Εθνικό Μετσόβιο Πολυτεχνείο
National Technical University of Athens

Misinformation containment in social network platforms

Diploma Thesis

Iliana Maria Xygkou
Supervisor: Symeon Papavassiliou

Athens, June 2022

Presentation structure

- Basic concepts
- Literature review
- Cautious Misinformation Minimization (CMM) Problem
- Hardness of problem
- Solution algorithms
- Experiments
- Conclusion

Basic concepts

- A social network is represented mathematically with a graph $G = (V, E)$, where V the set of vertices which correspond to the users and E the set of edges which correspond to the connections between the users.
- Complex Networks are those which are used the most in everyday life in the form of telecommunication networks, social networks, biological networks, etc., and are represented mathematically with artificial networks.
- The Social Networks can be simulated with great accuracy by the Scale-Free Networks.

Scale-Free Networks characteristics

- The node degree follows the power-law distribution – few nodes are connected with a lot of nodes while a lot of nodes are connected with few nodes.
- The average path length is small – the distance between two nodes is small.
- The clustering coefficient is relatively high – the neighbors of the nodes create cliques between them.

Node centralities

They indicate how important a node is in the network.
The most significant of them are:

- Degree centrality: $$C_D(u) = \frac{deg(u)}{|V| - 1}$$
- Closeness centrality: $$C_P(u) = \frac{|V| - 1}{\sum_{v \in V, v \neq u} d(u, v)}$$
- Betweenness centrality: $$C_B(u) = \frac{2 \cdot \sum_{s \neq u \neq t} \frac{\sigma_{st}(u)}{\sigma_{st}}}{(|V| - 1)(|V| - 2)}$$
- Eigenvector centrality: $$v(u) = \frac{1}{\lambda} \sum_{k \in V} A[u, k] v(k)$$

Edge betweenness centrality

It indicates how important an edge is to the diffusion of a piece of information in a network, if it's travelling through the shortest paths.

$$C_B(e) = \frac{2 \cdot \sum_{s \neq t} \frac{\sigma_{st}(e)}{\sigma_{st}}}{|V|(|V|-1)}$$

Information Diffusion

The most significant diffusion models:

- Independent Cascade (IC): Let's consider a set of initial adopters A (seed set). Considering discrete time steps $t \in Z$ at each time step t every node v who becomes active for the first time attempts only once to activate each out-neighbor w with success probability $p(v, w)$. The process stops at time step t', when no node becomes active.
- Linear Threshold (LT): Let's consider a set of initial adopters A (seed set) and a random according to the uniform distribution in the interval $[0,1]$ selection of thresholds $\theta(u)$: $u \in V$. Considering discrete time steps $t \in Z$ at each time step t every node u is activated if the following holds: $\sum_{w \in N_{in}(u)} b(w, u) \geq \theta(u)$. The process stops at time step t', when no node becomes active.

Influence of a node

- It shows how important a node is in a network with respect to the diffusion of a piece of information.
- It is defined as the expected number of nodes which will be active at the end given that the node u is the initial adopter of the idea, and is mathematically expressed as $\sigma(u, G)$.
- The exact calculation of the influence of a node is #P-Hard under the IC and LT models.

Literature review

The mitigation of misinformation comprises of 2 steps:

- Detection of the fake news item
- Prevention of the propagation of the fake news item

Towards the detection of fake news, machine learning tools are employed.

Towards the containment of its diffusion, algorithms which remove edges or nodes, or start the propagation of the corresponding true piece of information are used.

Node Blocking Methods

The goal is to find a set of nodes of constrained size, whose removal will result in the minimization of the propagation of the misinformation.

- Static methods: The blocking of the nodes happen in the beginning without taking into consideration the temporal evolution of the dissemination of the piece of information. They are computationally efficient, but lack in accuracy.
- Adaptive methods: The blocking of the nodes happens in time steps according to the evolution of the information’s propagation. They are more efficient than the static methods, but they are computationally expensive.

