

Úvod do analýzy sociálních sítí

Dynamické sítě

2023-24

Vývoj sítě v čase

- Dynamic, evolving, temporal networks?
- Jde o sítě, jejichž stav se v čase mění a o procesy, které se v čase realizují.
- V nejjednodušším případě se předpokládají tzv. rostoucí sítě (growing networks), kdy v čase přibývají vrcholy i hrany.
- Jak „evolving“ tak „temporal“ sítě se dají chápat jako dynamické. Liší s v úlohách, které se řeší a které mají jako jeden z parametrů **CAS**.

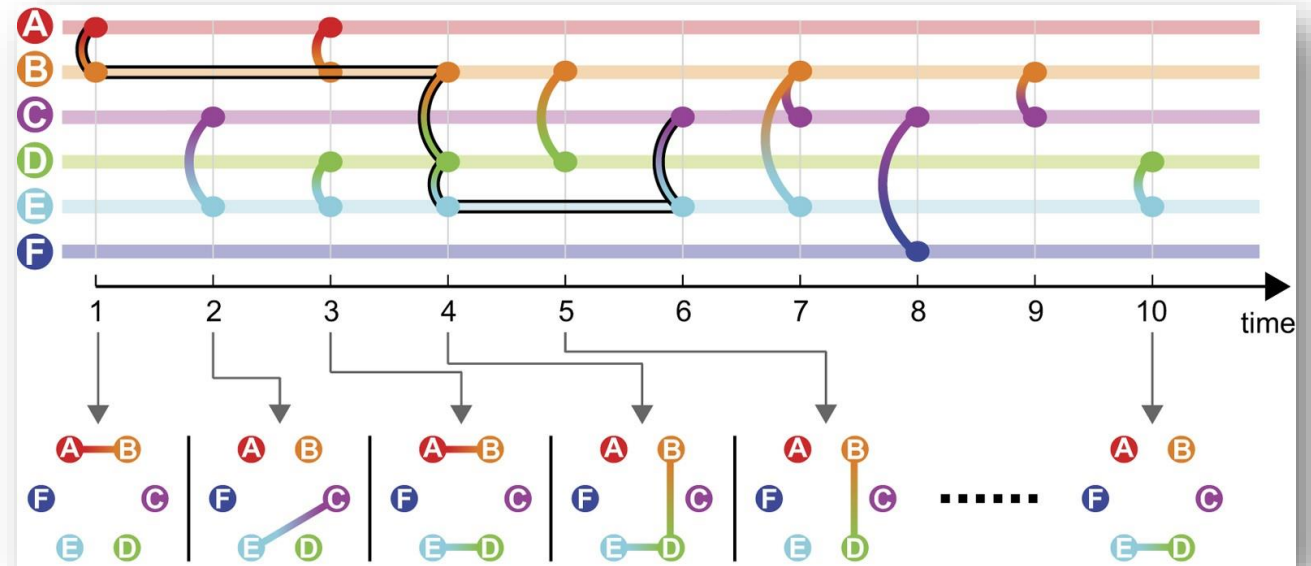
Dynamické sítě

- Základní charakteristikou je stochastický přístup k analýze.
- Pracuje se pravděpodobností existence (nebo mírou aktivity) vrcholu resp. hrany v daném časovém okamžiku.
- Jevy se často v dynamických sítích modelují pomocí simulace.
- Často se také předpokládá, že vrcholy sítě mohou být různého typu (heterogenní síť).

Evolving (vyvíjející se) sítě

- Sítě, ve kterých v čase přibývají vrcholy a hrany.
 - V některých případech mohou jak vrcholy, tak hrany ubývat („úmrtí“, „stárnutí“).
 - Obvykle se pracuje s neváženými sítěmi.
- Existují poměrně jednoduché modely popisující principy růstu a umožňující simulace (např. Barabasi-Albert model).
 - Modely jsou zaměřené např. na růst webu (zaměřené na velká centra).
- K přečtení: <http://networksciencebook.com/chapter/6>.

