

# *PB050: Modelování a predikce v systémové biologii*

David Šafránek

27.09.2011

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.



# *Obsah*

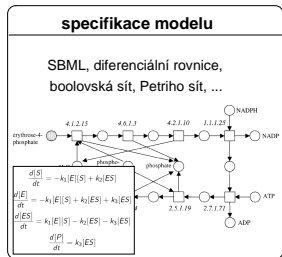
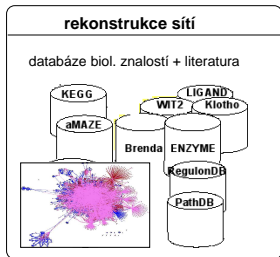
*Pojem modelu a simulace*

*Biologické sítě a dráhy*

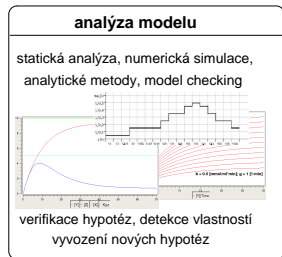
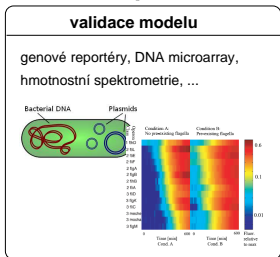
## *Shrnutí předchozí přednášky*

- biologický systém definován interakcemi mezi jeho komponentami
- interakce jsou omezeny základními zákony chemie ale i evolučním vývojem
- syntaxí organismu-systému je síť komponent
- sémantikou organismu-systému je jeho funkce (dynamika)
- základní koncepty systémové biologie
  - důraz na interakci, součinnost
  - hierarchie
  - modelování a simulace

# Workflow systémové biologie – opakování



biologická síť  
 →  
 hypotézy  
 →



objevené vlastnosti  
 ←  
 dotazy na model  
 →

# *Obsah*

*Pojem modelu a simulace*

*Biologické sítě a dráhy*

## *Proč dělat model?*

### *Cíl*

Pochopit vlastnost organismu (případně jeho reakci na dané podmínky prostředí) se všemi příčinami. Predikovat možné důsledky.

## *Proč dělat model?*

### *Cíl*

Pochopit vlastnost organismu (případně jeho reakci na dané podmínky prostředí) se všemi příčinami. Predikovat možné důsledky.

### *Problém*

Těžko realizovatelné *in vitro*, kde jsou jen omezené možnosti určitých experimentů.

## *Proč dělat model?*

### *Cíl*

Pochopit vlastnost organismu (případně jeho reakci na dané podmínky prostředí) se všemi příčinami. Predikovat možné důsledky.

### *Problém*

Těžko realizovatelné *in vitro*, kde jsou jen omezené možnosti určitých experimentů.

### *Řešení*

Vytvořit model zachycující všechny známé vztahy související se zkoumanou vlastností. Analýzou a simulací odvodit/potvrdit experimenty a dosavadní hypotézy. Předvídat nové hypotézy.

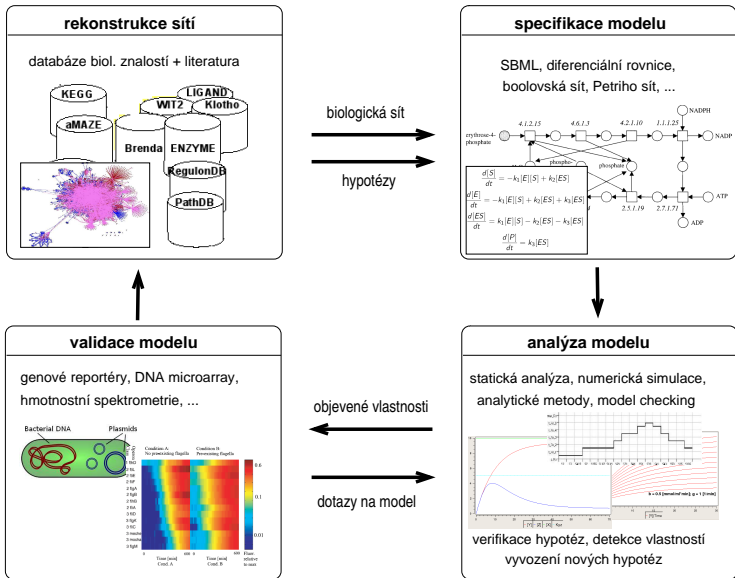


## *Typy modelů*

- statické modely
  - tzv. network biology: metabolic control analysis, flux balance analysis, ...
  - statická analýza biologické sítě jako grafu/matice
  - statistické srovnání topologie s náhodnými grafy
- dynamické modely
  - aplikace teorie dynamických systémů
  - simulace vývoje v čase
  - nutná znalost dynamických zákonů (např. reakční kinetika)

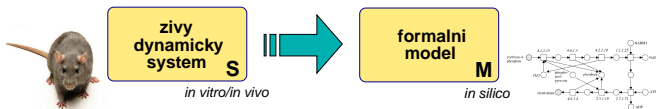
Budeme-li hovořit o modelu (bez přívlastků), budeme myslet dynamický model.

