UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat d'Informàtica de Barcelona

# A GAME THEORY APPROACH TO SET THRESHOLDS FOR INFLUENCE SPREAD

JAYA GARCÍA FERNÁNDEZ

**Thesis supervisor**
MARIA JOSE SERNA IGLESIAS (Department of Computer Science)

**Degree**
Master's Degree in Innovation and Research in Informatics (Advanced Computing)

**Master's thesis**

**Facultat d'Informàtica de Barcelona (FIB)**

**Universitat Politècnica de Catalunya (UPC) - BarcelonaTech**

**28/01/2026**

**Acknowledgments**

I would like to express my deepest gratitude to my advisor, Maria, whom I could not have done this work without. All her cunning ideas and incredible support, have given me the perfect fuel to develop the research, and to constantly push myself to deliver greater results. Besides delivering a project that I am very proud of, she has shed more light into what being a researcher and a computer scientist means. She has taught me how to look closer into details, and that there is always a conclusion lying beneath the thick surface of a theorem.

I would also like to thank my previous advisor Carme, who helped me finish my degree thesis. Without her inspiration and teachings, I would not have chosen to take a masters, and I am truly thankful for having done so, as it has shown me many things, not only from Computer Science. Her view of the world, has helped me greatly in understanding a bit more mine. She taught me to be more rigorous, and to always have a solid argument behind my ideas.

Finally, I want to dedicate this work to my family and all my friends, who have always been and are, the greatest support for me. Within five minutes and some laughs, I can always forget all my problems easily, and they give me the inspiration and strength to finish the challenges I put on myself. It is a blessing to have people that help me when I'm at my low and that inspire me to improve and be better.

**Abstract**

Social networks are the natural space for the spreading of information and influence and have become a media themselves. Several models capturing that diffusion process have been proposed, most of them based on the Independent Cascade (IC) model or on the Linear Threshold (LT) model. Although the LT-based models contemplate an individual threshold for each actor in the network, the existing experimental studies so far have almost always considered identical thresholds for all the actors. Our main objective in this work is to initiate the study on how the dissemination of information on networks behaves when those thresholds are set strategically by the actors. For doing so, we plan to design different non-cooperative games incorporating characteristics of the game, and analyze problems related to the expansion of influence in their Nash equilibrium network or by the use of dynamics on the best response graph.

**Key words:** Influence Network, Influence Maximization, Target Set Influence, Linear Threshold, threshold, Game Theory, Best Response, Game Dynamics, Potential Games, Nash Equilibrium.

**Resum**

Les xarxes socials han esdevingut l'espai natural per difondre informació i influencia. Diferents models han sigut proposats per representar aquest procés de difusió, basats majoritàriament en l'Independent Cascade (IC) o el Linear Threshold (LT). Tot i que els models basats en LT consideren un llindar individual per cada actor en la xarxa, els estudis experimentals fins ara han considerat un llindar idèntic per a tots els actors. El nostre objectiu principal és iniciar l'estudi en com la divulgació d'informació en xarxes es comporta quan els llindars són escollits estratègicament pels actors. Per a fer això, planegem dissenyar diferents jocs no cooperatius, incorporant característiques dels jocs i analitzant problemes relacionats amb l'expansió d'influència en les seves xarxes en equilibri de Nash o utilitzant el graf de la dinàmica de la millor resposta.

**Paraules Clau:** Xarxes d'Influència, Maximització d'Influència, Maximització d'Influència a un Objectiu, Llindar Lineal, Llindar, Teoria de Jocs, Millor Resposta, Dinàmica de Jocs, Jocs Potencials, Equilibris de Nash.

**Resumen**

Las redes sociales se han convertido en el espacio natural para difundir información e influencia. Varios modelos han sido propuestos para representar este proceso de difusión, basados mayoritariamente en el Independent Cascade (IC) o el Linear Threshold (LT). Aunque los modelos basados en LT consideran un umbral individual para cada actor en la red, los estudios experimentales hasta ahora han considerado siempre un umbral idéntico para todos los actores. Nuestro objetivo principal es iniciar el estudio de cómo la diseminación de información en redes se comporta cuando los umbrales son escogidos estratégicamente por los actores. Para ello, planeamos diseñar diferentes juegos no cooperativos, incorporando características del juego y analizando problemas relacionados con la expansión de influencia en sus redes en equilibrio de Nash o usando el grafo de la dinámica de mejor respuesta.



**Palabras Clave:** Redes de influencia, Maximización de Influencia, Maximización de Influencia a un Objetivo, Umbral Lineal, Umbral, Teoría de Juegos, Mejor Respuesta, Dinámica de Juego, Juegos Potenciales, Equilibrio de Nash.

**Disclaimer**

This thesis has been prepared with limited assistance of generative AI tools. The text has been written by us, and the usage of AI was limited strictly to improving its grammar and vocabulary. Regarding the code, we used it for debugging purposes, to make the code more readable, and to help us with the generation of the plots, especially to make them prettier. For the proofs, the definition of the models and theorems and the analysis of the obtained results, no generative AI tool was used.

# Contents

# 1 Introduction

Since the dawn of time, rumors and gossip have held power in societies. They have been shown to shift the political opinion of countries [1], impact the economic market [2], and disregard facts such as climate change [3]. This is more evident today, with social networks, where users' opinions can spread quickly with little restraint on the content of their messages. One of the key points of spreading rumors is influence. We, as humans, are more inclined to believe the opinions of people whom we admire and find familiar [4], and that is what usually causes harmful information to spread through a society. The more influence one has over others, the easier it will be to convince them.

One way to represent this effect is with the Influence Network model, which simulates the spread of influence, or rumors, on a graph. Different diffusion processes can be defined to simulate spread, such as probabilistic models, for example, Independent Cascade [5], epidemic models, namely SIS or SIR [6], physical models based on energy concepts [7], and many more. Each of these centers around different aspects of the spread. We will focus on the Linear Threshold Model (LT) [8, 9], which represents the effect of social pressure, where agents are more likely to be convinced if a proportion of their influencers are convinced. In this model, we have a *threshold* value, different for each agent, that represents how sensitive they are to peer pressure.

There are many interesting problems that can be defined in these models. Domingos et al. [10] introduced the Influence Maximization (IM) problem, where, given a social network, the goal is to find the minimum set of *seed nodes* such that starting the spread from them maximizes the influence to the rest of the network. This problem is the basis for different real life challenges, such as finding key users in product promotion, analyzing epidemic expansion, identifying influences in rumor spreading in electoral campaigns or social networks, the diffusion of technological innovations to find key users, and many more [11].

The applications go far beyond rumor spreading, and for this reason, it is a widely studied model. In [9] it was proven that for a random uniform distribution of thresholds, the IM problem is NP-hard and the authors give a $(1-1/e)$-approximation greedy algorithm. Later, it was proven that the exact computation of the spread when the thresholds are random is #P-hard [12], which makes the original solution proposed by Kempe et al. inefficient. Different greedy algorithms were designed to avoid this complexity issue and make the computation more efficient [13, 14, 15]. Subsequent works are based on different techniques to improve the running time and approximation of the results in static and dynamic networks [16, 17], or variations of the original problem, such as IM of targeted sets of nodes [18, 19], or budgeted versions [20]. Another different branch of studies, is the usage of centrality measures to compute these seed sets and the definition of new centrality measures related to the spread of influence [21, 22] to solve the problem of IM.

Many studies have been conducted on the IM problem. However, in [23] it was observed that for the LT model, most of the experimental studies assigned equally fixed thresholds for all nodes in the network. This is a reasonable decision, since it helps simplify the studies. Nevertheless, it is not realistic, since in social networks different users will have different interests and different levels of influence. It is safe to assume that non-uniform distributions of thresholds are more likely in real networks. Furthermore, different players will be influenced differently by different agents. A very recent work [24], introduces a new framework that generalizes the LT model to an arbitrary distribution of weights and thresholds. The authors prove that a greedy approximation algorithm can also solve the IM problem in this more realistic context. However, to our knowledge, no work has been done to investigate how fixed threshold values affect influence, especially when the thresholds are selected strategically by the agents to help or fight against the influence, and which role the thresholds play in these scenarios.

For this work, we will focus on a variation of the IM problem, the Target Set Influence (TSI) problem. Given a set of nodes called *target nodes*, we want to find the minimum initial seed that maximizes the influence of the target set. Just like IM, this is also an NP-hard problem [9], therefore we will look for approximate solutions.

Our goal is to make an experimental study on two models that solve the TSI problem by using concepts from Algorithmic Game Theory [25]. Game Theory is a branch of mathematics that models situations as competitive or cooperative games, where different players act against or along each other to reach a certain goal [26]. The motivation behind using these models is that they are helpful in representing social interactions between different agents. We define two models; first, the Seed Selection Game, where the agents aim to find an initial minimal set of influence, and second, the Threshold Selection Game, where the agents aim to select a value for their threshold that allows for the influence to happen. First, we prove that both games are potential games and give theoretical properties on the models, and then we use them to run experiments on different social networks and obtain metrics from the influence topology. Comparing the results of both models will allow us to see which impact the thresholds have on the influence. Our study will be a simplified one, we will only consider graphs with unweighted edges, since our main goal is to put a focus on the thresholds and to initiate a study on the role they have in the TSI problem with LT expansion.

# 2 Preliminaries

Some concepts must be known to understand this work. In this section, we will talk about strategic game theory and non-cooperative games and the model of influence on social networks, precisely the Linear Threshold Model.

## 2.1 Strategic Games

Strategic Games are a theoretical framework that model situations as competitive games between players. Each player, or agent, has its own selfish goal, which is described through a preference function over the state of the game, that they aim to minimize or maximize. If the function is to be maximized, we will call it a *utility* function; otherwise, it will be a *cost* function. To achieve their goal, they choose a strategy from a set of possible actions they can take. The key point is that each player chooses their strategy without knowing what the others will do. Formally, a *strategic game* is a tuple $\Gamma = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$, where:

- $N = \{1, ..., n\}$ is a finite set of players.
- For each player $i \in N$, we have a finite set of non-empty actions $A_i$ that the player can take. The action chosen from this set is called the *strategy* of the player $s_i \in A_i$.
- A strategy profile $s = (s_1, ..., s_n)$ is the vector with the strategies of each agent. It describes the state of the game. We denote with $S_i$ all the possible strategies of player $i \in N$ and with $S = S_1 \times ... \times S_n$ the set of all possible strategy profiles in the game.
- For each player $i \in N$, a utility or (cost) function $u_i : A_1 \times ... \times A_n \to \mathbb{R}$ is defined to measure its pay-off or benefit. We say that the *outcome* of an agent $i \in N$ in a strategy profile $s = (s_1, ..., s_n)$ is the utility or cost achieved by $i$, i.e., $u_i(s)$.

For each $s = (s_1, ..., s_n)$ and a strategy $s'_i \in A_i$, we denote a strategy change as:

$$(s_{-i}, s'_i) = (s_1, ..., s_{i-1}, s'_i, s_{i+1}, ..., s_n)$$

Note that in a strategy change, only one player is allowed to change the chosen action.

We say that a strategy $s^*_i$ for agent $i$ is dominant if for any strategy profile $s$ with $s_i \neq s^*_i$ we have:

$$u_i(s_{-i}, s^*_i) \geq u_i(s)$$

In other words, the strategy $s^*_i$ yields to $i$ a better utility (or cost) than any other strategy.

Given a strategy profile $s = (s_1, ..., s_n)$, the set of *Best Responses* for an agent $i \in N$ in this profile are those strategies that optimize the utility or cost function of the agent, provided that the others do not change their strategy. Formally,

$$BR(s_{-i}) = \{a_i \in A_i : u_i(s_{-i}, a_i) = \max_{a'_i \in A_i} u_i(s_{-i}, a'_i)\}$$

Notice that if we had a cost function, we would want to minimize instead of maximize. A Best Response is the best action that an agent can take, but sometimes we are interested only in a *Better Response*, which aims only to improve the outcome for the agent.

A *Pure Nash Equilibrium* (NE) describes a situation in which every agent is *happy*, meaning that each agent individually cannot improve their utility or cost function by changing its strategy. Formally,

A strategy profile $s = (s_1, ..., s_n)$ is a NE iff $\forall i \in N \;\; \forall s'_i \in A_i$ s.t. $s'_i \neq s_i \quad u_i(s_{-i}, s'_i) \leq u_i(s)$

Nash Equilibriums might not exist depending on the game definition. This concept was first introduced by John F. Nash [27].

The *Best or Better Response Dynamics* is the process of traversing the space of all strategy profiles for a given game. At each iteration of the process, one player chooses the action that yields a Best or Better response according to the current strategy profile. Formally, given a strategic game $\Gamma$, its corresponding dynamics graph $G = (V, E)$ is defined as:

- $V = S_1 \times ... \times S_n$, the set of all possible strategy profiles.
- $E = \{(s_{-i}, s_i), (s_{-i}, s'_i) : u_i(s_{-i}, s_i) < u_i(s_{-i}, s'_i)\}$, the set of improving changes between profiles.

This process follows a similar idea to that of a Local Search, where we traverse a space of solutions using local improving changes. The difference here is that the changes between states are related to agents individually improving their own utility or cost. Therefore, an improving strategy for an agent might be a decreasing strategy for another, which can lead to *Improving Response Cycles* (IRC) in the Dynamics. If the Dynamics does not have IRCs, we say it has the *Finite Improving Property* (FIP). Notice that the Dynamics is a solution search algorithm, it is not equivalent to letting the agents play the game.

If the Dynamics process reaches a node that has no out-going edges, then we will have reached a Nash Equilibrium, as no agent has any other action that will improve its current utility. The main heuristics applied to change between states are the Best Response or Better Response algorithms. If the cost function is defined correctly, the process follows the scheme shown in Pseudo-code 1. Here, the input functions are related to cost or utility.

**Algorithm 1** General Game Dynamics Algorithm

**Input**: set of agents $N$, set of actions $(A_i)_{i \in N}$, set of functions $(f_i)_{i \in N}$, initial profile $(s_i)_{i \in N}$
**Output**: Strategy profile $s$

**while** $\exists u \in N$ that can change to improving strategy **do**
    **for** $i \in N$ **do**
        $s_i \leftarrow$ `best_response(`$i, A_i, f_i$`)`
    **end for**
**end while**
**return** $s$

Every iteration of the first loop is called a round. The number of rounds played or the computational complexity of the algorithm will depend on, the definition of the game, the complexity of the better or best response algorithm, and the ability to avoid cycles. Notice that the number of steps needed to reach a Nash Equilibrium can be exponential, independently of the complexity of the Best or Better response. Therefore, the Game Dynamics strategy does not always stop in polynomial time.

We say a game is a *Potential Game* [28] if the incentives of all players to change strategy can be represented with a global function called potential function. Formally, a function $\phi : S \to \mathbb{R}$ is an exact potential function of game $\Gamma$ if

$$\forall i \in N \; \forall s \in S \; \forall s' \in A_i,\; u_i(s) - u_i(s_{-i}, s'_i) = \phi(s) - \phi(s_{-i}, s'_i)$$

In other words, variations in the utilities of the agents are captured exactly by the corresponding variation in the potential function.

Potential games have very useful properties on the Nash Equilibrium. A game is potential if and only if the Best Response graph is acyclic. Furthermore, any potential game has a PNE. Both properties can be easily proven [28]. Notice that if a game is potential, the Best Response Dynamics is equivalent to a Local Search over the different strategy profiles with the potential function as the heuristic.

To better grasp these concepts, we will review a classical Game Theory problem, the Prisoners Dilemma [29].

Two members of a criminal organization have been arrested and sent to prison. The police do not have enough evidence to convict them, hence they would have to charge them with a lesser crime and convict them to 1 year in prison. However, looking for a confession, they separate the prisoners and propose the following to both: if one of them confesses, she will be released while the other gets charged with 3 years. But there is a catch. If *both* partners testify against each other, they will both get 2 years in prison. With this in mind, they are left to consider what they want to do, taking into account that they cannot communicate with each other and they aim to minimize their time spent in prison individually.

Let us see the game theory formulation:

- $N = \{\text{Suspect 1, Suspect 2}\}$: we have two players, both partners.
- $A_1 = A_2 = \{\text{Quiet, Fink}\}$: they both have the same actions, either to stay quiet or to betray the other.
- $S_1 \times S_2 = \{(\text{Quiet, Quiet}), (\text{Quiet, Fink}), (\text{Fink, Quiet}), (\text{Fink, Fink})\}$: all possible strategy profiles in this game.
- Here we have a cost function, as they aim to minimize the number of years they get. This is not easy to summarize in a mathematical function, but Table 1 shows the cost for each agent in each possible case.

Table 1: Summary of the cost for each agent in each strategy profile and their Best Responses. The actions in blue imply that the curret strategy is the Best Response already.

| $a_1$ | $a_2$ | $c_1$ | $c_2$ | $BR(s_{-1})$ | $BR(s_{-2})$ |
|---|---|---|---|---|---|
| Quiet | Quiet | 1 | 1 | Fink | Fink |
| Quiet | Fink | 3 | 0 | Fink | Fink |
| Fink | Quiet | 0 | 3 | Fink | Fink |
| Fink | Fink | 2 | 2 | Fink | Fink |

Each row of the table corresponds to a strategy profile $s = (s_1, s_2)$, with its corresponding cost for each agent and also what their Best Response action would be, assuming the other agent does not change hers. In this case, the only time when both agents are happy is when they both fink, as no individual change in their strategy will give them a lower cost, and this is precisely the only Nash Equilibrium in the game. Notice that this game has no dominant strategy, since there is no action that yields minimum cost over the other actions of a player.

Regarding the Dynamics, we get a graph like the one shown in Figure 1. See how the state where they both fink does not have out-going edges, which means that this strategy profile is a Nash Equilibrium, as no agent can improve their cost. Furthermore, a Best or Better Response algorithm starting from any initial configuration always leads to a Nash Equilibrium, assuming they both could take a peek at what the other is thinking. Notice that this game does not have Improving Response Cycles.

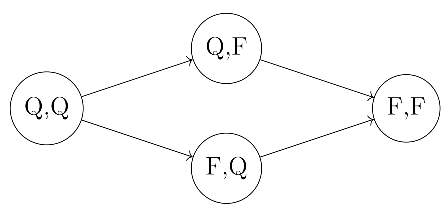


Figure 1: Game Dynamics graph of the Prisoners Dilemma. The labels in the nodes represent the current strategy profile $s = (s_1, s_2)$.

Lastly, it can be seen that the Prisoner Dilemma is a potential game, since the Best Response graph is acyclic. Furthermore, Table 2 shows an exact potential function for the game. It can be easily checked that this is a valid potential function with the cost table.

Table 2: Bi-matrix representation of the potential function (left) and the costs function (right) in the Prisoner's Dilemma Game. The rows are the actions of Player 1 and the columns the actions of Player 2.

| $\phi(s)$ | Quiet | Fink |
|---|---|---|
| Quiet | 3 | 2 |
| Fink | 2 | 1 |

| $c_1(s)\backslash c_2(s)$ | Quiet | Fink |
|---|---|---|
| Quiet | $1\backslash 1$ | $3\backslash 0$ |
| Fink | $0\backslash 3$ | $2\backslash 2$ |

## 2.2 Linear Threshold Model (LT)

An influence graph is a tuple $(G, w, f)$, where $G = (V, E)$ is a directed or undirected graph formed by a set of nodes $V$ and a set of edges $E$, $w : E \to \mathbb{N}$ is a weight function on the edges and $f : V \to \mathbb{N}$ is a labeling function that represents how easily each node is influenced. In this definition, node $i \in V$ influences another node $j \in V \setminus \{i\}$ with strength $w_{ij}$ if $(i, j) \in E$. Figure 2 shows a network of this type.

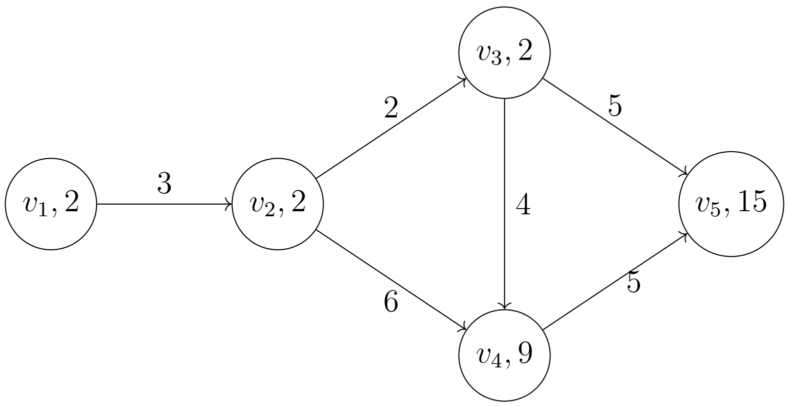


Figure 2: Example of an influence network. Each node $v$ contains the name of the node and its influence threshold $f(v)$ and the edges have the weighted value of their influence.

Given an influence graph $(G, w, f)$ and an initial influence set, or seed set, $I \subseteq V$, the *spread of influence* follows an iterative process in the Linear Threshold Model. In this setting, each node can either be active or inactive, and when they are active, they spread influence throughout their neighbors. The LT model assumes that an active node cannot go back to being inactive [9]. We use influenced and activated interchangeably. Let $F_t(S) \subseteq V$ be the set of nodes activated at some iteration $t$ of the algorithm. Initially, at step $t = 0$, only the nodes in $I$ are activated, i.e. $F_0(I) = I$. At iteration $t + 1$, node $i \in V$ is activated if and only if

$$\sum_{j \in F_t(I)} w_{ji} \geq f(i)$$

In other words, a node $i$ is influenced when the sum of weights of its incoming activated neighbors exceeds the threshold of the node. This process will stop when no more activation occurs in consecutive iterations. Formally, the set of influenced nodes by $I$ will be:

$$F(I) = \bigcup_{t=0}^{k} F_t(I) = F_0(I) \cup \ldots \cup F_k(I)$$

Where $k = \min\{t \in \mathbb{N} : F_t(I) = F_{t+1}(I)\}$. Notice that $k \leq |V|$, as at most we can have $|V|$ influenced nodes.

It is clear to see that $F$ is a monotonically increasing function. Given any node $v \in V$ and a set $I \subseteq V$, such that $v \notin I$, $F(I) \subseteq F(I \cup \{v\})$. This is true, as adding one node more to the initial set, either makes the influence the same as it was without it, or it increases the number of nodes influenced. Notice that the number of influenced nodes will never decrease with the addition of new nodes to the seed set, as the weight influence is additive and the weights are positive.

Let us see how this process works with the example on Figure 3.

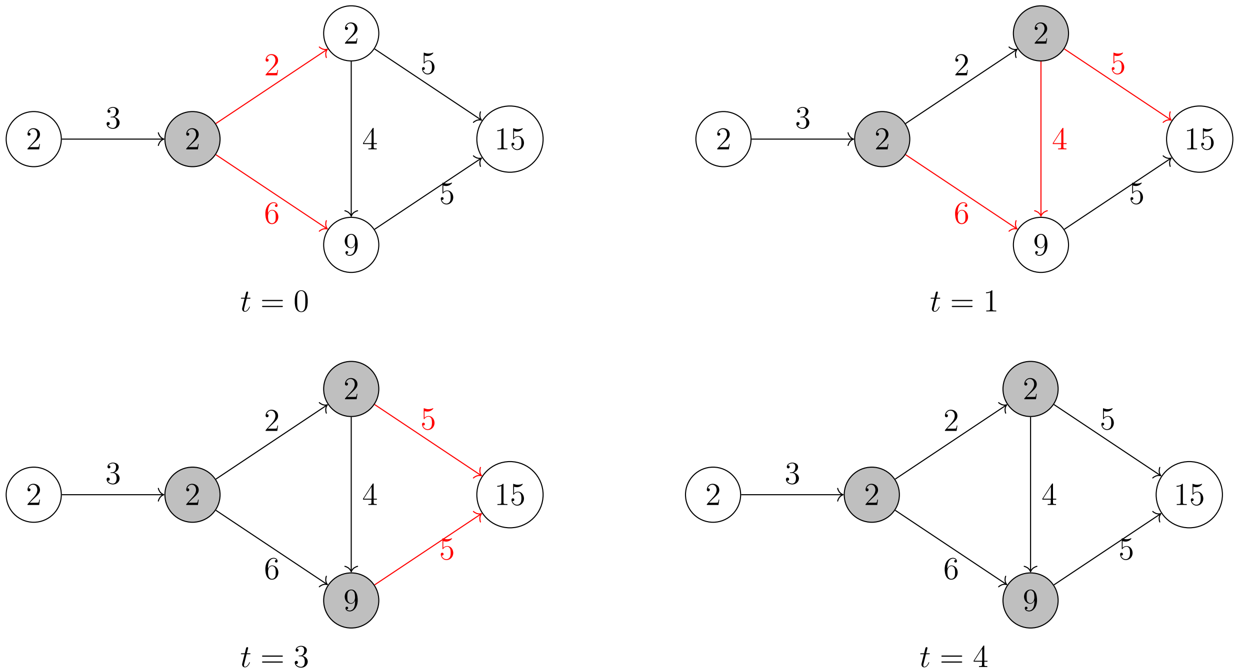


Figure 3: Example of the influence spread process with initial set $I = \{v_2\}$. The gray nodes are those in the set $F_t(I)$ and the edges in red are influences that have not been activated yet at time step $t$. The node names have been removed for better readability.

The influence propagates only when the sum of weights of the incoming influenced edges are greater than the thresholds. In this case, $F_3(I) = F_4(I)$, therefore the diffusion process stops at time $t = 3$ as no more nodes will be influenced after this.

Notice how, depending on the given influence graph, some nodes will never be convinced. In this case, node $v_5$ needs the sum of the weights of its influences to be greater than 15, but that will never happen since the highest influence she can receive is 10. Node $v_1$, on the other hand, will only be influenced if it is part of the initial set, as it has no incoming influence.

The algorithm for computing $F_t(I)$ can be seen in pseudo-code 2.

**Algorithm 2** Computing the influence set $F_t(I)$ [30]

**Input**: $G = (V, E)$, $w : E \to \mathbb{N}$, $f : V \to \mathbb{N}$, $I \subseteq V$
**Output**: $T \subseteq V$

```
total ← 0
Q ← ∅
for all v ∈ V \ I do
    v.influenced ← FALSE
    v.influence ← 0
    v.lsl ← −1
end for
for all v ∈ I do
    v.influenced ← TRUE
    v.influence ← 0
    v.lsl ← 0
    total ← total + 1
    PUSH(Q, v)
end for
while Q ≠ ∅ do
    v ← POP(Q)
    for all u ∈ E(v) do
        if not u.influenced then
            u.influence ← u.influence + f_i(w(v, u), v, u)
            if u.influence ≥ f(u) then
                u.influence ← TRUE
                total ← total + 1
                PUSH(Q, u)
            end if
            u.lvl ← v.lvl + 1
        end if
    end for
end while
return total
```

This algorithm was implemented in [30] and has been proven to have a cost of $O(n + m)$ as the input graph is assumed to be represented as an adjacency list. The main idea is to focus the spread on the influenced nodes. For each influenced node, we activate their spread to their neighbors, updating a variable on the number of incoming influences of a given node. When this threshold is exceeded, the node is set to be influenced and added to the queue. This process continues to iterate until there are no more nodes left in the queue. At each step, we can have at most $n$ nodes in the queue and we traverse all their neighbors; therefore, we get the cost of $O(n + m)$.

# 3 Target Set Influence Maximization

In this section, we present the models to solve the Target Set Influence (TSI) maximization problems in different contexts. First, we present a model that solves the problem of finding an initial influence set, and then we present a model that makes the agents choose a threshold such that influence happens. We will assume throughout the rest of this work that the weights on the edges are one, i.e., $\forall e \in E, w(e) = 1$. This choice is made to simplify the analysis to have some base study for starters. Notice that the algorithms presented in previous sections work the same when the edge weights are one. Since this is an experimental study, we are not interested in improving the efficiency of existing algorithms for the TSI maximization problems, therefore we will not argue if our algorithms could be improved nor if they are better than the state of the art.

## 3.1 The Problem

Let $G = (V, E, f)$ be a directed network, where $V$ is the set of nodes, $E$ the set of edges, and $f : V \to \mathbb{N}$ the threshold function.

