specifická	dynamické
Prohlížení	široká	nespecifikované
Zábava	nespecifikovaná	nespecifikované

1.1.1 Informační zdroje

Filtrování informací je určeno pro uspokojení relativně stálých a konkrétních informačních potřeb z měnících se informačních zdrojů. Tyto zdroje jsou většinou nestrukturované a obtížně strojově zpracovatelné. To znamená, že jde např. o texty v přirozeném jazyce nebo obrazové a zvukové informace.

Vyhledávání informací je naproti tomu zaměřeno na poskytování informací pro měnící se potřeby z relativně neměnných informačních zdrojů.
Při vyhledávání informací se mění dotaz (informační potřeba), ale kolekce dokumentů, která je prohledávána, zůstává relativně stabilní. Při filtrování informací je to naopak. Dotaz zůstává delší dobu stejný, ale „proud“zpracovávaných dat se rychle mění.
Pro některé oblasti se nehodí ani filtrování, ani vyhledávání; protože jak zdrojové informace, tak i uživatelské potřeby, se rychle mění. Příkladem můžou být burzovní informace.

1.1.2 Informační potřeby

Po informační zdrojích je druhým hlavním vstupem filtrování informací informační potřeba.

Informační potřebu vyjádřenou pro systém filtrování informací můžeme nazvat *pokračujícím dotazem* (continuous query). Filtrování informací nejlépe odpovídá na relativně stabilní informační potřeby, které lze reprezentovat například pomocí uživatelských profilů.
Změny informačních potřeb mohou probíhat rychle, pokud jsou zapříčiněny určitou událostí, nebo pomalu, kdy je jejich příčinou déle trvající proces. Změny můžou být náhlé, postupné nebo periodické. Pomalou změnu může způsobit například stárnutí, rychlou změnu nálada.
Na systém filtrování informací je proto kladen požadavek do jisté míry se přizpůsobovat změnám informačních potřeb. Zpětná vazba, kterou poskytne uživatel systému, může posloužit k řešení pomalých změn; umožnění manuálních úprav uživatelských profilů k řešení náhlých změn.

1.1.3 Cíl

Cílem filtrování informací je zpracovávat rozsáhlé objemy dynamicky tvořených informací tak, že jsou uživateli prezentovány jen ty informace, které pravděpodobně uspokojí jeho informační potřebu.

Na každoroční konferenci *TREC* (*T*ext *RE*trieval *C*onference) je úloha informačního filtrování je pojmenována jako *„směrování“* (routing). V tomto smyslu je účelem filtrování informací nasměrovat k uživateli jen ty informace, které jsou pro něj žádoucí.
Umberto Eco nazývá filtrování informací *„uměním decimace“* (Coppock, 1995). Filtr přirovnává k dveřníkovi, který k uživateli pouští pouze žádoucí informace.

1.1.4 Příbuzné obory

Mnoho postupů si filtrování informací vypůjčuje z příbuzných disciplín jako je vyhledávání informací, modelování uživatelů nebo strojového učení. Jedná se například o způsoby reprezentace dokumentů, měření jejich podobnosti nebo booleovský výběr.

Původ filtrování informací lze najít ve službách *adresního rozšiřování informací* (ARI)[3]. Jde o služby upozornění, které se uplatňovaly převážně v minulosti ve výzkumných a speciálních knihovnách.

3*selective dissemination of information* (SDI)

1.2 Komunikace informací v prostředí internetu

Standardní algoritmy pro vyhledávání informací jsou navrženy pro relativně malé a koherentní kolekce informací jako jsou články v časopisech nebo knihovní katalogy. Web byl naproti tomu označen jako *„prolinkovaná anarchie“*(García-Barriocanal, 2005). Je výrazně dynamické, proměnlivé povahy, což z něj činí optimální informační prostor pro aplikaci metod filtrování informací.

Internet byl původně navržen, aby tvořil komunikační síť, která bude fungovat i po jaderném útoku. Od toho se odvozuje jeho svobodomyslná (někdy až anarchistická) povaha, která považuje každou regulaci za poškození, jež se systém musí pokusit napravit.
Zákony, které se o takovou regulaci pokoušejí, jsou však otupeným nástrojem pro kontrolu obsahu internetu. Proto potřebují technologické prostředky.

1.2.1 Rysy informační společnosti

Současná společnost je založena na informacích, jejichž množství neustále narůstá.

S rozvojem společnosti lze pozorovat trend nárůstu „volného času“, tj. času, který není zapotřebí věnovat činnostem zajišťujícím bezprostřední přežití. Věda patří k činnostem, jejichž cílem je dlouhodobé přežití. Nárůst množství volného času umožňuje nárůst objemu vědeckého výzkumu.
Spolu s exponenciálním růstem znalostí, exponenciálně roste počet otázek, které s novými poznatky souvisí. Za každou odpovědí jsou dvě otázky (Huleatt, 2009b).
Stejně jako objem vědecko-technických informací roste objem informací, která produkují média. Věda a média tak představují hlavní tvůrce informací.
Ještě více se však díky rozvoji komunikačních technologií zvětšuje počet způsobů komunikace těchto informací. Díky tomu, že je více komunikace než informací, vzniká obrovská masa bezobsažné komunikace.
Na základě těchto podmínek vzniká u občanů informační společnosti stav informačního zahlcení.

1.2.1.1 Přehlcení informacemi

Informační zahlcení je stav, kdy jedinec má k dispozici tolik informací, že na nich není schopen založit akci. Spousta informací pro jedince není relevantní a nemůže s nimi nic dělat (Good, 2006). Lidé nejsou schopni se efektivně rozhodnout, kterým informacím mají věnovat pozornost.

Avšak ne všichni se shodují na tom, že informační přehlcení existuje a že je to nový problém. Clay Shirky tvrdí, že *„to není přehlcení informacemi: je to selhání filtrů.“* (Shirky, 2008). Podle něj není stav informačního přehlcení žádnou novinkou – je stavem přirozeným přinejmenším od doby vzniku knihtisku, kdy měl jedinec poprvé k dispozici více informací, než byl schopen zpracovat. Tedy podle Shirkyho trvá přehlcení informacemi nejméně od 15. století. V jednom interview (Juskalian, 2008) Shirky posunul vznik přehlcení informacemi až do doby existence Alexandrijské knihovny, která představuje první archeologický doklad toho, že bylo na jednom místě nashromážděno více informací než mohl jeden člověk za celý život přečíst.
Novinku však Shirky spatřuje v selhání filtrů informací, které až donedávna omezovaly

negativní následky informačního přehlcení. Dříve roli filtrů informací zastávali vydavatelé, kterým záleželo, aby na trhu byly pouze prodejné (a tedy relativně kvalitní) publikace (Assay, 2009). S tím jak publikační model v prostředí internetu snížil vliv vydavatelů, se přesunula nutnost filtrování informací až na koncové uživatele. V minulosti bylo filtrování prováděno u zdrojů informací, nyní se přesunulo spíše na jejich příjemce.
Dochází tak k decentralizaci filtrování a hodnocení informací, což je podle Shirkyho zcela oprávněný vývoj. Proto lze pozorovat výrazný příklon k systémům sociálního filtrování, které umožňují efektivní využívání silných stránek tohoto modelu. Vzhledem k množství informací nelze v prostředí internetu předpokládat efektivitu centralizovaného hodnocení a klasifikace obsahu, které obstarávají například profesionální katalogizátoři. Avšak jde o rozhodnutí mezi tím, zdali lze filtrování svěřit davu (crowd-sourcing) anebo důvěřovaným odborníkům.
Proti názoru Claye Shirkyho se vyslovil David Crotty (Crotty, 2009a), který se Shirkym souhlasí v případě mediálních informací, avšak odmítá neexistenci informačního přehlcení v oblasti vědy. S tím, jak více lidí dělá ve výzkumu, prudce vzrostl počet publikovaných časopisů. To však neznamená, že jsou publikovány nekvalitní informace, které by správně měly být odfiltrovány. Naopak, model současné společnosti je natolik pokročilý, že dovoluje provádět takový objem kvalitního vědeckého výzkumu, který byl v předchozích obdobích nemyslitelný. Problém většiny vědců není v tom, že by nenacházeli dostatek kvalitních informací. Mají své zavedené kmenové časopisy, které i tak nestíhají číst (Crotty, 2009b). Pro ně je informační přehlcení velice reálné.

1.2.1.2 Ekonomie pozornosti

Spolu s informačním přehlcením bezprostředně souvisí oblast *ekonomie pozornosti*. Každý jedinec má k dispozici omezené množství pozornosti, které může věnovat zvoleným informacím. Přebytek informací však vytváří nedostatek pozornosti. Informační filtrování slouží k *efektivní alokaci pozornosti*.

Nedostatek pozornosti vyvolává tlak na kvalitu systémů filtrování informací. Nesprávně filtrované informace jsou pro uživatele důvodem okamžitě přesunout svou pozornost jinam. Pokud se uživatel setká s nesprávně doporučenou informací, ihned se obrací k jinému zprostředkovateli, protože díky přebytku informací nemá co ztratit (Iskold, 2007b).

1.2.2 Model publikování v prostředí internetu

Internet představuje jednoduchý způsob, jakým lze informace distribuovat. Umožňuje publikovat bez prostředníků a takřka zdarma. Oproti ostatním komunikačním médiím má velice nízké vstupní bariéry.

- takřka nulové vstupní náklady
- dostupnost připojení - kvalitní síťová infrastruktura zajišťuje dobré pokrytí, nízká cena

snadnost publikování - systémy pro správu obsahu, záměrná jednoduchost HTML, hostování webových stránek zdarma

- bezprostřednost - k publikování není třeba prostředníka, jehož roli v tradičních prostředích zastává vydavatel

Na internetu je snadné publikovat nejen pro autory webových stránek, ale také pro jejich

uživatele. Velkou část informací na webu představuje *obsah vytvářený uživateli* (user-generated content). Pro stránky, které umožňují svým uživatelům publikovat, bylo dokonce navrženo doporučení *UGC Principles*, z něhož vyplývá nutnost implementovat filtrovací software.

1.2.2.1 Následky

Z těchto předpokladů vyplynulo, že se informace stala *komoditou*. Kvůli obrovským objemům, v nichž se vyskytuje, ztratila na vzácnosti a tím i na ceně.

Vzrůstem objemu publikovaných informací, klesla průměrná kvalita informací. Avšak tím, že se otevřely možnosti publikování, mohou informace zveřejňovat také ti, kteří by se v tradičních podmínkách k publikování nedostali. Tím se web dostávají také informace hodnotné pro minoritní uživatelské skupiny.

Jelikož jsou však náklady na publikování na webu relativně nízké, klesla zodpovědnost autorů za publikovaný obsah. Díky tomu, že publikování je tak levné, není třeba se obávat nedostatku zájmu uživatelů, kteří autorům přinášejí zisk (např. skrze reklamu). Pokud bude informace špatná a nebude o ni zájem, nepřinese to autorovi žádnou škodu.

Jinak tomu je v klasickém modelu publikování, kdy riziko špatného prodeje na sebe bere vydavatel. To jej motivuje k tomu fungovat jako základní filtr informací a dohlížet na to, aby publikoval pouze díla, která mají vyšší naději na dobrý odbyt.

1.2.2.2 Zodpovědnost za publikované informace

Protože je publikování na internetu zároveň levné a bez rizika, klesá zodpovědnost za publikované informace. Ty často nejsou nikým vlastněny. Vlastnictví obvykle vytváří mezi vlastníkem a majetkem vztah zodpovědnosti. Na internetu však často chybí lidé, kteří by za publikované informace byli zodpovědní.

1.3 Uplatnění

Na základě stavu, v němž se informační prostor internetu nachází, vzniklo mnoho důvodů pro jeho filtrování. Filtrování informací našlo uplatnění v mnoha oblastech pro různé druhy informací.

Aplikace filtrování informací se v zásadě dají rozdělit do 2 skupin:

1. ty, které odstraňují nežádoucí informace,
2. ty, které naopak uživatelům získávají informace pro ně žádoucí.

1.3.1 Odstranění nežádoucích informací

Požadavek na odstranění informací, které nejsou pro uživatele žádoucí, vzniká z nadbytku informací a z jejich průměrně nízké kvality. Jde o pasivní postoj, kdy systém filtrování informací pouze reaguje na příchozí komunikaci.

Aplikace sloužící pro odstraňování informací jsou velmi užitečné, jelikož lidé mají omezenou schopnost nežádoucí či irelevantní informace ignorovat (filtrovat). S vyšším věkem navíc tato schopnost klesá.

Snížení počtu informací může mít rovněž pozitivní obchodní přínos. Kvůli tomu, že uživatelé

webových stránek (potenciální zákazníci) nemohou najít žádané informace, se obracejí na majitele stránek (firmy) například tím, že zavolají na telefonickou podporu, což vytváří firmě náklady. Špatná zkušenost s webovými stránkami může také vyústit ve ztrátu možného zákazníka.

1.3.1.1 Deduplikace

Typickou nežádoucí informací je *informace duplicitní*. K jejímu odstranění lze využít metody filtrování informací jako je výpočet podobnosti mezi dokumenty, kdy při překročení určitého prahu podobnosti jsou 2 dokumenty označeny jako duplikáty. Nejprve však mohou být vytvořeny shluky potenciálně podobných dokumentů, jejichž členové jsou poté porovnáváni každý s každým a testováni na duplicitu.

Pro uživatele může odstranění zdvojení přinést vyřazení duplicit z výsledků vyhledávání, pro správce databází se hodí úspora prostoru získaná odstraněním duplicitního obsahu.

1.3.1.2 Antispam

Nechvalně známým příkladem nežádoucí informace je *spam*. Spam je definován jako nevyžádané obchodní sdělení.

Termín „spam“ve významu, v němž je dnes používán, pochází ze *Spam Song* Terryho Jonese a Erica Idle, která se objevila 15.12. 1970 v televizním pořadu *Monty Python's Flying Circus*. V této písni zpívající Vikingové neustále opakovali „spam, spam, spam“. Slovem „spam“prokládali všechnu komunikaci, čímž naprosto *znemožnili cokoli sdělit*.

Zahlcení komunikačních kanálů spamem velmi poškozuje legitimní komunikaci a omezuje možnosti efektivního sdělování. Spam proto může být chápán jako *negativní externalita* (podobně jako znečištění) komunikačních technologií, jejichž silných stránek zneužívá.

První případ spamu pochází již z roku 1978, kdy Digital Equipment Corporation rozeslala několika stovkám výzkumníkům podílejícím se na Arpanetu reklamní sdělení o jejich novém produktu (Templeton, 2003).

Hlavní rozšíření spamu však nastalo až po roce 1993. Toho roku začali Laurence Canter a Martha Siegel, manželský pár imigračních právníků z Arizony, rozšiřovat po síti nechvalně proslulou reklamu „Green Card Lottery“. Tito první spammeři nakonec o tom, jak pomocí spamu zbohatnout, napsali knihu.

Spam využívá snadnosti distribuce informací ve velkém měřítku pomocí internetu. Typická spamová „kampaň“spočívá v rozeslání miliónů zpráv.

Spam se vyznačuje mnoha charakteristickými znaky, čehož je intenzivně využíváno při vytváření antispamových řešení. Vztah mezi rozesilateli spamu a tvůrci antispamových filtrů je vztahem akce-reakce. Antispam se snaží využít typických znaků spamu proti němu, kdežto jeho původci usilují o to, aby spam nebyl filtry rozpoznán.

1.3.1.2.1 E-mailový spam

Pravděpodobně nejznámějším představitelem spamu je e-mailový spam. Není divu, že všichni najednou dostávají stejné e-maily, když 92,6 % ze všech e-mailů je spam (Han, 2008).

E-mailový spam je používán například k propagaci farmaceutických výrobků pochybné kvality, pornografických stránek, vylákávání peněz, krádežím osobních a autentizačních

údajů[4] nebo šíření malware (viry, trojské koně atp.). Viry mohou proměnit počítač v *„zombie“*, čímž spammer získá nad počítačem kontrolu a může jej využívat k rozesílání nevyžádané pošty.

1.3.1.2.2 Komentářový spam

Obsah na webu je nyní tvořen nejen tvůrci stránek, ale také jejich uživateli. Příkladem obsahu vytvářeného uživateli mohou být komentáře. Bohužel mezi uživatele píšící komentáře, patří také spammovací roboti. Pro šiřitele spamu jsou komentáře dalším distribučním kanálem, jímž mohou šířit svá nevyžádaná sdělení.

1.3.1.2.3 Ostatní

Spam není vymezen technologií, pomocí níž je šířen. Některé komunikační technologie pro něj však jsou vhodnější. Mezi hlavní předpoklady, které musí vhodná komunikační technologie splňovat, patří nízké náklady na šíření velkých objemů zpráv, anonymita původce zprávy a možnost automatizace rozesílání.

- **SMS spam**
  Komunikační protokol SMS (short message service) je využíván k vyměňování krátkých textových zpráv mezi mobilními zařízeními. Díky možnosti rozesílání SMS přes různé brány bývá tato komunikační technologie zneužívána k šíření spamu. Prozatím to však není tak rozšířená metoda, neboť stále obnáší jisté náklady spojené s přenosem zprávy.
- **IM spam**
  Spam pomocí chatovacích služeb *instant messaging* (IM) se v současnosti objevuje poměrně zřídka. Obvykle obsahuje odkaz na placený obsah.
- **VoIP spam**
  Voice over Internet Protocol (VoIP) pro přenos hlasů pomocí paketů není na rozdíl od obvyklých způsobů šíření spamu asynchronní a proto vyžaduje okamžitou reakci uživatele. Původci spamu mohou této komunikační technologie zneužít k šíření předem nahraných reklamních sdělení. Tento způsob prozatím není příliš rozšířený.