Temporální síť 1/2



- Také se označují jako „time-varying networks“.
- Jsou charakteristické tím, že hrany existují jen v jistém čase a mohou mít měnící se váhu. To prakticky znamená, že hrany mohou existovat „přerušovaně“, tedy v určitých časových intervalech v síti jsou a v jiných ne.
- K přečtení: <https://arxiv.org/pdf/1108.1780.pdf>

Temporální sítě 2/2

- S tím, jak se temporální síť v čase mění, souvisí i další vlastnosti.
 - Hustota lokálního propojení.
 - Stupně či jiné centrality vrcholů.
 - apod.
- Obvykle se tyto sítě analyzují tak, že se vybere určitý časový úsek a v něm se temporální síť transformuje na obyčejnou.

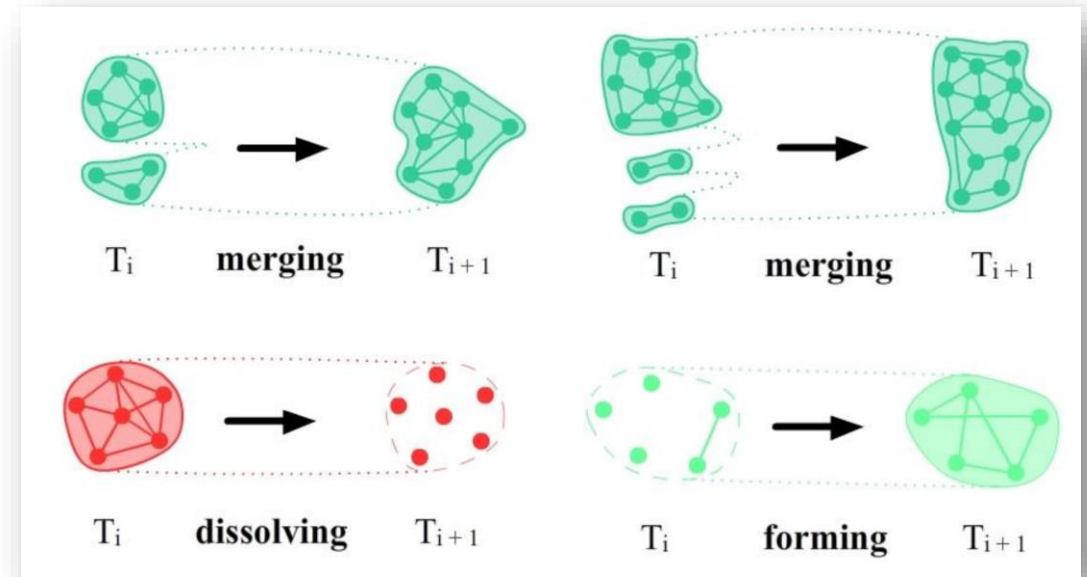
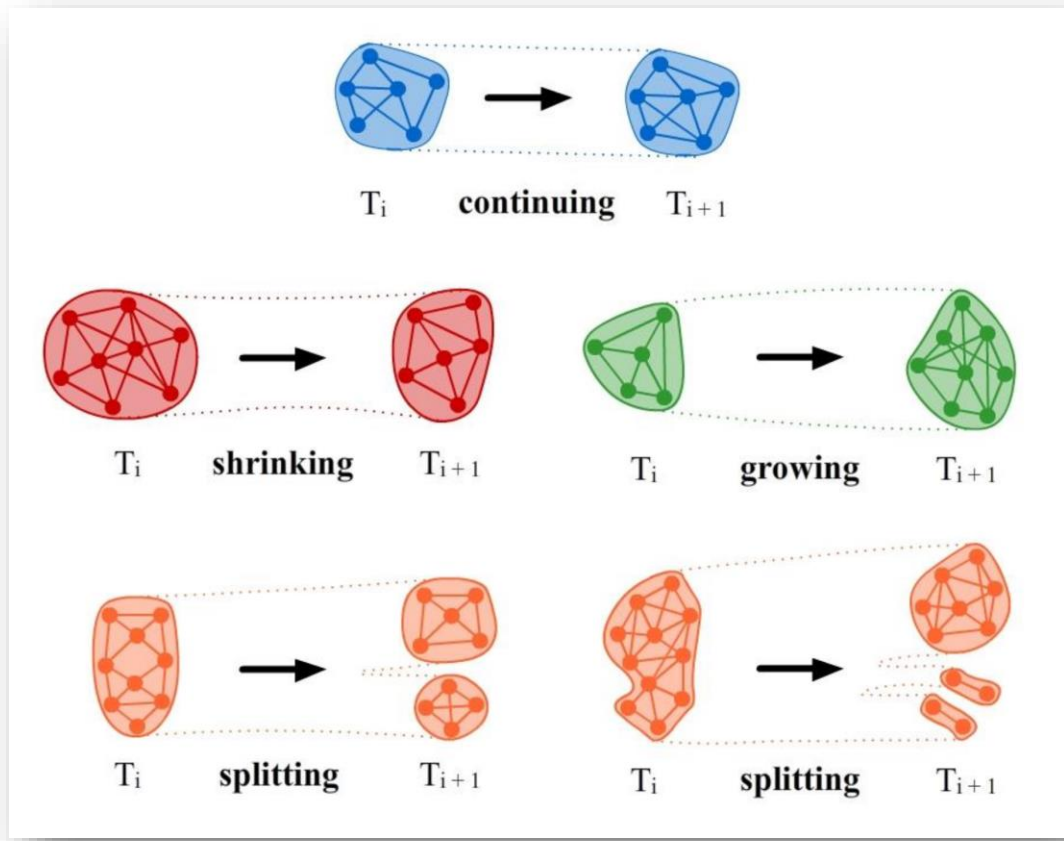
Příklady dynamických sítí

