Mining temporal networks

Lecture in course "Social networks and online markets"

Sapienza, Wednesday, April 10, 2024

Aristides Gionis, KTH Royal Institute of Technology, Sweden

argioni@kth.se

agenda

part I : introduction and motivation

part II : models of temporal networks

part III : algorithmic frameworks

part IV : data mining problems

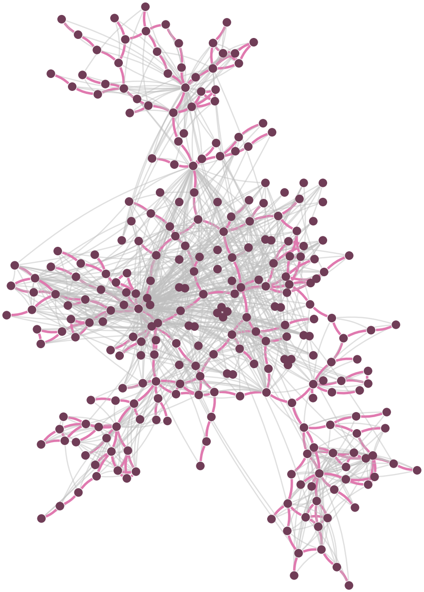
part V : conclusions and future challenges

part I

introduction and motivation

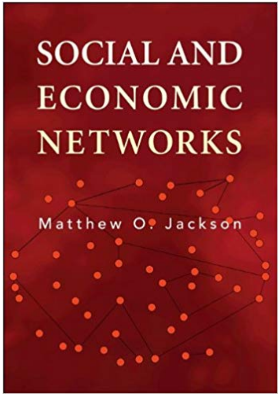
interconnected world

- networks model objects and their relations
- many different network types
  - social
  - informational
  - technological
  - biological
  - ...



impact of network science

- online communication networks and social media
- implications in
  - knowledge creation
  - information sharing
  - education
  - democracy
  - society as a whole



research questions in network science

- structure discovery
  - communities, summarization, events, role mining
- study complex dynamic phenomena
  - evolution, information diffusion, opinion formation, structural prediction
- develop novel applications
- design efficient algorithms

traditional view

- networks represented as pure graph-theory objects

  no additional vertex / edge information

- emphasis on static networks
- dynamic settings model structural changes

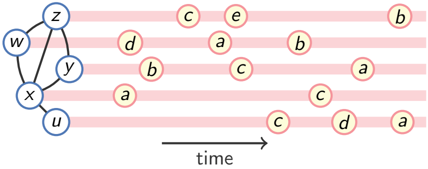
  vertex / edge additions / deletions

temporal networks

- ability to collect and store large volumes of network data
- available data have fine granularity
- lots of additional information associated to vertices/edges
- network topology is relatively stable, while lots of activity and interaction is taking place
- giving rise to new concepts, new problems, and new computational challenges

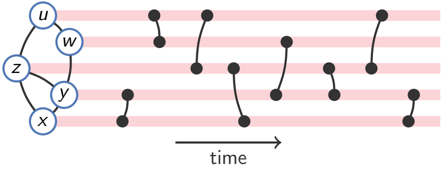
modeling activity in networks

1. network nodes perform actions (e.g., posting messages)

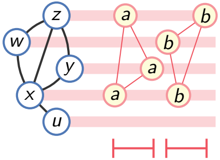


2. network nodes interact with each other

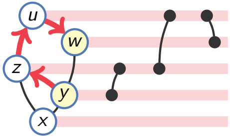
   (e.g., a "like", a repost, or sending a message to each other)



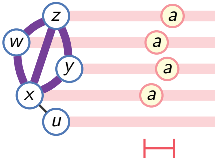
many novel and interesting concepts



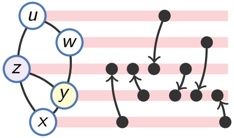
new pattern types



temporal information paths



new types of events



network evolution

temporal networks — objectives

- identify new concepts and new problems
- develop algorithmic solutions
- demonstrate relevance to real-world applications

terminology

- we use term "temporal networks", but terminology is not standardized
- term "X Y" can be encountered in the literature, where

X :	Y :
temporal	networks
dynamic	graphs
(time-)evolving	
time-varying	
time-dependent	
evolutionary	

examples of temporal networks

[Holme, 2015]

- human communication networks
  - phone, email, text messages, etc.
- human proximity networks
  - recorded by various sensors and devices, e.g., bluetooth, wifi, etc.
  - patient-referral networks, i.e., how patients are transferred between wards of a hospital system
  - sexual contact networks
- animal proximity networks
  - obtained via RFID devices
  - lifestock or wildlife

examples of temporal networks — cont'd

[Holme, 2015]

- bibliographic networks
  - collaboration and citation networks
- economic networks
  - credit card transactions
  - trade networks of countries
  - bitcoin transcations
- travel and transportation networks
  - airline connections, bus transport, bike-sharing systems

examples of temporal networks — cont'd

[Holme, 2015]

- brain networks
  - temporal correlations of the oxygen levels of brain regions as measured by fMRI scanning
- biological networks
  - genes involved in different interactions that change over time
  - current challenges, as one cannot measure precisely when two proteins interact with each other, but technology is improving

agenda

part I : introduction and motivation

part II : models of temporal networks

part III : algorithmic frameworks

part IV : data mining problems

part V : conclusions and future challenges

part II

models of temporal networks

representation of temporal networks

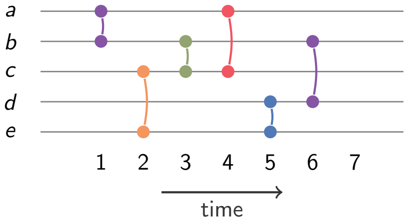
1. sequence of interactions

- a temporal network is represented as $G = (V, E)$
  - with set of nodes V, and set of edges $E = \{(u, v, t)\}$, with $u, v \in V$ and $t \in \mathbb{R}$
  - if interactions have duration, then $E = \{(u, v, t, \lambda)\}$
- this is a lossless representation — no information is lost
- also known as sequence of contacts, or sequence of (temporal) edges

representation of temporal networks

1. sequence of interactions

- visual representation of a temporal network as a sequence of interactions



representation of temporal networks

2. sequence of static graphs

- sequence $G_1, , \ldots, G_T$

  where $G_t = (V_t, E_t)$, with $t = 1, \ldots, T$

  typically assume that nodes are fixed, i.e., $V_t = V$

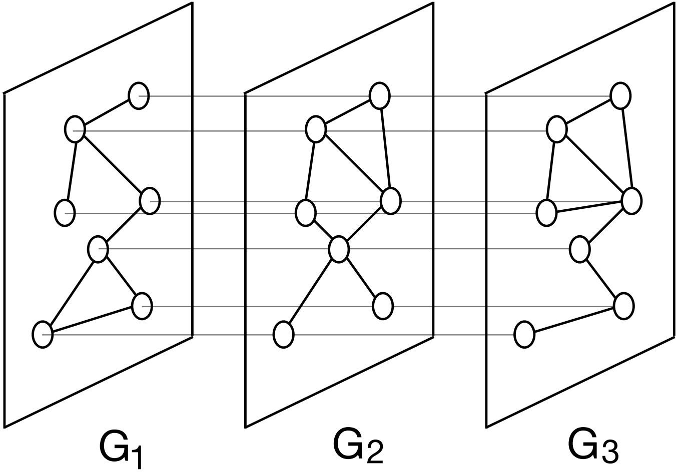
  E_t are the edges that occur in time interval t

- advantages: static graph analysis methods can be applied
- disadvantages: the representation assumes quantization into time intervals
  - thus, representation depends on quantization parameters, e.g., seconds, minutes, hours, days, etc.
  - coarse resolution may lead to information loss
  - fine resolution may lead to sparse (or even empty) static graphs

representation of temporal networks

2. sequence of static graphs

- visual representation of a temporal network as a sequence of static graphs



representation of temporal networks

3. time series of contacts

- a time-series for each pair of nodes in the network
- equivalent representation with sequence of interactions

4. tensor representation

- tensor representing $\text{node} \times \text{node} \times \text{time}$ information
- can apply powerful tensor-algebra techniques
- a complication is that time is directed, while tensor algebra assumes that indices can be relabeled (breaking the time ordering)

representation of temporal networks

[Casteigts et al., 2012]