As mentioned, we will focus on the Target Set Influence problem. Given a subset $T \subseteq V$, we want to find the minimum set $I \subseteq V \setminus T$ such that the start of the influence from these nodes maximizes the influence of the nodes in $T$. Finding the minimum set is NP-hard [9], so we might be more interested in finding an approximate solution. We say that a set of nodes $I$ is of *minimal influence* if

$$\forall u \in I,\ |F(I - \{u\}) \cap T| < |F(I) \cap T|$$

In other words, all nodes in $I$ are critical for influence, as eliminating one of them would decrease the number of nodes convinced in $T$ by at least one. We say that a minimal influence set is optimal if it minimizes the size of the set in comparison to other solutions. We will not discuss how far a minimal set is from a minimum set, since the properties of a minimal set are relevant enough for the study.

Notice that computing the influence of the nodes in $T$ can be done in polynomial time, since we only need to compute $F(I)$, which can be done in $O(n + m)$, and then compute the intersection $F(I) \cap T$, which can also be done in $O(\min(|T|, |F(I)|))$ in the worst case.

To clarify the notation, from now on we will use $n = |V|$ and $m = |E|$ for a given network $G = (V, E, f)$. The number of agents will be explicitly defined every time to avoid confusion.

### 3.2 Seed Selection Game

Given a graph $G = (V, E, f)$, a target set $T \subseteq V$, and a constant $\alpha \in \mathbb{R}$, we define the non-cooperative *Seed Selection Game* $\Gamma_I(G, T, \alpha)$ to model the competitiveness of seed selection with respect to finding an initial influence set. To simplify the notation, we will use $\Gamma_I(G, T, \alpha)$ and $\Gamma_I$ interchangeably. The game $\Gamma_I$ is defined as follows: there is one player for each node in $V \setminus T$. Each of them has two actions $A_u = \{0, 1\}$ and they strive to minimize the following cost function:

$$c_u(s) = |T| - |F(I_s) \cap T| + \alpha s_u$$

where $u \in V \setminus T$, $I_s = \{u | s_u = 1\}$ and $F(I_s)$ is the set of influenced nodes in the outcome of the spread. Conceptually, the players are the set of nodes that don't belong to the target set, and they can choose if they want to belong to the initial set or not. Their goal is to minimize the number of non-influenced nodes in $T$, while keeping the initial set small enough. To have a better picture, we will first analyze the properties of our model and then the optimality of its solutions.

**Theorem 1.** Let $\Gamma_I$ be a Seed Selection Game. Given a strategy profile $s$ and an agent $u \in V \setminus T$, $BR(s_{-u})$ can be computed in $O(n + m)$ time in the game.

*Proof.* Consider the Algorithm given in Pseudo-code 3.

**Algorithm 3** Best Response for a player in the game $\Gamma_I$

**Input**: $G = (V, E, f)$, $s = (s_1, ..., s_{|V \setminus T|})$, $u \in V \setminus T$
**Output**: $s_u \in \{0, 1\}$

```
if c_u(s_{-u}, 0) < c_u(s_{-u}, 1) then
    return {0}
else if c_u(s_{-u}, 0) > c_u(s_{-u}, 1) then
    return {1}
else
    return {0, 1}
end if
```

**Correctness:** This is a very simple scheme where we apply a brute-force trying first one action and then the other, and seeing which one yields the better cost. As we explore all the possible actions, it is clear that we will obtain the Best Response.

**Cost:** Since the algorithm only computes the cost function, we must see if the cost can be computed in polynomial time. $F(S)$ can be computed in $O(n+m)$ time and the intersection between $F(I_s)$ and $T$ can also be computed in $O(\min(|T|, |F(I_s)|))$. It is also clear that the number of elements in the sets and the addition of $\alpha s_u$ can be computed in constant time; hence the best response can be computed in $O(2(n + m)) = O(n + m)$, which is polynomial. □

**Theorem 2.** Let $\Gamma_I$ be a Seed Selection Game. For any given strategy profile $s = (s_1, ..., s_{|V \setminus T|})$, the function

$$\phi(s) = |T| - |F(I_s) \cap T| + \alpha |I_s|$$

is an exact potential function for the game $\Gamma_I$.

*Proof.* Fix a strategy profile $s = (s_1, ..., s_{|V \setminus T|})$ and an agent $u \in V \setminus T$. To prove this, we need to show that any change in the cost function between different strategies of a given player $u \in V \setminus T$ is the same change in the potential function. This can be done by comparing both the cost and the potential functions and seeing that they grow in the same way.

$$\begin{aligned} c_u(s) &= |T| - |F(I_s) \cap T| + \alpha s_u \\ \phi(s) &= |T| - |F(I_s) \cap T| + \alpha |I_s| \end{aligned}$$

Since there are only two actions, the easiest way to show this is by proving all possible cases. We will only prove one case, as the other is symmetric.

$$\begin{aligned} c_u(s_{-u}, 0) - c_u(s_{-u}, 1) &= (|T| - |F(I_{(s_{-u},0)}) \cap T| + \alpha 0) - (|T| - |F(I_{(s_{-u},1)}) \cap T| + \alpha 1) \\ &= |T| - |T| - |F(I_{(s_{-u},0)}) \cap T| + |F(I_{(s_{-u},1)}) \cap T| - \alpha \\ &= |F(I_{(s_{-u},1)}) \cap T| - |F(I_{(s_{-u},0)}) \cap T| - \alpha \end{aligned}$$

$$\begin{aligned} \phi(s_{-u}, 0) - \phi(s_{-u}, 1) &= (|T| - |F(I_{(s_{-u},0)}) \cap T| + \alpha |I_{(s_{-u},0)}|) - (|T| - |F(I_{(s_{-u},1)}) \cap T| + \alpha |I_{(s_{-u},1)}|) \\ &= |T| - |T| - |F(I_{(s_{-u},0)}) \cap T| + |F(I_{(s_{-u},1)}) \cap T| + \alpha |I_{(s_{-u},0)}| - \alpha |I_{(s_{-u},1)}| \\ &= |F(I_{(s_{-u},1)}) \cap T| - |F(I_{(s_{-u},0)}) \cap T| + \alpha |I_{(s_{-u},0)}| - \alpha(|I_{(s_{-u},0)}| + 1) \\ &= |F(I_{(s_{-u},1)}) \cap T| - |F(I_{(s_{-u},0)}) \cap T| + \alpha |I_{(s_{-u},0)}| - \alpha |I_{(s_{-u},0)}| - \alpha \\ &= |F(I_{(s_{-u},1)}) \cap T| - |F(I_{(s_{-u},0)}) \cap T| - \alpha \end{aligned}$$

It is clear to see that changing from strategy $s_v = 1$ to $s_v = 0$ will yield the same result but with the signs changed. Seeing that the players are symmetric, we can conclude that this is an exact potential function for the game $\Gamma_I$.

□

**Theorem 3.** Seed Selection Games are potential games.

*Proof.* By Theorem 2 we know that a game $\Gamma_I$ accepts a potential function, therefore the game is a potential one.

□

Up until now we have proven very powerful properties of the game. From Theorem 1 we know that the Best Response given a strategy profile can be computed in polynomial time and from Theorem 3 we know that the Best Response Dynamics has no cycles and that a PNE always exists. Now it remains to see if a PNE can be reached in polynomial time with the Best Response Dynamics, as we will show.

**Theorem 4.** Let $\Gamma_I$ be a Seed Selection Game. Given an initial strategy profile $s = (s_1, ..., s_{|V \setminus T|})$, a minimal influence set $I \subseteq V \setminus T$ can be computed in $O(|V \setminus T|)$ steps with the Best Response Dynamics in the game $\Gamma_I$.

*Proof.* We know that $\phi(s) = |T| - |F(I_s) \cap T| + \alpha|I_s|$ is a potential function for the game $\Gamma_I$. Therefore, the cost function for the agents and $\phi(s)$ have the exact same shape and will be equally bounded. We know that given a strategy profile $s$, $|I_s| \leq |V| - |T|$, by definition of the game $\Gamma_I$. Therefore, $\phi(s) \leq O(|T| + \alpha|I_s|) \leq O(|T| + \alpha(|V| - |T|))$, which means that $\phi(s)$ is upper bounded by the number of agents in the game. This implies that a search using this heuristic cannot take more than $O(|V \setminus T|)$ steps.

□

We have analyzed the properties of our model, and we have seen that reaching a PNE in this game can be done in polynomial time with the Best Response Dynamics. What remains to see now, is whether we can use this process to compute a minimal set of influences. For this, we need to argue what happens for different values of the parameter $\alpha$ of the cost function. This constant can be seen as the relevance that the influence set has in the minimization function. With high $\alpha$, the players prioritize the initial set to be smaller, and with small $\alpha$ they prioritize the influence of the nodes. Since our priority is to maximize influence while keeping the set small, we need to study which values are best. To do so, we can analyze when one strategy is better than the other for a given agent $u \in V \setminus T$.

$$\begin{aligned} c_u(s_{-u}, 1) < c_u(s_{-u}, 0) &\iff |T| - |F(I_{(s_{-u},1)}) \cap T| + 1 \cdot \alpha < |T| - |F(I_{(s_{-u},0)}) \cap T| + 0 \cdot \alpha \\ &\iff \alpha < |F(I_{(s_{-u},1)}) \cap T| - |F(I_{(s_{-u},0)}) \cap T| \end{aligned} \tag{1}$$

As Equation 1 shows, participation in the set is a better option if and only if the difference between the number of nodes influenced with $s_v = 0$ and with $s_v = 1$ is strictly greater than $\alpha$. Depending on the values of $\alpha$, we will see different behaviors. We will not consider the case $\alpha = 0$, since then all agents maximize the same function and there will be no better or worse strategy for each agent, which removes the concepts of game theory from the model.

When $\alpha < 0$, the difference in 1 will always be greater, since $F$ is a monotonically increasing function and participating in the set means subtracting $\alpha$ from the cost. Therefore, $s_v = 1$ is a dominant strategy as being in the set will always produce a better cost for the agents than not being in it. Regarding the solutions, we reach the following Lemma.

**Lemma 1.** In a Seed Selection Game $\Gamma_I$, when $\alpha < 0$, the unique Nash Equilibrium is $\forall u \in V \setminus T$, $s_u = 1$.

*Proof.* Since $s_u = 1$ is a dominant strategy, for any agent $u \in V \setminus T$, the agents will always prefer it; therefore, the only time they are all happy is when they all choose to participate in the set. □

This Lemma implies that the set obtained through the Dynamics is $I_s = V \setminus T$ when $\alpha < 0$, which is feasible but trivial, since we are not minimizing the size of the initial set.

For $\alpha > 0$, we get a restricted version of the previous one. Staying in the set will only be a better option when doing so strictly influences more than $\alpha$ new nodes. The higher the value of $\alpha$, the less the agents will be eager to participate in the set, as that will result in a worse cost for them. For this case, we can show the following Lemma.

**Lemma 2.** In a Seed Selection Game $\Gamma_I$, when $\alpha > 0$, if a strategy profile $s$ is a Nash Equilibrium in game $\Gamma_I$, then $I_s$ is a minimal influence set.

*Proof.* Let us assume that there exists a strategy profile $s$ that is a Nash Equilibrium but such that $I_s$ is not a minimal influence set, and reach a contradiction. If $I_s$ is not minimal:

$$\begin{aligned}
\exists u \in I, |F(I) \cap T| = |F(I - \{u\}) \cap T| &\implies |T| - |F(I) \cap T| = |T| - |F(I - \{u\}) \cap T| \\
&\implies |T| - |F(I_{(s_{-u},1)}) \cap T| = |T| - |F(I_{(s_{-u},0)}) \cap T| \\
&\implies c_u(s_{-u}, 1) - \alpha = c_u(s_{-u}, 0) \\
&\implies c_u(s_{-u}, 1) = c_u(s_{-u}, 0) + \alpha \\
&\implies c_u(s_{-u}, 1) > c_u(s_{-u}, 0)
\end{aligned}$$

Since $\alpha > 0$, the last implication is true, hence we reach a contradiction as we assumed that $I_s$ is a Nash Equilibrium, but the action $s_u = 0$ is strictly better. □

One might assume that the reciprocal is also true. However, we see that for specific values of $\alpha$, it is not.

**Lemma 3.** In a Seed Selection Game $\Gamma_I$, when $\alpha \geq 2$, a minimal influence set $I \subseteq V \setminus T$, might not be a Nash Equilibrium in its corresponding strategy profile $s$.

*Proof.* Fix $\alpha = 2$ and take a look at Figure 4.

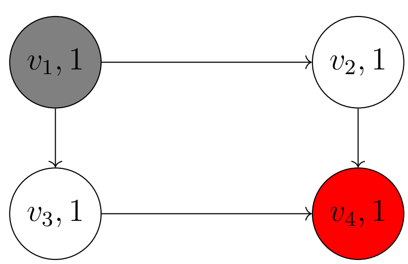


Figure 4: Example network for Lemma 3. Each node $v$ contains the name of the node and its influence threshold $f(v)$ and the edges have weight 1. $T = \{v_4\}$, $I_s = \{v_1\}$, $s = (1, 0, 0)$ and $\alpha = 2$.

It is clear to see that the set $I = \{v_1\}$ is of minimal influence, since the removal of $v_1$ from the set will decrease the number of influenced nodes by one. However, let us take a look at the costs of agent $v_1$ for each of its strategies.

$$c_{v_1}(s_{-v_1}, 0) = |T| - |F(I_s) \cap T| + \alpha 0 = 1 - 0 + 2 \cdot 0 = 1$$
$$c_{v_1}(s_{-v_1}, 1) = |T| - |F(I_s) \cap T| + \alpha 1 = 1 - 1 + 2 \cdot 1 = 2$$

Since the cost of removing itself from the set is lower than staying, the agent is not happy, hence this is not a Nash Equilibrium. By induction, we can see that higher values of $\alpha$ render the same result. □

See that for $\alpha \geq 2$, we can get solutions that are not feasible, since in some cases the agents would rather not participate in the set than to influence new nodes in the target set. This can lead to solutions where the minimization of the set if prioritized over the influence, which is not desirable. Another thing that can happen, is that the addition of a node to the set only influences new $\alpha$ nodes. In the previous example, if $\alpha = 1$, the cost of staying or leaving is the same, but leaving the set breaks the influence, which is not desirable. Notice that this can happen for $\alpha \geq 1$. Therefore, we need a finer value for $\alpha$, so that the Best Response for the agents is unambiguous.

We analyze now the case where $0 < \alpha < 1$. In this case, if we look at Equation 1, an agent $u \in V \setminus T$ will rather stay in the set when:

$$|F(I_{(s_{-u},1)}) \cap T| - |F(I_{(s_{-u},0)}) \cap T| \geq 1$$

Here the decision of staying or not is more clear than before, as the participation of a node at least influences one node, and if it does, the cost function will be lower when participating than when not. Regarding the Nash equilibrium, notice that the same result as Lemma 2 applies, as $\alpha$ is still strictly greater than 0. Furthermore, since $\alpha < 1$, we can now prove the following lemma to make the characterization stronger.

**Lemma 4.** In a Seed Selection Game $\Gamma_I$, when $0 < \alpha < 1$, if a set $I \subseteq V \setminus T$ is of minimal influence then it is also a Nash Equilibrium.

*Proof.* Let us assume that $I$ is a minimal influence set but it is not a Nash Equilibrium. If $I$ is not a Nash Equilibrium then

$$\begin{aligned} \exists u \in I, c_u(s_{-u}, 0) < c_u(s_{-u}, 1) &\implies |T| - |F(I_{(s_{-u},0)}) \cap T| < |T| - |F(I_{(s_{-u},1)}) \cap T| + \alpha \\ &\implies |F(I_{(s_{-u},1)}) \cap T| - |F(I_{(s_{-u},0)}) \cap T| < \alpha \\ &\implies |F(I) \cap T| - |F(I - \{u\}) \cap T| < \alpha \end{aligned}$$

As $0 < \alpha < 1$, the last inequality is true only when $|F(I) \cap T| = |F(I - \{u\}) \cap T|$, since these values are integers and $F$ is monotonically increasing. However, this means that $I$ is not minimal, as $u$ could be removed from the set. □

**Theorem 5.** In a Seed Selection Game $\Gamma_I$, when $0 < \alpha < 1$, a set $I$ is a Nash Equilibrium if and only if $I \subseteq V \setminus T$ is a minimal influence set.

*Proof.* The statement follows from Lemmas 2 and 4. □

Notice that this theorem can easily be generalized to any integer interval.

**Theorem 6.** In a Seed Selection Game $\Gamma_I$, for $k \in \mathbb{Z}, k \geq 0$, when $k < \alpha < k+1$, a set $I$ is a Nash equilibrium if and only if $I \subseteq V \setminus T$ is a minimal influence set.

*Proof.* Since $\alpha > 0$, Lemma 2 still holds. For the other direction, note that in the proof of Lemma 4, we relied on the difference $|F(I) \cap T| - |F(I - \{u\}) \cap T| < \alpha$. As $\alpha < 1$ this was true only when $|F(I) \cap T| = |F(I - \{u\}) \cap T|$, but notice that this is also true when $k < \alpha < k+1$, since the difference will be smaller than one. □

Although this generalization gives a stronger characterization, we have seen that values of $\alpha \geq 1$ may give non-feasible results, hence, the best option is to have $0 < \alpha < 1$.

To summarize, we have seen that the game $\Gamma_I$ is a potential game and that a Nash Equilibrium can always be reached in polynomial time using the Best Response Dynamics. Regarding the cost function, when $0 < \alpha < 1$, we proved that a minimal influence set is equivalent to a Nash Equilibrium. We have not discussed details about the initial configuration of the Dynamics or the order in which the agents play; however, these parameters only affect the size of the set but not the feasibility of the solutions.

## 3.3 Threshold Selection Game

In this section, we switch the focus to the thresholds instead of the initial seed set, since our original goal was to study their effect and find a way to compute these values strategically. Given a network $G = (V, E, f)$, a set of target nodes $T \subseteq V$ and a set of initial nodes $I \subseteq V$ such that $I \cap T = \varnothing$, we want to find an assignment of values to each node in $V - T - I$ for their threshold. Notice that the values for the thresholds of the nodes in $I$ do not matter, as the agents are already influenced from the start. Regarding the thresholds in $T$, we will argue later which values they should get, for now we can discuss the model without defining them. The *Threshold Selection Game* $\Gamma_f(G, T, I)$ is defined as follows: one player for each node $v \in V - T - I$ and each can choose their threshold value $a_v \in \{1, .., d_i(v)\}$ based on the following utility function.

$$u_v(s) = \begin{cases} f(v) & \text{if } T \subseteq F(I) \\ -f(v) & \text{otherwise} \end{cases} \tag{2}$$

In this section, we will use $s_v$ and $f(v)$ interchangeably and denote the incoming degree of a node by $d_i(v)$. Like before, we will use $\Gamma_f(G, T, I)$ and $\Gamma_f$ interchangeably.

Recall that now we have a utility function; hence agents here aim to maximize the threshold while allowing for the influence of the target set. Before going into the game analysis, we will prove the following Theorems on influence networks.

**Theorem 7.** Given a graph $G = (V, E, f)$, a target set $T \subseteq V$, an initial set $I \subseteq V$, such tat $I \cap T = \varnothing$ and a node $v \in V - T - I$, if $f(v) = 1$ and $T \nsubseteq F(I)$ then it holds that $\forall k \in \mathbb{N}, 1 < k \leq d_i(v)$ if $f(v) = k$, then $T \nsubseteq F(I_{s_{-v},k})$.

*Proof.* A value of one for the threshold means that the node needs only one influence to get convinced. Therefore, making the threshold greater implies making the influence harder, as we need more incoming nodes to influence us. If the lowest possible value for the threshold does not allow for influence, no greater value will allow it either. □

**Theorem 8.** Given a graph $G = (V, E, f)$ where $\forall v \in V, f(v) = 1$ , an initial set $I \subseteq V$ and a target set $T \subseteq V$ such that $I \cap T = \varnothing$, if $T \nsubseteq F(I)$ then there exists no different threshold assignment that allows the influence of the nodes in $T$ from $I$.

*Proof.* When all nodes have their threshold set to one, we are saying that we only need one influence for each agent to convince the nodes. This is the minimum possible value. Therefore, if the nodes in $T$ cannot be influenced by the influences of $I$, no greater threshold assignment will allow it either, as increasing the thresholds makes the influence harder. □

We will say that the target set cannot be influenced if the conditions of Theorem 8 are fulfilled. Let us now analyze the game theoretic properties of game $\Gamma_f$ and the complexity of the Best Response Dynamics.

**Theorem 9.** Let $\Gamma_f$ be a Threshold Selection Game. Given a strategy profile $s = (s_1, ..., s_{|V-I-T|})$ and an agent $v \in V - I - T$, $BR(s_{-u})$ can be computed in $O(m(n + m))$ time in the game.

*Proof.* Recall that Equation 2 shows that the agents aim to approach their threshold to their incoming degree as much as possible when $T \subseteq F(I_s)$. Consider the Algorithm in Pseudo-Code 4.

**Algorithm 4** Best Response for a player in the game $\Gamma_f$

**Input**: $G = (V, E, f)$, $I \subseteq V$, $T \subseteq V$ s.t. $I \cap T = \varnothing$, $s = (s_1, ..., s_{|V-T-I|})$, $v \in V - T - I$
**Output**: $s_v \in \{1, .., d_i(v)\}$

$s_v \leftarrow 1$
**if** $T \subseteq F(I_{(s_{-v},1)})$ **then**
  **for** $ths \in [2, ..., d_i(v)]$ **do**
    **if** $T \nsubseteq F(I_{(s_{-v},ths)})$ **then**
      **return** $s_v$
    **end if**
    $s_v \leftarrow ths$
  **end for**
**end if**
**return** $s_v$

**Correctness:** due to Theorem 7, if a threshold of 1 does not allow for influence, we know that no greater value will either, and since we would be in the second case of Equation 2, this value is precisely the one that maximizes the utility. On the other hand, if $f(v) = 1$ allows for influence, we can perform a linear search over the possible values of $f(v)$ and get the highest one that still allows for influence. It is clear that in both cases the Best Response for the agent is found.

**Cost:** at the start, the algorithm computes $F(I_{(s_{-v},1)})$ which takes $O(n + m)$ time. Then, for each value between 2 and $d_i(v)$, we recompute the spread of influence with increasing threshold values. The cost of this would be $O(d_i(v)(n + m))$ for an agent $v$ which leads to a total worst-case cost of $O(m(n + m))$. □

This algorithm could be reduced to $O((n + m) \log m)$ time by using a binary search over the possible values of the thresholds. By monotonicity, sorting the possible values for the thresholds in ascending order also gives a sorting order for the possibility of influencing $T$. Recall that the higher the threshold, the harder it is to influence the group. However, due to the size of our networks, we will use the linear search since it is easier to implement and, on average, it will not make a huge difference in performance.

**Theorem 10.** Let $\Gamma_f$ be a Threshold Selection Game. For any given strategy profile $s = (s_1, ..., s_{|V-I-T|})$, the function

$$\phi(s) = \sum_{v \in V-I-T} u_v(s)$$

is an exact potential function for the game.

*Proof.* Like in the previous model, we must prove that the difference between the utility and the potential function between different strategies of a player $v \in V - I - T$ is the same. To do so, we take a look at the possible changes in strategy.

The utility function is split according to whether the target set $T$ is influenced or not. We must check four cases: changes in strategy in the first case of function 2, changes in the second case, and changes between both. Fix an agent $v \in V - I - T$.

**Case 1:** Take $a_1, a_2 \in A_v$ such that both actions make $T \subseteq F(I)$:

$$\begin{aligned}
u_v(s_{-v}, a_1) - u_v(s_{-v}, a_2) &= a_1 - a_2 \\
\phi(s_{-v}, a_1) - \phi(s_{-v}, a_2) &= \sum_{w \in V-I-T} u_w(s_{-v}, a_1) - \sum_{w \in V-I-T} u_w(s_{-v}, a_2) \\
&= (\sum_{w \in V-I-T-\{v\}} u_w(s) + a_1) - (\sum_{u \in V-I-T-\{v\}} u_w(s) + a_2) \\
&= a_1 - a_2
\end{aligned}$$

**Case 2:** Take $a_1, a_2 \in A_v$ such that neither of the actions make $T \subseteq F(I)$:

$$\begin{aligned}
u_v(s_{-v}, a_1) - u_v(s_{-v}, a_2) &= -a_1 + a_2 \\
\phi(s_{-v}, a_1) - \phi(s_{-v}, a_2) &= \sum_{w \in V-I-T} u_w(s_{-v}, a_1) - \sum_{w \in V-I-T} u_w(s_{-v}, a_2) \\
&= (\sum_{w \in V-I-T-\{v\}} u_w(s) - a_1) - (\sum_{u \in V-I-T-\{v\}} u_w(s) - a_2) \\
&= -a_1 + a_2
\end{aligned}$$

**Case 3:** Take $a_1, a_2 \in A_v$ such that $a_1$ makes $T \subseteq F(I)$ and $a_2$ does not.

$$\begin{aligned}
u_v(s_{-v}, a_1) - u_v(s_{-v}, a_2) &= a_1 - a_2 \\
\phi(s_{-v}, a_1) - \phi(s_{-v}, a_2) &= \sum_{w \in V-I-T} u_w(s_{-v}, a_1) - \sum_{w \in V-I-T} u_w(s_{-v}, a_2) \\
&= (\sum_{w \in V-I-T-\{v\}} u_w(s) + a_1) - (\sum_{u \in V-I-T-\{v\}} u_w(s) - a_2) \\
&= a_1 - a_2
\end{aligned}$$

It is clear to see that the remaining case is symmetric to the third. Since the agents are symmetric and the difference between the strategy changes in the cost and potential functions is exactly the same, we conclude that $\phi(s)$ is an exact potential function. □

**Theorem 11.** Threshold Selection Games are potential games.

*Proof.* By Theorem 10 we know that a game $\Gamma_f$ accepts a potential function, therefore the game is a potential one. □

Since $\Gamma_f$ is a potential game, we know that the Best Response Dynamics has the FIP, now we need to see how many steps are needed to reach a NE through Best Response changes starting from any strategy profile.

**Theorem 12.** Let $\Gamma_f$ be a Threshold Selection Game. Given a strategy profile $s = (s_1, ..., s_{|V-I-T|})$, the Best Response Dynamics in game $\Gamma_f$ needs $O(|V| - |I| - |T|)$ steps to reach a Nash Equilibrium in the game $\Gamma_f$.

*Proof.* Since $\Gamma_f$ is a potential game, we know that a pure Nash Equilibrium always exists and that the Best Response Dynamics does not have any IRC. Therefore, an improvement in the dynamics means a step closer to a NE. Let us look at what the improved changes look like in the current profile $s$.

**Case 1:** assume that $f$ is distributed in such a way that $T \subseteq F(I)$ and let us focus on an agent $v \in V - I - T$. Currently we are in the first case of Equation 2. In this case, agent $v$ will aim to maximize its threshold until $f(v) = d_i(v)$ or until some value for its threshold breaks the influence. It is clear that she will never be inclined to break the influence, since $f(v) > -f(v)$. Hence, the agent will set their threshold to be the maximum possible and will give its turn to the next agent. Since the influence does not change, each agent will choose its threshold at most once, which means that we will at most have $|V| - |I| - |T|$ improvement changes in strategy.

**Case 2:** assume that $f$ is distributed in such a way that influence is not possible, no matter what the agents choose. Therefore, we are in the second case of Equation 2. In this case, the best value that agents can get is 1, as $\forall v \in V - I - T, \max(-f(v)) = 1$. Since there is no threshold value that will make $T \subseteq F(I)$, the agents will change their threshold to 1 at most once. Hence, we will have at most $|V| - |I| - |T|$ improvement strategy changes.

**Case 3:** the last case is when we get a threshold distribution that does not allow for influence, but some set of agents can change their threshold to allow influence. Notice that the other way around will never happen, as agents would go from a positive utility to a negative one, which is worse for them. Initially, the target set cannot be influenced; therefore, we are in the second case. At some point, some agent will change its threshold and allow the influence. After this, we will be in the first case, where agents aim to maximize their threshold without breaking the influence. Therefore, in the worst case, the dynamics will need $2(|V| - |I| - |T|)$ improvement changes to reach a NE, the first to set all thresholds to 1 and the second to assign the threshold values without breaking the influence.

Since in any case we need $O(|V| - |I| - |T|)$ steps to reach a NE, we conclude that the time needed to reach a Nash Equilibrium is polynomial.

□

Seeing that a NE always exists and that a Best Response Dynamics always reaches it in polynomial time, now we only need to see how the thresholds will be shaped in the output of this model.

When the target set cannot be influenced, the only possible equilibrium is when all agents $v \in V - I - T$ choose $s_v = 1$, because they want to maximize $-f(v)$.

