

ZÁPADOČESKÁ UNIVERZITA V PLZNI
FAKULTA APLIKOVANÝCH VĚD
KATEDRA MATEMATIKY

DIPLOMOVÁ PRÁCE

MATEMATICKÉ MODELY PENETRACE TRHU

PLZEŇ, 2020

BC. MICHAELA ŠIŠKOVÁ

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně a výhradně s použitím odborné literatury a pramenů, jejichž úplný seznam je její součástí.

V Plzni, dne 18. června 2020

.....
vlastnoruční podpis

Poděkování

Tímto bych ráda poděkovala vedoucímu diplomové práce doc. RNDr. Petru Stehlíkovi, Ph.D. za vstřícný přístup, cenné připomínky a pomoc při psaní této práce. Děkuji rovněž své rodině za podporu během celého studia.

Abstrakt

K modelování tržní penetrace po uvedení nového produktu na trh existuje mnoho různorodých matematických modelů. Obvykle uvažují nekonečnou, homogenní populaci a spojitý čas. Uvedené předpoklady mohou být v rozporu s reálným světem, ve kterém je zjevně struktura sociálních sítí jemnější a komplexnější. Kromě nereálného předpokladu nekonečné populace zmiňme naivní předpoklad homogenity. Tendence jedinců uvedený produkt zakoupit však může být ovlivněna různými faktory, např. věkem, pohlavím nebo ekonomickou situací.

Diplomová práce se proto věnuje návrhu alternativních matematických modelů uvažujících diskrétní stavový prostor a diskrétní čas, umožňujících modelování tržní penetrace (nebo obecněji šíření inovace) ve strukturované a nehomogenní populaci. Postupné šíření inovace na zvolených grafových strukturách je modelováno za pomoci Markovského řetězce s diskrétním stavovým prostorem. Jsou formulována dvě pravidla, podle kterých jedinci inovaci přijímají. V simulacích pro nehomogenní populaci je použita teorie small-world předpokládající sdružování lidí do shluků (rodina, stát, pracovní tým) a krátký řetězec osob mezi libovolnými dvěma jedinci v populaci (v řetězci jsou vždy spojeni takoví jedinci, kteří se vzájemně znají).

Simulace (a v některých jednodušších případech rovněž analytická odvození) naznačují, že inovace se obecně šíří rychleji v grafech s menším počtem vrcholů a vyšší konektivitou, tj. v menších sociálních sítích s vyšším počtem vazeb mezi jednotlivci. Rychlost šíření naopak snižuje přítomnost imunních členů v populaci (v některých případech se vlivem imunních jedinců inovace nerozšíří ke všem potenciálním zájemcům) nebo rozdíl mezi pravděpodobnostmi přijetí inovace v nehomogenních skupinách.

Klíčová slova: penetrace trhu, teorie šíření inovací, diskrétní matematické modely, sociální síť, small-world

Abstract

Many different mathematical models exist to describe market penetration after a new product is launched on a market. They usually assume an infinite continuous population and continuous time. These premises might be at odds with the real world, as the population is probably much more diverse and complex. Apart from the unrealistic assumption of an infinite population, the naive premise of homogeneous individuals should also be considered. The tendency of each individual to buy the product might vary depending on their age, gender, financial situation, and many other factors.

The goal of this master's thesis is to build alternative mathematical models considering discrete state space and discrete time, that will allow the description of market penetration (or, more generally, the diffusion of innovation), in a structured and heterogeneous population. The gradual spread of the innovation on selected graph structures is modelled using Markov chains with discrete state space. Two rules of how individuals adopt the innovation are defined. The theory of small-world networks assumes that people tend to form dense groups (family, state, project team) and there is a short chain of people connecting any two members in the social network (when people that know each other are connected in the chain).

Simulations (and analytical derivation in some simpler cases) suggest that innovation generally spreads faster in graphs with high connectivity and a smaller number of vertices (i.e. in smaller social webs with a higher number of links between individuals). On the other hand, the speed of adoption is reduced by the presence of immune individuals (the innovation will not even reach all potential adopters in some cases, when immune members are present) or by differences between probabilities of adoption in heterogeneous groups.

Keywords: market penetration, diffusion of innovation, discrete mathematical models, social network, small-world

Obsah

Prohlášení	i
Abstrakt	ii
Obsah	iv
1 Úvod	1
2 Penetrace trhu a šíření inovace	3
2.1 Teorie šíření inovací	3
2.2 Bassův model	5
2.3 Analogie marketingových a epidemiologických modelů	8
2.4 Empirické poznámky	12
3 Základní modely šíření inovace	15
3.1 Diskrétní Markovský řetězec	15
3.2 Diskrétní Markovský řetězec na úplném grafu	16
3.3 Diskrétní Markovský řetězec na k -regulárním grafu	23
3.4 Diskrétní Markovský řetězec na cestě	27
3.5 Diskrétní Markovský řetězec na hvězdě	30
3.6 Simulační srovnání	30
3.7 Konvergence k Bassovu modelu	33
4 Modely šíření inovace uvažující nehomogenní populaci	36
4.1 Nehomogenní populace	36
4.2 Small-world a šest stupňů separace	37
4.3 Přítomnost imunních jedinců v populaci	42
4.4 Nehomogenní pravděpodobnosti adopce	47
5 Závěr	50
Literatura	52

Úvod 1

Od 19. století jsme, zejména díky průmyslové revoluci, svědky obrovského technologického pokroku. Mnoho dnes běžně dostupných technologií se od prvotní fáze, ve které byly využívány jen vybraným počtem jedinců, postupně rozšířilo téměř do každé domácnosti po celém světě.

Jako příklad uveďme oblast telekomunikace a konkrétně rozmach internetu na přelomu 20. a 21. století. Podle zdroje [17] vzrostl mezi lety 2000 až 2020 počet uživatelů internetu o 1167 % a k počátku roku 2020 využívá internet 58.7 % světové populace.

Internetové připojení zjevně můžeme intuitivně označit za „inovaci“, nicméně byla by chyba omezovat tento pojem pouze na technologie. V populaci se mohou šířit i jiné typy inovací, např. informace, produkt nebo zvyk. Principům, podle kterých se obvykle inovace šíří, se věnuje tzv. teorie šíření inovací.

Této teorii se od druhé poloviny 20. století dostává velké pozornosti v mnoha sférách akademického světa. Relativně jednoduchá prvotní myšlenka z oblasti sociologie, že během jednotlivých fází šíření inovace tuto inovaci přijímají (adoptují) různé skupiny lidí, a že vývoj adopce v čase má obvykle specifický tvar, významně ovlivnila i další oblasti výzkumu.

V případě *marketingu* je rozdělení potenciálních zákazníků do kategorií jednou z cest, jak může firma vhodně sestavit marketingovou kampaň a reklamu s cílem dosáhnout s daným produktem co nejvyššího tržního podílu (tržní penetrace). Pokud by firma zároveň dokázala odhadnout, jak bude přibližně vypadat křivka popisující prodeje daného produktu v čase, umožnilo by jí to, v závislosti na aktuální fázi adopce, vhodně alokovat zdroje.

Příbuzná teorie a modely se používají také v *epidemiologii*, kde lze tuto teorii využít nejen ke zkoumání možných epidemiologických scénářů (tj. jakým způsobem se v populaci nemoc rozšíří), ale také k využití daných poznatků ke zpomalení šíření nákazy, jako tomu bylo podle zdroje [29] v případě preventivního programu s názvem „STOP AIDS“. Program byl v polovině 80. let 20. století zorganizován gay komunitou v San Franciscu a v tomto případě nemluvíme o šíření nemoci jako takové, ale o šíření informací edukativního charakteru o nezbytné prevenci, kdy tato aktivita dle zdroje [8] vedla ke dramatickému zpomalení šíření AIDS/HIV.

K modelování šíření inovací existuje množství matematických modelů, [19]. Jedním z nejznámějších je tzv. Bassův model, který z původní teorie šíření inovací vychází a předpokládá nestrukturovanou, nekonečnou a homogenní populaci a spojitý čas. Ačkoliv se Bassův model a jeho různá rozšíření, [36], prokázaly v mnohých případech jako užitečný nástroj, [30], předpoklady kladené na tyto modely mohou být v rozporu s reálným světem.

Podle zdroje [37] struktura populací velmi blízce odpovídá tzv. modelu malého světa, který předpokládá, že lidé se obvykle sdružují do menších shluků (tj. znáší určitého jedince se prav-

děpodobně znají navzájem), zároveň však délka řetězce osob mezi libovolnými dvěma jedinci na světě je překvapivě malá (v řetězci jsou vždy spojeni takoví jedinci, kteří se navzájem znají). V některých případech se uvádí, že tento řetězec je tvořen šesti osobami, o čemž pojednává tzv. teorie šesti stupňů separace, [14].

Stejně tak předpoklad, že ve vztahu k přijetí nějaké inovace je populace homogenní, je zjevně značně zjednodušený. Z pohledu marketingu by bylo naivní domnívat se, že pravděpodobnost zakoupení produktu je u všech potenciálních zákazníků stejná. Vliv na tuto pravděpodobnost mohou mít různé faktory, ať už se jedná o pohlaví cílového jedince, věk, sociální vazby nebo jeho ekonomickou situaci. Rovněž z hlediska epidemiologie se v populaci mohou vyskytovat jedinci, kteří jsou vůči dané nemoci přirozeně imunní, [3], a kteří tedy nikdy ne onemocní.

Cílem této diplomové práce je proto konstrukce alternativních matematických modelů uvažujících diskrétní stavový prostor a diskrétní čas, které umožní popis šíření inovace ve strukturované a nehomogenní populaci.

V kapitole 2 je představena teorie šíření inovací a její souvislost s marketingem a tržní penetrací. Jako jeden z modelů penetrace trhu je uveden Bassův model a je naznačena analogie s epidemiologickými modely.

Kapitola 3 definuje základní alternativní modely, které uvažují diskrétní stavový prostor. Pro představení přístupu k diskrétním modelům jsou uvažovány jednoduché grafové struktury (úplný graf, k -regulární graf, cesta, hvězda), na kterých je naznačen princip modelování šíření inovace za pomoci diskrétního Markovského řetězce. Rychlost šíření inovace je zkoumána simulačně a v některých případech také analyticky. Je zkoumán vliv parametrů simulací (graf modelující sociální síť a jeho konektivita, velikost populace, pravděpodobnost adopce, pravidlo adopce) na rychlost rozšíření inovace v populaci. Pro diskrétní Markovský řetězec na úplném grafu je naznačena souvislost s Bassovým modelem a jsou uvedeny okolnosti jeho konvergence ke spojitému modelu.

V kapitole 4 je přístup představený ve třetí kapitole rozšířen na složitější grafové struktury. Je představena teorie small-world, popisující strukturu mnoha sociálních sítí, a na základě této teorie jsou generovány odpovídající simulace. V neposlední řadě jsou uvažovány heterogenní populace s nehomogenní pravděpodobností adopce pro jednotlivé členy.

Penetrace trhu a šíření inovace 2

V roce 1957 uvedl rusko-americký matematik Igor Ansoff penetraci trhu jako jednu z možných tržně-produktových strategií. Jedná se o strategii firmy navýšit prodeje existujícího produktu ve snaze zvýšit svůj podíl na stávajícím trhu, [2].

Penetrace trhu může kromě výše zmíněné strategie označovat míru rozšíření daného produktu na trhu. Cambridgeský slovník vysvětluje pojem „market penetration“ jako „*the degree to which a product or brand is bought, used, or known by a particular group of customers*“, tedy jako „*stupeň, nakolik jsou daný produkt či značka kupovány, používány a rozpoznávány v konkrétní skupině zákazníků*“, [23].

V této práci chápeme tržní penetraci výhradně ve smyslu pronikání konkrétního produktu na trh. Trh je tvořen sítí potenciálních zákazníků a snahou firmy je rozšířit produkt k co největšímu počtu jedinců.

2.1 Teorie šíření inovací

Na penetraci trhu se lze podívat o něco obecněji. Předmětem šíření nemusí být přímo daný produkt. Populací se může šířit rovněž nová myšlenka, zpráva, zvyk či obecně nějaký objekt, který by jedinci v populaci označili za „nový“ (tj. dosud nerozšířený). Obecně lze mluvit o *inovaci*.

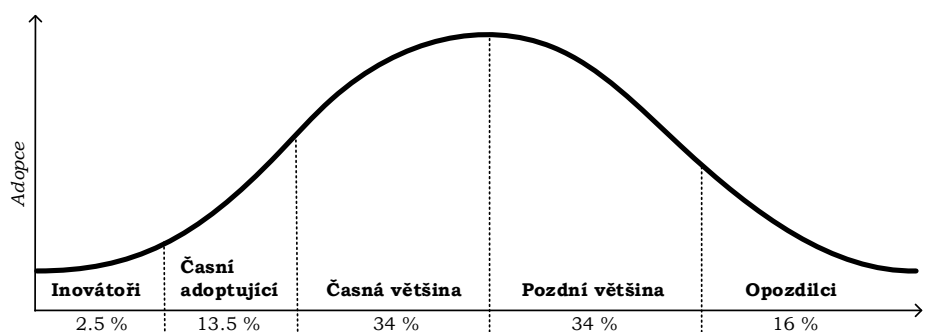
Jednou ze zásadních sociálních teorií, která o šíření inovací v populaci pojednává, je teorie šíření inovací, kterou v roce 1962 ve své knize *Diffusion of Innovations* [29] formuloval sociolog Everett M. Rogers.

Rogers popisuje šíření inovace jako proces, ve kterém je inovace postupně komunikována napříč jedinci v populaci, kdy tito jedinci komunikovanou inovaci postupně adoptují (tj. přijímají myšlenku, zprávu, osvojují si zvyk nebo kupují produkt). Protože jedinci v populaci obvykle inovaci neadoptují všichni naráz, jsou na základě času, kdy danou inovaci adoptují, rozdělení do pěti kategorií:

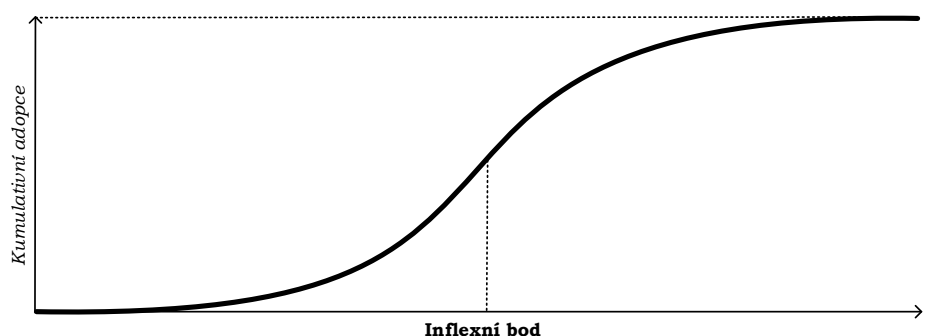
1. Inovátoři,
2. Časní adoptující,
3. Časná většina,
4. Pozdní většina,

5. Opozdilci.

Na Obrázku 2.1 je naznačeno rozdělení počtů adoptujících v závislosti na čase včetně vyznačení jednotlivých kategorií adoptujících a jejich přibližného procentuálního zastoupení, jak je uvádí Rogers. Podle Rogerse prozatímní výzkum difúze inovací z empirických dat naznačuje, že se počet adoptujících v závislosti na čase řídí normálním rozdělením.



Obrázek 2.1: Šíření inovace dle Rogerse [29]: Rozložení adopce v čase a kategorie adoptujících s přibližným procentuálním zastoupením v jednotlivých kategoriích.



Obrázek 2.2: Šíření inovace dle Rogerse [29]: Rozložení kumulativní adopce v čase ve tvaru S-křivky.

Přestože je zřejmě vyčlenění členů populace do pěti skupin zjednodušením v realitě složitější struktury, tento přístup poskytuje možnost srovnání. Rogers ve své knize popisuje „ideální“ zástupce jednotlivých kategorií. Pokud by firma vhodně odhadla, v jaké fázi pronikání na trh se nachází, může schopnost charakterizovat cílovou skupinu významně přispět k vhodně nastavené marketingové kampani. Například inovátoři jsou popisováni jako jedinci se zálibou ve zkoušení nových, neokoukaných, a v některých případech dokonce riskantních, věcí. Naproti tomu opozdilci jsou charakterizováni jako skeptičtí, opatrní jedinci, kteří inovaci adoptují až poté, co tak učinila většina populace.

Kumulativní verze šíření inovace (tj. rozložení kumulativního počtu adoptujících v čase) je naznačena na Obrázku 2.2. Tato funkce nabývá ve většině případů tvaru S-křivky (jedná

se o tzv. sigmoidální funkci). Ze začátku celkový počet adoptujících roste pomalu (adoptují jedinci z nejmenší skupiny inovátorů), ale se zvětšujícím se počtem adoptujících roste i rychlost adopce. Po překonání inflexního bodu S-křivky se rychlost šíření opět zpomaluje, protože v populaci zbývá menší počet jedinců, kteří dosud neadoptovali (pozdní většina nebo opozdilci). Nakonec se počet adoptujících ustálí (velikost populace je omezena) a proces šíření je dokončen.

2.2 Bassův model

Nevýhodou Rogersovy teorie šíření inovace je, že se nedá použít k predikci. Ačkoliv díky této teorii tušíme přibližný tvar rozdělení adoptujících v čase, nedá se snadno určit, v jaké části procesu se aktuálně firma s daným produktem nachází. Tento nedostatek pomohl v roce 1969 vyřešit matematický model sestavený Frankem M. Bassem v jeho článku [4] představujícím spojitý difúzní model popisující pronikání nového produktu na trh. Jak Bass sám uvádí, [5], model vychází z původní Rogersovy idey teorie šíření inovace, kterou kvantifikuje. Model je znám jako „Bassův model“.

Bassův model popisuje šíření nového produktu (penetraci trhu) pro první pořízení daného produktu. Bass ve své práci vychází z Rogersovy teorie, podle níž se část populace rozhodne inovaci adoptovat nezávisle na ostatních. To je typické pro 1. kategorii – kategorii inovátorů, která dle Rogerse pojímá zhruba 2.5 % z celkové populace.

Jedinci ze zbývajících kategorií jsou ve své adopci do nějaké míry ovlivněni tlakem společnosti, kdy s rostoucím počtem jedinců, kteří inovaci již adoptovali, roste také tlak k adoptování na zbývajících, dosud neadoptujících, jedince. Tato část populace je označována jako imitátoři a Bass uvádí, že se jedná o jedince ze skupin 2 až 5 v původní Rogersovo teorii.

Bassův model vychází z předpokladu, že pravděpodobnost $P(t)$, že daný jedinec ze skupiny imitátorů inovaci adoptuje v čase t , jestliže ji dosud neadoptoval, je lineární funkce počtu jedinců, kteří inovaci dosud adoptovali, ve tvaru

$$P(t) = p + \frac{q}{m} A(t). \quad (2.1)$$

Výrazy $p \geq 0$, $q > 0$ a $m > 0$ jsou konstanty, $A(t)$ je funkce kumulativního počtu adoptujících do času t . Protože $A(0) = 0$, parametr p udává pravděpodobnost adopce inovátorů v čase $t = 0$. Bass uvádí, že velikost p reflektuje vliv inovátorů ve společnosti. Výraz $\frac{q}{m} A(t)$ modeluje zvyšující se tlak k adopci na imitátory při růstu počtu jedinců, kteří již adoptovali. V literatuře bývá parametr p označován jako „koeficient inovace“ a parametr q jako „koeficient imitace“. Konstanta m označuje počet potenciálních adoptujících.

Bassův model má dva klíčové předpoklady. V první řadě uvažuje, že všichni potenciální adoptující inovaci časem adoptují. Za druhé, opětovná adopce (např. opětovné pořízení daného produktu) není v modelu uvažována, a v datech by proto neměla být zahrnuta.

Podmíněnou pravděpodobnost $P(t)$, že daný jedinec v čase t adoptuje, jestliže dosud neadoptoval, můžeme napsat ve tvaru

$$\frac{f(t)}{1 - F(t)} = P(t) = p + \frac{q}{m} A(t) = p + qF(t), \quad (2.2)$$

kde $f(t)$ je pravděpodobnost adopce v čase t , $F(t) = \int_0^t f(s) ds$ (kde $F(0) = 0$), je pravděpodobnost adopce do času t , m je velikost trhu (počet potenciálních adoptujících) a $A(t)$ a $F(t)$ jsou svázány vztahem

$$A(t) = mF(t).$$

Hodnoty funkce $F(t)$ jsou tedy normalizovány do intervalu $[0, 1]$.

Pro model procesu šíření inovací navrhl Bass nelineární diferenciální rovnici prvního řádu ve tvaru

$$\frac{dA(t)}{dt} = \left(p + \frac{q}{m}A(t)\right)(m - A(t)), \quad t > 0. \quad (2.3)$$

Připojíme-li přirozenou počáteční podmínku $A(0) = 0$ (a přejdeme-li současně ke stručnějším značení pro derivaci $A'(t) = \frac{dA(t)}{dt}$), dostaneme Bassovu počáteční úlohu

$$\begin{cases} A'(t) = \left(p + \frac{q}{m}A(t)\right)(m - A(t)), & t > 0, \\ A(0) = 0, \end{cases}$$

kteřou lze pomocí separace proměnných explicitně vyřešit. Řešení má tvar

$$A(t) = m \left(\frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}} \right). \quad (2.4)$$

Rovnice (2.4) popisuje výše zmíněnou S-křivku, kterou lze použít k modelování celkové (kumulativní) adopce v závislosti na čase, [30].