5. time-varying graphs defined as $G = (V, E, T, p, \lambda)$, where
   - V : set of nodes
   - $E \subseteq V \times V$: set of edges
   - T : a time domain
   - $p : E \times T \to \{0, 1\}$: a presence function
   - $\lambda : E \times T \to \mathbb{R}$: a latency function

- general definition that can be used to model graph datasets in different applications
  - transportation networks, communication networks, social networks

temporal networks vs. dynamic graphs

- dynamic graphs is a standard model typically studied in theoretical computer science
  - e.g., [Henzinger et al., 1999, Thorup, 2000]
- dynamic graphs are represented as a sequence of edge additions and/or edge deletions
- G_0 is the initial graph, and G_i is the graph resulting after the i-th edge addition/deletion operation
- objective: efficient maintenance of graph properties
  - e.g., connectivity, shortest paths, spanners, etc.

temporal networks vs. dynamic graphs

- in studies of dynamic graphs, the properties of interest refer to individual graph snapshots G_i, not considering the whole graph evolution
- emphasis on computational efficiency
  - computation time per operation
  - e.g., cost of maintaining a minimum spanning tree per edge additions/deletions
  - or, cost of maintaining a data structure that allows to answer short-path queries
- dynamic graph model captures topological changes, not interactions
  - e.g., dynamic graphs can be used to model friendship additions/deletions in a social network, but not discussions or other interactions

temporal networks vs. dynamic graphs

- dynamic graphs resemble sequence of interactions model
- main difference lies on which graph properties we study
- for dynamic graphs we typically consider properties on graph snapshots
  - i.e., minimum spanning tree on the current snapshot
- for temporal graphs we typically consider properties that span a time interval
  - i.e., a temporal pattern
- note : here we focus more on temporal networks, not so much on dynamic graphs

graph streams

- setting inspired by data streams [Muthukrishnan et al., 2005]
- recall the data-stream model:
  - data are presented as a sequence of data items (potentially infinite)
  - assume a small number of passes typically constant or just one pass
  - assume small memory compared to data size, e.g., poly-logarithmic
  - assume fast computation per data item processed, e.g., constant or poly-logarithmic

graph streams

- a graph stream is a graph dataset in the data-stream model
  - it can be either a sequence of interactions (temporal network)
    or sequence of edge additions/deletions (dynamic graph)
- thus, a graph stream is not a representation model, instead it refers to the underlying computational model
- thus, we can study questions of mining temporal networks in the graph-stream model

dynamic graph algorithms on streaming model

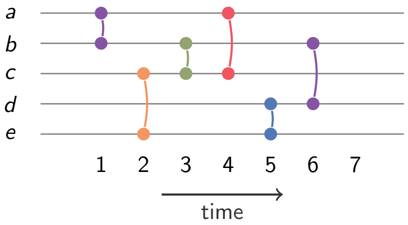
- well-studied model
- extensive survey by [McGregor, 2014]
- different settings considered
  - node/edge additions (incremental)
  - node/edge additions/deletions (fully-dynamic)
  - updating weights/labels is a special case of the fully-dynamic model
  - sliding-window setting: consider only edges from latest interval of fixed length
  - algorithms can be deterministic or randomized

time-respecting paths

time-respecting paths

- a fundamental concept in analysis of temporal networks
  - used in studies of information propagation, or epidemics spreading
- a time-respecting path is a sequence of temporal edges, such that
  - consecutive edges share a common node, and
  - time stamps of temporal edges are non-decreasing
- intuitively, a piece of information (or disease) can propagate in the network only over time-respecting paths

time-respecting paths — example



$(c, e, 2), (e, d, 5), (d, b, 6)$ is a time-respecting path from e to b

$(c, b, 3), (b, a, 1)$ is not a time-respecting path

static expansion of a temporal network

- a transformation of a temporal network to a directed (static) network so that time-respecting paths in the temporal network correspond to directed (static) paths in the directed (static) network

- how to create such a transformation?

static expansion of a temporal network

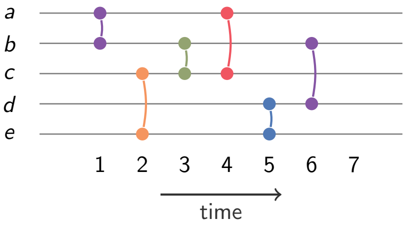
- a transformation of a temporal network to a directed (static) network so that time-respecting paths in the temporal network correspond to directed (static) paths in the directed (static) network

- how to create such a transformation?

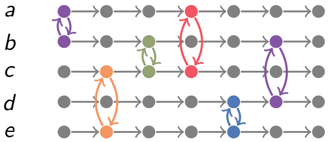
1. create a copy of each node for each time instance
2. create a directed edge from the $(t-1)$-th copy of u to the t-th copy of u,
   for all nodes u, and all time instances t
3. create directed edges for the temporal edges

static expansion of a temporal network

example



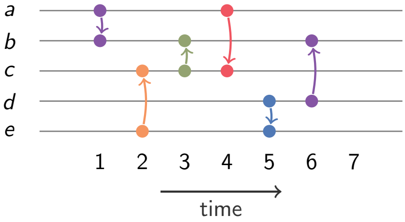
(a) representation of a temporal network



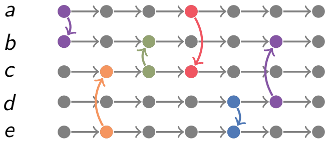
(b) static expansion of temporal network;

static expansion of a temporal network

example



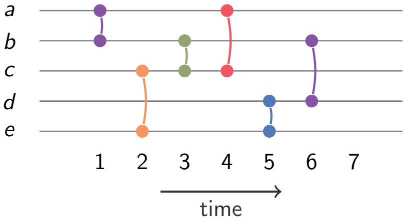
(a) representation of a temporal network



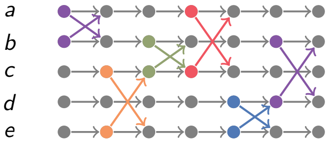
(b) static expansion of temporal network;
directed edges

static expansion of a temporal network

example



(a) representation of a temporal network



(b) static expansion of temporal network; delays

reachability, connectivity, and connected components

- defined as in static graphs, but using time-respecting paths
- reachability :
  - used to study infection spreading and information cascades
- connectivity : as in directed (static) graphs is not symmetric
  - distinguish strong and weak connectivity
  - in addition, we can define transitive connectivity:
    a subgraph is transitively connected if time-respecting paths from u to v and v to w imply a time-respecting path from u to w

minimum temporal paths

different notions of minimum temporal paths rely on time-respecting paths

- earliest-arrival path : a path from x to y with earliest arrival time
- latest-departure path : a path from x to y with latest departure time
- fastest path : path from x to y with minimum elapsed time
- shortest path : fastest path from x to y in terms of overall traversal time required on edges

[Wu et al., 2014]

diameter, network efficiency

- diameter : shortest latency of time-respecting paths over connected pairs [Chaintreau et al., 2007]
  - restricted on connected pairs, as real data have many disconnected pairs
- network efficiency : the harmonic mean of latency over all pairs [Tang et al., 2009]
  - discussion : what is the motivation for harmonic mean?

diameter, network efficiency

- diameter : shortest latency of time-respecting paths over connected pairs [Chaintreau et al., 2007]
  - restricted on connected pairs, as real data have many disconnected pairs
- network efficiency : the harmonic mean of latency over all pairs [Tang et al., 2009]
  - discussion : what is the motivation for harmonic mean?
  - it combines average latency and reachability ratio

centrality measures

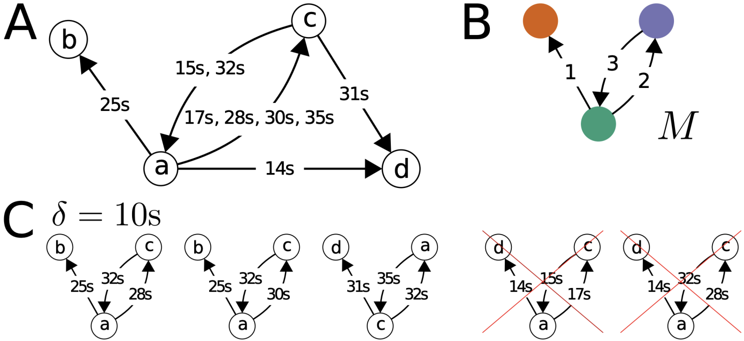
- many centrality measures on static graphs use distances
- closeness centrality: $C_c(u) = \frac{n-1}{\sum_{v \neq u} d(u,v)}$
- betweenness centrality: $C_b(u) = \frac{\sum_{v \neq u \neq w} p_u(v,w)}{\sum_{v \neq u \neq w} p(v,w)}$
- for temporal networks we replace distance with shortest latency time-respecting path
- analogues of Katz centrality and PageRank have also been defined
- discussion: how do these centrality measures on temporal networks compare with their static analogues?
  - main difference is that the temporal information becomes relevant

    e.g., betweenness centrality refers to a given node at a given time

temporal motifs

- temporal motif counting
  [Paranjape et al., 2017, Kovanen et al., 2013]:
  - temporal motif is a small subgraph with temporally ordered edges (and/or interval or delay constraints)

temporal motifs



δ-temporal motif: a sequence of directed temporally ordered edges which appear within a time window δ