On the other hand, when the target set can be influenced, there will be some nodes that are critical for influence and some that will not affect it. Being critical here means that the value that the agent chooses has an effect on the number of nodes influenced in $T$. For nodes that are not critical, their threshold will always be $d_i(v)$, as it is the value that maximizes the utility when $T \subseteq F(I)$. For the other nodes, their threshold will be equal to their number of incoming influencers. Notice that that is the maximum value possible without breaking the influence of $T$.

To better understand this, let us look at an arbitrary network where $|I| = |T| = 1$. When the agents' thresholds are 1, it is clear that the node in $T$ will be influenced if there is a directed path starting from $i \in I$ and ending in $t \in T$. Therefore, under the utility function of Equation 2, any network where the agents on some $i - t$ path have their threshold set to the number of incoming influences and where all other agents have threshold 1 will be a Nash Equilibrium. This could be generalized to any number of nodes in $I$ or $T$, with multiple paths.

Notice that we have not discussed the values of the thresholds of the nodes in $T$. This is because their values are initially fixed and do not change. Regardless of the thresholds, the agents always aim to maximize their threshold while maintaining influence. If that is not possible, due to topology, then their goal will be to set their threshold to the minimum value. In any case, equilibrium will still be reached, and the dynamics and the best response will be computed the same.

To summarize, we have proposed two models, one aimed at understanding how to strategically solve the problem of finding a minimal initial seed set, and the second to set the values of the thresholds of the agents to increase the influence of the target nodes. The first model provides as a Nash Equilibrium stable initial sets, in the sense that all agents are necessary for the current influence, while the second computes the minimum threshold that allows for influence of the target set. Conceptually, the second model computes a threshold that helps the influence to occur while keeping it high enough to avoid receiving influence from other parts of the network. The idea is to represent a somewhat real agent that would try to help the influence of a target set but still avoid being influenced by other rumors.

# 4 Experiments

In this section, we introduce the designed experiments and discuss the different networks and metrics used to measure the spread of influence. We will conduct two different experiments; the idea in both is to take a set of target nodes and a set of initial nodes from a given social network and simulate the Best Response Dynamics of the Seed Selection Game $\Gamma_I$, and the Threshold Selection Game $\Gamma_f$ on them. The first experiment will help us to see how the initial nodes are characterized and how the thresholds change when the influence already occurs. The second one, will help us to see how the thresholds affect the influence and how far from a Nash Equilibrium a given initial configuration of thresholds and initial nodes is. To compare the results, we will use different metrics, ones proposed by us and ones based on centrality measures.

## 4.1 Methodology

All experiments have been executed in the Computer Science department cluster *rdlab*, on a `HP ProLiant DL380p server` with two `Intel(R) Xeon(R) E5-2660`, using up to 32 cores. Each experiment will be executed 5 times, and the resulting metrics will be an average of these repetitions. Since some random choices are made in the executions, we will fix a seed value of 2000 for the random number generator, so that the experiments can be reproduced. For simplicity purposes, throughout this section we will use a constant $\epsilon = 10^{-6}$ for those values that are significantly small.

The experiments are executed in 10 different real networks, a portion of the ones used in [30] and [31]. Their attributes and properties are displayed in Table 3. Some of these networks have weights on the edges, but we will ignore them in this work to maintain simplicity as argued.

- **Dining Table [32]:** network with the seating preferences of 26 girls around a dining table. Every guest has two edges, each corresponding to their partners preference.
- **Dolphins [33]:** network with the relationships between bottlenose dolphins. An edge between two dolphins $(i, j)$ means a frequent relationship between the two.
- **Human Brain [34] [35]:** network with 477 connectomes of different humans obtained with magnetic resonances in the *Human Connectome Project*. A connectome is a map of neuronal connections of the brain modeled as an undirected graph. The nodes are regions of the brain and the edges connections between them.
- **ArXiv [36]:** network of scientific collaboration in the field of general relativity and quantum cosmology, where the nodes are scientists and the edges represent co-authorship in some paper.

- **Wikipedia** [37]**:** represents the vote for the new administrator of Wikipedia. An edge $(i, j)$ indicates that user $i$ voted for user $j$ to become an administrator.

- **Caida** [38]**:** network of different Internet Service Providers (ISP), where the nodes are actors and the edges contracts and professional relationships between them. It represents the state of the CAIDA project in November 2007.

- **ENRON** [39] [40]**:** non-directed network representing the internal e-mail communication between the employees of the ENRON company. An edge between two employers means that they sent at least one e-mail to each other.

- **Gnutella** [36] [41]**:** state of the peer to peer (p2p) Gnutella network on the 31th August 2002. Any node in the network can work either as a server or as a client and the edges represent the connections between users.

- **Epinions** [42]**:** directed network with trust relationships between users of the *Epinions.com* review website. An edge $(i, j)$ represents whether user $i$ trusts user $j$.

- **Higgs** [43]**:** network of *tweets* made in the first week of July 2012, regarding the discovery of the Higgs boson. The nodes are users, and the edges represent whether the account $i$ has retweeted some post from the account $j$.

Table 3: Attributes of the used networks. The line separates between what we consider the small, the medium and the big networks of this work.

| Network | $\|V\|$ | $\|E\|$ | Directed | Minimum In-Degree | Average In-Degree | Maximum In-Degree |
|---|---|---|---|---|---|---|
| **Dining Table** | 26 | 52 | Y | 0 | 2 | 6 |
| **Dolphins** | 62 | 159 | N | 1 | 5.1290 | 12 |
| **Human Brain** | 480 | 1000 | N | 1 | 4.1667 | 96 |
| **ArXiv** | 5242 | 14496 | N | 2 | 11.0568 | 162 |
| **Wikipedia** | 7115 | 106762 | Y | 0 | 14.5733 | 457 |
| **Caida** | 26475 | 106762 | Y | 1 | 4.0326 | 2628 |
| **ENRON** | 36692 | 183831 | N | 2 | 20.0404 | 2766 |
| **Gnutella** | 62586 | 147892 | Y | 0 | 2.3630 | 68 |
| **Epinions** | 75879 | 508837 | Y | 0 | 6.7059 | 3035 |
| **Higgs** | 256491 | 328132 | Y | 0 | 1.2793 | 14060 |

The output of our experiments will be a set of Nash Equilibriums. For both models, this will mean a set of initial nodes, target nodes, and threshold assignment. To compare the sets, we propose three new global metrics related to the influence graphs and use three well known centrality measures. For the thresholds, we will generate different plots with the computed values. Given a graph $G = (V, E)$ we design the following metrics:

- **Influence Size** $\pi_I$: measures how large the initial influence set is with respect to the network size. Given an initial set $I \subseteq V$, this is computed as follows:

$$\pi_I = \frac{|I|}{|V|}$$

  We say the influence is *small*, if the number of initial nodes is small in comparison to the network size.

- **Influence Expansion** $\pi_F$: metric that indicates how strong the influence of the initial nodes is. Given an initial set $I \subseteq V$ we compute the metric as follows:

$$\pi_F = \frac{|F(I)|}{|V|}$$

  We say that an initial set $I$ is *expansive* if the proportion of influenced nodes is high.

- **Influence Efficiency** $\pi_T$: we say that an initial set $I \subseteq V$ is efficient with respect to a target set $T \subseteq V$ such that $I \cap T = \varnothing$, if the proportion of nodes influenced in $T$ with respect to the size of $T$ is high. This is summarized in the following function.

$$\pi_T = \frac{|T \cap F(I)|}{|T|}$$

  Therefore, an initial set $I$ is *efficient* if it is able to influence a large number of nodes in $T$.

As mentioned, besides using our metrics, we also use three well known centrality measures.

- **Degree:** a local centrality metric based on the degree of a node. When the network is directed, we will refer to the *in-degree* $d_i(v)$ or the out-degree $d_o(v)$ of the node.

$$d(v) = |\{(u, v) \in E\}|$$

- **Betweenness** [44]**:** a global centrality metric based on paths and distances. A node will be important if it lies in many shortest paths between different nodes. If we take nodes as message routers, the betweenness measure of a node will represent how important that node is to transmit information between two endpoints. Given a node $v \in V$, its betweenness is computed as follows:

$$Betweenness(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

  where $\sigma_{st}$ is the number of shortest paths between $s, t \in V \times V$ and $\sigma_{st}(v)$ the number of paths between $s$ and $t$ that cross $v$. It is equal to the average proportion of paths where $v$ is included in.

- **Pagerank** [45]**:** designed by Lawrence Page and Sergey Brin with Google, this is the most well known algorithm for ranking web pages according to their relevance. The algorithm classifies nodes according to the probability that a user will visit them when randomly traversing the network. The process can be viewed as a random walk on a Markov Chain, where we let a user randomly traverse the network for a set of iterations until the probability distribution converges to a stable distribution. The number of iterations will depend on whether the agent decides to stop or not, which can be represented with a probability variable called amortization factor. Therefore, a page, or node, will be ranked as important if its resulting probability is high enough. Notice that an important page will have important pages that point to it. The computation of the metric can be summarized in the following formula:

$$Pagerank(v) = (1 - \alpha) + \alpha \sum_{(u,v) \in E} \frac{Pagerank(u)}{d_o(u)}$$

  where $d_o(v)$ is the out-degree of a node $u \in V$ and $\alpha$ is the amortization factor where $0 < \alpha \leq 1$. See how the metric is computed recursively.

For the centrality measures, we compute the average, maximum, and minimum of a given set. The average will help us compare the sets, and the minimum and maximum will help us see the dispersion of the measures.

To complete the analysis and study how long the models take to converge to a Nash Equilibrium, we will also measure the number of rounds each Dynamics takes. This will help us see the complexity of obtaining an answer for each model.

As mentioned before, we will also study the distribution of the thresholds selected by the Threshold Selection Game. To do so, we will plot the proportion of nodes that lie in specific ranges of threshold values. To have a better comparison between networks, we will use the threshold as a proportion of neighbors needed to get influenced, so that the values are normalized to the interval $(0, 1]$.

## 4.2 Implementation

Most of the code has been reused from [30] and [31], since we will be using the same functionalities. We will review the relevant implementation details; for a more in depth description, one can read the corresponding works. The data processing and computation pipeline is basically the same in all these works. The complete code can be checked in our Github repository [46].

The simulation is implemented entirely in C++. The Graph and the input data processing have been implemented by [30] and [31], while the code related to our models has been implemented by us. We had to introduce some tweaks to the code, to adapt it to our functionalities, especially to the thresholds management. There are also some Python scripts, that are mainly used to generate the plots of this document and to draw different graphs. To do so, we used the *Networkx* package [47].

For the implementation of the networks, we use Graphs represented as adjacency lists. Since we work with sparse social networks, this choice makes the computations much more efficient. The algorithms for the expansion computation, the Game Dynamics of both models, and their Best Response are implemented exactly like shown in the previous pseudo-codes. To improve the running time of the experiments, we used OpenMP [48] to parallelize some parts of the code. We also reuse the results obtained in [31], where different metrics are computed for the networks, precisely the Pagerank and Betweennes. This will help us save time and run more experiments on our models.

Lastly, for the random choices in the algorithms, we use uniform distributions. For random sampling, we implement the Fisher Yates algorithm [49], to randomly shuffle an array of nodes and then take in order the number of nodes needed. The advantage of this implementation is that we only need $O(n)$ steps. The implementation can be viewed in the source code.

## 4.3 Experiment 1

Our first goal is to study how the initial and influenced nodes are shaped and what relevance do the thresholds have when influence can already happen. Therefore, for a fixed social network, the first experiment has the following outline:

1. Uniformly fix the threshold of all nodes. We repeat the experiments for 4 different values of thresholds to see how they relate to each other: 0.25, 0.50, 0.75, and 0.95.
2. Randomly select a proportion of target nodes $T$ ($\approx 20\%$).
3. Compute the initial set of nodes $I \subseteq V$ with the Best Response Dynamics of $\Gamma_I$.
4. Compute the thresholds of the nodes $v \in V - I - T$ with the Best Response Dynamics of $\Gamma_f$.
5. Analyze the metrics of the found NE of $\Gamma_I$ and $\Gamma_f$.

The main idea here is to use the first model to compute a seed set and then, use this seed set with the second model to compute the thresholds. By computing the seed set first, we can see how the initial set is distributed in the network when given a random set of target nodes. After running the second model, we will obtain a set of thresholds that allow for influence. Here we shall study whether changing the initial thresholds makes any change in the influence and if the selected values are far or not from the initially fixed threshold.

Recall that the Best Response Dynamics starts from an initial configuration. Therefore, we will repeat this experiment twice, each with different initial configurations, to see if starting from different points leads to different solution qualities. We will try with a *complete* initial set and a *random* initial set.

Since we have many results and many tables to compare, in this section we will only outline the most relevant differences and similarities between each type of experiment. For each initial configuration, we will compare the results starting with different thresholds for those relevant values. The complete set of tables and plots can be viewed in the Appendix.

To shorten the explanation, in this section we will refer to the Nash Equilibrium obtained with the Best Response Dynamics of the corresponding game as the output of the game.

### 4.3.1 Complete initial configuration

Here we start the Game Dynamics of the first model with an initial strategy profile $s = (1, ..., 1)$, which implies $I_s = V \setminus T$. We will refer to this configuration as the *complete* configuration. Recall that here we execute both models, first $\Gamma_I$ and then $\Gamma_f$ with the output of the first. The complete configuration only applies to the Best Response Dynamics of $\Gamma_I$, the outcome of the game will then be passed to the game $\Gamma_f$ in step 4 of the outline.

The first result we notice is that the metrics obtained from the output of $\Gamma_I$ and the ones obtained from the output of $\Gamma_f$ are exactly the same, as can be seen in Tables 4 and 5. This can be a bit surprising, since the second model tries to improve the influence by changing the thresholds of the nodes that are not either seed or target nodes. This implies that the thresholds of the nodes $v \in V - I - T$ do not affect the influence when we already have found a minimal initial seed set. We have seen this behavior throughout all the results in the first experiment, hence, we will only include the output of the second model from now on. The complete set of tables can be checked in the Appendix. Notice that the only difference between the two experiments is the number of rounds, which is reasonable, since both models take different times to reach a NE. The positive aspect is that the number of rounds taken by both models is smaller than 4 on average, which is constant.

Table 4: Table with the metrics of the first experiment with initial threshold 0.25 and initial complete configuration for the initial and influence sets. This table is the outcome of the $\Gamma_I$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Average Out-Degree Initial | Average In-Degree Influence | Average Betweenness Initial | Average Betweenness Influence | Average Pagerank Initial | Average Pagerank Influence |
|---|---|---|---|---|---|---|---|---|---|---|
| Dining Table | 0.0615 | 0.7615 | 0.8400 | 2.0000 | 0.2000 | 2.0000 | 0.0000 | 0.0399 | 0.0125 | 0.0458 |
| Dolphins | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 3.4000 | 5.1290 | 0.0000 | 0.0393 | 0.0118 | 0.0161 |
| Human Brain | 0.0067 | 0.9883 | 0.9958 | 2.0000 | 1.1900 | 4.2000 | 0.0000 | 0.0079 | 0.0034 | 0.0021 |
| ArXiv | 0.0286 | 0.8954 | 0.9813 | 2.0000 | 3.2808 | 11.9604 | 0.0000 | 0.0012 | 0.0002 | 0.0002 |
| Wikipedia | 0.0017 | 0.3289 | 0.3455 | 2.0000 | 0.0000 | 24.9066 | 0.0000 | 0.0001 | 0.0001 | 0.0001 |
| Caida | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 4.0326 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| ENRON | 0.0120 | 0.9589 | 0.9921 | 2.0000 | 3.4052 | 20.7663 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| Gnutella | 0.0021 | 0.9881 | 0.9916 | 2.0000 | 0.1353 | 2.3614 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Epinions | 0.0086 | 0.6547 | 0.6778 | 2.0000 | 0.3959 | 9.4864 | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 |
| Higgs | 0.0231 | 0.0780 | 0.1436 | 2.0000 | 0.0365 | 1.4466 | 0.0000 | $< \epsilon$ | $< \epsilon$ | $< \epsilon$ |

Table 5: Table with the metrics of the first experiment with initial threshold 0.25 and initial complete configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Average Out-Degree Initial | Average In-Degree Influence | Average Betweenness Initial | Average Betweenness Influence | Average Pagerank Initial | Average Pagerank Influence |
|---|---|---|---|---|---|---|---|---|---|---|
| Dining Table | 0.0615 | 0.7615 | 0.8400 | 1.4000 | 0.2000 | 2.0000 | 0.0000 | 0.0399 | 0.0125 | 0.0458 |
| Dolphins | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 3.4000 | 5.1290 | 0.0000 | 0.0393 | 0.0118 | 0.0161 |
| Human Brain | 0.0067 | 0.9883 | 0.9958 | 1.8000 | 1.1900 | 4.2000 | 0.0000 | 0.0079 | 0.0034 | 0.0021 |
| ArXiv | 0.0286 | 0.8954 | 0.9813 | 1.0000 | 3.2808 | 11.9604 | 0.0000 | 0.0012 | 0.0002 | 0.0002 |
| Wikipedia | 0.0017 | 0.3289 | 0.3455 | 1.0000 | 0.0000 | 24.9066 | 0.0000 | 0.0001 | 0.0001 | 0.0001 |
| Caida | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 4.0326 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| ENRON | 0.0120 | 0.9589 | 0.9921 | 1.0000 | 3.4052 | 20.7663 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| Gnutella | 0.0021 | 0.9881 | 0.9916 | 1.0000 | 0.1353 | 2.3614 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Epinions | 0.0086 | 0.6547 | 0.6778 | 1.0000 | 0.3959 | 9.4864 | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 |
| Higgs | 0.0231 | 0.0780 | 0.1436 | 1.0000 | 0.0365 | 1.4466 | 0.0000 | $< \epsilon$ | $< \epsilon$ | $< \epsilon$ |

With respect to the values obtained, we see different things. Starting with the influence metrics taken on the computed NE, we see that $\pi_I < 0.1$ in all networks, which means that the minimal sets found are very small. If we look at the proportion of nodes influenced $\pi_F$ and $\pi_T$, in general more than half of the network ends up being influenced. Furthermore, $\pi_T > \pi_F$, which is expected as we are trying to maximize the number of nodes influenced in $T$. However, for the Wikipedia and Higgs networks, we see that $\pi_F, \pi_T < 0.5$, which slightly contradicts the results of the other networks. Our hypothesis is that influence does not perform well in these networks, most likely due to their topology. As a matter of fact, both networks are not connected, and Higgs has a very low average in-degree. Therefore, it is very likely that these networks have some specific sets of nodes that are very hard to influence, or that they need to be a part of the initial seed set to be influenced, which our model does not allow.

Turning to the centrality measures, we see a variety of results. For the initial sets, we see that their average out-degree and their global centrality metrics are very low. This means that these nodes are not lying in centric parts of the network and that they are slightly isolated, with very few neighboring nodes. Regarding the influence nodes, we see that their in-degree average is quite high, taking values that are very close to the average in-degree of the whole network. Since in the minimal sets found are highly expansive, this is not a surprising result. For the global metrics, we get higher values than those of the initial sets, which makes sense since the influenced nodes will lie in more central parts of the network. Nevertheless, for both types of sets, we get generally low values for the global centrality metrics, which is reasonable if we take into account that we are working with social networks that are very sparse, and hence, will not have many central nodes.

To simplify the tables, here we only display the average measure of the sets. In Tables 17 and 18 of the Appendix, we also get the maximum and minimum value of each metric. We see that the metrics show high dispersion in these sets, being closer to the minimum values than to the maximum, which can also be explained by the sparsity of the networks.

From all these results, we can generally say that a very small subset of nodes is able to influence the greater part of the networks and that these seed nodes are in non-central parts of the network. One explanation for this might be due to the networks being non-weighted, where the strength of influence of the agents comes from the singular connections and not from individual power over their neighbors.

Regarding the thresholds fixed at the start of the experiment, we see that starting from other threshold values does not significantly change the results. Let us look at Tables 6, 7 and 8.

In the smaller networks, the difference between $\pi_I$, $\pi_F$, and $\pi_T$ is more noticeable when starting from different thresholds than in the larger networks. The main equivalences that we see in the metrics are when the thresholds are fixed to 0.25 and to 0.5, and when they are fixed to 0.75 and 0.95. This can be explained by the sparsity of the networks. The threshold measures the number of incoming neighbors that a node needs to be convinced of a rumor. However, we have seen that the average degree of the nodes in these networks is very small and that most of the samples are closer to the minimum value. Therefore, the difference between a lower threshold and a higher threshold will be very small in sparse networks. We do see some differences between the results obtained starting with threshold 0.25 and threshold 0.95, but they are not significant.

Lastly, in Figures 9, 10, 11, 12, 13, 14, 15, and 16, we have the plots for the threshold distribution. We have not included them here, since they do not show any relevant data. The main result we observe is that the distributions are very different for each network and do not seem to follow any pattern. Some seem to follow a type of skewed normal distributions, while others have a bit of a random distribution. Regarding whether they are close to the initially fixed threshold values or not, we do not see any general behavior that can be expressed, since some networks get thresholds closer to the initial value and some others get very different values. What we do observe is that the results are consistent with the same networks.

Table 6: Table with the metrics of the first experiment with initial threshold 0.50 and initial complete configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| **Network** | $\pi_I$ | $\pi_F$ | $\pi_T$ | **Rounds** | **Average Out-Degree Initial** | **Average In-Degree Influence** | **Average Betweenness Initial** | **Average Betweenness Influence** | **Average Pagerank Initial** | **Average Pagerank Influence** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Dining Table** | 0.0462 | 0.6923 | 0.8800 | 1.4000 | 0.1000 | 2.0000 | 0.0000 | 0.0417 | 0.0060 | 0.0485 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 3.0000 | 5.1290 | 0.0000 | 0.0393 | 0.0110 | 0.0161 |
| **Human Brain** | 0.0071 | 0.9917 | 1.0000 | 2.0000 | 1.2000 | 4.1916 | 0.0000 | 0.0078 | 0.0027 | 0.0021 |
| **ArXiv** | 0.0312 | 0.9057 | 0.9832 | 1.0000 | 3.3092 | 11.8734 | 0.0000 | 0.0012 | 0.0002 | 0.0002 |
| **Wikipedia** | 0.0017 | 0.3290 | 0.3331 | 1.0000 | 0.0000 | 24.8701 | 0.0000 | 0.0001 | 0.0001 | 0.0001 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 4.0326 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **ENRON** | 0.0120 | 0.9588 | 0.9902 | 1.0000 | 3.3798 | 20.7659 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **Gnutella** | 0.0021 | 0.9877 | 0.9913 | 1.0000 | 0.1420 | 2.3617 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **Epinions** | 0.0084 | 0.6544 | 0.6765 | 1.0000 | 0.3855 | 9.4913 | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 |
| **Higgs** | 0.0231 | 0.0783 | 0.1441 | 1.0000 | 0.0333 | 1.4506 | 0.0000 | $< \epsilon$ | $< \epsilon$ | $< \epsilon$ |

Table 7: Table with the metrics of the first experiment with initial threshold 0.75 and initial complete configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| **Network** | $\pi_I$ | $\pi_F$ | $\pi_T$ | **Rounds** | **Average Out-Degree Initial** | **Average In-Degree Influence** | **Average Betweenness Initial** | **Average Betweenness Influence** | **Average Pagerank Initial** | **Average Pagerank Influence** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Dining Table** | 0.0538 | 0.7077 | 0.8400 | 1.2000 | 0.4000 | 2.0000 | 0.0000 | 0.0398 | 0.0098 | 0.0483 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 3.0000 | 5.1290 | 0.0000 | 0.0393 | 0.0110 | 0.0161 |
| **Human Brain** | 0.0063 | 0.9888 | 1.0000 | 2.0000 | 1.0500 | 4.1999 | 0.0000 | 0.0079 | 0.0028 | 0.0021 |
| **ArXiv** | 0.0316 | 0.9057 | 0.9849 | 1.0000 | 3.2583 | 11.8685 | 0.0000 | 0.0012 | 0.0002 | 0.0002 |
| **Wikipedia** | 0.0013 | 0.3282 | 0.3376 | 1.0000 | 0.0000 | 24.8795 | 0.0000 | 0.0001 | 0.0001 | 0.0001 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 4.0326 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **ENRON** | 0.0119 | 0.9592 | 0.9914 | 1.0000 | 3.3842 | 20.7626 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **Gnutella** | 0.0021 | 0.9882 | 0.9915 | 1.0000 | 0.1468 | 2.3614 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **Epinions** | 0.0085 | 0.6544 | 0.6766 | 1.0000 | 0.3875 | 9.4906 | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 |
| **Higgs** | 0.0232 | 0.0781 | 0.1443 | 1.0000 | 0.0333 | 1.4533 | 0.0000 | $< \epsilon$ | $< \epsilon$ | $< \epsilon$ |

Table 8: Table with the metrics of the first experiment with initial threshold 0.95 and initial complete configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| **Network** | $\pi_I$ | $\pi_F$ | $\pi_T$ | **Rounds** | **Average Out-Degree Initial** | **Average In-Degree Influence** | **Average Betweenness Initial** | **Average Betweenness Influence** | **Average Pagerank Initial** | **Average Pagerank Influence** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Dining Table** | 0.0385 | 0.6538 | 0.7600 | 1.4000 | 0.0000 | 2.0000 | 0.0000 | 0.0433 | 0.0058 | 0.0500 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 3.0000 | 5.1290 | 0.0000 | 0.0393 | 0.0110 | 0.0161 |
| **Human Brain** | 0.0071 | 0.9917 | 0.9958 | 1.8000 | 1.1167 | 4.1917 | 0.0000 | 0.0078 | 0.0025 | 0.0021 |
| **ArXiv** | 0.0301 | 0.9019 | 0.9847 | 1.0000 | 3.2663 | 11.9023 | 0.0000 | 0.0012 | 0.0002 | 0.0002 |
| **Wikipedia** | 0.0016 | 0.3288 | 0.3361 | 1.0000 | 0.0000 | 24.9591 | 0.0000 | 0.0001 | 0.0001 | 0.0001 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 4.0326 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **ENRON** | 0.0118 | 0.9586 | 0.9923 | 1.0000 | 3.4338 | 20.7726 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **Gnutella** | 0.0021 | 0.9884 | 0.9921 | 1.0000 | 0.1453 | 2.3615 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **Epinions** | 0.0084 | 0.6544 | 0.6768 | 1.0000 | 0.3837 | 9.4907 | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 |
| **Higgs** | 0.0230 | 0.0778 | 0.1428 | 1.0000 | 0.0332 | 1.4581 | 0.0000 | $< \epsilon$ | $< \epsilon$ | $< \epsilon$ |

### 4.3.2 Random initial configuration

Now, instead of choosing $s = (1, 1, 1, ..., 1)$ as initial strategy, we pick a subset of a random set of players to participate in the initial configuration of the Seed Selection Game. With this experiment, we will check how relevant the initial configuration is in finding a minimal influence set with the Best Response Dynamics.

Like before, we see that the metrics obtained from the outcome of the game $\Gamma_I$ and $\Gamma_f$ are quite the same, which can be checked in the Appendix. In addition to that, we see that the difference between the results starting with a random configuration and a complete one does not change much the values of the metrics, especially $\pi_T$. We see that in some networks, $\pi_T$ is slightly decreased; therefore, we can conclude that starting from a random configuration of initial nodes or a complete one does not have a meaningful impact on the influence of the target set, although a complete one does have slightly better results.

Regarding the thresholds distribution, we see the same behavior as with the complete configuration.