Alternativně můžeme sestavit Bassův model pro normalizovanou funkci $F(t)$. Funkce $F(t)$ popisuje procentuální penetraci trhu v populaci. Vydělíme-li rovnici (2.3) počtem potenciálních adoptujících $m > 0$, dostaneme

$$\begin{aligned} \frac{1}{m}A'(t) &= \frac{1}{m} \left(p + \frac{q}{m}A(t)\right)(m - A(t)), \\ \frac{1}{m}A'(t) &= \left(p + \frac{q}{m}A(t)\right) \left(1 - \frac{A(t)}{m}\right), \end{aligned}$$

a následně dostáváme normalizovanou Bassovu diferenciální rovnici

$$F'(t) = (p + qF(t))(1 - F(t)), \quad t > 0. \quad (2.5)$$

Řešení počáteční úlohy s $F(0) = 0$ má následně tvar

$$F(t) = \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}}. \quad (2.6)$$

Funkci $f(t)$, popisující podíl adoptujících v čase t , získáme derivací $F(t)$ podle času

$$F'(t) = f(t) = \frac{p(p+q)^2 e^{-(p+q)t}}{(p + qe^{-(p+q)t})^2}. \quad (2.7)$$

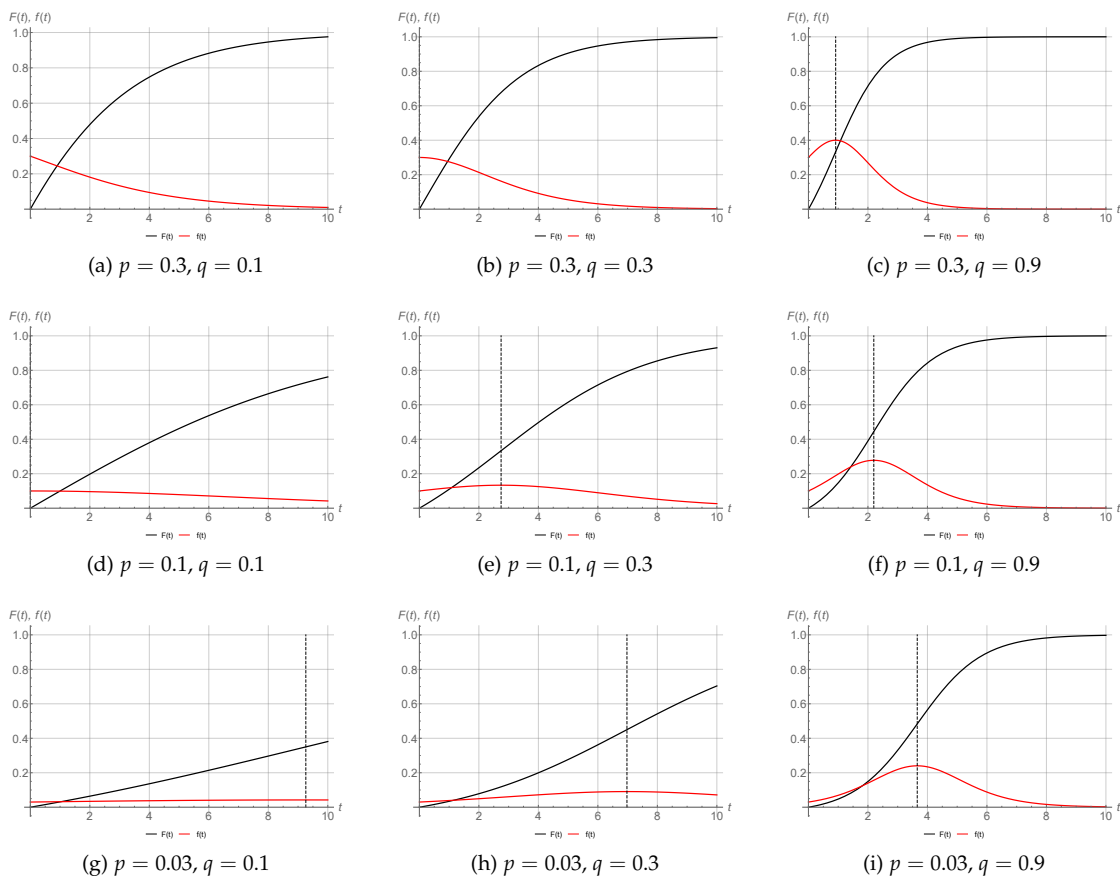
Bod nejrychlejšího růstu funkce $F(t)$, odpovídající bodu maxima $f(t)$, označíme jako t^* . Jedná se o bod, ve kterém jsou funkce $F''(t)$, resp. $f'(t)$, rovny nule, tj. dostáváme rovnici

$$F''(t^*) = f'(t^*) = \frac{p(p+q)^3 e^{-(p+q)t^*} (qe^{-(p+q)t^*} - p)}{(p + qe^{-(p+q)t^*})^3} = 0 \quad (2.8)$$

a jejím řešením je bod

$$t^* = -\frac{\ln \frac{p}{q}}{(p+q)}. \quad (2.9)$$

Bod t^* lze interpretovat jako bod nejrychlejší penetrace či jako bod vrcholu (tzv. „peaku“) adopcce. Na Obrázku 2.3 jsou funkce $F(t)$ a $f(t)$ vykresleny pro různé hodnoty parametrů p a q . Poloha bodu t^* je naznačena přerušovanou čarou v případech, ve kterých to dává smysl (tj. $p < q$).



Obrázek 2.3: Vykreslení funkcí $F(t)$ a $f(t)$ popsaných rovnicemi (2.6), resp. (2.7), pro různé hodnoty parametrů p a q se znázorněným bodem peaku t^* (přerušovaná čára).

Normalizovaná Bassova diferenciální rovnice (2.5) má přirozený úzký vztah na jeden z nejvýznamnějších růstových populačních modelů, logistickou diferenciální rovnici.

Poznámka 2.1. Uvažujeme-li speciální případ normalizovaného Bassova modelu (2.5) s $p = 0$, dostáváme (normalizovanou) logistickou diferenciální rovnici

$$F'(t) = qF(t)(1 - F(t)), \quad t > 0,$$

resp. její nenormalizovanou verzi

$$A'(t) = qA(t)(m - A(t)), \quad t > 0. \quad (2.10)$$

Rovnice (2.10) se nazývá logistická diferenciální rovnice a představuje jeden z nejpoužívanějších modelů populačního růstu. Byla poprvé odvozena již v roce 1838 belgickým matematikem Pierrem Françoisem Verhulstem. Parametr $m > 0$ zde má přirozenou interpretaci kapacity prostředí, více v [27]. Bassův model lze tedy vnímat jako drobné zobecnění logistického růstu.

Tato vazba nám odkrývá důležitou vlastnost Bassova modelu, a tou je spojitý čas. Fakt, že je pro modelování penetrace trhu používána diferenciální, a ne diferenční, rovnice, plyne jednoduše z chaotického chování diskrétní analogie logistické diferenciální rovnice (2.10) (viz např. [32]), které přímo implikuje i chaotické chování diskrétní analogie Bassova modelu (2.5) mající tvar např.

$$F_{n+1} = F_n + (p + qF_n)(1 - F_n), \quad n \in \mathbb{N}.$$

Vzhledem k diskrétní povaze dat, která slouží k odhadování parametrů Bassova modelu, byly navrženy i jiné diskrétní verze, [30], zaručující nechaotické chování. Jejich nepřirozené vlastnosti (jako nemonotonie trajektorií či absence přirozené interpretace) nicméně objasňují drtivou převahu používání spojitých modelů.

Kromě Bassova modelu, který je založen na logistickém růstu, lze k modelování procesu šíření inovace a celkové adopce v závislosti na čase použít i jiné modely, mající mírně odlišné analytické vlastnosti, [18]. Jeden z takových přístupů je založen na tzv. Gompertzově růstu

$$F(t) = me^{-Ae^{-Bt}},$$

kde m je maximální velikost populace a $A, B > 0$ jsou parametry udávající tvar růstové křivky.

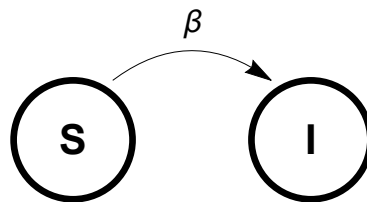
2.3 Analogie marketingových a epidemiologických modelů

Přestože byly modely šíření inovací doposud popisovány především v kontextu marketingu a penetrace trhu, dané modely mají úzkou souvislost s modelováním procesu šíření nákazy v populaci, [35].

Proto i v následujících kapitolách mluvíme obecně o „inovaci“ (myšlenka, produkt, nemoc, atd.), „adopci“ (jedinec přijme myšlenku, koupí produkt, nakazí se nemocí) a jestliže daný jedinec inovaci adoptoval, mluvíme o něm jako o „nakaženém“ jedinci.

Abychom vazbu marketingových a epidemiologických modelů dobře objasnili, představme si základní epidemiologické modely, [12]. Jedná se o tzv. přihrádkové modely (angl. *compartmental models*), kde každý jedinec populace je v jedné ze dvou (tří, obecně $n \in \mathbb{N}$) skupin dle fáze nakažení, tzv. přihrádek.

Nejjednodušším příkladem přihrádkového epidemiologického modelu je SI model, který uvažuje dvě přihrádky. SI model se schematicky popisuje tzv. diagramem (schématem) přihrádkového modelu, viz Obr. 2.4. Každý jedinec (nekonečně a homogenní) populace se buď může



Obrázek 2.4: Schéma přihrádkového SI modelu (2.11).

nacházet ve stavu S – susceptible (zdravý a náchylný k nakažení), resp. I – infected (nakažený).

Matematicky je model SI popsán soustavou dvou diferenciálních rovnic

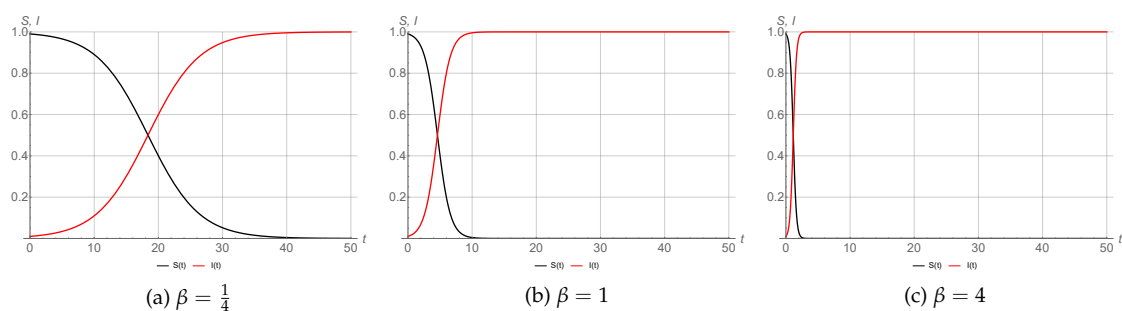
$$\begin{cases} S'(t) = -\frac{\beta I(t)S(t)}{N}, & t > 0, \\ I'(t) = \frac{\beta I(t)S(t)}{N}, \end{cases} \quad (2.11)$$

kde $S(t), I(t)$ je počet náchylných a nakažených v čase t ,

$$N = S(t) + I(t) \quad (2.12)$$

je konstantní velikost populace a $\beta > 0$ je konstanta popisující intenzitu šíření nemoci.

Na Obrázku 2.5 je řešení modelu SI v závislosti na čase t pro různé hodnoty koeficientu β .



Obrázek 2.5: Epidemiologický model SI. Počáteční podmínka odpovídá 1 % nakažených v populaci v $t = 0$.

Z epidemiologického hlediska je tento základní SI model nepraktický, protože nakaženým neumožňuje uzdravení, nakažení ani neumírají. Je ale ideální pro ukázání vazby mezi marketingovým Bassovým modelem (2.5), populačním Volterrovým modelem (2.10) a epidemiologickými modely, [31]. Využijeme-li totiž vztahu (2.12), který lze přepsat jako $S(t) = N - I(t)$, můžeme ho dosadit do druhé rovnice v (2.11) a získáme jednu diferenciální rovnici pro počet nakažených

$$I'(t) = \frac{\beta I(t)S(t)}{N} = \frac{\beta I(t)(N - I(t))}{N}.$$

Označíme-li si novou konstantu $\bar{\beta} = \frac{\beta}{N}$, tak získáváme

$$I'(t) = \bar{\beta} I(t)(N - I(t)), \quad (2.13)$$

což je přímo model ekvivalentní s populačním Volterrovým modelem (2.10) a ve smyslu Poznámky 2.1 i s Bassovým modelem (2.5) s $p = 0$.

Vztah marketingových a epidemiologických (či populačních) modelů je tedy velice úzký, i když není zcela identický, zejména kvůli přítomnosti konstanty p popisující podíl inovátorů v Bassově modelu (2.5). Nicméně, široká rozpracovanost např. přihrádkových modelů v epidemiologii a jejich rozmanitých variant a rozšíření poskytuje nepřehledné množství možností rozšiřování modelů.

Uveďme si několik dalších možných směrů pro rozšíření marketingových modelů na základě epidemiologických modelů, pro přehled zejména spojených epidemiologických modelů odkazujeme např. na knihy [1, 12] a přehledový článek [6].

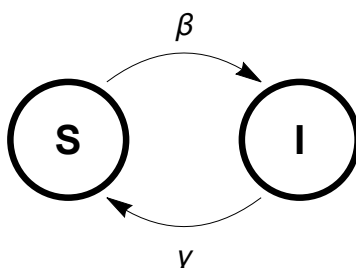
- SIS model umožňující opakovanou nákazu (sezónní chřipka) odpovídající např. opakovaným nákupům, viz schema na Obr. 2.6,

$$\begin{cases} S'(t) = -\frac{\beta I(t)S(t)}{N} + \gamma I(t), & t > 0, \\ I'(t) = \frac{\beta I(t)S(t)}{N} - \gamma I(t), \end{cases} \quad (2.14)$$

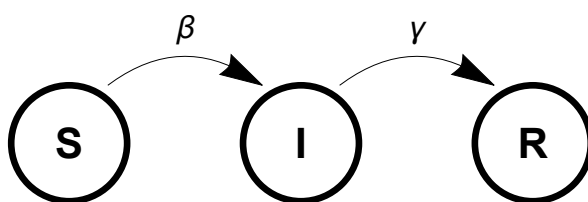
- SIR model umožňující úplné vyléčení (nemoci, které lze prodělat jen jednou a vytvářející protilátky, např. spalničky, zarděnky, příušnice, ...) odpovídající např. jednorázovému předplatnému či přechodu k lepší technologii, viz schema na Obr. 2.7,

$$\begin{cases} S'(t) = -\frac{\beta I(t)S(t)}{N}, & t > 0, \\ I'(t) = \frac{\beta I(t)S(t)}{N} - \gamma I(t), \\ R'(t) = \gamma I(t), \end{cases} \quad (2.15)$$

- buněčné automaty typu Greenberg-Hastings (viz např. [12, Kap. 6.2.]) umožňující studium malých populací,
- prostorově heterogenní modely šíření epidemií (viz např. [20]), umožňující studium vlivu struktury vazeb uvnitř dané (konečné) populace.

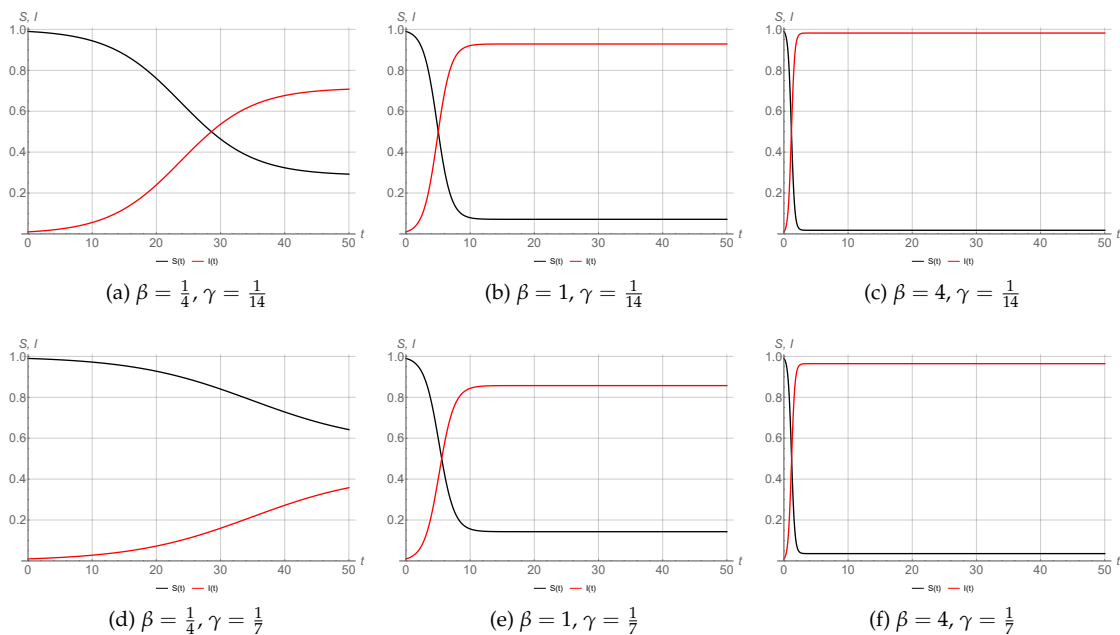


Obrázek 2.6: Schéma přihrádkového SIS modelu (2.14).

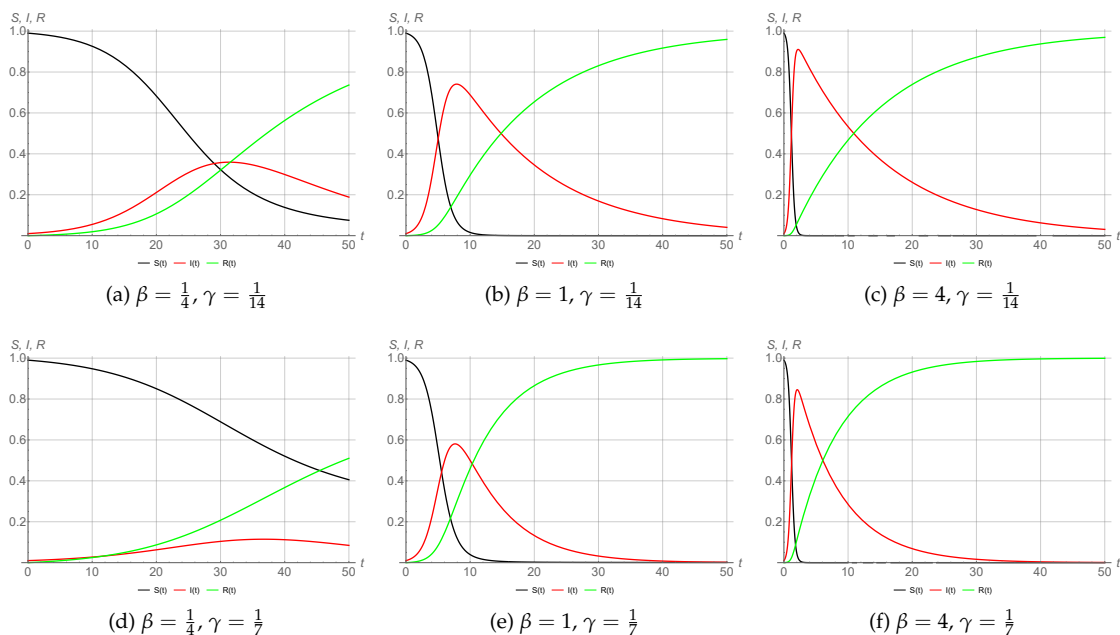


Obrázek 2.7: Schéma přihrádkového SIR modelu (2.15).

Na Obrázcích 2.8 a 2.9 jsou vykreslena řešení modelů SIS a SIR v závislosti na čase t pro různé hodnoty parametrů β a γ . Hodnota parametru $\beta = \frac{1}{4}$ modeluje situaci, kdy nakažený jedinec z přihrádky I nakazí v průměru jednoho zdravého člověka za čtyři časové jednotky, uvažujme dále dny. V případě $\gamma = \frac{1}{14}$ trvá průměrně 14 dní, než se nakažený jedinec vyléčí (přesune se z přihrádky I do S v případě SIS modelu nebo se přesune z I do R v případě SIR modelu).



Obrázek 2.8: Epidemiologický model SIS. Počáteční podmínka odpovídá 1 % nakažených v populaci v $t = 0$.



Obrázek 2.9: Epidemiologický model SIR. Počáteční podmínka odpovídá 1 % nakažených v populaci v $t = 0$.

2.4 Empirické poznámky

Díky velké aplikovatelnosti Bassova modelu existuje množství článků zabývajících se odhady parametrů p , q a m Bassova modelu napříč různými produkty na trhu, např. [24], [25], [30], [34].

Podle Liliena, Rangaswama a De Bruyna [21] je jednou z výhod Bassova modelu jeho „flexibilita“ ve smyslu modelování různých scénářů prodeje daného produktu přes různé hodnoty parametrů p a q :

- Jestliže $q > p$, dominuje efekt imitace nad efektem inovace a funkce adopce $f(t)$ má tvar převráceného U.
- Pro $q < p$ naopak dominuje efekt inovace, tj. největší prodeje jsou zaznamenány při představení produktu a funkce adopce $f(t)$ postupně klesá.
- Jestliže p i q jsou vysoké, pak produkt zaznamená rychlý nárůst prodeje a zároveň jejich rychlý pokles po dosažení maximální hodnoty (peaku).

Zdroj [21] dále zmiňuje, že k predikci prodeje je vhodné nejprve samostatně odhadnout tržní potenciál m (např. průzkumem dlouhodobých obchodních záměrů) a následně pro dané m použít nelineární regresi k odhadu parametrů p a q Bassova modelu.

Satoh [30] uvádí, že pro odhad parametrů p a q se používá např. metoda nejmenších čtverců, metoda maximální věrohodnosti nebo metoda nelineárních nejmenších čtverců. Zatímco metoda nelineárních čtverců obecně dosahuje uspokojivých výsledků při použití k predikci, výhoda obyčejných nejmenších čtverců je v jejich relativně jednoduché implementaci.

V Tabulce 2.1 je uvedeno několik příkladů průměrných odhadů parametrů pro různé produkty vč. zdrojů těchto odhadů. Odhady p a q ze zdroje [24] pocházejí z dat prodeje v USA, v případě zdroje [25] jde o odhady provedené na základě prodeje na německém trhu.

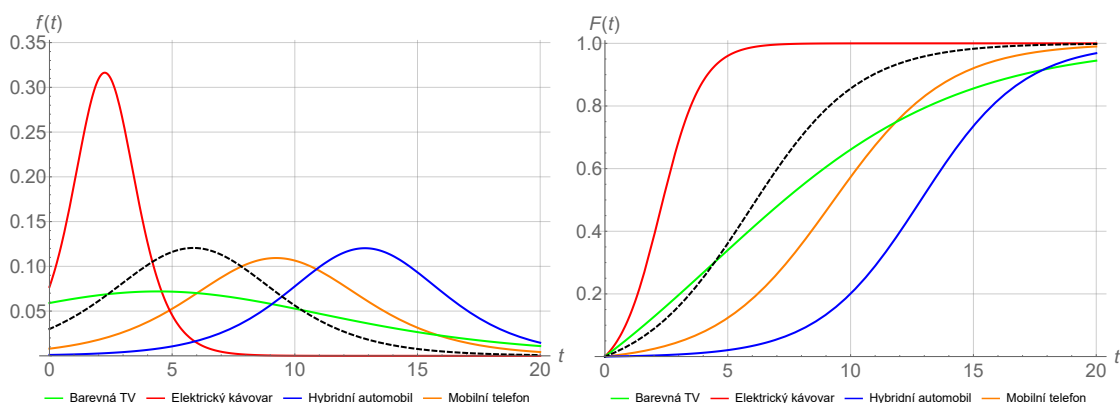
	p	q
Barevná TV, [24]	0.059	0.146
Elektrický kávovar, [24]	0.077	1.106
Hybridní automobil, [25]	0.001	0.479
Mobilní telefon, [24]	0.008	0.421

Tabulka 2.1: Odhady parametrů p a q Bassova modelu pro různé produkty.

Chandrasekaran a Tellis [11] uvádí, že

- průměr parametru inovace p napříč širším spektrem produktů leží mezi hodnotami 0.027 a 0.03 s relativně malým rozptylem pro různé typy produktů, zatímco
- průměr parametru imitace q napříč širším spektrem produktů leží mezi hodnotami 0.38 a 0.42 s velkým rozptylem odhadu závislejícím na typu produktu a dalších faktorech.

