

# Středoškolská odborná činnost

**Obor:** 1. Matematika a statistika

## The best title ever

**Autoři:** Jakub Dostál, Slovanské gymnázium Olomouc  
Nela Spurná, Slovanské gymnázium Olomouc

**Konzultant:** Chceme nějakého?, UP Olomouc

Olomouc 2016

## **Poděkování**

Neděkujeme nikomu, jsme prostě nejlepší sami o sobě

Prohlašuji, že jsem tuto práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Prohlašuji, že tištěná verze a elektronická verze soutěžní práce SOČ jsou shodné.

V Olomouci dne .....

.....

Title: New ways to simulate epidemic spreading in human society

Author: Jakub Dostál

School: Slavonic grammar school Olomouc

Supervisor: Tomáš Fürst, Department of Mathematical Analysis and Applications of Mathematics, Faculty of Science, Palacký University Olomouc

Abstract:

Society needs to fight against dangerous diseases. However, disease may stand for much more abstract process, for instance a fashion wave.

The goal of the work is to show the need to separate the topology of the network of human relations from the behaviour of the disease itself. This will reveal us substantive dependence of the simulation results on the structure of society.

First basic ways of modelling epidemic spreading are shown, both deterministic and stochastic. Traditional methods of simulation on complex networks are recalled. Then a new algorithm is presented separating behaviour of the disease from the features of the network itself.

Keywords: epidemiology, complex networks, SI, SIS, SIR, SIRS

Název práce: Nový způsob simulace šíření epidemie ve společnosti

Autor: Jakub Dostál

Škola: Slovanské gymnázium Olomouc

Konzultant: Tomáš Fürst, Katedra matematické analýzy a aplikací matematiky, Přírodovědecká fakulta, Univerzita Palackého Olomouc

Abstrakt:

I v dnešní době společnost neustále pocítuje potřebu bojovat proti nebezpečným chorobám. Pod pojmem choroba si však můžeme představit i mnohem abstraktnější pojmy, jako je třeba módní trend.

Cílem práce je ukázat potřebu oddělení vlastností topologie vztahů lidí ve společnosti, kde se epidemie šíří, a samotného charakteru choroby. Tím zároveň ukážeme velkou závislost výpočtů právě na struktuře společnosti.

Nejprve uvádíme základní způsoby modelování, a to jak deterministické, tak stochastické. Ukazujeme tradiční způsoby simulací na komplexních sítích. Poté představujeme nový algoritmus oddělující vlastnosti sítě od vlastností samotné choroby.

Klíčová slova: epidemiologie, komplexní sítě, SI, SIS, SIR, SIRS

# Obsah

<b>1</b>	<b>Co je to infomační bublina</b>	<b>3</b>
1.1	Filter bubble . . . . .	3
<b>2</b>	<b>Data a metody (nematematicky)</b>	<b>5</b>
2.1	Twitter data . . . . .	5
<b>3</b>	<b>Sentimentální analýza textu</b>	<b>5</b>
<b>4</b>	<b>Závěr</b>	<b>6</b>
	<b>Reference</b>	<b>7</b>
	<b>Přílohy</b>	<b>9</b>

# Úvod

# 1 Co je to informační bublina

Ve světě plném technického pokroku a masového používání internetu se snadno můžeme ztrácet v množství přijímaných informací. Nepřeborné množství zdrojů nás dennodenně zahlcuje spoustou nových zpráv ať už na sociálních sítích, či kdekoli jinde. Je složité se v nich orientovat, natož tak hledat a nalézat relevantní souvislosti, neboť nikde není psáno, které zprávy považovat za pravdivé a které naopak za nepravdivé zprávy anonymních komentátorů.

Role, tedy působnost žurnalistiky samotné se razantně mění. Dnešní novinář nechodí s poznámkovým blokem a tužkou v kapse, získané informace a souvislosti nepsuje doma do celistvých článků poutajících pozornost čtenářů svou kvalitou, nýbrž se smartphonem v ruce fotí, natáčí a dělá živé reportáže přímo z místa.

Toto online vysílání a sdílení má kromě spousty výhod, jakožto rychlé informovanosti uživatelů od dění v silničním provozu, po nejaktuálnější výsledky voleb, také svá úskalí. Představme si kupříkladu nebezpečí atentátu, či atentát již spáchaný. Sociální sítě se začnou naplňovat spoustou příspěvků. Odhady počtu obětí, identita pachatelů a možnost dalšího ohrožení jsou hlavními tématy. Přispívat do společné online diskuze smí každý, ať už kvalifikovaný publicista nebo náhodný autor. Zájem davů se upíná na dechberoucí události a čerstvé zprávy, proto každý příspěvek, zejména z místa činu, bude dále sdílen a zmiňován jako důvěryhodný. Kde je však záruka pravdivosti příspěvku? Nemůže nakonec dojít k situaci, kdy ten nejzajímavější, leč nepodložený fakty, bude brán ve vážnost u mas lidí právě více než ten fakty podložený, jen ne tak senzační?