Table 9: Table with the metrics of the first experiment with initial threshold 0.25 and initial random configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Average Out-Degree Initial | Average In-Degree Influence | Average Betweenness Initial | Average Betweenness Influence | Average Pagerank Initial | Average Pagerank Influence |
|---|---|---|---|---|---|---|---|---|---|---|
| Dining Table | 0.0538 | 0.5923 | 0.7200 | 1.2000 | 1.9000 | 2.0000 | 0.0000 | 0.0469 | 0.0147 | 0.0572 |
| Dolphins | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 6.0000 | 5.1290 | 0.0000 | 0.0393 | 0.0141 | 0.0161 |
| Human Brain | 0.0063 | 0.9904 | 1.0000 | 2.0000 | 2.6667 | 4.1961 | 0.0000 | 0.0078 | 0.0033 | 0.0021 |
| ArXiv | 0.0298 | 0.8993 | 0.9815 | 1.0000 | 3.9118 | 11.9239 | 0.0000 | 0.0012 | 0.0002 | 0.0002 |
| Wikipedia | 0.0014 | 0.3285 | 0.3398 | 1.0000 | 0.0000 | 24.9744 | 0.0000 | 0.0001 | 0.0001 | 0.0001 |
| Caida | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.8000 | 4.0326 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| ENRON | 0.0118 | 0.9587 | 0.9914 | 1.0000 | 3.9092 | 20.7703 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| Gnutella | 0.0021 | 0.9883 | 0.9922 | 1.0000 | 0.1418 | 2.3605 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Epinions | 0.0087 | 0.6543 | 0.6777 | 1.0000 | 0.7067 | 9.4954 | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 |
| Higgs | 0.0229 | 0.0786 | 0.1444 | 1.0000 | 0.0743 | 1.5031 | 0.0000 | $< \epsilon$ | $< \epsilon$ | $< \epsilon$ |

Table 10: Table with the metrics of the first experiment with initial threshold 0.50 and initial random configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Average Out-Degree Initial | Average In-Degree Influence | Average Betweenness Initial | Average Betweenness Influence | Average Pagerank Initial | Average Pagerank Influence |
|---|---|---|---|---|---|---|---|---|---|---|
| Dining Table | 0.0462 | 0.6077 | 0.8000 | 1.4000 | 1.1000 | 2.0000 | 0.0000 | 0.0439 | 0.0092 | 0.0533 |
| Dolphins | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 6.6000 | 5.1290 | 0.0000 | 0.0393 | 0.0191 | 0.0161 |
| Human Brain | 0.0058 | 0.9883 | 0.9958 | 1.8000 | 2.7800 | 4.2017 | 0.0000 | 0.0079 | 0.0030 | 0.0021 |
| ArXiv | 0.0308 | 0.9036 | 0.9826 | 1.0000 | 3.8809 | 11.8849 | 0.0000 | 0.0012 | 0.0002 | 0.0002 |
| Wikipedia | 0.0015 | 0.3286 | 0.3324 | 1.0000 | 0.0000 | 25.0136 | 0.0000 | 0.0001 | 0.0001 | 0.0001 |
| Caida | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.8000 | 4.0326 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| ENRON | 0.0122 | 0.9592 | 0.9917 | 1.0000 | 3.8071 | 20.7595 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| Gnutella | 0.0022 | 0.9885 | 0.9912 | 1.0000 | 0.1556 | 2.3616 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Epinions | 0.0085 | 0.6538 | 0.6783 | 1.0000 | 0.7213 | 9.5006 | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 |
| Higgs | 0.0231 | 0.0796 | 0.1453 | 1.0000 | 0.1002 | 1.5203 | 0.0000 | $< \epsilon$ | $< \epsilon$ | $< \epsilon$ |

Table 11: Table with the metrics of the first experiment with initial threshold 0.75 and initial random configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Average Out-Degree Initial | Average In-Degree Influence | Average Betweenness Initial | Average Betweenness Influence | Average Pagerank Initial | Average Pagerank Influence |
|---|---|---|---|---|---|---|---|---|---|---|
| **Dining Table** | 0.0538 | 0.6385 | 0.7600 | 1.2000 | 1.5000 | 2.0000 | 0.0000 | 0.0458 | 0.0196 | 0.0535 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 6.0000 | 5.1290 | 0.0000 | 0.0393 | 0.0194 | 0.0161 |
| **Human Brain** | 0.0054 | 0.9850 | 0.9958 | 1.8000 | 2.1467 | 4.2101 | 0.0000 | 0.0079 | 0.0016 | 0.0021 |
| **ArXiv** | 0.0300 | 0.8992 | 0.9884 | 1.0000 | 3.6723 | 11.9172 | 0.0000 | 0.0012 | 0.0002 | 0.0002 |
| **Wikipedia** | 0.0015 | 0.3287 | 0.3244 | 1.0000 | 0.0000 | 25.1070 | 0.0000 | 0.0001 | 0.0001 | 0.0001 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 2.6000 | 4.0326 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **ENRON** | 0.0119 | 0.9590 | 0.9907 | 1.0000 | 3.8303 | 20.7652 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **Gnutella** | 0.0022 | 0.9885 | 0.9917 | 1.0000 | 0.1378 | 2.3614 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **Epinions** | 0.0084 | 0.6537 | 0.6755 | 1.0000 | 0.7317 | 9.5030 | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 |
| **Higgs** | 0.0229 | 0.0787 | 0.1446 | 1.0000 | 0.0790 | 1.4962 | 0.0000 | $< \epsilon$ | $< \epsilon$ | $< \epsilon$ |

Table 12: Table with the metrics of the first experiment with initial threshold 0.95 and initial random configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Average Out-Degree Initial | Average In-Degree Influence | Average Betweenness Initial | Average Betweenness Influence | Average Pagerank Initial | Average Pagerank Influence |
|---|---|---|---|---|---|---|---|---|---|---|
| **Dining Table** | 0.0538 | 0.6769 | 0.8000 | 1.4000 | 0.7000 | 2.0000 | 0.0000 | 0.0422 | 0.0103 | 0.0494 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 5.6000 | 5.1290 | 0.0000 | 0.0393 | 0.0185 | 0.0161 |
| **Human Brain** | 0.0054 | 0.9879 | 0.9958 | 1.8000 | 3.2333 | 4.2033 | 0.0000 | 0.0079 | 0.0016 | 0.0021 |
| **ArXiv** | 0.0294 | 0.8985 | 0.9803 | 1.0000 | 3.7559 | 11.9364 | 0.0000 | 0.0012 | 0.0002 | 0.0002 |
| **Wikipedia** | 0.0013 | 0.3283 | 0.3280 | 1.0000 | 0.0000 | 25.0014 | 0.0000 | 0.0001 | 0.0001 | 0.0001 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 7.8000 | 4.0326 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **ENRON** | 0.0119 | 0.9585 | 0.9919 | 1.0000 | 3.8488 | 20.7732 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **Gnutella** | 0.0022 | 0.9886 | 0.9915 | 1.0000 | 0.1431 | 2.3608 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **Epinions** | 0.0085 | 0.6535 | 0.6783 | 1.0000 | 0.6837 | 9.5036 | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 |
| **Higgs** | 0.0231 | 0.0790 | 0.1455 | 1.0000 | 0.0721 | 1.4963 | 0.0000 | 0.0000 | $< \epsilon$ | $< \epsilon$ |

## 4.4 Experiment 2

For the second experiment, we change the focus and we want to study how the selection of thresholds shape the initial set. In this case, we have the following outline:

1. Randomly select a proportion of target nodes $T$ ($\approx 20\%$). For all $v \in T$, set $f_v = d_i(v)$, i.e., the maximum possible value for the threshold.
2. Randomly select a proportion of initial nodes $I$ ($\approx 20\%$) such that $T \cap I = \varnothing$.
3. Compute a set of thresholds for the agents $v \in V - I - T$ with the Best Response Dynamics of the game $\Gamma_f$.
4. With the chosen thresholds and the same target, compute an initial set $I \subseteq V$ that allows for influence with the Best Response Dynamics of the game $\Gamma_I$.
5. Analyze the metrics of the found NE of $\Gamma_f$ and $\Gamma_I$.

By comparing the outcome of both models, we can see how relevant the selection of thresholds made by the agents is and study if it does affect influence or not. To fix an initial configuration for the Best Response Dynamics of the game $\Gamma_f$, we will set the thresholds of all agents $v \in V - I - T$ to one, i.e., the lowest threshold possible. Recall that in this way, the agents only need one influenced neighbor to be convinced. Notice that the nodes in the target set have the highest possible values for the threshold. This decision is made to see how the influence behaves in the worst case, i.e., when the target nodes are very suspicious about any incoming rumor.

We will make the agents select the thresholds in two different ways. In the first case, they will use the model defined in $\Gamma_f$, which is more *cooperative*, where the agents aim to select the threshold that helps influence. In the second case, we define a *malicious* type of agent that wants to avoid the influence by setting their threshold as high as possible. To do so, we define a variant of this model $\Gamma'_f$, with the following cost function:

$$c_v(s) = \begin{cases} f(v) & \text{if } T \nsubseteq F(I) \\ -f(v) & \text{otherwise} \end{cases} \tag{3}$$

In this variant, agents want to minimize their threshold when influence cannot occur and maximize it when it can, to avoid the influence. Notice that this model is the same as $\Gamma_f$, but negated, in the sense that we minimize the negated function. Conceptually, the agents try to break the influence by setting their threshold as high as possible, and when it is already broken, they try to minimize it, to avoid breaking other types of influence. Since the only thing we are doing is changing the signs of the optimization function, we can safely assume that the results shown for the game $\Gamma_f$ will also hold for this variant.

### 4.4.1 Cooperating agents

Let us first look at the influence metrics. Table 13 shows the metrics obtained from the outcome of the game $\Gamma_f$ while Table 14 the ones obtained from $\Gamma_I$. In Table 13 we see that $\pi_I = 0.2$ as expected, since we take 20% of the network as initial nodes. Regarding $\pi_F$ and $\pi_T$ we see that a random initial set produces strangely high values of influence. Furthermore, $\pi_F$ is higher in some networks like Higgs than in the previous experiment. This happens because we are choosing 20% of the nodes to be initial and recall that they also count as influenced. Therefore, it is natural that $\pi_F$ is also higher in this experiment. In fact, in Table 14 we see that the results are more similar to those of the first experiment, which proves some consistency, after we select a minimal initial seed set.

Another noticeable difference between the two outcomes is that $\pi_T$ is increased in the outcome of $\Gamma_I$ with respect to the outcome of $\Gamma_f$, but $\pi_F$ is dramatically reduced in some networks such as Epinions or Higgs. This can be explained by the fact that, the outcome of the game $\Gamma_I$ is a minimal set that maximizes the influence to the target set while minimizing the size of the seed set. Therefore, the size of the initial set is reduced from 20% to less than 1% of the networks' size, which also reduces the value of $\pi_F$.

Regarding the outcome of $\Gamma_f$, we see in Table 14, that the value of $\pi_T$ is a bit better than in the first experiment, which indicates that selecting the thresholds strategically before choosing a minimal influence set does have a positive impact on the influence of the target set, even though it is a small one.

For the centrality metrics, we do not see any new result from what we see in the first experiment; they are quite consistent.

Table 13: Table with the metrics of the second experiment with cooperative agents for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Average Out-Degree Initial | Average In-Degree Influence | Average Betweenness Initial | Average Betweenness Influence | Average Pagerank Initial | Average Pagerank Influence |
|---|---|---|---|---|---|---|---|---|---|---|
| **Dining Table** | 0.1923 | 0.6692 | 0.6800 | 1.2000 | 2.1200 | 2.0000 | 0.0000 | 0.0406 | 0.0486 | 0.0507 |
| **Dolphins** | 0.1935 | 1.0000 | 1.0000 | 2.0000 | 6.0333 | 5.1290 | 0.0000 | 0.0393 | 0.0165 | 0.0161 |
| **Human Brain** | 0.2000 | 0.9833 | 0.9750 | 1.0000 | 3.8021 | 4.2136 | 0.0000 | 0.0079 | 0.0020 | 0.0021 |
| **ArXiv** | 0.1999 | 0.9101 | 0.8845 | 1.0000 | 10.9992 | 11.8289 | 0.0000 | 0.0012 | 0.0002 | 0.0002 |
| **Wikipedia** | 0.2000 | 0.4648 | 0.3403 | 1.0000 | 14.1966 | 20.3516 | 0.0000 | 0.0001 | 0.0001 | 0.0001 |
| **Caida** | 0.2000 | 1.0000 | 1.0000 | 2.0000 | 4.4241 | 4.0326 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **ENRON** | 0.2000 | 0.9592 | 0.9476 | 1.0000 | 20.3321 | 20.7587 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **Gnutella** | 0.2000 | 0.9828 | 0.9783 | 1.0000 | 2.3599 | 2.3563 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **Epinions** | 0.2000 | 0.7170 | 0.6479 | 1.0000 | 6.8474 | 8.7791 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **Higgs** | 0.2000 | 0.2552 | 0.0684 | 1.0000 | 1.2798 | 1.2411 | 0.0000 | $< \epsilon$ | $< \epsilon$ | $< \epsilon$ |

Table 14: Table with the metrics of the second experiment with cooperative agents for the initial and influence sets. This table is the outcome of the $\Gamma_I$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Average Out-Degree Initial | Average In-Degree Influence | Average Betweenness Initial | Average Betweenness Influence | Average Pagerank Initial | Average Pagerank Influence |
|---|---|---|---|---|---|---|---|---|---|---|
| **Dining Table** | 0.0462 | 0.6308 | 0.8800 | 2.4000 | 1.1000 | 2.0000 | 0.0000 | 0.0407 | 0.0136 | 0.0512 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 4.2000 | 5.1290 | 0.0000 | 0.0393 | 0.0136 | 0.0161 |
| **Human Brain** | 0.0067 | 0.9912 | 1.0000 | 2.0000 | 1.2500 | 4.1934 | 0.0000 | 0.0078 | 0.0020 | 0.0021 |
| **ArXiv** | 0.0307 | 0.9018 | 0.9851 | 2.0000 | 3.6437 | 11.8982 | 0.0000 | 0.0012 | 0.0002 | 0.0002 |
| **Wikipedia** | 0.0018 | 0.3291 | 0.3460 | 2.0000 | 0.0000 | 24.8671 | 0.0000 | 0.0001 | 0.0001 | 0.0001 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 4.0326 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **ENRON** | 0.0120 | 0.9589 | 0.9904 | 2.0000 | 3.6343 | 20.7659 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| **Gnutella** | 0.0022 | 0.9889 | 0.9919 | 3.0000 | 0.0849 | 2.3612 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| **Epinions** | 0.0083 | 0.6534 | 0.6780 | 3.0000 | 0.5779 | 9.5042 | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 |
| **Higgs** | 0.0228 | 0.0786 | 0.1434 | 3.0000 | 0.0742 | 1.5071 | 0.0000 | $< \epsilon$ | $< \epsilon$ | $< \epsilon$ |

Turning now to the threshold distributions, we can see their plots in Figures 5 and 6. The first observation we make is that, in general, these networks have most of their nodes with values in the range $(0.9, 1]$, especially with larger networks. This is reasonable since the target set problem is a local problem, hence there will be many nodes that are useless, in the sense that the value they set to their threshold will not matter for the influence of the target set. Therefore, they will choose a threshold of one, in terms of proportion of neighbors, which is the maximum utility for them. In some networks like ArXiv we get values more fairly distributed than in other networks, probably due to the fact that ArXiv is not directed, has a higher average in-degree and the network has a topology better suited to get mostly influenced, as seen before. However, in general, it seems that thresholds tend to follow a kind of skewed normal distribution, centered at the interval $(0.4, 0.5]$. This is interesting since it shows a certain order in which the agents select their thresholds. This is not completely true for all networks, since some distributions are a bit more random, but they do seem to follow a certain order. In any case, nothing strong can be said about this.

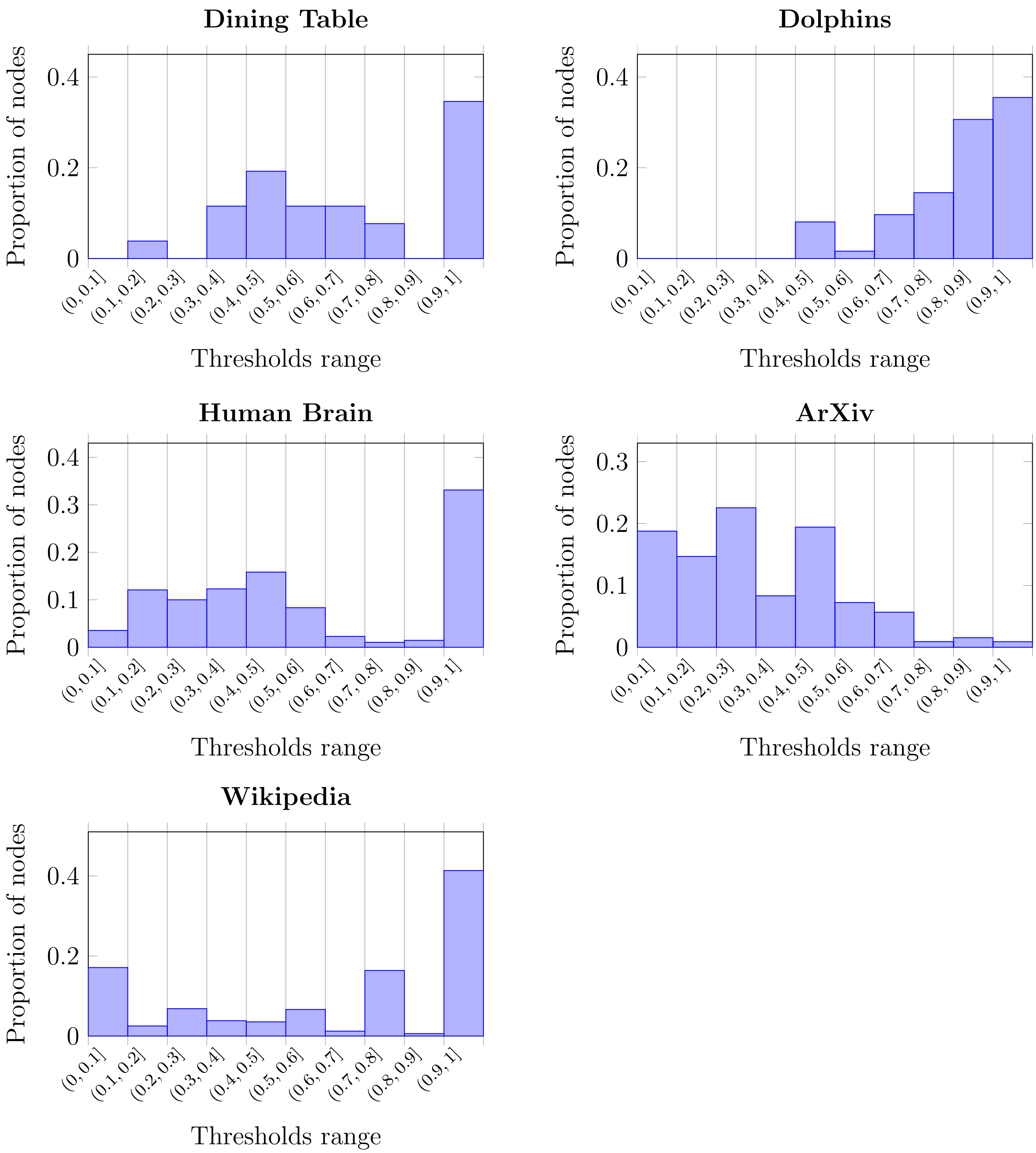


Figure 5: Threshold value ranges for the second experiment outcome with cooperative agents. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the small datasets: Dining Table, Dolphins, Human Brain, ArXiv, and Wikipedia. The results are an average of 5 repetitions of the experiment.

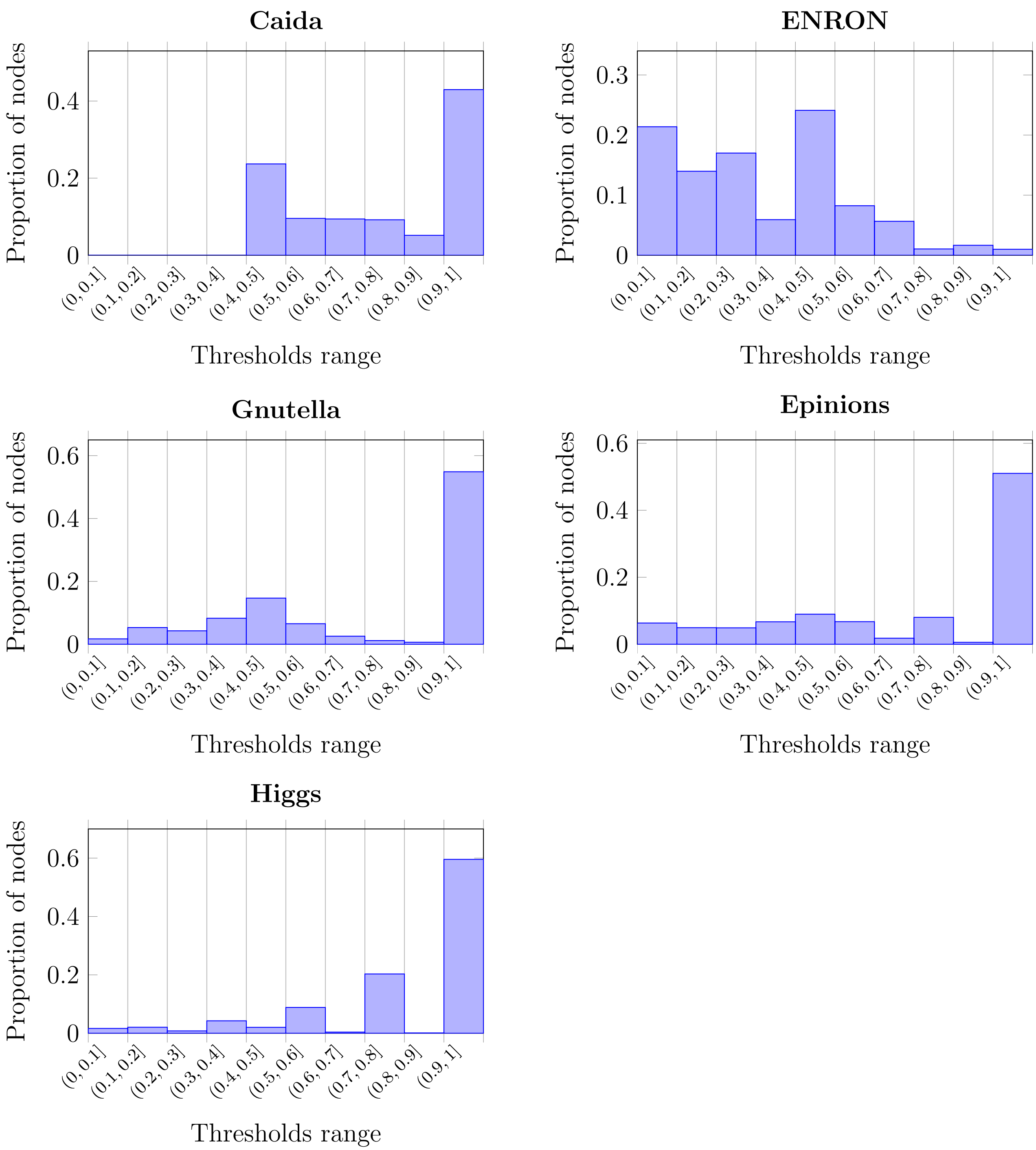


Figure 6: Threshold value ranges for the second experiment outcome with cooperative agents. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the big datasets: Caida, ENRON, Gnutella, Epinions and Higgs. The results are an average of 5 repetitions of the experiment.

### 4.4.2 Malicious agents

Recall that now the goal of the agents changes, becoming *malicious*, and aiming to minimize the influence of the target set. The metrics for the outcome of the game $\Gamma_f$ can be found in table 15 and the ones for the game $\Gamma_I$ can be found in 16. From Table 15 we see the same results as with the cooperative agents, which is consistent with what we have seen in the first experiment, since changing the thresholds with a fixed initial set does not affect the influence.

On the other hand, in Table 16 we can observe that the results are not very different from the ones obtained with the cooperative model. This is very interesting, since it shows that a selection of thresholds made to impact negatively the influence does not have any real effect. What we conclude from here is that the thresholds that the agents select are not as relevant as the selection of the initial seed set. Furthermore, it appears that even if we shield the target set from the influence of a specific seed set, another one can be found that maximizes it anyway. Apart from this, the other metrics have similar values as in the other experiments.

Table 15: Table with the metrics of the second experiment with malicious agents for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Average Out-Degree Initial | Average In-Degree Influence | Average Betweenness Initial | Average Betweenness Influence | Average Pagerank Initial | Average Pagerank Influence |
|---|---|---|---|---|---|---|---|---|---|---|
| Dining Table | 0.1923 | 0.6692 | 0.6800 | 2.0000 | 2.1200 | 2.0000 | 0.0000 | 0.0406 | 0.0486 | 0.0507 |
| Dolphins | 0.1935 | 1.0000 | 1.0000 | 2.0000 | 6.0333 | 5.1290 | 0.0000 | 0.0393 | 0.0165 | 0.0161 |
| Human Brain | 0.2000 | 0.9833 | 0.9750 | 1.0000 | 3.8021 | 4.2136 | 0.0000 | 0.0079 | 0.0020 | 0.0021 |
| ArXiv | 0.1999 | 0.9101 | 0.8845 | 1.0000 | 10.9992 | 11.8289 | 0.0000 | 0.0012 | 0.0002 | 0.0002 |
| Wikipedia | 0.2000 | 0.4648 | 0.3403 | 2.0000 | 14.1966 | 20.3516 | 0.0000 | 0.0001 | 0.0001 | 0.0001 |
| Caida | 0.2000 | 1.0000 | 1.0000 | 2.0000 | 4.4241 | 4.0326 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| ENRON | 0.2000 | 0.9592 | 0.9476 | 1.0000 | 20.3321 | 20.7587 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| Gnutella | 0.2000 | 0.9828 | 0.9783 | 2.0000 | 2.3599 | 2.3563 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Epinions | 0.2000 | 0.7176 | 0.6489 | 2.0000 | 6.3974 | 8.7724 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| Higgs | 0.2000 | 0.2556 | 0.0702 | 2.0000 | 1.3305 | 1.2405 | 0.0000 | $< \epsilon$ | $< \epsilon$ | $< \epsilon$ |

Table 16: Table with the metrics of the second experiment with malicious agents for the initial and influence sets. This table is the outcome of the $\Gamma_I$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Average Out-Degree Initial | Average In-Degree Influence | Average Betweenness Initial | Average Betweenness Influence | Average Pagerank Initial | Average Pagerank Influence |
|---|---|---|---|---|---|---|---|---|---|---|
| Dining Table | 0.0462 | 0.6308 | 0.8800 | 2.4000 | 1.1000 | 2.0000 | 0.0000 | 0.0407 | 0.0136 | 0.0512 |
| Dolphins | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 4.2000 | 5.1290 | 0.0000 | 0.0393 | 0.0136 | 0.0161 |
| Human Brain | 0.0067 | 0.9912 | 1.0000 | 2.0000 | 1.2500 | 4.1934 | 0.0000 | 0.0078 | 0.0020 | 0.0021 |
| ArXiv | 0.0307 | 0.9018 | 0.9851 | 2.0000 | 3.6437 | 11.8982 | 0.0000 | 0.0012 | 0.0002 | 0.0002 |
| Wikipedia | 0.0018 | 0.3291 | 0.3460 | 2.0000 | 0.0000 | 24.8671 | 0.0000 | 0.0001 | 0.0001 | 0.0001 |
| Caida | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 4.0326 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| ENRON | 0.0120 | 0.9589 | 0.9904 | 2.0000 | 3.6343 | 20.7659 | 0.0000 | 0.0001 | 0.0000 | 0.0000 |
| Gnutella | 0.0022 | 0.9889 | 0.9919 | 3.0000 | 0.0849 | 2.3612 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Epinions | 0.0087 | 0.6547 | 0.6796 | 3.0000 | 0.5763 | 9.4882 | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 |
| Higgs | 0.0228 | 0.0787 | 0.1442 | 3.0000 | 0.0716 | 1.5088 | 0.0000 | 0.0000 | $< \epsilon$ | $< \epsilon$ |

Lastly, Figures 7 and 8 show the plots of the threshold distribution. Interestingly enough, we see again results similar to those with the cooperative agents, except for the Wikipedia and Higgs networks. When agents want to help influence, most of the nodes set their threshold to one, while when they want to break it, most of the nodes would set their threshold as close to zero as possible. However, this effect is only observed in the Wikipedia and Higgs networks.