Odlišnost tvaru křivek Bassova modelu pro produkty uvedené v Tabulce 2.1 je pro funkci adopce $f(t)$ na Obrázku 2.10a, pro funkci kumulativní adopce $F(t)$ na Obrázku 2.10b. Přerušovanou čarou je vykreslen tvar křivek pro průměrné hodnoty odhadů (resp. horní hranici odhadovaných intervalů) $p = 0.03$ a $q = 0.42$ podle Chandrasekarana a Tellise, [11].



(a) Funkce adopce $f(t)$ pro produkty uvedené v Tabulce 2.1. (b) Funkce kumulativní adopce $F(t)$ pro produkty uvedené v Tabulce 2.1. Křivka pro uváděné průměrné hodnoty $p = 0.03$ a $q = 0.42$ je naznačena přerušovanou čarou.

Faktory ovlivňující parametry Bassova modelu

Hodnoty odhadů parametrů Bassova modelu závisí na různých faktorech. Významným faktorem je *kategorie typu produktu*, kdy např. průmyslové či zdravotnické produkty mají obecně vyšší koeficient inovace než jiné typy zboží, [11].

Zmíňme také závislost na *regionu* (např. státu), ve kterém je produkt uveden. Obecně platí, že jak průměrný koeficient inovace, tak průměrný koeficient imitace, jsou pro rozvinuté země vyšší, než pro země rozvojové. Zároveň Van den Bulte [34] uvádí, že v zemích s tradičně kolektivistickou mentalitou je hodnota parametru imitace q vyšší, než v zemích s mentalitou individualistickou (jako příklad je uvedeno Japonsko jako zástupce spíše kolektivistické kultury a USA jako zástupce spíše individualistické kultury).

V neposlední řadě uvedme faktor *času*, ve kterém je produkt představen. Z empirických dat se ukazuje, že některé fáze penetrace trhu se postupně zkracují. Chandrasekaran a Tellis [11] uvádí, že zatímco produkt, který v roce 1946 v USA dosahoval 5% penetrace trhu, potřeboval k rozšíření z 10 % na 90 % odhadované maximální kumulativní adopce cca 14 let, v roce 1980 se tento čas zkrátil v průměru na polovinu, tj. 7 let.

Výhody a nevýhody použití Bassova modelu

Ačkoliv se Bassův model v mnoha případech prokázal jako užitečný nástroj k odhadům prodejů nově představených produktů, objevují se i kritické názory na použití difúzních modelů a jejich spolehlivost k odhadování penetrace trhu. Heeler a Hustad [16] uvádí, že difúzní modely dosahují v některých případech podprůměrných výsledků, Bernhardt a MacKenzie [7] dokonce uvádí, že nadprůměrné výsledky těchto modelů jsou v některých případech zapříčiněny „*prozíravým výběrem situace, populace, inovace a časového intervalu pro vyhodnocení výsledků odhadů dat*“.

Zdroj [22] uvádí, že přestože Bassův model obvykle dobře aproximuje roční prodeje, v případě čtvrtletních či dokonce měsíčních dat bývá jeho použití vzhledem k možné sezónnosti v datech méně účinné. Jestliže by však firma chtěla daný model použít k predikci úspěšnosti nového produktu, v drtivém množství případů je použití ročních dat příliš pozdě na vyhodnocení klíčových investic.

Nesporná výhoda použití Bassova modelu je v jeho relativní jednoduchosti a snadné implementaci. Nicméně tato jednoduchost souvisí s mnohdy nereálnými předpoklady na model. Ve spojitém modelu je každý imitátor okamžitě ovlivněn zvyšujícím se počtem adoptujících, populace je homogenní a nestrukturovaná.

Tyto předpoklady jsou zjevně v rozporu s reálným světem, ve kterém jsou někteří jedinci v populaci do jisté míry „izolováni“ od ostatních (ať už fyzickou nebo sociální vzdáleností), někteří potenciální adoptující mají vyšší pravděpodobnost adopce než jiní (např. u nové technologie může mít vyšší pravděpodobnost adopce jedinec z ekonomicky aktivní skupiny obyvatel), atd.

Vzhledem k uvedeným nevýhodám spojitého Bassova modelu se dále zabýváme modely uvažujícími konečnou a strukturovanou populaci a diskrétní čas, které však, stejně jako Bassův model, z myšlenky šíření inovací vycházejí.

Základní modely šíření inovace 3

3.1 Diskrétní Markovský řetězec

V následující kapitole budeme šíření inovace napříč konečnou a strukturovanou populací modelovat za pomoci diskrétního Markovského řetězce. V nejjednodušším případě úplného grafu půjde o Markovský řetězec s množinou stavů $S = \{1, 2, \dots, n\}$, kde $n \in \mathbb{N}$ je maximální velikost populace. Jednotlivé stavy řetězce $i \in S$ jsou počty jedinců, kteří inovaci již adoptovali. V případě obecného grafu (obecné heterogenní struktury) bude množina stavů mnohem bohatší. V nejextrémnějším případě může obsahovat 2^n stavů. Počáteční rozdělení $\mu^{(0)}$ popisuje počáteční podmínku stavu adopce (přijmutí) dané inovace. V grafech je, bez újmy na obecnosti, počáteční podmínka na indexu $t = 1$.

K popisu daných populačních struktur použijeme neorientovaný graf $G = (V, E)$, kde V je množina vrcholů a E je množina neorientovaných hran. Populaci velikosti n modelujeme grafem na $n = |V|$ vrcholech, kde každý vrchol označuje konkrétního člena populace a přítomnost hrany $\{u, v\} \in E$ označuje, že se jedinci u a v navzájem znají (tj. mohou se vzájemně ovlivňovat v adoptování dané inovace). Graf G budeme popisovat maticí sousednosti $A(G)$ jak pro analytické, tak simulační postupy.

Každý jedinec (resp. vrchol grafu $v \in V$) se může nacházet v jednom ze dvou stavů. Stav 0 odpovídá situaci, kdy daný jedinec uvažovanou inovaci ještě neadoptoval a budeme v návaznosti na souvislost s epidemiologickými modely mluvit o nenakaženém jedinci. Stav 1 naopak bude odpovídat jedinci, který již danou inovaci adoptoval, a budeme hovořit o nakaženém jedinci.

Formulujme dále uvažovaná „pravidla“ adopce. Prvním pravidlem (označme ho jako pravidlo A) je přístup, ve kterém pravděpodobnost adopce závisí na počtu nakažených sousedů. Pokud má nenakažený vrchol m nakažených sousedů, pak ho každý z m sousedů může v daném kroku nakazit s pravděpodobností p . Pravděpodobnost adopce tedy závisí na počtu nakažených sousedů.

Druhým pravidlem (označme ho jako pravidlo B) je přístup, ve kterém se jedinec nakazí s pravděpodobností p v případě, že má alespoň jednoho nakaženého souseda. Pravděpodobnost adopce tedy nezávisí na počtu nakažených sousedů (v případě, že má jedinec alespoň jednoho nakaženého souseda). Pravidla adopce tedy jsou:

- *Pravidlo A*: Každý nakažený soused může daného jedince nakazit s pravděpodobností p .

S každým nakaženým sousem tedy roste pravděpodobnost, že jedinec inovaci rovněž adoptuje.

- *Pravidlo B*: Pokud má jedinec alespoň jednoho nakaženého sousem, jeho pravděpodobnost adopce je p . Vyšší počet nakažených sousemů nezvyšuje pro daného jedince pravděpodobnost adopce.

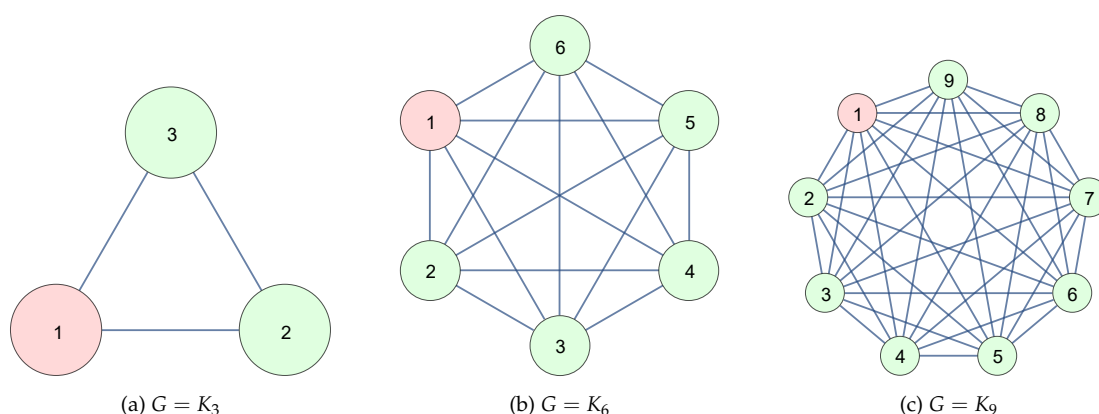
Na základě zvoleného pravidla adopce pak definujeme pravděpodobnost adopce p . V případě pravidla A jde o pravděpodobnost, s jakou má každý nakažený sousem šanci „zdravého“ jedince nakazit. V případě pravidla B jde o pravděpodobnost, s jakou se zdravý jedinec nakazí, pokud má alespoň jednoho nakaženého sousem.

Mezi parametry modelu šíření inovace tedy zahrneme

- graf $G = (V, E)$ na n vrcholech reprezentovaný maticí sousednosti $A(G)$,
- velikost populace $n = |V|$,
- počáteční rozdělení inovace $\mu^{(0)}$,
- pravidlo adopce (pravidlo A nebo B),
- pravděpodobnost adopce p .

3.2 Diskrétní Markovský řetězec na úplném grafu

Jedním z nejjednodušších uvažovaných modelů struktury konečné populace je úplný graf K_n na n vrcholech. V úplném grafu vede hrana mezi každými dvěma vrcholy, tj. modelujeme situaci, kdy se každý jedinec zná s každým (např. uží rodina, studijní kruh, projektový tým, ...). Vizualizace takové struktury pro různá n je na Obrázku 3.1. Pro jednoduchost uvažujme situaci, kdy se na počátku v dané populaci vyskytne právě jeden nakažený člen a bez újmy na obecnosti uvažujme, že tento nakažený je popsán vrcholem na pozici 1, tj. $\mu^{(0)} = (1, 0, \dots, 0)^T$. Nakažený vrchol je v grafech zvýrazněn červeně.



Obrázek 3.1: Modely populace pro různá n , úplný graf $G = K_n$ s naznačenou počáteční podmínkou jednoho nakaženého v populaci (červený vrchol).

Počáteční rozdělení Markovského řetězce je vektor délky n s jedničkou na pozici nakaženého vrcholu a nulami jinde. Kromě stavu, kdy je celá populace nakažena, jsou všechny stavy přechodné. Situace, kdy je nakaženo všech n jedinců v populaci, je stav absorpční, protože neuvažujeme možnost vyléčení.

Pro úplný graf na n vrcholech a pravidlo A (tj. zvyšující se počet nakažených sousedů zvyšuje pravděpodobnost, že daný vrchol inovaci adoptuje) bude matice pravděpodobností přechodu $P^A = \{p_{ij}^A, i, j \in S\}$ matice $n \times n$ s prvky na pozici p_{ij}^A ve tvaru

$$p_{ij}^A = \binom{n-i}{j-i} (1 - (1-p)^i)^{j-i} ((1-p)^i)^{n-j}, \quad i \leq j. \quad (3.1)$$

Jedná se o horní trojúhelníkovou matici, protože neuvažujeme možnost vyléčení – při stavu i , kdy i je počet nakažených, můžeme zůstat v současném stavu (nikdo se v daném kroku nenakazí) nebo nakazit až $n - i$ jedinců. Při přechodu ze stavu i do stavu j , kde $i \leq j$, tedy z nenakažených $n - i$ jedinců vybíráme $j - i$ takových, kteří se při přechodu z i do j nakazí. Kombinační číslo $\binom{n-i}{j-i}$ popisuje počet všech kombinací, jak tyto jedince vybrat. Výraz ve tvaru $(1 - (1-p)^i)^{j-i}$ označuje pravděpodobnost nákazy $j - i$ jedinců, kteří se při daném přechodu z i do j nakazí, výraz $((1-p)^i)^{n-j}$ označuje pravděpodobnost toho, že se zbývajících $n - j$ jedinců při přechodu z i do j nenakazí.

Například pro populaci o velikosti $n = 3$, znázorněnou na Obrázku 3.1a, je matice pravděpodobností přechodu matice 3×3 ve tvaru

$$P_{K_3}^A = \begin{pmatrix} (1-p)^2 & 2p(1-p) & p^2 \\ 0 & (1-p)^2 & 1 - (1-p)^2 \\ 0 & 0 & 1 \end{pmatrix}. \quad (3.2)$$

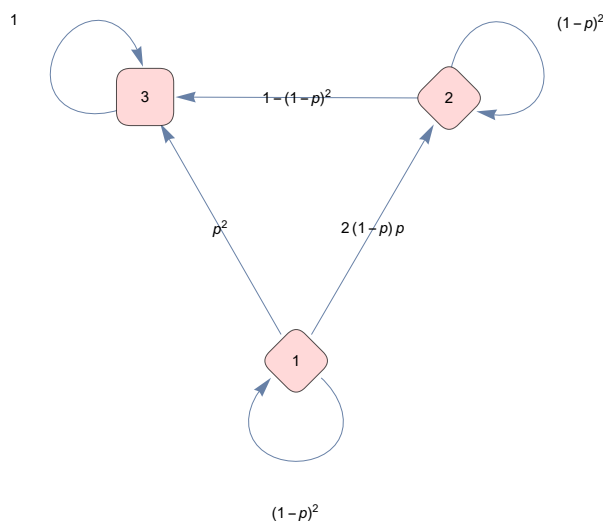
Na Obrázku 3.2 je pro danou populaci znázorněn diskretní Markovský řetězec popisující pravděpodobnosti přechodu mezi jednotlivými stavy pro pravidlo A (závisí na počtu nakažených sousedů). Stavy daného řetězce jsou následující:

- Stav 1 (nakažen jeden člen populace): přechodný a výchozí stav,
- Stav 2 (nakaženi dva členi populace): přechodný stav,
- Stav 3 (nakaženi tři členi populace): absorpční stav.

Na Obrázku 3.2 vidíme, že šipky vedou vždy od stavu s nižším číslem ke stejnému nebo vyššímu. Tato skutečnost souvisí s předpokladem, že jedince nelze vyléčit. Pokud tedy začínáme s jedním nakaženým, v dalším kroku se nemusí nakazit žádný jedinec (zůstáváme ve stavu 1) nebo se nakazí jeden jedinec (přecházíme do stavu 2) nebo se nakazí oba zbývajcí jedinci (přecházíme do stavu 3) a řetězec je absorbován.

V případě pravidla B je matice pravděpodobností přechodu $P^B = \{p_{ij}^B, i, j \in S\}$ matice $n \times n$ s prvky na pozici p_{ij}^B ve tvaru

$$p_{ij}^B = \binom{n-i}{j-i} p^{j-i} (1-p)^{n-j}, \quad i \leq j. \quad (3.3)$$



Obrázek 3.2: Markovský řetězec pro K_3 , pravidlo A a dané pravděpodobnosti přechodu shrnuté v matici $P_{K_3}^A$ v rovnici (3.2).

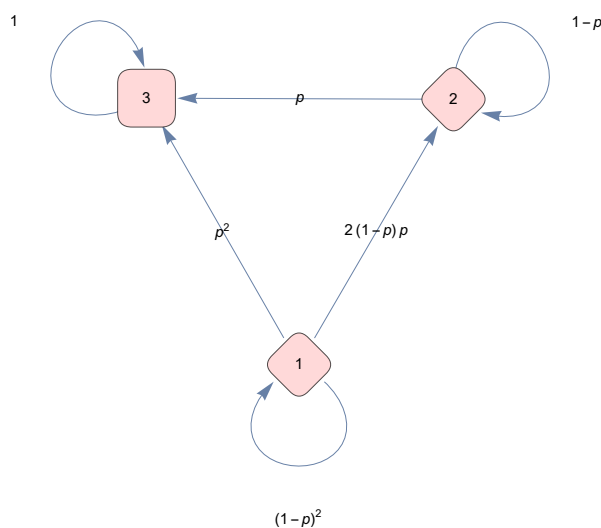
Stejně jako v případě pravidla A se jedná o horní trojúhelníkovou matici, protože neuvažujeme možnost vyléčení. Oproti pravidlu A nezvyšuje pravděpodobnost adopce počet nakažených sousedů i . Výraz p^{j-i} označuje pravděpodobnost nákazy $j-i$ jedinců, kteří se při daném přechodu ze stavu i do stavu j nakazí, výraz $(1-p)^{n-j}$ označuje pravděpodobnost toho, že se zbývajících $n-j$ jedinců při přechodu z i do j nenakazí, a kombinační číslo $\binom{n-i}{j-i}$ dává počet všech kombinací, jak je možné při daném přechodu nakazit $j-i$ jedinců z $n-i$ zdravých.

Uvažujme opět populaci o velikosti $n=3$ znázorněnou na Obrázku 3.1a. Matice pravděpodobností přechodu pro pravidlo adopce B je matice 3×3 ve tvaru

$$P_{K_3}^B = \begin{pmatrix} (1-p)^2 & 2p(1-p) & p^2 \\ 0 & 1-p & p \\ 0 & 0 & 1 \end{pmatrix}. \quad (3.4)$$

Na Obrázku 3.3 je pro danou populaci znázorněn diskretní Markovský řetězec popisující pravděpodobnosti přechodu mezi jednotlivými stavy pro pravidlo B (nezávisí na počtu nakažených sousedů). Stavy řetězce jsou totožné jako v případě pravidla A.

Jak je patrné při srovnání rovnice (3.2), popisující matici pravděpodobností přechodu pravidla A, a rovnice (3.4), popisující matici pravděpodobností přechodu pro pravidlo B, jediný řádek, který se pro $n=3$ liší, je řádek druhý. V případě, kdy jsme ve stavu $i=1$ (popsán prvním řádkem), jsou pravděpodobnosti přechodů do stavů $j=1$, $j=2$ a $j=3$ stejné pro obě pravidla, protože každý zdravý vrchol má pouze jednoho nakaženého souseda. Pro stav $i=2$ jsou odpovídající řádky obou matic odlišné, protože pro pravidlo A je vyšší pravděpodobnost, že se jediný zdravý vrchol nakazí (a analogicky nižší, že se nenakazí). Pro stav $i=3$ jsou odpovídající si řádky opět shodné, neboť nakaženi jsou již všichni členi populace a zůstáváme tedy ve stavu $i=3$ s pravděpodobností 1.



Obrázek 3.3: Markovský řetězec pro K_3 , pravidlo B a dané pravděpodobnosti přechodu shrnuté v matici $P_{K_3}^B$ v rovnici (3.4).

Čas absorpce Markovského řetězce na úplném grafu

Zajímavým atributem šíření inovace napříč jednotlivými populacemi je průměrný čas do absorpce. V případě uvažovaných Markovských řetězců, znázorněných na Obrázku 3.2 pro pravidlo A a na Obrázku 3.3 pro pravidlo B, lze rozložit množinu stavů S na množinu stavů přechodných (nakaženo méně jedinců než n) a jednoprvkovou množinu stavů trvalých, tj. tento stav je absorpční (nakaženo n jedinců). Matici pravděpodobností přechodu lze pak uspořádat do tzv. kanonického tvaru, [28],

$$P = \begin{pmatrix} P^* & 0 \\ Q & R \end{pmatrix}, \quad (3.5)$$

kde $P^* = \{p_{ij}, i, j \in T^c\}$, $Q = \{p_{ij}, i \in T, j \in T^c\}$, $R = \{p_{ij}, i, j \in T\}$, T je množina přechodných stavů a T^c je množina obsahující trvalý (absorpční) stav.

Matice P^* odpovídá přechodům mezi absorpčními stavy i a j . V našem případě je množina absorpčních stavů vždy jednoprvková. Pro pravidlo A s maticí pravděpodobností přechodu (3.2) je $P^{*A} = p_{33}^A = 1$. Všimněme si, že oproti (3.2) dochází k přeuspořádání stavů a absorpční stav budeme psát do prvního řádku dle zvyků z teorie Markovských procesů, [28].

Matice Q odpovídá přechodu z přechodného stavu i do trvalého (zde absorpčního) stavu j . V případě matice (3.2) je

$$Q^A = \begin{pmatrix} p^2 \\ 1 - (1-p)^2 \end{pmatrix}. \quad (3.6)$$

Matice R je část matice přechodu v kanonickém tvaru s přechodnými stavy, v matici (3.2) je tvaru

$$R^A = \begin{pmatrix} (1-p)^2 & 2p(1-p) \\ 0 & (1-p)^2 \end{pmatrix}. \quad (3.7)$$

Vzhledem k diskrétní povaze jednotlivých stavů je časem do absorpce myšlen celkový počet časových okamžiků strávených v množině přechodných stavů T předtím, než je řetězec absorbován. K analytickému určení průměrného času do absorpce použijeme tzv. fundamentální matici ve tvaru

$$F = (I - R)^{-1}, \quad (3.8)$$

kde I je jednotková matice rozměru odpovídajícímu rozměru R . Vektor

$$\bar{t} = F \cdot c, \quad (3.9)$$

kde c je sloupcový vektor jedniček odpovídající délky, potom na pozici \bar{t}_i obsahuje průměrný čas absorpce, pokud řetězec začal ve stavu i .

Pro populaci $n = 3$ s maticí pravděpodobností přechodu (3.2) je F ve tvaru

$$F^A = \begin{pmatrix} \frac{1-(1-p)^2}{4p^2-p^3+p^4} & \frac{2(1-p)p}{4p^2-p^3+p^4} \\ 0 & \frac{1-(1-p)^2}{4p^2-p^3+p^4} \end{pmatrix} \quad (3.10)$$

a vektor \bar{t}^A je

$$\bar{t}^A = \left(\frac{4-3p}{(-2+p)^2p}, \frac{1}{2p-p^2} \right)^T. \quad (3.11)$$

Prvek $\bar{t}_1^A = \frac{4-3p}{(-2+p)^2p}$ je průměrný čas do absorpce, jestliže řetězec začal ve stavu $i = 1$ (jeden nakažený), $\bar{t}_2^A = \frac{1}{2p-p^2}$ je průměrný čas do absorpce, jestliže řetězec začal ve stavu $i = 2$ (dva nakažení). Pro $p = 0.1$ je \bar{t}^A rovno přibližně $\bar{t}^A = (10.25, 5.26)^T$.

