Finding simple temporal cycles in an interaction network

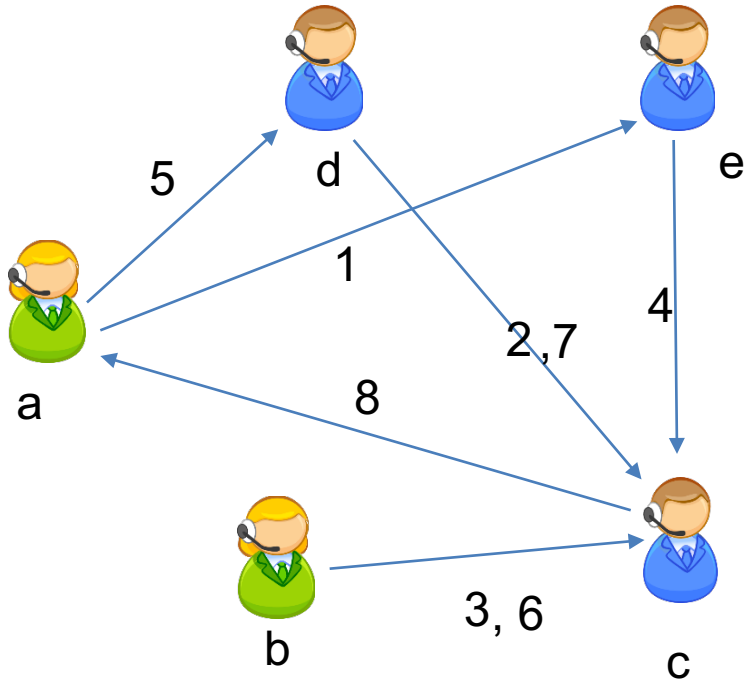
Rohit Kumar, Toon Calders

18-09-2017

Interaction Network/Graph

- Interaction Network: sequence of timestamped interactions ϵ over edges of a static graph G = (V,E).
- For example
  - Social interaction in a social network;
  - Email/ Message or call interaction in a communication network;
  - Data exchange in a computer network;
  - Money transactions in a financial network.

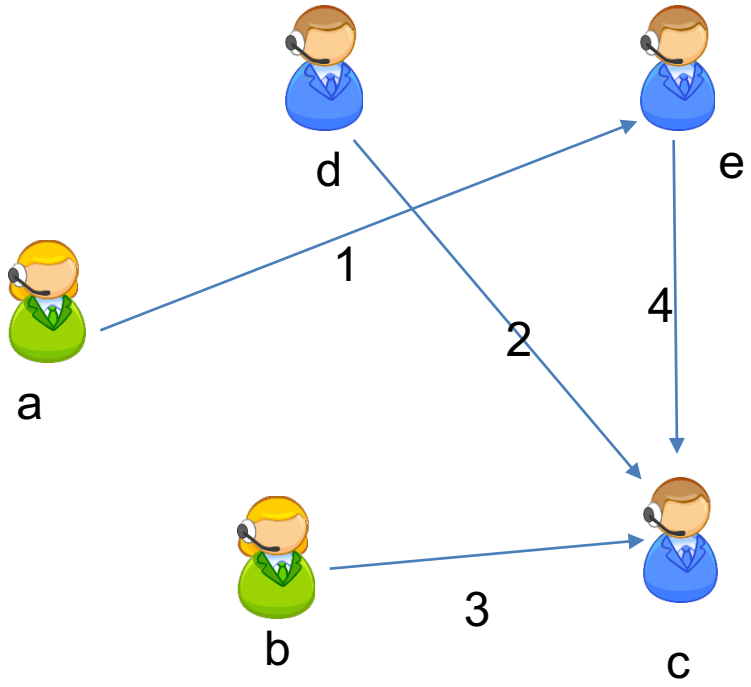
Example



(a , e,1) (d, c, 2) (b , c, 3) (e , c, 4) (a , d, 5) (b , c, 6) (d, c, 7) (a , c, 8) . . . .