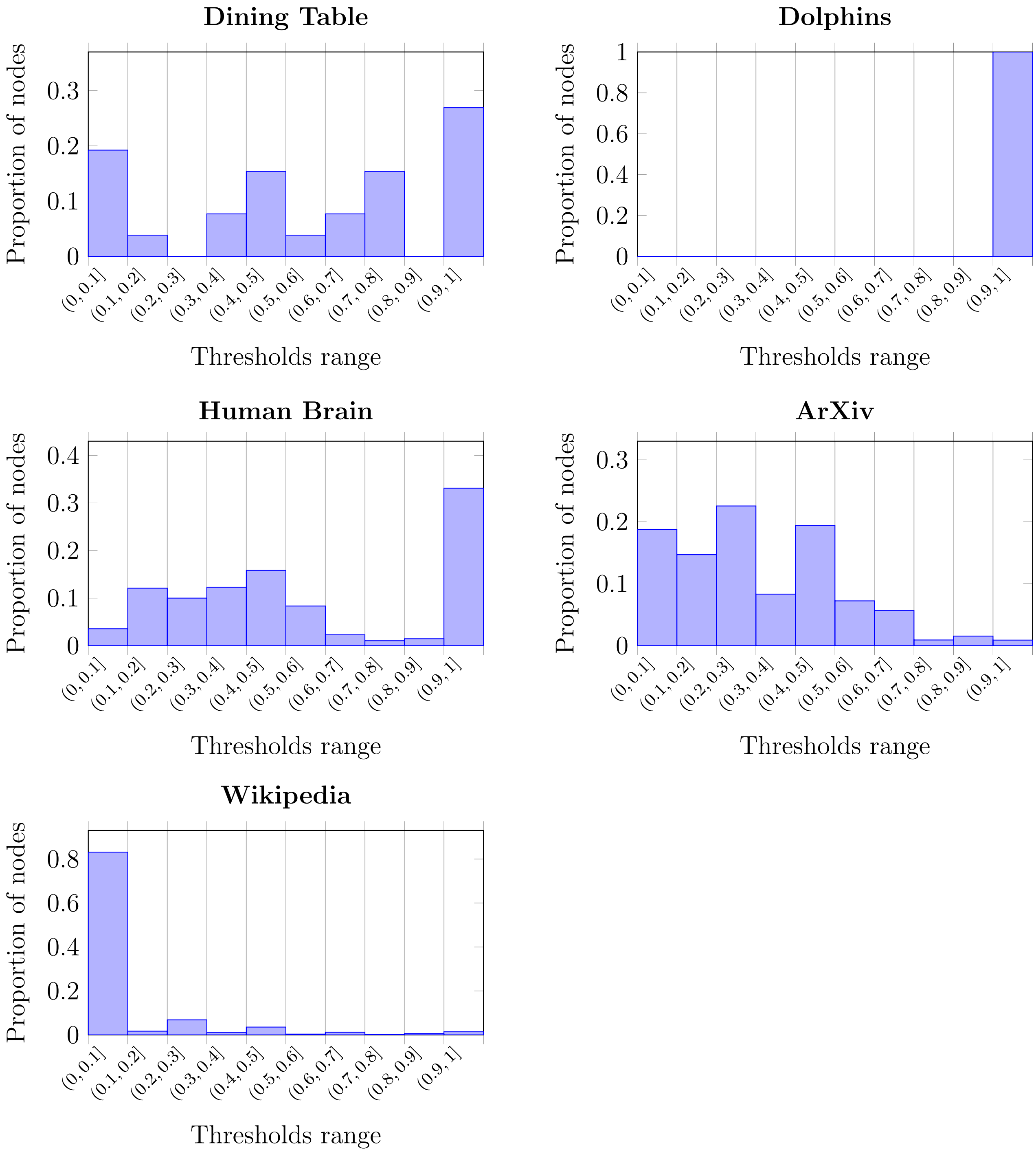


Figure 7: Threshold value ranges for the second experiment outcome with malicious agents. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the small datasets: Dining Table, Dolphins, Human Brain, ArXiv, and Wikipedia. The results are an average for 5 repetitions of the experiment.

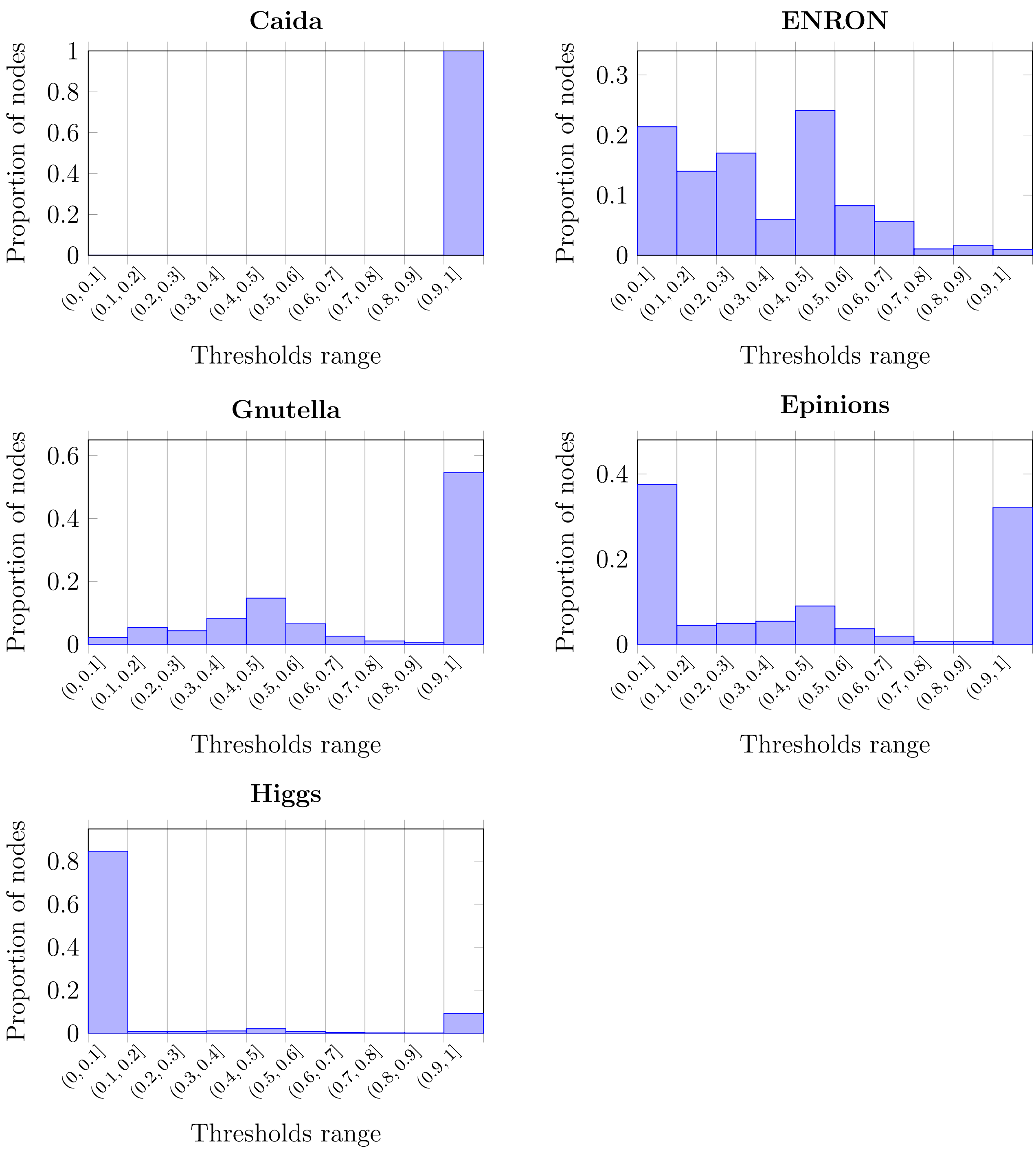


Figure 8: Threshold value ranges for the second experiment outcome with malicious agents. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the big datasets: Caida, ENRON, Gnutella, Epinions and Higgs. The results are an average for 5 repetitions of the experiment.

## 4.5 Results Summary

We have observed different behaviors in these experiments. The main result we see is that in the found NE, influence appears to depend more on the network and the selected initial set than on the thresholds. This can be observed by the fact that the experiments yield very consistent metrics and threshold distributions across the different networks. Some networks like Higgs or Wikipedia appear to be harder to influence in equilibrium than the others. This can be explained through the fact that maybe they have target nodes that can only be influenced if they are chosen as initial seed sets. Since we have considered the problem of finding an initial set that is disjoint with the target set, some nodes might never be influenced, especially in not connected or directed networks.

Regarding influence metrics, we have seen that the minimal influence set obtained with the outcome of the Seed Selection Game always achieves that $\pi_I < 0.1$, which means that the algorithm is good in minimizing the set. For the influence of the set, we cannot say something general, since the influence depends on the topology of the network, but it does seem that in most networks $\pi_F \geq 0.5$ and $\pi_T \geq 0.6$, which makes it quite efficient. See that this is not true for the Wikipedia and Higgs networks.

For the centrality metrics, we see that the initial sets have very small values, which indicate that these nodes are not lying in central parts of the network, and that their influence is reduced to a small set of neighbors. This is interesting since it shows that our model finds initial sets that are not important, in a topology sense, since on average, they have a low out-degree, and globally they are located in non-central parts of the network, which is counterintuitive. On the other hand, the average in-degree of the influenced nodes approaches the average in-degree of the whole network, which makes sense since the influence is very expansive in most instances. Regarding the betweennes and pagerank metrics, their values for the influenced nodes are a bit higher than with the initial sets, which means that they lie in a more central part of the network. Nevertheless, global metrics yield very small values due to the sparsity of the networks.

Concerning the thresholds distribution, we have plotted the proportion of nodes that select their threshold in specific ranges of values. These figures, in general, seemed to follow a sort of skewed normal distribution, except most of the values lie in the interval $(0.9, 1]$, in most cases. Since most of the nodes are useless for influence, their best response is to set their threshold to one. Nevertheless, nothing strong can be said about the distributions, since in some networks they are not as clear. For malicious agents, the thresholds seem to follow the same distribution, since we have seen that their effect is negligible, except for Wikipedia and Higgs where most agents set their threshold to a value $\leq 0.1$, as it would be expected from malicious agents.

Regarding the Seed Selection Game, we see that starting from a complete initial configuration yields slightly better NE than a random one, in terms of influence metrics. For the Threshold Selection Game, we see that the threshold selected by the agents does not have a great effect on the influence in comparison to the topology of the network and the initial set.

# 5 Conclusions

The initial goal of this project was to introduce concepts from Algorithmic Game theory into the study of Influence Maximization. In particular, we wanted to focus on the Target Set Influence problem, to deepen the knowledge on the threshold effect in the Linear Threshold expansion in unweighted social networks.

We have defined two non-cooperative games. First, the Seed Selection Game $\Gamma_I$, to solve the problem of finding an initial seed set to influence a target set of users in a network, and second, the Threshold Selection game $\Gamma_f$ to let the agents choose a threshold so as to help maximize the influence of a seed set over the target set. For both games, we proved that they are Potential Games; therefore, a Nash Equilibrium always exists, and it can be reached through a Best Response Dynamics. Furthermore, we showed that the Best Response can be computed in polynomial time, and that the number of steps needed to reach a NE starting from any configuration is polynomial in both games. For the Seed Selection Game, we proved that a strategy profile is a NE if and only if the seed set defined by the strategy profile $I_s$ is a minimal influence set when $0 < \alpha < 1$. We also proved different theorems for different values of the parameter $\alpha$. For the Threshold Selection Game, we argued that the thresholds selected by the agents, depend on whether the agents are critical nodes for the influence of the target set. Although we did not prove it, we discussed that agents that are not necessary set their thresholds equal to their incoming degree, whereas agents that are necessary set their threshold to be the number of incoming influencers. Being necessary in this context, means lying in some critical path between the seed set and the target set. Conceptually, this showcases a model in which agents try to help the influence of the target set while setting the threshold high enough, to stay critical about influences that might come from other parts of the network.

In summary, we showed that a minimal influence set and a set of thresholds that help maximize influence can be computed in polynomial time with the corresponding Best Response Dynamics of each game. Although these algorithms are not state of the art in terms of efficiency, they are very interesting, since they convey characteristics of social cooperation.

Regarding the experimental section, we have designed two experiments where we simulate the Best Response Dynamics of both models, and then gather different metrics over the influenced sets, initial sets and thresholds. We have seen different behaviors over a set of social networks of different sizes. The main result we obtained, is that influence is more dependent on the topology of the network and the initial seed set than on the value of the thresholds. We see this effect throughout all the experiments, where the networks hold very consistent values in different equilibriums, especially with the Higgs and Wikipedia networks where the influence always performs badly.

In practice, the Dynamics of the Seed Selection Game selects an initial set that makes the target highly influenced in most cases, while keeping the size of the set very small with respect to the size of the network. Furthermore, we have seen that these initial nodes, generally, have a very low out-degree, and that they are not relevant in the network, in terms of centrality. This is surprising, since it tells us that our model can find a set of nodes that greatly expand influence to the target set, but are not expected to do so, seeing that they influence a very small set of neighbors.

Regarding the thresholds, we have seen that the values strategically chosen by the agents in the Threshold Selection Game, tend to form a skewed uniform distribution, where the most chosen value is the maximum possible threshold. As argued, this is a reasonable value, since that is the Best Response for most agents in the networks. We have used two versions of the Threshold Selection Game, one where the agents try to help reach the influence of the target set, and one where they try to break the influence. We have seen that both models yield very similar equilibriums, except with the Higgs and Wikipedia networks, hence, the selection of the threshold does not have a great impact in general. It appears that, the effect of the topology of the network and the selection of the initial seed set, have much more impact on the influence than the selection of thresholds.

However, these results are generally very local, since we are putting a focus on a target set, which corresponds to a specific fraction of the network. Therefore, although our results are positive, they convey local properties, since the seed sets and the thresholds are oriented to influence the target set.

## 5.1 Future work

Since this was an introductory work, there are many new paths of study. The most basic line of work, would be to generalize the same models and experiments to a model where we have weights on the edges. For this model, it should be checked if our games hold the same properties, and if the experimental results yield the same values. For more representative results, another line of work would be to test the experiments with more networks. We wanted to use a large Amazon network [50] about item sales, however, the computation needed to obtain the metrics for this network was too high, and we could not get the results in time. Another interesting work would be to define social cost and social utilities for the games, and analyze what happens with the Price of Anarchy and the Price of Stability. These values measure how the best and worst equilibriums, in terms of social utility, are from a solution where the utility is optimized. This way, one could define metrics to measure social stability and to argue how influence relates to social work in this field. For a more in-depth study, another line of work would be to experiment on the IM problem without fixing a target set, and study how the possible NE relate to optimal or suboptimal results. Our models could be easily adapted to this idea, but it should be proved if the properties still hold. Finally, to complete our study, it would be better to repeat the experiments with different percentages of selected initial and target sets, to have a better picture.

# References

[1] Serge Galam. Modelling rumors: the no plane pentagon french hoax case. *Physica A: Statistical Mechanics and its Applications*, 320:571–580, March 2003.

[2] Michael Kosfeld. Rumours and markets. *Journal of Mathematical Economics*, 41(6):646–664, 2005.

[3] Kathie Treen, Hywel Williams, and Saffron O'Neill. Online misinformation about climate change. *WIREs Climate Change*, 11, 06 2020.

[4] Gordon Pennycook and David G. Rand. The psychology of fake news. *Trends in Cognitive Sciences*, 25(5):388–402, May 2021.

[5] Jacob Goldenberg and Barak Libai. Using complex systems analysis to advance marketing theory development: Modeling heterogeneity effects on new product growth through stochastic cellular automata. *Acad. Market. Sci. Rev*, 9, 01 2001.

[6] D. J. DALEY and D. G. KENDALL. Stochastic rumours. *IMA Journal of Applied Mathematics*, 1(1):42–55, 03 1965.

[7] Shuo Han, Fuzhen Zhuang, Qing He, Zhongzhi Shi, and Xiang Ao. Energy model for rumor propagation on social networks. *Physica A: Statistical Mechanics and its Applications*, 394:99–109, 2014.

[8] Mark Granovetter. Threshold models of collective behavior. *American Journal of Sociology*, 83(6):1420–1443, 1978.

[9] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '03, page 137–146, New York, NY, USA, 2003. Association for Computing Machinery.

[10] Pedro Domingos and Matt Richardson. Mining the network value of customers. In *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '01, page 57–66, New York, NY, USA, 2001. Association for Computing Machinery.

[11] Sunil Kumar Meena, Shashank Sheshar Singh, and Kuldeep Singh. Exploring influence maximization: State-of-the-art methods, taxonomies, and trends. *ACM Trans. Knowl. Discov. Data*, November 2025. Just Accepted.

[12] Wei Chen, Yifei Yuan, and Li Zhang. Scalable influence maximization in social networks under the linear threshold model. In *2010 IEEE International Conference on Data Mining*, pages 88–97, 2010.

[13] Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne VanBriesen, and Natalie Glance. Cost-effective outbreak detection in networks. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '07, page 420–429, New York, NY, USA, 2007. Association for Computing Machinery.

[14] Amit Goyal, Wei Lu, and Laks V.S. Lakshmanan. Celf++: optimizing the greedy algorithm for influence maximization in social networks. In *Proceedings of the 20th International Conference Companion on World Wide Web*, WWW '11, page 47–48, New York, NY, USA, 2011. Association for Computing Machinery.

[15] Suqi Cheng, Huawei Shen, Junming Huang, Guoqing Zhang, and Xueqi Cheng. Staticgreedy: solving the scalability-accuracy dilemma in influence maximization. In *Proceedings of the 22nd ACM International Conference on Information & Knowledge Management*, CIKM '13, page 509–518, New York, NY, USA, 2013. Association for Computing Machinery.

[16] Hui Li, Susu Yang, Mengting Xu, Sourav S Bhowmick, and Jiangtao Cui. Influence maximization in social networks: A survey, 2023.

[17] Myriam Jaouadi and Lotfi Ben Romdhane. A survey on influence maximization models. *Expert Systems with Applications*, 248:123429, 2024.

[18] Fa-Hsien Li, Cheng-Te Li, and Man-Kwan Shan. Labeled influence maximization in social networks for target marketing. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, pages 560–563, 2011.

[19] Wei Lu and Laks V.S. Lakshmanan. Profit maximization over social networks. In *2012 IEEE 12th International Conference on Data Mining*, pages 479–488, 2012.

[20] Furkan Gursoy and Dilek Gunnec. Influence maximization in social networks under deterministic linear threshold model. *Knowledge-Based Systems*, 161:111–123, 2018.

[21] Fabián Riquelme, Pablo Gonzalez-Cantergiani, Xavier Molinero, and Maria Serna. Centrality measure in social networks based on linear threshold model. *Knowledge-Based Systems*, 140:92–102, 2018.

[22] Xavier Molinero, Fabián Riquelme, and Maria Serna. Power indices of influence games and new centrality measures for agent societies and social networks. In Carlos Ramos, Paulo Novais, Céline Ehrwein Nihan, and Juan M. Corchado Rodríguez, editors, *Ambient Intelligence - Software and Applications*, pages 23–30, Cham, 2014. Springer International Publishing.

[23] Maria J. Blesa and Maria Serna. Playing with thresholds on the forward linear threshold rank, 2022.

[24] Alexander Kagan, Elizaveta Levina, and Ji Zhu. Flexible modeling of information diffusion on networks with statistical guarantees, 2025.

[25] Christos Papadimitriou. Algorithms, games, and the internet. page 749–753, 2001.

[26] Roger B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, 1991.

[27] John Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.

[28] Dov Monderer and Lloyd S. Shapley. Potential games. *Games and Economic Behavior*, 14(1):124–143, 1996.

[29] Steven Kuhn. Prisoner's Dilemma. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2025 edition, 2025.

[30] Pau García Rodríguez. Estudio experimental del forward linear threshold rank, June 2019.

[31] Alejandro Domínguez Besserer. Estudio de mecanismos de ponderación de influencia y su efecto en el forward linear threshold rank, January 2020.

[32] Wouter Nooy, Andrej Mrvar, and Vladimir Batagelj. *Exploratory Social Network Analysis with Pajek: Revised and Expanded Edition for Updated Software*. 07 2018.

[33] David Lusseau, Karsten Schneider, Oliver Boisseau, Patti Haase, Elisabeth Slooten, and Stephen Dawson. The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. *Behavioral Ecology and Sociobiology*, 54:396–405, 01 2003.

[34] Balázs Szalkai, Csaba Kerepesi, Bálint Varga, and Vince Grolmusz. The budapest reference connectome server v2.0. *Neuroscience Letters*, 595:60–62, 2015.

[35] Balázs Szalkai, Csaba Kerepesi, Bálint Varga, and Vince Grolmusz. Parameterizable consensus connectomes from the human connectome project: The budapest reference connectome server v3.0, 2016.

[36] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Trans. Knowl. Discov. Data*, 1(1):2–es, March 2007.

[37] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Signed networks in social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, page 1361–1370, New York, NY, USA, 2010. Association for Computing Machinery.

[38] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. KDD '05, page 177–187, New York, NY, USA, 2005. Association for Computing Machinery.

[39] Bryan Klimt and Yiming Yang. Introducing the enron corpus. In *CEAS 2004 - First Conference on Email and Anti-Spam, July 30-31, 2004, Mountain View, California, USA*, 2004.

[40] Jure Leskovec, Kevin J. Lang, Anirban Dasgupta, and Michael W. Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters, 2008.

[41] Matei Ripeanu, Adriana Iamnitchi, and Ian T. Foster. Mapping the gnutella network. *IEEE Internet Comput.*, 6(1):50–57, 2002.

[42] Matthew Richardson, Rakesh Agrawal, and Pedro Domingos. Trust management for the semantic web. volume 2870, pages 351–368, 10 2003.

[43] M Domenico, A Lima, Paul Mougel, and Mirco Musolesi. The anatomy of a scientific rumor. *Scientific reports*, 3:2980, 10 2013.

[44] Linton C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41, 1977.

[45] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1):107–117, 1998. Proceedings of the Seventh International World Wide Web Conference.

[46] García Jaya. Tfm-targetinfluencemaximizationgames. https://github.com/jayagarsi/TFM-TargetInfluenceMaximizationGames, 2026.

[47] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. Exploring network structure, dynamics, and function using networkx. In Gaël Varoquaux, Travis Vaught, and Jarrod Millman, editors, *Proceedings of the 7th Python in Science Conference*, pages 11 – 15, Pasadena, CA USA, 2008.

[48] Leonardo Dagum and Ramesh Menon. Openmp: an industry standard api for shared-memory programming. *Computational Science & Engineering, IEEE*, 5(1):46–55, 1998.

[49] Frank Yates and Ronald A. Fisher. Statistical tables for biological agricultural and medical research. *Nature*, 144(3647):533–533, 1939.

[50] Jaewon Yang and Jure Leskovec. Defining and evaluating network communities based on ground-truth. In *2012 IEEE 12th International Conference on Data Mining*, pages 745–754, 2012.

# A Appendix

This section contains the full Tables for both experiments. These also include the maximum and minimum values for each metric and each different set. To easily find the appropiate Tables and Figures we insert an index for each of them.

## List of Figures

# List of Tables

Table 17: Full Table with the metrics of the first experiment with initial threshold 0.25 and initial complete configuration for the initial and influence sets. This table is the outcome of the $\Gamma_I$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0615 | 0.7615 | 0.8400 | 2.0000 | 0.0000 | 0.2000 | 0.4000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0399 | 0.1331 | 0.0058 | 0.0125 | 0.0192 | 0.0058 | 0.0458 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 3.4000 | 3.4000 | 3.4000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0118 | 0.0118 | 0.0118 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0067 | 0.9883 | 0.9958 | 2.0000 | 1.0000 | 1.1900 | 1.6000 | 1.0000 | 4.2000 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0079 | 0.3551 | 0.0007 | 0.0034 | 0.0082 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0286 | 0.8954 | 0.9813 | 2.0000 | 2.0000 | 3.2808 | 12.4000 | 2.0000 | 11.9604 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0001 | 0.0002 | 0.0008 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0017 | 0.3289 | 0.3455 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 24.9066 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0003 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0120 | 0.9589 | 0.9921 | 2.0000 | 2.0000 | 3.4052 | 14.0000 | 2.0000 | 20.7663 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0021 | 0.9881 | 0.9916 | 2.0000 | 0.0000 | 0.1353 | 1.0000 | 0.0000 | 2.3614 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0086 | 0.6547 | 0.6778 | 2.0000 | 0.0000 | 0.3959 | 6.2000 | 0.0000 | 9.4864 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0231 | 0.0780 | 0.1436 | 2.0000 | 0.0000 | 0.0365 | 4.0000 | 0.0000 | 1.4466 | 84.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0037 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

Table 18: Full Table with the metrics of the first experiment with initial threshold 0.25 and initial complete configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| **Network** | $\pi_I$ | $\pi_F$ | $\pi_T$ | **Rounds** | **Out-Degree Initial** | | | **In-Degree Influence** | | | **Betweenness Initial** | | | **Betweenness Influence** | | | **Pagerank Initial** | | | **Pagerank Influence** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0615 | 0.7615 | 0.8400 | 1.4000 | 0.0000 | 0.2000 | 0.4000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0399 | 0.1331 | 0.0058 | 0.0125 | 0.0192 | 0.0058 | 0.0458 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 3.4000 | 3.4000 | 3.4000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0118 | 0.0118 | 0.0118 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0067 | 0.9883 | 0.9958 | 1.8000 | 1.0000 | 1.1900 | 1.6000 | 1.0000 | 4.2000 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0079 | 0.3551 | 0.0007 | 0.0034 | 0.0082 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0286 | 0.8954 | 0.9813 | 1.0000 | 2.0000 | 3.2808 | 12.4000 | 2.0000 | 11.9604 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0001 | 0.0002 | 0.0008 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0017 | 0.3289 | 0.3455 | 1.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 24.9066 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0003 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0120 | 0.9589 | 0.9921 | 1.0000 | 2.0000 | 3.4052 | 14.0000 | 2.0000 | 20.7663 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0021 | 0.9881 | 0.9916 | 1.0000 | 0.0000 | 0.1353 | 1.0000 | 0.0000 | 2.3614 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0086 | 0.6547 | 0.6778 | 1.0000 | 0.0000 | 0.3959 | 6.2000 | 0.0000 | 9.4864 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0231 | 0.0780 | 0.1436 | 1.0000 | 0.0000 | 0.0365 | 4.0000 | 0.0000 | 1.4466 | 84.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0037 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

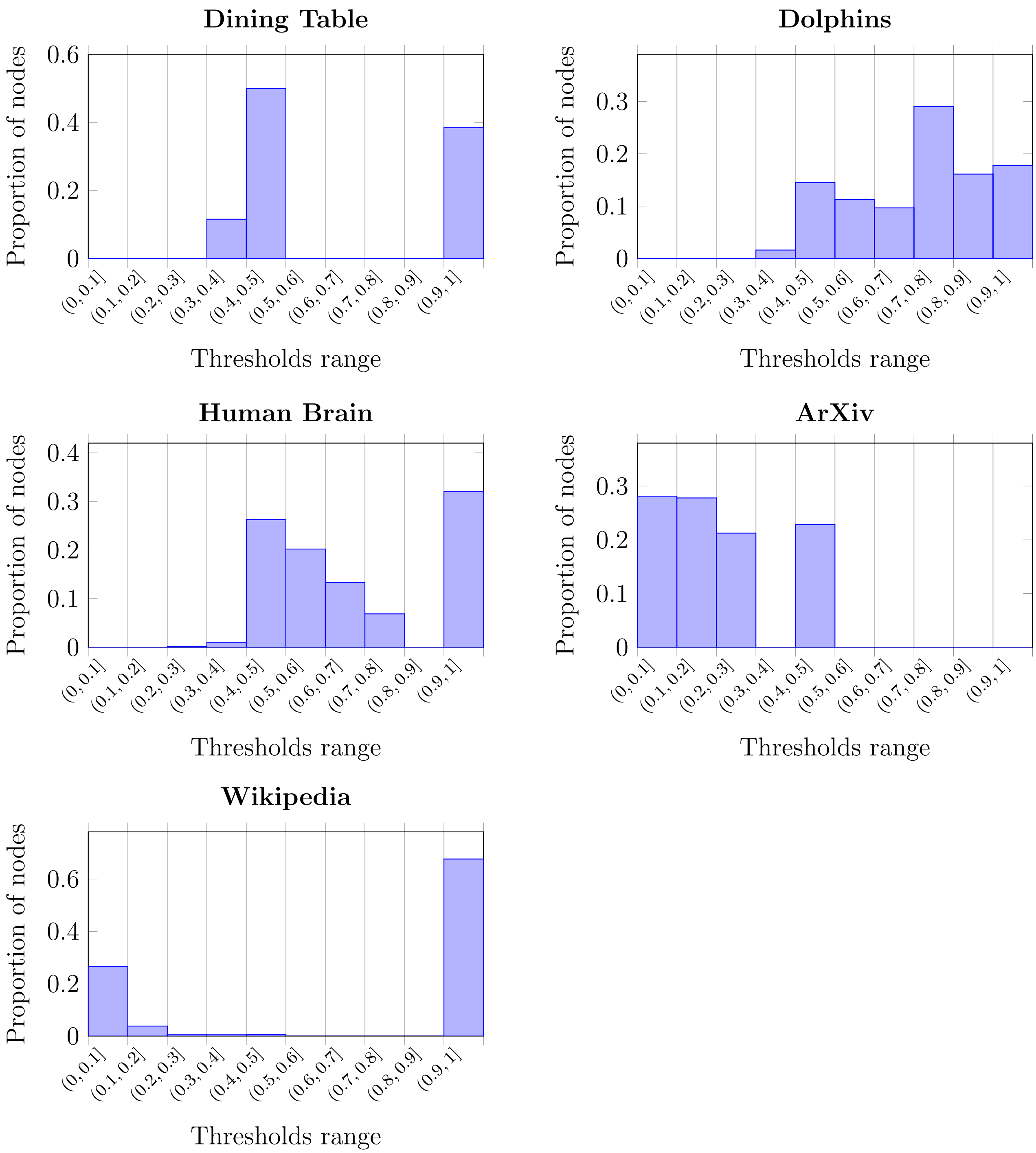


Figure 9: Threshold value ranges for the first experiment outcome starting with thresholds 0.25 and initial complete configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the small datasets: Dining Table, Dolphins, Human Brain, ArXiv, and Wikipedia. The results are an average of 5 repetitions of the experiment.