1.3.1.3 Blokování webových stránek

Vzhledem k množství a povaze webových stránek není divu, že také některé z nich bývají shledávány nežádoucími.

Nežádoucí webové stránky jsou filtrovány ani ne tak kvůli jejich chabé kvalitě, jako spíš kvůli nemorálnosti, protiprávnosti, politické nekorektnosti nebo kontextu, v němž jsou přijímány.

1.3.1.3.1 Ochrana nezletilých

Nejčastěji používaným zdůvodněním blokování webových stránek je ochrana nezletilých. Ta spočívá v zamezení přístup k těm stránkám, které by podle všeobecně přijímaných měřítek mohly mít negativní dopad na psychický vývoj nezletilého. Mezi typické příklady stránek,

---

4tzv. *phishing*

jejichž prohlížení je nezletilým odepřeno, patří pornografie, stránky šířící rasismus a nesnášenlivost nebo stránky vybízející k porušování zákona (např. autorských práv).

Nezletilí jsou také chránění před „on-line predátory“, čímž jsou obvykle míněni pedofilové.

Nejčastěji jsou tato omezení aplikována tam, kde je přístup k internetu k dispozici veřejnosti se sníženou možností kontroly uživatelů. Také instituce, které jsou financovány ze státního rozpočtu, jako školy a knihovny, jsou v některých zemích povinny zavést blokování webových stránek v určité formě. Uplatňuje se také řešení webového portálu nebo vyhledávače, který je přizpůsoben tak, že nabízí jen odkazy na stránky, které jsou pro děti bezpečné a neprezentují závadný obsah.

V tomto případě se často namítá, že mnohem účinnější než technologické řešení, je výuka k informační gramotnosti zahrnující návyky kritického hodnocení informací, zachovávání soukromí na internetu a vyvarování se ilegálního obsahu.

Technologie navíc není nestranná ve svých účincích a hodnotách, které šíří nebo omezuje. Svádí ke zneužití k jiným účelům než je ochrana dětí.

1.3.1.3.2 Knihovny

Ačkoliv jsou veřejné knihovny obvykle chápány jako nástroj prosazování práva jedince na informace, musí v některých zemích toto právo naopak omezovat nainstalováním filtrovacího software.

Knihovny v USA, které chtějí dostávat federální finance[5], musí dodržovat *Children Internet Protection Act* (CIPA). To pro ně znamená, že musí mít stanoveny bezpečnostní zásady pro používání internetu a musí na své počítače nainstalovat *ochrannou technologii*, která zablokuje přístup k materiálům, které jsou obscénní nebo škodlivé pro děti.

Každá knihovna musí mít od 1.7. 2004 stanoveny bezpečnostní zásady pro internet, které musí obsahovat:

- používání technologické ochrany, jako je filtrovací software
- tento software nainstalovaný na každý počítač v knihovně

postup pro odblokování webové stránky na základě žádosti dospělého

Přestože je CIPA podle svého názvu zaměřen na ochranu dětí, vyžaduje filtrování internetu také pro dospělé a dokonce i pracovníky knihovny.

Pro některé Američany představují veřejné knihovny jedinou možnost přístupu k internetu. Tuto skupinu by nadměrné filtrování postihlo nejvíce. Neefektivně vedené filtrování by mohlo vést k většímu prohloubení rozdílu mezi skupinou uživatelů přistupujících k internetu výhradně ve veřejných knihovnách a ostatními, kteří mají jinde přístup k nefiltrovanému internetu.

Podle CIPA z roku 2000 může dospělý v knihovnách přistupovat k jakýmkoli webovým stránkám (s výjimkou právně definované obscenity a dětské pornografie). Pokud je mu přístup ke stránkám zablokován, má právo požádat pracovníka knihovny, aby mu byl přístup umožněn. Přestože je právně možné si v knihovně prohlížet pornografické stránky, nutnost požádat knihovníka nejspíš většinu uživatelů odradí. Podobný účinek může mít upozornění na to, že veškeré prohlížené stránky jsou monitorovány a ukládány.

V knihovnách a státních institucích, které poskytují přístup k internetu se jako funkční

5 *e-rate* – státní příspěvek na připojení k internetu

ukazuje řešení umístit počítače s připojením k internetu do otevřených prostor. Vzniká tak sociální tlak na přijatelné využívání internetu, neboť je uživatel vystaven zrakům ostatních lidí.

1.3.1.3.3 Blokování podle zamýšleného účelu využití internetu

Poskytovatel přístupu k internetu má právo omezit dostupné webové stránky tak, aby zamezil jeho využívání pro účely jiné než jím stanovené. Poskytovatel vymezí okruh povolených způsobů použití internetu a jakýkoli jiný způsob nechá zablokovat. Případně, pokud je benevolentnější, může uživatelům nastavit kvótu; tzn. počet přístupů za den na určité stránky jako jsou zprávy nebo počasí, které nespadají mezi jím podporované způsoby využití internetu.

Mezi typické případy, kdy mají uživatelé využívat internetu pouze určitým způsobem, patří školy a zaměstnání.

- **Školy**
  Účelem připojení k internetu ve školách a jiných vzdělávacích zařízeních je umožnění přístupu k webovým stránkám, které mají vzdělávací hodnotu jako jsou stránky zpřístupňující vzdělávací materiály a e-learningové aplikace.
  Aby se podpořil tento způsob využití internetu, jsou webové stránky bez vzdělávácí hodnoty zablokovány.
  V USA musí 83% studentů veřejných škol podepsat prohlášení, že nebudou přístupu k internetu zneužívat k zakázaným aktivitám (Messmer, 2005).
- **Zaměstnání**
  Účelem filtrování webových stránek zaměstnavateli je zamezení aktivitám zaměstnanců, které nesouvisejí s jejich pracovní náplní.
  Blokované stránky jsou zaměstnavatelem považovány za ztrátu času (pornografie, gambling, sport). Internetové rádio a televize navíc představují velké datové přenosy, které mohou zablokovat nebo omezit dovolené způsoby využití. Existují však také stránky, které jsou nebezpečné samy o sobě: zaznamenávají stisknuté klávesy ve snaze zachytit nebo jinak vylákat autentizační údaje uživatelů. Filtrovány mohou být také stránky pro hledání zaměstnání.

V některých firmách zaměstnanci podepisují *Internet User Policy* (IUP), která stanovuje, co je a co není bráno jako přijatelné využití internetu.

1.3.1.3.4 Cenzura

Filtrování informací se v krajních případech může proměnit v cenzuru. Nástrojům určeným k blokování webových stránek se někdy dokonce říká *„censorware“*.

Přestože panuje souhlas v tom, že ten, kdo zprostředkovává přístup k internetu, má právo jej filtrovat, pokud je zprostředkovatelem vláda, neměla by filtrování aplikovat.
Cenzura internetu se vyskutuje jak v zemích demokratických, tak nedemokratických autoritativních režimech potlačujících lidská práva a svobody. Hlavní rozdíl mezi fungováním filtrování jakožto nástroje cenzury v těchto zemích, spočívá ve způsobu, jakým je prováděno.
Předmětem cenzury může být pornografie, gambling, náboženské a politické stránky nebo stránky propagující hnutí potlačující svobodu jednotlivce.

- **Cenzura v demokratických zemích**
  Cenzura v demokratických zemích probíhá jako centrálně řízené filtrování internetu se zaměřením na obsah, který podle legislativy daného státu není legální. Vyznačuje se tím, že je transparentní, což znamená, že je zřejmé co a proč je blokováno.
  Mezi země, v nichž se uplatňuje tento druh cenzury, patří kupříkladu Jižní Korea nebo Francie, kde jsou např. zablokovány stránky *Yahoo!* , které nabízejí na prodej nacistické relikvie.
- **Cenzura v nedemokratických zemích**
  Cenzura v zemích nedemokratických je často zaměřena na politicky nebo nábožensky nekorektní obsah. Filtrování probíhá netransparentním způsobem. Takřka nikdy není jasné, co a proč bylo odfiltrováno. Neexistují spolehlivé seznamy toho, co je blokováno.
  Cenzurní zákony se často tváří, že potírají nikoliv závadný obsah, nýbrž následky (sekundární efekty) jeho dostupnosti, a tedy jsou *„obsahově neutrální“*, díky čemuž vyhovují podmínce svobodného projevu.
  Tento druh cenzury se uplatňuje v zemích jako Barma, Severní Korea, Čína, Írán nebo Saudská Arábie. Nejpropracovanější filtrování internetu je pravděpodobně zavedeno v Číně.
  Čína filtruje výsledky internetových vyhledávačů, takže např. dotaz na *„nezávislost Taiwanu“* nic nenalezne. V Číně jsou rovněž jsou filtrovány SMS, které obsahují podezřelé výrazy, fráze a čísla.

1.3.2 Získání žádoucích informací

Pro uživatele, kteří aktivně hledají nové informace, jsou určeny aplikace filtrování informací, které se zaměřují na získání žádoucích informací. Uplatňují se v zásadě ve 2 případech: pokud uživatel hledá informace podle *přibližných požadavků*, a pokud uživatel chce *doporučit vhodné informace*.

1.3.2.1 Hledání podle přibližných požadavků

Tradiční systémy vyhledávání informací požadují po uživateli, aby svou informační potřebu formalizoval do podoby rešeršního dotazu. V některých případech však takové zpřesnění není možné nebo je pro uživatele obtížné. V systémech filtrování informací jsou informační potřeby uživatelů zachyceny prostřednictvím jejich *profilů*, v nichž jsou uvedeny například jejich zájmy.

Pokud uživatel přesně ví, co chce, použije vyhledávání. Naopak, pokud jeho informační potřeba není blíže specifikovaná, použije prohlížení (Iskold, 2007a). V případech, kdy uživatel hledá něco, co by se mu mohlo líbit, se uplatní systémy filtrování informací. Stejně tak v oblastech, kde neví, co hledá. Typickým příkladem takové oblasti je seznamka.

1.3.2.2 Doporučování

Systémy filtrování informací vytvářející doporučení, vznikly jako řešení rozšiřujících se možností ve výběru nabízených produktů a služeb. Čím širší je sortiment nabízeného zboží, tím obtížnější je pro uživatele výběr. Přebytek možných voleb vede k dezorientaci a

nespokojenosti uživatele. Navíc v prostředí elektronických obchodů chybí osoba prodejce, který je schopen působit jako odborný poradce při výběru.

Doporučení v e-shopech má za účel maximalizovat pravděpodobnost nákupu. 60 % nákupů na Netflix.com vznikne na základě doporučení (Stockwell, 2009).

Systémy doporučení obohacují elektronické obchody 3 způsoby:

1. mění návštěvníky v zákazníky
2. podporují větší nákupy (cross-selling) tím, že doporučují další jednotky, které by si uživatel mohl koupit
3. budují loajalitu - zákazník věnuje určitý čas tomu, aby se systém doporučení naučil, jak mu vytvářet správná doporučení, a proto nemá zájem přecházet ke konkurenci, protože jejich systém by nejprve vyžadoval určitý čas, aby se naučil zákazníkův profil

2 Základy

Tato kapitola pojednává o základních prvcích a postupech, které jsou v metodách filtrování informací používány. Prvky, s nimiž tyto metody pracují, jsou především vstupní data. Na ně aplikují postupy sloužící k jejich předzpracování. Nejprve je však uvedeno použité rozdělení metod, jimž se věnují následující kapitoly.

2.1 Dělení metod filtrování informací

Rozdělení metod, které jsou používány k filtrování informací, není jednoznačné. Mají mnoho společného a používají podobné postupy. Navíc jsou v praxi často implementovány v kombinaci nebo hybridním propojení.

Původně byly rozlišeny 3 druhy filtrování informací (Malone, 1987):

1. *kognitivní* - filtrování založené na obsahu
2. *ekonomické* - dokumenty jsou filtrovány na základě nákladů potřebných k jejich získání ve srovnání s užitkem plynoucím z jejich použití[6]
3. *sociální* - založeno na důležitosti odesílatele pro příjemce zprávy

Později se rozšířilo zejména rozdělení na *filtrování kolaborativní* a *filtrování založené na obsahu*. Tyto dvě skupiny byly brány takřka jako protiklady. Začaly se však také objevovat formy, které využívaly předností obou.

Pro účely tohoto textu bylo použito rozdělení metod filtrování informací podle pravidel, která používají.

2.1.1 Rozdělení podle pravidel

Všechny metody filtrování informací jsou *založeny na pravidlech*. Pravidla stanovují, za jakých podmínek je informační jednotka vyhodnocena jako žádoucí nebo nežádoucí.

Rozdíl mezi jednotlivými druhy filtrování spočívá v tom, na co jimi užívaná pravidla kladou podmínky. Tak například jestliže je podmínkou pro označení informační jednotky za žádoucí kvalitní a relevantní obsah, jedná se o filtrování založené na obsahu. Pokud je podmínkou vysoké hodnocení ostatních uživatelů systému, jde o filtrování založené na spolupráci.

- *filtrování založené na atributech* klade podmínky na *atributy* informační jednotky (např. IP adresa, způsob prezentace)
- *filtrování založené na obsahu* klade podmínky na *obsah* informační jednotky (např. slova v textu)
- *filtrování založené na spolupráci* klade podmínky na *hodnocení* informační jednotky

Předmětem podmínek jsou vždy *data*. Ta však musí být ve strojově zpracovatelné podobě, aby s nimi mohly algoritmy filtrování informací pracovat. Proto je zapotřebí nejprve data

6 např. dostupná kapacita telekomunikační linky (šířka pásma), velikost dokumentu

do této formy předpřipravit.

2.2 Reprezentace uživatelů

Prvním typem vstupních dat systémů filtrování informací jsou údaje o uživatelích. Tyto informace je třeba reprezentovat ve strojově zpracovatelné podobě, která umožní zachytit informační potřeby uživatelů. K tomuto účelu jsou vytvářeny *uživatelské profily*.

2.2.1 Uživatelské profily

Profil reprezentuje dlouhodobé informační potřeby uživatele. Představuje záznam o jeho *„životě“* v systému.

Může být strukturován na 3 úrovních:

1. **Schopnosti** Explicitně vyjádřené schopnosti uživatele, které určují, jaký typ informačních jednotek dovede využít (např. jeho jazykové kompetence nebo úroveň odbornosti).
2. **Zájmy** Vymezení oblastí zájmu uživatele (např. pomocí mapy konceptů).
3. **Stopy** Stopy vznikají zaznamenáváním chování uživatele v systému. Jde o druh implicitního hodnocení.

Profily mohou být tvořeny z:

1. *klíčových slov* - poměrně jednoduchý způsob, který přibližně vyjadřuje informační potřebu uživatele
2. *jednotek* - jde o jednotky, o nichž uživatel určitým způsobem prohlásí, že jsou pro něj zajímavé
3. *hodnocení* - profil se základá na souboru hodnocení, která uživatel poskytl systému

Aby mohla být zjišťována vhodnost informační jednotky pro uživatele, musí být oboje zachyceno tak, aby bylo možné vzájemné porovnání. Proto je často uživatelský profil přeložen na reprezentaci „pseudo-dokumentu“- tzn. jako vektor.

2.2.2 Reprezentace důvěryhodnosti uživatelů

Součástí profilů může v některých případech být hodnocení uživatele ostatními. Tímto způsobem lze do systému vpravit prvek „důvěry“a zlepšit tím jeho kvalitu. Tak může být měřena reputace a autorita uživatelů. Evidence důvěry umožňuje částečně odstínit negativní vliv uživatelů snažících se zkreslit doporučení systému.[7]

„Důvěra“může vyjadřovat přesnost předvídaného hodnocení. Pokud je předvídané a skutečné hodnocení shodné (nebo blízce podobné), zvyšuje se tím „důvěryhodnost“zdroje předvídaného hodnocení (uživatele).

Pokud uživatel A hodnotil kladně uživatele B a ten hodnotil kladně uživatele C, lze se domnívat, že A by C hodnotil také kladně.

Jde o princip *„web of trust“*, který způsobuje propagaci důvěry sítí. Díky tomu lze vytvářet

---

7Typickým příkladem začlenění hodnocení důvěryhodnosti je Epinions.com (O’Donovan, 2006), systém pro doporučení lyžařských lokalit *MoleSkiing.it* nebo aplikace *TrustMail*, která vytváří síť důvěry pro e-mailovou komunikaci.

vztahy příbuznosti ne pouze na základě podobnosti uživatelských profilů, ale také na základě blízkosti v síti vzájemně se hodnotících uživatelů.
Přesto (nebo možná právě proto) jsou systémy založené na prvku důvěry náchylné ke zkreslení manipulativními uživateli, jelikož spolu s „důvěrou“se šíří hodnocení důvěřovaných uživatelů, což zakládá na možnost posilování šíření manipulativních hodnocení. Možná také z toho důvodu se ukazuje, že uživatelé by rovněž ocenili, kdyby měli možnost kromě důvěry v systémů vyjadřovat také *nedůvěru* neboli negativní hodnocení (Lee, 2008).

2.3 Reprezentace jednotek

Stejně jako je třeba zachytit uživatele tak, aby jejich reprezentace mohla být strojově zpracovávána, je zapotřebí reprezentovat informační jednotky, s nimiž systém pracuje. Ty jsou často v nestrukturované podobě, jako texty v přirozeném jazyce, obrázky nebo multimédia.

Jednotky lze reprezentovat na základě informací získaných z jejich obsahu nebo informací získaných zvenčí jako jsou metadata nebo hodnocení. Popis založený na obsahu jednotky se uplatňuje zejména v případě textových jednotek. U zvukových nebo multimediálních jednotek se prozatím většinou spoléhá na metadata.