Edge Blocking Methods

The goal is to find a set of edges of constrained size, whose removal will result in the minimization of the propagation of the misinformation.

- <u>Source-ignorant methods</u>: The blocking of the edges aims at the containment of the flow of information in the network. They are not effective but can be applied in the general case.
- <u>Source-aware methods</u>: The blocking of the edges aims at the containment of the diffusion of the information given that it is initially propagated by a known set of nodes. They are more effective than the above ones, but they demand the accurate and fast detection of the seed set.

Clarification Methods

The goal is to select a set of nodes which will initiate the diffusion of the true news item so as to limit the number of nodes which will adopt the corresponding fake news item.

- <u>Campaign-oriented methods</u>: The nodes are selected with the goal of minimization of the diffusion of the misinformation, based on either the network's topology or in combination with the graph's structure, the individual behavior (preferences, personal benefit, location, etc.). The second case is more effective, but requires information associated with the users, which might not be available.
- <u>Protection-oriented methods</u>: The nodes are selected with the goal of preventing the adoption of the fake news item at least by a specific percentage of users.

Cautious Misinformation Minimization Problem (CMM)

- Given a social network expressed with a simple directed graph $G = (V, E, w)$, with w: $V^2 \rightarrow [0,1]$, where $w(u, v)$ the probability of propagation of information from node u to node v.
- Given I_T and I_F the classes of information of true and fake content respectively related to topic I. It's considered that the diffused news items are not competitive but independent.
- Given e: $V \rightarrow [0,1]$ the metric of expertise of a user with respect to the topic of the propagated news item, where the values 0 or 1 indicate the absolute ignorance or the ultimate expertise respectively.

Diffusion Models

- IC Model:

The expertise e of one user is considered in the following way:

Given X_u random variable such that: $X_u = \begin{cases} 1 & \text{αν ο χρήστης } u \text{ υιοθετήσει την είδηση.} \\ 0 & \text{αλλιώς.} \end{cases}$

Then, $p_T(u) = Pr\{X_u = 1 | i \in I_T\} = \max\{e(u), 1 - e(u)\}$, $p_F(u) = Pr\{X_u = 1 | i \in I_F\} = 1 - e(u)$

and the activation probability of a node u at time step $t+1$ by a newly activated at time step t in-neighbor $v \in N_{in}(u)$ is calculated as: $p_C(v, u) = w(v, u) \cdot p_C(u)$, $C \in \{T, F\}$.

This way, it's feasible to simulate the diffusion processes of news items which belong to the classes I_T και I_F by defining 2 new graphs $G_T = (V, E, p_T)$ and $G_F = (V, E, p_F)$ respectively, and by applying the steps of the known IC model.

- LT Model:

The expertise of a user is not considered. As a result, the evolution of the propagation of a piece of information in the network is with respect to dynamics the same regardless if the information belongs to the class I_T or the class I_F.

- Deterministic LT Model (DLT):

It’s different from the above probabilistic LT model because of the deterministic selection of the nodes’ thresholds θ. Considering the expertise e of a user, the thresholds for the acceptance of the information item $i \in I_T$ ή $i \in I_F$ are calculated as follows: $\theta_T(u) = \min\{e(u), 1 - e(u)\}$, $\theta_F(u) = e(u)$

This way, it’s feasible to simulate the diffusion processes of news items which belong to the classes I_T και I_F by defining 2 new graphs $G_T = (V, E, w)$ with threshold $\theta_T(u), \forall u \in V$ and $G_F = (V, E, w)$ with threshold $\theta_F(u), \forall u \in V$ respectively, and by applying the steps of the known LT model.