Interaction network in a window

Window=4

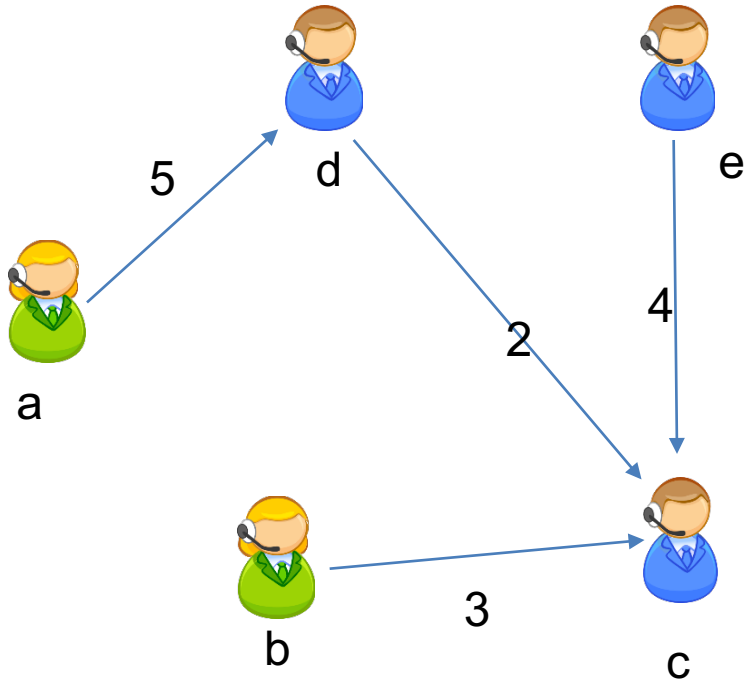


Future Edges

(a , e,1) (d, c, 2) (b , c, 3) (e , c, 4) (a , d, 5) (b , c, 6) (d, c, 7) (a , c, 8) . . . .

Interaction network in a window

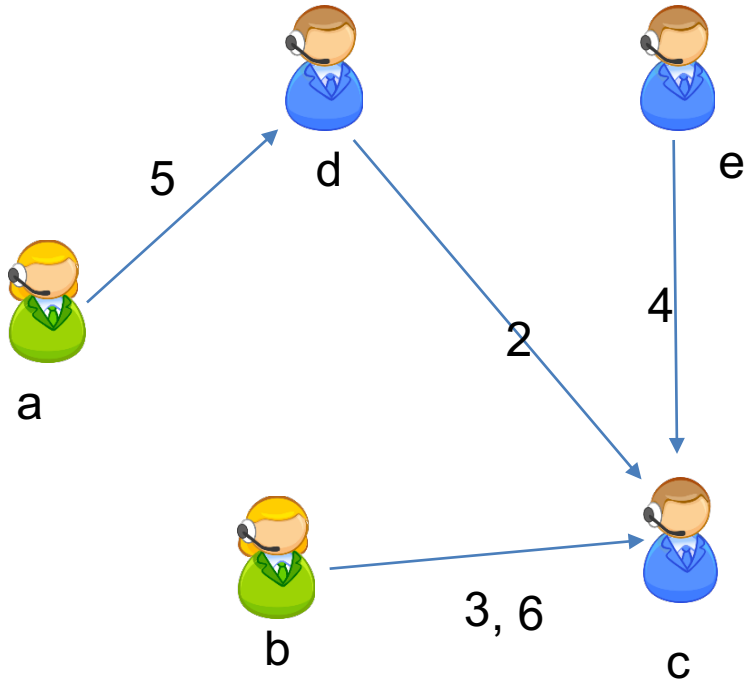
Window=4



(a , e,1) (d, c, 2) (b , c, 3) (e , c, 4) (a , d, 5) (b , c, 6) (d, c, 7) (a , c, 8) . . . .