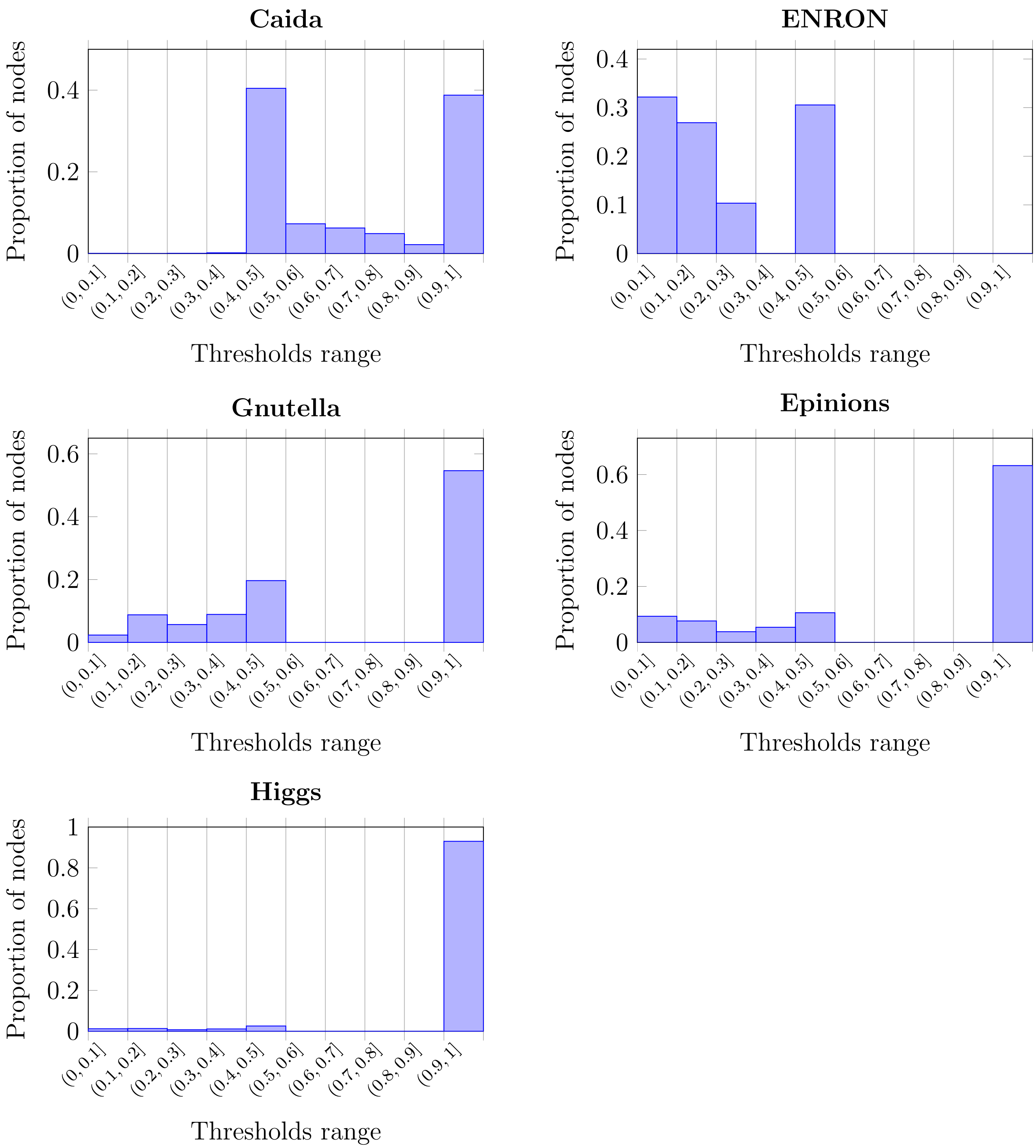


Figure 10: Threshold value ranges for the first experiment outcome starting with thresholds 0.25 and initial complete configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the big datasets: Caida, ENRON, Gnutella, Epinions and Higgs. The results are an average of 5 repetitions of the experiment.

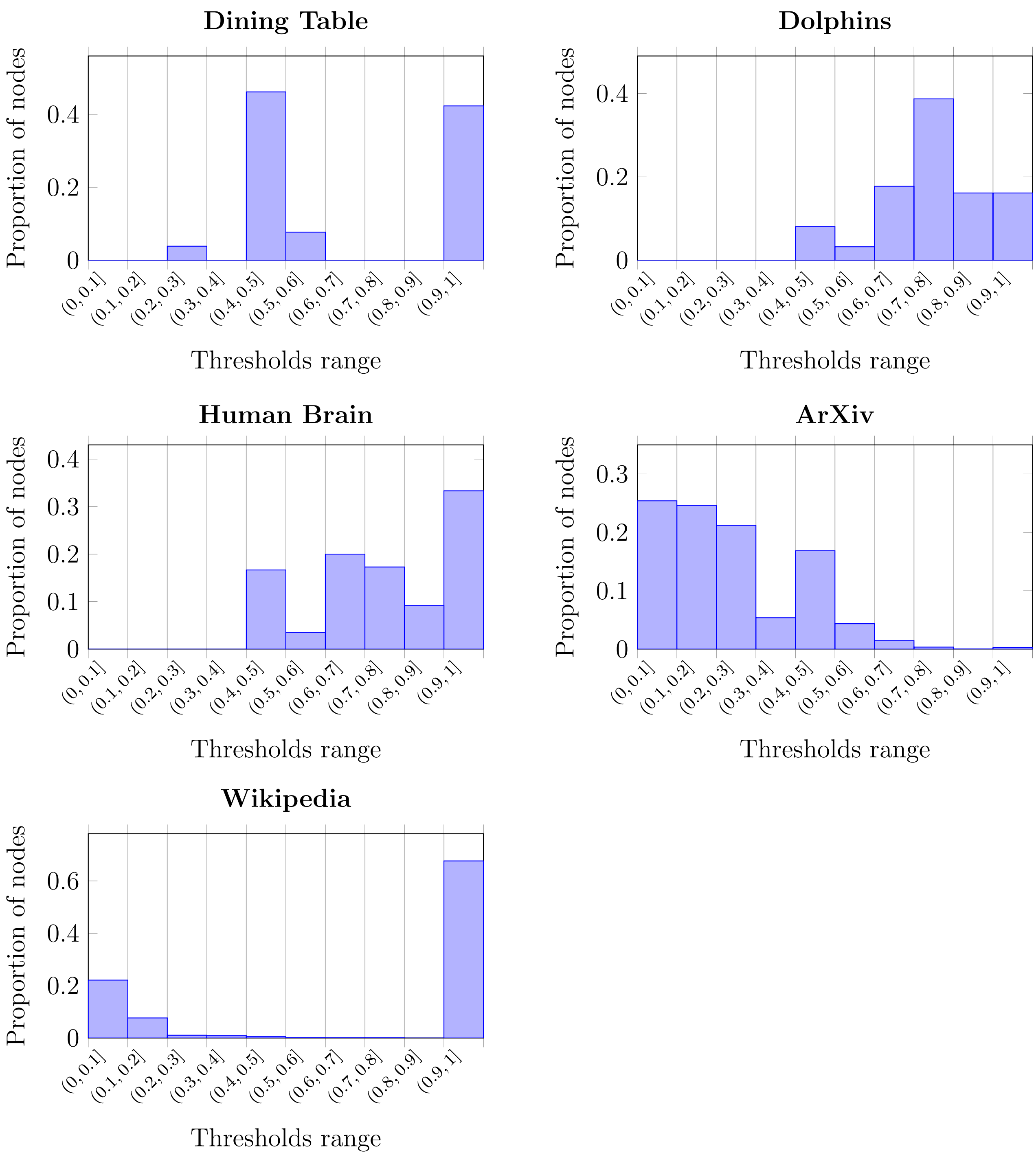


Figure 11: Threshold value ranges for the first experiment outcome starting with thresholds 0.50 and initial complete configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the small datasets: Dining Table, Dolphins, Human Brain, ArXiv, and Wikipedia. The results are an average of 5 repetitions of the experiment.

Table 19: Full Table with the metrics of the first experiment with initial threshold 0.50 and initial complete configuration for the initial and influence sets. This table is the outcome of the $\Gamma_I$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0462 | 0.6923 | 0.8800 | 2.0000 | 0.0000 | 0.1000 | 0.2000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0417 | 0.1331 | 0.0058 | 0.0060 | 0.0062 | 0.0058 | 0.0485 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 3.0000 | 3.0000 | 3.0000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0110 | 0.0110 | 0.0110 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0071 | 0.9917 | 1.0000 | 2.0000 | 1.0000 | 1.2000 | 1.6000 | 1.0000 | 4.1916 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0078 | 0.3551 | 0.0007 | 0.0027 | 0.0057 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0312 | 0.9057 | 0.9832 | 2.0000 | 2.0000 | 3.3092 | 12.0000 | 2.0000 | 11.8734 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0001 | 0.0002 | 0.0008 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0017 | 0.3290 | 0.3331 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 24.8701 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0003 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0120 | 0.9588 | 0.9902 | 2.0000 | 2.0000 | 3.3798 | 14.0000 | 2.0000 | 20.7659 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0021 | 0.9877 | 0.9913 | 2.0000 | 0.0000 | 0.1420 | 1.2000 | 0.0000 | 2.3617 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0084 | 0.6544 | 0.6765 | 2.0000 | 0.0000 | 0.3855 | 7.0000 | 0.0000 | 9.4913 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0231 | 0.0783 | 0.1441 | 2.0000 | 0.0000 | 0.0333 | 3.6000 | 0.0000 | 1.4506 | 84.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0015 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

Table 20: Full Table with the metrics of the first experiment with initial threshold 0.50 and initial complete configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0462 | 0.6923 | 0.8800 | 1.4000 | 0.0000 | 0.1000 | 0.2000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0417 | 0.1331 | 0.0058 | 0.0060 | 0.0062 | 0.0058 | 0.0485 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 3.0000 | 3.0000 | 3.0000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0110 | 0.0110 | 0.0110 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0071 | 0.9917 | 1.0000 | 2.0000 | 1.0000 | 1.2000 | 1.6000 | 1.0000 | 4.1916 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0078 | 0.3551 | 0.0007 | 0.0027 | 0.0057 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0312 | 0.9057 | 0.9832 | 1.0000 | 2.0000 | 3.3092 | 12.0000 | 2.0000 | 11.8734 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0001 | 0.0002 | 0.0008 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0017 | 0.3290 | 0.3331 | 1.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 24.8701 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0003 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0120 | 0.9588 | 0.9902 | 1.0000 | 2.0000 | 3.3798 | 14.0000 | 2.0000 | 20.7659 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0021 | 0.9877 | 0.9913 | 1.0000 | 0.0000 | 0.1420 | 1.2000 | 0.0000 | 2.3617 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0084 | 0.6544 | 0.6765 | 1.0000 | 0.0000 | 0.3855 | 7.0000 | 0.0000 | 9.4913 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0231 | 0.0783 | 0.1441 | 1.0000 | 0.0000 | 0.0333 | 3.6000 | 0.0000 | 1.4506 | 84.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0015 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

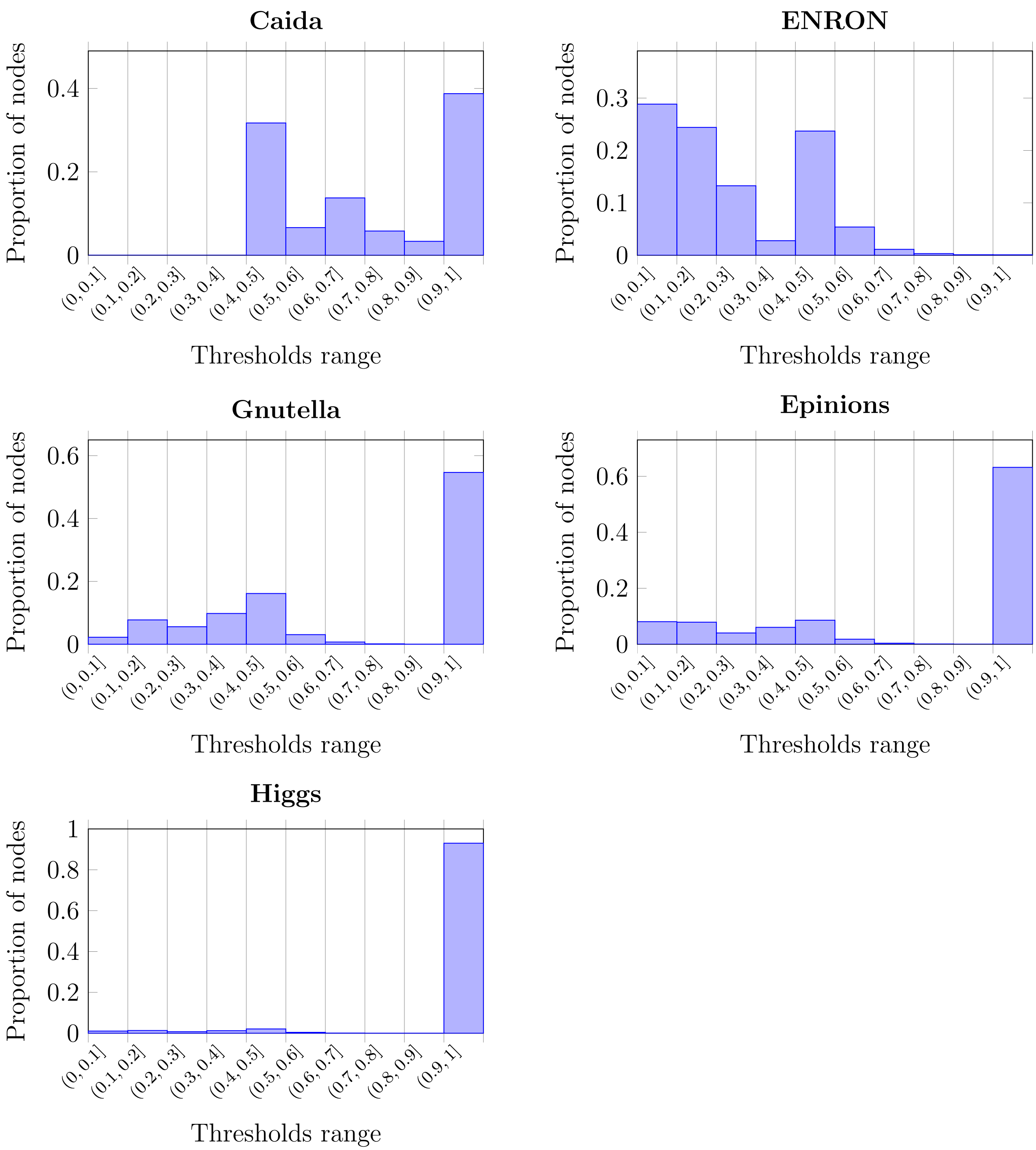


Figure 12: Threshold value ranges for the first experiment outcome starting with thresholds 0.50 and initial complete configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the big datasets: Caida, ENRON, Gnutella, Epinions and Higgs. The results are an average of 5 repetitions of the experiment.

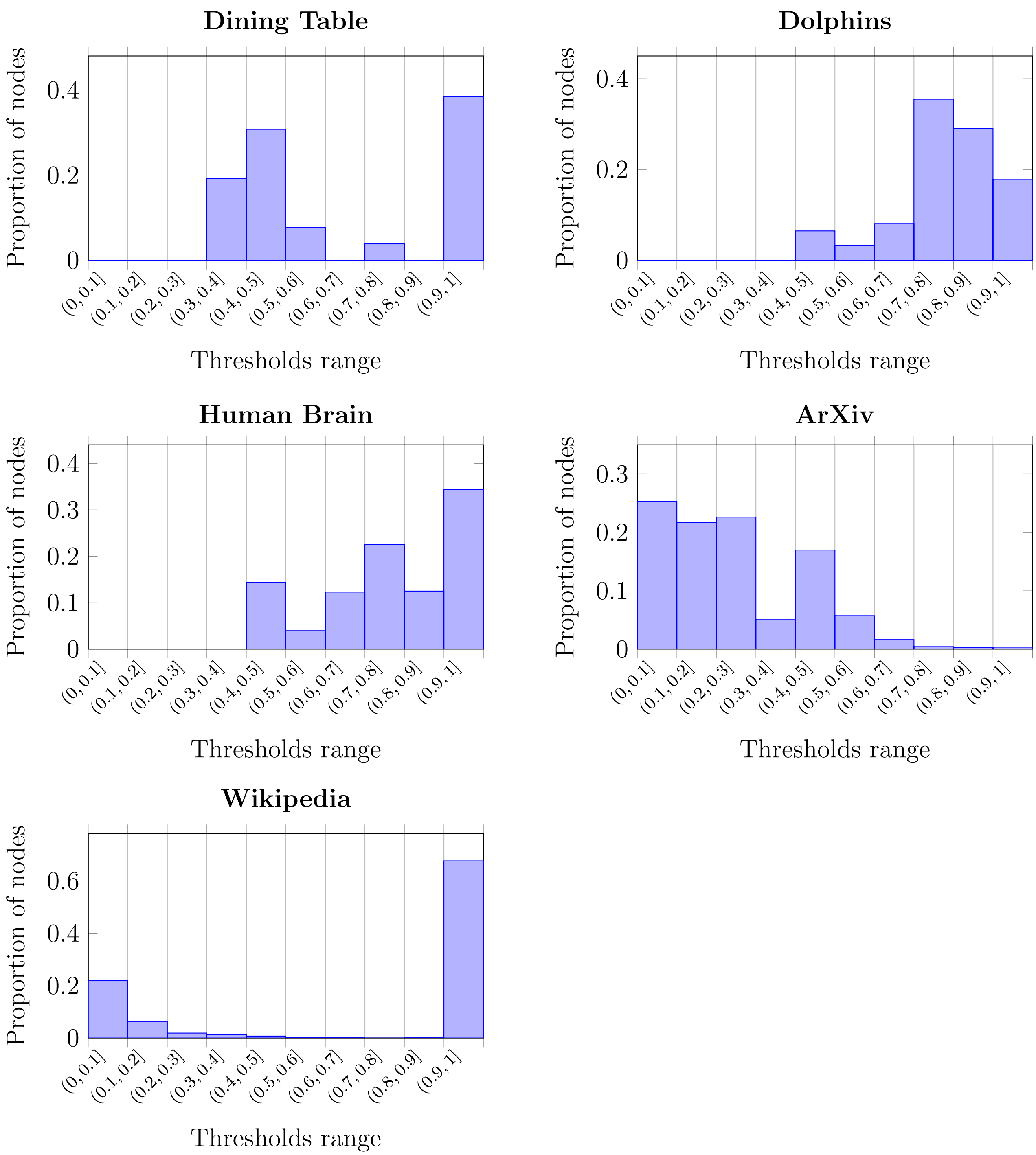


Figure 13: Threshold value ranges for the first experiment outcome starting with thresholds 0.75 and initial complete configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the small datasets: Dining Table, Dolphins, Human Brain, ArXiv, and Wikipedia. The results are an average of 5 repetitions of the experiment.

Table 21: Full Table with the metrics of the first experiment with initial threshold 0.75 and initial complete configuration for the initial and influence sets. This table is the outcome of the $\Gamma_I$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0538 | 0.7077 | 0.8400 | 2.0000 | 0.2000 | 0.4000 | 0.6000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0398 | 0.1331 | 0.0062 | 0.0098 | 0.0134 | 0.0058 | 0.0483 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 3.0000 | 3.0000 | 3.0000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0110 | 0.0110 | 0.0110 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0063 | 0.9888 | 1.0000 | 2.0000 | 1.0000 | 1.0500 | 1.2000 | 1.0000 | 4.1999 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0079 | 0.3551 | 0.0008 | 0.0028 | 0.0057 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0316 | 0.9057 | 0.9849 | 2.0000 | 2.0000 | 3.2583 | 12.0000 | 2.0000 | 11.8685 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0001 | 0.0002 | 0.0010 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0013 | 0.3282 | 0.3376 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 24.8795 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0005 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0119 | 0.9592 | 0.9914 | 2.0000 | 2.0000 | 3.3842 | 13.6000 | 2.0000 | 20.7626 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0021 | 0.9882 | 0.9915 | 2.0000 | 0.0000 | 0.1468 | 1.0000 | 0.0000 | 2.3614 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0085 | 0.6544 | 0.6766 | 2.0000 | 0.0000 | 0.3875 | 6.8000 | 0.0000 | 9.4906 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0232 | 0.0781 | 0.1443 | 2.0000 | 0.0000 | 0.0333 | 3.4000 | 0.0000 | 1.4533 | 76.8000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0015 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

Table 22: Full Table with the metrics of the first experiment with initial threshold 0.75 and initial complete configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0538 | 0.7077 | 0.8400 | 1.2000 | 0.2000 | 0.4000 | 0.6000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0398 | 0.1331 | 0.0062 | 0.0098 | 0.0134 | 0.0058 | 0.0483 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 3.0000 | 3.0000 | 3.0000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0110 | 0.0110 | 0.0110 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0063 | 0.9888 | 1.0000 | 2.0000 | 1.0000 | 1.0500 | 1.2000 | 1.0000 | 4.1999 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0079 | 0.3551 | 0.0008 | 0.0028 | 0.0057 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0316 | 0.9057 | 0.9849 | 1.0000 | 2.0000 | 3.2583 | 12.0000 | 2.0000 | 11.8685 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0001 | 0.0002 | 0.0010 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0013 | 0.3282 | 0.3376 | 1.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 24.8795 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0005 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0119 | 0.9592 | 0.9914 | 1.0000 | 2.0000 | 3.3842 | 13.6000 | 2.0000 | 20.7626 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0021 | 0.9882 | 0.9915 | 1.0000 | 0.0000 | 0.1468 | 1.0000 | 0.0000 | 2.3614 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0085 | 0.6544 | 0.6766 | 1.0000 | 0.0000 | 0.3875 | 6.8000 | 0.0000 | 9.4906 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0232 | 0.0781 | 0.1443 | 1.0000 | 0.0000 | 0.0333 | 3.4000 | 0.0000 | 1.4533 | 76.8000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0015 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

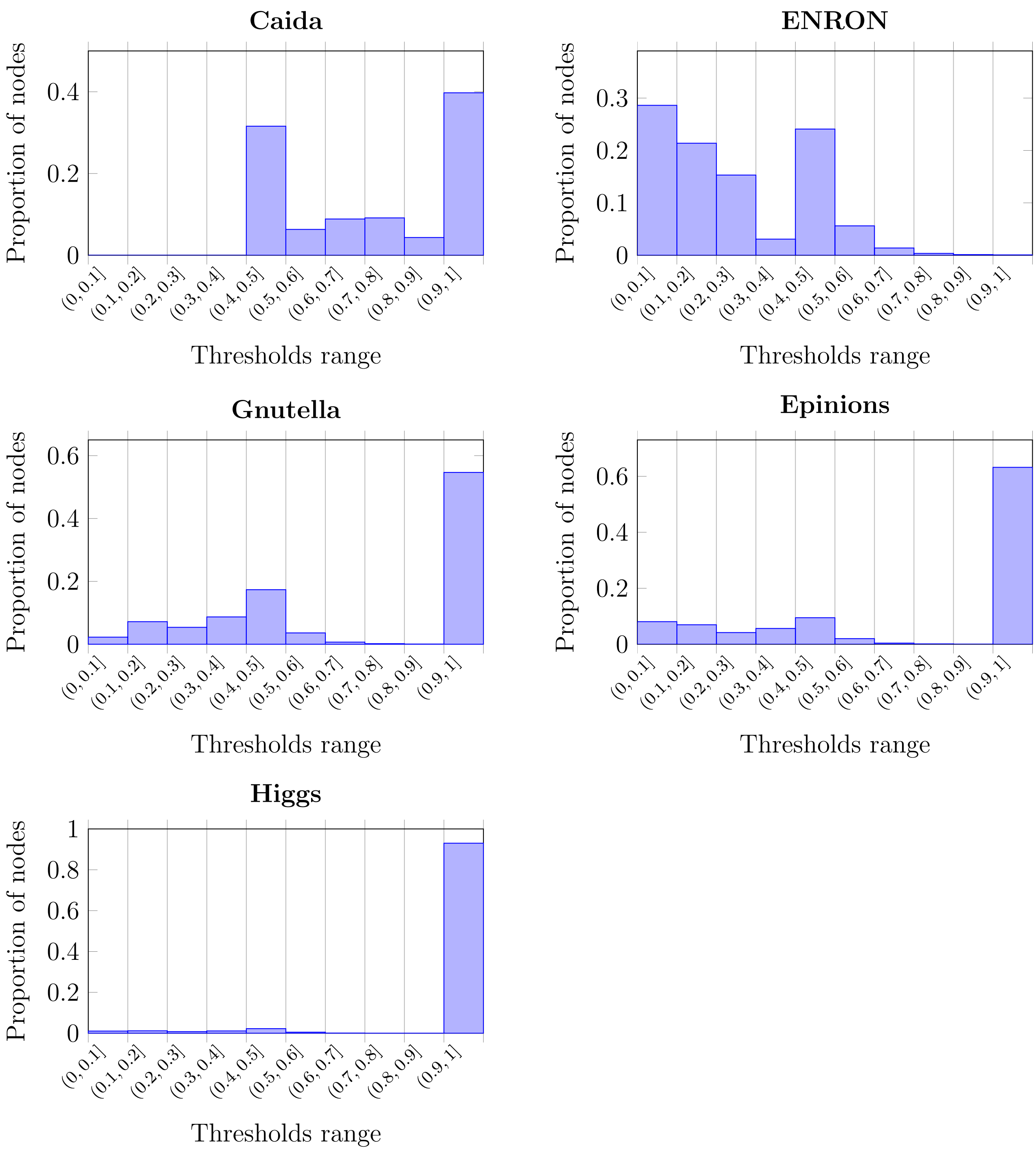


Figure 14: Threshold value ranges for the first experiment outcome starting with thresholds 0.75 and initial complete configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the big datasets: Caida, ENRON, Gnutella, Epinions and Higgs. The results are an average of 5 repetitions of the experiment.