- Emailové sítě
- Telefonní sítě
- Technické sítě
- Transportní sítě
- Obecně všechny komunikační sítě...

Příklad – spoluautorské sítě

- Autoři článků v určité doméně (např. v informatice) publikují v různých týmech či komunitách a s různou mírou pravidelnosti.
- Mohou publikovat v několika týmech současně.
- Týmy mění svou strukturu, mohou zanikat a vznikat nové, např. podle toho, jakým tématům se věnují.
- Vazby mezi některými autory se posilují opakovaným publikováním, mohou však také slábnout a zanikat.

V čase se mění i komunity



Sítě pro analýzu

- email-dnc
 - <http://networkrepository.com/email-dnc.php>
- ia-reality-call
 - <http://networkrepository.com/ia-reality-call.php>
 - <http://realitycommons.media.mit.edu/realitymining.html>
- college-msg
 - <https://snap.stanford.edu/data/CollegeMsg.html>

Předzpracování sítí

- Pro každou síť jsme vytvořili jsme 5 snímků v čase, každý z nich odpovídá připojení 20% nových vrcholů do sítě.
 - Hrany jsou v souborech uspořádány podle času, kdy na těchto hranách proběhla interakce.
 - Každý snímek může obsahovat „opakované“ hrany. Důvodem je, že hrana zde reprezentuje jednu interakci. Součet opakování hrany lze chápat jako váhu hrany.
- Všechny sítě jsou rostoucí (vrcholy a hrany pouze přibývají).

email-dnc

- Jde o orientovanou e-mailovou síť vytvořenou z uniklých mailů Demokratické strany v USA v roce 2016. Vrcholy sítě odpovídají osobám (v rolích odesílatelů i příjemců), hrany odeslaným mailům.
- Síť jsme převedli na neorientovanou.

ia-reality-call

- Data vznikla z hovorů z mobilních telefonů mezi malou skupinou uživatelů na Massachusetts Institute of Technology (MIT), kterým byly přiděleny mobilní telefony a jejich hovory byly monitorovány.
- Data také obsahují volání s uživateli mimo tuto malou skupinu uživatelů do jiných telefonů jednotlivců, a proto jejich vrcholy mají méně hran než ostatní vrcholy.
- Data byla shromážděna za dobu 9 měsíců v rámci experimentu Reality Mining provedeného v roce 2004 v rámci projektu Reality Commons.