Interaction network in a window

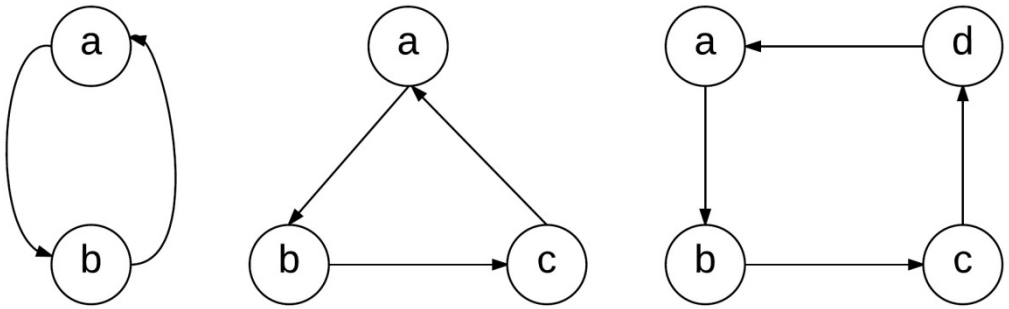
Window=4

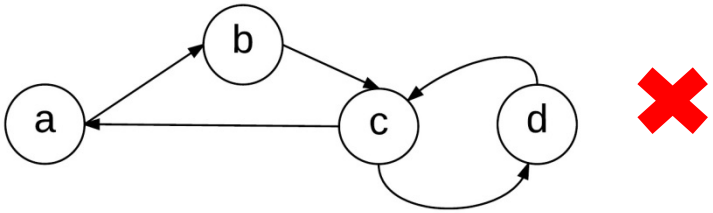


(a , e,1) (d, c, 2) (b , c, 3) (e , c, 4) (a , d, 5) (b , c, 6) (d, c, 7) (a , c, 8) . . . .

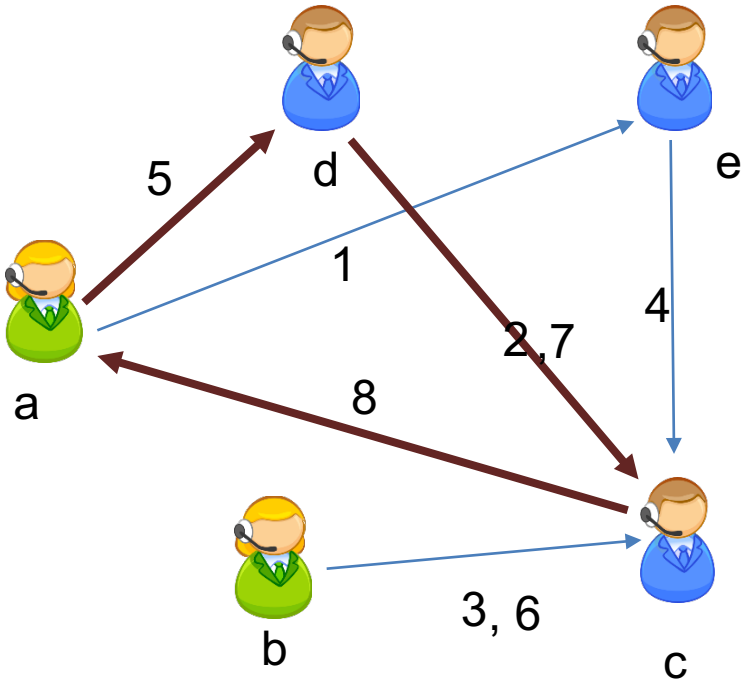
Simple Cycle

"A simple cycle is a closed path/walk with no repetitions of vertices or edges allowed, other than the repetition of starting and end vertex."





Simple Temporal Cycle



Duration = t_n-t_1 = 8-5=3

Length = #edges = 3

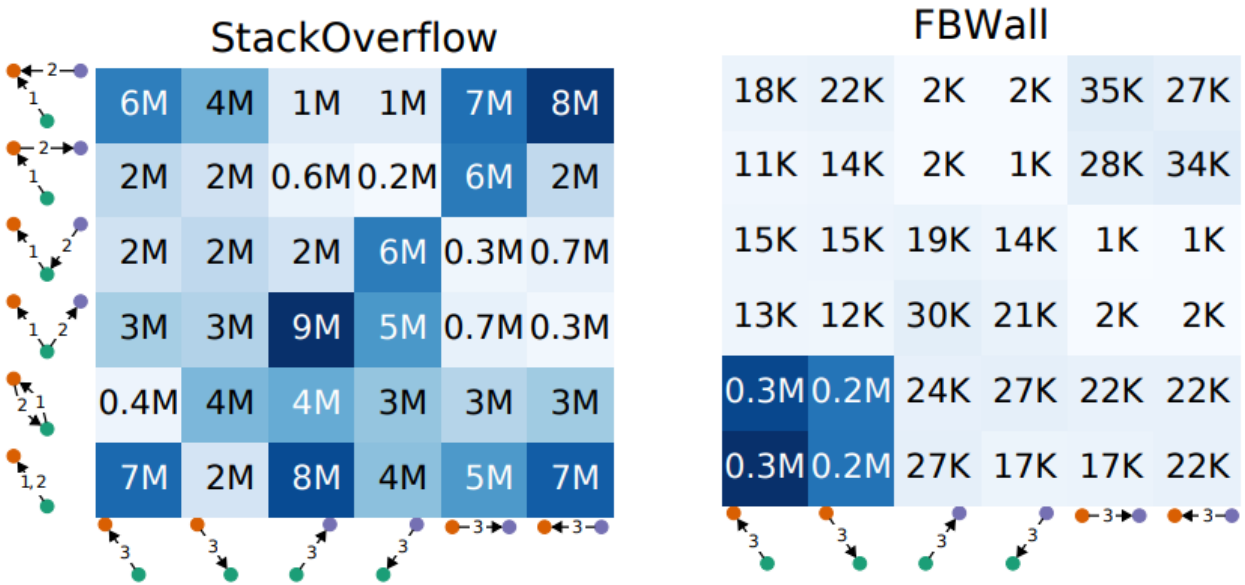
(a , d, 5) →(d, c, 7)→(c , a, 8)