Definition of the CMM Problem

- Under the LT model:

Given the network $G = (V, E, w)$, the two information classes I_T, I_F and their corresponding diffusion processes under the LT model, the seed sets of the class I_T and the class I_F, S_T and S_F respectively, and a positive integer number k, the goal is to find a set of edges $E' \subseteq E, |E'| \leq k$ whose removal will cause the maximum possible reduction of the diffusion of the class I_F and the concurrent minimum possible reduction of the diffusion of the class I_T. Mathematically the problem is expressed as:

$$E' = \underset{E', E' \subseteq E, |E'| \leq k}{\arg\min} \{ \sum_{a \in S_T} \sigma(a, (V, E)) - \sum_{a \in S_T} \sigma(a, (V, E \backslash E')) + \sum_{b \in S_F} \sigma(b, (V, E \backslash E')) \}$$

, with objective function: $$f_{LT}(E') = \sum_{a \in S_T} \sigma(a, (V, E)) - \sum_{a \in S_T} \sigma(a, (V, E \setminus E')) + \sum_{b \in S_F} \sigma(b, (V, E \setminus E'))$$

- Under the DLT and IC models:

Given the network $G = (V, E, w)$, the two information classes I_T, I_F and their corresponding diffusion processes under the DLT or IC model, the seed sets of the class I_T and the class I_F, S_T and S_F respectively, and a positive integer number k, the goal is to find a set of edges $E' \subseteq E, |E'| \leq k$ whose removal will cause the maximum possible reduction of the diffusion of the class I_F and the concurrent minimum possible reduction of the diffusion of the class I_T. Mathematically the problem is expressed as:

$$E' = \underset{E', E' \subseteq E, |E'| \leq k}{\arg\min} \{ \sigma(S_T, (V, E)) - \sigma(S_T, (V, E \setminus E')) + \sigma(S_F, (V, E \setminus E')) \}$$

, with objective function: $f_{IC/DLT}(E') = \sigma(S_T, (V, E)) - \sigma(S_T, (V, E \setminus E')) + \sigma(S_F, (V, E \setminus E'))$

Hardness of the CMM Problem

The CMM Problem is NP-Hard under all 3 models:

- In the case of the IC and DLT models a sequence of reductions from NP-Complete Problems is conducted:

$$Max\ Cut\ \leq_P Max\ Directed\ Cut \leq_P\ Max\ Bisection \leq_P\ Bisection\ Width \leq_P CMM$$

- In the case of the LT model we define an instance of the CMM problem with $S_T = \emptyset$. Then, the objective function becomes simpler as:

$$f_{LT}(E') = \sum_{b \in S_F} \sigma(b, (V, E \setminus E'))$$

and we define an equivalent function, which is non-decreasing, submodular and non-negative:

$$h(E') = \sum_{b \in S_F} \sigma(b, (V, E)) - \sum_{b \in S_F} \sigma(b, (V, E \setminus E'))$$

The CMM problem requires the maximization of this function under the constraint in the solution's size, which is an NP-Hard problem.

Solution Algorithm

- Under the LT model:

Greedy iterative algorithm with at most k iterations.

In each iteration, the edge that satisfies the criterion below is removed:

$$e = (u, v) = \arg\max_{e \in E \setminus E_t} \{\Delta(e|E_t)\} = \arg\max_{e \in E \setminus E_t} \{f_{LT}(E_t) - f_{LT}(E_t \cup \{e\}\}$$

Point of interest: How can the value of the function f_{LT} be calculated efficiently and approximately?

By using live-edge graphs which are constructed as follows: Independently, for $\forall v \in V$ at most one incoming edge (u, v) is selected with probability $w(u, v)$, while no incoming edge is selected with probability $1 - \sum_{u:(u,v) \in E} w(u, v)$. This way, we create the graph $X = (V, E_X)$, where $E_X \subseteq E$ the set of the chosen (live) edges. In this case, the diffusion process initiated by an initial node u is deterministic and consists of all the paths beginning from node u and containing live edges. Given $r(u, X)$ the number of nodes which are part of these paths, then:

$$\sigma(u, G) = \mathbb{E}_X[r(u, X)] = \sum_{X \in \mathscr{X}_G} Pr[X|G] \cdot r(u, X)$$

The following approximate formula is used:

$$\sigma(u, G) \approx \bar{\sigma}(u, G) = \frac{1}{|\mathscr{X}_S|} \cdot \sum_{X_i \in \mathscr{X}_S} r(u, T^u_{X_i})$$

, where $\mathrm{X}_S = \{X_i \colon 1 \leq i \leq x_S\} \subseteq X$ the set of the sampled live-edge graphs of the graph G and $\mathrm{T}^u_{\mathrm{X}_i}$ the tree with root the node u which ensues from the execution of BFS on the graph X_i.