Mají stále žurnalisté moc ovlivnit čtenáře a oslovit je svými pracemi, či je čtenářova pozornost upínána na senzace vyskytující se na zdech sociálních sítí? Facebook a především Twitter, fungující na rychlém tweetování zpráv, se dostává do popředí sdělovacích prostředků. Na jejich zdech sledujeme mnohé odkazy na komentáře, videa a články, formující se názory jedinců i davů, vznikající i zánikající politické myšlenky. Nakolik jsou tato samozvaná masmédia relevantní a kde mají své nedostatky, obzvláště v rovnoměrném šíření informací napříč společností, se budeme zabývat dále.

## 1.1 Filter bubble

*Filter bubble* nebo také *informační bublina* je jeden z mnoha fenoménů dnešní doby. Jako první na ni upozornil a popsal ji Eli Pariser [4, 3]. Jde o jev vyskytující se na sociálních sítích, kdy uživatel každou online aktivitou<sup>1</sup> poukazuje na oblasti jeho zájmů. Tyto algoritmy snažící se usnadnit život pomocí personalizace viděných informací se každou další činností uživatele zdokonalují, což v konečném výsledku znamená, že jsou to právě ony, jež rozhodují, co bude pro uživatele viditelné a co naopak konsekvencí prací algoritmů bude uživateli podáno v menší míře, či úplně skryto v celkovém proudu informací [2]. Celkově to tedy může vést jedince do situace, kdy místo širokého spektra příspěvků na sobě obsahově nezávislých, vidí příspěvky jen takové, jež byly vybrány preferenčními algoritmy na základě jeho předešlé činnosti a tedy velmi zúžené škále informací.

Vezmeme-li v úvahu, kolik času lidé tráví na sociálních sítích, docházíme k závěru, že jejich názory a postoje se primárně vytvářejí zrovna zde [2, 1].

---

<sup>1</sup>Tím rozumějmě kliknutí, sdílení, komentování a obdobné činnosti.

Sleduje-li uživatel pouze názorově shodné příspěvky může toV demokratických společnostech by Filter bubble mohla představovat značné ohrožení systému, neboť uživatelům předkládá již vyfiltrované příspěvky a to zejména takové, jež by podpořili názor uživatele samotného, nikoli

Potřebujeme popsát jak funguje twitter (ze je tam omezeny počet slov, jak funguje following a podobne) [5]

a taky proc je twitter relevantni ke studiu [7]



## 2 Data a metody (nematematicky)

Jak jsme již v předchozích kapitolách zmínili, s pokrokem v oblasti technologie probíhá i obrovský pokrok v oblasti čerpání zpráv a informací. Narozdíl od minulosti, drtivá většina obyvatel vyspělých států má přístup k novinkám v průběhu celého dne. Denní počet příspěvků na *Twitteru* dosahoval v roce 2010 neskutečných 6 milionů [8]. Je zřejmé, že není možné používat stejné postupy pro analýzu takto rozsáhlého množství dat jako dříve. Naštěstí se v již nějakou dobu významným tempem posouvá i oblast *Big Data*, *Machine Learningu* a jejich podoblastí. Tyto obory nacházejí uplatnění téměř ve všech oblastech dnešní vědy [12] a stále větší pozornosti se jim dostává i v sociologii [9, 10, 11].

### 2.1 Twitter data

Právě data z Twitteru jsou pro nás velmi vhodná. Jak se ukazuje [6], narozdíl od *Facebooku*, Twitter užívá mnoho lidí jako zdroj informací a zpráv o aktuálním dění ve světě. Samozřejmě můžeme zpochybňovat validitu a přesnost informací, které se na takovýchto sítích objevují. Ať už jsou takovéto obavy oprávněné, nebo ne, vlivem sociologických jevů, který v této práci provádíme to nijak neovlivňuje.

Dalším důležitým důvodem, který vedl k výběru Twitteru jako média, ze kterého budeme stahovat informace je snadná přístupnost k datům pomocí služby API poskytované právě přímo Twitterem [14]. Konkrétně pro naše účely jsme tuto službu nepoužívali přímo, ale za pomoci balíčku *tweepy* [13] pro programovací jazyk *python*. Ten umožňuje velmi snadné ovládání a filtrování proudu dat, které si vyžádáme Twitteru a také jejich okamžitou analýzu a zpracování v *pythonu*.