2.3.1 Vektorová reprezentace

Zavedenou formou reprezentace textových jednotek jsou vektory. Tento způsob je využíván především při vyhledávání informací.

Dejme tomu, že v databázi jsou 2 hypotetické dokumenty:

- Dokument A: *„Každý upír má rád kefír.“*
- Dokument B: *„Drákula je upír.“*

Mřížka (rozměry vektorového prostoru) je určena výrazy, které se nacházejí v korpusu všech dokumentů. Výrazy mohou být zjednodušeny na kořen slova, což umožní výrazně snížit počet potřebných rozměrů.

(„Drákula“, „být“, „každý“, „kefír“, „mít“, „rád“, „upír“)

Dále lze mřížku zjednodušit vypuštěním nevýznamových slov jako jsou předložky, spojky nebo částice. Stejně tak je možné se zbavit výrazů, které se vyskytují velice často, což by v tomto případě byly zřejmě „být“, „každý“, „mít“a „rád“. K přesnějšímu určení váhy výrazu se používá technika *TF×IDF*, která bude popsána v kapitole o filtrování založeném na obsahu.

Dokumenty můžeme reprezentovat jako vektory, kde hodnota pro každý rozměr určuje počet výskytů daného výrazu v dokumentu:

- A=(0,0,1,1,1,1,1)
- B=(1,1,0,0,0,0,1)

Jakmile jsou informační jednotky reprezentovány v této podobě, lze s nimi provádět efektivní algoritmické operace.

2.3.2 Latentní sémantické indexování

Klasický model vektorového prostoru reprezentuje jednotky pomocí výrazů přirozeného jazyka. Avšak i po sjednocení výrazů odvozených ze stejného kořenu slova a odstranění nevýznamových a častých výrazů, zůstává vektorová reprezentace plná výrazů vyjadřujících stejný význam.

Mnohem efektivnější by bylo, kdybychom mohli jednotky reprezentovat přímo pomocí významů. Tím pádem by rozměry vektorového prostoru nebyly jednotlivé výrazy, nýbrž sémantické koncepty, s podobnou funkcí jakou mají předmětová hesla nebo klasifikační notace.

Pro tento účel slouží pokročilá metoda založená na lineární algebře nazvaná *latentní sémantické indexování,* které využívá skryté struktury ve vzorcích používání slov (Papadimitriou, 1998). Tímto způsobem lze potlačit negativní vlivy jevů vysktujících se v přirozeném jazyce jako je synonymie, homonymie nebo polysémie.

2.4 Shlukování

V případě, kdy by každé slovo tvořilo jednu dimenzi vektorového prostoru, nacházely by se textové dokumenty v prostorech s mnoha tisíci rozměry. Vícedimenzionální prostory jsou výpočetně náročné, a proto se přistupuje k zmenšení počtu rozměrů pomocí jejich sdružování.

Jeho účelem je sjednotit podobné prvky do shluků - vysledovat v datech odlišné skupiny. Míra podobnosti členů shluku se nazývá jeho „těsností“. Požadovanou těsnost je nutné vhodně stanovit, aby vznikl rozumný počet shluků. Pokud by totiž byl požadavek na velmi vysokou míru podobnosti, vzniklo by velké množství malých shluků, což by vedlo k potlačení původního účelu shlukování. Shlukování je aplikováno jak na uživatele, tak na jednotky. V základní formě shlukování je prvek (uživatel, jednotka) přiřazen k právě jednomu shluku. To například znamená, že při vytváření doporučení pro uživatele jsou bráni v potaz jenom členové shluku, do nějž přísluší. Novější přístupy využívající fuzzy logiku umožňují účast prvku ve více shlucích, kdy je míra účasti určena pomocí koeficientu. Díky tomu lze vytvářet přesnější doporučení, které je založeno na více uživatelích pocházejících z více shluků, k nimž prvek náleží.

Metody založené na vytváření shluků obvykle nabízejí méně osobní doporučení, přičemž jejich přesnost může být i horší než při metodách založených na nejbližších sousedech. Výhodu shlukování je jeho rychlost a nenáročnost, jakmile jsou shluky vytvořeny, což může být provedeno v off-line režimu. Shlukování se hodí jako první krok před použitím některých náročnějších metod.

Mezi základní postupy určené pro shlukování patří:

1. *hierarchické shlukování* - vytvářející hierarchicky organizované shluky
    (a) *divizní*- postupující shora dolů

(b) *aglomerační* - postupující zdola vzhůru

2. *k-středové shlukování* - využívací přiřazování k nejbližším středům

2.5 Hodnocení

Hodnocení představují metadata, která umožňují prvky systému filtrování informací (jednotky, uživatele) organizovat. Tento typ dat přidává systému další strukturu.

2.5.1 Dělení hodnocení

Hodnocení lze rozdělit podle předmětu, ke kterému se vztahují. Tím se v systému vytváří několik úrovní hodnocení.

1. *hodnocení jednotek*
2. *hodnocení uživatelů*

Kromě jednotek lze hodnotit uživatele systému. Tím mohou uživatelé vyjádřit, nakolik jsou pro ně hodnocení uživatele užitečná.

3. *hodnocení hodnocení*

   Další stupeň představuje namísto hodnocení uživatelů hodnotit přímo jejich jednotlivá hodnocení. Takže namísto kladného hodnocení určitého uživatele, lze hodnotit jeho hodnocení konkrétní jednotky.[8]

4. *hodnocení hodnocení hodnocení* atd.
   Teoreticky by hodnocení mohlo pokračovat i na vyšších úrovních, kdy by uživatel hodnotil hodnocení jiného hodnocení. V praxi se tato úroveň podrobnosti nepoužívá.

Hodnocení může mít více vrstev:

1. **základní slovník** pro autory informačních jednotek
   Ten by měl být schopen popsat:
   (a) *obsah* - např. nahota, vulgarismy, násilí
   (b) *kontext* - např. medicínské zpracování, sportovní násilí
   (c) *médium* - např. obraz, text, video
2. **šablony** hodnotící výrazy základního slovníku, které jim přiřazují různou váhu; mohou být vytvářeny např. neziskovými organizacemi nebo firmami
3. **pomocné blacklisty a whitelisty** sloužící pro zachycení výjimek, které by sice podle pravidel z předchozích vrstev byly zablokovány, avšak poskytují hodnotný obsah

Hodnocení informací můžou vytvářet:

1. **autoři informací**
   Jsou zastoupeni zejména jednotlivci nebo korporacemi (např. vydavatelé). Mají však pro hodnocení nedostatek motivace, jelikož jim nepřináší bezprostřední užitek. Navíc jsou tato hodnocení často nespolehlivá, jelikož se nepodřizují žádným standardům.

8V některých systémech doporučení je hodnocení posunuto na vyšší úroveň tím, že dovolují hodnotit recenze. Je tomu tak např. na *Amazon.com*, kde lze označit, zdali byla recenze uživateli užitečná nebo nikoli.

2. **zainteresované třetí strany**
   Kromě původců informačních jednotek mohou hodnocení vytvářet ty subjekty, které pro to mají zájem nebo odbornost. Tato skupina hodnotících však není rozsáhlá ve srovnání s množstvím dostupných informací (např. učitelé, sdružení jako ACLU[9], firmy vytvářející blacklisty nebo katalogizátoři).
3. **uživatelé informací**
   Využívání hodnocení uživateli se osvědčilo v metodách kolaborativního filtrování. Hlavní výhodou je počet uživatelů, který umožňuje vytvořit dostačující objem hodnocení.

2.5.1.1 Dělení podle počtu rozměrů

Hodnocení lze také rozdělit podle počtu rozměrů, které jsou schopna vyjádřit.

2.5.1.1.1 Jednorozměrná

Jednorozměrná hodnocení mají velice základní popisnou schopnost, ale lze je jednoduše zpracovávat.

1. **Skalární** - 1–5 hvězdiček, hodnocení ∈ {silně souhlasí, souhlasí, neutrální, nesouhlasí, silně nesouhlasí}
2. **Binární** - líbí/nelíbí, souhlasí/nesouhlasí, dobré/špatné
3. **Unární** - uživatel si objekt prohlížel, koupil, nebo jiným způsobem pozitivně hodnotil

2.5.1.1.2 Vícerozměrná

Hodnocení mohou být kromě jednorozměrných, také vícerozměrná neboli *sémantická*.

Jednorozměrná hodnocení postrádají *důvod*. Pokud uživatel hodnotil danou jednotku například záporně, nelze zjistit, co ho k tomu vedlo. Například pokud se mu nelíbí film „Smrtonosná past“, není zřejmé, zdali se mu nelíbí, protože jej považuje za příliš násilný nebo naopak málo násilný.

Systémy využívající vícerozměrná hodnocení, jako je například systém pro doporučování restaurací *Entree* (Burke, 2000), však uživatelům dovolují hodnotit na více úrovních. Tato hodnocení jsou nazývána *„kritiky“*. Uživatel například může uvést, že daná restaurace je příliš drahá, nebo příliš tradiční.

Vícerozměrná hodnocení lze převádět na hodnocení jednorozměrná. Takže například kritiku „příliš drahé“lze interpretovat jako negativní hodnocení. Tím se však přichází o významnou informační hodnotu hodnocení.

Vícerozměrná hodnocení jsou příkladem zlepšení funkce systémů filtrování informací pomocí obohacení vstupu. Druhou možností je vyvinout pokročilejší algoritmy pro zpracování sémanticky chudých údajů.

2.5.1.2 Dělení podle způsobu získávání

Hodnocení lze rovněž rozdělit podle způsobu, jímž je získáváno. Rozlišuje se tak hodnocení

9American Civil Liberties Union

implicitní, které uživatel poskytuje systému nevědomě, a hodnocení *explicitní*, které poskytuje vědomě.

2.5.1.2.1 Implicitní hodnocení

Implicitní hodnocení je automaticky vyvozováno z chování uživatele. Potenciálně může každá interakce uživatele se systémem vytvářet implicitní data. Aby bylo možné přiřadit takto získané hodnocení k uživateli, je nutné, aby se uživatel při vstupu do systému identifikoval. Data získaná bez předchozího určení uživatele slouží spíše jako podklad pro webovou statistiku.

Při tvorbě implicitních hodnocení se může zohlednit například doba, kterou uživatel strávil prohlížením jednotky. Jako kladné hodnocení lze brát případy, kdy si uživatel jednotku zakoupil. Sekvence vyhledávacích dotazů, které uživatel pokládá, může být zaznamenávána a interpretována jako postupné vylepšování dotazu.

Nevýhodou tohoto způsobu získávání hodnocení je možnost jeho různé interpretace. Pokud např. zjistíme, že uživatel strávil prohlížením určité jednotky hodinu, může to také znamenat, že si mezitím zašel na oběd. Stejně tak může víceznačně vyznít zakoupení jednotky. Lze se domnívat, že ji uživatel kupuje ze zájmu, avšak nelze vyloučit ani situace, kdy ji kupuje pro někoho jiného jako dar nebo je pro něj koupě nezbytná.

U implicitního hodnocení je však obtížnější úmyslné zkreslení, zvláštně v případech, kdy není uživateli zřejmé, co je sledováno a jak je s takovými údaji nakládáno.

2.5.1.2.2 Explicitní hodnocení

Explicitní hodnocení se zakládá na konkrétní vědomé akci uživatele. Při jeho získávání jsou uživatelé výslovně požádáni o zhodnocení jednotky. Vytvoření takového hodnocení vyžaduje od uživatele akci „navíc“, kdy např. musí zvolit hodnocení na škále od 1 do 5, napsat komentář či recenzi.

Jedním z hlavních problémů explicitních hodnocení je cena jejich získávání a motivace uživatelů k jejich poskytování. Explicitní hodnocení vyžaduje po uživateli kognitivní úsilí. To samo o sobě není špatná věc - podobně jako se předpokládá, že učitelé budou nejprve přemýšlet o známkách, které dají svým studentům.

Výhodou explicitního hodnocení je, že má ve většině případů jednu jednoznačnou možnost interpretace. Takže v případě, kdy uživatel ohodnotil jednotku plným počtem hvězdiček, je zřejmé, že tím vyjadřuje kladné hodnocení.

Tím, že je akt explicitního hodnocení vědomý, však zakládá na možnost záměrného zkreslování. To je však poměrně teoretická možnost, jelikož poskytování zkreslených dat obvykle není v zájmu uživatele. Pokud by uživatel vytvářel zkreslená hodnocení, dostával by zkreslená doporučení a užitečnost systému filtrování informací by pro něj poklesla. Jestliže však cílem uživatele není pouze přesnost doporučení, která od systému dostává, může být motivován ke zkreslování údajů, které systému poskytuje. To se může stát v případě různých veřejných hlasování nebo soutěží.

2.5.2 Zachycení kontextu

Problémem jak explicitního, tak implicitního hodnocení, je jeho ochuzenost o kontext.

Kontextuální informace, jako je například původ hodnocení, by v systému měly být zahrnuty. Většinou jsou však hodnocení chápána tak, že jejich hodnota je nezávislá na okolnostech, při nichž byla vytvořena. To představuje objektivistický přístup, který předpokládá, že hodnocení mají smysl i mimo kontext.

Hodnocení by mělo obsahovat *kontextuální operátory* reflektující kontext informační jednotky. Například násilí může být bezdůvodné (u masových vrahů) nebo „legitimní“(film z 2. světové války). V sociologické literatuře se rozlišuje účinek násilí na pozorovatele podle toho, zdali je násilí odměněno či potrestáno (Balkin, 1999). Nahota v uměleckém kontextu (např. antické sochy) může být tolerována, na rozdíl od nahoty v pornografii. Stejně tak je velký rozdíl, zdali je dílo realistické nebo komiksové.
Tyto kontextuální informace však není lehké zachytit ve formalizované podobě, proto se navrhuje přenechat řešení pro whitelisty obsahující seznam výjimek (Balkin, 1999).

2.5.3 Citace a odkazy

Speciálním druhem hodnocení informační jednotky je její citace, která v prostředí internetu nabývá podoby *hyperlinkového odkazu*. Analýzou citací poměřují citační rejstříky kvalitu informačních jednotek. Na podobném principu fungují internetové vyhledávače, které podle počtu a kvality odkazů vedoucích na informační jednotku odvozují její důležitost, kterou pak zohledňují při řazení výsledků vyhledávání.

Tento způsob hodnocení se může uplatnit také v systémech filtrování informací. Stejně jako u citačních rejstříků však musí brát v potaz faktory, které zkreslují interpretaci citací. Je zapotřebí neomezit se pouze na kvantitu citací (odkazů), ale vyhodnocovat také jejich *kvalitu* (García-Barriocanal, 2005). Starší informační jednotky obvykle mají poměrně více citací oproti novějším, proto je vhodné zavést citační skóre relativní ke stáří jednotky.

2.5.4 Formalizace

Hodnocení je po jeho vytvoření potřeba zachytit ve strojově čitelném formátu tak, aby je mohl filtrovací software *rozpoznat a respektovat*. Pro tento účel bylo vytvořeno několik druhů formátů. Ke známějším patří schéma PICS (Platform for Internet Content Selection), PICS Rules a formáty založené na RDF (Resource Description Framework). Dá se však říci, že žádný z těchto formátů se příliš neuchytil a povětšinou se v systémech filtrování informací používají formáty nestandardizované.

2.6 Měření podobnosti

Pokud má systém filtrování informací k dispozici data ve strojově zpracovatelné podobě, může s nimi začít provádět výpočetní operace. Základní druh výpočtů, které provádějí systémy filtrování informací, slouží k *určení podobnosti* prvků, které obsahuje.

Míra podobnosti bývá obvykle vyjadřována jako koeficient $k \in \langle 0,1 \rangle$. V případě, když $k=1$, jsou porovnávané jednotky identické. Když $k=0$ jsou zcela odlišné.
Podobné prvky se mohou vyznačovat tím, že obsahují výskyty podobných klíčových slov (tagů). V tom případě by se pro určení podobnosti použila množinová reprezentace prvků.
Nejdůležitějším problémem tohoto a mnoha zavedených způsobů měření podobnosti však je,

že probíhají spíše na syntaktické rovině, přestože jejich účelem je zjistit podobnost v rovině sémantické, na úrovni významů.

2.6.1 Podobnost ve vektorovém prostoru

Pokud rozměry prostoru reprezentují významy, lze blízkost jednotek v takovém prostoru chápat jako jejich podobnost. V případě modelu vektorového prostoru tvoří jednotlivé rozměry buď dílčí výrazy nebo latentní významy.

Rozšířeným způsobem výpočtu podobnosti je měření ve vektorovém prostoru pomocí kosinu úhlu, který svírají vektory jednotek. Podobnost jednotek *u* a *v* můžeme vypočítat tak, že skalární součin vektorů, které je reprezentují, vydělíme součinem jejich délek.

2.6.2 Vzdálenost v ontologii

Pokud jsou jednotky klasifikovány podle určité ontologie, jako jsou systémy věcného pořádání, lze jejich podobnost odvodit z jejich vzdálenosti v hierarchii této ontologie. Ve stromovém zobrazení ontologie je vzdálenost, která dělí jednotky, rovna počtu hran, které tvoří jejich spojnici. Podobnější jsou ty jednotky, které sdílí bližšího předka.

2.6.3 Další způsoby měření podobnosti

Existuje mnoho dalších způsobů určování podobnosti. Mezi nim jsou to například:

- Heuristická podobnost
- Pearsonova korelace
- Jaccardův koeficient
- Manhatannská vzdálenost
- Tanimotovo skóre

3 Filtrování založené na atributech

Pravidla, která používá filtrování založené na atributech, zkoumají atributy informační jednotky. Podle výskytu sledovaných atributů je rozhodnuto, zdali se jedná o informaci žádoucí či nežádoucí.