- Under the DLT model:

Greedy iterative algorithm with at most k iterations.

In each iteration, the edge that satisfies the criterion below is removed:

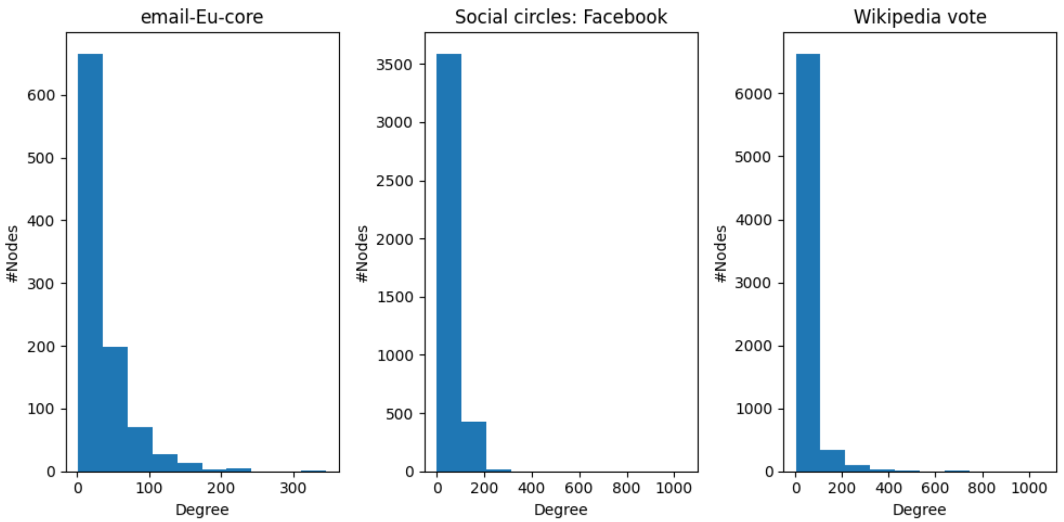
$$e = (u, v) = \arg\min_{e \in E \setminus E_t} \{f_{DLT}(E_t \cup \{e\})\}$$

The calculation of the function f_{DLT} is executed deterministically in linear time.

Experiments

- 3 real social networks:
  - email-Eu-core: email exchange inside a university.
  - Social circles: Facebook: friends lists on Facebook.
  - Wikipedia vote: ballots for the election of administrators on the Wikipedia platform.
- They can be categorized as scale-free networks:

Δίκτυο	Γράφος	Κόμβοι	Ακμές	Μέσο μήκος μονοπατιού	Μέσος συντελεστής ομαδοποίησης
email-Eu-core	Κατευθυνόμενος	986	16064	2.5869	0.4071
Social circles: Facebook	Μη κατευθυνόμενος	4039	88234	3.6925	0.6055
Wikipedia vote	Κατευθυνόμενος	7115	100762	3.247	0.1409



- Comparative methods:
  - Random: Removal of k random edges, which start from a node included in the seed set of the fake information item.
  - Weighted: Removal of k edges, which start from a node included in the seed set of the fake information, with the highest probability of fake information transmission, $w(u,v)$.
  - DistanceDiff (only under the LT model): Removal of k edges with the lowest difference of the distance from seed set of only the true information item, from the distance from the seed set of only the fake information item.

$$diff(u,v) = \min_{sf \in S_F \setminus S_T} \{d(sf,u)\} - \min_{st \in S_T \setminus S_F} \{d(st,u)\}$$