(a , d, 5) →(d, c, 2)→(c , a, 8) ✖

Motivation

Interaction networks: many interesting patterns.

Patterns capture differences in use of networks



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

Perspective

- **Simple Cycle in Retweet network**
  - **Discussion among peers**
  - **Group of fake account to promote advertisement messages.**
- **Financial transaction network**
  - **Indicator for fraud or tax invasion.**

What we want to study

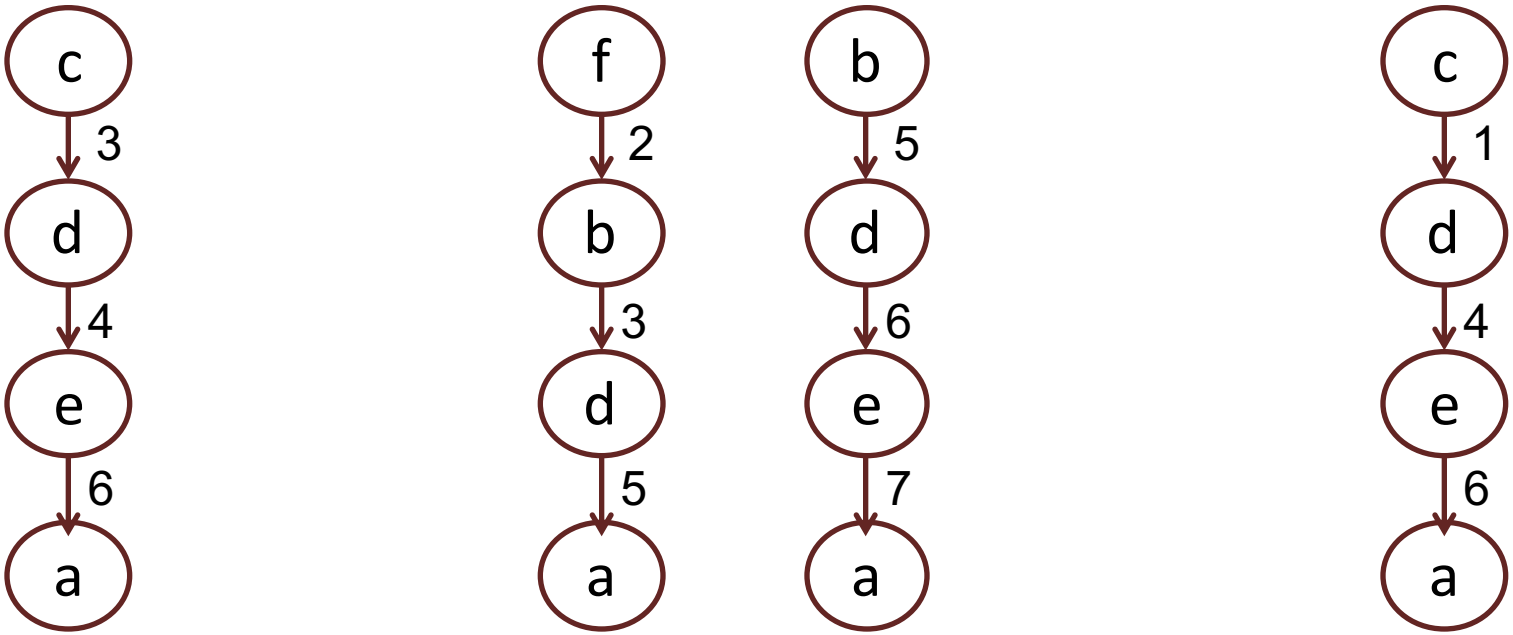
The main focus of this study is, given an temporal network and a time window (w) :

