

# UŽITÍ SÍŤOVÉ ANALÝZY V MARKETINGU

Ladislav Beránek

Jihočeská universita v Českých Budějovicích

## Abstrakt

**Porozumění vztahů se zákazníkem je důležitým aspektem firemního marketingu. Zatímco teorie sítí je v současné široce používaná v oblastech sociální a organizační vědy, formální techniky analýzy sítí nejsou využívány tak široce v marketingovém výzkumu. Hlavní důvod je ten, že získání potřebných dat není jednoduché a také nejsou nástroje, které by byly uplatnitelné v praktickém marketingu firem. Analýza sociální sítě však umožňuje studovat charakteristických vlastností sítí zákazníků a i jednotlivých aktérů v síti podle jejich pozice v síti, studovat vztahy v síti (například rozpoznat aktéry, kteří jsou schopni ovlivňovat další, atd.). Cíl tohoto příspěvku je ukázat některé koncepty použití síťové analýzy v oblasti marketingu. Naším cílem je vývoj nástrojů pro praktické použití síťové analýzy pro praktické úlohy porozumění chování spotřebitelů a reprezentaci chování zákazníků pomocí orientovanými grafy. V této fázi využíváme prostředí MATLAB pro testování základních přístupů, protože toto prostředí umožňuje rychlý vývoj potřebných nástrojů a vizualizačních schémat.**

## 1 Úvod

Sociální sítě ve své standardní podobě byly do nedávné doby spíše popisné, obsahovaly několik málo globálních parametrů a jen obtížně bylo možné generovat z nich předpovědi pro budoucí chování sítě. To bylo způsobeno velkou měrou nedostatkem dat o sítích, na základě kterých by bylo možné provádět experimentální studium a navrhovat a verifikovat modely těchto sítí.

V současné době s nárůstem komunikačních technologií umožňující nejen komunikaci, ale i sběr dat o této komunikaci, se situace změnila. V současné době jsou dostupná velká množství dat o velkých sociálních sítích. Jsou to data zpracovaná z blogů, sítí pro sdílení znalostí, sítí spolupracujících osob, sítí online her, diskusní skupiny, chaty, webové stránky a další. Tyto sítě mají desítky tisíc až dokonce milióny uzlů, a často obsahují značná množství informací na úrovni jednotlivých uzlů a umožňují tak zkoumání a verifikaci různých modelů sítí a dynamických procesů, které mohou na této struktuře probíhat. Tyto modely a nové metody analýzy sítí umožňují předpovídat vývoj sítí a i použití v rozhodování, to i v rámci zákaznických sítí.

## 2 Síťová hodnota zákazníků

Hodnota zákazníka je obvykle definována jako očekávaný zisk z prodeje tomuto zákazníkovi, a to přes celý životní cyklus vztahu zákazníka a dané společnosti. O tuto hodnotu se společnosti velice zajímají, protože ona určuje, zda má význam utrácet peníze, aby společnost získala a udržela si tohoto zákazníka. Avšak tyto tradiční měřítka hodnoty zákazníka ignorují skutečnost, že v okamžiku, kdy si zákazník koupí produkt dané společnosti, může, pokud je spokojen s produktem a službami, ovlivnit jiné, aby i oni si zakoupili daný produkt nebo služby. Například pokud si koupím dobrý výrobek, pak mohu přesvědčit několik svých přátel, aby si i oni zakoupili tento produkt. Potom moje hodnota jako zákazníka je pro výrobce tohoto produktu několikanásobně vyšší a společnost vyrábějící daný produkt může vydat více peněz na podporu prodeje výrobku a to cíleně na moji osobu, což by jinak neudělala. A obráceně, jestli se spíše nechám ovlivnit svými přáteli v oblasti nákupu specifických produktů, pak společnost vyrábějící tyto produkty udělá lépe, pokud bude cílit marketingové nástroje na mého přítele než na mě, kde by to bylo plýtvání zdrojů. Síťovou hodnotu zákazníka nazýváme tedy očekávané zvýšení prodeje jiným osobám, který bude důsledkem cíleného marketingu na tohoto zákazníka.