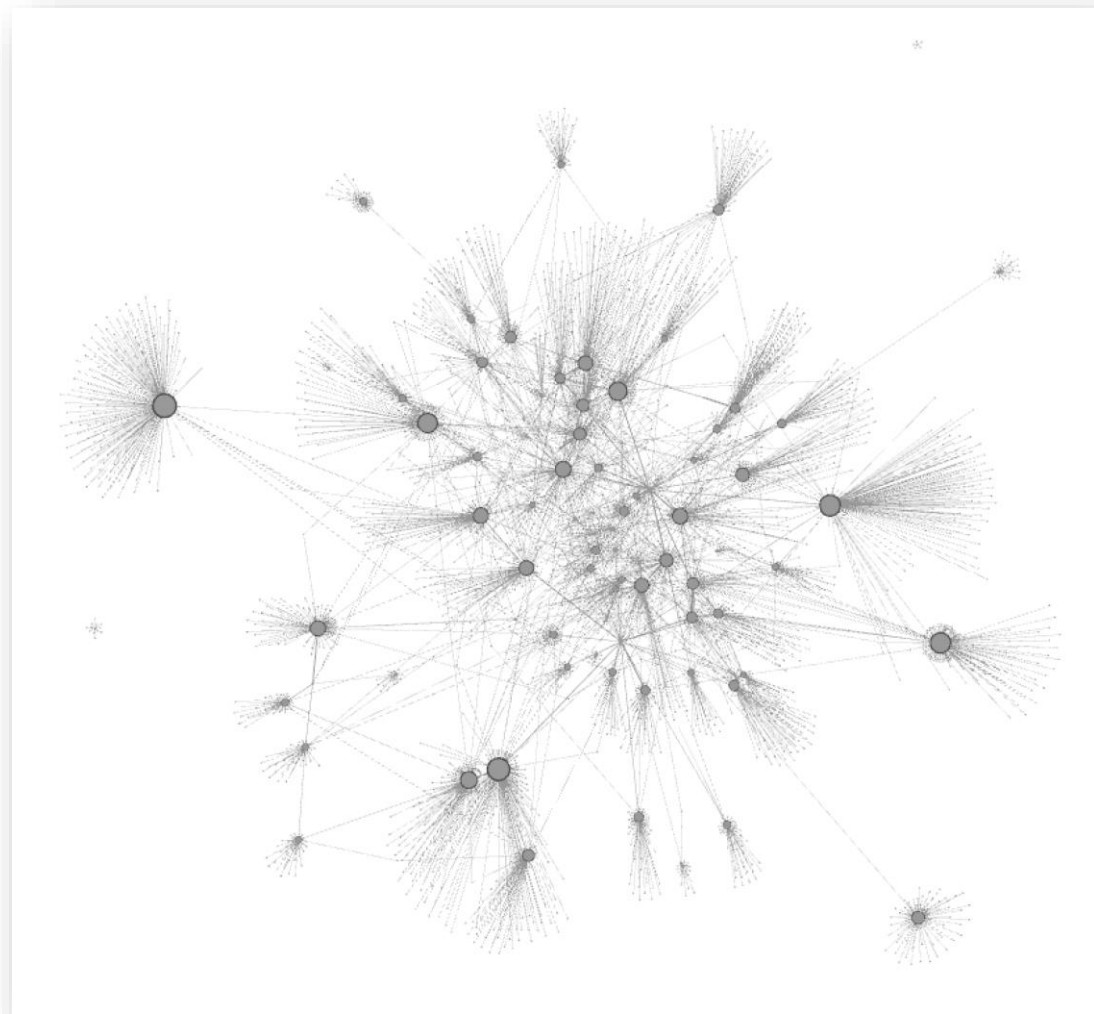
college-msg

- Základem je datová sada skládající se ze soukromých zpráv zasílaných na online sociální síti na Kalifornské univerzitě v Irvine.
- Hrany mezi uživateli s podobným profilem jsou výsledkem konverzací mezi nimi.

Úkoly

- Napište program (C++, C#, Java, Python), který ze zdrojových dat vytvoří definovaný počet časových snímků (sítí měnících se v čase).
- Zjistěte, jak se v čase (v jednotlivých snímcích) mění průměrný stupeň a průměrný vážený stupeň sítě.
- Zjistěte, jak se v čase mění počet komunit a souvislých komponent, a jak se mění průměrná a maximální velikost komunity.
- Vizualizujte jednotlivé sítě a popřemýšlejte, jaké typické vzorce se v sítích vyskytují a v jakém množství (kliky a hvězdy, malé souvislé komponenty a outliers apod.).
- Ve vizualizaci najděte nějaké vrcholy s vysokým stupněm resp. s vysokým váženým stupněm a popřemýšlejte, proč se v některých případech nejedná o tytéž vrcholy a zkuste zformulovat, proč tomu tak je.
- V alespoň jedné síti vyberte alespoň jeden vrchol, který má na posledním snímku vysoký stupeň resp. vážený stupeň a na něm ukažte, jak se vyvíjí čase (postupné změny ve stupni resp. váženém stupni, velikost komunity popř. komponenty, do které patří).

Stupeň



Vážený stupeň