[Paranjape et al., 2017]

agenda

part I : introduction and motivation

part II : models of temporal networks

part III : algorithmic frameworks

part IV : data mining problems

part V : conclusions and future challenges

part III

algorithmic frameworks for temporal network analysis

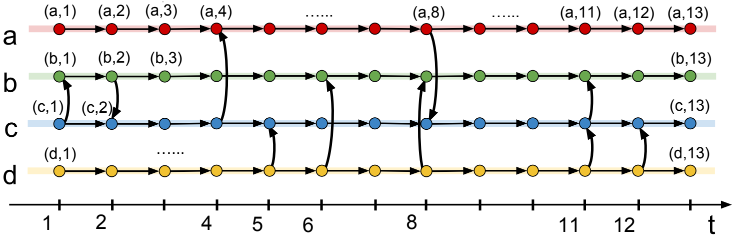
frameworks

adopted traditional frameworks

- static expansion graphs
- dynamic graphs
- time-series

static expansion graphs

- static graph of time-stamped nodes and time-forwarding edges $G_e = (V_e, E_e)$
- $V_e = \{(v, t) \mid v \in V,\ t \in T\}$, where T is the set of all possible timestamps
- edges E_e : interactions between the temporal nodes V_t



static expansion graphs

- static expansion graph is a directed graph
- standard graph algorithms (BFS, DFS, Dijkstra, Bellman-Ford) can be adopted for finding:
  - fastest temporal paths, shortest temporal paths, and weighted combinations
  - walks (revisiting of vertices and edges is allowed)
  - journeys (revisiting of vertices is allowed, but not edges)
- upstream, downstream reachability sets

minimum temporal paths

different notions of minimum temporal paths rely on time-respecting paths

- earliest-arrival path : a path from x to y with earliest arrival time
- latest-departure path : a path from x to y with latest departure time
- fastest path : path from x to y with minimum elapsed time
- shortest path : fastest path from x to y in terms of overall traversal time required on edges

[Wu et al., 2014]

earliest-arrival path

- temporal graph $G = (V, E)$
- source vertex x, starting time t_s
- array T of size $|V|$ to record arrival times to each node
- $T[x] = t_s$ and $T[v] = \infty$, for nodes other than source
- process edges (u, v, t, λ) in temporal order
  - if $t \geq T[u]$ (u is already reached from x)
  - check if the edge creates the earliest-seen-so-far path from x to v and update $T[v]$: $T[v] = \min(T[v], t + \lambda)$

[Wu et al., 2014]

latest-departure path

- temporal graph $G = (V, E)$
- sink vertex x, ending time t_s
- same process as for earliest-arrival path, but
  - process edges in reverse temporal order
- add new interaction to the path if it does not violate temporal order

[Wu et al., 2014]

minimum spanning trees

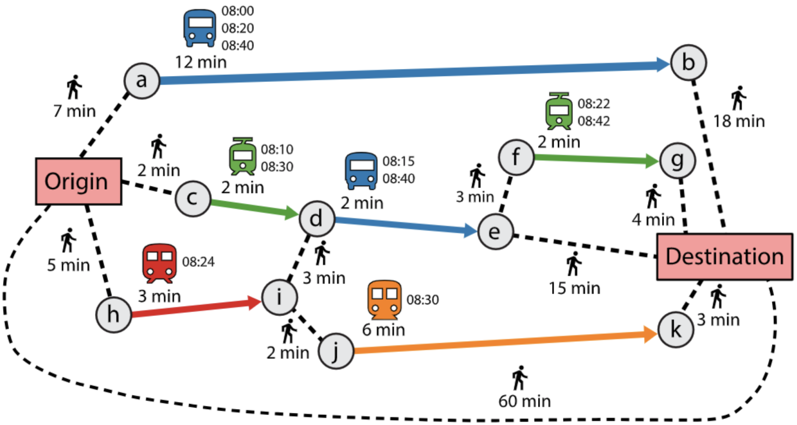
- mst_a : minimum spanning tree with earliest-arrival times
  - each temporal path from the root to the node is the earliest arrival path
  - can be solved in linear time
- mst_w : minimum spanning tree with smallest total weight
  - or with the smallest number of hops: directed Steiner tree
  - NP-hard, can adapt approximation algorithm from directed Steiner tree

[Huang et al., 2015]

applications of temporal paths

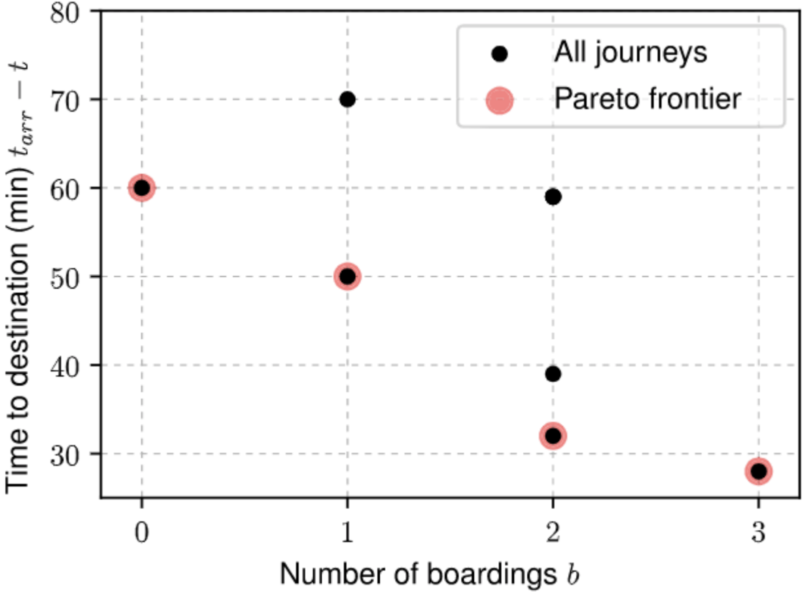
- temporal reachability problems
  - diffusion simulation, centrality measures
- directed spanning trees or Steiner trees
  - reconstruction of diffusion
- drawback: large size of expansion graph may lead to high computational complexity and large memory consumption
- challenge: scalable algorithms and approximations

applications — transportation temporal networks



[Kujala et al., 2018]

Pareto-optimal journeys



[Kujala et al., 2018]

dynamic graph algorithms on streaming model

- well-studied model
- extensive survey by [McGregor, 2014]
- different settings considered
  - node/edge additions (incremental)
  - node/edge additions/deletions (fully-dynamic)
  - updating weights/labels is a special case of the fully-dynamic model
  - sliding-window setting: consider only edges from latest interval of fixed length
  - algorithms can be deterministic or randomized

dynamic graph algorithms on streaming model

[McGregor, 2014]

	Insert-Only	Insert-Delete	Sliding Window (width w)
Connectivity	Deterministic [27]	Randomized [5]	Deterministic [22]
Bipartiteness	Deterministic [27]	Randomized [5]	Deterministic [22]
Cut Sparsifier	Deterministic [2, 8]	Randomized [6, 31]	Randomized [22]
Spectral Sparsifier	Deterministic [8, 46]	Randomized $\tilde{O}(n^{5/3})$ space [7]	Randomized $\tilde{O}(n^{5/3})$ space [22]
$(2t-1)$-Spanners	$O(n^{1+1/t})$ space [11, 23]	Only multiple pass results known [6]	$O(\sqrt{wn^{(1+1/t)}})$ space [22]
Min. Spanning Tree	Exact [27]	$(1+\epsilon)$-approx. [5] Exact in $O(\log n)$ passes [5]	$(1+\epsilon)$-approx. [22]
Unweighted Matching	2-approx. [27] 1.58 lower bound [42]	Only multiple pass results known [3, 4]	$(3+\epsilon)$-approx. [22]
Weighted Matching	4.911-approx. [25]	Only multiple pass results known [3, 4]	9.027-approx. [22]

Table 1: Single-Pass, Semi-Streaming Results: Algorithms use $O(n \operatorname{polylog} n)$ space unless noted otherwise.