V případě pravidla B a populaci o velikosti $n = 3$ s maticí pravděpodobností přechodu (3.4) je fundamentální matice F ve tvaru

$$F^B = \begin{pmatrix} \frac{1}{2p-p^2} & \frac{1}{p} + \frac{1}{p-2} \\ 0 & \frac{1}{p} \end{pmatrix} \quad (3.12)$$

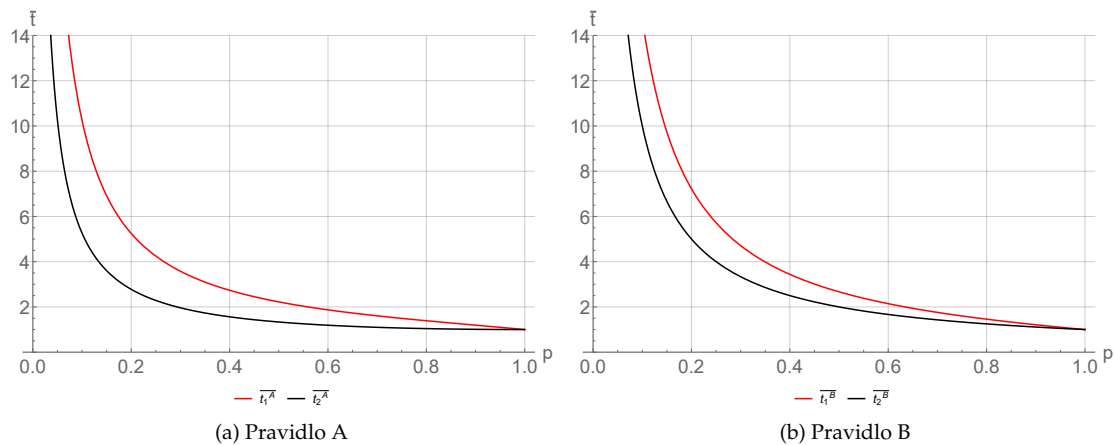
a vektor \bar{t}^B je

$$\bar{t}^B = \left(\frac{-3+2p}{(-2+p)p}, \frac{1}{p} \right)^T. \quad (3.13)$$

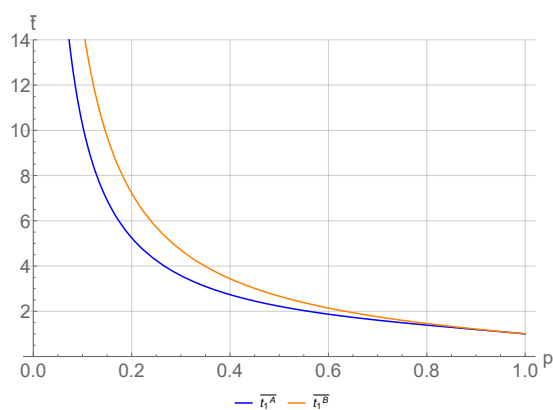
Pro $p = 0.1$ je \bar{t}^B rovno přibližně $\bar{t}^B = (14.74, 10)^T$. Jak pro start ze stavu $i = 1$ (jeden nakažený), tak pro start z $i = 2$ (dva nakažení) je odpovídající průměrný čas absorpce vyšší než v případě pravidla A, což souvisí s tím, že vyšší počet nakažených sousedů zvyšuje v případě pravidla A pro daný vrchol pravděpodobnost adopce.

Na Obrázku 3.4 je znázorněna závislost průměrného času absorpce \bar{t} na pravděpodobnosti absorpce p pro obě uvažovaná pravidla adopce.

Na Obrázku 3.5 je srovnání závislosti času absorpce \bar{t}_1 na pravděpodobnosti absorpce p při startu ze stavu $i = 1$ pro obě uvažovaná pravidla adopce. Z grafu je patrné, že průměrný čas absorpce při startu ze stavu $i = 1$ \bar{t}_1 je pro danou pravděpodobnost vyšší pro pravidlo B, protože pro něj platí, že zvyšující se počet nakažených nezvyšuje pro zdravé vrcholy pravděpodobnost adopce. V případě pravidla A zvyšující se podíl nakažených v populaci zvyšuje pravděpodobnost, že zbývající zdraví jedinci inovaci adoptují, a průměrný čas absorpce je nižší než v případě pravidla B.



Obrázek 3.4: Závislost průměrného času absorpce na pravděpodobnosti absorpce p pro $G = K_3$ a počáteční stavy řetězce $i = 1$ (červená křivka) a $i = 2$ (černá křivka).



Obrázek 3.5: Srovnání závislosti času absorpce \bar{t}_1 na pravděpodobnosti absorpce p při startu ze stavu $i = 1$ (jeden nakažený) pro pravidlo A (modrá křivka) a pro pravidlo B (oranžová křivka).

Tímto způsobem jsme schopni získat čas absorpce pro libovolnou velikost $n \in \mathbb{N}$ úplného grafu K_n . Označme si $a_n(p)$ průměrný čas absorpce v populaci popsané úplným grafem s počáteční podmínkou danou jedním nakaženým jedincem,

$$\mu^{(0)} = (1, 0, \dots, 0)^T.$$

Pak můžeme opět rozložením na kanonický tvar (3.5) rozlišit přechodné a absorpční stavy a získat obecný tvar matice R_n^A řádu $(n-1) \times (n-1)$ reprezentující přechody mezi jednotlivými přechodnými stavy $1, \dots, n-1$

$$R_n^A = \begin{pmatrix} p_{11}^A & p_{12}^A & \cdots & p_{1,n-1}^A \\ 0 & p_{22}^A & \cdots & p_{2,n-1}^A \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{n-1,n-1}^A \end{pmatrix},$$

kde prvky p_{ij}^A jsou dány vztahy (3.1). Následně můžeme vyjádřit čas absorpce analyticky jako první složku vektoru časů absorpce \bar{t} .

LEMMA 3.1. Průměrný čas absorpce Markovského řetězce na úplném grafu K_n , $n \in \mathbb{N}$, $n \geq 2$ řízeného pravidlem A s počáteční podmínkou $\mu^{(0)} = (1, 0, \dots, 0)^T$ je dán

$$a_n^A(p) = \bar{t}_1^A,$$

kde vektor \bar{t}^A je určen

$$\bar{t}^A = (I - R_n^A)^{-1}c.$$

Důkaz. Důkaz plyne z [15, Theorem 11.5]. □

Samozřejmě můžeme získat podobný výsledek pro Pravidlo B, kdy jednotlivé členy matice

$$R_n^B = \begin{pmatrix} p_{11}^B & p_{12}^B & \cdots & p_{1,n-1}^B \\ 0 & p_{22}^B & \cdots & p_{2,n-1}^B \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{n-1,n-1}^B \end{pmatrix},$$

LEMMA 3.2. Průměrný čas absorpce Markovského řetězce na úplném grafu K_n , $n \in \mathbb{N}$, $n \geq 2$ řízeného pravidlem B s počáteční podmínkou $\mu^{(0)} = (1, 0, \dots, 0)^T$ je dán

$$a_n^B(p) = \bar{t}_1^B, \tag{3.14}$$

kde vektor \bar{t}^B je určen

$$\bar{t}^B = (I - R_n^B)^{-1}c.$$

Poznámka 3.3. Uzavřený tvar časů absorpce $a_n^A(p)$ a $a_n^B(p)$ pro libovolné $n \in \mathbb{N}$ neumíme získat, ale díky postupu popsanému Lemmaty 3.1–3.2 můžeme vypočítat konkrétní hodnoty pro pevně zadané $n \in \mathbb{N}$. Prvních několik hodnot je vypočteno v Tabulce 3.1.

Všimněme si jednak toho, že pro $n = 2$ platí $a_2^A(p) = a_2^B(p)$, což plyne jednoduše z faktu, že neadoptující jedinec může mít pouze jednoho nakaženého souseda. Podobně hodnoty $a_3^A(p)$ a $a_3^B(p)$ jsme přímo vypočetli v (3.11) a (3.13).

n	$a_n^A(p)$	$a_n^B(p)$
2	$\frac{1}{p}$	$\frac{1}{p}$
3	$\frac{4-3p}{(-2+p)^2 p}$	$\frac{-3+2p}{(-2+p)p}$
4	$\frac{4p^5-31p^4+90p^3-130p^2+99p-33}{(p-2)p(p^2-3p+3)^2(p^2-2p+2)}$	$\frac{p(3(p-4)p+19)-11}{(p-2)p((p-3)p+3)}$
\vdots	\vdots	\vdots

Tabulka 3.1: Průměrné časy absorpce Markovských řetězců na úplných grafech K_n určených pravidly A a B, viz Lemmata 3.1–3.2.

Srovnání analytického a simulačního přístupu určení času absorpce pro úplný graf

Pro strukturovanou populaci popsanou úplným grafem na n vrcholech je analytické určení průměrného času absorpce (popsáno výše Lemmaty 3.1–3.2) porovnáno s průměrným časem do absorpce získaného za pomoci simulací. V obou případech uvažujeme počáteční podmínku $i = 1$, tj. v populaci se na počátku nachází jeden nakažený jedinec.

Hlavní motivací pro tento odstavec je navázání na následující odstavce, kde již není vždy možno pro obecný graf (neúplný a s jemnější strukturou) získat obecná tvrzení typu Lemmat 3.1–3.2 a simulační přístup se stane důležitějším článkem naší analýzy.

Za účelem srovnání analytických a simulačních výsledků bylo pro obě pravidla adopce pro jednotlivá n od 5 do 50 provedeno 10000 simulací. Srovnání analytického řešení a odpovídajícího výsledku simulací je na Obrázku 3.6.

V případě pravidla A se zvyšujícím se n klesá průměrný čas do absorpce, zatímco u pravidla B je tomu naopak.

Poměr nakažených v populaci v závislosti na čase

Dalším zkoumaným ukazatelem pro různé struktury populace je poměr průměrného počtu nakažených i v populaci o velikosti n v závislosti na čase.

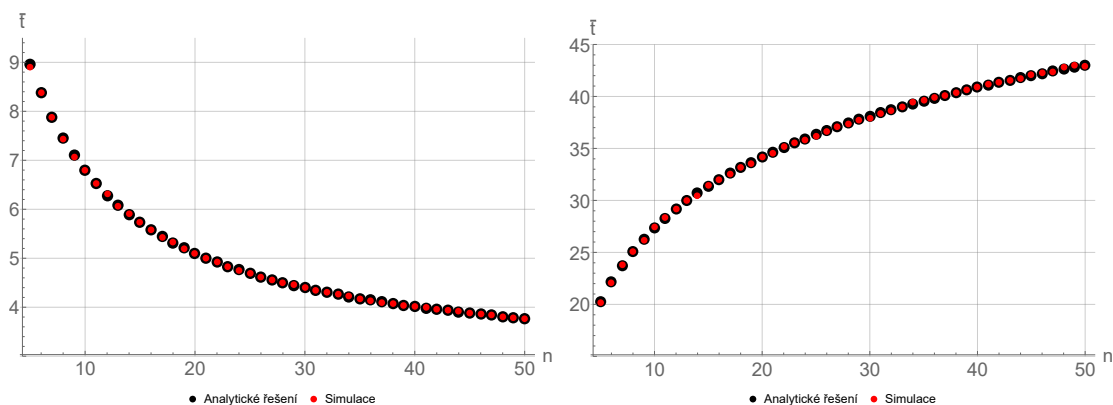
Průměrný počet nakažených v daném čase je určen simulačně. Pro každou z uvažovaných pravděpodobností adopce $p = 0.1, p = 0.2, \dots, p = 1$ je provedeno 1000 simulací a na Obrázku 3.7 je pak vykreslen průměrný poměr nakažených pro uvažovaná pravidla adopce.

Získané křivky popisují kumulativní adopci, stejně jako tomu bylo u normalizovaného Bassova modelu představeného v kapitole 2 a funkce kumulativní adopce $F(t)$ popsané rovnicí (2.7).

3.3 Diskrétní Markovský řetězec na k -regulárním grafu

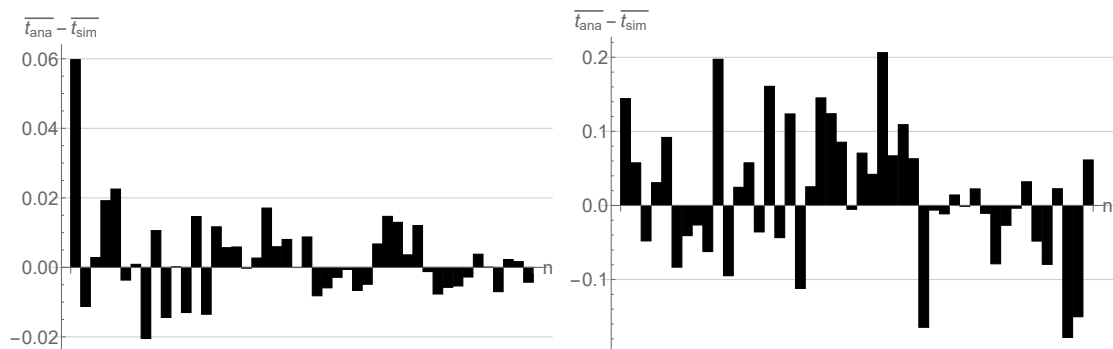
Jako další model strukturované populace uvažujme k -regulární graf. Graf je k -regulární, jestliže stupeň každého vrcholu $d(v)$ je roven číslu k pro všechna $v \in V$, [10]. Dle této definice je i výše uvedený úplný graf K_n regulární s $k = n - 1$ (každý vrchol má stupeň $n - 1$).

V případě libovolného k je však nutné ověřit podmínky existence daného grafu, neboť, na rozdíl od úplného grafu na n vrcholech, k -regulární graf na n vrcholech nebude pro libovolnou kombinaci n, k existovat vždy. Například, pro lichý počet vrcholů n nebude existovat k -regulární graf, kde k je liché.



(a) Průměrný čas absorpce při startu ze stavu $i = 1$ (jeden nakažený) pro různá n , úplný graf, pravidlo A, $p = 0.1$, 10000 simulací.

(b) Průměrný čas absorpce při startu ze stavu $i = 1$ (jeden nakažený) pro různá n , úplný graf, pravidlo B, $p = 0.1$, 10000 simulací.



(c) Rozdíl analytického řešení a simulací pro různá n , úplný graf, pravidlo A, $p = 0.1$, 10000 simulací.

(d) Rozdíl analytického řešení a simulací pro různá n , úplný graf, pravidlo B, $p = 0.1$, 10000 simulací.

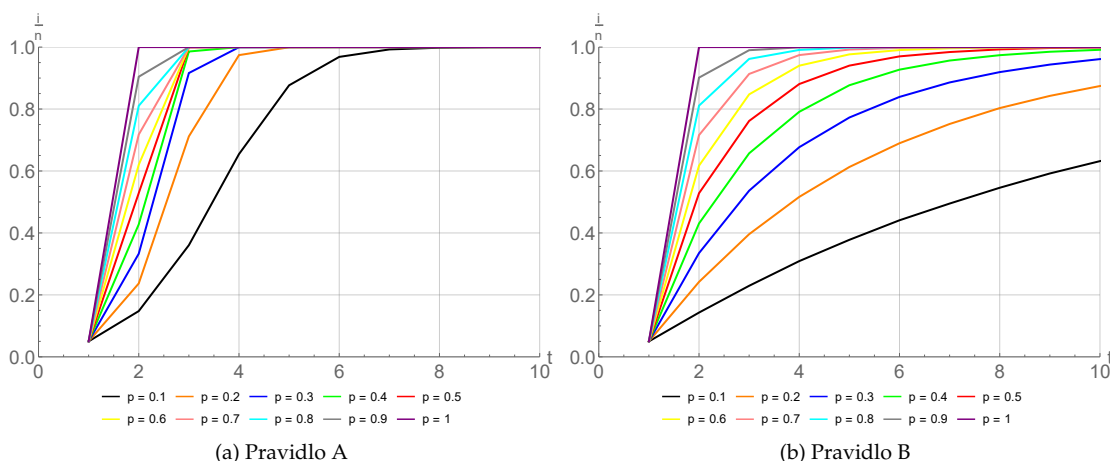
Obrázek 3.6: Srovnání analytického a simulačního přístupu určení času do absorpce při startu ze stavu $i = 1$ (jeden nakažený).

Pro ověření existence k -regulárního grafu je použita *Erdős-Gallaiova* věta, [10]:

VĚTA 3.4. Nutnou a postačující podmínkou toho, aby posloupnost $d = (d_1, d_2, \dots, d_n)$, kde $d_1 \geq d_2 \geq d_n$, byla grafová, je sudý součet stupňů $\sum_{i=1}^n d_i$ a splnění nerovnosti

$$\sum_{i=1}^m d_i \leq m(m-1) + \sum_{i=m+1}^n \min\{m, d_i\} \quad \text{pro } 1 \leq m \leq n. \quad (3.15)$$

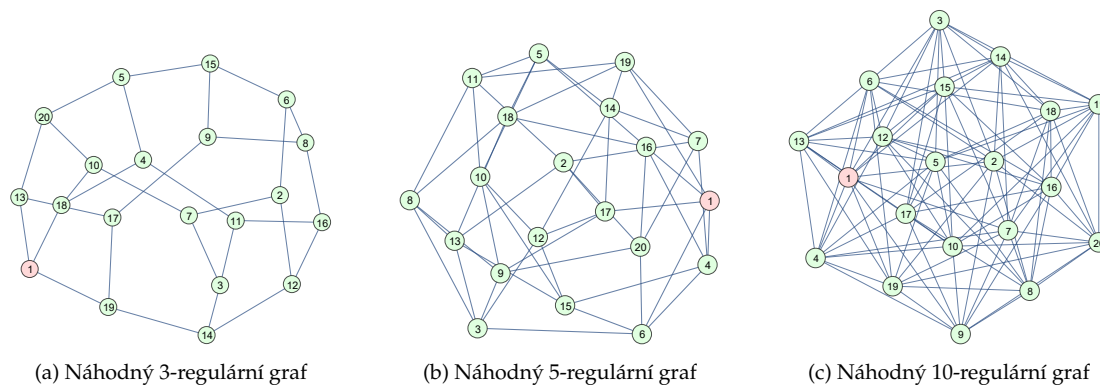
Pro daný k -regulární graf obecně neplatí, na rozdíl od úplných grafů, že by procesy s různými nakaženými jedinci byly vzájemně ekvivalentní. Navíc, pro danou dvojici n, k může existovat mnoho vzájemně nepřevoditelných (neizomorfních) k -regulárních grafů na n vrcholech, čemuž opět budou odpovídat rozdílné a vzájemně nepřevoditelné Markovské řetězce. Proto se omezíme na simulační přístup.



Obrázek 3.7: Poměr průměrného počtu nakažených i v populaci velikosti $n = 20$ v závislosti na čase t pro různé pravděpodobnosti adopce p , úplný graf K_{20} , 1000 simulací.

Čas absorpce

Pro obě pravidla adopce a pravděpodobnost adopce $p = 0.1$ byl simulačně spočten průměrný čas absorpce. Pro každé k od 2 do 10 bylo vygenerováno 1000 grafů pro $n = 20$ a $n = 50$. Příklady takto vygenerovaných grafů na $n = 20$ vrcholech pro $k = 3$, $k = 5$ a $k = 10$ jsou na Obrázku 3.8. Výsledné průměrné časy absorpce jsou uvedeny v Tabulce 3.2.



Obrázek 3.8: Modely populace pro $n = 20$, náhodně vygenerovaný k -regulární graf G s označenou počáteční podmínkou jednoho nakaženého v populaci (červený vrchol).

Na Obrázku 3.9 jsou vizualizovány průměrné časy absorpce uvedené v Tabulce 3.2. Z Tabulky 3.2 je patrné, že se zvyšujícím se k se zvětšuje rozdíl mezi odpovídajícími průměrnými časy absorpce pravidla A a pravidla B. V případě $k = 2$ je struktura populace popsána pomocí jediného souvislého 2-regulárního grafu na n vrcholech, kterým je cyklický graf (kružnice) C_n . Inovace se v populaci šíří „postupně“, tj. vrchol, který není spojen s jiným nakaženým vrcholem, se nemůže nakazit. V případě kružnice C_n se tedy inovace bude pro obě pravidla adopce šířit

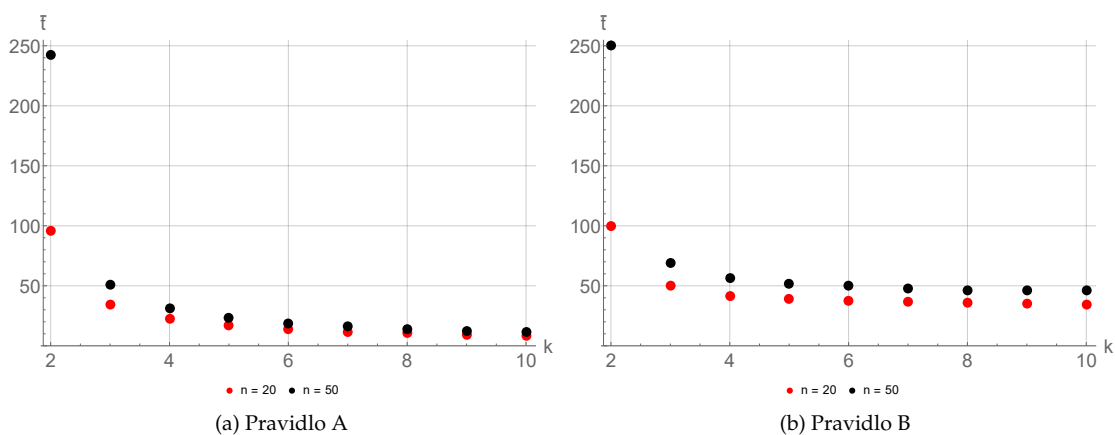
$n \backslash k$	2	3	4	5	6	7	8	9	10
20	95.673	34.716	22.560	16.958	13.978	11.894	10.414	9.490	8.521
50	242.298	50.725	31.226	23.469	18.960	16.202	14.090	12.632	11.444

(a) Pravidlo A

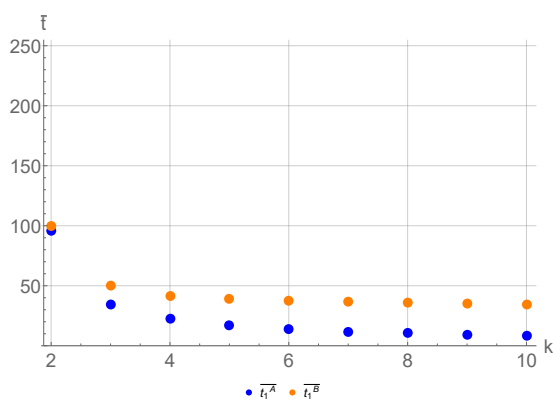
$n \backslash k$	2	3	4	5	6	7	8	9	10
20	100.126	50.026	41.582	39.090	37.665	36.843	35.550	34.964	34.651
50	250.351	69.307	56.796	51.623	49.882	47.743	46.557	46.556	46.125

(b) Pravidlo B

Tabulka 3.2: Průměrný čas absorpce v závislosti na k pro $n = 20$ a $n = 50$, $p = 0.1$, 1000 simulací.



Obrázek 3.9: Průměrný čas absorpce v závislosti na k pro $n = 20$ a $n = 50$, $p = 0.1$, 1000 simulací.