Tyto metody jsou aplikovány zejména v oblasti filtování spamu a blokování webových stránek. Je pro ně typické vytváření *seznamů odesílatelů* informačních jednotek. Pro nežádoucí jsou to *blacklisty*, pro žádoucí *whitelisty*.
Nedostatkem těchto metod je, že původci nevyžádaných e-mailů se rychle naučí, které atributy jsou sledovány, a podle toho přizpůsobí spam, který rozesílají, aby těmto kritériím vyhověl a skrze filtr prošel. Velká část spammerů se však nepřizpůsobuje tak rychle, takže tento způsob filtrování obvykle zachytí značný podíl nelegitimní pošty.

3.1 Filtrování podle IP adresy

V prostředí internetu se většina komunikace odehrává podle pravidel sady protokolů TCP/IP[10], podle níž je každý účastník komunikace jednoznačně identifikován pomocí IP adresy. Pokud se chceme zbavit určitých informací, můžeme zablokovat jejich odesílatele. Na základě informací, které odesílatel rozesílá, určíme zdali je nežádoucí, zjistíme jeho IP adresu a zablokujeme jej.

3.1.1 Blacklisty

Blacklist je seznam odesílatelů identifikovaných pomocí IP adres nebo URL, kteří jsou zablokováni. Kvůli tomu bývají blacklisty někdy také nazývány *blocklisty*.

Blacklisty jsou vytvářeny manuálně nebo poloautomaticky zaměstnanci firem zabývajících se tvorbou software pro filtrování webových stránek. Pro zjednodušení často použijí webový vyhledávač jako Google, v němž provedou vyhledání výrazů a frází jako *„live sex chat rooms“*. Seznam výsledků je poté očištěn o adresy vzdělávacích (`.edu`) nebo vládních (`.gov`) stránek a další chybně zahrnuté adresy manuálně odstraní zaměstnanci producenta. Zbytek vytvoří základ pro blacklist, k němuž jsou časem přidávány další položky.

3.1.1.1 DNS blacklisty

Typickým příkladem blacklistů jsou *DNS blacklisty*, které využívají systému Domain Name System (DNS). Tyto blacklisty je možné integrovat takřka s jakýmkoli poštovním serverem. Uplatnění na straně poštovního serveru (MTA[11]) je vhodné z toho důvodu, že server zná IP adresu odesílatele již na začátku přenosu zprávy, takže při odmítnutí spojení nevznikají žádné náklady na přenos zprávy. Filtrování přímo na serveru je proto velice rychlé.

DNS blacklisty poskytují službu, která umožňuje ptát se, zdali je určitá IP adresa na blacklistu. Evidují servery, které nedodržují pravidla stanovená protokoly pro zasílání e-

10 Transmission Control Protocol/Internet Protocol
11 Mail Transfer Agent

mailů. Velkou část blokovaných adres tvoří anonymní otevřené proxy servery nebo *open relays*, což jsou servery nakonfigurované takovým způsobem, jenž umožňuje přeposílání spamu.
Pokud chce poštovní server zjistit, zdali je určitá adresa vedena na blacklistu, připojí ji před adresu DNS blacklistu. Pokud jde o IP, adresu, před spojením nejprve převrátí pořadí částí (oktetů) IP adresy. Takže pokud je hledaná adresa `169.254.100.5`, je převrácena na `5.100.254.169` a připojena k adrese DNS blacklistu jako třeba `spam.services.net`. Pokud má jméno `5.100.254.169.spam.services.net` záznam v DNS, pak je zkoumaná IP adresa na blacklistu.
Tento způsob distribuce blokovaných adres poprvé v praxi uplatnil Paul Vixie v *Realtime Blackhole Listu* (RBL). Nyní existují stovky DNS blacklistů, z nichž některé jsou vedeny profesionálně, jiné představují nepříliš kvalitní osobní seznamy.[12]

3.1.1.2 Behaviorální blacklisty

IP adresy lze relativně snadno měnit a v některých případech jsou dokonce přidělovány dynamicky. Co je ale poměrně neměnné, je chování spammerů. U nich lze vysledovat určité *vzorce chování* - jako je rozesílání nadměrných objemů zpráv - které nejsou přítomny u legitimních odesílatelů.

Na tomto předpokladu jsou založeny *behaviorální blacklisty*. IP adresa se na tento blacklist dostane, pokud je její chování v rozesílání pošty identifikováno s často se vyskytujícími se vzory u šíření spamu. Tyto seznamy mohou být distribuovány podobně jako DNS blacklisty nebo prostřednictvím peer-to-peer (P2P) sítí (Ramachandran, 2007).

3.1.2 Whitelisty

Zatímco blacklisty představují seznamy nežádoucích odesílatelů, whitelisty slouží k evidenci odesílatelů *žádoucích*. Jejich účelem je shromažďovat důvěryhodné kontakty. Mohou se také uplatnit jako *seznamy výjimek* pro hodnotné informační jednotky, které se však vyskytují v takovém kontextu, který by způsobil jejich zařazení na blacklist.

Pokud uživatel zasílá zprávu zákazníkovi poskytovatele whitelistů, dostanu ji zpět s tím, že je třeba, aby potvrdil svou identitu (tzn. že je odesílatelem člověk a nikoli automat). Pokud projde tímto ověřením, e-mail bude doručen a adresa odesílatele se přesune na whitelist.
Pokud je adresa odesílatele na whitelistu, e-mail nemusí procházet dalšími stupni kontroly, čímž se velice získá na rychlosti. Díky tomu jde o velmi efektivní metodu pro zastavení spamu.
Má však nesporné nevýhody v tom, že vyžaduje na odesílateli provedení akce. Při používání whitelistu může dojít k tomu, že někteří legitimní odesílatelé budou zablokování z toho důvodu, že nechtějí kliknout na verifikační adresu nebo že na to prostě zapomněli. Další možností je, že jim výzva k autentizaci nebyla vůbec doručena, protože byla jejich

12Například Spamhaus Block List, Composite Blocking List, Relay Stop List, Open Relay Database (ORDB), Not Just Another Bogus List, Easynet Lists, SpamCop BL a mnoho dalších.

filtrovacím systémem označena za spam.

3.1.3 Greylisty

Greylisty jsou seznamy odesílatelů, které je třeba prověřit. Prověření je založeno na dodržování přijatých standardů komunikačních protokolů. Původci spamu se totiž často vyznačují nestandardním chováním, kdy z důvodu rychlosti nerespektují pravidla e-mailové komunikace.

Spam je obvykle rozesílán specializovaným software, který je zaměřen na dosažení co největší rychlosti rozesílání, takže si nevšímá chybových hlášení. Komunikace mezi odesílatelem a příjemcem (resp. jejich servery) probíhá v předvídatelné posloupnosti kroků (příkazů), proto kvůli rychlosti šiřitelé spamu někdy posílají všechny příkazy najednou, aniž by čekali na odpověď.

Greylisting může být na serveru implementován tak, že tehdy, když je přijata zpráva neznámého původu, odešle poštovní server status `temporary error` a uloží si IP adresu odesílatele včetně dalších detailů o spojení. Programy postupující podle standardů se mají pokusit e-mail znovu odeslat po přiměřené prodlevě. Programy pro rozesílaní spamu však z důvodu rychlosti vrácené zprávy znovu neposílají. Pokud se odesílatel pokusí zprávu znovu doručit v přijatelné době, server ji přijme bez zdržení stejně jako všechny následující e-maily od tohoto odesílatele.

V praxi jsou tímto způsobem filtrování 4% legitimních e-mailů pozdržena přibližně o hodinu, přičemž je však zablokována většina příchozího spamu.

3.1.3.1 Techniky výzva-odpověď

Techniky výzva-odpověď jsou používány v implementacích whitelistů a greylistů. Slouží k ověření odesílatele. Může se jednat o variace na Turingův test, sloužící k rozlišení lidí a automatů.

Tyto metody jsou rozšířeny například pro blokování komentářového spamu. Podmínku pro ověření často představuje CAPTCHA[13] kdy je odesílatel požádán o akci, kterou dovede pouze člověk, jako je opsání znaků z obrázku nebo cokoli jiného, co nelze jednoduše naučit automaty.

Podobně fungují i další metody, které využívají odlišností chování automatů. Zajímavým způsobem, jak zamezit šíření komentářového spamu, je použití skrytého pole, jehož vyplnění kvalifikuje komentář jako spam. Toto pole v odesílacím formuláři je viditelné ve zdrojovém kódu webové stránky, avšak v prohlížeči je před uživatelem pomocí kaskádových stylů (CSS) skryto. Pro uživatele, jejichž prohlížeč nepodporuje CSS, je k poli připojeno varování, aby jej nevyplňoval. Toto varování je však formulováno v přirozeném jazyce, a proto jej automat není schopen adekvátně interpretovat. Protože je takový robot naprogramován tak, aby ve formulářích vyplnil všechna pole, vyplní i toto skryté pole, z čehož je ihned zřejmé, že formulář byl odeslán robotem.[14]

---

13 Completely Automated Public Turing test to tell Computers and Human Apart

14 Tuto techniku poprvé navrhl Aral Balkan - http://aralbalkan.com/

3.2 Filtrování podle výskytu jednotek

Spam lze rozpoznat podle počtu exemplářů, v němž byl rozeslán. Legitimní zprávy jsou rozesílány obvykle malému počtu příjemců, kdežto spam může mít milióny adresátů. Aby bylo možné evidovat počty shodných jednotek, je zapotřebí je nějakým způsobem jednoznačně identifikovat. K tomutu účelu jsou používány kontrolní součty. Pokud se objeví mnoho jednotek se shodným kontrolním součtem, jedná se pravděpodobně o spam.

3.2.1 Porovnávání kontrolních součtů

Kontrolní součet (hash) je kód, který je vypočten z obsahu jednotky, kterou jednoznačně identifikuje. Unikátnost kontrolních součtů nebyla sice dokázána, avšak ani vyvrácena.[15] Hashe zpráv, jejichž výskyt překročí určitou mez, jsou zařazeny do sdílených nebo lokálních databází. Poštovní servery poté mohou kontrolovat příchozí zprávy dotazováním se takové databáze. Pokud se hash přijaté zprávy v databázi vyskytuje, je to nejspíše spam.

Příkladem sdílené databáze hashů je *Distributed Checksum Clearinghouse* (DCC).[16] V této síti je začleněno zhruba 250 serverů, přes které denně projde okolo 300 miliónů e-mailů. Díky velkému pokrytí je spam rychle identifikován.

3.2.1.1 Lokálně sensitivní kontrolní součty

Nevýhodou běžných hashovacích algoritmů je to, že stačí pozměnit jediný znak zdrojové jednotky, aby dostala zcela jiný hash. V praxi stačí do spamu přidat náhodně generované znaky, aby měl každý e-mail unikátní hash.

Z tohoto důvodu byly vyvinuty *lokálně sensitivní hashe*, které drobné odchylky a nepravidelnosti zdrojových jednotek tolerují a vytvoří pro ně shodný nebo podobný hash. Lokálně sensitivní hashovací funkce vypočítávají hashe pro jednotlivé segmenty jednotky (odstavce, řádky), což způsobuje, že by jednotka musela být náhodná v celém svém rozsahu, aby jí byl přidělen unikátní hash.

Modelovou aplikací lokálně sensitivních hashů může být například Nilsimsa. Nilsimsa je velice odolná vůči náhodně přidaným znakům (může být přidáno i 250% náhodného textu). Nahrazování synonym ve zprávě podle slovníku je z hlediska tvůrce spamu účinnější, ale přesto ani s touto technikou nemá Nilsimsa problém.

Šanci na únik má však spam, který používá náhrady znaků za vizuálně podobné - např. „s3cur1ty“namísto „security“. Pokud je tímto způsobem substituováno více než 20% znaků zprávy, není odhalena jako identická. Avšak tato míra obvykle postačí, protože větší podíl nahrazených znaků má silně negativní vliv na čitelnost a srozumitelnost zprávy, což jí činí z hlediska spammera neužitečnou.

3.3 Analýza parametrů

Při snaze o oddělení nežádoucích jednotek lze sledovat relativně „drobné“atributy, které však vypovídají v prospěch nebo neprospěch jednotky. Těmto parametrům jsou přidělena bodová

15 Zjišťování jednoznačnosti probíhá na distribuované infrastruktuře BOINC v testu *SHA-1 collision test*.

16 http://www.rhyolite.com/dcc/

hodnocení. Kladná, pokud indikují nežádoucí jednotku, záporná, pokud se vyskytují často u legitimních jednotek. Každá jednotka projde sadou testů na tyto atributy a po sečtení dílčích bodových hodnocení dostane celkové skóre, na jehož základě se rozhodne, ke které skupině jednotka přísluší.

Jednotka bude označena jako nežádoucí v tom případě, kdy překročí určitý předem nastavený práh. Ten obvykle nastavuje správce poštovního serveru nebo přímo koncový uživatel podle toho, jakou úroveň přísnosti chce zvolit. Nízký práh sice s velkou pravděpodobností zablokuje většinu spamu, ale může přibrat také několik legitimních e-mailů. Zato při vysokém prahu určitá část spamu projde, ale uživatelé se nemusí tolik obávat, že by přišli o legitimní e-maily. Aplikace pro filtrování e-mailů na základě analýzy jejich parametrů využívají toho, že spam je typický tím, že nedodržuje standardy, které stanovují RFC (Request for Comments) pro formát e-mailů. Spam obvykle nerespektuje formální náležitosti, které jsou pro e-maily vyžadovány. Například může postrádat prázdný řádek mezi hlavičkou a tělem zprávy, mít prázdnou položku `From:` identifikující odesílatele nebo chybné kódování znaků.

Stejně tak kvůli svému reklamnímu účelu se spam obvykle vyznačuje určitými předvídatelnými rysy, protože se musí prezentovat tak, aby zaujal příjemcovu pozornost a motivoval jej k nákupu propagovaného výrobku nebo služby. Proto spam často nadužívá kapitálky nebo křiklavé barvy, obsahuje výrazy jako „No prescription needed“nebo obsahuje odkazy, které vypadají, jako by vedly na zabezpečené připojení (HTTPS).

Ve spamovém e-mailu musí být odkaz na stránky prodejce nabízeného produktu nebo služby. Pokud by odkaz přítomen nebyl, postrádal by spam možnost zpěněžení. Díky tomuto atributu lze filtrování založit na zjišťování přítomnosti blacklistovaných adres v těle zprávy.

Zajímavým návrhem je, aby byly za spam označovány všechny e-maily, v nichž se vyskytuje fráze *„Click here“*. Podle anonymního autora knihy *Inside the SPAM cartel* by se tak mohlo dosáhnout až 75% úspěšnosti, poněvadž jistě existuje jen málo legitimních e-mailů, které tuto frázi obsahují (Spammer X, 2004).

Některé atributy spam získal v reakci na fungování e-mailových filtrů. Ve snaze zamaskovat druh zprávy přidávají původci spamu např. 80–90% prázdných řádků, zakódovávají části textu v datovém formátu Base64 nebo klíčová slova spamu jako „cialis“, „levitra“, „soma“, „valium“nebo „xanax“zapisují pomocí *děravých* verzí obsahujících znaky navíc.

Všechny tyto atributy však napovídají systémům filtrování informací, že se jedná o nevyžádanou poštu, a proto podle nich lze s poměrnou jistotou spam oddělit.

3.4 Finanční podmínky

Jedním z atributů informační jednotky může být také peněžní částka, které je k ní přiložena.

Jedním z důvodů, proč je e-mailový spam tolik rozšířen na rozdíl od nevyžádané běžné pošty, je ve vzniku minimálních nákladů, které jsou spojeny s jeho zasíláním. Proto bylo několikrát navrženo, aby byly e-maily zpoplatněny minimální částkou, která by byla pro běžného uživatele zanedbatelně nízká, avšak pro původce spamu, který rozesílá e-maily v milionových objemech, by představovala výrazné omezení. Například za každý odeslaný e-mail by odesílatel musel zaplatit 0,001 dolaru. To by při malých objemech pošty, kterou odesílá běžný uživatel, nepředstavovalo problém. Avšak pro spammera, který rozesílá běžně

několik miliónů e-mailů za den, by toto zpoplatnění vytvořilo finanční bariéru[17] (Costales, 2005).

Lessig a Resnick navrhují, aby ke každému e-mailu byla připojená určitá minimální peněžní částka v elektronické podobě, tak, že pokud by příjemce nebyl se zprávou spokojen, peníze by si nechal. Pokud by se jednalo o legitimní e-mail, předpokládají, že příjemce by peníze vrátil. Pokud by k e-mailu nebyla přiložena dostatečná částka, byl by automaticky odmítnut se sdělením minimální částky (Lessig, 1998).

Tato řešení se však neujala, neboť mají podobně jako i jiné způsoby filtrování negativní dopad nejen na nežádoucí spam, ale také na legitimní komunikaci.

17Viz např. http://www.goodmailsystems.com

4 Filtrování založené na obsahu

Metody filtrování založeného na obsahu se snaží rozhodnout o tom, zdali je informační jednotka žádoucí či nikoli přímo na základě jejího obsahu. Tento způsob filtrování provádějí lidé, když se snaží z obsahu informace vyvodit, zdali je pro ně žádoucí nebo ne. Počítačové nástroje se do jisté míry snaží tento postup napodobit. Filtrování spamu se může učit z následnosti akcí uživatele analyzujícího, zdali se jedná o spam či nikoli. V tomto případě se informační filtry snaží imitovat pochody v lidské mysli, které vedou k tomu, že je určitý e-mail identifikován jako spam (Han, 2008).