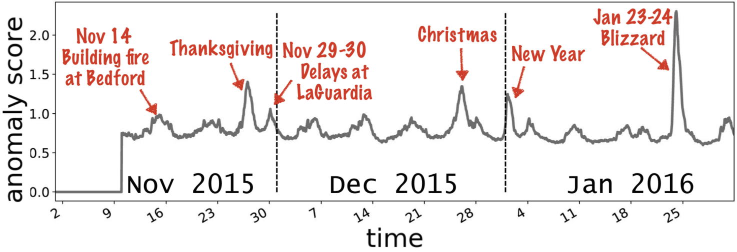
time-series analysis

- view a temporal network as a (multivariate) time series
  - calculate temporal profile of nodes, edges, or a whole network
  - calculate distance between adjacent snapshots and analyze the resulting time series
- distance: edit distance, node-profile distances, vector-space distance
- applications in change-point detection, anomaly detection, evolutionary pattern mining

event detection in time series

- given a sequence of graphs G_t
- a function to calculate the vertex affinity matrix S, where s_{ij} indicates the influence vertex i has on vertex j
- a set of time points for detected events is $\{t \in T \mid d(G_t, G_{t+1}) \geq \delta\}$ where

$$d(G_t, G_{t+1}) = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \left(\sqrt{S_{ij}^{(t)}} - \sqrt{S_{ij}^{(t+1)}}\right)^2}$$



[Eswaran et al., 2018]

time-series analysis

- anomaly detection survey [Ranshous et al., 2015]
- approach does not solve all the problems, as it does not capture the network topology
- possible work-around: use more topology embeddings metrics
  - larger neighborhoods, influence measures, eigenvectors, . . .

agenda

part I : introduction and motivation

part II : models of temporal networks

part III : algorithmic frameworks

part IV : data mining problems

part V : conclusions and future challenges

part IV

data mining problems

data mining problems

- community detection
- event detection
- finding important nodes
- epidemics analysis and influence spreading

community detection

community detection in static graphs

- static graphs: extensive survey [Fortunato and Hric, 2016]
- standard community definitions
  - a community is a set of nodes, which are closer to each other than to the rest of the network
  - a community is a dense network subgraph
- general definition [Coscia et al., 2011]
  - a community in a complex network is a set of entities that share some closely correlated sets of actions with the other entities of the community
- typical problem settings
  - a single community vs. network partition
  - overlapping vs. non-overlapping communities

community detection in static graphs

partition measures

- modularity : the difference between the actual number of edges and the expected
- cut : number of edges between a community and the rest of the graph
- ratio cut : cut normalized by the number of edges of community nodes
- ...

single-community measures

- average degree : $\frac{|E(S)|}{2|S|}$
- density : $\frac{2|E(S)|}{|S|(|S|-1)}$
- conductance : $\frac{cut(S,\bar{S})}{min\{vol(S),vol(\bar{S})\}}$
- ...

community detection in temporal networks

temporal information gives rise to several issues

- temporal localization: concise time interval or intervals, whole time history
- behaviour: single-appearance, recurring, persistent, evolutionary patterns, smoothness
- partition of the topology network vs. partition of the time history
- online vs. offline
- application-specific settings

temporal communities : temporal assumptions

often assume a prior model, which describes what is the temporal behavior of interesting community structures, e.g.,

- small/large covering intervals of community interactions
- frequent patterns
- persistent patterns

evolutionary patterns : vocabulary

evolutionary patterns of communities in the network [Dakiche et al., 2019]

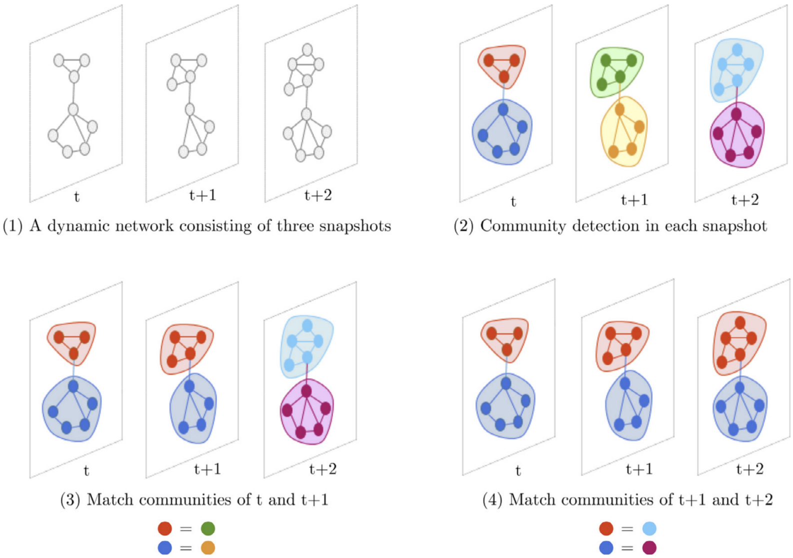
- birth
- death
- growth
- contraction
- merge
- split
- continue
- resurgence

temporal communities: idea #1

we follow a recent survey on community detection [Dakiche et al., 2019]

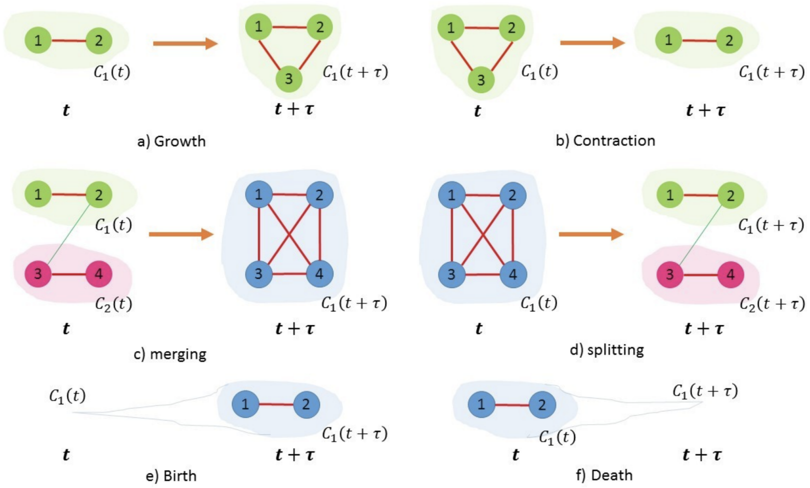
- independent community detection and matching
  - first detect communities at each timestamp
  - then match them across different timestamps

independent community detection and matching



[Dakiche et al., 2019]

typical evolutionary patterns



[Sun et al., 2015]

independent community detection and matching

advantages

- reuse of standard community detection methods
- use existing similarity measures

disadvantages

- instability of community-detection algorithms

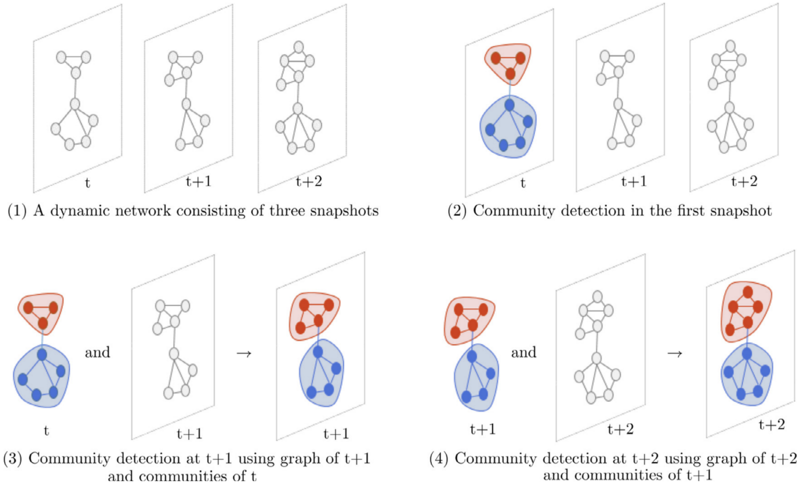
temporal communities: idea #2

- dependent community detection [Dakiche et al., 2019]

  detect communities at time t based on

  - network topology at t, and
  - communities at time $t - 1$

dependent community detection



[Dakiche et al., 2019]