# Uplatnění modelu



# Model

- abstraktní (formální, výpočetní) model
- teoretický (idealizovaný) obraz skutečného organismu  
→ problém vztahu mezi modelem a modelovaným objektem
- sestává z množiny proměnných a množiny logických a kvantitativních relací mezi proměnnými



## *Simulace modelu*

- simulace představuje “spuštění” modelu pro dané výchozí nastavení proměnných v daném prostředí
- simulace imituje skutečné chování modelovaného objektu na určité úrovni abstrakce
- simulace umožňuje predikci hypotéz
- simulace zobecňuje a doplňuje in vivo/in vitro experimenty
  - predikce má smysl pouze pro validovaný model!

# *Obsah*

*Pojem modelu a simulace*

*Biologické sítě a dráhy*

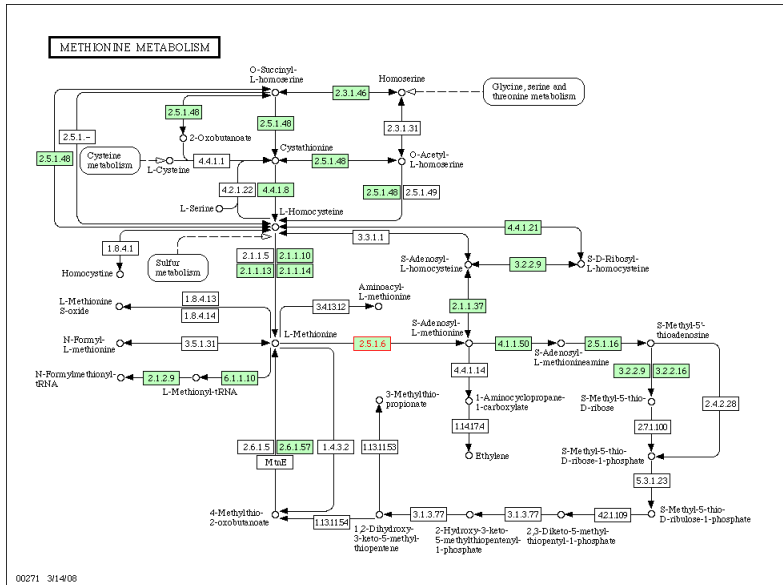
## *Biologická síť jako způsob reprezentace modelu*

- biologická síť – komplexní systémový popis organismu
  - neexistuje jednoznačná definice
  - orientovaný nebo neorientovaný graf
  - uzly představují typicky proměnné
  - hrany představují typicky (funkční) relace mezi proměnnými
- k uzlům a relacím jsou přiřazeny kvalitativní i kvantitativní informace potřebné k simulaci (dynamická analýza)
- biologické sítě lze strukturně zkoumat – statická analýza
  - srovnávání sítí různých organismů
  - vyhledávání alternativních cest
  - zkoumání měřitelných vlastností sítí
  - zkoumání změn v sítích při evoluční selekci

## *Dráhy vs. sítě*

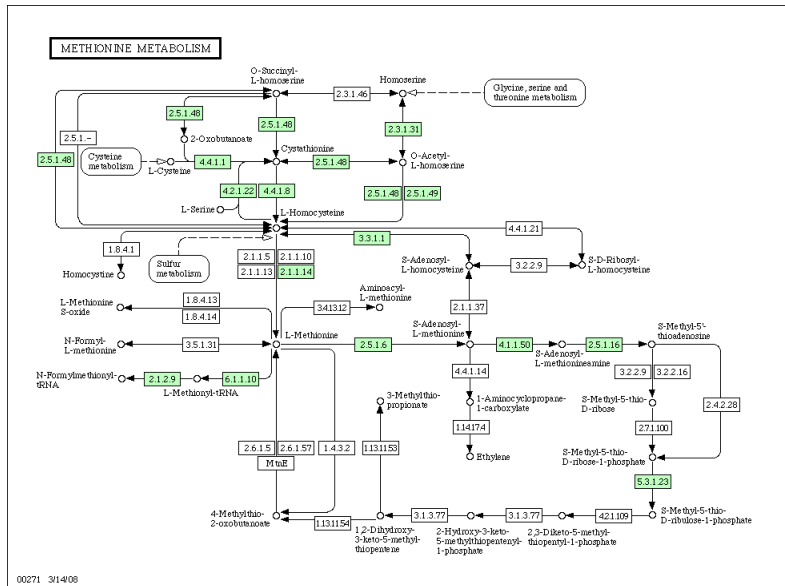
- dráhy jsou podsítě lineárního tvaru
  - sekvence metabolických reakcí
  - specifické zaměření na určité proměnné
  - analyzované problémy: délka dráhy, existence alternativních drah
- sítě reprezentují komplexní data (zohledňují širokou množinu proměnných a všech relevantních interakcí)
  - sítě interakcí určitého charakteru (transkripce, metabolismus, protein-protein, ...)
  - analyzované jevy: stupeň větvení, délka nejkratší dráhy, modularita, motivy, ...
- příklady zdrojů:
  - KEGG (<http://www.genome.jp/kegg/>)
  - RegulonDB (<http://regulondb.ccg.unam.mx/>)