Filtrování založené na obsahu obvykle bere v potaz obsah a relevanci informace, ale již ne její kvalitu. Ukazuje se že kvalita informace souvisí s její popularitou a známostí mezi uživateli, a proto jsou pro její odhalení lepší systémy filtrování založeného na spolupráci.

Tento druh filtrování se hodí obzvláště pro doporučování méně známých jednotek uživatelům s vyhraněným vkusem, kdy není dostatek informací o ostatních uživatelích, avšak obsah jednotek lze efektivně zpracovat. Nehodí se do oblastí, kde je obsahu velmi málo nebo jej lze omezeně reprezentovat ve strojově čitelné formě (např. doporučování restaurací).

Pro filtrování založené na obsahu je klíčová kvalitní reprezentace informačních jednotek. Používá se proto především pro textové dokumenty, které lze převést do numerické nebo symbolické struktury, s níž lze dále počítačově manipulovat. Důležité je, aby byl k dispozici plný text jednotky nebo alespoň její abstrakt. Obsah vizuálních, zvukových či multimediálních jednotek prozatím není možné efektivně strojově vyhodnocovat, ačkoli se již v této oblasti několik způsobů objevilo.[18]

Reprezentace obsahu je porovnávána s uživatelskými profily k určení vhodnosti a relevance jednotky pro daného uživatele. Zjednodušeně lze uživatelské profily chápat jako množiny klíčových slov, které určitým způsobem vystihují informační potřeby uživatelů. Lehce pokročilejší možností je určovat profil uživatele podle jednotek, které označí za zajímavé a relevantní. Pro určení míry podobnosti je zapotřebí převést uživatelský profil do podobné formy, jakou má reprezentace jednotek.

Jakmile jsou jak uživatelé, tak jednotky, zachyceny v podobné struktuře, lze provádět jejich porovnávání a zjišťovat nejvhodnější jednotky pro daného uživatele. Filtrování založené na analýze obsahu je však metodou nejvíce náročnou na výpočetní výkon, a proto je vhodné, aby jednotky nebyly vyhodnocovány na základě požadavku v reálném čase, ale spíše dávkově v off-line režimu.

4.1 Textová analýza

Filtrování založené na obsahu využívá k vyhodnocování informačních jednotek postupy textové analýzy. Zkoumá výskyty různých jazykových jevů, jejich kontext a použitou syntax. Automatizované metody filtrování informací však kladou důraz spíše na jevy, které lze kvantifikovat. Vyhodnocování kvalit jednotek je pak obvykle vyvozováno z analýzy těchto jevů.

Příkladem může být hledání vzorů v textu. K tomu lze využít například *regulární výrazy*,

18 Například vyhodnocování podílu tělových barev na fotografiích pro odhalení pornografického obsahu.

které slouží k testování na shodu vzorů. Ty jsou však schopny analyzovat pouze *syntaktickou strukturu* jednotky, zatímco při filtrování založeném na obsahu jde především o *strukturu sémantickou*. Proto se regulární výrazy hodí spíše pro testování, zdali je určitý text v požadovaném formátu (např. e-mailová adresa).

4.1.1 Normalizace

Před prováděním analýzy obsahu jednotky je vhodné její reprezentaci normalizovat. Díky tomu lze ušetřit výpočetní kapacitu a zvýšit účinnost. Typickým postupem vedoucím k normalizaci je odstranění nevýznamových slov jako jsou předložky, spojky nebo částice.

Dále je možné se zbavit nejčastěji se vyskytujících slov v daném jazyce. V praxi se to provádí tak, že se z korpusu daného jazyka získá např. 100 nejfrekventovanějších výrazů, které jsou pak z analyzovaných jednotek odstraněny.

Jedním z cílů normalizace je odstranění náhodně generovaných znaků, odchylek a ostatního šumu. K tomuto účelu je využívána *bayesovská redukce šumu*, která se učí ze slohu a formátu nežádoucích jednotek (spamu) a snaží se vysledovat vzory v používání náhodných znaků, frází, mezer nebo značek jazyka HTML, které byly nesprávně použity pro zmatení filtrů.

Následovat může *lemmatizace* slov jednotky, která spočívá v jejich redukci na kořen slova.

Pomocí těchto způsobů lze dosáhnout toho, že v reprezentaci jednotky budou zachyceny především ty výrazy, pomocí nichž lze specifikovat její obsah. Zdůraznění významu výrazů, které dobře popisují obsah jednotky, lze provést pomocí metody *Term Frequency×Inverse Document Frequency*.

4.1.1.1 Term Frequency×Inverse Document Frequency

Term Frequency×Inverse Document Frequency (TF×IDF) je metoda sloužící k určení váhy výskytů slov. Vychází z předpokladu, že význam výskytu výrazu v jednotce je tím vyšší, čím méně se výraz vyskytuje u ostatních jednotek. Tato technika slouží k zvýšení váhy vzácných a specifických výrazů a snížení váhy četných a obecných výrazů, čímž se dosáhne praktičtější reprezentace jednotky (Morgan, 2009a,b).

TF×IDF je stanoveno jakožto součin četnosti výrazu v jednotce (TF) a převrácené četnosti v kolekci všech jednotek.

$$TF \times IDF = TF \cdot IDF$$

Četnost výrazu (TF) je část, kterou v jednotce zabírají výskyty daného slova.

$$TF = C/T$$

- C …počet výskytů daného výrazu v jednotce
- T …celkový počet všech výrazů v jednotce

Převrácená četnost výrazu (IDF) znázorňuje ojedinělost slova v kolekci všech jednotek.

$$IDF = D/DF$$

- D …celkový počet jednotek v kolekci
- DF …počet jednotek v kolekci, v nichž se vyskytuje daný výraz

Tento základní vzorec bývá často upravován s použitím přirozeného logaritmu.

$$TF \times IDF = C/T \cdot ln(D/DF)$$

4.2 Automatická klasifikace

Pokud jsou k dispozici algoritmicky zpracovatelné reprezentace jednotek, lze provádět jejich *automatickou klasifikaci*. Jejím účelem je roztřídit jednotky do skupin podle obsahu jako je např. spam a legitimní pošta. Automatická klasifikace může sloužit k určení jazyka jednotky nebo k jejímu zatřídění podle tématu do složek v e-mailové schránce.

Při klasifikace jde o zařazení jednotky do kategorie. V případě filtrování informací máme obvykle 2 kategorie: žádoucí a nežádoucí. Tyto kategorie jsou však pro každého uživatele definovány jiným způsobem. Žádoucí znamená relevantní vůči informačním potřebám uživatele, proto automatická klasifikace usiluje o předmětové zařazení jednotky.

Jednotky jsou klasifikovány podle *rysů*, jež obsahují. Rysy mohou mít různou velikost - jednotlivá slova, fráze nebo témata. Velikost rysu by měla být určena podle jeho užitečnosti. Čím je rys větší, tím je vzácnější, a proto lépe charakterizující. Malé rysy jako jsou jednotlivá slova se naopak vyskytují v mnoha jednotkách a nejsou pro klasifikaci tak užitečné. Vhodné je volit rozsah rysu podle kontextu - například pro údaj „autor“bude rysem celé jméno autora a nikoli jednotlivá slova, z nichž se skládá.

Při automatické klasifikaci jsou nejprve shromážděny všechny výskyty rysů v jednotkách. Výskyty jsou rozděleny na výskyty v nežádoucích jednotkách a výskyty v žádoucích jednotkách. Následuje převod výskytů na pravděpodobnost. Například slovo „viagra“se objeví 4× v nežádoucích jednotkách a pouze 1× v žádoucích, takže pravděpodobnost je 0,8 pro nežádoucí jednotky a 0,2 pro žádoucí. Protože můžou být tyto výskyty na počátku klasifikace jednotek zkreslené, je vhodné vycházet z předpokládané neutrální pravděpodobnosti 0,5. Jinou možností je použít jako výchozí pravděpodobnosti z dříve nashromážděných dat ostatními uživateli.

Pravděpodobnost spoluvýskytu výrazů se vypočítá vynásobením jejich pravděpodobností. Tedy, pokud je pravděpodobnost slova „viagra“v nežádoucích jednotkách 0,8 a pravděpodobnost slova „money“v nežádoucích jednotkách 0,2, tak je pravděpodobnost spoluvýskytu „viagra“a „money“v nežádoucích jednotkách rovna $0{,}8 \cdot 0{,}2 = 0{,}16$. Pravděpodobnost příslušnosti ke sledované kategorii pro celou jednotku získáme na základě dílčích pravděpodobností pro její rysy.

4.2.1 Adaptivní algoritmy

Klasické metody filtrování založeného na obsahu provádějí jednorázovou statistickou analýzu informační jednotky, většinou s ohledem na klíčová slova. Na rozdíl od těchto metod se adaptivní algoritmy učí z dříve nashromážděných dat. „Učení“probíhá zkoumáním množiny vstupů a očekávaných výstupů. Výkonnost adaptivních algoritmů se zvyšuje s časem, po který jsou nasazeny. Zpočátku nejsou příliš účinné, avšak později, jakmile mají k dispozici dostatek trénovacích dat, mohou dosáhnout velmi dobrých výsledků.

Učení probíhá tak, že uživatel systému filtrování informací poskytuje příklady správné klasifikace a *zpětnou vazbu* pro klasifikaci, kterou systém provádí automaticky. Například tím, že uživatel označí určitý e-mail jako spam, ukáže systému příklad, podle jehož charakteristik lze identifikovat další spamové zprávy. Při filtrování spamu na základě obsahu jsou obvykle shromažďovány legitimní e-maily (pozitivní příklady). Filtrovací algoritmus se pak „učí“podle jejich obsahu rozpoznávat další legitimní e-maily. Stejně tak je možné sbírat

spam (negativní příklady) a podle jeho obsahu určovat další spam. „Učením“se postupně upravují pravděpodobnosti výskytu rysů ve sledovaných skupinách (např. spam/legitimní e-maily), takže s růstem objemu shromážděných jednotek roste přesnost filtrování.
Adaptivní algoritmy využívají například bayesovské pravděpodobnostní metody, neurální sítě nebo rozhodovací stromy. Nejrozšířenějším postupem je použití naivního bayesovského klasifikátoru.

4.2.1.1 Naivní bayesovský klasifikátor

Naivní bayesovský klasifikátor vychází z bayesovské analýzy, která se zakládá na myšlence, že pravděpodobnost jevu lze určit sledováním okolností, za nichž nastal, a četnosti správných rozhodnutí o jeho výskytu. Jinými slovy pravděpodobnost jevu je *podmíněna* jevy předešlými.

Jméno získala bayesovská analýza podle reverenda Thomase Bayese, který v roce 1776 navrhl metodu výpočtu pravděpodobnosti události na základě matematického a statistického rozboru předchozích událostí.
Bayesův teorém lze zapsat:

$$P(A|B) = P(B|A)\cdot P(A)/P(B)$$

- P(A) …pravděpodobnost A
- P(B) …pravděpodobnost B
- P(A|B) …pravděpodobnost A, pokud nastalo B
- P(B|A) …pravděpodobnost B, pokud nastalo A

Ze základu, který poskytuje bayesovský teorém pro výpočet podmíněné pravděpodobnosti, jsou odvozeny algoritmy, jenž využívají bayesovské klasifikátory.

Bayesovské klasifikátory fungují nezávisle na účelu, pro nějž jsou využívány. Lze je aplikovat na klasifikaci textových jednotek podle jakéhokoli znaku. Například lze klasifikátor naučit rozpoznávat jazyk, v němž je obsah jednotky napsán.
Proto je výhodou bayesovských filtrů jejich přizpůsobitelnost konkrétnímu druhu spamu, který příjemce na rozdíl od ostatních dostává. Jejich nevýhodou je nutnost nejprve se naučit spam rozpoznávat, což od uživatele vyžaduje aktivní přístup a trpělivost.
Naivní bayesovský klasifikátor je „naivní“nazýván, protože předpokládá, že pravděpodobnosti jednotlivých rysů jsou na sobě nezávislé. Což je očividně mylný předpoklad, jelikož výskyty určitých rysů spolu mohou souviset, a pokročilejší automatické klasifikátory již toto berou v potaz.
Přes tento zřejmý nedostatek se ve skutečnosti naivní klasifikátory překvapivě prokázaly jako vysoce učinné. Navíc dosahují přijatelné rychlosti a jejich implementace je poměrně jednoduchá. Producenti antispamových řešení založených na bayesovských metodách tvrdí, že jejich produkty při základu přibližně 4000 e-mailů, které byly označeny jako spam, dosahují při rozpoznávání spamu až 99,9% přesnosti. Pro srovnání se uvádí, že člověk dosahuje při rozlišení spamu a legitimních e-mailů přesnosti 99,84% (Yerazunis, 2004).

4.2.1.2 Fisherova metoda

Na podobných principech jako bayesovské metody je založena alternativní Fischerova

metoda. Liší se zejména ve způsobu, jakým nakládá s pravděpodobnostmi dílčích rysů jednotky a tím, že je porovnává s množinou náhodných pravděpodobností. Představuje komplexnější a výpočetně náročnější variantu bayesovského klasifikátoru, avšak dosahuje lepších výsledků.

4.2.1.3 Rocchiův algoritmus

Mezi další metody, které využívají adaptivní algoritmy, patří Rocchiův algoritmus pocházející z roku 1971. Bývá používán k výpočtům se zpětnou vazbou, kterou uživatel poskytne pro hodnocení relevance jednotek, které mu jsou prezentovány systémem filtrování informací. Rocchiův algoritmus poskytuje nástroj k modifikaci vektorové reprezentace uživatelského profilu začleněním informací získaných zpětnou vazbou.

4.3 Analýza kontextu

Starší metody filtrování založeného na obsahu kontextuální informace nejprve ignorovaly. Systémy pro blokování webových stránek byly velice primitivní, kdy jediný nežádoucí výraz způsobil zablokování celé stránky. Při běžné statistické analýze klíčových slov, kterou tyto systémy prováděly, docházelo ke ztrátě informací obsažených v kontextu. Jednotky byly chápány jako neuspořádané množiny znakových řetězců, čímž byly zcela opomíjeny syntaktické a sémantické vztahy mezi slovy.

To vyvolalo kritiku ze strany organizací zabývajících se prosazováním svobody projevu. Následující generace software již disponovala více schopnostmi k rozpoznávání kontextu, v němž se hledané výrazy nacházely. Avšak jak narůstalo pokrytí blokovaných výrazů, bylo pro filtrovací software stále obtížnější správně analyzovat jejich kontext.

Začlenění kontextuálních informací lze v zásadě řešit dvěma způsoby. Prvním je přesunout pozornost z jednotlivých výrazů na významové koncepty, které jednotka obsahuje, a použít například *latentní sémantické indexování*.

Druhou možností je rozšířit velikost rysů, které jsou při analýze obsahu sledovány. Tento způsob předpokládá platnost *distribuční hypotézy*, která považuje slova vyjádřená v podobných kontextech za sémanticky příbuzná. V takovém případě je při analýze jednotka čtena skrze posuvný průzor, v němž se nachází několik slov v bezprostředním okolí, jejichž vzájemné vztahy jsou brány v potaz. Poté, co je tato skupina vyhodnocena, průzor se posune dále o jedno slovo. Jednotlivé pohledy se tak v následujících krocích překrývají, čímž je dosahováno určité návaznosti analýzy. Pro efektivní analýzu je zapotřebí zvolit vhodnou velikost průzoru, jelikož při malé velikosti dochází ke ztrátě kontextuálních informací, a naopak při nadměrné velikosti je obtížné a výpočetně náročné jednotlivé pohledy analyzovat. Tyto techniky bývají nazývány *šindelové*, protože význam analyzovaných skupin sousedících výrazů se podobně jako u šindelů na střeše překrývá.

Metody, které se snaží vyhodnocovat kontextuální informace, mohou posloužit k rozlišení polysémických výrazů, homonym a celkovému snížení víceznačnosti. Šindelové techniky se navíc používají k určování duplicitních dokumentů (Hasnah, 2006).

5 Filtrování založené na spolupráci

Metody filtrování založeného na spolupráci se snaží rozhodnout, zdali je informační jednotka žádoucí či nikoli na základě rozhodnutí uživatelů.

Systémy filtrování založeného na spolupráci jsou úspěšné, protože modelují běžné sociální situace, kdy se obracíme na své přátele pro jejich hodnocení nám neznámých jednotek.

Například pokud přátelé hodnotí určitý film kladně, je pravděpodobné, že o něj budme mít větší zájem. Platí to i naopak, znechucení známých může od rozhodnutí jít film shlédnout odradit. Můžeme zjistit, že jednomu našemu známému se líbí filmy, které se líbí také nám, a proto budeme jeho doporučením přikládat větší váhu. Naopak časem rozlišíme, kteří přátelé doporučují filmy, které se nám nelíbí, a kteří doporučují úplně všechny filmy. Nakonec přijdeme na to, kterým přátelům se vyplatí naslouchat. Stejným průběhem prochází výběr uživatelů v systémech filtrování založeného na spolupráci.

Díky internetu však lze shromáždit doporučení tisíců lidí, kteří se navzájem neznají. S pomocí efektivních algoritmů můžeme na základě těchto údajů vytvářet doporučení, která mohou být stejně užitečná jako od známého.

Filtrování založené na spolupráci se obzvlášť hodí tehdy, kdy jednotky nelze dobře popsat slovy.

Filtrování založené na spolupráci lze rozdělit na sociální a kolaborativní filtrování. Sociální filtrování je založeno na explicitně stanovených vazbách mezi uživateli. Rozšířeným způsobem zachycení takových vazeb je označení uživatele jakožto *„přítele“*. Naproti tomu v kolaborativním filtrování jsou většinou vazby mezi uživateli vypočítávány, čímž vznikají *„sousedé“*.