Louvain algorithm

- a fast greedy approach based on modularity optimization
- reminder: the modularity objective

$$Q = \frac{1}{2m} \sum_{u,v} \left[A_{uv} - \frac{d_u d_v}{2m} \right] \delta(c_u, c_v)$$

- two phases repeated iteratively
  - initially, each node in network is a community
  - then, for each node i, consider its neighbor j and compute the gain of modularity of putting i into the community of j
  - node i is placed into the community with the largest gain, if the gain is positive

[Blondel et al., 2008]

Louvain algorithm

- on the second phase, each community is considered as a super-node
  - the edges between these super-nodes are contracted and re-weighed by the number of edges between them
- the two phases are repeated until there is no improvement in modularity
- the algorithm is extremely fast

[Blondel et al., 2008]

history-dependent approach

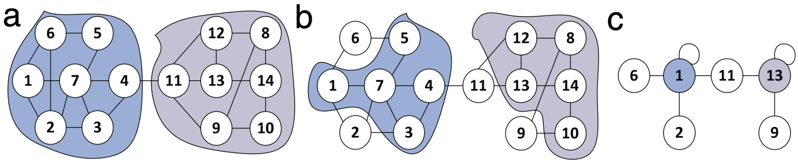
idea

- for two consecutive time steps, there only few edges that affect the community structure
- if the connections of all the nodes in the same community at time step $t-1$ keep unchanged at time step t, they are still in the same community at time step t
- thus, no need to break that super-node

[He and Chen, 2015]

history-dependent approach

- find all communities in snapshot $t = 1$
- for $t = 2$:
  - if a node's connection change, then remove it from its super-node and add as single node
  - leave all other nodes inside the super-node
  - re-weight the edges



[He and Chen, 2015]

dependent community detection

advantages

- a solution for the problem of instability
- improved computational complexity

disadvantages

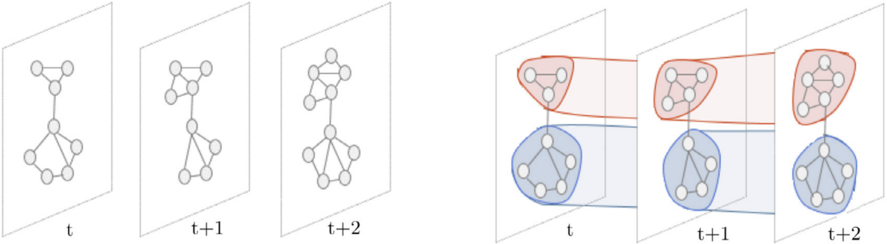
- traditional community detection methods are no longer directly applicable

temporal communities: idea #3

simultaneous community detection on all snapshots [Dakiche et al., 2019]

- construct a static expansion graph
  - add edges between instances of nodes in different timestamps
- run a standard community detection on the resulting graph

simultaneous community detection on all snapshots



(1) A dynamic network consisting of three snapshots (2) Community detection on all snapshots

[Dakiche et al., 2019]

simultaneous community detection

costs

- switching cost : each node u incurs cost C_{sw} when changing community affiliation
- false negative cost : two nodes incur cost C_{fn} when belong to the same community but do not interact
- false positive cost : two nodes incur cost C_{fp} when belong to different communities but do interact

resulting problem

- find a partition into clusters that minimizes the total cost of switching, false negative, and false positive

[Tantipathananandh and Berger-Wolf, 2011]

simultaneous community detection on all snapshots

advantages

- provides a solution for the problem of instability

disadvantages

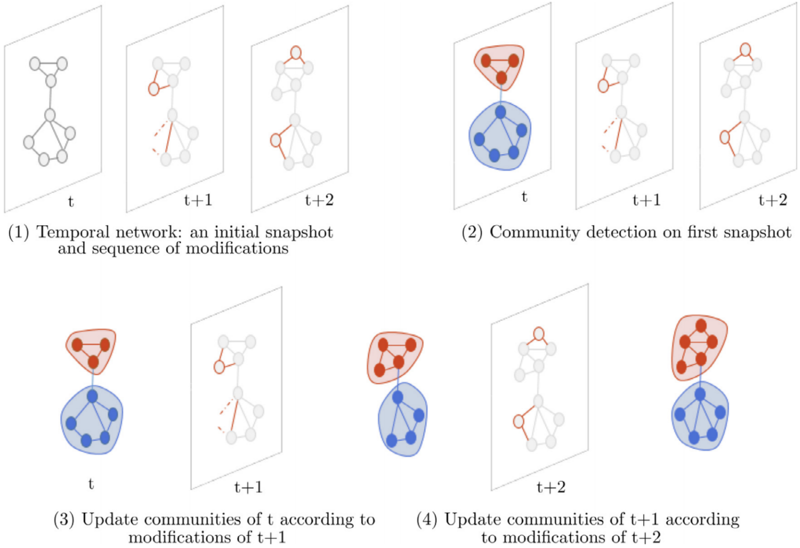
- no possibility to track community evolution in a network evolving in real time

temporal communities: idea #4

dynamic community detection [Dakiche et al., 2019]

- update previously discovered communities according to network modifications

dynamic community detection



[Dakiche et al., 2019]

dynamic community detection

advantages

- provides a solution for the problem of instability
- light-weight methods to track communities

disadvantages

- possibility to drift towards invalid communities

event detection

event detection

- given a network representing some kind of activity
  - network of social interactions
  - social-media feed
  - transportation network
- an event can be generally defined as an activity with some prominent qualitative or quantitative difference from the background activity
  - bursting news about major natural disasters
  - abnormally high traffic in the city
  - an emerging new discussion topic in social media

temporal event detection: standard approaches

abnormality score

- the likelihood that an interval contains an event can be estimated by comparing an abnormality score on the interval

[Heins and Stern, 2014]

predictive models

- learn a predictive model and find intervals and time points whose behavior differ from the predicted one

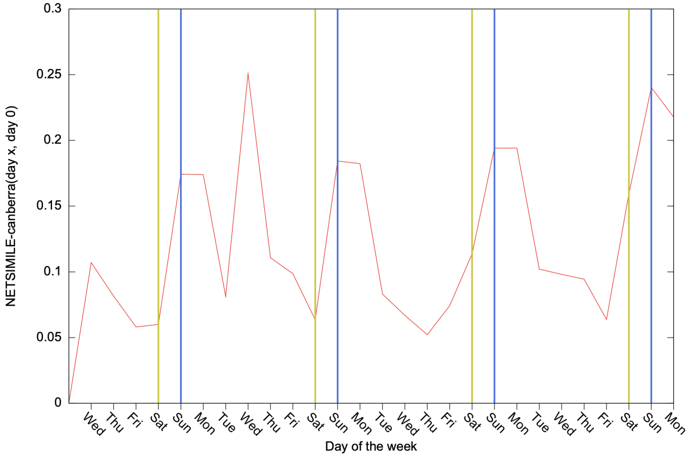
[Hunter and McIntosh, 1999, Gensler and Sick, 2017]

Netsimile

- an event exists in G_{j+1}, if G_{j+1} is very different than G_j
- for each node calculate 7 local and ego-network-based measures
  - degree
  - clustering coefficient
  - average degree of neighbours
  - average clustering coefficient of neighbours
  - number of edges in the ego-network
  - number of edges outgoing from the ego-network
  - number of neighbours of the ego-network
- combine into a signature vector and compare

[Berlingerio et al., 2012]

Netsimile algorithm



(a) NetSimile between each day and day 0 in Yahoo! IM

[Berlingerio et al., 2012]

spatiotemporal event detection

detailed survey by [Shi and Pun-Cheng, 2019]