Obrázek 3.10: Srovnání závislosti času absorpce \bar{t}_1 na k pro $n = 20$ při startu ze stavu $i = 1$ (jeden nakažený) pro pravidlo A (modré body) a pro pravidlo B (oranžové body), 1000 simulací.

stejně až do stavu $i = n - 1$, protože každý zdravý vrchol má nejvýše jednoho nakaženého souseda. Ve stavu $i = n - 1$ je pravděpodobnost, že zbývající vrchol inovaci adoptuje, vyšší pro pravidlo A. Zdravý vrchol má právě dva nakažené sousedy a v případě pravidla A se s každým nakaženým sousedem zvyšuje pravděpodobnost adopce.

Se zvyšujícím se k klesá průměrný čas absorpce pro pravidlo A výrazněji, než pro pravidlo B, což je patrné z Obrázku 3.10, na kterém jsou pro $n = 20$ vykresleny časy absorpce pro různá k .

Poměr nakažených v populaci v závislosti na čase

Podívejme se nyní na to, jak se liší poměr průměrného počtu nakažených jedinců i v populaci velikosti n v závislosti na čase pro různé hodnoty k regulárního grafu.

Na Obrázku 3.11 je vykreslen poměr $\frac{i}{n}$ v závislosti na čase pro $n = 20$, $p = 0.1$ a $k = 2$, $k = 3$ až $k = 10$ pro obě pravidla adopce.

Pro pravidlo A na Obrázku 3.11a je celá populace nakažena (tj. poměr $\frac{i}{n}$ je roven jedné) nejrychleji pro $k = 10$. Se snižujícím se k se zpomaluje také rychlost šíření inovace.

Stejně je tomu pro pravidlo B na Obrázku 3.11b, nicméně v případě pravidla B je pro odpovídající si k šíření pomalejší než v případě pravidla A. Rozdíl mezi rychlostí šíření inovace pro odpovídající si k se prohlubuje se zvyšujícím se k . Křivky pro obě pravidla adopce vypadají na první pohled odlišně (např. pro $k = 10$ má v případě pravidla A šíření inovace strmější růst, než v případě pravidla B). Oproti tomu pro $k = 2$ vypadají křivky v případě obou pravidel adopce téměř totožně.

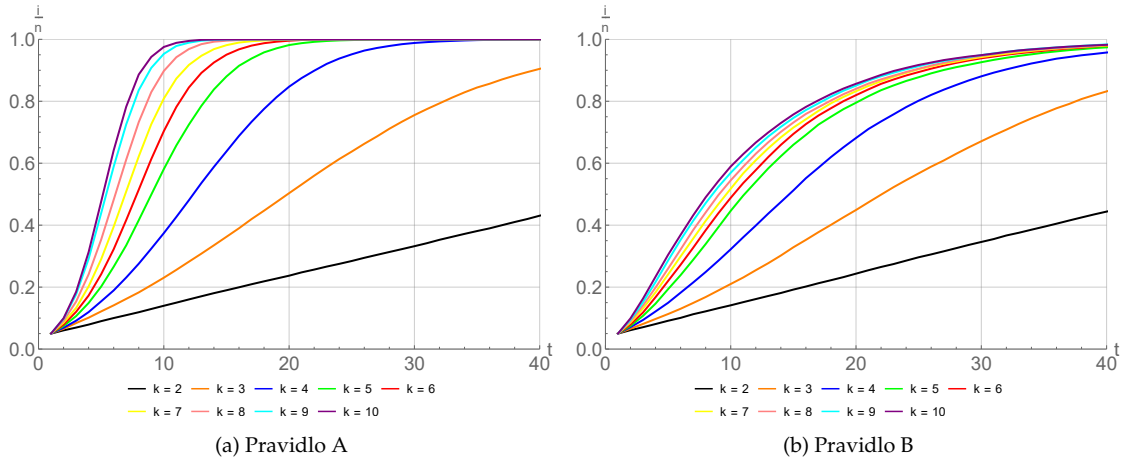
Pro regulární graf s $k = 2$ je grafem popisujícím populaci kružnice na n vrcholech C_n . Protože na počátku šíření uvažujeme v populaci jednoho nakaženého, každý zdravý vrchol má nejvýše jednoho nakaženého souseda a pravděpodobnost adopce je tedy stejná pro pravidlo A i pro pravidlo B. Jedinou výjimkou je situace, kdy v populaci zbývá jediný zdravý vrchol, který má dva nakažené sousedy. V tomto případě se bude pravděpodobnost adopce pro daný vrchol lišit s ohledem na zvolené pravidlo adopce. Pokud je uvažováno pravidlo A, pro vrchol bude vyšší pravděpodobnost, že inovaci adoptuje, protože má dva nakažené sousedy a každý ho může nakazit s pravděpodobností p . U pravidla B bude pravděpodobnost, že poslední vrchol inovaci adoptuje, stejná jako v případě všech předchozích vrcholů s nakaženým sousedem.

Závislost poměru nakažených $\frac{i}{n}$ na čase pro různé pravděpodobnosti adopce p je pro 4-regulární graf vizualizována na Obrázku 3.12. Pro obě pravidla adopce platí, že se zvyšující se pravděpodobností adopce p má poměr nakažených v závislosti na čase strmější růst.

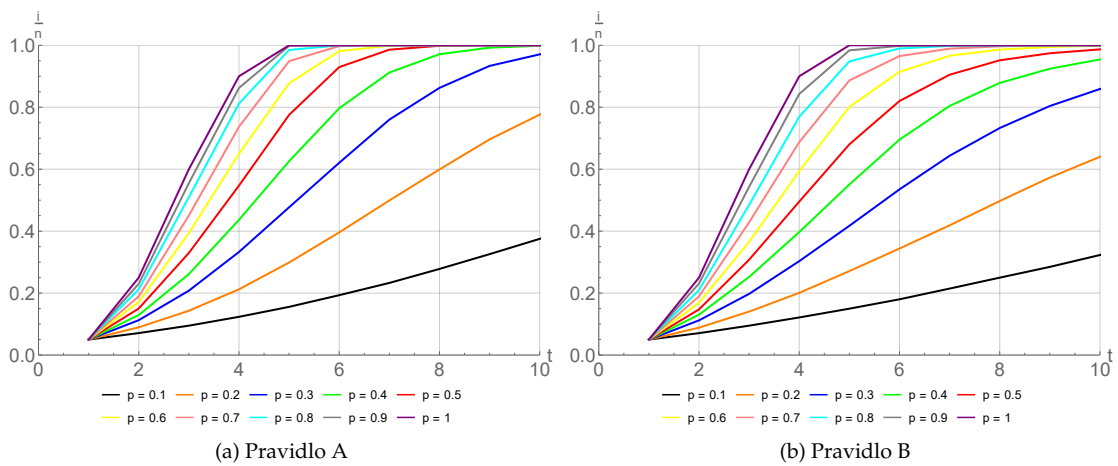
3.4 Diskrétní Markovský řetězec na cestě

V případě struktury populace popsané cestou na n vrcholech se jedná o graf P_n , ve kterém, kromě krajních vrcholů, má každý vrchol dva sousedy. S ohledem na uvažovanou počáteční podmínku jednoho nakaženého v populaci bude pro šíření inovace na cestě P_n záležet na umístění nakaženého vrcholu. V případě, že první nakažený vrchol bude vždy jeden ze dvou krajních vrcholů, nebude záležet na zvoleném pravidlu adopce, protože nakažený jedinec může v daném kroku nakazit nejvýše jednoho dalšího jedince, resp. nenakažený jedinec může být v každém kroku nakažen nejvýše jedním sousedem.

Předpokládejme pro jednoduchost extrémní případ, kdy první nakažený jedinec je vždy umístěn do počátečního (krajního) vrcholu, obě pravidla adopce by tedy měla poskytovat obdobné výsledky šíření inovace v dané populaci.



Obrázek 3.11: Poměr průměrného počtu nakažených $\frac{i}{n}$ v populaci velikosti $n = 20$ v závislosti na čase t pro různé k -regulární grafy, $p = 0.1$, 1000 simulací.



Obrázek 3.12: Poměr průměrného počtu nakažených i v populaci velikosti $n = 20$ v závislosti na čase t pro různé pravděpodobnosti adopce p , 4-regulární graf, 1000 simulací.

V tomto jednoduchém případě jsme schopni dokázat analyticky tvrzení o úplném nakažení populace.

VĚTA 3.5. *Uvažujeme-li cestu P_n s počátečním stavem jediného nakaženého jedince na okraji a $p \in (0, 1]$, pak k úplnému nakažení populace dojde v průměru po $\frac{n-1}{p}$ krocích jak v případě pravidla A, tak v případě pravidla B.*

Důkaz. Matice přechodu je dvoudiagonální matice ve tvaru

$$P = \begin{pmatrix} 1-p & p & 0 & \dots & 0 & 0 \\ 0 & 1-p & p & \dots & 0 & 0 \\ 0 & 0 & 1-p & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1-p & p \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix},$$

a tedy

$$R = \begin{pmatrix} 1-p & p & \dots & 0 & 0 \\ 0 & 1-p & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1-p & p \\ 0 & 0 & \dots & 0 & 1-p \end{pmatrix}.$$

Fundamentální matice je následně horní trojúhelníková matice ve tvaru

$$F = \begin{pmatrix} \frac{1}{p} & \frac{1}{p} & \dots & \frac{1}{p} & \frac{1}{p} \\ 0 & \frac{1}{p} & \dots & \frac{1}{p} & \frac{1}{p} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{1}{p} & \frac{1}{p} \\ 0 & 0 & \dots & 0 & \frac{1}{p} \end{pmatrix}$$

A proto snadno ověříme, že v každém stavu (počtu nakažených) setrvá proces v průměru $\frac{1}{p}$ kroků. Následně

$$\bar{t} = F \cdot c = \left(\frac{n-1}{p}, \frac{n-2}{p}, \dots, \frac{2}{p}, \frac{1}{p} \right)^T,$$

z čehož plyne tvrzení Věty 3.5. □

Pro model šíření inovace, informace, atd., se jedná o model ve smyslu hry „tichá pošta“. Informace se šíří napříč danou skupinou velikosti n od prvního jedince a každý další jedinec předá informaci tomu následujícímu, který ji adoptuje s pravděpodobností p .

Cesta P_{20} se znázorněnou počáteční podmínkou jednoho nakaženého v populaci je na Obrázku 3.13a.

3.5 Diskrétní Markovský řetězec na hvězdě

Graf S_n označovaný jako „hvězda“ je graf na n vrcholech, kde právě jeden vrchol je spojen hranou se všemi ostatními vrcholy grafu, zatímco každý ze zbývajících vrcholů (listů) je spojen pouze s tímto jedním (centrálním) vrcholem.

Stejně jako v případě cesty P_n bude u hvězdy S_n záležet na umístění prvního nakaženého jedince. Protože proces s jedním nakaženým v centrálním vrcholu nebude nikdy obsahovat vrcholy s více než jedním nakaženým sousedem, Pravidla A a B budou odpovídat stejným Markovským řetězcům. Navíc existuje přímá vazba na Markovské procesy na úplném grafu s pravidlem B.

LEMMA 3.6. *Markovský řetězec s pravidlem A i B na hvězdě s počáteční podmínkou odpovídající nakaženému centrálnímu uzlu*

$$\mu^{(0)} = (1, 0, \dots, 0)^T,$$

je ekvivalentní s Markovským řetězcem na úplném grafu K_n a průměrný čas absorpce je dán $a_n^B(p)$ ze vztahu (3.14).

Markovský řetězec s pravidlem A i B na hvězdě s počáteční podmínkou odpovídající nakaženému libovolnému listu, např.

$$\mu^{(0)} = (0, 1, \dots, 0)^T,$$

má průměrný čas absorpce daný

$$\frac{1}{p} + a_{n-1}^B(p).$$

Důkaz. První tvrzení plyne z faktu, že u pravidla B na úplném grafu nezávisí na počtu sousedů a všechny vrcholy jsou od začátku sousedy centrálního vrcholu, který je již nakažen.

Druhé tvrzení vyplývá z toho, že centrální vrchol se nakazí v průměru za $1/p$ kroků a dále proces odpovídá, díky stejnému argumentu, Markovskému procesu na úplném grafu K_{n-1} s $n-1$ vrcholy. \square

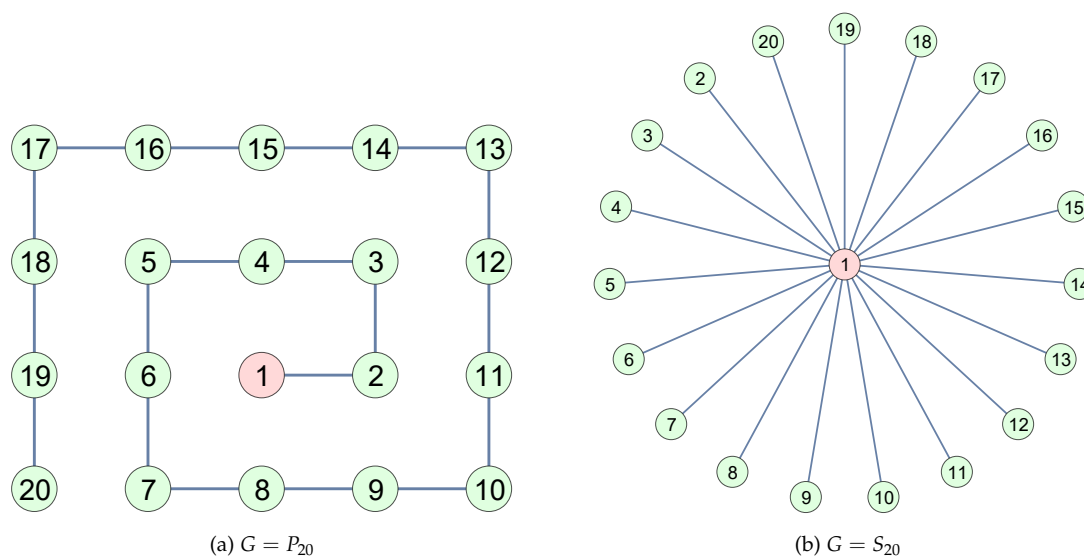
Pro model šíření inovace lze takovým grafem modelovat např. reklamu na sociální síti zaměřenou na určitou část populace. Centrální vrchol grafu S_n je daná reklama, která se opakovaně zobrazuje jedincům, kteří se vzájemně neznají a neovlivňují (ve smyslu adoptování dané informace/inovace), tj. zbývajícím vrcholům grafu S_n .

Hvězda S_{20} se znázorněnou počáteční podmínkou jednoho nakaženého v populaci je na Obrázku 3.13b.

3.6 Simulační srovnání

V předchozích kapitolách byly popsány jednotlivé uvažované modely struktury populace popsané grafem G na n vrcholech. Zabývali jsme se

- úplným grafem K_n ,
- k -regulárním grafem na n vrcholech,
- cestou P_n ,
- hvězdou S_n .



Obrázek 3.13: Modely populace pro $n = 20$, cesta P_{20} a hvězda S_{20} s naznačenou počáteční podmínkou jednoho nakaženého v populaci (červený vrchol).

V případě k -regulárního grafu nejsme, na rozdíl od ostatních možností, schopni získat analytické vyjádření, proto se pokusíme o simulační srovnání.

Předmětem dalšího zkoumání může být otázka, pro jakou strukturu se inovace nejrychleji rozšíří do celé populace a jaký vliv na to mají uvažovaná pravidla adopce.

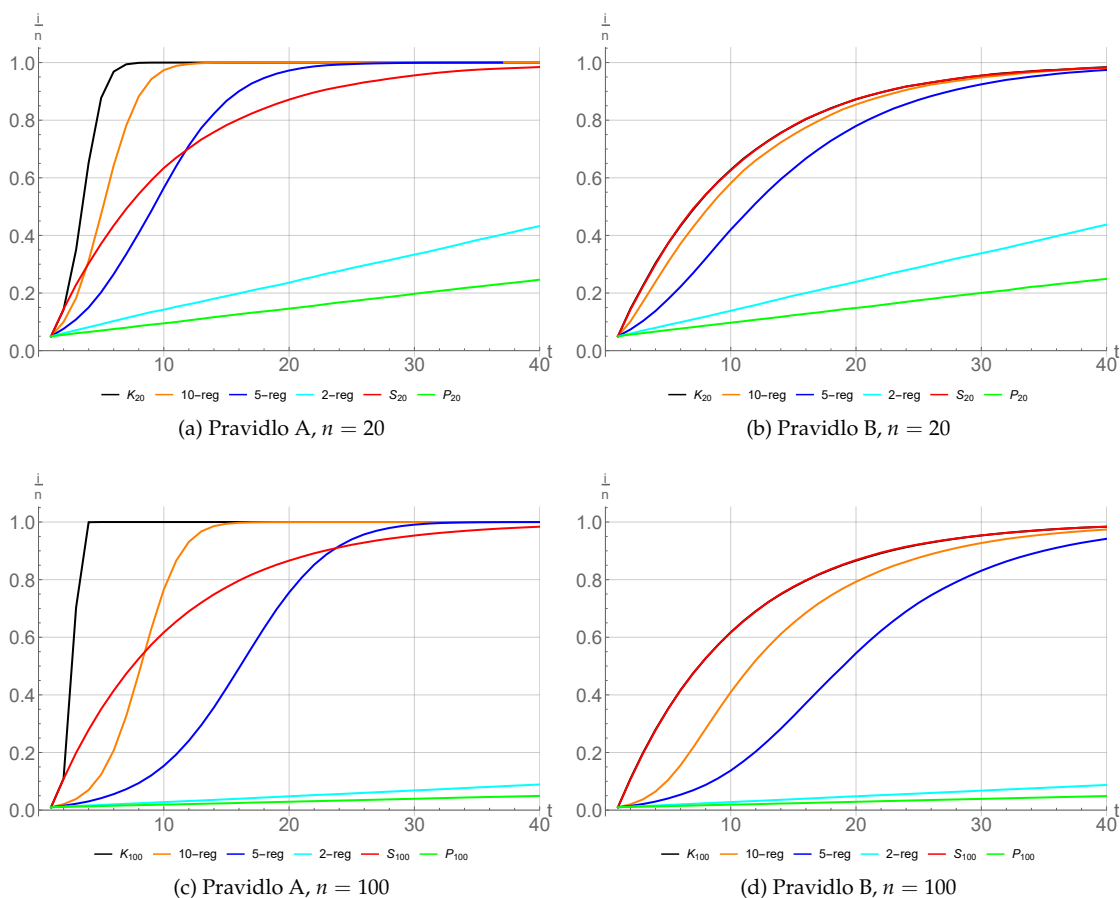
Za účelem srovnání vlivu struktury populace na šíření inovace bylo provedeno pro každý typ grafu 1000 simulací pro obě pravidla adopce. Pravděpodobnost adopce je v simulacích nastavena jako $p = 0.1$ a je uvažováno několik velikostí populace, konkrétně $n = 20, 30, \dots, 100$.

Pro populaci velikosti $n = 20$ a $n = 100$ je na Obrázku 3.14 vykreslen poměr nakažených v populaci v závislosti na čase pro obě pravidla adopce. V dané grafické vizualizaci si lze všimnout několika skutečností.

Začněme poznatky týkajícími se úplného grafu K_n . Pro pravidlo A (počet nakažených sousedů zvyšuje pravděpodobnost, že daný vrchol inovaci adoptuje) zobrazené na Obrázku 3.14a pro $n = 20$ a na Obrázku 3.14c pro $n = 100$ se pro poměr nakažených v populaci v závislosti na čase jedná o S-křivku.

V případě pravidla B (počet nakažených sousedů nezvyšuje pravděpodobnost, že daný vrchol inovaci adoptuje) na Obrázku 3.14b pro $n = 20$ a na Obrázku 3.14d pro $n = 100$ je vývoj počtu nakažených pro K_n téměř totožný s grafem pro hvězdu S_n , viz Lemma 3.6. V případě hvězdy S_n je na počátku nakažen centrální vrchol, ze kterého vede hrana do každého zdravého vrcholu. Každý nenakažený vrchol je tedy od počátku simulace spojen hranou s prvním nakaženým vrcholem, stejně je tomu u úplného grafu K_n . Protože u pravidla B zvyšující se počet nakažených sousedů nezvyšuje pravděpodobnost adopce daného vrcholu, poskytuje simulace pro K_n téměř totožné výsledky, jako v případě S_n , kde mezi vrcholy, které jsou na počátku simulace zdravé, hrany neexistují.

U 10-regulárního grafu na n vrcholech je tvar křivek pro obě pravidla podobného charakteru, jako v případě úplného grafu K_n . Šíření inovace je v průměru rychlejší pro úplný graf



Obrázek 3.14: Poměr průměrného počtu nakažených i v populaci velikosti n v závislosti na čase t pro různé struktury populace, $p = 0.1$, 1000 simulací.

K_n , než pro 10-regulární graf. V případě $n = 20$ vede u úplného grafu K_{20} z každého vrcholu $20 - 1 = 19$ hran, v případě $n = 100$ dokonce $100 - 1 = 99$ hran, zatímco u 10-regulárního grafu pro obě uvažované velikosti populace $n = 20$ a $n = 100$ vede z každého vrcholu 10 hran.

Podobně jako u 10-regulárního grafu je tomu u 5-regulárního grafu, kde je však průměrné šíření inovace pomalejší, než v případě úplného grafu K_n nebo 10-regulárního grafu, protože z každého vrcholu vede „pouze“ 5 hran a inovace se tedy napříč populací šíří pomaleji.

V případě hvězdy S_n poskytují pro dané n obě pravidla adopce téměř totožné výsledky. Protože na počátku je nakažený jedinec vždy umístěn do centrálního vrcholu, každý zdravý vrchol má na počátku právě jednoho nakaženého souseda, kterým je onen centrální vrchol.

Pro 2-regulární graf a cestu P_n je šíření inovace v dané populaci nejpomalejší ze všech uvažovaných struktur. Na rozdíl od 2-regulárního grafu, kde se inovace může šířit na obě strany od množiny nakažených vrcholů, pro cestu P_n se v daném časovém kroku může nakazit nejvýše jeden vrchol. Proto se pro cestu P_n šíří inovace nejpomaleji. U cesty P_n je první nakažený umístěn do počátečního (krajního) vrcholu, proto v případě cesty P_n obě pravidla adopce poskytují obdobné výsledky.

Podrobné srovnání průměrných časů absorpce pro různé grafové struktury G , velikosti populace n a obě pravidla adopce je v Tabulce 3.3.

$G \backslash n$	20	30	40	50	60	70	80	90	100
K_n	5.120	4.390	4.018	3.782	3.564	3.354	3.206	3.105	3.043
10-reg.	8.482	9.810	10.620	11.435	11.913	12.475	12.823	13.187	13.370
5-reg.	16.801	19.470	22.140	23.160	25.370	25.790	27.030	27.180	27.920
S_n	34.287	38.168	40.724	42.861	44.967	46.732	47.736	48.475	49.532
2-reg.	96.384	145.215	194.839	246.37	293.159	346.174	395.431	446.798	492.490
P_n	188.647	288.707	391.116	486.922	590.198	692.388	788.686	885.617	993.182

(a) Pravidlo A

$G \backslash n$	20	30	40	50	60	70	80	90	100
K_n	34.661	37.408	41.130	42.305	45.023	46.494	47.565	48.004	50.429
10-reg.	35.465	40.729	42.750	46.252	47.291	49.402	52.125	52.944	53.933
5-reg.	38.267	44.303	48.583	51.138	55.356	56.977	58.493	58.340	61.068
S_n	34.843	38.341	40.705	43.270	44.222	46.454	47.567	48.881	49.593
2-reg.	99.432	150.304	198.229	248.337	300.327	349.654	398.412	448.772	503.123
P_n	189.921	289.379	389.838	491.334	589.587	689.009	795.472	889.334	991.146

(b) Pravidlo B

Tabulka 3.3: Průměrný čas absorpce pro různé G a n , $p = 0.1$, 1000 simulací.