Ve skutečnosti jsou však prvky sociálního a kolaborativní filtrování s úspěchem kombinovány, jelikož si navzájem dokážou kompenzovat své slabší stránky, jako je například „studený start“kolaborativního filtrování.

5.1 Sociální filtrování

Sociální filtrování vyzdvihuje sociální vazby mezi uživateli nad jejich podobnost. „Sociální“prvky filtrování zahrnují reputaci, důvěru a motivaci k recipročnímu sdílení.

Sociální filtrování spočívá ve využívání úsudku ostatních lidí pro volbu informací. Doporučení informačních jednotek se zakládají na údajích od důvěřovaných uživatelů. Mimo prostředí internetu můžou být příkladem filtrování založeného na reputaci časopisy, kdy jejich čtenáři důvěřují editorům v tom, že vyberou jen kvalitní články.

Označením dobrých zprostředkovatelů informací (přátel, autorit, odborníků) vyjadřujeme zájem o informace, které buď oni sami poskytují nebo zprostředkovávají. Zprostředkovatel informací je osoba, která má znalosti o informacích a aktivně tyto informace šíří mezi svými známými. Tuto funkci umožňují systémy filtrování informací plnit zasíláním přímých doporučení nebo sdílením jednotek. Mimo aktivní vyhledávání nových zajímavých informací tak mohou uživatelé využít ostatních lidí k nalézání a filtrování informací pro ně.

Výzkumy ukázaly, že neformální síť spolupracovníků, kolegů nebo přátel je jedním z nejefektivnějších způsobů šíření informací uvnitř organizace (Kautz, 1997).
Využívání sociálních vazeb se však osvědčilo především spíše v úzce vymezených doménách. Pokud systém zprostředkovává obsah z mnoha oblastí, je lepší informační jednotky nejprve hodnotit podle obsahu a významu sociálních vazeb dávat menší váhu. Podobnost mezi uživateli se totiž v různých oblastech může značně lišit. Uživatelé mohou shodnout, kdo je uznávaným odborníkem určitého vědního odvětví, ale již spolu nesouhlasí v subjektivně založených oblastech - například která hudební skupina je nejlepší.
Sociální filtrování je založeno na platnosti dvou principů.

1. uživatelům se často líbí informační jednotky, které sdílí jejich přátelé
2. uživatelům se často líbí informační jednotky, které se líbí jejich přátelům

Pokud srovnáme doporučení založená na neznámých podobných uživatelích a podobných uživatelích z řad přátel, jsou doporučení přátel obvykle chápána jako užitečnější a důvěryhodnější. Vztah mezi původcem a příjemcem doporučení má zásadní vliv na to, jak bude doporučení přijato. Vliv má nejenom známost, ale také podobnost v demografických údajích nebo podobné zájmy.

V mnoha sociálních sítích však není kladné hodnocení primárně vyjádřením toho, že se uživateli daná jednotka líbila. Uživatel mnohdy chce spíše, aby tuto jednotku ostatní (jeho přátelé) také viděli (Scoble, 2009).
Skrze sociální sítě obvykle to, co je hodnotné, „probublá“k pozornosti uživatelů. Prostřednictvím přeposílání a virálního šíření se tak informace rozšíří na základě své kvality. Čím delší je řetězec lidí, přes než se informace dostala, tím je pravděpodobnější, že informace je kvalitní. Využívá se tím memetický charakter informace v síťovém prostředí (Huleatt, 2009a).
Rozšířeným druhem aplikace sociálního filtrování jsou systémy sdílení záložek jako *Digg*[19] nebo *Delicious*[20]. Sdílené záložky umožňují využívat ostatní uživatele jakožto „lidské filtry“. Tím, že uživatelé určitý dokument vyberou a označí ho štítkem, umožní ostatním použít jejich záložky jako doporučení. Datové struktury, které tvoří podklad systémů sdílení záložek, se nazývají *folksonomie*. Sestávají se z množin uživatelů, štítků, zdrojů a vazeb mezi štítky a zdroji.
Kvůli využívání štítků (tagů) je tento druh sociálního filtrování označován jako *kolaborativní tagování*. V prostředí internetu může jít o velmi efektivní způsob klasifikace obsahu, jelikož dovoluje zapojení mas uživatelů, čímž může získat pokrytí velkého objemu informačních jednotek. Systémy sdílených záložek pak lze využít jako alternativu internetových vyhledávačů. Stačí hledané téma najít mezi štítky, které uživatelé použili.
V systémech sociálního filtrování lze vytvářet různě zaměřené *sub-komunity*, čímž je možné se vyhnout „tyranii většiny“, která způsobuje, že systém primárně uspokojuje informační potřeby většinových uživatelů. Příkladem tyranie většiny je třeba titulní stránka služby Digg, kam se dostanou jen ty jednotky, které si oblíbilo nejvíce uživatelů, což se podobá seznamu bestsellerů.

19 http://digg.com
20 http://delicious.com

5.2 Kolaborativní filtrování

Kolaborativní filtrování je proces filtrování a hodnocení jednotek pomocí názorů dalších lidí. Principy kolaborativního filtrování vychází z předpokladu, že lidé s podobným vkusem hodnotí informační jednotky podobně. Kolaborativní filtrování kombinuje „know-how“podobných uživatelů.

Striknί kolaborativní filtrování je založeno na předpokladu, že k funkčnímu doporučování nejsou sociální vazby mezi uživateli zapotřebí, jelikož si systém vystačí s daty o jejich podobnosti.

Kolaborativní filtrování nevyžaduje obsah. Umožňuje vzít v potaz rysy jednotek, které nelze jednoduše automaticky extrahovat, jelikož hodnocení těchto rysů vytvářejí lidé a ne automaty. Jednotkami, s nimiž systém pracuje, proto může být cokoli, co lze hodnotit - například restaurace, knihy, výtvarné umění, časopisecké články nebo prázdninové destinace. Kolaborativní filtrování vede k objevení neočekávaných jednotek, a proto se hodí k nasazení do oblastí, kde je důležitá novost, jednotky se rychle mění nebo jich je veliké množství.

5.2.1 Historický vývoj

Termín „kolaborativní filtrování“byl poprvé použit v roce 1992 v práci *Using collaborative filtering to weave an information tapestry* (Goldberg, 1992) Davida Goldberga z výzkumného střediska Xerox PARC v Palo Alto, kde byl vyvinut první systém kolaborativního filtrování nazvaný *Tapestry*. Umožňoval ukládat *„anotace“* k dokumentům, jejichž prostřednictvím uživatelé mohli vyjádřit kvalitu jednotky. Aby uživatel získal doporučení jednotek, musel aktivně zadat dotaz. Novější systémy kolaborativního filtrování doporučení nabízejí samy bez nutnosti formulace explicitního „dotazu“.

Limitací starších systémů byl fakt, že požadovaly komunitu lidí, kteří se navzájem znají, jelikož každý uživatel musel vědět, kterým hodnocením může věřit.

Mezi rané automatizované kolaborativní systémy lze zařadit *GroupLens* pro články na Usenetu, *Ringo* pro oblast hudby nebo *Bellcore Video Recommendation* pro oblast filmu.

5.2.2 Úlohy a funkce

Systémy kolaborativního filtrování jsou používány pro řešení typizovaných úloh vedoucích k uspokojení informační potřeby uživatele.

1. Najdi nové jednotky, které by se uživateli mohly líbit.
2. Je jednotka podle hodnocení ostatních uživatelů dobrá nebo špatná?
3. Najdi uživatele, kteří by se uživateli mohli líbit. To může pomoci při vytváření diskusních skupin nebo k propojování uživatelů tak, aby si doporučení mohli vyměnovat sociálně.
4. Najdi jednotky, které by se mohly líbit skupině uživatelů. Například pokud se přátelé chtějí domluvit, na který film jít.
5. Najdi vyváženou směs „nových“a „starých“jednotek. Například uživatel chce občas zajít do restaurací, které již navštívil a hodnotil kladně. Jindy by rád navštívil restaurace, v nichž ještě nebyl.

6. Pomož uživateli s úlohami specifickými pro určitou oblast, určitý kontext. Například doporuč zdroje, které by měl citovat článek, na němž uživatel pracuje, nebo vhodný film pro rodiny s malými dětmi.

Technologickou odpověď na tyto požadavky představují implementované funkce systémů kolaborativního filtrování.

1. *Doporuč jednotky* – vytvoř seznam doporučených jednotek řazených podle relevance. Tento postup je často chápán jako predikce hodnocení, která by uživatel jednotce přiřkl, a prezentace těch jednotek, u nichž je předpokládáno nejlepší hodnocení. Mnoho aplikací však vůbec predikovaná hodnocení nepočítá (např. systém, který používá Amazon.com, zobrazuje průměrné hodnocení ostatních uživatelů).
2. *Předpověz hodnocení jednotky* – předpovídání hodnocení pro každou jednotku může být výpočetně náročnější než doporučení, navíc predikce pro řídce hodnocené jednotky představují velmi obtížný úkol.
3. *Doporuč z vymezené množiny jednotek* – uživatel stanoví kritéria pro žádanou podmnožinu jednotek (např. komedie přístupné dětem do 12 let), z níž poté systém vytvoří doporučení. Pro tento účel je navrhován jazyk podobný SQL, který by disponoval schopnostmi určit žádaný zdroj doporučení a podmnožinu jednotek, z níž by měl být proveden výběr. Dotaz by pak mohl vypadat například takto (Schafer, 2007):

```
RECOMMEND Movie TO User BASED ON Rating FROM
MovieRecommender WHERE Movie.Length < 120 AND
Movie.Rating < 3 AND User.City = Movie.Location
```

5.2.3 Předpoklady

Metody kolaborativního filtrování jsou účinnější pro případy z oblastí, které splňují určité předpoklady. Jsou užitečné pro doporučování poměrně známých jednotek, zejména v komunitách uživatelů společných zájmů, kdy nejsou k dispozici efektivně zpracovatelné informace o obsahu jednotky. Oblasti, v nichž se kolaborativní filtrování uplatní nejlépe obvykle splňují následující podmínky.

1. **Distribuce dat**
   (a) V systému se nachází mnoho jednotek. Pokud je jednotek málo, rozhodování mezi nimi lépe provede uživatel, a nepotřebuje tak podporu automatizovaného systému.
   (b) Ke každé jednotce je k dispozici mnoho hodnocení. Pokud je hodnocení málo, systém nemá dost informací k vytváření užitečných doporučení.
   (c) V systému je více uživatelů než jednotek. Tato podmínka vychází z předchozích požadavků. Pokud by systém například měl doporučovat webové stránky, byl by tento požadavek nesplnitelný. Google uvádí, že indexuje 8 biliónů stránek, což je více než lidí na Zemi (a mnohem více než lidí, kteří mají přístup k internetu).
   (d) Uživatelé hodnotí více jednotek. Pokud hodnotili jen málo jednotek, měla by jejich hodnocení spíše statistický význam.

2. **Perzistence dat**
   (a) Jednotky zůstávají dlouho relevantní. Například v oblasti novinových zpráv trvá jejich zajímavost obvykle pouze několik dní. K efektivnímu fungování kolaborativního filtrování je zapotřebí získat hodnocení mnoha uživatelů, což není v tak krátkém čase dosažitelné, a proto se na doporučování novinek a pomíjivých informací nehodí.
   (b) Vkus a informační potřeby uživatele se mění pomalu a spíše přetrvávají. Systémy kolaborativního filtrování jsou nejúspěšnější v případech, kdy se vkus uživatele nemění rychle (hudba, filmy, knihy). V opačném případě, kdy se vkus uživatele mění často, jsou starší hodnocení méně užitečná (např. oblečení).
3. **Struktura dat**
   (a) Pro každého uživatele se v systému najdou jiní *uživatelé s podobnými potřebami* nebo vkusem. Pokud by vkus uživatele byl natolik specifický, že by jej nesdílel s žádným jiným uživatelem, systém kolaborativního filtrování by mu nic nepřinesl.
   (b) Hodnocení jednotek vyžaduje *osobní vkus*. Kolaborativní filtrování se hodí zejména do oblastí, kde je hodnocení jednotek vysoce subjektivní (např. hudba, film), nebo tehdy, pokud se jednotky vyznačují velkým množstvím objektivně kvalifikovatelných charakteristik, kterým však uživatelé přikládají subjektivní váhu (např. automobily).
   (c) Jednotky jsou *homogenní*, což znamená, že jednotky jsou v objektivně zhodnotitelných kritériích podobné a hlavní rozdíl mezi nimi vzniká až na základě osobního vkusu. Například hudební nosiče jsou podobné ceny nebo délky. Tomuto požadavku však nevyhovují produkty v supermarketu, protože nejsou homogenní - některé jsou levné, jiné velmi drahé.

5.2.4 Algoritmy

Kolaborativní filtrování používá mnoho typů algoritmů. V novějších systémech se často uplatňují pravděpodobnostní algoritmy, které využívají principy pravděpodobnostního počtu. Kolaborativní filtrování uplatňuje postupy využívající bayesovské sítě, analýzu grafů nebo neurální sítě. Algoritmy používané metodami kolaborativního filtrování jsou děleny do dvou skupin.

1. metody založené na *paměti* – berou v potaz a do paměti ukládají všechna hodnocení, jednotky a uživatele, a proto dosahují vyšší přesnosti, avšak nejsou dobře škálovatelné pro reálné aplikace.
2. metody založené na *modelu* – periodicky v offline režimu vytváří shrnutí vzorů v hodnocení. Dosahují sice nižší přesnosti, avšak jsou lépe rozšiřitelné.

5.2.5 Vytváření doporučení

Primární funkcí systémů kolaborativního filtrování je vytváření doporučení jednotek pro uživatele. Doporučení mohou být vytvářena periodickým spouštěním filtrovacího algoritmu nebo v reálném čase na základě požadavku. Vznikají algoritmickým

vyhodnocováním dat, která jsou v systému shromažďována. Při tom se uplatňují metody měření podobnosti nebo vyvozování na základě odpozorovaných pravidel.

5.2.5.1 Zjišťování nejbližších sousedů

Často používané jsou algoritmy určující nejbližší sousedy mezi jednotkami nebo mezi uživateli. Nejbližší sousedé tvoří množinu, jejíž členové jsou si navzájem nejvíce podobni. Na obsahu skupiny nejbližších sousedů jsou poté založena doporučení.

5.2.5.1.1 Nejbližší sousedé mezi uživateli

Nejbližší sousedé mezi uživateli jsou určováni podle hodnocení, která sdílí. Doporučení jsou generována z hodnocení, která jednotce přidělili sousedé uživatele.

Tento přístup je velice rozšířený, ačkoli trpí několika slabinami. Pro uživatele s malým počtem hodnocení mohou vyjít sousedské vazby zkresleně.
Také často opomíjí variace mezi uživateli. Existují „optimističtí“uživatelé, kteří soustavně hodnotí výše než průměrný uživatel. Jsou také „pesimističtí“uživatelé, kteří hodnotí nadměrně kriticky. Přesto mohou jejich hodnocení znamenat totéž. 5 hvězdiček od „optimisty“se může rovnat 3 hvězdičkám od „kritika“a obojí hodnocení lze interpretovat jako „jeden z mých nejoblíbenějších filmů“. Tyto odchylky lze odhalit porovnáním s trendy v hodnoceních uživatele.
Výpočet sousedství je rovněž výpočetně náročný, protože vyžaduje porovnání se všemi ostatními uživateli. Pokud by se k řešení přistupovalo tímto způsobem, výpočetní náročnost by rostla přímo úměrně s počtem uživatelů a jednotek. Proto bylo navrženo několik postupů, které adresují tento nedostatek.

1. z*menšování vzorku* – pro doporučení je brána v potaz pouze omezená podmnožina uživatelů. Způsob, jakým se tato podmnožina volí, je inteligentně navržen tak, aby bylo dosaženo takřka stejné přesnosti jako v případě, kdy jsou zvažováni všichni uživatelé.
2. *shlukování* – shlukovací algoritmy jsou navrženy k rychlému vyhledání sousedů uživatele. Namísto toho, aby byli uživatelé porovnávání jako jednotlivci, je uživatel porovnáván s celými skupinami ostatních uživatelů. Nejprve jsou vytvořeny shluky podobných uživatelů a poté jsou nejbližší sousedé vybráni z nejpodobnějších shluků.

5.2.5.1.2 Nejbližší sousedé mezi jednotkami

Nejbližší sousedé mezi jednotkami jsou podobné sousedům mezi uživateli. Vytváří doporučení na základě podobnosti mezi jednotkami.

Opět se zde vyskytuje vysoký nárok na výkon, pokud by byly počítány korelace mezi všemi jednotkami. V praxi se to řeší výpočtem korelací pouze pro jednotky, které mají více než k hodnocení. Z těchto korelací se může v systému ponechat jen n nejvyšších (nejpodobnějších) pro jednotku. Díky těmto úpravám jsou algoritmy nejbližšího souseda založené na jednotkách poměrně efektivní jak z hlediska využívání paměti, tak nároky na výpočetní výkon. Samozřejmě je však třeba počítat s tím, že odstranění některých korelací zvyšuje náročnost vytváření doporučení pro určité řídce hodnocené jednotky a uživatele s malým počtem hodnocení.

5.2.5.2 Využívání asociativních pravidel

Techniky využívání asociativních pravidel se zakládají na častých vzorech v maticích hodnocení. Například lze vypozorovat, že uživatelé, kteří hodnotili jednotku *A* kladně, hodnotí jednotku *B* také kladně. Z toho lze odvodit pravidlo, které se skládá z vstupní podmínky (jednotka *A* hodnocena kladně) a výsledku (jednotka *B* hodnocena kladně). *Podporu* pravidla představují uživatelé, kteří hodnotili jak vstupní jednotku, tak jednotku výstupní. *Jistotu* pravidla představuje podíl uživatelů, kteří hodnotili v souladu s pravidlem. Doporučení může vycházet z těchto pravidel, teprve tehdy, když je v systému dostatek údajů, z nichž lze tyto pravidla odvodit.