- **Find all simple cycles in the time window.**
- **Find most Frequent root nodes.**
- **Using simple cycle length frequency distribution to categories the type of network.**

Naïve algorithm

Window = 7 timestamps

$\overrightarrow{a, b, 9}$ Step 1: Fetch all temporal path ending at “a”

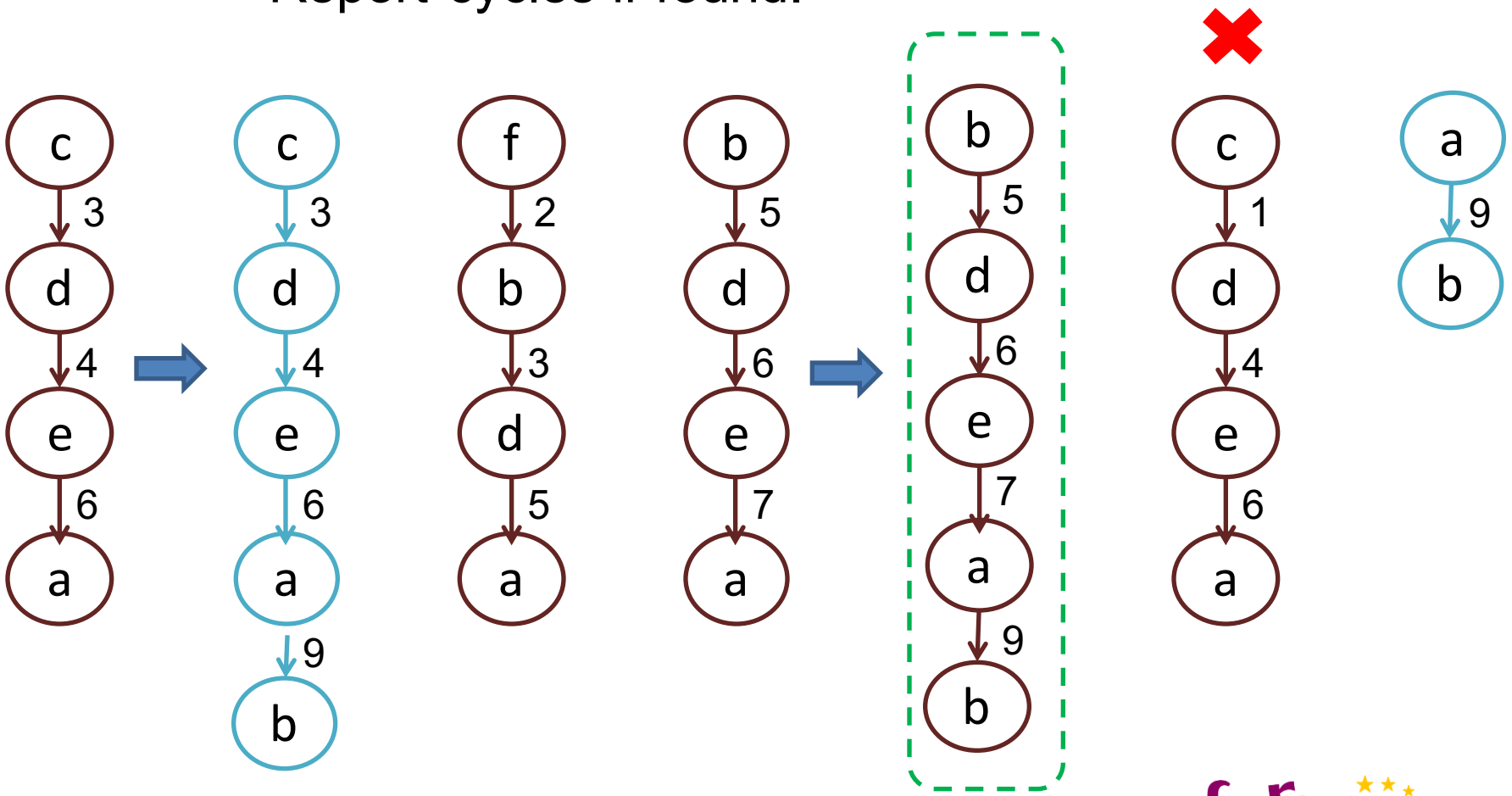


Naïve algorithm

Window = 7 timestamps

$\overrightarrow{a, b, 9}$

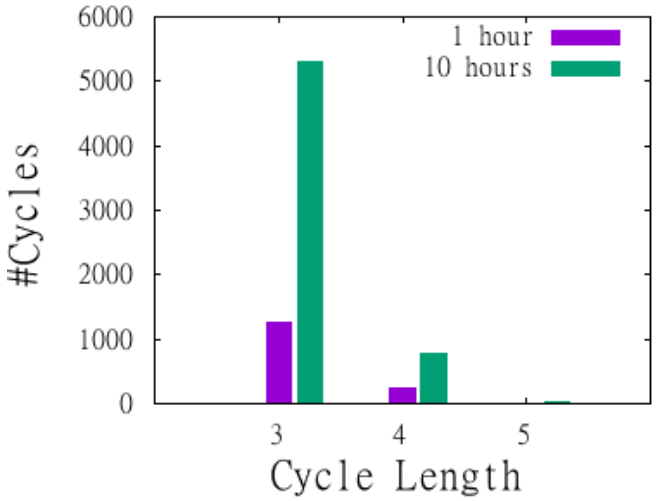
Step 2: Create new path by extending if possible
Report cycles if found.



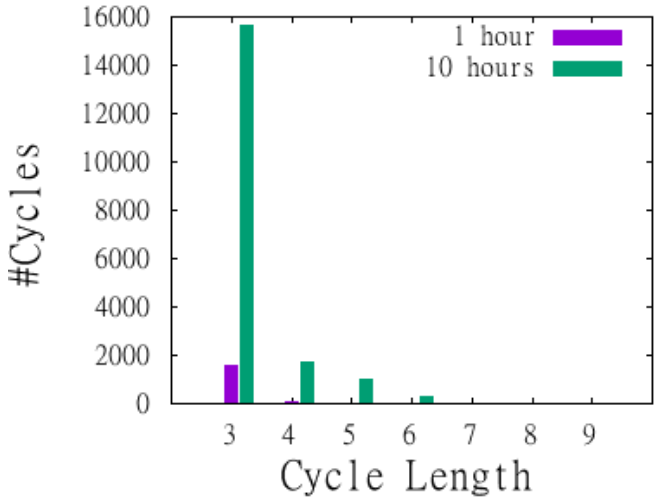
Performance

Dataset	#Nodes	Edges	Processing time (min)		Memory (MB)	
			w=1 hr	w=10 hr	w=1 hr	w=10 hr
Facebook WSON	46,952	876,993	6.4	7.0	14	19
Higgs Twitter	304,691	526,167	35.2	187.3	156	1815