V případě pravidla A, kde vyšší počet nakažených sousedů zvyšuje pravděpodobnost adopce, se v provedených simulacích inovace nejrychleji rozšířila v populaci popsané úplným grafem K_n . Úplný graf modeluje populaci, ve které se každý jedinec „zná“ se všemi zbývajícími členy populace a může je v adopci dané inovace ovlivnit. V rychlosti šíření inovace pro pravidlo A po úplném grafu K_n následují 10-regulární a 5-regulární grafy na n vrcholech, dále hvězda S_n , 2-regulární graf a nejpomaleji se inovace šíří v případě cesty P_n .

U pravidla B, kde vyšší počet nakažených sousedů nezvyšuje pravděpodobnost adopce, je, stejně jako u pravidla A, úplný graf K_n jedním z grafů, pro které se inovace šíří nejrychleji. Nicméně obdobné průměrné časy adopce, jako má graf K_n , poskytuje v daných simulacích hvězda S_n . To je způsobeno skutečností, že každý zdravý vrchol je od začátku simulace spojen hranou s nakaženým vrcholem. Po úplném grafu K_n a hvězdě S_n v rychlosti šíření inovace následují 10-regulární, 5-regulární a 2-regulární grafy a strukturou, pro kterou se inovace v populaci rozšíří nejpomaleji, je, stejně jako v případě pravidla A, cesta P_n .

3.7 Konvergence k Bassovu modelu

Diskrétní Markovské řetězce, které jsme studovali v této kapitole, mají přes svůj odlišný charakter (diskrétní čas, diskrétní a konečná populace) úzký vztah k Bassovu modelu, resp. ke spojitým modelům popsaných pomocí diferenciálních rovnic. Tento vztah se nyní pokusíme nastínit.

Budeme uvažovat nejjednodušší konfiguraci, úplný graf K_n , a bude nás zajímat chování námi představených diskrétních Markovských řetězců pro n zvyšující se do nekonečna.

Vezměme v potaz nejprve pravidlo B a podívejme se na diferenční rovnici, která zkoumá vývoj průměrného počtu nakažených $x(t)$.

LEMMA 3.7. *Průměrná trajektorie diskrétního Markovského řetězce na K_n , který se řídí pravidlem B, splňuje diferenční rovnici*

$$x(t+1) = x(t) + p(1 - x(t)).$$

Důkaz. V čase t má populace $1 - x(t)$ dosud nenakažených jedinců, pravděpodobnost nakažení každého z nich je p . Počet nakažených se tedy řídí binomickým rozdělením $Bi(1 - x(t), p)$ a průměrný počet nově nakažených je $p(1 - x(t))$. Proto je počet nakažených v čase $t + 1$ roven počtu nakažených v čase t a počtu nově nakažených, tedy

$$x(t+1) = x(t) + p(1 - x(t)).$$

□

Za účelem odvození spojitého modelu, nyní uvažujme jemnější diskretizaci (diskrétní časovou škálu) s krokem h

$$\mathbb{T}_h = \{0, h, 2h, \dots\} \quad (3.16)$$

a upravme příslušně pravděpodobnost nakažení na

$$\tilde{p} = hp. \quad (3.17)$$

Diskrétní Markovský řetězec na \mathbb{T}_h splňuje

$$\begin{aligned} x(t+h) &= x(t) + hp(1 - x(t)), \\ \frac{x(t+h) - x(t)}{h} &= p(1 - x(t)). \end{aligned} \quad (3.18)$$

VĚTA 3.8. *Diskrétní Markovský řetězec řídící se pravidlem B s jemnějším časovým krokem (3.16) a modifikovanou pravděpodobností (3.17) konverguje ke spojitému modelu*

$$x'(t) = p(1 - x(t)). \quad (3.19)$$

Důkaz. Tvrzení plyne přímo z limitního přechodu $h \rightarrow 0$ v diferenční rovnici (3.18). □

Přejdeme nyní k pravidlu A (pravděpodobnost se zvětšuje s počtem nakažených), kde bude vztah komplexnější, zejména kvůli tomu, že obdoba Lemmatu 3.7 je složitější.

LEMMA 3.9. *Je-li $x(t)$ podíl nakažených jedinců v čase t diskrétního Markovského řetězce na K_n , který se řídí pravidlem A, pak $x(t+1)$ splňuje*

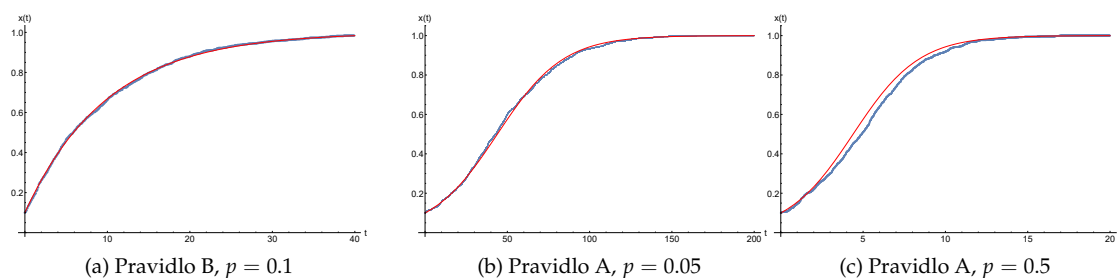
$$x(t+1) = x(t) + px(t)(1 - x(t)) + O(p^2). \quad (3.20)$$

Důkaz. Pravděpodobnost nakažení jednoho nenakaženého jedince v závislosti na podílu nakažených jedinců $x \in (0, 1)$ je dána

$$P(x) = 1 - (1 - p)^x.$$

Pro dostatečně malé p můžeme použít aproximaci Taylorovým rozvojem

$$P(x) = 1 - (1 - p)^x = px - \frac{1}{2}p^2((x-1)x) + \frac{1}{6}p^3(x-2)(x-1)x + O(p^4).$$



Obrázek 3.15: Ilustrace konvergence diskrétního Markovského řetězce ke spojitému Bassovu modelu. Červené křivky zobrazují spojitě modely (3.19) a (3.22), modré body pak řešení modifikovaných diferencí rovnic (3.18) a (3.21), při uvažovaném grafu $G = K_{10}$ a diskretizačním kroku $h = \frac{1}{10}$.

Protože podíl nakažených je $1 - x(t)$, průměrný počet nakažených v každém kroku získáme jako střední hodnotu binomického rozdělení, tedy

$$px(t)(1 - x(t)) + O(p^2),$$

z čehož plyne (3.20). □

Uvažujeme-li opět proces s jemnějším diskrétním časem \mathbb{T}_h daným (3.16), můžeme odvodit modifikovanou diferencí rovnici

$$\frac{x(t+h) - x(t)}{h} = px(t)(1 - x(t)) + O(hp^2), \quad t \in \mathbb{T}. \quad (3.21)$$

VĚTA 3.10. *Diskrétní Markovský řetězec říditý se pravidlem A s jemnějším časovým krokem (3.16) a modifikovanou pravděpodobností (3.17) konverguje ke spojitému modelu*

$$x'(t) = px(t)(1 - x(t)). \quad (3.22)$$

Důkaz. Tvrzení plyne z limitního přechodu $h \rightarrow 0$ v diferencí rovnici (3.21). □

Poznámka 3.11. Kvůli členům vyšších řádů v (3.21) je konvergence Markovského řetězce s pravidlem A pomalejší pro větší pravděpodobnosti p . Zároveň oproti pravidlu B pozorujeme větší odchylky od spojitě modelu pro větší hodnoty pravděpodobnosti nakažení p , viz Obr. 3.15.

Dále si všimněme faktu, že limitní rovnice (3.22) přímo odpovídá normalizovanému Bassovu modelu (2.5) bez inovací.

Přestože jsme se nyní zabývali konvergencí k Bassovu modelu, uvědomme si výhody, které, oproti spojitě Bassovu modelu, použití diskrétních Markovských řetězců poskytuje. V uvažovaných diskrétních modelech rozlišujeme každého jedince v populaci. Použitím grafových struktur k modelování populace můžeme každému vrcholu, představujícímu daného jedince, přiřadit jak unikátní hodnotu pravděpodobnosti adopce, tak konektivitu. Diskrétní Markovské řetězce tak umožňují modelování rozmanitější a komplexnější populace. V další kapitole se proto budeme zabývat těmito diskrétními řetězci a jejich použitím k modelování šíření inovace v nehomogenní populaci, jejíž struktura odpovídá empirickým studiím sociálních sítí.

Modely šíření inovace uvažující nehomogenní populaci 4

4.1 Nehomogenní populace

V předchozí kapitole byl pro model šíření inovace použit diskrétní Markovský řetězec. Jako model struktury dané populace byl použit graf G na n vrcholech a byla uvažována dvě různá pravidla adopce.

V případě pravidla A se se zvyšujícím se počtem nakažených sousedů zdravého vrcholu zvyšuje pravděpodobnost, že zdravý vrchol inovaci adoptuje, protože v daném kroku ho může nakazit každý z nakažených sousedů s pravděpodobností p .

Pro pravidlo B zvyšující se počet nakažených sousedů nezvyšuje pravděpodobnost, že zdravý vrchol inovaci adoptuje. Pokud se v množině jeho sousedů vyskytuje alespoň jeden nakažený, pak daný vrchol inovaci adoptuje s pravděpodobností p .

Jedním z nedostatků těchto (do jisté míry „naivních“) modelů je skutečnost, že je na všechny vrcholy v populaci v jistém smyslu nahlíženo stejně.

U pravidla A mají každé dva vrcholy, které mají stejný počet nakažených sousedů, stejnou pravděpodobnost adopce.

U pravidla B mají stejnou pravděpodobnost adopce všechny vrcholy, které mají alespoň jednoho souseda, který přijal danou inovaci.

To je značně zjednodušený pohled na ve skutečnosti mnohem komplexnější a složitější realitu. Uvažujeme-li dané modely ve smyslu šíření inovace, v reálném světě zřejmě nebude pro každého jedince v populaci platit stejná pravděpodobnost, že danou inovaci adoptuje.

Dalším v realitě těžko uplatnitelným předpokladem je popis populace za pomoci pravidelné a symetrické struktury. V předchozí kapitole byly pro jednoduchost uvažovány grafy, ve kterých na počátku nenakažené vrcholy byly stejného stupně (kromě cesty P_n , ve které je poslední nenakažený vrchol stupně 1 a ostatní na počátku nenakažené vrcholy jsou stupně 2).

Nehomogenní populaci lze v případě simulací modelovat několika způsoby. Prvním způsobem je popis populace pomocí grafu G s různými stupni vrcholů, tj. pomocí neregulárních grafů.

Druhým způsobem, jak zajistit nehomogenní populaci, je rozdílnost vrcholů ve smyslu jejich

pravděpodobnosti adopce. Místo jednoho fixního parametru p pro celou populaci velikosti n , jako to bylo uvažováno v předchozí kapitole, můžeme každému i -tému vrcholu, kde $i = 2, \dots, n$ (uvažujeme, že prvním nakaženým je vrchol na pozici 1), nastavit unikátní hodnotu $p_i \in [0, 1]$.

4.2 Small-world a šest stupňů separace

Small-world network (volně přeloženo jako „sít malého světa“) je fenomén, podle kterého jsou, zjednodušeně řečeno, libovolní dva lidé ve světové populaci propojeni relativně krátkým řetězcem osob, kdy v řetězci jsou vždy spojeni takoví jedinci, kteří se navzájem znají.

Jednou z prvních studií, které myšlenku small-world formulovaly, je práce Stanleyho Milgrama [26] z roku 1967. V rámci této práce, s cílem prozkoumat sociální vazby v americké populaci, provedl Milgram experiment. Tzv. „startující“ jedinci dostali za úkol doručit zásilku tzv. „cílovému“ jedinci. Náhodně vybraný vzorek startujících jedinců žijících ve městě Wichita, Kansas, obdržel zásilku se základními informacemi o náhodně určeném cílovém jedinci žijícím v Bostonu, Massachusetts.

Společně s informacemi o cílovém jedinci byl v zásilce přiložen popis jednoduchých pravidel experimentu – jestliže daný startující jedinec cílového člověka zná osobně (kde „znát se osobně“ znamená, že se vzájemně oslovují křestním jménem), může mu zásilku poslat přímo. Jestliže ho osobně nezná, má se zamyslet nad člověkem z okruhu svých známých, který by cílového jedince znát mohl, a poslat zásilku jemu. Tento vybraný známý má postupovat obdobným způsobem.

Přestože mnoho ze započatých řetězců předání zásilky nebylo dokončeno, z těch, které se časem dostaly k cílovému jedinci, vyplynulo, že průměrný počet prostředníků (tj. délka řetězce mezi startujícím a cílovým jedincem) je mezi pěti a šesti. Teorie, podle které je počet prostředníků roven právě šesti, je potom označována jako „šest stupňů separace“, [14].

Ačkoliv reálný počet stupňů separace mezi světovou populací nemusí být avizovaných šest, skutečnost, že libovolní dva jedinci jsou propojeni relativně krátkým řetězcem prostředníků, byla pozorována i v dalších studiích. V roce 2011 byla provedena studie analyzující sociální síť 721 miliónů v té době aktivních uživatelů Facebooku. Pro danou strukturu byla mimo jiné určena průměrná nejkratší cesta délky 4.71, [33]. To pro danou sociální síť odpovídá průměrnému stupni separace 3.71.

Protože naší snahou je, na rozdíl od předchozí kapitoly, sestaveným modelem populace co nejdříve reflektovat realitu, použijeme dále v simulacích šíření inovací jako danou grafovou strukturu G právě small-world síť.

Modelování small-world

Myšlenku small-world kvantifikuje článek Wattse a Strogatz [37] z roku 1998. V daném článku Watts a Strogatz upozorňují, že k modelování sítí (např. elektrické soustavy, neuronových sítí či sociálních sítí) se obvykle používají buď struktury s pravidelnými vazbami, nebo struktury s vazbami čistě náhodnými. S ohledem na teorii small-world však tyto dva extrémní případy nemusejí být vhodné.

Graf G označují jako small-world síť, jestliže se vyznačuje

- vysokou mírou shlukování,
- krátkou průměrnou nejkratší cestou.

Populaci, ve které se známí konkrétního jedince pravděpodobně navzájem znají, lze popsat sítí se shluky, [14]. Uvnitř shluků je struktura provázanější, jednotlivé shluky jsou zároveň menším počtem hran propojeny mezi sebou. Přítomnost shluků může být v realitě zapříčiněna fyzickou vzdáleností, sociální vzdáleností, jazykovou bariérou, pohlavím, zemí původu, atd. Ve výše uvedené studii uživatelů Facebooku bylo zjištěno, že 84.2 % hran popisujících propojení uživatelů vede v rámci jednoho státu, zbytek hran propojuje uživatele napříč různými státy.

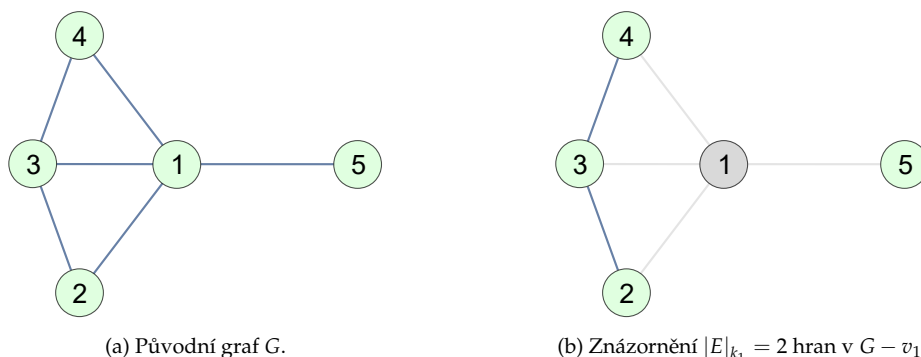
Krátká průměrná nejkratší cesta koresponduje s ideou šesti stupňů separace. Je určena délka tzv. „charakteristické cesty“ L , což je průměrný počet hran v nejkratší cestě mezi libovolnými dvěma vrcholy v dané síti.

K měření míry shlukovosti v síti je použit shlukový koeficient C . Vrchol v má v uvažované síti k_v sousedních vrcholů. Jestliže samotný vrchol v vynecháme, mezi danými k_v vrcholy může existovat nejvýše $\frac{k_v(k_v-1)}{2}$ hran. Lokální shlukový koeficient C_v je potom počet existujících hran mezi k_v sousedy uzlu v ku jejich maximálnímu možnému počtu. Označíme-li si množinu těchto hran E_v a jejich počet $|E_v|$, pak platí

$$C_v = \frac{|E_v|}{\frac{k_v(k_v-1)}{2}} = \frac{2|E_v|}{k_v^2 - k_v}. \quad (4.1)$$

Shlukový koeficient C dané sítě je určen jako průměrný lokální shlukový koeficient.

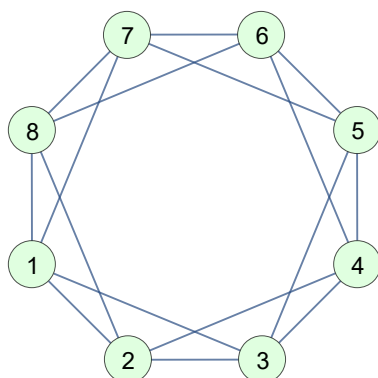
Příklad určení lokálního shlukového koeficientu C_1 pro vrchol v_1 je na Obrázku 4.1. Z původního grafu G na Obrázku 4.1a odstraníme v_1 . Vrchol v_1 má v původním grafu $k_1 = 4$ sousedů. Na čtyřech vrcholech může existovat nejvýše $\frac{4(4-1)}{2} = 6$ hran. Po odstranění v_1 (znázorněno šedou barvou) zbývají mezi k_1 sousedy $|E_{v_1}| = 2$ hrany. Tyto hrany jsou znázorněny na Obrázku 4.1b. Lokální shlukový koeficient pro vrchol v_1 je proto $C_1 = \frac{2}{6} = \frac{1}{3}$.



Obrázek 4.1: Znázornění určení lokálního shlukového koeficientu C_1 .

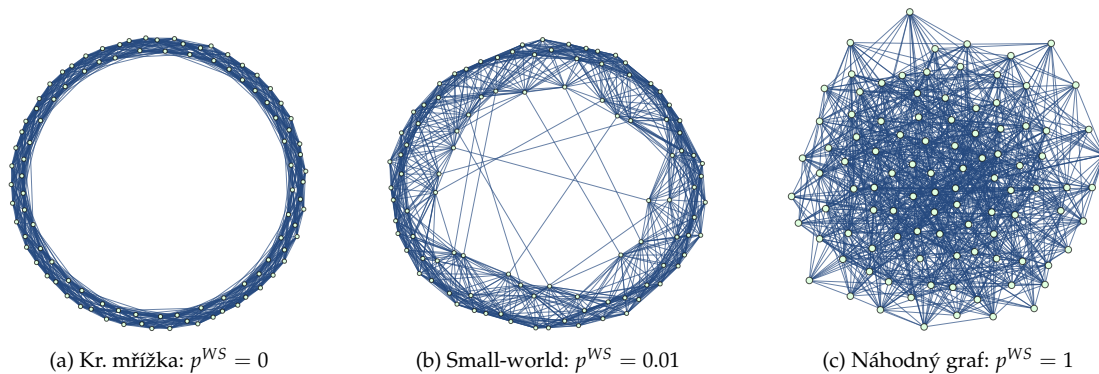
Small-world struktura G je získána zakomponováním určitého stupně náhodnosti do pravidelné struktury, tj. jedná se kombinaci pravidelného a náhodného grafu. Tato náhodnost je určena pravděpodobností p^{WS} .

Při tvorbě modelu small-world struktury dle Wattse a Strogatze začínáme s tzv. k -regulární kruhovou mřížku na n vrcholech. V k -regulární kruhové mřížce na n vrcholech je každý vrchol spojen s vrcholy do vzdálenosti k v kružnici C_n , tj. jedná se o k -tou mocninu kružnice C_n , [9]. Pro $n = 8$ a $k = 4$ je tato pravidelná, nenáhodná struktura na Obrázku 4.2.

Obrázek 4.2: Příklad pravidelné 4-regulární kruhové mřížky na $n = 8$ vrcholech.

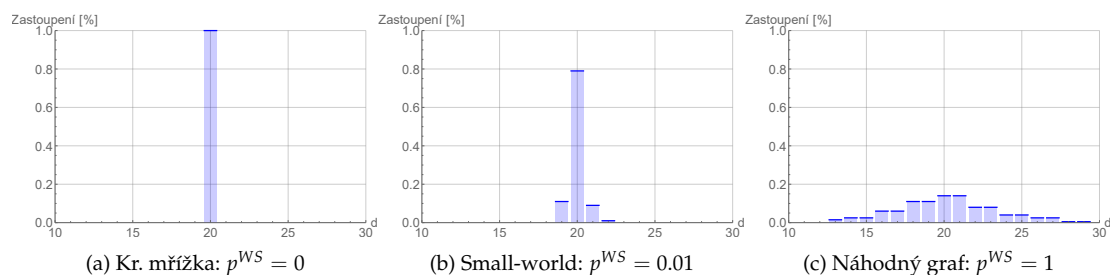
V kruhové mřížce zvolíme vrchol v a hranu, která v spojuje s jeho nejbližším sousedem u_1 (ve směru hodinových ručiček). Následně s pravděpodobností p^{WS} tuto hranu „přepojíme“ z u_1 do libovolného jiného vrcholu v síti, s tím, že násobné hrany nejsou povoleny. Po směru hodinových ručiček takto postupujeme u každého vrcholu v síti a jeho nejbližšího souseda. Dále uvažujeme vrchol v a jeho druhého nejbližšího souseda u_2 a opět hranu s pravděpodobností p^{WS} přepojíme z u_2 do jiného vrcholu v síti. Obdobným způsobem postupujeme, dokud není každá hrana v původní mřížce „podrobena“ pokusu o přepojení.

Příklad sítě získané z pravidelné mřížky s $n = 100$ a $k = 20$ na Obrázku 4.3a s rozdělením stupňů znázorněným na Obrázku 4.4a při pravděpodobnosti přepojení $p^{WS} = 0.01$ je na Obrázku 4.3b, rozdělení stupňů vrcholů dané struktury je na Obrázku 4.4b.

Obrázek 4.3: Struktury získané v závislosti na pravděpodobnosti p^{WS} , která pro každou hranu v pravidelné kruhové mřížce s $n = 100$ a $k = 20$ udává pravděpodobnost, že dojde k náhodnému přepojení.