Tradiční marketing nebere v úvahu síťovou hodnotu zákazníků a tím rozhodnutí která jsou v tradičním marketingu dělána, nejsou optimální. V literatuře o marketingu je často zmiňována existence síťových efektů, ale až donedávna se tato oblast považovala za oblast, kde není možné nastavit nějaká kvantitativní měřítka. V současné době však existují nástroje, modely, které umožňují měřit síťovou

hodnotu zákazníka. Pro každého zákazníka lze namodelovat pravděpodobnost, že určitý zákazník koupí daný produkt jako funkci reálných vlastností zákazníka a produktu se započítáním vlivu sousedů zákazníka v síti. Určením pravděpodobnostních závěrů na model sítě všech zákazníků můžeme odpovědět na otázky typu: obchodujeme s touto specifickou množinou zákazníků, jaký je očekávaný zisk pro z celé sítě poté, co jsme celou množinou zákazníků ovlivnili pouze prostřednictvím „vybraných vlivných zákazníků“. Tuto schopnost můžeme použít pro to, abychom určili specifickou malou množinu „vybraných vlivných“ zákazníků, na které budeme cílit naše marketingové zdroje. Tato množina zákazníků bude efektivní z hlediska vynaložených nákladů na marketing. Hledáme tedy zákazníky s nejvyšší sítovou hodnotou, cílíme na ně marketingové zdroje a získáváme ze schopností těchto zákazníků ovlivnit druhé.

### 3 Faktory ovlivňující sítovou hodnotu

Jaké jsou faktory, které určují zákazníka s vysokou sítovou hodnotou? V první řadě vysoký počet vazeb v sociální síti. Ale v našem modelu jsou další důležité faktory. Především je důležité, zda zákazník má rád produkt. Zákazníci, kteří mají vysoký počet vazeb, ale nemají rádi produkt, mají pro společnost zápornou sítovou hodnotu, a marketing by se jim měl vyhnout. Další klíčový aspekt je, zda zákazník s mnoha vazbami ovlivňuje své známe více (teoreticky mnohem více) než tito známi ovlivňují zpátky tohoto zákazníka. Jestli by stupeň ovlivňování byl souměrný, nezískáme žádnou výhodu, pokud budeme oslovovat tohoto zákazníka s mnoha vazbami. Naštěstí, nesouměrné ovlivňování je v praktickém životě daleko rozšířenější a této zkušenosti využívá i náš přístup. Zatímco v různých oblastech jsou známí názoroví vůdci (například různé osobnosti, celebrity), náš přístup umožňuje identifikovat názorové vůdce na lokální úrovni [5]. Třetí (a možná nejdůležitější) aspekt je to, že zákaznickova sítová hodnota nekončí u jeho bezprostředních známých. Tito bezprostřední známí postupně ovlivňují další lidi, a tak dále až je potenciálně dosaženo celé sítě. Tito známi by tedy měli také mít rádi produkt a měli by ovlivnit mnoho dalších lidí. Zákazník, který nemá mnoho vazeb, může mít ve skutečnosti vysokou sítovou hodnotu, jestliže jeden z jeho známých je osoba s mnoha vazbami (například, rádce názorového vůdce). V experimentální webové síti sdílení znalostí [1] měl nejcennější zákazník sítovou hodnotu přes 20,000. To znamená, že marketing směřující k tomuto zákazníkovi bude tak efektivní jako marketing cílený na 20,000 lidí, kteří nemají takový sítový efekt. Množství vazeb tohoto zákazníka počet přímých spojení k jiným aktérům v síti (i.e., lidé, kteří čtou jeho recenze) však bylo mnohem menší. Ústní marketing však nemusí být efektivní na určitých trzích, protože potřebné sítě vlivů nejsou prostě přítomné. Mnoho začínajících společností neuspělo ve snaze rozprout sítový efekt. Obráceně, rozšíření nějakých produktů, jako pokladní karta a interaktivní televize, měly za následek neúspěch; protože produkt byl dán malému vzorku izolovaných zákazníků, který neměl za následek vznik sítového efektu [1]. Když jsou k dispozici data, modely umožňují měřit tyto efekty přesně a udělat odpovídající lepší rozhodnutí.

Další zajímavý následek sítových modelů je simulace finanční ztráty, kterou lze utrpět na základě negativního hodnocení produktu určitých zákazníčích, kteří však jsou dost vlivní. Ve vřevém marketingu, jestliže se poskytne produkt zdarma pro dobře voleného zákazníka, může se to vyplatit v mnoha prodejších dalším zákazníkům.

### 4 Výsledky experimentů

Máme tedy model sociální sítě a máme dobře definovaný optimalizační problém: vybrat si množinu zákazníků, se kterými navážeme užší obchodní (marketingový) kontakt s cílem maximalizace zisku. Tento problém je ale NP-složitý [2]. Je nutno použít přibližných metod. V našich experimentech byly použity různé heuristické metody s dobrými výsledky vyhledávání. Např. každý zákazník je přidán k aktuálně stanovené „marketinkové množině“, dokud toto přidání zlepšuje se celkový přínos (zisk). Měření efektu přidávání zákazníka k množině „vlivných“ není jednoduché. Sice velká většina zákazníků má velmi malou sítovou hodnotu; jejich vliv v síti se nešíří velmi daleko, a tak výpočet v těchto případech rychle konverguje. Pro málo zákazníků, kteří mají velký počet vazeb, však může výpočet trvat značný čas.