Pravidla lze také extrapolovat z transakčních dat zaznamenávajících chování uživatele. Například lze vytvořit pravidlo, že uživatel, který si koupil gril na dřevěné uhlí bude za měsíc kupovat hasící přístroj.
Asociační pravidla mohou vyjadřovat, že určitý produkt je často kupován s dalšími produkty. Rozšířeným využitím asociačních pravidel je například rozložení zboží na regálech.

5.2.6 Konverzační systémy doporučení

Poté, co jsou pro daného uživatele vytvořena a prezentována doporučení, lze s nimi dále pracovat prostřednictvím zpětné vazby. K tomu slouží *konverzační systémy doporučení*, které vedou s uživatelem konverzaci, během níž jim uživatel poskytuje zpětnou vazbu. Dosahuje se tak optimálního propojení lidského a strojového rozhodování. Vytvořená doporučení jsou postupně s přihlédnutím k reakcím uživatele upravována. Každý krok v konverzaci zpřesňuje výběr doporučení. Objevují se 2 způsoby, jakými se lze konverzačními doporučujícími systémy navigovat.

1. **Navigace pomocí dotazování**
   V případě, kdy navigace v systému probíhá pomocí dotazování, jsou uživateli prezentovány výčty hodnot atributů s žádostí, aby jednu z nich vybral, nebo mohou být uživateli předloženy dotazy s volbou ANO/NE.
2. **Navigace pomocí návrhů**
   Druhý způsob navigace probíhá pomocí návrhů, které jsou uživateli předkládány. Návrhy mají podobu mezi-doporučení, která jsou postupně zpřesňována. Po uživateli může být požadována reakce obvykle v jedné ze tří forem.
   (a) **Hodnocení**
   Uživatel je vybídnut, aby navržená doporučení zhodnotil; např. přiřadil 1–5 hvězdiček.
   (b) **Kritika**
   Uživatel vyjadřuje kritiku navržených doporučení. Výběr omezuje podle určitého rysu (atributu), takže např. může stanovit, že chce doporučit jednotky, které jsou tišší, levnější nebo blíže k jeho bydlišti.
   (c) **Preference**
   Uživatel zvolí jedno z navržených doporučení, které preferuje. Tento způsob získávání zpětné vazby se považuje za vhodný v případech, kdy má uživatel malé znalosti o oblasti, z níž je doporučováno (např. výběr svatebních šatů).

Tím, jak systém postupně získává odezvu uživatele na prezentované návrhy, roste přesnost doporučení, která jsou jím vytvářena. Po určitém, omezeném počtu kroků jsou uživateli představena finální doporučení.

5.2.7 Rizika

Pokud jsou metody kolaborativního filtrování vhodně aplikovány, dosahují velice uspokojivých výsledků s vysokou mírou přesnosti. Na druhou stranu je však zapotřebí počítat s riziky, která se tímto typem filtrování informací pojí.

Patří k nim již zmíněná *tyranie většiny*, kdy jsou uspokojovány primárně potřeby většinových uživatelů s populárním vkusem. Kolaborativní filtrování se také nevyhne *cíleným útokům*. Tyto záměrně vedené akce slouží ke změně chování systému doporučení ve prospěch útočníka a jsou příčinou zkreslených doporučení.

Protože ke své funkci potřebují hodnocení informačních jednotek, musí čelit problémům, které se s nimi pojí, jako jsou přiměřené škály, motivace a pobídky k hodnocení, předpojatí uživatelé vytvářející zkreslující hodnocení, problému nehodnotících uživatelů a dosažení kritického množství uživatelů.

5.2.7.1 Problém studeného startu

Problém studeného startu spočívá v tom, že systémy kolaborativního filtrování nemohou efektivně fungovat, pokud nemají dostatek údajů. Tento stav nastává při spuštění systému, kdy není dostupné minimální množství akumulovaných hodnocení a dalších uživatelských vstupů, které jsou nezbytné pro správnou funkci systému.

Klíčové je „kritické množství uživatelů“a jejich vkladů do systému (hodnocení, komentářů apod.), které představuje okamžik, kdy se v systému nachází dostatek údajů, aby vytvářená doporučení byla kvalitní.

Kolaborativní filtrování obvykle funguje dobře nad hustě provázanými daty. V případě nedostatku dat, tzv. *„řídké matice“*, se jeho výkonnost významně snižuje. Mezi navrhovanými opatřeními se objevuje adaptivní přizpůsobování metody filtrování podle analýzy hustoty údajů, které jsou k dispozici. Tak lze plynule přecházet mezi kolaborativním filtrováním v případě dostatku dat, filtrováním založeným na obsahu v situacích, kdy je dat nedostatek, a hybridními metodami kombinujícími postupy více typů filtrování.

5.2.7.2 Nízká rozmanitost doporučení

Dalším obvyklým problémem, s nímž se mohou setkat uživatelé systémů kolaborativního filtrování, je nízká rozmanitost poskytovaných doporučení. Jedná se o stav, kdy je prvních pár doporučených jednotek natolik podobných, že se mezi nimi uživatelé neumí rozhodnout.

Pokud by se například jednalo o systém pro doporučování dovolené, výsledky odkazující na jediný hotel, třebaže v různé termíny, nepřinesou uživateli žádný užitek. Proto je v rámci podobnosti výsledků třeba zavést vzájemnou rozmanitost.

Řešením může být začlenění náhodnosti pro řazení výsledků nebo uvolnění nároků na podobnost doporučených jednotek.

6 Závěr

Tato práce představuje úvod do oblasti automatizovaného filtrování informací v prostředí internetu. Jejím cílem bylo vytvořit strukturovaný přehled základních metod, které jsou pro filtrování využívány. Ten se v mnoha ohledech podařilo naplnit, přestože kvůli svému rozsahu musela práce na mnoha místech slevit z požadavku na zevrubný popis všech stránek filtrování informací.

Podle situace popsané v úvodní kapitole této práce je zřejmé, že budoucnost musí nezbytně být filtrovaná. Je však zapotřebí, aby budoucí implementace filtrování informací braly ohledy na několik klíčových požadavků.

Namísto toho, aby byly zneužívány k zajištění uniformity pomocí cenzury, by měly decentralizací kontroly nastavení filtrů a poskytováním přístupu k rozmanitým informačním zdrojům podporovat ideologickou různorodost. Tím by se vyhnuly riziku centralizovaného řízení filtrování, které spočívá v tom, že ten, kdo jej ovládá, do jeho nastavení začne promítat vlastní morálku a svá měřítka.

Aby jim mohli jejich uživatelé důvěřovat, musí být systémy filtrování informací naprosto transparentní a umět „vysvětlit“, proč jsou informace odfiltrovány. Důvěru uživatelů si musí získat také tím, že znemožní narušování jejich soukromí.

K dosažení transparence je optimální zveřejnit zdrojový kód systému filtrování informací, takže bude zřejmý způsob, jakým funguje. Při zveřejňování software pro filtrování pod některou z open-source licencí je možné, že dojde ke standardizaci slovníku hodnocení a popisu informačních jednotek, díky čemuž by mohla být vytvořena určitá úroveň interoperability a vzájemné kompatibility mezi jednotlivými aplikacemi.

Díky použití pokročilých technik, které si půjčují postupy z oblastí vyhledávání informací, strojového učení nebo lineární algerby, dosahují již dnešní systémy filtrování informací vysoké účinnosti. Jejich přesnost může za určitých podmínek dosahovat přesnosti, které je při odhalování žádoucích a nežádoucích informací schopen člověk.

Problémem, který se doposud nepodařilo zcela odstranit, jsou *falešné pozitivy* - nesprávně filtrované jednotky. V tomto případě, kdy technologie blokuje legitimní obsah, představuje omezení svobody projevu a práva na informace.

Filtry umožňují blokovat nežádoucí způsoby reklamy (spam) a zároveň mohou poskytovat žádoucí způsob propagace. To ilustruje fakt, že z doporučení navržených systémy filtrování informací vzniká významná část nákupů na internetu.

Po splnění těchto požadavků a dosažení efektivity mohou být metody automatizovaného filtrování informací užitečnou a spolehlivou pomocí každému uživateli internetu.

Seznam použitých zkratek

ACLU American Civil Liberties Union
ARI Adresní rozšiřování informací
CAPTCHA Completely automated public Turing test to tell computers and human apart
CIPA Children internet protection act
CSS Cascading stylesheets
DCC Distributed checksum clearinghouse
DNS Domain name system
HTML Hypertext markup language
HTTPS Hypertext transfer protocol secure
IM Instant messaging
IP Internet protocol
IUP Internet user policy
MTA Mail transfer agent
P2P Peer to peer
PICS Platform for internet content selection
RBL Realtime blackhole list
RDF Resource description framework
RFC Request for comments
SQL Structured query language
TCP Transmission control protocol
TREC Text retrieval conference
SDI Selective dissemination of information
SHA Secure hash algorithm
SMS Short message service
TF×IDF Term frequency × inverse document frequency
UGC User-generated content
VoIP Voice over Internet Protocol

Seznam použité literatury

- AARDSMA, T.L. Keep your inbox clean with spam-filtering software. *Inside the Internet.* August 2002, vol. 9, no. 8, s. 7–10. ISSN 1075-7902.
- ABERNETHY, Jacob [et al.]. Collaborative filtering with attributes. In *The Snowbird learning workshop, Snowbird, USA, April , 2008.* Snowbird, 2008.
- ABRAMOWICZ, Witold; KALCZYŃSKI, Pawel; WECEL, Krzystof. *Filtering the web to feed data warehouses*. London : Springer, 2002. xii, 267 s. ISBN 1-85233-579-3.
- ABRAMOWICZ, Witold [ed.]. *Knowledge-based information retrieval and filtering from the web*. Boston (MA) : Kluwer, 2003. xvi, 303 s. The Kluwer international series in engineering and computer science. ISBN 1-4020-7523-5.
- ADAMS, Helen R. Filters and access to information : part I. *School Library Media Activities Monthly.* September 2008, vol. 25, no. 1, s. 55. ISSN 0889-9371.
- ADAMS, Helen R. Filters and access to information : part II. *School Library Media Activities Monthly.* October 2008, vol. 25, no. 2, s. 54. ISSN 0889-9371.
- ALA to monitor Internet filter implementation, provide support to library staff and users as CIPA deadline approaches. *Newsletter on intellectual freedom.* 2004, vol. 53, no. 5, s. 173. ISSN 0028-9485.
- ANG, Peng Hwa. Censorship of the Internet. In *Encyclopedia of library and information science.* ed. New York : Marcel Dekker, c2003, s. 475-483. ISBN 978-0-8247-2071-1 (elektronická verze).
- ASSAY, Matt. *Shirky : problem is filter failure, not info overload* [online]. January 19, 2009 [cit. 2009-05-13]. Dostupný z WWW: <http://news.cnet.com/8301-13505_3-10142298-16.html>.
- BALKIN, J.M.; NOVECK, Beth Simone; ROOSEVELT, Kermit. Filtering the internet : a best practices model [online]. Yale Law School, 1999 [cit. 2009-05-03]. Dostupný z WWW: <http://www.yale.edu/lawweb/jbalkin/articles/Filters0208.pdf>.
- BATES, Mary Ellen. Building a better search engine. *Online.* 2007, vol. 31, no. 2, s. 64. ISSN 0146-5422.
- BELL, Mary Ann. The elephant in the room. *School Library Journal.* January 2007, vol. 53, no. 1, s. 40–42. ISSN 0362-8930.
- BELLUCK, Pam [et al.]. Caught in the web. *New York Times Upfront.* Feb 14, 2005, s. 5. ISSN 1525-1292.
- BERGHEL, Hal [et al.]. Cyberbrowsing : information customization on the web. *Journal of the American Society for Information Science.* 1999, vol. 50, no. 6, s. 505–513. ISSN 0002-8231.
- BIGGS, Maggie. Reining in the web. *Federal Computer Week.* February 4, 2002, vol. 16, no. 3, s. 33–35. ISSN 0893-052X.
- BODARD, Katia. Free access to information challenged by filtering techniques.

Information & Communication Technology Law. October 2003, vol. 12, no. 3, s. 263–279. ISSN 1360-0834.

- BRADLEY, Keith; SMYTH, Barry. Information ordering vs information filtering. In *European conference on case based reasoning, Aberdeen, Scotland, UK, September 4–7, 2002 : workshop proceedings.* Berlin : Springer, 2002, s. 13–14. Lecture notes in computer science, Lecture notes in artificial intelligence, vol. 2416. ISBN 978-3-540-44109-0.
- BROŽOVSKÝ, Lukáš.*Recommender system for a dating service.* Praha, 2006. 66 s. Diplomová práce (Mgr.). Univerzita Karlova v Praze. Matematicko-fyzikální fakulta. Dostupný také z WWW: <http://colfi.wz.cz/colfi.pdf>.
- BURKE, Robin. Hybrid web recommender systems. In BRUSILOVSKY, Peter; KOBSA, Alfred; NEJDL, Wolfgang [eds.]. *The adaptive web : methods and strategies of web personalization*. Berlin; Heidelberg : Springer, 2007, s. 377–408. Lecture notes in computer science, 4321/2007. ISBN 978-3-540-72078-2.
- BURKE, Robin. Semantic ratings and heuristic similarity for collaborative filtering. In *Proceedings of the national conference on artificial intelligence.* Menlo Park (CA) : AAAI Press, c2000. ISBN 0-262-51112-4.
- CAMPBELL, Shugana. Politics, religion, images and abortion : do internet filters block controversial sources of information? *Mississippi Libraries.* 2001, vol. 65, no. 4, s. 107–108. ISSN 0194-388X.
- CARMAGNOLA, Francesca [et al.]. Towards a tag-based user model : how can user benefit from tags? In *User modeling 2007 : proceedings of the international conference, Corfu, Greece, July 25–29, 2007.* Berlin : Springer, 2007, s. 445–449. Lecture notes in computer science, 4511/2007. ISBN 978-3-540-73077-4.
- CATTUTO, Ciro [et al.]. Semantic analysis of tag similarity measures in collaborative tagging systems. In *Proceedings of the workshop on ontology learning and population at the European conference on artificial intelligence, July , 2008 - Patras, Greece.* Amsterdam : IOS Press, 2008. Dostupný také z WWW: <http://arxiv.org/abs/0805.2045>.
- COPPOCK, Patrick. A conversation on information : an interview with Umberto Eco. *Multimedia World*. December 1995, vol. 3. Dostupný také z WWW: <http://carbon.cudenver.edu/~mryder/itc_data/eco/eco.html>. ISSN 1073-4759.
- COSTALES, Bryan; FLYNT, Marcia. *Sendmail milters : a guide for fighting spam*. Upper Saddle River (NJ) : Addison-Wesley, 2005. ISBN 0-321-21333-5.
- CROTTY, David. 2009a. *Information overload is not filture failure* [online]. January 14, 2009 [cit. 2009-05-13]. Dostupný z WWW: <http://www.cshblogs.org/cshprotocols/2009/01/14/information-overload-is-not-filture-failure/>.
- CROTTY, David. 2009b. *Information overload part 2* [online]. January 16, 2009 [cit. 2009-05-13]. Dostupný z WWW: <http://www.cshblogs.org/cshprotocols/2009/01/16/information-overload-part-2/>.

- DAILY, Geoff. A case of delivering family-friendly entertainment. *EContent.* May 2006, vol. 29, no. 4, s. 45–47. ISSN 1525-2531.
- DAMIANI, E. [et al.]. An open digest-based technique for spam detection. In *Proceedings of the 2004 International workshop on security in parallel and distributed systems.* Washington (D.C.) : IEEE, 2004, s. 15–17.
- DEEPAK, P.; JYOTHI, John; SANDEEP, Parameswaran. A community based approach for spam filtering. In *Proceedings of the 1 International conference on information and communication technologies : from theory to applications*. Washington (D.C.) : IEEE, 2004, s. 611–612. ISBN 0-7803-8482-2.
- GARCÍA, Kimberly. Kid-safe surfing : how to protect your child online. *Hispanic.* March 2003, vol. 16, no. 3, s. 52.
- GARCÍA-BARRIOCANAL, Elena; SICILIA, Miguel-Angel. Filtering information with imprecise social criteria : a FOAF-based backlink model. In MONTSENY, Eduard; SOBREVILLA, Pilar (eds.). *Proceedings of the 4 conference of the European Society for Fuzzy Logic and Technology.* Barcelona : Universidad Polytecnica de Cataluña, 2005, s. 1094-1098. ISBN 84-7653-872-3.
- GLEICK, James. Tangled up in spam. *New York Times Magazine.* Feb 9, 2003, s. 42–47. Dostupný také z WWW: <http://www.nytimes.com/2003/02/09/magazine/09SPAM.html>. ISSN 0028-7822.
- GOLDBERG, David [et al.]. Using collaborative filtering to weave an information tapestry. *Communications of the ACM*. Special issue on information filtering. December 1992, vol. 35, is. 12, s. 61–70. Dostupný také z WWW: <http://www.ischool.utexas.edu/~i385d/readings/Goldberg_UsingCollaborative_92.pdf>. ISSN 0001-0782.
- GOOD, Robin. *Information overload : blogs as content navigators, information filters, trusted niche guides* [online]. February 10, 2006 [cit. 2009-05-13]. Dostupný z WWW: <http://www.masternewmedia.org/news/2006/02/10/information_overload_blogs_as_content.htm>.
- HAN, Asung [et al.]. Semantic analysis of user behaviors for detecting spam mail. In *IEEE international workshop on semantic computing and applications*. Washington (D.C.) : IEEE, 2008, s. 91–95. ISBN 978-0-7695-3317-9.
- HARRISON, Carrie. When spam-blocking filters labels „good“e-mail „bad“. *The Canadian Manager.* Summer 2003, vol. 28, no. 2, s. 23–24. ISSN 0045-5156.
- HASNAH, Ahmad M. A new filtering algorithm for duplicate document based on concept analysis. *Journal of Computer Science.* 2006, vol. 2, no. 5, s. 434–440. ISSN 1549-3636.
- HÖFFERER, Max; KNAUS, Bernd; WINIWARTER, Werner. Adaptive information filtering by monitoring user behavior. In *8 international symposium on methodologies for intelligent systems*. Vienna : Institute of Applied Computer Science and Information Systems, 1994.
- HÖFFERER, Max; KNAUS, Bernd; WINIWARTER, Werner. Using genetics in information filtering. In *Proceedings of the annual conference of the Canadian*

Association for Information Science. Montreal : Canadian Association for Information Science, 1994, s. 435–445.