Pro $p^{WS} = 1$ dostáváme náhodný graf. Příklad náhodného grafu získaného z pravidelné mřížky s $n = 100$ a $k = 20$ je na Obrázku 4.3c, jeho rozdělení stupňů vrcholů je na Obrázku 4.4c.

Watts a Strogatz ve své práci zkoumají tři typy reálných sítí. Jako zástupce sociální struktury testují síť herců, kde mezi vrcholy, označující jednotlivé herce, vede hrana, pokud daní herci společně účinkovali ve filmu. Dále je zkoumána síť popisující elektrické vedení na západě



Obrázek 4.4: Procentuální zastoupení stupňů vrcholů ve struktuře získané v závislosti na pravděpodobnosti p^{WS} , která pro každou hranu v pravidelné kruhové mřížce s $n = 100$ a $k = 20$ udává pravděpodobnost, že dojde k náhodnému přepojení.

USA a v neposlední řadě síť modelující nervovou soustavu červa *Caenorhabditis elegans* (česky *Háďátka obecné*).

Pro všechny uvedené reálné struktury jsou Wattsem a Strogatzem určeny hodnoty L a C a tyto hodnoty jsou porovnány s hodnotami L_R , resp. C_R pro náhodný graf získaný z mřížky na stejném počtu vrcholů n a se stupněm vrcholů k , kde k je průměrný počet hran v odpovídající reálné síti. Získané hodnoty, uvedené v Tabulce 4.1, dle Wattse a Strogatze naznačují, že ve všech případech jde o small-world struktury. Délka charakteristické cesty reálné sítě se hodnotově blíží délce charakteristické cesty v náhodném grafu, zatímco shlukový koeficient je v případě reálné sítě řádově vyšší, než v případě odpovídajícího náhodného grafu.

	L	L_R	C	C_R
Síť herců	3.650	2.990	0.790	0.001
Elektrické vedení	18.700	12.400	0.080	0.005
Háďátka obecné	2.650	2.250	0.280	0.050

Tabulka 4.1: Empirické příklady small-world sítí, [37, Table 1].

V článku Wattse a Strogatze [37] je vykreslen normalizovaný shlukový koeficient $\frac{C(p^{WS})}{C(0)}$ a normalizovaná charakteristická cesta $\frac{L(p^{WS})}{L(0)}$ pro různé pravděpodobnosti přepojení p^{WS} (normalizace je provedena hodnotami pro odpovídající regulární mřížku $C(0)$ a $L(0)$). Z obrázku [37, Figure 2] je patrné, že hodnota $p^{WS} = 0.01$ poskytuje jak vysokou hodnotu shlukového koeficientu, tak relativně krátkou charakteristickou cestu. V simulacích, kde jsou z pravidelné regulární mřížky generovány small-world struktury, je proto dále použita hodnota pravděpodobnosti přepojení $p^{WS} = 0.01$.

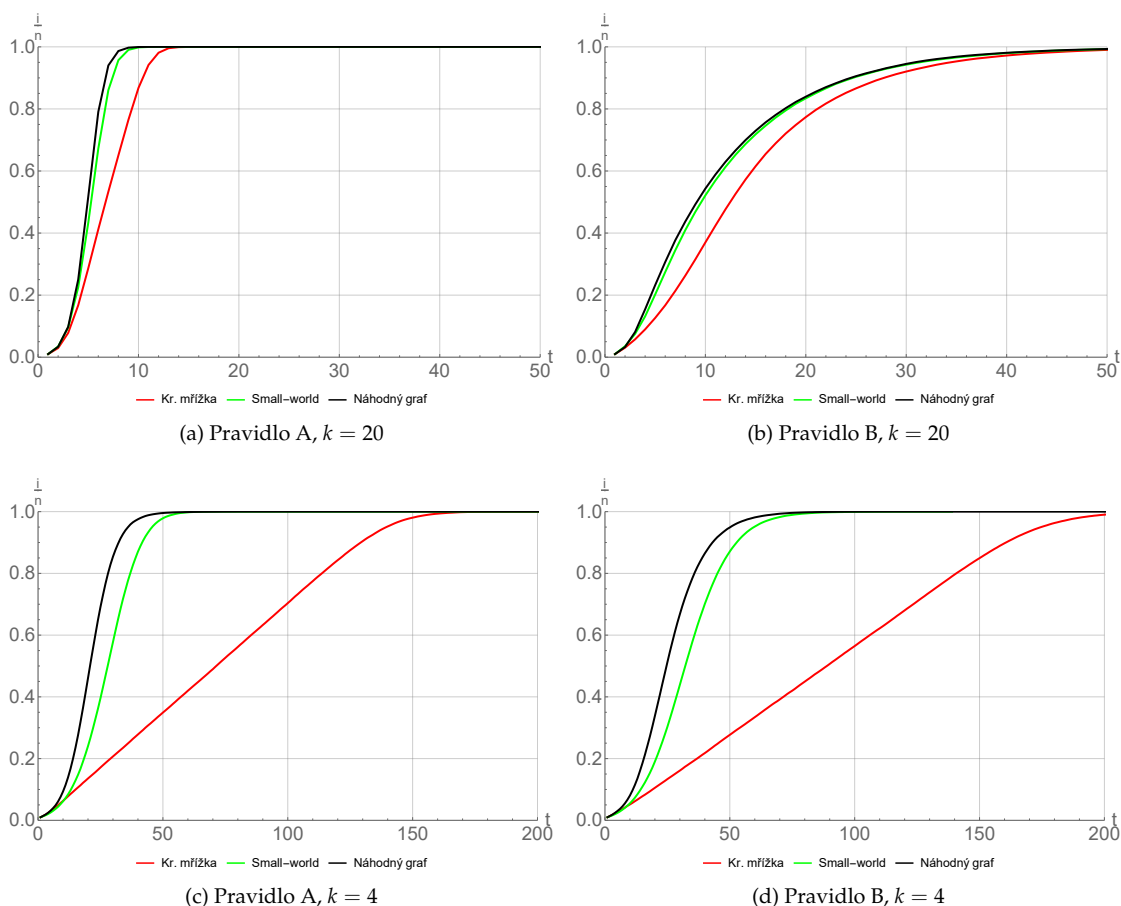
Simulační srovnání

Podívejme se nyní na simulační srovnání rychlosti adopce pro pravidelnou k -regulární mřížku, small-world strukturu, která vznikne z pravidelné k -regulární mřížky při pravděpodobnosti přepojení $p^{WS} = 0.01$ a náhodný graf, který vznikne z pravidelné k -regulární mřížky při pravděpodobnosti přepojení $p^{WS} = 1$.

Na Obrázku 4.5a je vykreslen poměr nakažených v závislosti na čase pro pravidlo adopce A, na Obrázku 4.5b je poměr nakažených vykreslen pro pravidlo B. Jedná se o srovnání rychlosti

šíření ve strukturách vzniklých z pravidelné 20-regulární kruhové mřížky s výše popsanými pravděpodobnostmi přepojení p^{WS} .

Obdobným způsobem je vykreslen poměr nakažených na Obrázcích 4.5c pro pravidlo A a 4.5d pro pravidlo B, jestliže small-world a náhodnou strukturu vytváříme z pravidelné 4-regulární kruhové mřížky.



Obrázek 4.5: Poměr počtu nakažených i v populaci velikosti $n = 100$ v závislosti na čase t pro různé struktury populace získané z pravidelné k -regulární mřížky, $p = 0.1$, 1000 simulací.

Z Obrázku 4.5 je patrné, že šíření inovace v populaci je nejrychlejší pro náhodný graf. Nicméně inovace se šíří relativně rychle i ve struktuře small-world, což souvisí s vlastností krátké charakteristické cesty L , přestože si tato struktura zachovává poměrně vysoký shlukový koeficient C . Nejpomaleji je inovace šířena v pravidelné k -regulární mřížce, která se sice, stejně jako small-world, vyznačuje vysokým shlukovým koeficientem C , nicméně v této struktuře je charakteristická cesta L nejdelší. Délka charakteristické cesty L a shlukový koeficient C jsou pro struktury G , vzniklé z k -regulární mřížky, uvedeny v Tabulce 4.2.

Simulačně jsou určeny průměrné časy absorpce pro grafové struktury vycházející z pravidelných mřížek s $k = 20$, $k = 10$, $k = 6$ a $k = 4$, zkoumána jsou obě pravidla adopce

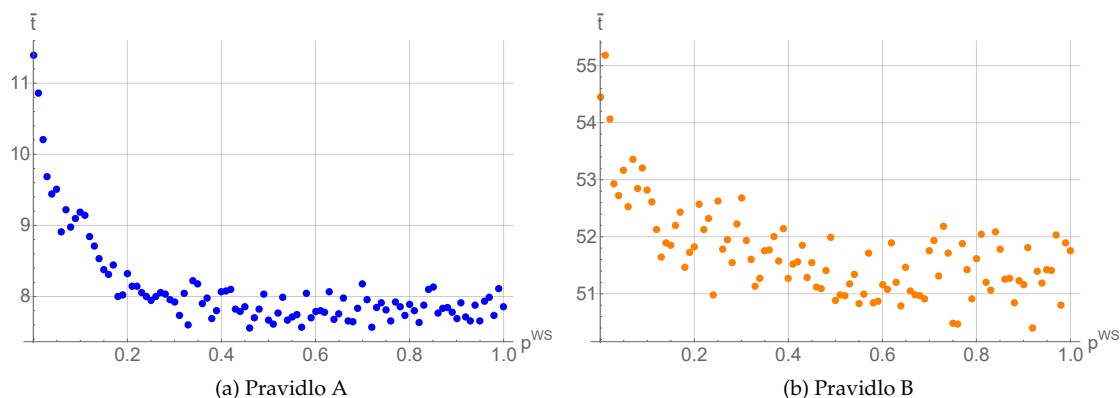
G	L	C	G	L	C
Kr. mřížka	2.980	0.711	Kr. mřížka	12.878	0.500
Small-world	2.036	0.562	Small-world	4.819	0.392
Náhodný graf	1.809	0.201	Náhodný graf	3.488	0.009

(a) $k = 20$ (b) $k = 4$

Tabulka 4.2: Délka charakteristické cesty L a shlukový koeficient C pro pravidelnou k -regulární mřížku, small-world síť, která vznikne z pravidelné k -regulární mřížky při pravděpodobnosti přepojení $p^{WS} = 0.01$ a náhodný graf, který vznikne z pravidelné k -regulární mřížky při pravděpodobnosti přepojení $p^{WS} = 1$.

s pravděpodobností adopce $p = 0.1$. Výsledky simulací jsou uvedeny v Tabulce 4.3a pro pravidlo A a v Tabulce 4.3b pro pravidlo B.

Všimněme si, že v případě obou pravidel se inovace šíří tím rychleji, čím vyšší je hodnota k . Zároveň rychlost šíření ovlivňuje velikost populace daná parametrem n , kdy pro větší populaci trvá rozšíření inovace delší dobu. Stejně jako v kapitole 3 je čas adopce menší, pokud k modelování šíření inovace používáme pravidlo A. Na Obrázku 4.6 je průměrný čas absorpce v závislosti na pravděpodobnosti přepojení p^{WS} pro obě uvažovaná pravidla adopce.



Obrázek 4.6: Průměrný čas absorpce pro populaci velikosti $n = 100$ v závislosti na pravděpodobnosti přepojení p^{WS} pro různé struktury populace získané z pravidelné 20-regulární mřížky, $p = 0.1$, 1000 simulací.

4.3 Přítomnost imunních jedinců v populaci

Ať už je na daný model nahlíženo z marketingového či epidemiologického pohledu, v obou případech může populace velikosti n obsahovat jak potenciálně adoptující členy, tak přirozeně „imunní“ jedince.

V případě marketingu zjevně nemusí být předpoklad, že každý člen populace inovaci časem adoptuje, správný. V populaci se mohou vyskytovat jedinci, kteří si daný výrobek nikdy nekoupí.

$G \backslash n$	40	60	80	100
Kr. mřížka, $k = 20$	6.445	8.007	9.687	11.493
Small-world, $k = 20$	6.300	7.472	7.875	8.215
Náhodný graf, $k = 20$	6.246	6.843	7.342	7.662
Kr. mřížka, $k = 10$	13.955	19.234	24.575	29.790
Small-world, $k = 10$	11.260	14.370	16.284	17.545
Náhodný graf, $k = 10$	10.900	13.144	15.951	15.884
Kr. mřížka, $k = 6$	29.977	43.251	56.531	69.965
Small-world, $k = 6$	24.639	24.776	35.938	32.879
Náhodný graf, $k = 6$	20.103	23.183	29.675	26.081
Kr. mřížka, $k = 4$	58.483	86.864	114.795	148.313
Small-world, $k = 4$	52.265	55.587	66.911	48.817
Náhodný graf, $k = 4$	39.678	41.329	44.560	46.655

(a) Pravidlo A

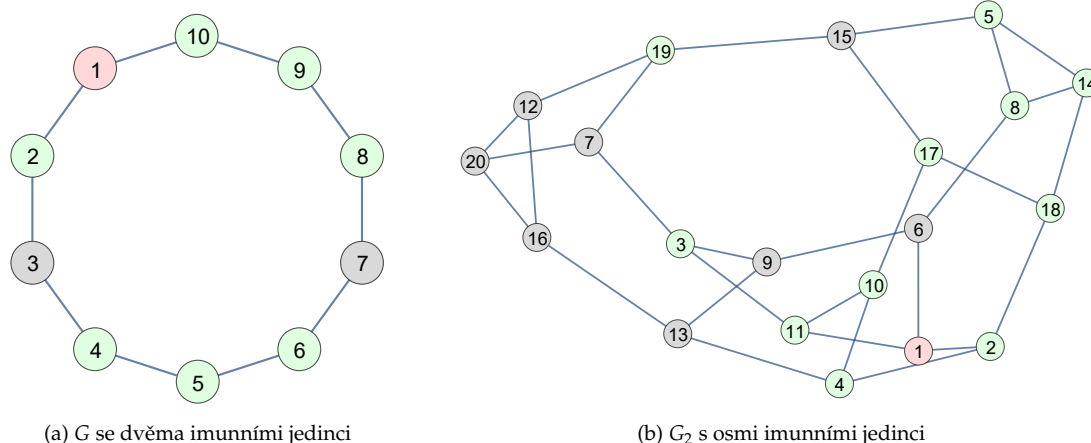
$G \backslash n$	40	60	80	100
Kr. mřížka, $k = 20$	41.718	47.284	51.301	54.996
Small-world, $k = 20$	42.245	46.523	49.309	51.384
Náhodný graf, $k = 20$	41.913	45.458	49.155	51.959
Kr. mřížka, $k = 10$	46.356	55.471	64.056	71.211
Small-world, $k = 10$	43.977	50.292	53.329	56.646
Náhodný graf, $k = 10$	42.659	47.813	52.233	54.920
Kr. mřížka, $k = 6$	59.508	76.625	95.467	112.588
Small-world, $k = 6$	53.364	56.381	70.169	65.946
Náhodný graf, $k = 6$	46.675	51.697	60.638	58.844
Kr. mřížka, $k = 4$	84.763	119.004	153.271	187.594
Small-world, $k = 4$	76.842	81.526	99.448	77.793
Náhodný graf, $k = 4$	58.208	61.069	64.743	69.245

(b) Pravidlo B

Tabulka 4.3: Časy absorpce při pravděpodobnosti adopce $p = 0.1$ pro pravidelnou k -regulární mřížku, small-world síť, která vznikne z pravidelné k -regulární mřížky při pravděpodobnosti přepojení $p^{WS} = 0.01$ a náhodný graf, který vznikne z pravidelné k -regulární mřížky při pravděpodobnosti přepojení $p^{WS} = 1$, 1000 simulací.

Podobně je tomu u epidemiologie a simulování šíření určité nákazy. V populaci se mohou vyskytovat jedinci, kteří jsou vůči dané nemoci přirozeně imunní, a proto ne onemocní, ani pokud jsou dané infekci vystaveni, [3].

Počet imunních jedinců může ovlivnit rozšíření inovace napříč populací. V některých případech může umístění imunních členů zapříčinit, že někteří jedinci inovaci neadoptují, ačkoliv patří mezi potenciální adoptující (jejich pravděpodobnost adopce je nenulová). Jestliže takový jedinec přichází do kontaktu pouze s jedinci imunními (tj. odpovídající vrchol není spojen hranou s žádným dalším potenciálním adoptujícím), pak je „blokovan“ imunní bariérou a v daném modelu se inovace ke všem potenciálním adoptujícím nikdy zcela nerozšíří. Podívejme se na jednoduchý příklad takové situace, který je znázorněn na Obrázku 4.7a.



Obrázek 4.7: Znázornění imunních členů v populaci (šedé vrcholy) v grafu $G = C_{10}$ a v grafu $G_2 =$ náhodný 3-regulární graf na 20 vrcholech.

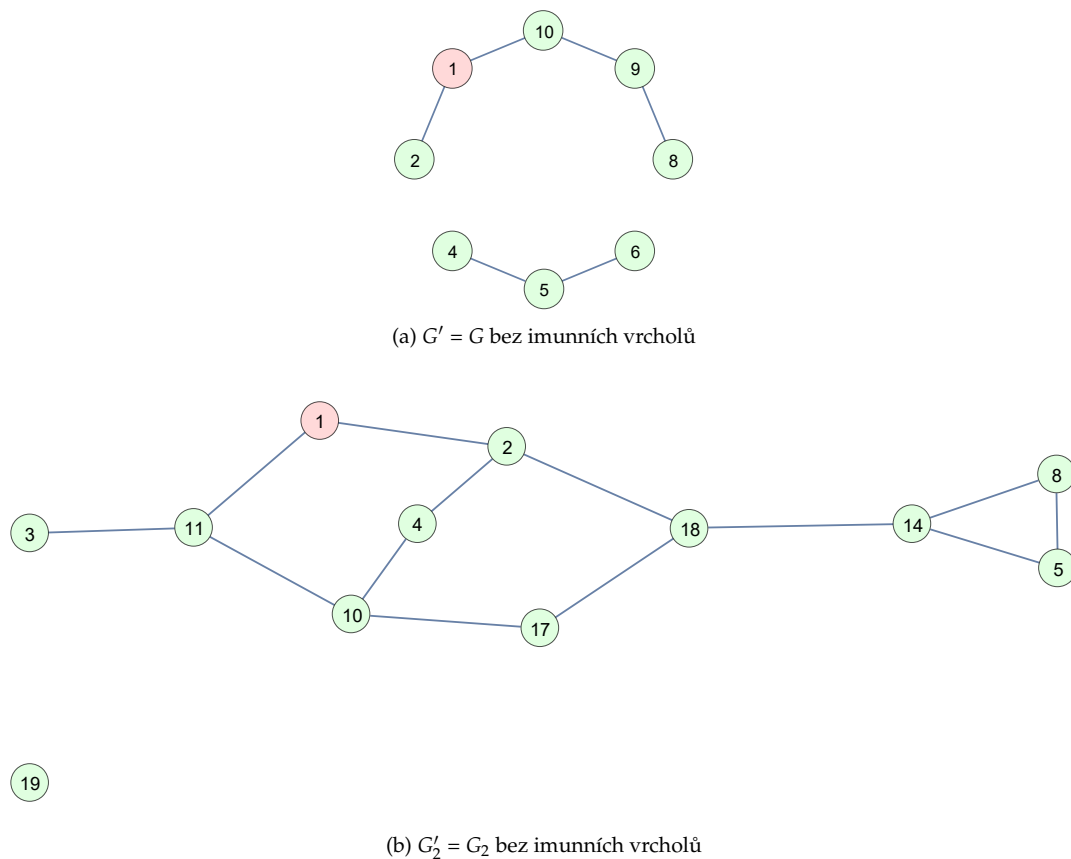
Uvažujeme nejprve populaci velikosti $n = 10$ se strukturou popsanou kružnicí C_{10} . Stejně jako v předcházejících případech umístíme do populace jednoho nakaženého (červený vrchol). V populaci se v tomto případě nacházejí dva imunní jedinci s nulovou pravděpodobností adopce, jedná se o vrcholy 3 a 7 (šedé vrcholy), tj. $p_3 = p_7 = 0$. Zbytek vrcholů odlišených zelenou barvou má nenulovou pravděpodobnost adopce. Ačkoliv vrcholy 4, 5 a 6 jsou potenciální adoptující (jejich pravděpodobnost adopce je nenulová), tito jedinci inovaci nikdy neadoptují, protože jsou obklopeni imunními členy populace.

Pro kružnici na n vrcholech je na první pohled vidět, jestli potenciální adoptující inovaci adoptují, nebo ne. V případě komplexnější grafové struktury však nemusí být na první pohled patrné, zda jsou potenciální adoptující izolováni imunní bariérou.

Už v případě náhodného 3-regulárního grafu G_2 na $n = 20$ vrcholech a s osmi imunními jedinci na pozicích 6, 7, 9, 12, 13, 15, 16 a 20, znázorněného na Obrázku 4.7b, nemusí být od pohledu zřejmé, zda se inovace může ke všem potenciálním adoptujícím rozšířit.

Pro obecné určení vrcholů, které inovaci nikdy neadoptují, ačkoliv mají nenulovou pravděpodobnost adopce (tj. vrcholů obklopených imunní bariérou), je proto v simulacích využít algoritmus prohledávání grafu do hloubky, [13]. V případě imunních jedinců je simulaci odpovídajícím vrcholům nastavena pravděpodobnost adopce $p_i = 0$.

Nejprve jsou z původních grafů G a G_2 odstraněny všechny imunní vrcholy. Získané podgrafy (označme je jako G' a G'_2) jsou na Obrázku 4.8a (podgraf G' získaný odstraněním imunních členů z G) a na Obrázku 4.8b (podgraf G'_2 získaný odstraněním imunních členů z G_2).



Obrázek 4.8: Podgrafy G' a G'_2 získané odstraněním imunních vrcholů z grafů G a G_2 .

Jestliže v získaných podgrafech je každý vrchol dosažitelný z prvního nakaženého vrcholu, potom časem inovaci adoptují všichni potenciální jedinci. Pro ověření dosažitelnosti prohledáváme podgrafy G' a G'_2 do hloubky a startujeme v prvním nakaženém vrcholu (tj. ve vrcholu na pozici 1). Jestliže jsou nalezeny všechny vrcholy mající nenulovou pravděpodobnost adopce, pak není v původních grafech G a G_2 žádný potenciální adoptující izolován imunní bariérou, a proto časem inovaci adoptují všichni potenciální jedinci.

Naproti tomu, pokud v získaném podgrafu existuje vrchol, který z prvního nakaženého vrcholu není dosažitelný (tj. podgraf získaný odstraněním imunních členů není souvislý), potom v populaci existují potenciální adoptující, kteří i přes nenulovou pravděpodobnost adopce danou inovaci neadoptují.

Na Obrázku 4.8 je vidět, že G' a G'_2 nejsou souvislé, tudíž v obou případech existují potenciální adoptující, kteří inovaci neadoptují. V případě G' se jedná o vrcholy na pozicích 4, 5 a 6 a v případě G'_2 se jedná o vrchol na pozici 19.