# Vyhledávání alternativních drah – *E. coli*

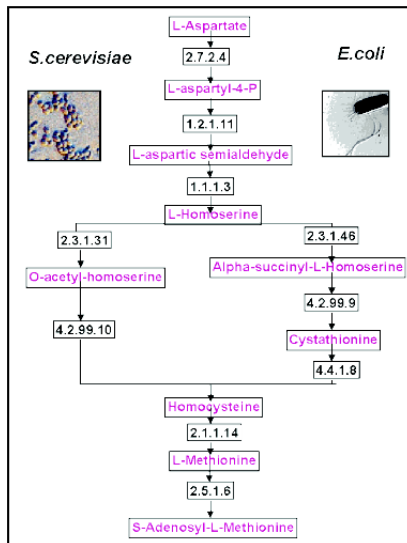




# Vyhledávání alternativních drah – *S. cerevisiae*



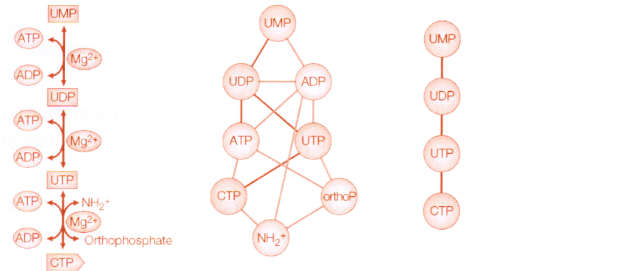
# Alternativní dráhy



## *Alternativní dráhy*

- význam v genetice a genomice
  - modifikace drah souvisí přímo s vývojem genomu (evoluce)
  - identifikace neznámých genů
- význam v biotechnologii
  - identifikace a implementace alternativních variant
- význam ve farmakologii
  - laterální náhrada genu
  - metabolicky-specifické medikamenty

# Biologické sítě



- různé typy sítí:
  - regulační síť (popis transkripční regulace)
  - proteinové síť (popis interakce proteinů)
  - metabolické síť (popis metabolismu)
  - signální síť (popis aktivačních/deaktivačních kaskád)
  - další typy (např. neuronové síť)

## Biologická síť jako graf

### Definition

Nechť  $V$  je konečná množina uzlů a  $E \subseteq V \times V$  relace.

*Biologickou sítí* nazveme graf  $G$  reprezentovaný uspořádanou dvojicí  $G \equiv (V, E)$ .

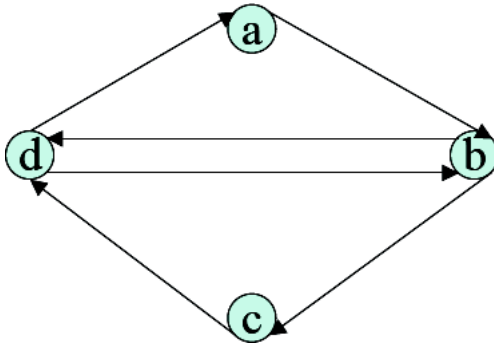
- Pokud  $\forall \langle a, b \rangle \in E. \langle a, b \rangle \in E \rightarrow \langle b, a \rangle \in E$ ,  $G$  nazýváme *neorientovaný*.
- V ostatních případech hovoříme o *orientovaném* grafu.

typ sítě	$V$	$E$	$G$
genová regulační	geny (resp. proteiny)	regulace exprese	or.
proteinová	proteiny	proteinové interakce	neor.
metabolická	metabolity, enzymy	enzymové reakce	or.
signální	molekuly	aktivace/deaktivace	or.

## *Cesty a kružnice*

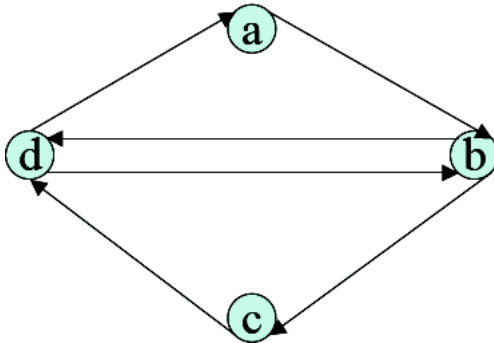
- *cesta* v grafu je libovolná sekvence uzlů  $[a_1, a_2, \dots, a_n]$  t.ž.  $\forall i \in \{1, \dots, n-1\}. \langle a_i, a_{i+1} \rangle \in E$ , číslo  $n-1$  nazýváme *délkou cesty* (počet hran)
- cestu nazveme *elementární* pokud se na ní každý vrchol vyskytuje právě jednou
- *kružnice* v grafu je libovolná elementární cesta  $[a_1, a_2, \dots, a_n]$  t.ž.  $a_1 = a_n$
- *smyčkou* nazýváme libovolnou kružnici délky 1

## *Cesty a kružnice*



- kolik kružnic ... 4
- kolik cest z *a* do *d* ... 2
- délka nejkratší cesty z *d* do *c* ...

## Cesty a kružnice



- kolik kružnic ... 4
- kolik cest z  $a$  do  $d$  ... 2
- délka nejkratší cesty z  $d$  do  $c$  ...  $d(d, c) = 2$



# Vlastnosti grafu

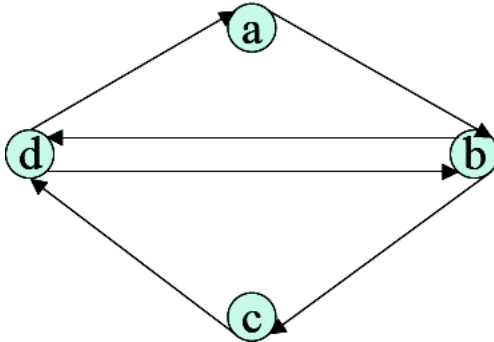
## Charakteristická délka cesty

- délku nejkratší cesty z  $a$  do  $b$  značíme  $d(a, b)$
- neexistuje-li cesta z  $a$  do  $b$ , uvažujeme  $d(a, b) = 0$
- *charakteristickou délku cesty* (orientovaného) grafu  $G \equiv (V, E)$  značíme  $L_G$  a definujeme:

$$L_G = \frac{\sum_{a,b \in V} d(a, b)}{|V|(|V| - 1)}$$

- průměrná délka cesty (přes všechny dvojice uzlů)