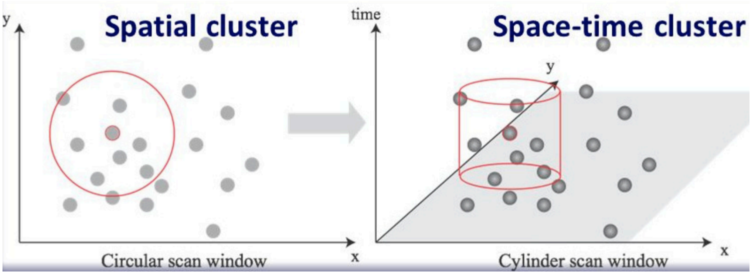
- consider time and the (geo-)location of an event
- sources of spatial data
  - GPS devices / smart phones
  - geo-tagged messages in online social networks
- typical approaches model the data as a set of geo-locations associated with activity measurements
- given a set of locations with activity measures, find a subset of locations that are close to each other and have abnormal activity pattern
- in spatiotemporal setting, one is also interested in finding the time interval (moment) of an event

spatiotemporal event detection: scan statistics

- a classic family of methods is spatial and spatiotemporal scan statistics
- scan over the space and time windows to identify regions of data generated by some process

spatiotemporal event detection: scan statistics

- a seminal paper : spatial scan statistics [Kulldorff, 1997]
- scan a circular spatial window and test the non-randomness of data against Poisson or Bernoulli baseline process
- later the approach was extended to spatiotemporal scans with cylindric windows



[Takahashi et al., 2004]

structural event

- structural event:
  - set of interconnected abnormal nodes
- e.g., the edge weights represent similarity of nodes
  - similarities between twitter users in preferences, language, visited locations, etc.
- scan extension to graph model [Liu et al., 2016]
- scan through a graph neighborhood — a set of interconnected nodes
- dense subgraph detection
  - e.g., [Charikar, 2000, Khuller and Saha, 2009]

finding important nodes

PageRank

- classic approach for measuring node importance
- listed in the top-10 most important data-mining algorithms [Wu et al., 2008]
- numerous applications
  - ranking web pages
  - trust and distrust computation
  - finding experts in social networks
  - ...

static PageRank

- graph $G = (V, E)$
- corresponding row-stochastic matrix $P \in \mathbb{R}^{n \times n}$
- personalization vector $\mathrm{h} \in \mathbb{R}^n$
- PageRank is the stationary distribution of a random walk,

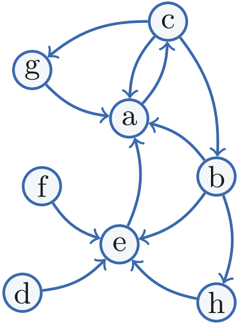
  with restart probability $(1 - \alpha)$

$$\pi(u) = \sum_{v \in V} \sum_{k=0}^{\infty} (1 - \alpha)\alpha^k \sum_{\substack{z \in \mathcal{Z}(v,u) \\ |z| = k}} h(v) \mathrm{Pr}[z \mid v]$$

where, $\mathcal{Z}(v, u)$ is the set of all paths from v to u

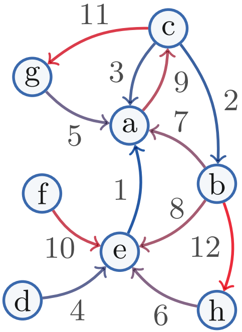
and $\mathrm{Pr}[z \mid v] = \prod_{(i,j) \in z} P(i, j)$

motivating example



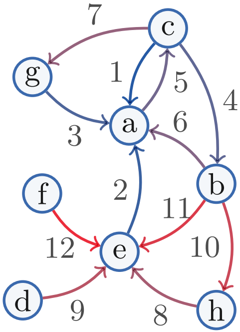
(a)

static network



(b)

temporal network



(c)

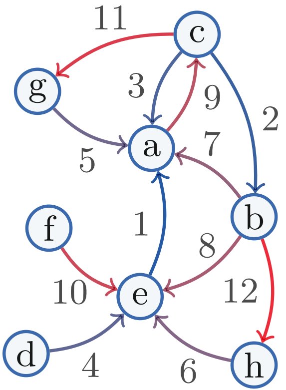
temporal network

temporal PageRank

- make a random walk only on temporal paths

  e.g., time-respecting paths

  time-stamps increase along the path



$c \to b \to a \to c$: time respecting

$a \to c \to b \to a$: not time respecting

temporal PageRank

- intuition : probability of visiting node u at time t,

  given a random walk on temporal paths

- need to model probability of following next temporal edge
  - we use an exponential distribution

- temporal PageRank definition

$$r(u,t) = \sum_{v \in V} \sum_{k=0}^{t} (1-\alpha)\alpha^k \sum_{\substack{z \in \mathcal{Z}^T(v,u|t) \\ |z|=k}} \Pr{}'[z|\ t]$$

$\mathcal{Z}^T(v, u \mid t)$ set of temporal paths from v to u until time t

static vs. temporal PageRank

- computation: simple online algorithm iterating over edges
- temporal PageRank is designed to capture changes in network dynamics and concept drifts
- proposition :

  if the edge distribution is stable, then
  as $T \to \infty$, the temporal PageRank on G
  converges to the static PageRank on G_S,
  with personalization vector equal to weighted out-degree

[Rozenshtein and Gionis, 2016]

diffusion analysis and influence spreading

diffusion analysis and influence spreading

- propagation models
  - used to study disease spreading or information cascade in the network
- activity spreading: virus, information, idea, rumor
- applications: epidemiology, information security, marketing
- why use models?
  - facilitate mathematical analysis of propagation processes
  - have intuitive interpretation
  - proven to be realistic by empirical studies
- extensive survey in the book [Shakarian et al., 2015]

standard models

- susceptible-infected (SI) model
  - SIR, SIRS, other variants
- independent cascade (IC) model
- linear threshold (LT) model
- shortest path (SP) model

susceptible-infected-recovered (SIR) model

- a popular model to analyze epidemics
- population is divided into three categories
  - susceptible: may be infected if comes in contact with an infectious individual
  - infectious: infected and capable of infecting susceptible individuals
  - recovered: either recovered and become immune, or deceased
- $S(t)$, $I(t)$, $R(t)$: number of susceptible, infectious, recovered individuals at time t
- ordinary differential equations describe the rate of growth of the three populations

susceptible-infected-recovered (SIR) model

- parameters of the SIR model
  - β: average number of contacts, multiplied by the probability of disease transmission
  - $\gamma = 1/D$, where D: an individual is infectious for an average time period D
  - $R_0 = \beta/\gamma$: basic reproduction ratio
  - λ: largest eigenvalue of stochastic system matrix, if network structure is considered
- model can be used to analyze whether the disease will persist or die out
- exhibits threshold phenomena behavior
- many variants of the basic model

static models: assumptions

all models have similar implicit assumptions on temporality:

1. uniform time steps
2. interactions happen at each time step and are independent

drawbacks of static models

- large heterogeneity in the time instances of real interactions

  [Barabasi, 2005, Candia et al., 2008, Leskovec and Horvitz, 2008]

  - burstiness in communication patterns
  - periodic activity changes
  - causal relationships between interactions

temporal propagation models

- intuitive extensions from static graphs to temporal graphs
- add distributions (e.g., Poisson or power-law) of the intervals between interactions

  [Vazquez et al., 2007, Min et al., 2011]
- continuous time, partially observed graph
- develop mathematical analysis for novel and generalized models

  [Harris, 2002, Fernández-Gracia et al., 2011]

typical problem formulations

- immunization strategies
- influence maximization
- seed and cascade reconstruction

static immunization strategies

- how to stop or prevent a viral diffusion?
- main aspects differentiating the research works:
  - assumptions about the spreading model
  - assumptions about the network structure
  - whether the whole network is observable
- both assumptions on the network structure and on the infection propagation are crucial
- results may not hold for any general network and real infection

[Newman, 2003, Pastor-Satorras and Vespignani, 2002a]

static immunization strategies

- simple model-blind strategies, such as random immunization, perform moderately well in different scenarios

  [Pastor-Satorras and Vespignani, 2002b, Madar et al., 2004]

- better results on real-world networks can be achieved by immunizing nodes with high connectivity

  [Pastor-Satorras and Vespignani, 2002b, Dezső and Barabási, 2002].