Simulační srovnání

Podívejme se nyní simulačně na to, kolik takových imunních jedinců musí být v populaci zastoupeno, aby se inovace ke členům s nenulovou pravděpodobností adopce nerozšířila. Tato otázka je zkoumána pro pravidelnou k -regulární mřížku, small-world síť vzniklou z dané mřížky s pravděpodobností přepojení $p^{WS} = 0.01$ a pro náhodný graf generovaný z mřížky s pravděpodobností přepojení $p^{WS} = 1$.

Procentuální zastoupení imunních členů v celkové populaci velikosti n popíšeme parametrem r_{IM} . V každé provedené simulaci je vygenerována odpovídající grafová struktura G . Do G jsou náhodně umístěni imunní jedinci, jejichž počet odpovídá parametru r_{IM} . Následně jsou vrcholy s nulovou pravděpodobností adopce odstraněny. Jestliže je získaný podgraf G' souvislý (testováno výše zmíněným algoritmem prohledávání grafu do hloubky), pak v dané simulaci adoptují všichni potenciální adoptující.

Získané procento simulací, ve kterých všichni potenciální jedinci adoptovali, ku počtu celkových provedených simulací, pak můžeme vnímat jako odhad pravděpodobnosti, že při daném poměru r_{IM} imunních jedinců adoptují všichni potenciální zájemci.

V Tabulce 4.4 jsou uvedeny odhady této pravděpodobnosti pro různá r_{IM} a pro různé grafy G . Čím nižší je poměr r_{IM} , tím vyšší je odhad pravděpodobnosti, že všichni potenciální jedinci inovaci časem adoptují.

Zároveň je z uvedených výsledků simulací patrné, že nižší konektivita k vede k nižšímu odhadu pravděpodobnosti, že se inovace rozšíří.

U small-world a náhodného grafu, které jsou z k -regulární mřížky generovány, není obecně zaručeno, že získaná struktura bude souvislá. Jestliže je tedy v dané simulaci G vygenerován nesouvislý, pak by se inovace rozšířila jen v případě, že by všichni potenciální jedinci byli umístěni ve stejné komponentě grafu G (společně s prvním nakaženým).

$r_{IM} \backslash G$	0.05	0.10	0.25	0.50	0.75	0.90	0.95
Kr. mřížka, $k = 20$	1	1	1	1	0.586	0.009	0.004
Small-world, $k = 20$	1	1	1	1	0.608	0.007	0.013
Náhodný graf, $k = 20$	1	1	1	1	0.905	0.222	0.093
Kr. mřížka, $k = 10$	1	1	0.999	0.555	0	0	0
Small-world, $k = 10$	1	1	1	0.613	0.001	0	0.002
Náhodný graf, $k = 10$	0.998	0.998	0.987	0.849	0.141	0.005	0.006
Kr. mřížka, $k = 6$	1	0.999	0.713	0	0	0	0
Small-world, $k = 6$	1	1	0.802	0.002	0	0	0
Náhodný graf, $k = 6$	0.900	0.852	0.659	0.143	0.002	0	0.001
Kr. mřížka, $k = 4$	0.993	0.853	0.004	0	0	0	0
Small-world, $k = 4$	0.995	0.869	0.025	0	0	0	0
Náhodný graf, $k = 4$	0.343	0.256	0.069	0.001	0	0	0.001

Tabulka 4.4: Odhad pravděpodobnosti, že při poměru r_{IM} náhodně umístěných imunních jedinců ve struktuře G na $n = 100$ vrcholech se časem nakazí všichni potenciálně adoptující v populaci, 1000 simulací.

4.4 Nehomogenní pravděpodobnosti adopce

Zaměřme se nyní na situaci, kdy jednotlivci v populaci popsané k -regulárním grafem mají nenulové a nehomogenní pravděpodobnosti adopce.

Zajímavou otázkou může být, v jakém typu populace se inovace rozšíří rychleji. Jestli v homogenní populaci s pravděpodobností adopce p , nebo v populaci nehomogenní, kdy průměr pravděpodobností adopce jednotlivých členů odpovídá pravděpodobnosti adopce p v homogenní populaci.

Mějme populaci s homogenní pravděpodobností adopce $p = 0.5$. Pro srovnání dále uvažujme populaci s nehomogenními pravděpodobnostmi adopce. V této populaci jsou přítomny dvě stejně velké skupiny jedinců. V první skupině jsou jedinci s nižší pravděpodobností adopce než p , ve druhé skupině s vyšší.

Uvažování pouhých dvou skupin s různou pravděpodobností adopce v celé populaci zjevně, stejně jako v případě homogenních pravděpodobností, značně zjednodušuje komplexnější realitu. Přesto tato situace intuitivně lépe modeluje reálný svět. Jako tyto dvě rozdílné skupiny můžeme uvažovat např. skupiny dané pohlavím jedinců, kdy ženy mohou mít u určitých produktů vyšší pravděpodobnost pořízení a naopak.

Rozdíl pravděpodobností nehomogenních skupin popíšme parametrem q . V první skupině mají členi populace pravděpodobnost adopce $p_1 = p - q$, analogicky pak ve druhé skupině je pravděpodobnost adopce $p_2 = p + q$.

Simulační srovnání

Protože jsou simulace prováděny pro k -regulární grafy na $n = 100$ vrcholech, v případě nehomogenní populace má vrchol na pozici 2 v simulacích nastavenou pravděpodobnost adopce p (aby obě nehomogenní skupiny měly stejný počet členů). Stejně jako v případě předchozích kapitol je vrchol na pozici 1 uvažován jako první nakažený jedinec v populaci.

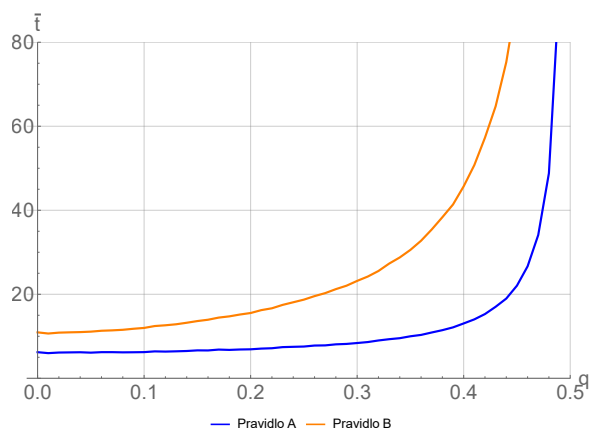
Pro zbývající vrcholy na pozicích 3, ..., n je náhodně určeno, zda daný vrchol patří do první skupiny s pravděpodobností adopce p_1 , nebo do skupiny druhé, ve které je pravděpodobnost adopce p_2 .

Rychlost rozšíření inovace je určena pro homogenní populaci (odpovídá $q = 0$). Dále je čas absorpce určen pro populaci nehomogenní s parametrem $q = 0.1$ (tj. pravděpodobnosti adopce nehomogenních skupin jsou $p_1 = 0.4$ a $p_2 = 0.6$), ..., $q = 0.4$ (tj. pravděpodobnosti adopce nehomogenních skupin jsou $p_1 = 0.1$ a $p_2 = 0.9$).

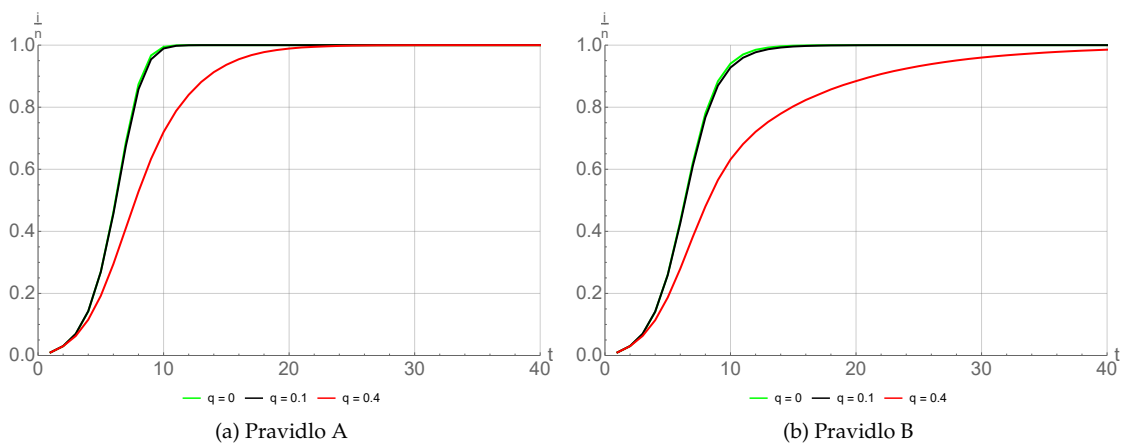
Přestože je průměrná pravděpodobnost adopce v případě homogenní i nehomogenní populace shodná, rychlost šíření inovace se může pro oba uvažované typy populace lišit, jak je naznačeno v Tabulce 4.5.

Získané výsledky indikují, že v případě nehomogenní populace vyšší hodnota parametru q (tj. větší rozdíl v pravděpodobnostech adopce nehomogenních skupin) vede k delšímu času absorpce. Závislost průměrného času absorpce na hodnotě parametru q pro 6-regulární graf je na Obrázku 4.9 pro obě pravidla adopce. Přestože nejprve průměrný čas absorpce roste s rostoucím q pomalu, po překonání určité hodnoty q průměrný čas absorpce prudce narůstá.

Na Obrázku 4.10 jsou vykresleny poměry počtu nakažených v čase pro homogenní populaci a pro nehomogenní populaci s $q = 0.1$ a $q = 0.4$. Vidíme, že v případě $q = 0.1$ pro obě pravidla je rozdíl v rychlosti šíření inovace mezi homogenní a nehomogenní populací menší, než pro $q = 0.4$.



Obrázek 4.9: Průměrný čas absorpce v závislosti na q pro 6-regulární graf na $n = 100$ vrcholech, 1000 simulací.



Obrázek 4.10: Poměr počtu nakažených i v populaci velikosti $n = 100$ v závislosti na čase t pro různá q popisující rozdíl mezi pravděpodobnostmi u nehomogenních skupin, 4-regulární graf, $p = 0.5$, 1000 simulací.

$G \backslash q$	0	0.1	0.2	0.3	0.4
6-reg.	6.200	6.424	7.131	8.616	13.005
4-reg.	8.924	9.325	10.342	12.819	20.294
2-reg.	99.461	103.259	117.750	154.372	273.792

(a) Pravidlo A

$G \backslash q$	0	0.1	0.2	0.3	0.4
6-reg.	10.771	12.218	15.869	23.017	45.912
4-reg.	12.842	14.099	17.525	25.280	48.462
2-reg.	99.621	71.297	87.465	126.697	251.625

(b) Pravidlo B

Tabulka 4.5: Srovnání času absorpce pro homogenní populaci s pravděpodobností adopce $p = 0.5$ (odpovídá sloupci s $q = 0$) s rychlostí adopce v populaci nehomogenní, kde parametr q udává rozdíl pravděpodobností adopce u nehomogenním skupin, $n = 100$, 1000 simulací.

Podívejme se nyní na parametr q s jemnějším rozlišením. V Tabulce 4.6 jsou uvedeny průměrné časy absorpce pro $q = 0$, $q = 0.01$, $q = 0.02$, ..., $q = 0.10$, uvažujeme-li 6-regulární graf na $n = 100$ vrcholech.

q	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
A	6.211	6.218	6.213	6.238	6.249	6.271	6.296	6.323	6.347	6.385	6.408
B	10.911	10.915	10.937	11.012	11.091	11.213	11.401	11.546	11.741	11.935	12.179

Tabulka 4.6: Srovnání času absorpce pro homogenní populaci s pravděpodobností adopce $p = 0.5$ (odpovídá sloupci s $q = 0$) s rychlostí adopce v populaci nehomogenní, kde parametr q udává rozdíl pravděpodobností adopce u nehomogenním skupin, uvažována pravidla A a B, 6-regulární graf, $n = 100$, 10000 simulací.

I v tomto případě se inovace rozšíří rychleji pro menší hodnoty parametru q , přestože průměrná pravděpodobnost adopce v populaci je ve všech případech stejná.

Závěr 5

V diplomové práci jsme se zabývali návrhem matematických modelů tržní penetrace po uvedení nového produktu na trh (nebo obecněji návrhem matematických modelů šíření inovace v populaci). Přestože v současné době existuje množství matematických modelů šíření inovací, tyto modely v mnoha případech předpokládají homogenní, nekonečnou populaci a spojitý čas. Představili jsme jeden z nejznámějších z těchto modelů, tzv. Bassův model.

S ohledem na teorii small-world však předpoklad nestrukturované populace nemusí být správný. Podle této teorie mají lidé tendenci sdružovat se do shluků (např. vlivem národnosti, pracovních týmů, rodiny, atd.), zároveň však řetězec osob mezi dvěma libovolnými jedinci na světě, kde v řetězci jsou spojeni lidé, kteří se vzájemně znají, je relativně krátký.

Stejně tak předpoklad homogenní populace je značně zjednodušený, protože mezi lidmi mohou existovat rozdíly zapříčínující odlišný přístup k adoptování inovace (např. odlišná ekonomická situace, vzdělání, věk, pohlaví jedinců, atd.). V populaci se rovněž mohou vyskytovat jedinci, kteří danou inovaci nikdy neadoptují.

Vzhledem k výše uvedeným nereálným předpokladům spojitých modelů bylo cílem diplomové práce sestavení modelů alternativních, uvažujících diskrétní stavový prostor, diskrétní čas, nehomogenní, konečnou a strukturovanou populaci. Tyto alternativní modely poskytují možnost reálněji popsat mnohdy velmi komplexní a nehomogenní sociální sítě. Ač v některých případech více odpovídají realitě, diskrétní modely bývají složitější nejen na implementaci, ale i na různá analytická odvození jejich chování. Proto jsou pro ně výsledky mnohdy určovány pouze za pomoci simulací.

Ačkoliv jsou v této práci některé vztahy diskrétních modelů odvozeny analyticky (průměrný čas absorpce Markovského řetězce na úplném grafu, na cestě a na hvězdě), pro šíření inovace na složitějších grafových strukturách byly výsledky určeny simulačně. Byla studována dvě pravidla, podle kterých mohou jedinci inovaci přijímat. Jestliže je inovace přijímána podle pravidla A, pak nenakaženého jedince může každý nakažený známý nakazit s pravděpodobností p . U pravidla B nenakažený jedinec inovaci přijme s pravděpodobností p , pokud je alespoň jeden jeho známý nakažen. Simulace postupného šíření inovace na zvolených grafech pro definovaná pravidla adopce byly implementovány v prostředí MATLAB. Samotné grafové struktury použité v simulacích pak byly generovány za pomoci Wolfram Mathematica.

Z analytických výsledků i provedených simulací je patrné, že inovace se rozšířila rychleji v grafech s vyšší konektivitou, vyšší průměrnou pravděpodobností adopce a menším počtem vrcholů (kromě pravidla A na úplném grafu, pro který se s větší populací čas absorpce snižuje). Z uvažovaných pravidel adopce se inovace šířila rychleji v případě pravidla A, protože zvyšující

se počet adoptujících známých daného jedince zvyšuje pravděpodobnost, že on sám inovaci rovněž adoptuje.

Kromě pravidelných a symetrických grafů byla rychlost šíření inovace zkoumána rovněž pro složitější struktury. Sociální sítě byly generovány na základě Wattsova-Strogatzova modelu. Tyto struktury mají vlastnosti odpovídající výše zmíněným small-world sítím. I v případě malé pravděpodobnosti připojení p^{WS} při generování small-world sítě z pravidelné kruhové mřížky se inovace ve small-world síti rozšířila výrazně rychleji než v odpovídající kruhové mřížce.

Rychlost šíření inovace snižuje přítomnost imunních členů populace, kdy v některých případech se vlivem bariéry vytvořené imunními jedinci (tj. potenciálního adoptujícího mohou ovlivnit pouze jedinci, kteří sami nikdy neadoptují) inovace nerozšířila ke všem potenciálním zájemcům o inovaci.

V neposlední řadě byla porovnána rychlost šíření inovace v populaci homogenní a v populaci nehomogenní. Přestože v každé simulaci byla průměrná pravděpodobnost adopce v populaci stejná, inovace se rychleji rozšířila v populaci homogenní. Průměrný čas absorpce u nehomogenní populace se zvyšuje s rozdílem mezi pravděpodobnostmi adopce nehomogenních skupin.

Kromě zkoumání vlivu parametrů alternativních modelů na šíření inovace v populaci byly uvedeny podmínky konvergence Markovského řetězce na úplném grafu k Bassovu modelu. Možným rozšířením diplomové práce by bylo zkoumání limitního chování Markovských řetězců na dalších grafových strukturách.

Literatura

- [1] Allen L., J. S., 2007, *An Introduction to Mathematical Biology*, Pearson/Prentice Hall.
- [2] Ansoff H. I., 1957, 'Strategies for Diversification', *Harvard Business Review*, 35(5), s. 113—124.
- [3] Arbor A., 2011, *Breakthrough: Researchers find wide gap in immune responses of people who did or didn't get the flu after exposure*, Michigan News: University of Michigan, přistoupeno 04.04.2020, dostupné z <https://news.umich.edu/breakthrough-researchers-find-wide-gap-in-immune-responses-of-people-who-did-or-didnt-get-the-flu-after-exposure/>.
- [4] Bass F. M., 1969, 'A new product growth for model consumer durables', *Management Science*, 15(5), s. 215—227.
- [5] Bass F. M., 2004, 'Comments on 'A New Product Growth for Model Consumer Durables': The Bass Model', *Management Science*, 50(12), s. 1833—1840.
- [6] Berek L., 2018, 'Když se matematika potká s biologií: matematická epidemiologie', *Pokroky matematiky, fyziky a astronomie*, 63, s. 91–107.
- [7] Bernhardt I., MacKenzie K.M., 2004, 'Some problems in using diffusion models for new products', *Management Science*, 19(1972), s 187–200.
- [8] Bertrand J. T., 2004, 'Diffusion of Innovations and HIV/AIDS', *Journal of Health Communication*, 9(1), s 113–121.
- [9] Bondy J. A., Murty U. S. R., 2008, *Graduate Texts in Mathematics 244: Graph Theory*, Springer, Londýn. ISBN 978-1-84628-969-9.
- [10] Bondy J. A., Murty U. S. R., 1976, *Graph Theory with Applications*, Elsevier, New York. ISBN 0-444-19451-7.
- [11] Chandrasekaran, D., Tellis, G. J., 2007, 'Diffusion and Growth of New Products: A Critical Review of Models and Findings', *Review of Marketing Research*, s. 39–80, přistoupeno 22.05.2020, dostupné z <https://ssrn.com/abstract=969775>.
- [12] de Vries G., Hillen T., Lewis M., Müller J., Schönfisch B., 2006, 'A course in mathematical biology', *Society for Industrial and Applied Mathematics*.

- [13] Even S., 1979, *Graph Algorithms*, Computer Science Press, Potomac. ISBN 0-914894-21-8.
- [14] Golbeck J., 2013, *Analyzing the Social Web*, Elsevier, Waltham. ISBN 978-0-12-405531-5.
- [15] Grinstead C. M., Snell J. L., 2009, *Introduction to probability*, Orange Grove Texts Plus. ISBN 978-1616100469.
- [16] Heeler R.M., Hustad T.P., 1980 'Problems in predicting new product growth for consumer durables', *Management Science*, 26(10), s. 1007–1020.
- [17] Internet World Stats, *Usage and Population Statistics*, přistoupeno 30.05.2020, dostupné z <https://www.internetworldstats.com/stats.htm>.
- [18] Kahn K. B., 2006, *A new product forecasting: an applied approach*, M.E. Sharpe, New York. ISBN 0-7656-1609-2.
- [19] Kijek A., Kijek T., 2010, 'Modelling of innovation diffusion', *Operations Research and Decisions*, 3-4, s. 53–68.
- [20] Kiss I. Z., Miller J. C., Simon P. L., 2017, *Mathematics of Epidemics on Networks*, Springer. ISBN 978-3-319-50804-7.
- [21] Lilien G. L., Rangaswamy A., De Bruyn A., 2007, 'The Bass Model: Marketing Engineering Technical Note', *Principles of Marketing Engineering*, přistoupeno 23.05.2020, dostupné z <http://faculty.washington.edu/sundar/NPM/BASS-Forecasting%20Model/Bass%20Model%20Technical%20Note.pdf>.
- [22] Marketing Analytics, *The Bass Diffusion Model*, přistoupeno 23.05.2020, dostupné z <https://www.ashokcharan.com/Marketing-Analytics/~pv-bass-diffusion-model.php>.
- [23] „Market Penetration“, 2020, *Cambridge Dictionary*, Cambridge University Press, přistoupeno 24.04.2020, dostupné z <https://dictionary.cambridge.org/dictionary/english/market-penetration>.
- [24] Mahajan V., Muller E., Wind Y., 2000, *New-product diffusion models*, Springer, New York. ISBN 0-7923-7751-6.
- [25] Massiani J, Gohs A., 2015, 'The choice of Bass model coefficients to forecast diffusion for innovative products: An empirical investigation for new automotive technologies', *Research in Transportation Economics*, 50, s. 17–28.
- [26] Milgram. S., 1967, 'The small world problem', *Psychology Today*, 1(1), s. 61–67.
- [27] Noonburg V. W., 2010, *Differential Equations: From Calculus to Dynamical Systems*, 2. edice, American Mathematical Society, Rhode Island. ISBN 781470444006.
- [28] Prášková Z., Lachout P., 2012, *Základy náhodných procesů I.*, MatfyzPress, Praha. ISBN 978-80-7378-210-8.
- [29] Rogers E. M., 2003, *Diffusion of innovations*, 5. edice, The Free Press, New York. ISBN 0-7432-5823-1.

- [30] Satoh D., 2001, 'A discrete Bass model and its parameter estimation', *Journal of the Operations Research Society of Japan*, 40(1), s. 1—18.
- [31] Sterman J. D., 2000, *Business Dynamics: Systems Thinking and Modeling for a Complex World*, McGraw-Hill/Irwin, Boston. ISBN 0-07-231135-5.
- [32] Strogatz S. H., 2001, *Nonlinear Dynamics and Chaos*, Westview Press.
- [33] Ugander J., Karrer B., Backstrom L., Marlow C., 2011, *The Anatomy of the Facebook Social Graph*, přístoupeno 09. 04. 2020, <https://arxiv.org/pdf/1111.4503.pdf>.
- [34] Van den Bulte C., 2002, 'Technical Report: Want to know how diffusion speed varies across countries and products? Try using a Bass model', *Visions*, 26(4), s. 12–15.
- [35] Vendrig N., 2013, *Epidemiology of innovations: Combining the Bass model with the SIR-model for prediction of diffusion of innovations*, Wageningen UR.
- [36] Wang W., Fergola P., Lombardo S., Mulone G., 2006, 'Mathematical models of innovation diffusion with stage structure', *Applied Mathematical Modelling*, 30(1), s. 129–146.
- [37] Watts D. J., Strogatz S. H., 1998, 'Collective dynamics of 'small-world' networks', *Nature*, 393(6684), s. 440–442.