## *Cesty a kružnice*



$$L_G = \frac{18}{4 \cdot 3} = 1.5$$

# Vlastnosti grafu

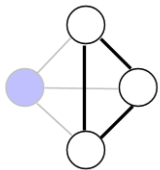
## Stupeň uzlu a koeficient seskupení

- množinu *sousedních uzlů* uzlu  $a$  značíme  $N_a$  a definujeme  $N_a = \{b \in V \mid \langle a, b \rangle \in E \vee \langle b, a \rangle \in E\}$
- *stupeň* uzlu  $a$  značíme  $k_a$  a definujeme jako počet všech *sousedních uzlů* uzlu  $a$ , tedy  $k_a = |N_a|$
- *koeficient seskupení uzlu* (clustering coefficient [Watts, Strogatz])  $a$  značíme  $C_a$  a definujeme:

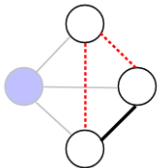
$$C_a = \frac{|\{\langle c, d \rangle \in E \mid c \in N_a \wedge d \in N_a\}|}{k_a(k_a - 1)}$$

# Vlastnosti grafu

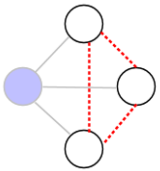
## Stupeň uzlu a koeficient seskupení



$$c = 1$$



$$c = 1/3$$

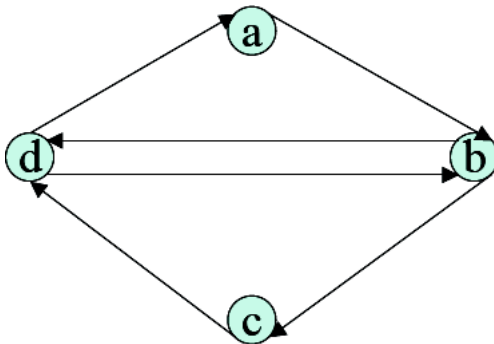


- *koeficient seskupení grafu  $G$  je značen  $C_G$  a definován jako průměr koeficientů seskupení všech uzlů:*

$$C_G = \frac{1}{|V|} \sum_{a \in V} C_a$$

# Vlastnosti grafu

*Stupeň uzlu a koeficient seskupení*



$$C_a = \frac{2}{2} = 1$$

$$C_b = \frac{2}{6} = \frac{1}{3}$$

$$C_G = \frac{1}{4} \cdot \left(2 + \frac{2}{3}\right) = 0.65$$

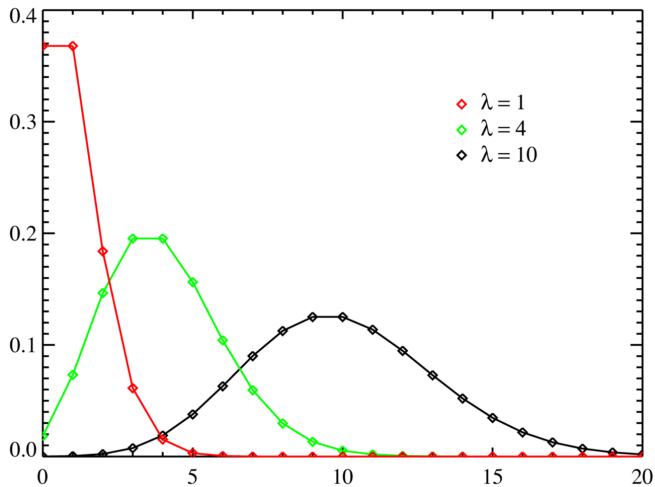
## Náhodný graf

- **náhodný graf** je definován pevným počtem uzlů a pravděpodobností  $p$  existence hrany mezi libovolnými dvěma uzly
- *alternativní definice*: zvolíme množinu vrcholů  $V$  a počet hran  $n$ , z množiny všech možných hran  $\binom{V}{2}$  vybereme náhodně  $n$  hran
- pravděpodobnost, že v náhodném grafu má daný uzel stupeň  $k$ , je charakterizována Poissonovým rozložením (s konst.  $\lambda$ ):

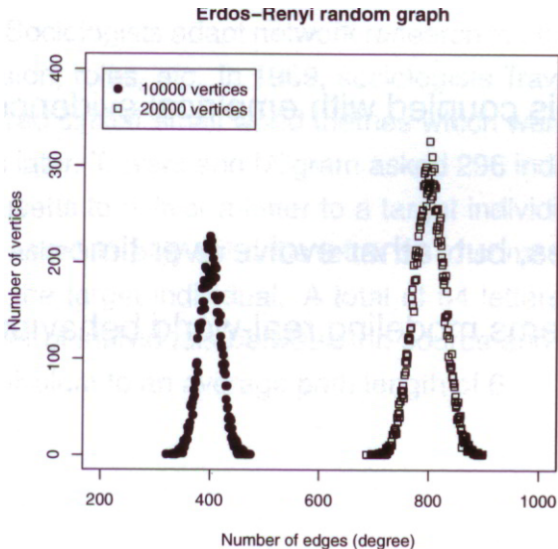
$$f(k|\lambda) = \frac{e^{-\lambda} \lambda^k}{k!}$$