- requires explicit knowledge of the network structure and it is impractical in real applications

static immunization strategies

- [Cohen et al., 2003] overcomes this drawback by employing acquaintance immunization strategy:
- immunization of random neighbors of randomly selected nodes leads to immunization of the most central nodes without knowing any global information about the network

temporal immunization strategies

- adjust successful static strategies
- e.g., Cohen's neighborhood vaccination scheme [Lee et al., 2012]
- two vaccination strategies
- recent :
  - ask a random individual i to name its most recent contact and vaccinate this person
- weight :
  - ask a random individual i to name its most frequent contact since some time t

static influence maximization

- how to select the initial set of infected nodes (seeds), such that the speed, size, or other spread characteristics are optimized
- applications in marketing and network design
- influence maximization problem was introduced by [Kempe et al., 2003] in the IC and LT models
- find a set of k seed nodes, such that the expected number of nodes activated by the infection cascade is maximized

static influence maximization

- NP-hard [Kempe et al., 2003]
- simple greedy algorithm with approximation guarantee
- influence maximization problem was been studied for many different variants of other models, constraints, and objective functions
- many practical heuristics and approximations

[Chen et al., 2009, Chen et al., 2010, Tang et al., 2014]

temporal influence maximization

- intuitive approach to capture temporality:
  - sequence of graphs (or snapshots)
  - each time step of propagation corresponds to propagation over the corresponding graph
  - all interactions within one time step happen simultaneously
- related papers by [Aggarwal et al., 2012, Zhuang et al., 2013, Gayraud et al., 2015]

temporal influence maximization

- another approach:
- incorporate time into the diffusion model as distribution of intervals between the interactions
- different types of models and interval distributions
  [Chen et al., 2012, Liu et al., 2012, Rodriguez and Schölkopf, 2012, Du et al., 2013]
- the most realistic approachable setting?
- an interesting research problem:
  - infer propagation model parameters from the data
    [Rodriguez et al., 2011, Gomez-Rodriguez et al., 2016]

seed and cascade reconstruction

- given some observed data about the infection
  - e.g., a small subset of infected nodes,

  the goal is to find the most probable seed nodes
- other versions:
  - find the most probable cascades
- the order of infection (who got infected from whom)
- these works are data-driven:
  - it is essential that the assumed propagation model matches the actual infection flow in the network

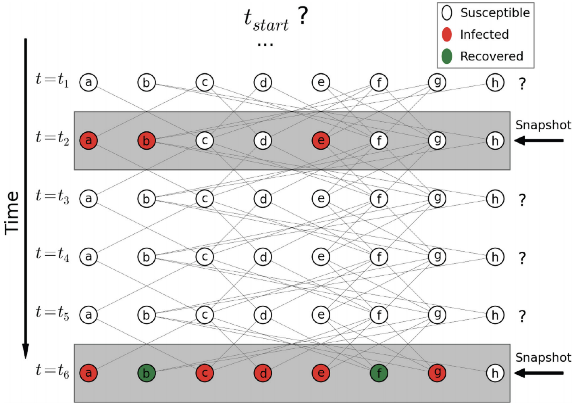
seed and cascade reconstruction

- applications:
  - epidemiology (who was the patient zero?)
  - influencer discovery (who was the source of information?)
- a number of different approaches
  - find a single source under the SI model [Shah and Zaman, 2011]
  - multiple seeds [Prakash et al., 2012]
  - k seeds under the IC model [Lappas et al., 2010]
  - take advantage of the recorded infection order [Sefer and Kingsford, 2016]
- the above papers are pre-covid, so hugely obsolete now

temporal reconstruction

- the problems formulated in this setting tend to be either
  - oversimplified versions of static reconstruction or
  - become too hard or ill-posed
- knowing the history of interactions allow to reconstruct feasible paths of infection and prune unfeasible
- any noise or missing information adds uncertainty
  - typically need more assumptions about the noise and missing information
- the knowledge of the diffusion model in crucial
- see survey paper by [Holme, 2015]

history reconstruction



[Sefer and Kingsford, 2016]

agenda

part I : introduction and motivation

part II : models of temporal networks

part III : algorithmic frameworks

part IV : data mining problems

part V : conclusions and future challenges

part V

conclusions and future challenges

temporal community detection: challenges

- large number of problem formulations and variants
- lack of fundamental theoretical treatment
  - most of the approaches are heuristics
  - many are combinations of several ideas and algorithms
  - require many parameters and attention to implementation details
- hard to compare methods and choose one for a specific application
  - few datasets with ground-truth temporal communities
  - synthetic generators are built on various assumptions
  - no standard benchmarks
- a large number of quality metrics to calculate and compare
- may be misleading if a method is not designed for that particular community definition

event detection: challenges

- actively evolving area, application- and data-oriented
- families of problems and methods are considered only for the specific sources of data
  - e.g., a large body of research is focused on the analysis of Twitter data

    [Atefeh and Khreich, 2015]
- no unified classification for problem settings, research questions, and data requirements
- speed and quality:
  - online streaming event-detection techniques are demanded for nearly real-time event detection
  - quality: both false events and missed events may have a high price
- methods should rely more multi-modal data, e.g., combining network structure with text

references I

Aggarwal, C. C., Lin, S., and Yu, P. S. (2012).
On influential node discovery in dynamic social networks.
In *Proceedings of the 2012 SIAM International Conference on Data Mining*, pages 636–647. SIAM.

Atefeh, F. and Khreich, W. (2015).
A survey of techniques for event detection in twitter.
Computational Intelligence, 31(1):132–164.

Barabasi, A.-L. (2005).
The origin of bursts and heavy tails in human dynamics.
Nature, 435(7039):207.

Berlingerio, M., Koutra, D., Eliassi-Rad, T., and Faloutsos, C. (2012).
Netsimile: A scalable approach to size-independent network similarity.
arXiv preprint arXiv:1209.2684.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008).
Fast unfolding of communities in large networks.
Journal of statistical mechanics: theory and experiment, 2008(10):P10008.

references II

- Candia, J., González, M. C., Wang, P., Schoenharl, T., Madey, G., and Barabási, A.-L. (2008).
  Uncovering individual and collective human dynamics from mobile phone records.
  Journal of physics A: mathematical and theoretical, 41(22):224015.
- Casteigts, A., Flocchini, P., Quattrociocchi, W., and Santoro, N. (2012).
  Time-varying graphs and dynamic networks.
  International Journal of Parallel, Emergent and Distributed Systems, 27(5):387–408.
- Chaintreau, A., Mtibaa, A., Massoulie, L., and Diot, C. (2007).
  The diameter of opportunistic mobile networks.
  In *Proceedings of the 2007 ACM CoNEXT conference*, page 12. ACM.
- Charikar, M. (2000).
  Greedy approximation algorithms for finding dense components in a graph.
  In *International Workshop on Approximation Algorithms for Combinatorial Optimization*, pages 84–95. Springer.
- Chen, W., Lu, W., and Zhang, N. (2012).
  Time-critical influence maximization in social networks with time-delayed diffusion process.
  In *AAAI*, volume 2012, pages 1–5.

references III

Chen, W., Wang, C., and Wang, Y. (2010).
Scalable influence maximization for prevalent viral marketing in large-scale social networks.
In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1029–1038. ACM.

Chen, W., Wang, Y., and Yang, S. (2009).
Efficient influence maximization in social networks.
In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 199–208. ACM.

Cohen, R., Havlin, S., and Ben-Avraham, D. (2003).
Efficient immunization strategies for computer networks and populations.
Physical review letters, 91(24):247901.

Coscia, M., Giannotti, F., and Pedreschi, D. (2011).
A classification for community discovery methods in complex networks.
Statistical Analysis and Data Mining: The ASA Data Science Journal, 4(5):512–546.

references IV

Dakiche, N., Tayeb, F. B.-S., Slimani, Y., and Benatchba, K. (2019).
Tracking community evolution in social networks: A survey.
Information Processing & Management, 56(3):1084–1102.

Dezső, Z. and Barabási, A.-L. (2002).
Halting viruses in scale-free networks.
Physical Review E, 65(5):055103.

Du, N., Song, L., Rodriguez, M. G., and Zha, H. (2013).
Scalable influence estimation in continuous-time diffusion networks.
In *Advances in neural information processing systems*, pages 3147–3155.

Eswaran, D., Faloutsos, C., Guha, S., and Mishra, N. (2018).
Spotlight: Detecting anomalies in streaming graphs.
In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1378–1386. ACM.