  - EdgeBetweennessDiff (only under the DLT model): Removal of k edges with the greatest difference of the edge betweenness centrality in the induced true information diffusion graph from the edge betweenness centrality in the induced fake information diffusion graph.

$$diff(e) = C_{B,false}(e) - C_{B,true}(e)$$

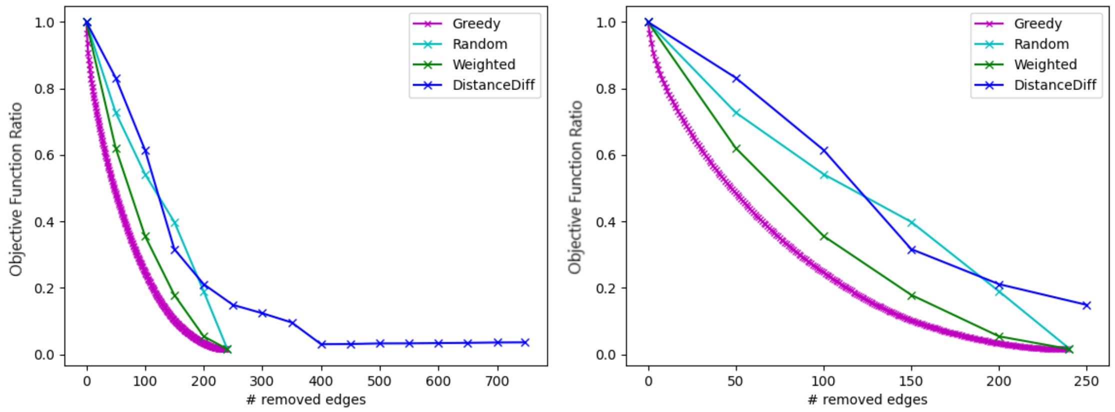
- The sets S_T, S_F are selected randomly according to the uniform distribution with size $|S_T| = |S_F| = \lceil 1\% \ \cdot \ |V\,| \rceil$.
- The integer k is defined as $k = \lceil 3\% \cdot |E| \rceil$.
- Under the LT model, we select $x_S = 5000$.
- The performance of the algorithms is measured after the removal of k edges as:

$$ratio(k) = \frac{f_{LT}(E')}{f_{LT}(\emptyset)}, \text{ με } |E'| = k$$

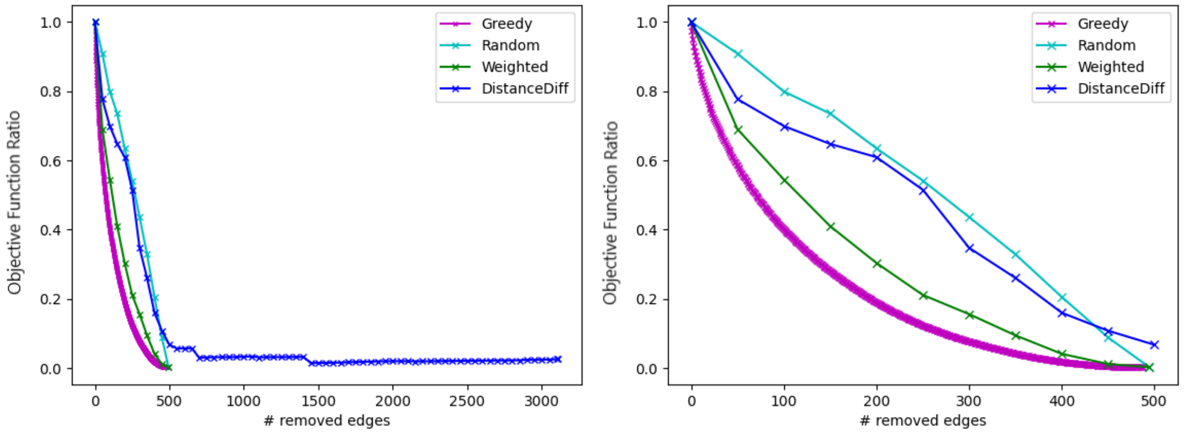
or

$$ratio(k) = \frac{f_{DLT}(E')}{f_{DLT}(\emptyset)}, \text{ με } |E'| = k$$

Results under the LT Model

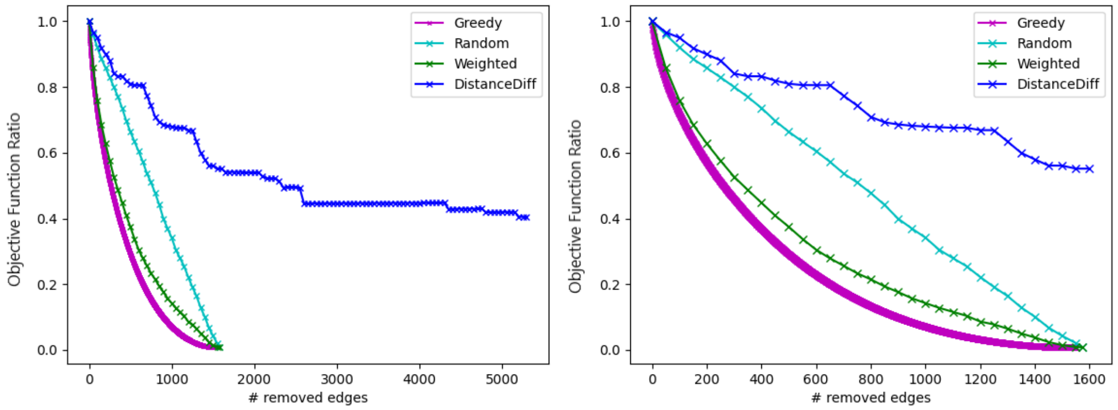


email-Eu-core



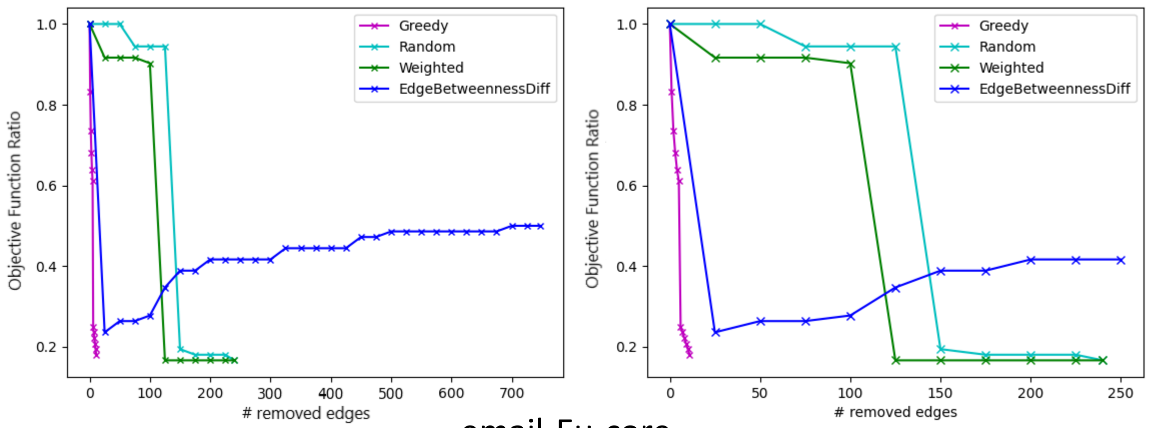
Wikipedia vote

For each figure, the chart on the left side depicts the complete results, while the one on the right side focuses on the interval of x-axis until the number of the edges which were removed through the application of all the methods except for "DistanceDiff", in order to observe the methods' impact on tackling the problem under study in greater detail.