[Erdős, Rényi, “On the evolution of random graphs”]

# Poissonovo rozložení

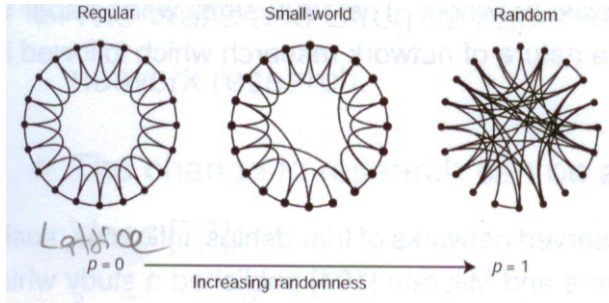


# Náhodný graf – Poissonovo rozložení stupně uzlů





# Vlastnosti náhodných grafů

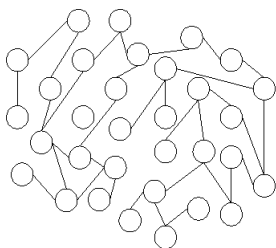


typ grafu	$C_G$	$L_G$
svaz	vysoké	dlouhé
náhodný graf	nízké	krátké
small-world	vysoké	krátké

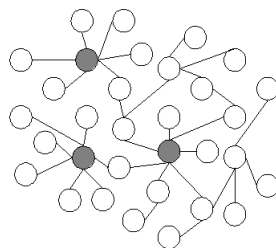
## *Small-world sítě*

- zavedeny Wattsem a Strogatzem, “Collective dynamics of ‘small-world’ networks”, Nature 393, 1998
- klíčem jsou lokální a globální metriky seskupení uzlů a metrika charakteristické délky cesty
- identifikovány jako grafy s vysokým koeficientem seskupení ale krátkou charakteristickou délkou cesty
- bylo prokázáno, že mnoho reálných sítí má tento charakter
  - např. graf filmových herců propojených dle společného účinkování
  - neuronové sítě v *C. elegans*
- výrazný posun v porozumění chování rozsáhlých dynamických systémů
- zavedení pojmu “real-world graphs”

## *Scale-free síť*



(a) Random network



(b) Scale-free network

- zavedl Barabási a Albert, "Emergence of Scaling in Random Networks", Science 286, 1999

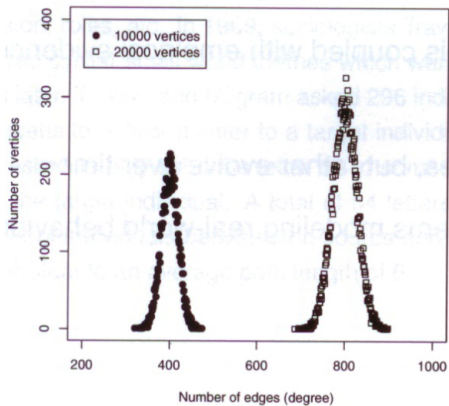
## *Scale-free sítě*

- reálné sítě nejsou statické (nemají pevný počet uzlů), ale vyvíjejí se dynamicky v čase, tzv. “rostou”
- nové uzly se napojují nejvíce k těm uzlům, které jsou se zbytkem sítě již dobře propojeny
- např. metabolické sítě *E. coli* jsou scale-free [Wagner, Fell, 2001]
- označíme-li  $P(k)$  pravděpodobnost, že libovolný uzel má stupeň  $k$ , pak pro scale-free sítě platí následující úměra (Mocninný zákon pro konst.  $\lambda$ ):

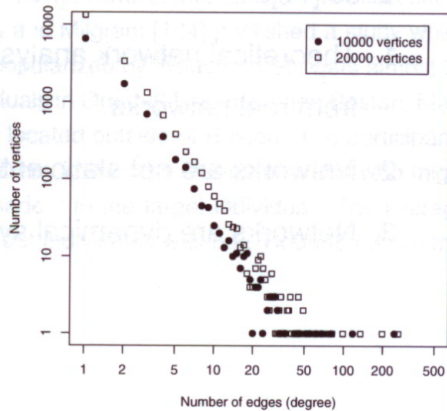
$$P(k) \sim k^{-\lambda}$$

# Scale-free síť

### Erdos-Renyi random graph



### Barabasi-Albert scale-free graph



## *Motivy ve scale-free sítích*

- ve scale-free sítích se vyskytují specifické uzly, tzv. huby – uzly s vysokým stupněm propojení na kostru síťové struktury
- ostatní uzly jsou lokálně napojeny k hubům
- objeveno např. při studiu proteinové sítě kvasinky pivovarské (*Saccharomyces cerevisiae*) [Jeong, Mason, 2001]
- díky hubům jsou sítě robustní proti náhodnému vyjmutí uzlu, ale naopak vyjmutí hubu znamená výrazné porušení sítě
- tato struktura vede k hierarchičnosti a modulárnímu charakteru
- jako moduly jsou identifikovány často opakující se výrazné podsítě (motivy) [Alon et.al., “Network Motifs: Simple Building Blocks of Complex Networks”, 2002]

http:

[//www.weizmann.ac.il/mcb/UriAlon/coliData.html](http://www.weizmann.ac.il/mcb/UriAlon/coliData.html)