## 3 Sentimentální analýza textu

## 4 Závěr

## Reference

- [1] LIAO, Q. Vera a Wai-Tat FU. Beyond the filter bubble: Interactive effects of perceived threat and topic involvement on selective exposure to information. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. New York, New York, USA: ACM Press, 2013, 2359-2368. DOI: 10.1145/2470654.2481326. ISBN 9781450318990. Dostupné také z: <http://dl.acm.org/citation.cfm?doid=2470654.2481326>
- [2] GOTTRON, Thomas a Felix SCHWAGEREIT. The Impact of the Filter Bubble – A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities. *ARXIV: Computer Science - Social and Information Networks*. 2016, **2016**. 2016arXiv161206551G. Dostupné také z: <https://arxiv.org/abs/1612.06551#>
- [3] PARISER, Eli. (2011). *Beware online "filter bubbles"* [online]. Dostupné z [https://www.ted.com/talks/eli\\_pariser\\_beware\\_online\\_filter\\_bubbles](https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles)
- [4] PARISER, Eli. *The filter bubble: what the Internet is hiding from you*. New York: Penguin Press, 2011. ISBN 15-942-0300-8.
- [5] Twitter. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2017, 2017-01-22 [cit. 2017-01-22]. Dostupné z: <https://en.wikipedia.org/wiki/Twitter>
- [6] BARTHEL, MICHAEL, ELISA SHEARER, JEFFREY GOTTFRIED a AMY MITCHELL. *The Evolving Role of News on Twitter and Facebook* [online]. In: . Pew Research Center, 2015 [cit. 2017-01-22]. Dostupné z: <http://www.journalism.org/2015/07/14/the-evolving-role-of-news-on-twitter-and-facebook/>
- [7] Twitter and the News: How people use the social network to learn about the world. *American Press Institute: Insights, tools and research to advance journalism* [online]. 2015, **2015** [cit. 2017-01-22]. Dostupné z: <https://www.americanpressinstitute.org/publications/reports/survey-research/how-people-use-twitter-news/>
- [8] MATHIOUDAKIS, Michael a Nick KOUDAS. TwitterMonitor. *Proceedings of the 2010 international conference on Management of data - SIGMOD '10*. New York, New York, USA: ACM Press, 2010, **2010**, 1155-1158. DOI: 10.1145/1807167.1807306. ISBN 9781450300322. Dostupné také z: <http://portal.acm.org/citation.cfm?doid=1807167.1807306>
- [9] TINATI, Ramine, Susan HALFORD, Leslie CARR a Catherine POPE. Big Data: Methodological Challenges and Approaches for Sociological Analysis. *Sociology*. 2014, **48**(4), 663-681. DOI: 10.1177/0038038513511561. ISSN 0038-0385. Dostupné také z: <http://journals.sagepub.com/doi/10.1177/0038038513511561>

- [10] MCFARLAND, Daniel A., Kevin LEWIS a Amir GOLDBERG. Sociology in the Era of Big Data: The Ascent of Forensic Social Science. *The American Sociologist*. 2016, **47**(1), 12-35. DOI: 10.1007/s12108-015-9291-8. ISSN 0003-1232. Dostupné také z: <http://link.springer.com/10.1007/s12108-015-9291-8>
- [11] SHAH, D. V., J. N. CAPPELLA a W. R. NEUMAN. Big Data, Digital Media, and Computational Social Science: Possibilities and Perils. *The ANNALS of the American Academy of Political and Social Science*. 2015, **659**(1), 6-13. DOI: 10.1177/0002716215572084. ISSN 0002-7162. Dostupné také z: <http://ann.sagepub.com/cgi/doi/10.1177/0002716215572084>
- [12] HUBERMAN, Bernardo A. Sociology of science: Big data deserve a bigger audience. *Nature*. 2012-2-15, **482**(7385), 308-308. DOI: 10.1038/482308d. ISSN 0028-0836. Dostupné také z: <http://www.nature.com/doifinder/10.1038/482308d>
- [13] HILL, Aaron a Joshua ROESSLEIN. Tweepy. *Github* [online]. [cit. 2017-01-22]. Dostupné z: <https://github.com/tweepy/tweepy>
- [14] API Overview. *Twitter Developer Documentation* [online]. Twitter, 2016 [cit. 2017-01-22]. Dostupné z: <https://dev.twitter.com/overview/api>

# Přílohy

## První kapitola příloh