Fernández-Gracia, J., Eguíluz, V. M., and San Miguel, M. (2011).
Update rules and interevent time distributions: Slow ordering versus no ordering in the voter model.
Physical Review E, 84(1):015103.

references V

Fortunato, S. and Hric, D. (2016).
Community detection in networks: A user guide.
Physics reports, 659:1–44.

Gayraud, N. T., Pitoura, E., and Tsaparas, P. (2015).
Diffusion maximization in evolving social networks.
In *Proceedings of the 2015 ACM on Conference on Online Social Networks*, pages 125–135. ACM.

Gensler, A. and Sick, B. (2017).
Performing event detection in time series with swiftevent: an algorithm with supervised learning of detection criteria.
Pattern Analysis and Applications, pages 1–20.

Gomez-Rodriguez, M., Song, L., Du, N., Zha, H., and Schölkopf, B. (2016).
Influence estimation and maximization in continuous-time diffusion networks.
ACM Transactions on Information Systems (TOIS), 34(2):9.

Harris, T. E. (2002).
The theory of branching processes.
Courier Corporation.

references VI

He, J. and Chen, D. (2015).
A fast algorithm for community detection in temporal network.
Physica A: Statistical Mechanics and its Applications, 429:87–94.

Heins, K. and Stern, H. (2014).
A statistical model for event sequence data.
In *Artificial Intelligence and Statistics*, pages 338–346.

Henzinger, M. R., King, V., and King, V. (1999).
Randomized fully dynamic graph algorithms with polylogarithmic time per operation.
Journal of the ACM (JACM), 46(4):502–516.

Holme, P. (2015).
Modern temporal network theory: a colloquium.
The European Physical Journal B, 88(9):234.

Huang, S., Fu, A. W.-C., and Liu, R. (2015).
Minimum spanning trees in temporal graphs.
In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, pages 419–430. ACM.

references VII

Hunter, J. and McIntosh, N. (1999).

Knowledge-based event detection in complex time series data.

In *Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making*, pages 271–280. Springer.

Kempe, D., Kleinberg, J., and Tardos, É. (2003).

Maximizing the spread of influence through a social network.

In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM.

Khuller, S. and Saha, B. (2009).

On finding dense subgraphs.

In *International Colloquium on Automata, Languages, and Programming*, pages 597–608. Springer.

Kovanen, L., Karsai, M., Kaski, K., Kertész, J., and Saramäki, J. (2013).

Temporal motifs.

In *Temporal Networks*, pages 119–133. Springer.

references VIII

Kujala, R., Weckström, C., Mladenović, M. N., and Saramäki, J. (2018).
Travel times and transfers in public transport: Comprehensive accessibility analysis based on pareto-optimal journeys.
Computers, Environment and Urban Systems, 67:41–54.

Kulldorff, M. (1997).
A spatial scan statistic.
Communications in Statistics-Theory and Methods, 26(6):1481–1496.

Lappas, T., Terzi, E., Gunopulos, D., and Mannila, H. (2010).
Finding effectors in social networks.
In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1059–1068. ACM.

Lee, S., Rocha, L. E., Liljeros, F., and Holme, P. (2012).
Exploiting temporal network structures of human interaction to effectively immunize populations.
PloS one, 7(5):e36439.

references IX

Leskovec, J. and Horvitz, E. (2008).
Planetary-scale views on a large instant-messaging network.
In *Proceedings of the 17th international conference on World Wide Web*, pages 915–924. ACM.

Liu, B., Cong, G., Xu, D., and Zeng, Y. (2012).
Time constrained influence maximization in social networks.
In *Data Mining (ICDM), 2012 IEEE 12th International Conference on*, pages 439–448. IEEE.

Liu, Y., Zhou, B., Chen, F., and Cheung, D. W. (2016).
Graph topic scan statistic for spatial event detection.
In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 489–498. ACM.

Madar, N., Kalisky, T., Cohen, R., Ben-avraham, D., and Havlin, S. (2004).
Immunization and epidemic dynamics in complex networks.
The European Physical Journal B, 38(2):269–276.

McGregor, A. (2014).
Graph stream algorithms: a survey.
ACM SIGMOD Record, 43(1):9–20.

references X

Min, B., Goh, K.-I., and Vazquez, A. (2011).
Spreading dynamics following bursty human activity patterns.
Physical Review E, 83(3):036102.

Muthukrishnan, S. et al. (2005).
Data streams: Algorithms and applications.
Foundations and Trends® in Theoretical Computer Science, 1(2):117–236.

Newman, M. E. (2003).
The structure and function of complex networks.
SIAM review, 45(2):167–256.

Paranjape, A., Benson, A. R., and Leskovec, J. (2017).
Motifs in temporal networks.
In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pages 601–610. ACM.

Pastor-Satorras, R. and Vespignani, A. (2002a).
Epidemics and immunization in scale-free networks.
arXiv preprint cond-mat/0205260.

references XI

Pastor-Satorras, R. and Vespignani, A. (2002b).
Immunization of complex networks.
Physical Review E, 65(3):036104.

Prakash, B. A., Vreeken, J., and Faloutsos, C. (2012).
Spotting culprits in epidemics: How many and which ones?
In *Data Mining (ICDM), 2012 IEEE 12th International Conference on*, pages 11–20. IEEE.

Ranshous, S., Shen, S., Koutra, D., Harenberg, S., Faloutsos, C., and Samatova, N. F. (2015).
Anomaly detection in dynamic networks: a survey.
Wiley Interdisciplinary Reviews: Computational Statistics, 7(3):223–247.

Rodriguez, M. G., Balduzzi, D., and Schölkopf, B. (2011).
Uncovering the temporal dynamics of diffusion networks.
arXiv preprint arXiv:1105.0697.

Rodriguez, M. G. and Schölkopf, B. (2012).
Influence maximization in continuous time diffusion networks.
arXiv preprint arXiv:1205.1682.

references XII

Rozenshtein, P. and Gionis, A. (2016).
Temporal pagerank.
In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 674–689. Springer.

Sefer, E. and Kingsford, C. (2016).
Diffusion archeology for diffusion progression history reconstruction.
Knowledge and information systems, 49(2):403–427.

Shah, D. and Zaman, T. (2011).
Rumors in a network: Who's the culprit?
IEEE Transactions on information theory, 57(8):5163–5181.

Shakarian, P., Bhatnagar, A., Aleali, A., Shaabani, E., and Guo, R. (2015).
Diffusion in social networks.
Springer.

Shi, Z. and Pun-Cheng, L. S. (2019).
Spatiotemporal data clustering: A survey of methods.
ISPRS International Journal of Geo-Information, 8(3):112.

Sun, Y., Tang, J., Pan, L., and Li, J. (2015).
Matrix based community evolution events detection in online social networks.
In *2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity)*, pages 465–470. IEEE.

Takahashi, K., Yokoyama, T., and Tango, T. (2004).
Flexscan: Software for the flexible spatial scan statistic.
National Institute of Public Health, Japan.

Tang, J., Musolesi, M., Mascolo, C., and Latora, V. (2009).
Temporal distance metrics for social network analysis.
In *Proceedings of the 2nd ACM workshop on Online social networks*, pages 31–36. ACM.

Tang, Y., Xiao, X., and Shi, Y. (2014).
Influence maximization: Near-optimal time complexity meets practical efficiency.
In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, pages 75–86. ACM.

references XIV

Tantipathananandh, C. and Berger-Wolf, T. Y. (2011).
Finding communities in dynamic social networks.
In *2011 IEEE 11th International Conference on Data Mining*, pages 1236–1241. IEEE.

Thorup, M. (2000).
Near-optimal fully-dynamic graph connectivity.
In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pages 343–350. Citeseer.

Vazquez, A., Racz, B., Lukacs, A., and Barabasi, A.-L. (2007).
Impact of non-poissonian activity patterns on spreading processes.
Physical review letters, 98(15):158702.

Wu, H., Cheng, J., Huang, S., Ke, Y., Lu, Y., and Xu, Y. (2014).
Path problems in temporal graphs.
Proceedings of the VLDB Endowment, 7(9):721–732.

references XV

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Philip, S. Y., et al. (2008).

Top 10 algorithms in data mining.

KAIS.

Zhuang, H., Sun, Y., Tang, J., Zhang, J., and Sun, X. (2013).

Influence maximization in dynamic social networks.

In *Data Mining (ICDM), 2013 IEEE 13th International Conference on*, pages 1313–1318. IEEE.