# Motivy

Network	Nodes	Edges	$N_{real}$	$N_{rand} \pm SD$	Z score	$N_{real}$	$N_{rand} \pm SD$	Z score	$N_{real}$	$N_{rand} \pm SD$	Z score
<b>Gene regulation (transcription)</b>				<b>Feed-forward loop</b>			<b>Bi-fun</b>				
<i>E. coli</i>	424	519	40	7 ± 3	10	203	47 ± 12	13			
<i>S. cerevisiae*</i>	685	1,052	70	11 ± 4	14	1812	300 ± 40	41			
<b>Neurons</b>				<b>Feed-forward loop</b>			<b>Bi-fun</b>			<b>Bi-parallel</b>	
<i>C. elegans</i> †	252	509	125	90 ± 10	3.7	127	55 ± 13	5.3	227	35 ± 10	20
<b>Food webs</b>				<b>Three chain</b>			<b>Bi-parallel</b>				
Little Rock	92	984	3219	3120 ± 50	2.1	7295	2220 ± 210	25			
Ythan	83	391	1182	1020 ± 20	7.2	1357	230 ± 50	23			
St. Martin	42	205	469	450 ± 10	NS	382	130 ± 20	12			
Chesapeake	31	67	80	82 ± 4	NS	26	5 ± 2	8			
Coachella	29	243	279	235 ± 12	3.6	181	80 ± 20	5			
Skipwith	25	189	184	150 ± 7	5.5	397	80 ± 25	13			
B. Brook	25	104	181	130 ± 7	7.4	267	30 ± 7	32			
<b>Electronic circuits (forward logic chips)</b>				<b>Feed-forward loop</b>			<b>Bi-fun</b>			<b>Bi-parallel</b>	
s15850	10,383	14,240	424	2 ± 2	285	1040	1 ± 1	1200	480	2 ± 1	335
s38584	20,717	34,204	413	10 ± 3	120	1739	6 ± 2	800	711	9 ± 2	320
s38417	23,843	33,661	612	3 ± 2	400	2404	1 ± 1	2550	531	2 ± 2	340
s9234	5,844	8,197	211	2 ± 1	140	754	1 ± 1	1050	209	1 ± 1	200
s13207	8,651	11,831	403	2 ± 1	225	4445	1 ± 1	4950	264	2 ± 1	200
<b>Electronic circuits (digital fractional multipliers)</b>				<b>Three-node feedback loop</b>			<b>Bi-fun</b>			<b>Four-node feedback loop</b>	
s208	122	189	10	1 ± 1	9	4	1 ± 1	3.8	5	1 ± 1	5
s420	252	399	20	1 ± 1	18	10	1 ± 1	10	11	1 ± 1	11
s838‡	512	819	40	1 ± 1	38	22	1 ± 1	20	23	1 ± 1	25
<b>World Wide Web</b>				<b>Feedback with two mutual dyads</b>			<b>Fully connected triad</b>			<b>Uplinked mutual dyad</b>	
nd.edu§	325,729	1.46e6	1.1e5	2e3 ± 1e2	800	6.8e6	5e4 ± 4e2	15,000	1.2e6	1e4 ± 2e2	5000

## *Predikce síťových motivů*

- *problém: jak výrazně je podgraf v dané reálné síti zastoupen?*



## *Predikce síťových motivů*

- *problém: jak výrazně je podgraf v dané reálné síti zastoupen?*
- *smysl: je toto zastoupení statisticky významné?*

## *Predikce síťových motivů*

- *problém: jak výrazně je podgraf v dané reálné síti zastoupen?*
- *smysl: je toto zastoupení statisticky významné?*
- *řešení: porovnání reálné sítě s dostatečným množstvím náhodných sítí náležících do vhodné reprezentativní třídy vzhledem k reálné síti*

## *Predikce síťových motivů – ER model*

- počet uzlů i hran stejný jako v reálné síti
- hrany náhodně rozmístěny mezi uzly
- mějme orientovaný graf  $G = (V, E)$
- počet všech možných dvojic uzlů pro umístění (orientované) hrany:

$$|V|(|V| - 1)$$

## *Predikce síťových motivů – ER model*

- počet uzlů i hran stejný jako v reálné síti
- hrany náhodně rozmístěny mezi uzly
- mějme orientovaný graf  $G = (V, E)$
- počet všech možných dvojic uzlů pro umístění (orientované) hrany:

$$|V|(|V| - 1)$$

- hrana může být smyčka

## *Predikce síťových motivů – ER model*

- počet uzlů i hran stejný jako v reálné síti
- hrany náhodně rozmístěny mezi uzly
- mějme orientovaný graf  $G = (V, E)$
- počet všech možných dvojic uzlů pro umístění (orientované) hrany:

$$|V|(|V| - 1)$$

- hrana může být smyčka  $\Rightarrow$  máme navíc  $|V|$  možností

## *Predikce síťových motivů – ER model*

- počet uzlů i hran stejný jako v reálné síti
- hrany náhodně rozmístěny mezi uzly
- mějme orientovaný graf  $G = (V, E)$
- počet všech možných dvojic uzlů pro umístění (orientované) hrany:

$$|V|(|V| - 1)$$

- hrana může být smyčka  $\Rightarrow$  máme navíc  $|V|$  možností
- celkem tedy dostáváme pro výběr dvojic uzlů v orientovaném grafu:

$$|V|(|V| - 1) + |V|$$

## *Predikce síťových motivů – ER model*

- počet uzlů i hran stejný jako v reálné síti
- hrany náhodně rozmístěny mezi uzly
- mějme orientovaný graf  $G = (V, E)$
- počet všech možných dvojic uzlů pro umístění (orientované) hrany:

$$|V|(|V| - 1)$$

- hrana může být smyčka  $\Rightarrow$  máme navíc  $|V|$  možností
- celkem tedy dostáváme pro výběr dvojic uzlů v orientovaném grafu:

$$|V|(|V| - 1) + |V| = |V|^2$$

## *Predikce autoregulačního motivu*

- pravděpodobnost existence (orientované) hrany mezi dvěma uzly:

$$p = \frac{|E|}{|V|^2}$$

- pravděpodobnost existence smyčky:

$$p_{self} = \frac{|V|}{|V|^2}$$



## *Predikce autoregulačního motivu*

- pravděpodobnost existence (orientované) hrany mezi dvěma uzly:

$$p = \frac{|E|}{|V|^2}$$

- pravděpodobnost existence smyčky:

$$p_{self} = \frac{|V|}{|V|^2} = \frac{1}{|V|}$$

## *Predikce autoregulačního motivu*

- pravděpodobnost existence (orientované) hrany mezi dvěma uzly:

$$p = \frac{|E|}{|V|^2}$$

- pravděpodobnost existence smyčky:

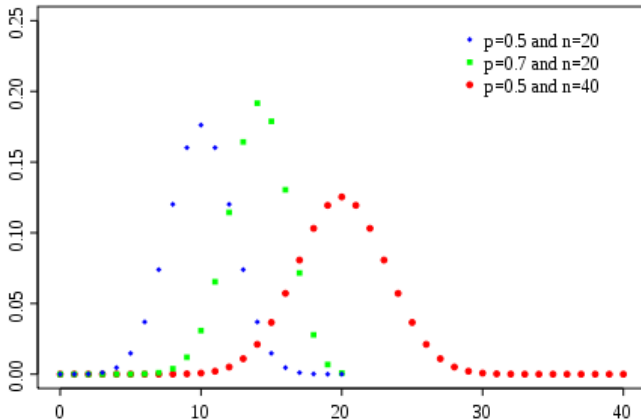
$$p_{self} = \frac{|V|}{|V|^2} = \frac{1}{|V|}$$

- pravděpodobnost existence právě  $k$  smyček lze vyjádřit binomicky:

$$P(k) = \binom{E}{k} p_{self}^k (1 - p_{self})^{E-k}$$

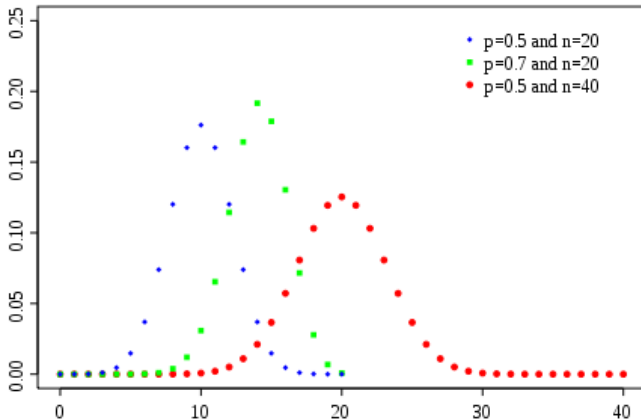
# *Binomické rozdělení*

$$P(K = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$



## Binomické rozdělení

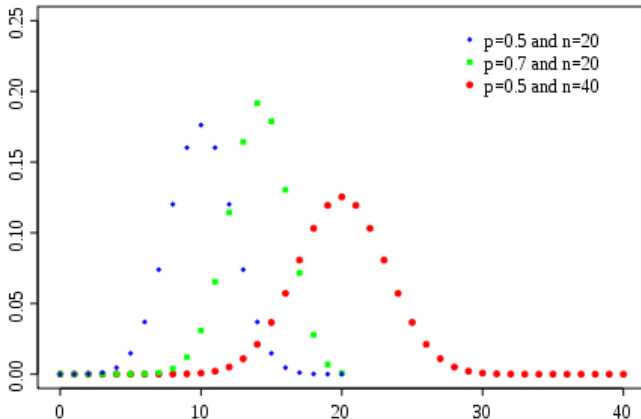
$$P(K = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$



- střední hodnota:  $m(k) = np$
- rozptyl:  $v(k) = np(1 - p)$

## Binomické rozdělení

$$P(K = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$



- střední hodnota:  $m(k) = np$
- rozptyl:  $v(k) = np(1 - p) \Rightarrow \varrho(k) = \sqrt{v(k)} = \sqrt{np(1 - p)}$