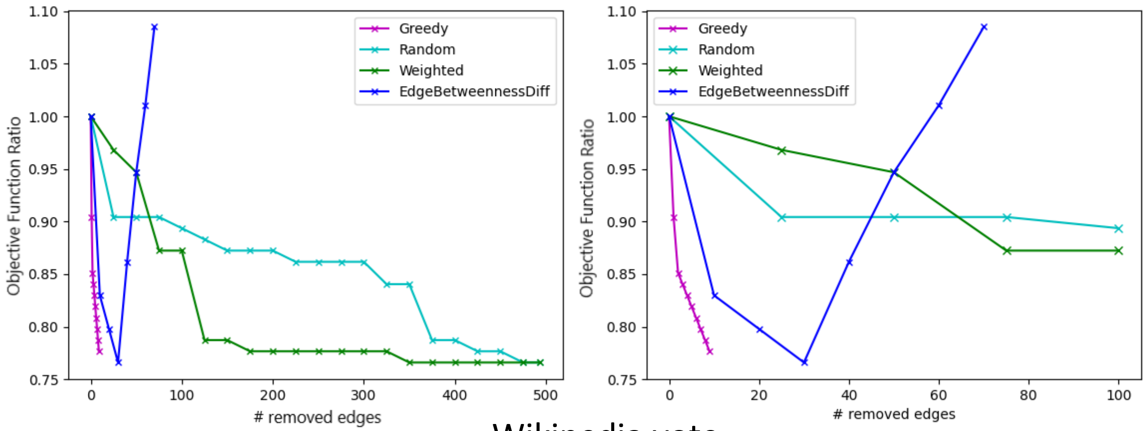


Social circles: Facebook

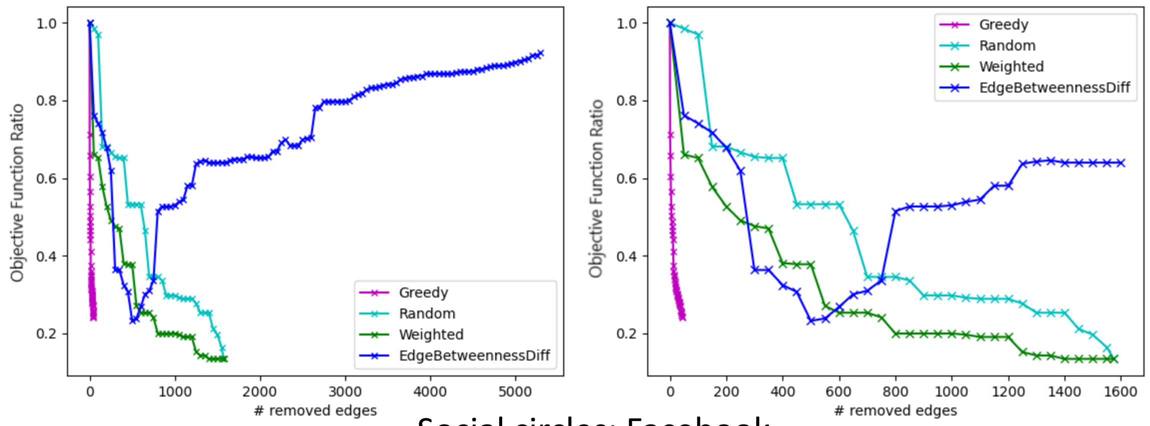
Results under the DLT Model



email-Eu-core



Wikipedia vote



Social circles: Facebook

Conclusion

- The proposed greedy methods surpass the rest of the comparative methods, since the objective function is decreased significantly with the removal of fewer edges.
- The “DistanceDiff” and “EdgeBetweennessDiff” methods, which exploit primarily the topology of the network, are not sufficiently efficient.
- In the future it seems appealing to:
  - Study experimentally the IC model too (in the context of this thesis the computational load was found prohibitive).
  - Learn the parameters of the diffusion models instead of assigning random values.
  - Study the dynamic networks besides the static ones.
  - Find the suitable stopping criterion of the “EdgeBetweennessDiff” method under the DLT model, since it seems to be effective when only a few edges are removed.

Thank you for your attention!