- HULEATT, Sam. 2009a. *Information day trading* [online]. February 20, 2009 [cit. 2009-05-13]. Dostupný z WWW: <http://www.leveragingideas.com/2009/02/20/information-day-trading/>.
- HULEATT, Sam. 2009b. *10 posts on data mining and information filtering* [online]. February 21, 2009 [cit. 2009-05-13]. Dostupný z WWW: <http://www.leveragingideas.com/2009/02/21/10-posts-on-data-mining-and-information-filtering/>.
- HUNTER, Christopher D. *Filtering the future? : software filters, porn, PICS, and the internet content conundrum.* Philadelphia (PA), 1999, 159 s. Diplomová práce (M.A.). University of Pennsylvania. Faculty of the Annenberg School for Communication.
- CHAU, Rowena; YEH, Chung-Hsing. Filtering multilingual web content using fuzzy logic and self-organizing maps. *Neural Computing & Applications.* 2004, vol. 13, issue 2, s. 140–148. ISSN 0941-0643.
- CHEN, Anne Yun-An; MCLEOD, Dennis. Collaborative filtering for information recommendation systems. In WANG, John [ed.] *Encyclopedia of data warehousing and mining.* Hershey (PA) : Idea Group, 2005. ISBN 1-59140-557-2.
- CHISLENKO, Alexander. Collaborative information filtering and semantic transports. In WOLF, Milton T.; ENSOR, Pat; THOMAS, Mary A. (ed.). *Information imagineering : meeting at the interface*. Chicago : American Library Association, 1998. xiv, 255 s. Dostupný také z WWW: <http://www.lucifer.com/~sasha/articles/ACF.html>. ISBN 0-8389-0729-6.
- ISKOLD, Alex. 2007a. *The art, science and business of recommendation engines* [online]. Janurary 16, 2007 [cit. 2009-05-13]. Dostupný z WWW: <http://www.readwriteweb.com/archives/recommendation_engines.php>.
- ISKOLD, Alex. 2007b. *The attention economy : an overview* [online]. March 1, 2007 [cit. 2009-05-13]. Dostupný z WWW: <http://www.readwriteweb.com/archives/attention_economy_overview.php>.
- JÁCSÓ, Péter. Savvy searching : citedness scores for filtering information and ranking search results. *Online Information Review*. 2004, vol. 28, no. 5, s. 371–376. ISSN 1468-4527.
- JANES, Joseph. Chuck chuck. *American Libraries.* October 2007, vol. 38, no. 9, s. 49. ISSN 0002-9769.
- JI, Ae-Ttie [et al.]. Collaborative tagging in recommender systems. In *AI 2007 : Advances in artificial intelligence : proceedings of 20 Australian joint conference on artificial intelligence, Gold Coast, Australia, December 2–6, 2007.* Berlin : Springer, 2007, s. 377–386. Lecture notes in computer science, 4830/2007. ISBN

978-3-540-76926-2. DOI 10.1007/978-3-540-76928-6\s\do5(3)9.

- JOHNSON, Doug. Freedom and filters. *Library Media Connection.* February 2003, vol. 21, no. 5, s. 107–108. ISSN 1542-4715.
- JUSKALIAN, Russ. *Interview with Clay Shirky : part I* [online]. December 19, 2008 [cit. 2009-05-13]. Dostupný z WWW: <http://www.cjr.org/overload/interview_with_clay_shirky_par.php>.
- KAUTZ, Henry; SELMAN, Bart; SHAH, Mehul. Referral web : combining social networks and collaborative filtering. *Communications of the ACM.* March 1997, vol. 40, no. 3, s. 63–65. ISSN 0001-0782.
- KELLAR, Melanie; WATTERS, Carolyn. Using web browser interactions to predict task. In *Proceedings of the international conference on World Wide Web.* New York : ACM, 2006, s. 843–844. ISBN 1-59593-323-9.
- LEE, Danielle Hyunsook. PITTCULT : trust-based cultural event recommender. In *Proceedings of the 2008 ACM conference on recommender systems.* New York : ACM, 2008, s. 311–314. ISBN 978-1-60558-093-7.
- LERMAN, Kristina. Social networks and social information filtering on Digg. In *Proceedings of International conference on weblog and social media, Boulder, Colorado, USA.* Menlo Park (CA) : AAAI Press, 2007.
- LESSIG, Lawrence; RESNICK, Paul. Zoning speech on the internet : a legal and technical model. *Michigan Law Review.* Summer 1998, issue 2, s. 395–431. ISSN 0026-2234.
- LIEBLER, Raizel. Beware the Mini-CIPA. *American Libraries.* August 2004, vol. 35, no. 7, s. 39. ISSN 0002-9769.
- LONGBOTTOM, Clive. Don't get a spasm over Spit or spam. *Computer Weekly.* Jan 9, 2007, s. 34. ISSN 0010-4787.
- LOSINSKI, Robert. Patrolling web 2.0. *T.H.E. Journal.* March 2007, vol. 34, no. 3, s. 50–52. ISSN 0192-592X.
- MALONE, Thomas W. [et al.]. Intelligent information sharing systems. *Communications of the ACM.* May 1987, vol. 30, issue 5, s. 390–402. ISSN 0001-0782.
- MARTIN, Malachi. Information access, libraries, and filtering : philosophical considerations. *Mississippi Libraries.* Summer 2003, vol. 67, no. 2, s. 41–43. ISSN 0194-388X.
- MARVANOVÁ, Eva. Filtrování internetu v zahraničí. In *Knihovny současnosti 2006 : sborník z 14. konference, konané ve dnech 12.–14. září 2006 v Seči u Chrudimi.* Brno : Sdružení knihoven ČR, 2006, s. 159–168. ISBN 80-86249-41-7.
- MCCARTHY, David. Internet filtering for schools. *Media and Methods.* May/Jun 2005, vol. 41, no. 6, s. 9–11. ISSN 0025-6897.
- MESSMER, Ellen. Web filtering tools handling ever-larger jobs. *Network World,* May 2005, vol. 22, no. 20, s. 24. ISSN 0887-7661.
- MINKEL, Walter. A filter that lets good information in. *School Library Journal.* Mar

2004, vol. 50, no. 3, s. 28–30. ISSN 0362-8930.

- MINKEL, Walter. Who's blocking whom? *School Library Journal.* June 2003, vol. 49, no. 6, s. 35. ISSN 0362-8930.
- Monroe City adopts tough net policy. *Library Journal.* June 2007, vol. 132, no. 11, s. 16–17. ISSN 0363-0277.
- MORGAN, Candace D. Internet filtering and individual choice. *Oregon Library Association Quarterly.* Winter 2004, vol. 10, no. 4, s. 5–7. ISSN 1093-7374.
- MORGAN, Eric Lease. 2009a. *TFIDF in libraries : part I of III (for librarians)* [online]. April 13, 2009 [cit. 2009-05-13]. Dostupný z WWW: <http://infomotions.com/blog/2009/04/tfidf-in-libraries-part-i-for-librarians/>.
- MORGAN, Eric Lease. 2009b. *TFIDF in libraries : part II of III (for programmers)* [online]. April 20, 2009 [cit. 2009-05-13]. Dostupný z WWW: <http://infomotions.com/blog/2009/04/tfidf-in-libraries-part-ii-of-iii-for-programmers/>.
- MUNROE, Mary H. To filter or not to filter, that is the question. *Illinois Library Association Reporter.* February 2006, vol. 24, no. 1, s. 38-39. ISSN 0018-9979.
- National report : internet filters can't replace parents. *School Library Journal.* June 2002, vol. 48, no. 6, s. 16. ISSN 0362-8930.
- NOTESS, Greg R. Community filtering : digg, Slashdot, and the social web. *Online.* Jan/Feb 2007, vol. 31, no. 1, s. 45–47. ISSN 0146-5422.
- OARD, Douglas W.; MARCHIONINI, Gary. *A conceptual framework for text filtering*. College Park (MD) : University of Maryland, 1996. 32 s.
- O'DONOVAN, John; SMYTH, Barry. Is trust robust? : an analysis of trust-based recommendation. In *Proceedings of the international conference on intelligent user interfaces.* New York : ACM, 2006, s. 101–108. ISBN 1-59593-287-9.
- O'RIORDAN, Colm; SORENSEN, Humphrey. *Information filtering and retrieval : an overview* [online]. Galway (IE) : National University of Ireland, 1999 [cit. 2009-05-03]. Dostupný z WWW: <http://ww2.it.nuigalway.ie/cirg/localpubs/ORiordanTechnical1999.ps>.
- PAPADIMITRIOU, Christos H. [et al.]. Latent semantic indexing : a probabilistic analysis. In *Proceedings of the ACM SIGACT-SIGMOD-SIGART symposium on principles of database systems.* New York : ACM, 1998, s. 159–168. ISBN 0-89791-996-3.
- PATIL, Chris; SIEGEL, Vivian. Drinking from the firehose of scientific publishing. *Disease Models & Mechanisms.* March/April 2009, vol. 2, no. 3–4, s. 100-102. DOI 10.1242/dmm.002758.
- PAZZANI, Michael J. Representation of electronic mail filtering profiles : a user study. In *Proceedings of the international conference on intelligent user interfaces.* New York : ACM, 2000, s. 202–206. ISBN 1-58113-134-8.
- PAZZANI, Michael J.; BILLSUS, Daniel. Content-based recommendation systems. In BRUSILOVSKY, Peter; KOBSA, Alfred; NEJDL, Wolfgang [eds.]. *The adaptive*

web : methods and strategies of web personalization. Berlin; Heidelberg : Springer, 2007, s. 325–341. Lecture notes in computer science, 4321/2007. ISBN 978-3-540-72078-2.

- PIKE, George H. Living in a post-CIPA world. *Information Today.* September 2003, vol. 20, no. 8, s. 17–20. ISSN 8755-6286.
- PIKE, George H. Myspace.com and library filters. *Information Today.* July/August 2006, vol. 23, no. 7, s. 15–19. ISSN 8755-6286.
- *Proceedings of the 5 DELOS workshop on filtering and collaborative filtering.* The European Research Consortium for Informatics and Mathematics. Pisa : ERCIM, 1997. Dostupný také z WWW: <http://www.ercim.org/publication/ws-proceedings/DELOS5/delos5.pdf>. ISBN 2-912335-04-3.
- PRYOR, Michael H. The effects of singular value decomposition on collaborative filtering. Hanover (NH) : Dartmouth College, 1998. 43 s. Technická zpráva. PCS-TR98-338.
- RAMACHANDRAN, Anirudh; FEAMSTER, Nick; VEMPALA, Santosh. Filtering spam with behavioral blacklisting. In *Proceedings of the ACM conference on computer and communications security.* New York : ACM, 2007, s. 342–351. ISBN 978-1-59593-703-2.
- RAO, K. Nageswara; TALWAR, V.G. Generating user profiles by translating content queries to concepts using thesaurus. *DESIDOC Bulletin of Information Technology.* July 2006, vol. 26, no. 4, s. 3–15. ISSN 0971-4383.
- REDDICK, Thomas M. Building and running a collaborative internet filter is akin to a Kansas barn raising. *Computers in Libraries.* April 2004, vol. 24, no. 4, s. 10–14. ISSN 1041-7915.
- SCOBLE, Robert. *Things I've learned by clicking „like"15,301 times* [online]. January 22, 2009 [cit. 2009-05-13]. Dostupný z WWW: <http://scobleizer.com/2009/01/22/things-ive-learned-by-clicking-like-15301-times/>.
- SEGARAN, Toby. *Programming collective intelligence : building smart web 2.0 applications*. Sebastopol : O'Reilly, 2007. ISBN 978-0-596-52932-1.
- SHARDANAND, Upendra; MAES, Pattie. Social information filtering : algorithms for automating „word of mouth". In *Proceedings of the SIGCHI conference on human factors in computing systems, May 7–11, 1995 - Denver, Colorado, USA.* Denver (CO) : ACM, c1995. Dostupný také z WWW: <http://www.cs.ubc.ca/~conati/532b/papers/chi-95-paper.pdf>. ISBN 0-201-84705-1.
- SHEPHERD, Michael; WATTERS, Carolyn. Content filtering technologies and internet service providers : enabling user choice [online]. Halifax (Nova Scotia) : Dalhousie University. Faculty of Computer Science. Web Information Filtering Lab, 2000 [cit. 2008-04-21]. Výzkumná zpráva. Kanada. Industry Canada. Dostupný z WWW: <http://users.cs.dal.ca/~shepherd/filtering/ISPweb.htm>.

- SHIRKY, Clay. *It's not information overload. It's filter failure* [online]. 18.9. 2008 [cit. 2009-05-13]. Dostupný z WWW: <http://web2expo.blip.tv/file/1277460>.
- SCHAFER, J. Ben; KONSTAN, Joseph A.; RIEDL, John. E-commerce recommendation applications. *Data Mining and Knowledge Discovery*. 2001, vol. 5, no. 1-2, s. 115–153. ISSN 1384-5810.
- SCHAFER, J. Ben [et al.]. Collaborative filtering recommender systems. In BRUSILOVSKY, Peter; KOBSA, Alfred; NEJDL, Wolfgang [eds.]. *The adaptive web : methods and strategies of web personalization*. Berlin; Heidelberg : Springer, 2007, s. 291–324. Lecture notes in computer science, 4321/2007. ISBN 978-3-540-72078-2.
- SCHMIDT, Cindy M. Those intefering filters! : how to deal with the reality of filters in your school library. *Library Media Connection*. March 2008, vol. 26, no. 6, s. 54–55. ISSN 1542-4715.
- SMYTH, Barry. Case-based recommendation. In BRUSILOVSKY, Peter; KOBSA, Alfred; NEJDL, Wolfgang [eds.]. *The adaptive web : methods and strategies of web personalization*. Berlin; Heidelberg : Springer, 2007, s. 342–376. Lecture notes in computer science, 4321/2007. ISBN 978-3-540-72078-2.
- SORTORE, Sam M. Filtering : a piece of the puzzle. *School Library Journal : part Netconnect*. Summer 2001, s. 20–21. ISSN 0362-8930.
- SPAMMER X. *Inside the spam cartel : trade secrets from the dark side*. Rockland (MA) : Syngress, 2004. 450 s. ISBN 1-932266-86-0.
- SPRING, Tom. Buying way into your inbox. *PC World*. May 2006, vol. 24, no. 5, s. 24. ISSN 0737-8939.
- STANNARD-STOCKTON, Sean. *Information filtering* [online]. February 20, 2009 [cit. 2009-05-13]. Dostupný z WWW: <http://tacticalphilanthropy.com/2009/02/information-filtering>.
- STOCKWELL, Laura Porto. *SXSW 2009 : collaborative filters, the evolution of recommendation engines* [online]. March 14, 2009 [cit. 2009-05-13]. Dostupný z WWW: <http://www.digitaldialogs.com/2009/03/sxsw-2009-collaborative-filters.html>.
- TEMPLETON, Brad. *Reflections on the 25 anniversary of Spam* [online]. May 2003 [cit. 2009-05-13]. Dostupný z WWW: <http://www.templetons.com/brad/spam/spam25.html>.
- WALKER, Andrew [et al.]. Collaborative information filtering : a review and an educational application. *International journal of artificial intelligence in education*. 2004, vol. 14, s. 1–26. ISSN 1560-4292 (print).
- WANG, Jun; DE VRIES, Arjen P.; REINDERS, Marcel J.T. A user-item relevance model for log-based collaborative filtering. In *Advances in information retrieval*. Berlin : Springer, 2006, s. 37–48. Lecture notes in computer science, 3936/2006. ISBN 978-3-540-33347-0.
- WEISER, Christine. New web filtering approaches. *Scholastic Administr@tor*.

January 2008, vol. 7, no. 4, s. 23–25. ISSN 1538-5191.

- YERAZUNIS, William S. The spam-filtering accuracy plateau at 99.9 percent accuracy and how to get past it. In *Proceedings of the MIT spam conference*. MIT Press : Cambridge (MA) : 2004.
- ZDZIARSKI, Jonathan A. *Ending spam.* San Francisco : No Starch, 2005. 312 s. ISBN 1-59327-052-6.
- ZHANG, Yi-Cheng. Towards a new information theory. *International Journal of Modern Physics B.* 2004, vol. 18, no. 17–19, s. 2361–2364. ISSN 0217-9792.

Evidence výpůjček

Prohlášení

Dávám svolení k půjčování této bakalářské práce. Uživatel potvrzuje svým podpisem, že bude tuto práci řádně citovat v seznamu použité literatury.

V Praze, 20.5. 2009

Jindřich Mynarz

Jméno	Katedra/pracoviště	Datum	Podpis