Bez ohledu na to, jak moc dat máme, úplně určení sítě sociálních interakcí mezi lidmi v skutečném světě nikdy nebude proveditelná. Takže vyvstává důležitá otázka, zda přístup k

maximalizování osobního ovlivňování ještě funguje, když neexistuje úplná znalost o síti. Toto bylo testováno náhodným odstraněním proměnného množství hran ze sítě před aplikací metod data miningu. Systém vzniklý odebráním uzlů však byl celkem robustní. Výsledky simulací ukazovaly, že cíleným marketingem došlo k 70% zvýšení zisku, přestože bylo známo jen 5% hran v síti. Model může být také použitý pro určení efektivního způsobu, jaké je potřeba získat dodatečné znalosti o síti. Z pokusů se ukázalo, že jednoduché dotazníky na zákazníky s nejvyšší sítí hodnotou v reálné síti, kdo jsou jejich známí, jsou využitelné [3].

Stávající modely jsou jenom začátkem velmi bohaté množiny možností, které lze z těchto sítí získat. Skutečné sociální sítě se vyvíjejí v čase, mají různé typy hran a uzlů, jsou ovlivňovány akcemi různých hráčů a při jejich vyhodnocování lze využít kombinace různých datových zdrojů. Protože referenční body nejsou nezávislé a identicky distribuované, má to za následek vznik problémů při statistickém vyhodnocování. Jeden v současnosti vyvíjený nástroj se jmenuje Markovovy logické sítě, kde se kombinuje pravděpodobnostní modelování Markovových náhodných polí s logikou prvního řádu [4].

## 5 Závěry

Tradiční marketing potřebuje nové metody. Zákazníci se stále více zvykají na teleshioping, direkt mailing atd. Zároveň společnosti jako Amazon, Google a Hotmail mají úspěch, přestože neprovádějí marketing, který by byl založen výhradně na osobním doporučení. Marketing založený na osobním doporučení má klíčovou výhodu v tom, že doporučení přítele nebo podobného zdroje jsou důvěryhodné, zatímco reklama nemusí pro zákazníka být důvěryhodná. Existuje tedy pákový efekt, kdy zákazníci sami dělají marketing, což může mít za následek velmi vysokou efektivnost investice do marketingu, pokud je správně cílen. Dosud takový marketing nebyl možný. Cíl naší práce je prozkoumat možné způsoby virového marketingu a dát firmám do ruky nástroj, kterým by mohly tento marketing provádět v běžné praxi.

Mimo oblast marketingu je tento způsob optimalizace šíření slovního doporučení v sociální síti možný v oblastech, kde potřebujeme dosáhnout rozšíření informací nebo postojů k problémům s jen velmi omezenými zdroji. Příkladem může být snižování šíření HIV, boj proti kouření dospívajících a podobně. Donedávna sociologie zaostávala za dalšími přírodními vědami v rozvoji výpočtových metod. Velké množství dat, které poskytují komunikační sítě a prostředky však dovoluje zpracovat některé otázky dosud výpočtově nepokryté.

Při testech užíváme prostředí MATLAB pro reprezentaci síťových dat (vrcholy, okraje), pro výpočet charakteristických vlastností sítí a pro vizuální zobrazování sítí. MATLAB je vhodný také pro analýzy často se vyskytujících vzorů spotřebitelského chování, vzhledem k rozsahu vizualizačních procedur a maticové algebře, která je jazykem grafů, na nichž je analýza sítí založena.

## Odkazy

[1] P. Domingos and M. Richardson. *Mining the network value of customers*. In Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 57{66, San Francisco, CA, 2001. ACM Press.

[2] D. Kempe, J. Kleinberg, and E. Tardos. *Maximizing the spread of inuence in a social network*. In Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 137{146, Washington, DC, 2003. ACM Press.

[3] M. Richardson and P. Domingos. *Mining knowledge-sharing sites for viral marketing*. In Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 61{70, Edmonton, Canada, 2002. ACM Press.

[4] M. Richardson and P. Domingos. *Markov logic networks*. Technical report, Department of Computer Science and Engineering, University of Washington, Seattle, WA, 2004. <http://www.cs.washington.edu/homes/pedrod/mln.pdf>.

[5] L. Beránek. *Znalostní síť v organizaci*. Konference Znalosti pro tržní praxi. Olomouc září 2007

---

Ladislav Beránek  
Katedra informatiky  
Jihočeská universita  
Jeronýmova 10  
České Budějovice  
370 01  
beranek@pf.jcu.cz