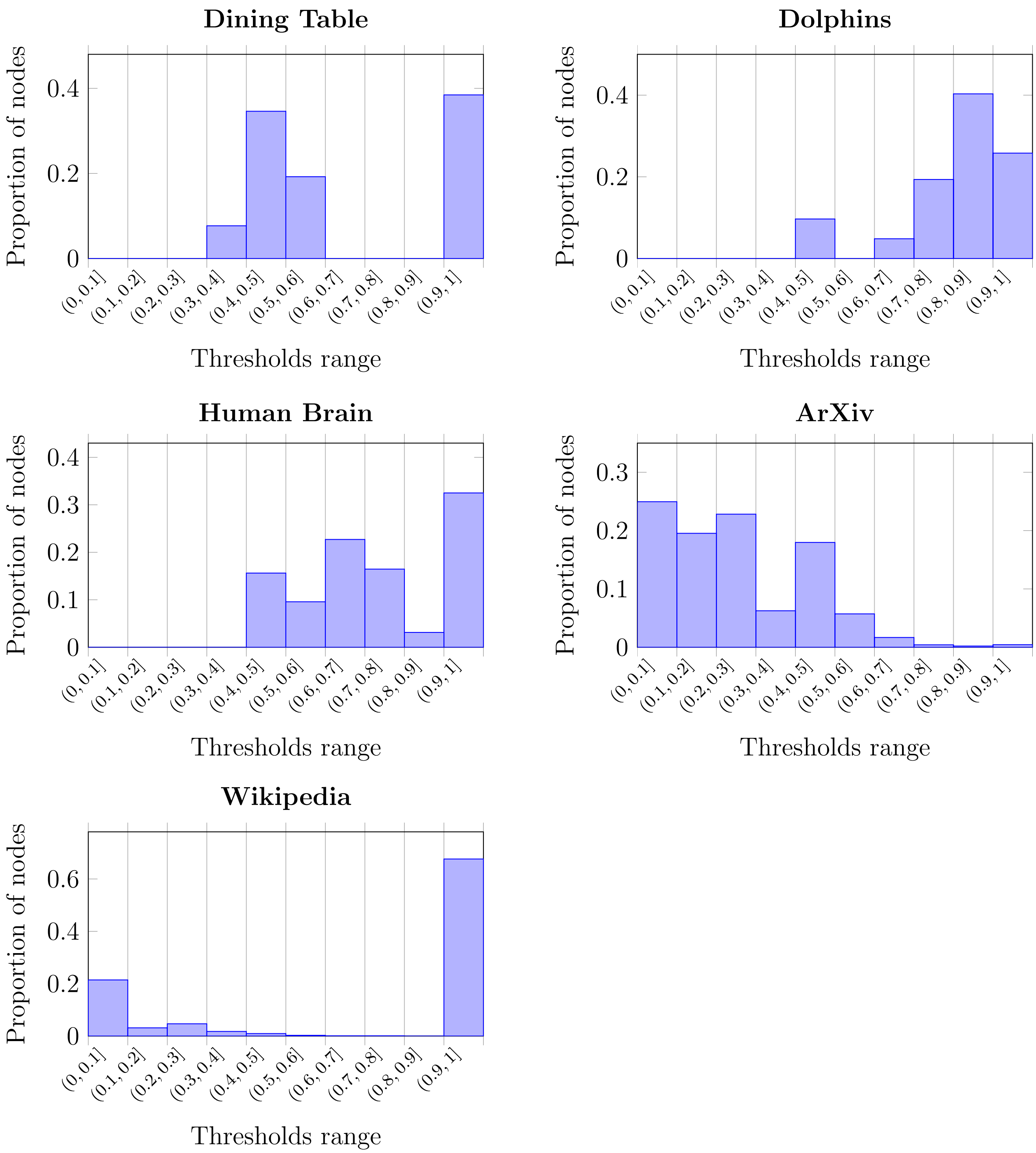


Figure 15: Threshold value ranges for the first experiment outcome starting with thresholds 0.95 and initial complete configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the small datasets: Dining Table, Dolphins, Human Brain, ArXiv, and Wikipedia. The results are an average of 5 repetitions of the experiment.

Table 23: Full Table with the metrics of the first experiment with initial threshold 0.95 and initial complete configuration for the initial and influence sets. This table is the outcome of the $\Gamma_I$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0385 | 0.6538 | 0.7600 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0433 | 0.1331 | 0.0058 | 0.0058 | 0.0058 | 0.0058 | 0.0500 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 3.0000 | 3.0000 | 3.0000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0110 | 0.0110 | 0.0110 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0071 | 0.9917 | 0.9958 | 2.0000 | 1.0000 | 1.1167 | 1.4000 | 1.0000 | 4.1917 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0078 | 0.3551 | 0.0008 | 0.0025 | 0.0057 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0301 | 0.9019 | 0.9847 | 2.0000 | 2.0000 | 3.2663 | 12.4000 | 2.0000 | 11.9023 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0000 | 0.0002 | 0.0007 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0016 | 0.3288 | 0.3361 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 24.9591 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0004 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0118 | 0.9586 | 0.9923 | 2.0000 | 2.0000 | 3.4338 | 13.6000 | 2.0000 | 20.7726 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0021 | 0.9884 | 0.9921 | 2.0000 | 0.0000 | 0.1453 | 1.0000 | 0.0000 | 2.3615 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0084 | 0.6544 | 0.6768 | 2.0000 | 0.0000 | 0.3837 | 7.0000 | 0.0000 | 9.4907 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0230 | 0.0778 | 0.1428 | 2.0000 | 0.0000 | 0.0332 | 3.6000 | 0.0000 | 1.4581 | 84.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0014 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

Table 24: Full Table with the metrics of the first experiment with initial threshold 0.95 and initial complete configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0385 | 0.6538 | 0.7600 | 1.4000 | 0.0000 | 0.0000 | 0.0000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0433 | 0.1331 | 0.0058 | 0.0058 | 0.0058 | 0.0058 | 0.0500 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 3.0000 | 3.0000 | 3.0000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0110 | 0.0110 | 0.0110 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0071 | 0.9917 | 0.9958 | 1.8000 | 1.0000 | 1.1167 | 1.4000 | 1.0000 | 4.1917 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0078 | 0.3551 | 0.0008 | 0.0025 | 0.0057 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0301 | 0.9019 | 0.9847 | 1.0000 | 2.0000 | 3.2663 | 12.4000 | 2.0000 | 11.9023 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0000 | 0.0002 | 0.0007 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0016 | 0.3288 | 0.3361 | 1.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 24.9591 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0004 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0118 | 0.9586 | 0.9923 | 1.0000 | 2.0000 | 3.4338 | 13.6000 | 2.0000 | 20.7726 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0021 | 0.9884 | 0.9921 | 1.0000 | 0.0000 | 0.1453 | 1.0000 | 0.0000 | 2.3615 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0084 | 0.6544 | 0.6768 | 1.0000 | 0.0000 | 0.3837 | 7.0000 | 0.0000 | 9.4907 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0230 | 0.0778 | 0.1428 | 1.0000 | 0.0000 | 0.0332 | 3.6000 | 0.0000 | 1.4581 | 84.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0014 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

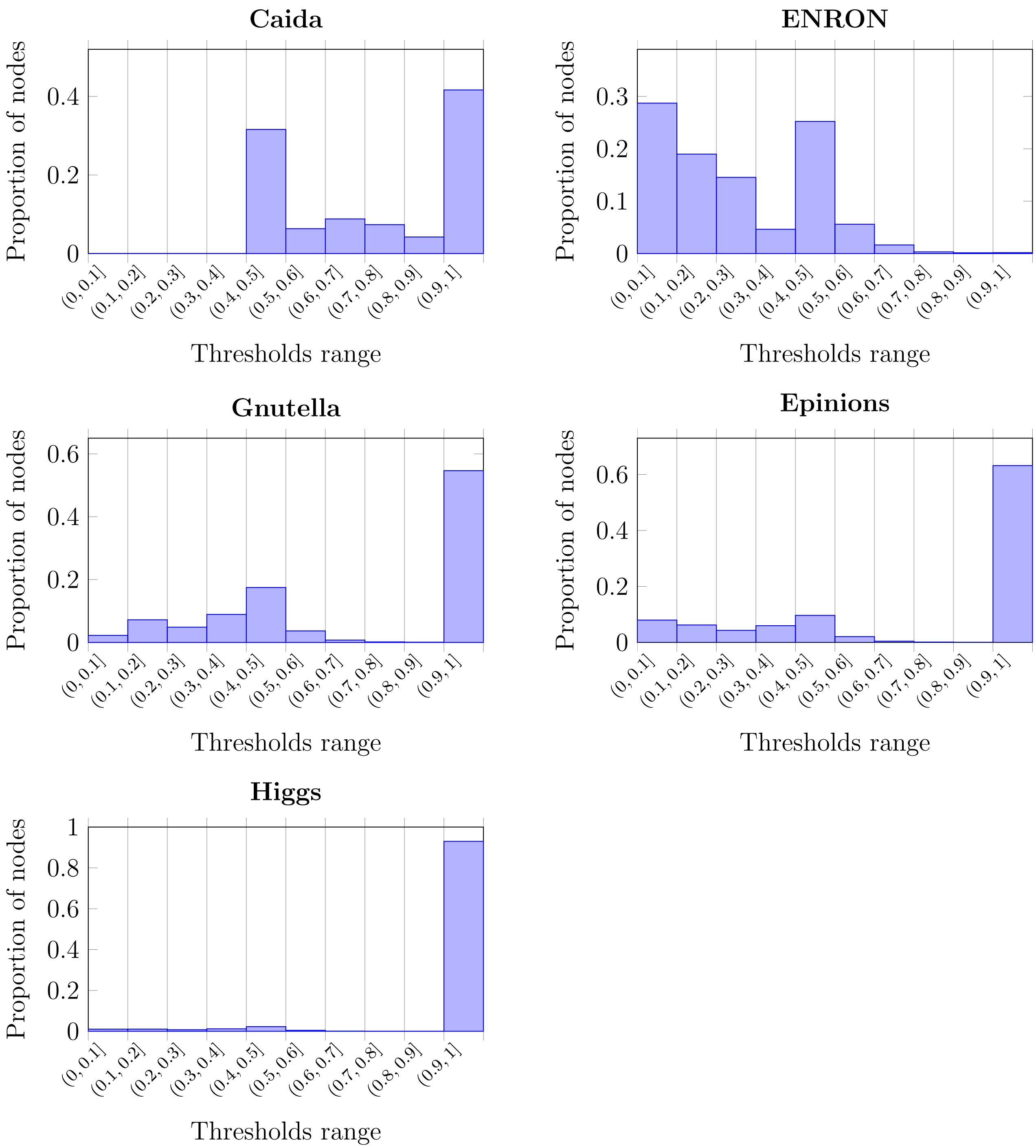


Figure 16: Threshold value ranges for the first experiment outcome starting with thresholds 0.95 and initial complete configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the big datasets: Caida, ENRON, Gnutella, Epinions and Higgs. The results are an average of 5 repetitions of the experiment.

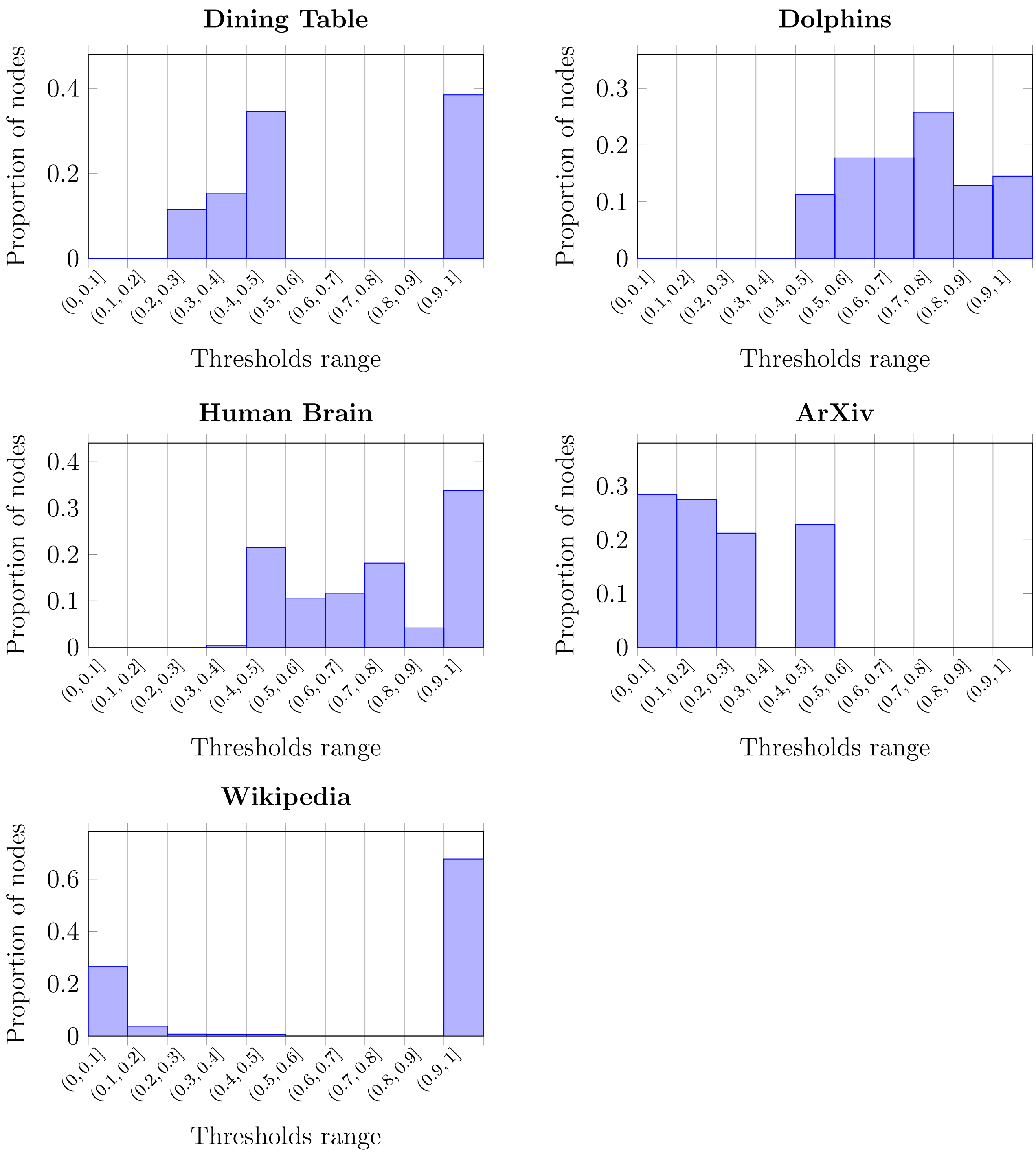


Figure 17: Threshold value ranges for the first experiment outcome starting with thresholds 0.25 and initial random configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the small datasets: Dining Table, Dolphins, Human Brain, ArXiv, and Wikipedia. The results are an average of 5 repetitions of the experiment.

Table 25: Full Table with the metrics of the first experiment with initial threshold 0.25 and initial random configuration for the initial and influence sets. This table is the outcome of the $\Gamma_I$ model. The results are an average for 5 repetitions of the experiment.

| **Network** | $\pi_I$ | $\pi_F$ | $\pi_T$ | **Rounds** | **Out-Degree Initial** | | | **In-Degree Influence** | | | **Betweenness Initial** | | | **Betweenness Influence** | | | **Pagerank Initial** | | | **Pagerank Influence** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | **Min** | **Avg** | **Max** | **Min** | **Avg** | **Max** | **Min** | **Avg** | **Max** | **Min** | **Avg** | **Max** | **Min** | **Avg** | **Max** | **Min** | **Avg** | **Max** |
| **Dining Table** | 0.0538 | 0.5923 | 0.7200 | 2.4000 | 1.8000 | 1.9000 | 2.0000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0016 | 0.0469 | 0.1331 | 0.0102 | 0.0147 | 0.0191 | 0.0062 | 0.0572 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 6.0000 | 6.0000 | 6.0000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0141 | 0.0141 | 0.0141 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0063 | 0.9904 | 1.0000 | 2.0000 | 1.2000 | 2.6667 | 4.4000 | 1.0000 | 4.1961 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0078 | 0.3551 | 0.0006 | 0.0033 | 0.0100 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0298 | 0.8993 | 0.9815 | 2.0000 | 2.0000 | 3.9118 | 18.8000 | 2.0000 | 11.9239 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0000 | 0.0002 | 0.0010 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0014 | 0.3285 | 0.3398 | 2.4000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 24.9744 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0003 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.8000 | 1.8000 | 1.8000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0118 | 0.9587 | 0.9914 | 2.0000 | 2.0000 | 3.9092 | 42.0000 | 2.0000 | 20.7703 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0021 | 0.9883 | 0.9922 | 2.0000 | 0.0000 | 0.1418 | 1.0000 | 0.0000 | 2.3605 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0087 | 0.6543 | 0.6777 | 3.0000 | 0.0000 | 0.7067 | 24.8000 | 0.0000 | 9.4954 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0229 | 0.0786 | 0.1444 | 3.0000 | 0.0000 | 0.0743 | 9.4000 | 0.0000 | 1.5031 | 80.4000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0018 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

Table 26: Full Table with the metrics of the first experiment with initial threshold 0.25 and initial random configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0538 | 0.5923 | 0.7200 | 1.2000 | 1.8000 | 1.9000 | 2.0000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0016 | 0.0469 | 0.1331 | 0.0102 | 0.0147 | 0.0191 | 0.0062 | 0.0572 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 6.0000 | 6.0000 | 6.0000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0141 | 0.0141 | 0.0141 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0063 | 0.9904 | 1.0000 | 2.0000 | 1.2000 | 2.6667 | 4.4000 | 1.0000 | 4.1961 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0078 | 0.3551 | 0.0006 | 0.0033 | 0.0100 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0298 | 0.8993 | 0.9815 | 1.0000 | 2.0000 | 3.9118 | 18.8000 | 2.0000 | 11.9239 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0000 | 0.0002 | 0.0010 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0014 | 0.3285 | 0.3398 | 1.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 24.9744 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0003 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.8000 | 1.8000 | 1.8000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0118 | 0.9587 | 0.9914 | 1.0000 | 2.0000 | 3.9092 | 42.0000 | 2.0000 | 20.7703 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0021 | 0.9883 | 0.9922 | 1.0000 | 0.0000 | 0.1418 | 1.0000 | 0.0000 | 2.3605 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0087 | 0.6543 | 0.6777 | 1.0000 | 0.0000 | 0.7067 | 24.8000 | 0.0000 | 9.4954 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0229 | 0.0786 | 0.1444 | 1.0000 | 0.0000 | 0.0743 | 9.4000 | 0.0000 | 1.5031 | 80.4000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0018 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

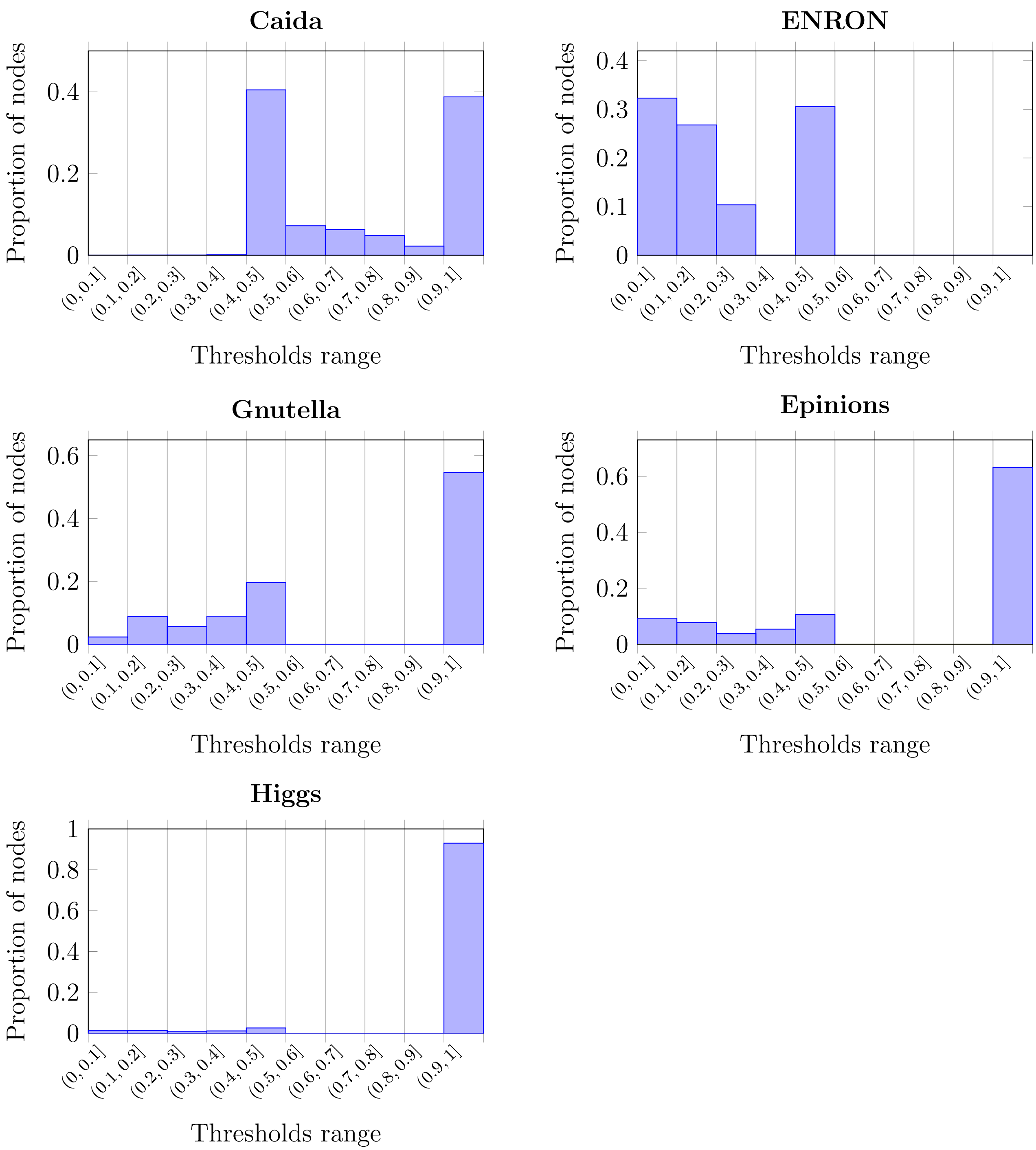


Figure 18: Threshold value ranges for the first experiment outcome starting with thresholds 0.25 and initial random configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the big datasets: Caida, ENRON, Gnutella, Epinions and Higgs. The results are an average of 5 repetitions of the experiment.

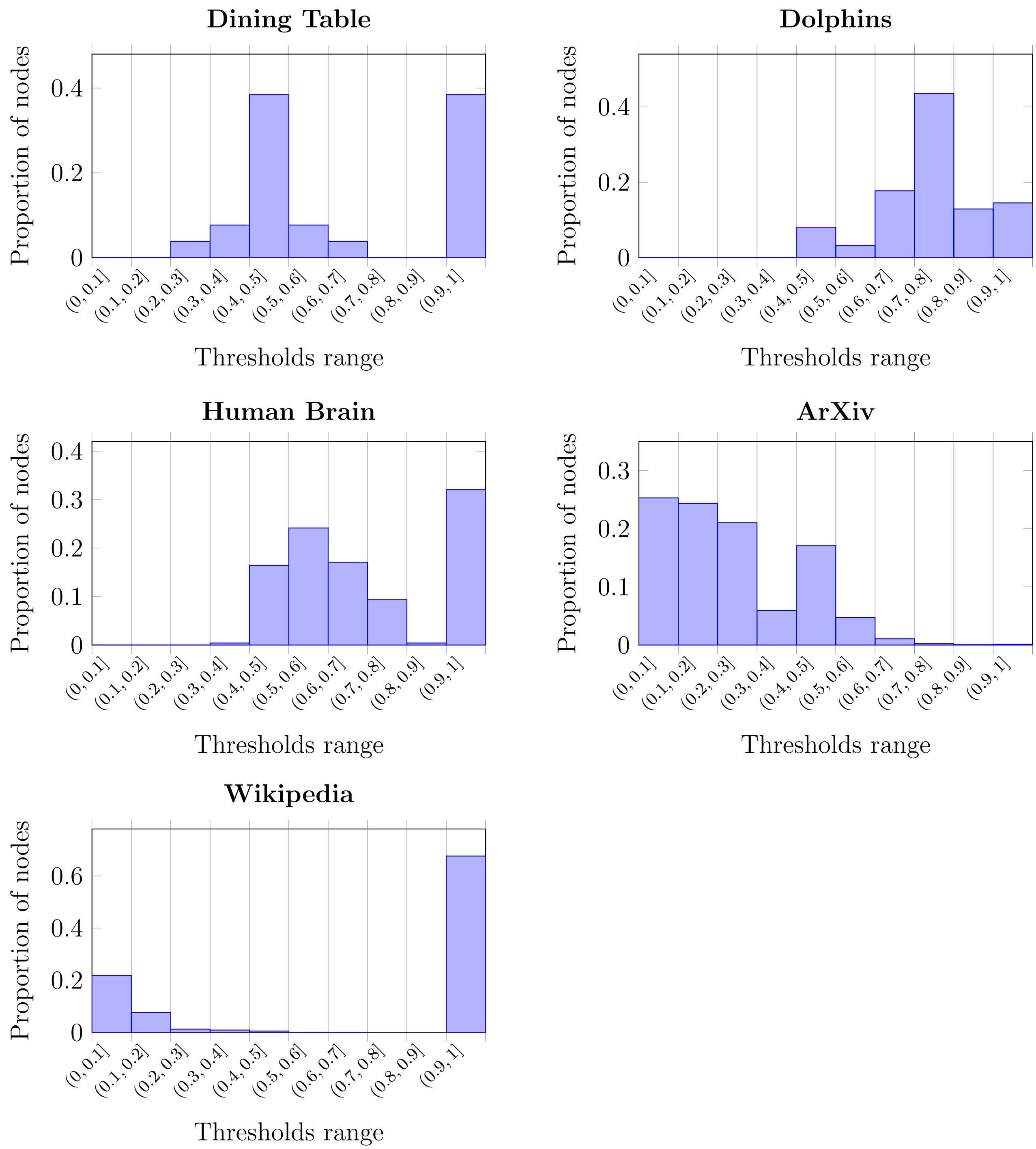


Figure 19: Threshold value ranges for the first experiment outcome starting with thresholds 0.50 and initial random configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the small datasets: Dining Table, Dolphins, Human Brain, ArXiv, and Wikipedia. The results are an average of 5 repetitions of the experiment.