SMS-A	44,100	545,000	11.3	50.9	29	777

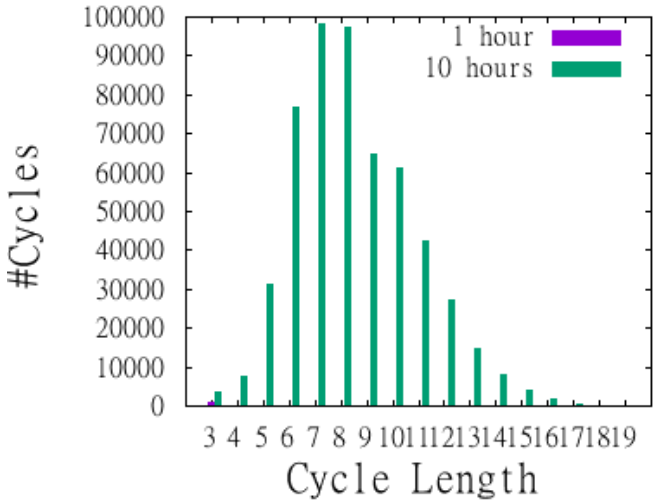
Cycle length Frequency Distribution



Facebook Chat



SMS Chat



Twitter retweet

Current work

- Analyze the text of tweets/retweets in the temporal cycles
  - Most of them are advertisements and the users looked like fake accounts used to promote message on social network.
- Smart algorithm using 2 or 3 pass over the data.
  - 1st Pass: get all the root nodes and edges which might be part of a cycle.
  - 2nd Pass: running smart DFS using the information gained in first pass.
- Using cycle length distribution to categories temporal network.

The art and science of asking questions is the source of all knowledge
- Thomas Berger