## *Predikce autoregulačního motivu*

- průměrný počet smyček v ER grafu  $G = (V, E)$ :

$$|e_{self}|_{ER} \sim |E|p_{self}$$

## *Predikce autoregulačního motivu*

- průměrný počet smyček v ER grafu  $G = (V, E)$ :

$$|e_{self}|_{ER} \sim |E|p_{self} \sim \frac{|E|}{|V|}$$

## *Predikce autoregulačního motivu*

- průměrný počet smyček v ER grafu  $G = (V, E)$ :

$$|e_{self}|_{ER} \sim |E| p_{self} \sim \frac{|E|}{|V|}$$

- standardní odchylka  $\varrho_{self ER}$ :

$$\varrho_{self ER} \sim \sqrt{\frac{|E|}{|V|}}$$

- např. v trnsc. síti E. coli máme  $|E| = 520$ ,  $|V| = 420$  a tedy pro náhodné grafy ER modelu dostáváme následující charakteristiku:

$$|e_{self}|_{ER} \sim 1.2 \quad \varrho_{self ER} \sim 1.1$$



## *Z-skóre motivu*

- Z-skóre kvantizuje statistickou signifikanci jevu
- dáno počtem standardních odchylek které odlišují reálnou síť od třídy náhodných grafů ER modelu

$$Z = \frac{|E_{self}|_{real} - |e_{self}|_{ER}}{\sigma_{self ER}}$$

## *Z-skóre motivu*

- Z-skóre kvantizuje statistickou signifikanci jevu
- dáno počtem standardních odchylek které odlišují reálnou síť od třídy náhodných grafů ER modelu

$$Z = \frac{|E_{self}|_{real} - |e_{self}|_{ER}}{\sigma_{self ER}}$$

- pro autoregulační motiv máme v síti E. coli 40 smyček, a tedy Z-skóre autoregulačního motivu v této síti je

$$Z = \frac{40 - 1.2}{1.1}$$

## *Z-skóre motivu*

- Z-skóre kvantizuje statistickou signifikanci jevu
- dáno počtem standardních odchylek které odlišují reálnou síť od třídy náhodných grafů ER modelu

$$Z = \frac{|E_{self}|_{real} - |e_{self}|_{ER}}{\sigma_{self ER}}$$

- pro autoregulační motiv máme v síti E. coli 40 smyček, a tedy Z-skóre autoregulačního motivu v této síti je

$$Z = \frac{40 - 1.2}{1.1} \sim 35$$

- to prokazuje signifikantní zastoupení tohoto podgrafu v reálné síti E. coli
- typicky považujeme za signifikantní  $Z > 2$

## Predikce víceuzlových motivů

- uvažujme podgraf  $S_G = (V_S, E_S)$  grafu  $G = (V, E)$  t.ž.  $V_S \subseteq V$ ,  $E_S \subseteq E$  a zaveďme značení  $v_S = |V_S|$  a  $e_S = |E_S|$
- předpokládejme  $v_S > 1$
- **problém:** *kolik je průměrně výskytů podgrafu  $S_G$  v náhodných sítích ER modelu vzhledem k  $G$  (až na izomorfismus)?*
- vybíráme  $v_S$  uzlů:  $|V| \cdot (|V| - 1) \cdots (|V| - v_S + 1) \sim |V|^{v_S}$
- mezi něž umístíme  $e_S$  hran:  $p^{e_S}$
- předpokládejme  $S_G$  t.ž. existuje  $\alpha$  izomorfních variant
- **řešení:** *průměrný výskyt podgrafu  $S_G$  v ER lze aproximovat:*

$$o(S_G, G) \sim \frac{1}{\alpha} |V|^{v_S} p^{e_S}$$

## *Predikce víceuzlových motivů – nástroje*

- NetMatch – <http://baderlab.org/Software/NetMatch>
  - plugin aplikace Cytoscape
  - verifikuje zda zadaný podgraf je motivem v dané síti
  - detekuje instance daného motivu přímo v grafu
  - reflektuje různé typy hran a uzlů
  - simuluje na základě randomizace grafu (Barabasi-Albert)

## *Predikce víceuzlových motivů – nástroje*





- mFinder – <http://www.weizmann.ac.il/mcb/UriAlon/groupNetworkMotifSW.html>
  - umožňuje plnou enumeraci i samplování
  - vizualizace grafů pomocí mDraw

Kashtan, N., et al., Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs. *Bioinformatics*, 2004. 20(11): p. 1746-58.
- FANMOD – <http://theinf1.informatik.uni-jena.de/~wernicke/motifs/index.html>
  - umožňuje plnou enumeraci i samplování
  - neumožňuje vizualizaci, vytváří HTML report





S. Wernicke and F. Rasche. FANMOD: a tool for fast network motif detection. *Bioinformatics*, 22(9):1152–1153, 2006.
- MAVisto – <http://mavisto.ipk-gatersleben.de/>
  - umožňuje vizualizaci výsledků
  - obsahuje editor grafů

Schreiber, F. and Schwöbbermeyer H.: MAVisto: a tool for the exploration of network motifs. *Bioinformatics*, 21, 3572-3574, 2005.

## *Základní literatura*

-  Alon, U. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall, 2006.
-  Kitano, H. *Looking beyond the details: a rise in system-oriented approaches in genetics and molecular biology*. Curr Genet., 2002.
-  Ellner, S.P. and Guckenheimer, J. *Dynamical Models in Biology*. Princeton University Press, 2006.
-  Bolouri, H. *Computational Modeling of Gene Regulatory Networks – a Primer*. Imperial College Press, 2008.

## *Doplňující literatura*

-  Palsson, B. *Systems Biology: Properties of Reconstructed Networks*. Cambridge University Press, 2006.
-  de Vries, G. et al. *A Course in Mathematical Biology: Quantitative Modeling with Mathematical and Computational Methods*. S.I.A.M., 2006.
-  Edelstein-Keshet, L. *Mathematical Models in Biology*. S.I.A.M., 2005.
-  Wilkinson, D.J. *Stochastic Modelling for Systems Biology*. Chapman & Hall/CRC Mathematical & Computational Biology, 2006.