Table 27: Full Table with the metrics of the first experiment with initial threshold 0.50 and initial random configuration for the initial and influence sets. This table is the outcome of the $\Gamma_I$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0462 | 0.6077 | 0.8000 | 2.4000 | 0.8000 | 1.1000 | 1.4000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0439 | 0.1331 | 0.0081 | 0.0092 | 0.0102 | 0.0058 | 0.0533 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 6.6000 | 6.6000 | 6.6000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0191 | 0.0191 | 0.0191 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0058 | 0.9883 | 0.9958 | 2.0000 | 1.8000 | 2.7800 | 4.4000 | 1.0000 | 4.2017 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0079 | 0.3551 | 0.0009 | 0.0030 | 0.0052 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0308 | 0.9036 | 0.9826 | 2.0000 | 2.0000 | 3.8809 | 16.0000 | 2.0000 | 11.8849 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0000 | 0.0002 | 0.0008 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0015 | 0.3286 | 0.3324 | 2.4000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 25.0136 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0004 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.8000 | 1.8000 | 1.8000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0122 | 0.9592 | 0.9917 | 2.0000 | 2.0000 | 3.8071 | 24.4000 | 2.0000 | 20.7595 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0022 | 0.9885 | 0.9912 | 2.2000 | 0.0000 | 0.1556 | 1.2000 | 0.0000 | 2.3616 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0085 | 0.6538 | 0.6783 | 3.0000 | 0.0000 | 0.7213 | 16.2000 | 0.0000 | 9.5006 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0231 | 0.0796 | 0.1453 | 3.0000 | 0.0000 | 0.1002 | 12.4000 | 0.0000 | 1.5203 | 86.8000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0046 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

Table 28: Full Table with the metrics of the first experiment with initial threshold 0.50 and initial random configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0462 | 0.6077 | 0.8000 | 1.4000 | 0.8000 | 1.1000 | 1.4000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0439 | 0.1331 | 0.0081 | 0.0092 | 0.0102 | 0.0058 | 0.0533 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 6.6000 | 6.6000 | 6.6000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0191 | 0.0191 | 0.0191 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0058 | 0.9883 | 0.9958 | 1.8000 | 1.8000 | 2.7800 | 4.4000 | 1.0000 | 4.2017 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0079 | 0.3551 | 0.0009 | 0.0030 | 0.0052 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0308 | 0.9036 | 0.9826 | 1.0000 | 2.0000 | 3.8809 | 16.0000 | 2.0000 | 11.8849 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0000 | 0.0002 | 0.0008 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0015 | 0.3286 | 0.3324 | 1.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 25.0136 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0004 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.8000 | 1.8000 | 1.8000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0122 | 0.9592 | 0.9917 | 1.0000 | 2.0000 | 3.8071 | 24.4000 | 2.0000 | 20.7595 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0022 | 0.9885 | 0.9912 | 1.0000 | 0.0000 | 0.1556 | 1.2000 | 0.0000 | 2.3616 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0085 | 0.6538 | 0.6783 | 1.0000 | 0.0000 | 0.7213 | 16.2000 | 0.0000 | 9.5006 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0231 | 0.0796 | 0.1453 | 1.0000 | 0.0000 | 0.1002 | 12.4000 | 0.0000 | 1.5203 | 86.8000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0046 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

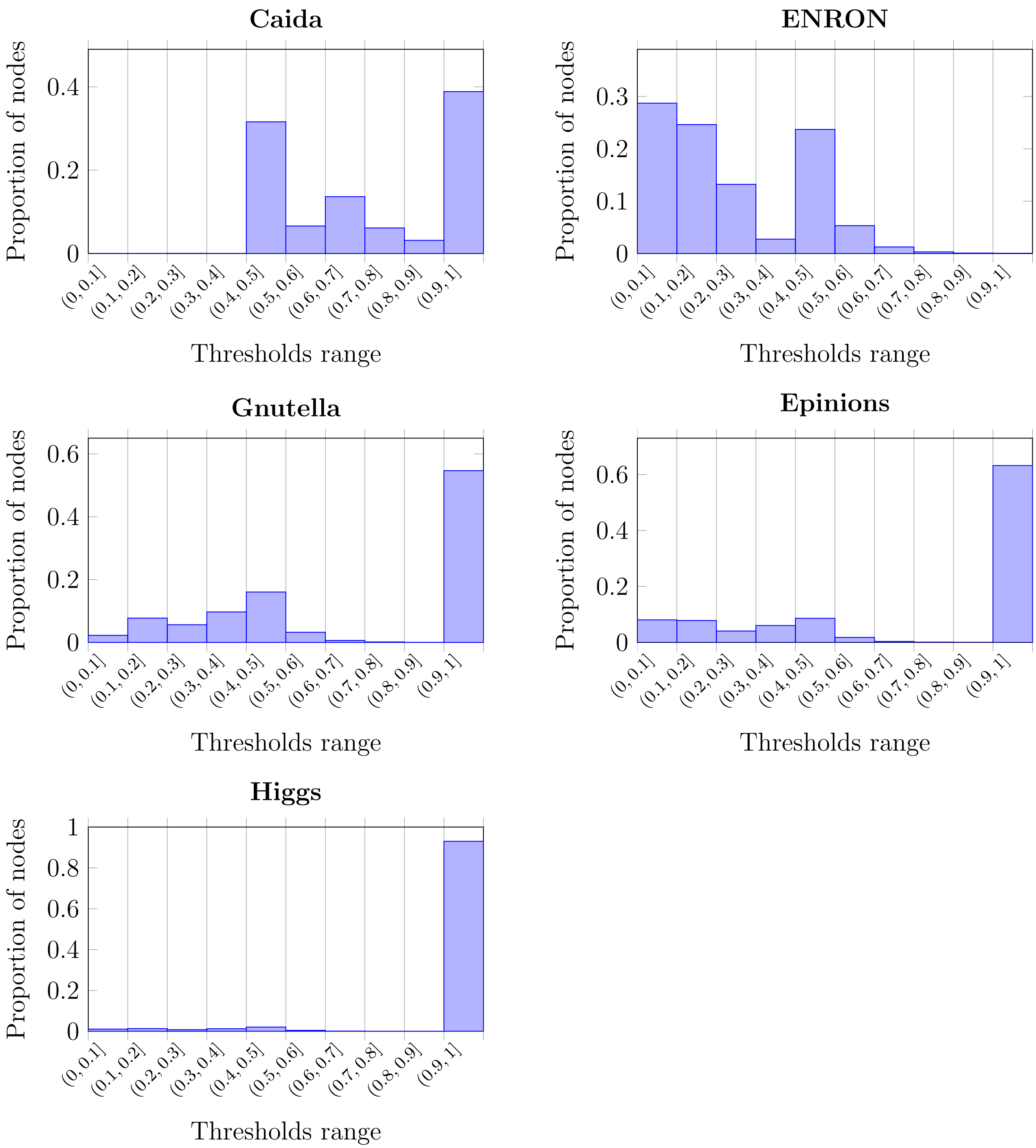


Figure 20: Threshold value ranges for the first experiment outcome starting with thresholds 0.50 and initial random configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the big datasets: Caida, ENRON, Gnutella, Epinions and Higgs. The results are an average of 5 repetitions of the experiment.

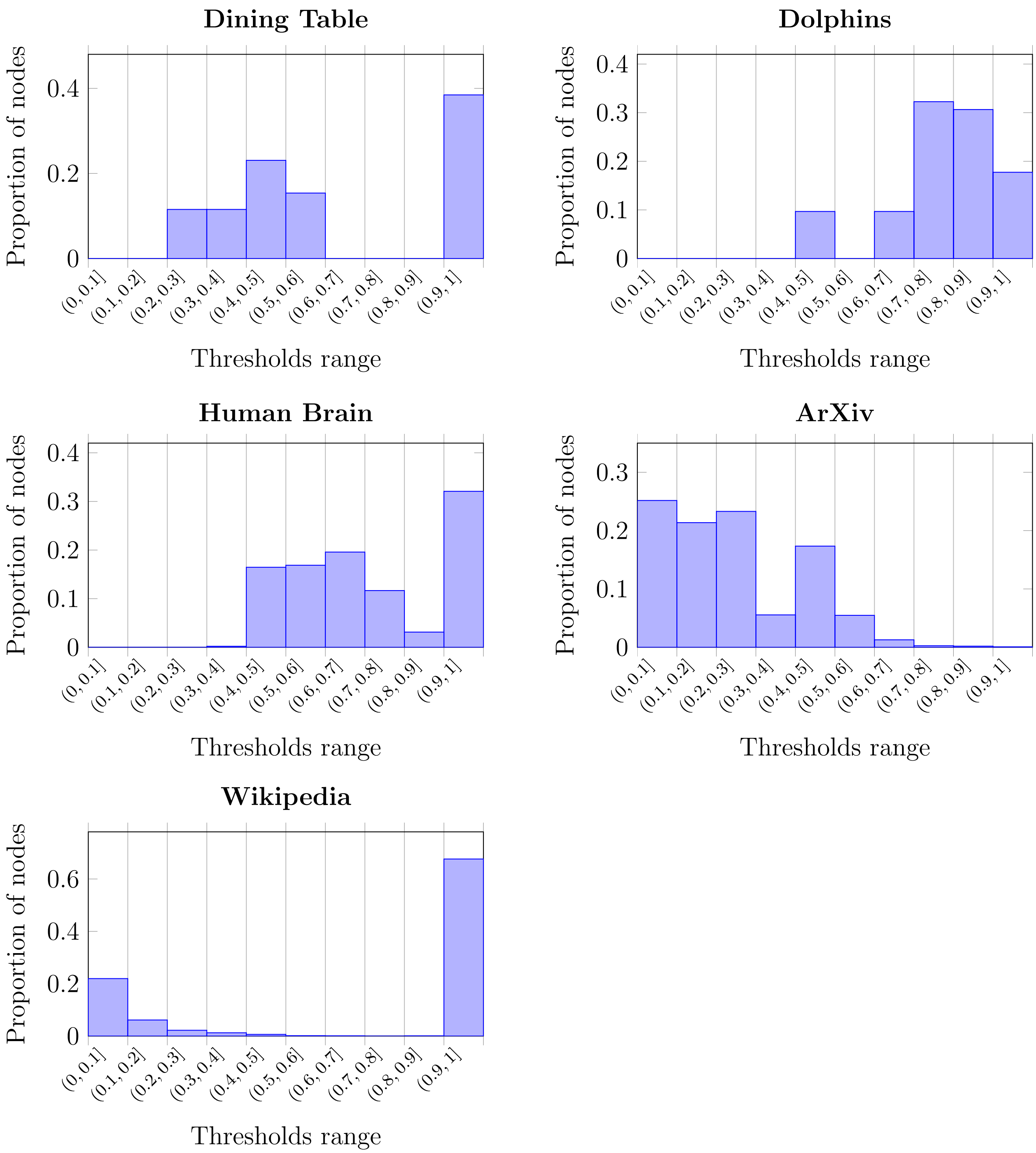


Figure 21: Threshold value ranges for the first experiment outcome starting with thresholds 0.75 and initial random configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the small datasets: Dining Table, Dolphins, Human Brain, ArXiv, and Wikipedia. The results are an average of 5 repetitions of the experiment.

Table 29: Full Table with the metrics of the first experiment with initial threshold 0.75 and initial random configuration for the initial and influence sets. This table is the outcome of the $\Gamma_I$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0538 | 0.6385 | 0.7600 | 2.2000 | 1.2000 | 1.5000 | 1.8000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0016 | 0.0458 | 0.1331 | 0.0129 | 0.0196 | 0.0262 | 0.0062 | 0.0535 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 6.0000 | 6.0000 | 6.0000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0194 | 0.0194 | 0.0194 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0054 | 0.9850 | 0.9958 | 2.0000 | 1.8000 | 2.1467 | 2.6000 | 1.0000 | 4.2101 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0079 | 0.3551 | 0.0010 | 0.0016 | 0.0024 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0300 | 0.8992 | 0.9884 | 2.0000 | 2.0000 | 3.6723 | 15.2000 | 2.0000 | 11.9172 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0001 | 0.0002 | 0.0009 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0015 | 0.3287 | 0.3244 | 2.2000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 25.1070 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0003 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 2.6000 | 2.6000 | 2.6000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0119 | 0.9590 | 0.9907 | 2.0000 | 2.0000 | 3.8303 | 22.8000 | 2.0000 | 20.7652 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0022 | 0.9885 | 0.9917 | 2.0000 | 0.0000 | 0.1378 | 1.2000 | 0.0000 | 2.3614 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0001 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0084 | 0.6537 | 0.6755 | 3.0000 | 0.0000 | 0.7317 | 10.8000 | 0.0000 | 9.5030 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0229 | 0.0787 | 0.1446 | 3.0000 | 0.0000 | 0.0790 | 8.2000 | 0.0000 | 1.4962 | 70.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0036 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

Table 30: Full Table with the metrics of the first experiment with initial threshold 0.75 and initial random configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0538 | 0.6385 | 0.7600 | 1.2000 | 1.2000 | 1.5000 | 1.8000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0016 | 0.0458 | 0.1331 | 0.0129 | 0.0196 | 0.0262 | 0.0062 | 0.0535 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 6.0000 | 6.0000 | 6.0000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0194 | 0.0194 | 0.0194 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0054 | 0.9850 | 0.9958 | 1.8000 | 1.8000 | 2.1467 | 2.6000 | 1.0000 | 4.2101 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0079 | 0.3551 | 0.0010 | 0.0016 | 0.0024 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0300 | 0.8992 | 0.9884 | 1.0000 | 2.0000 | 3.6723 | 15.2000 | 2.0000 | 11.9172 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0001 | 0.0002 | 0.0009 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0015 | 0.3287 | 0.3244 | 1.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 25.1070 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0003 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 2.6000 | 2.6000 | 2.6000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0119 | 0.9590 | 0.9907 | 1.0000 | 2.0000 | 3.8303 | 22.8000 | 2.0000 | 20.7652 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0022 | 0.9885 | 0.9917 | 1.0000 | 0.0000 | 0.1378 | 1.2000 | 0.0000 | 2.3614 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0001 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0084 | 0.6537 | 0.6755 | 1.0000 | 0.0000 | 0.7317 | 10.8000 | 0.0000 | 9.5030 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0229 | 0.0787 | 0.1446 | 1.0000 | 0.0000 | 0.0790 | 8.2000 | 0.0000 | 1.4962 | 70.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0036 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

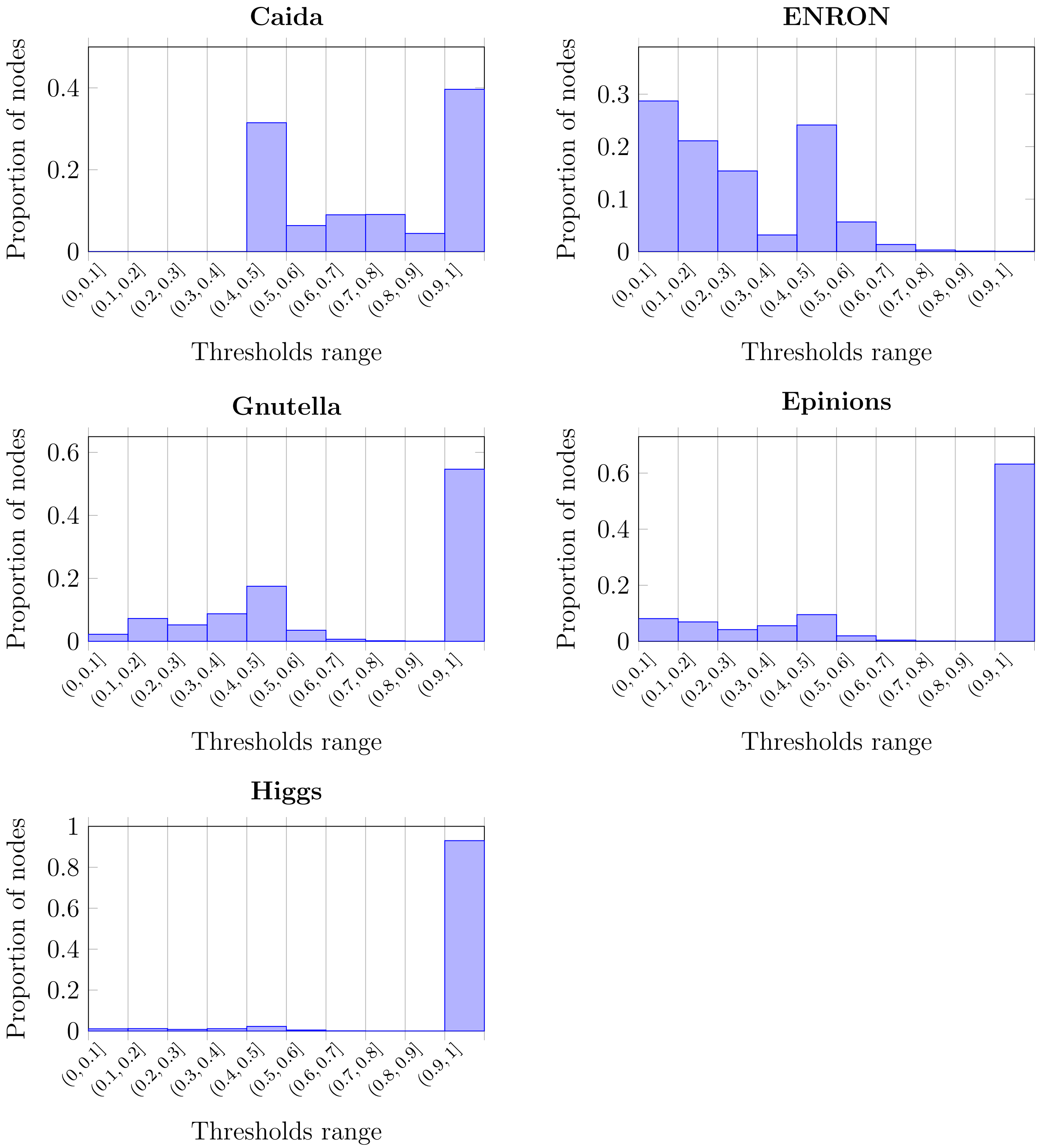


Figure 22: Threshold value ranges for the first experiment outcome starting with thresholds 0.75 and initial random configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the big datasets: Caida, ENRON, Gnutella, Epinions and Higgs. The results are an average of 5 repetitions of the experiment.

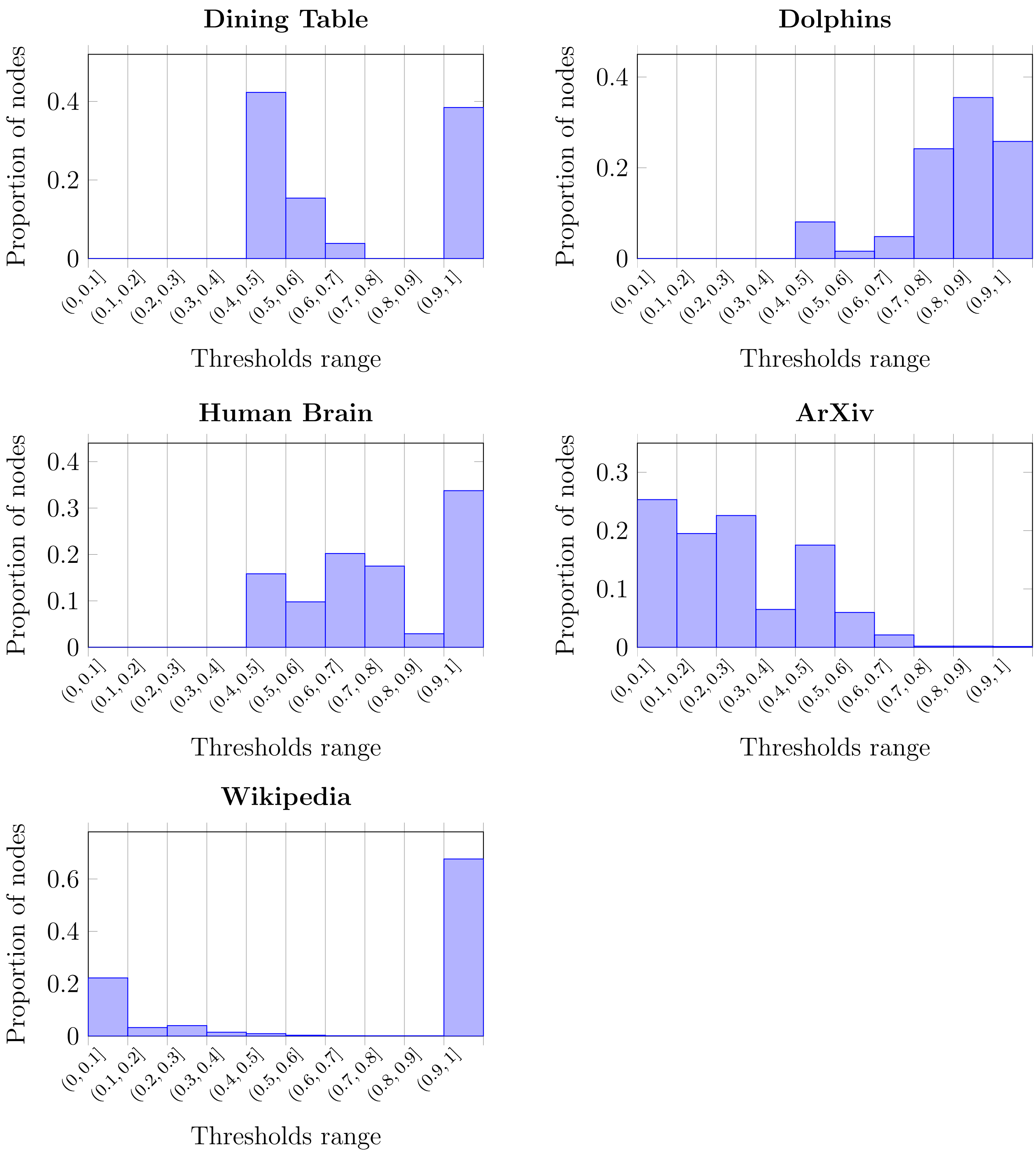


Figure 23: Threshold value ranges for the first experiment outcome starting with thresholds 0.95 and initial random configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the small datasets: Dining Table, Dolphins, Human Brain, ArXiv, and Wikipedia. The results are an average of 5 repetitions of the experiment.

Table 31: Full Table with the metrics of the first experiment with initial threshold 0.95 and initial random configuration for the initial and influence sets. This table is the outcome of the $\Gamma_I$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0538 | 0.6769 | 0.8000 | 2.6000 | 0.4000 | 0.7000 | 1.0000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0422 | 0.1331 | 0.0077 | 0.0103 | 0.0130 | 0.0058 | 0.0494 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 5.6000 | 5.6000 | 5.6000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0185 | 0.0185 | 0.0185 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0054 | 0.9879 | 0.9958 | 2.0000 | 1.4000 | 3.2333 | 5.8000 | 1.0000 | 4.2033 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0079 | 0.3551 | 0.0008 | 0.0016 | 0.0023 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0294 | 0.8985 | 0.9803 | 2.0000 | 2.0000 | 3.7559 | 14.4000 | 2.0000 | 11.9364 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0000 | 0.0002 | 0.0009 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0013 | 0.3283 | 0.3280 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 25.0014 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0004 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 7.8000 | 7.8000 | 7.8000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0119 | 0.9585 | 0.9919 | 2.0000 | 2.0000 | 3.8488 | 38.0000 | 2.0000 | 20.7732 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0022 | 0.9886 | 0.9915 | 2.4000 | 0.0000 | 0.1431 | 1.0000 | 0.0000 | 2.3608 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0085 | 0.6535 | 0.6783 | 3.0000 | 0.0000 | 0.6837 | 17.2000 | 0.0000 | 9.5036 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0231 | 0.0790 | 0.1455 | 3.0000 | 0.0000 | 0.0721 | 14.6000 | 0.0000 | 1.4963 | 86.8000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0033 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

Table 32: Full Table with the metrics of the first experiment with initial threshold 0.95 and initial random configuration for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| **Network** | $\pi_I$ | $\pi_F$ | $\pi_T$ | **Rounds** | **Out-Degree Initial** | | | **In-Degree Influence** | | | **Betweenness Initial** | | | **Betweenness Influence** | | | **Pagerank Initial** | | | **Pagerank Influence** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0538 | 0.6769 | 0.8000 | 1.4000 | 0.4000 | 0.7000 | 1.0000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0422 | 0.1331 | 0.0077 | 0.0103 | 0.0130 | 0.0058 | 0.0494 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 5.6000 | 5.6000 | 5.6000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0185 | 0.0185 | 0.0185 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0054 | 0.9879 | 0.9958 | 1.8000 | 1.4000 | 3.2333 | 5.8000 | 1.0000 | 4.2033 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0079 | 0.3551 | 0.0008 | 0.0016 | 0.0023 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0294 | 0.8985 | 0.9803 | 1.0000 | 2.0000 | 3.7559 | 14.4000 | 2.0000 | 11.9364 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0000 | 0.0002 | 0.0009 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0013 | 0.3283 | 0.3280 | 1.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 25.0014 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0004 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 7.8000 | 7.8000 | 7.8000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0119 | 0.9585 | 0.9919 | 1.0000 | 2.0000 | 3.8488 | 38.0000 | 2.0000 | 20.7732 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0022 | 0.9886 | 0.9915 | 1.0000 | 0.0000 | 0.1431 | 1.0000 | 0.0000 | 2.3608 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0085 | 0.6535 | 0.6783 | 1.0000 | 0.0000 | 0.6837 | 17.2000 | 0.0000 | 9.5036 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0231 | 0.0790 | 0.1455 | 1.0000 | 0.0000 | 0.0721 | 14.6000 | 0.0000 | 1.4963 | 86.8000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0033 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

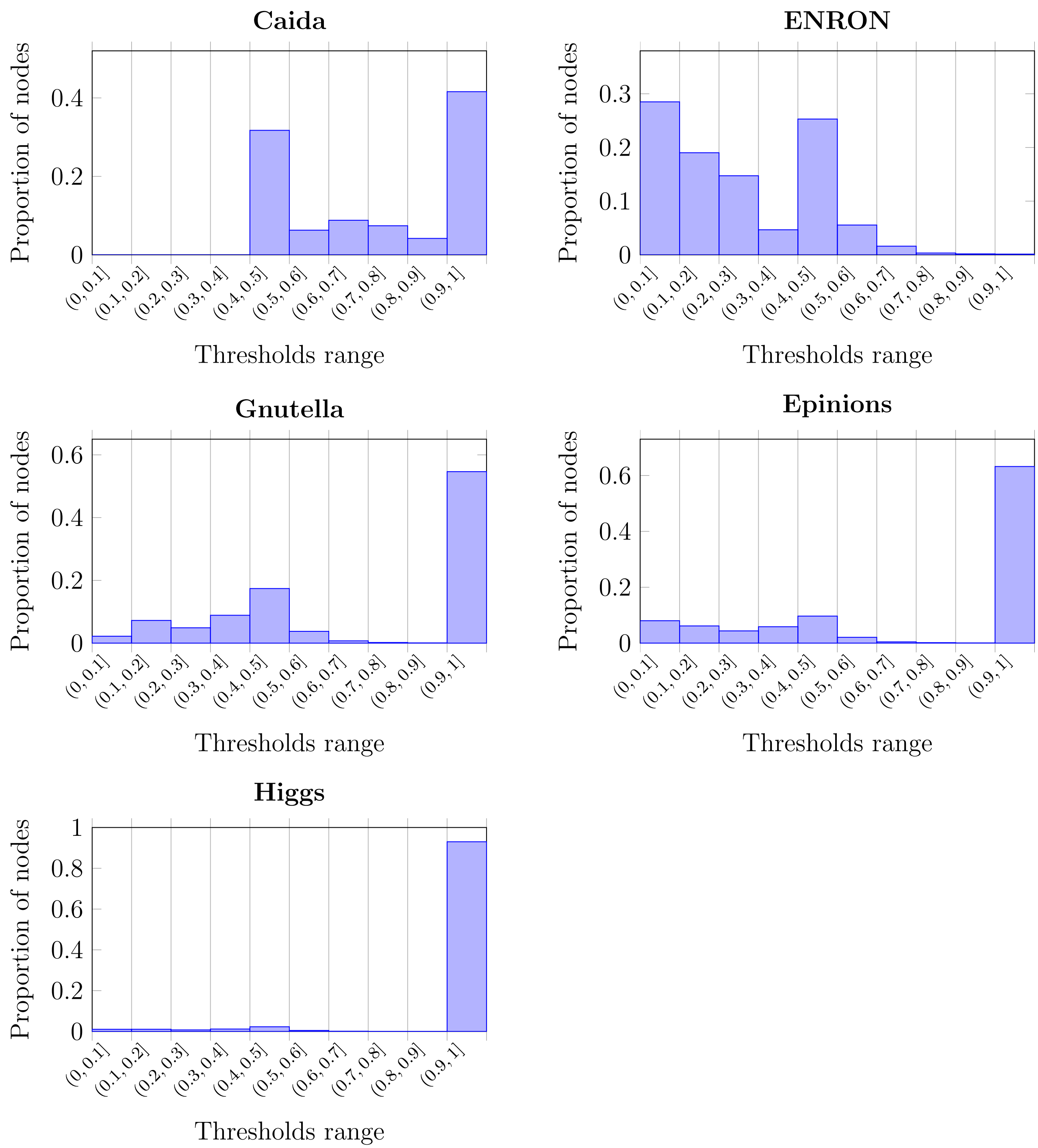


Figure 24: Threshold value ranges for the first experiment outcome starting with thresholds 0.95 and initial random configuration. Each bar is the number of nodes that have a threshold value lying in the corresponding range. These plots are for the big datasets: Caida, ENRON, Gnutella, Epinions and Higgs. The results are an average of 5 repetitions of the experiment.

Table 33: Full Table with the metrics of the second experiment with cooperative agents for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.1923 | 0.6692 | 0.6800 | 1.2000 | 0.4000 | 2.1200 | 4.2000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0406 | 0.1331 | 0.0077 | 0.0486 | 0.1313 | 0.0058 | 0.0507 | 0.2329 |
| **Dolphins** | 0.1935 | 1.0000 | 1.0000 | 2.0000 | 1.8000 | 6.0333 | 11.0000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0056 | 0.0165 | 0.0285 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.2000 | 0.9833 | 0.9750 | 1.0000 | 1.0000 | 3.8021 | 48.6000 | 1.0000 | 4.2136 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0079 | 0.3551 | 0.0004 | 0.0020 | 0.0281 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.1999 | 0.9101 | 0.8845 | 1.0000 | 2.0000 | 10.9992 | 140.8000 | 2.0000 | 11.8289 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0000 | 0.0002 | 0.0013 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.2000 | 0.4648 | 0.3403 | 1.0000 | 0.0000 | 14.1966 | 289.8000 | 0.0000 | 20.3516 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0030 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.2000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 4.4241 | 2207.4000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0119 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.2000 | 0.9592 | 0.9476 | 1.0000 | 2.0000 | 20.3321 | 2337.6000 | 2.0000 | 20.7587 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0046 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.2000 | 0.9828 | 0.9783 | 1.0000 | 0.0000 | 2.3599 | 42.6000 | 0.0000 | 2.3563 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0001 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.2000 | 0.7170 | 0.6479 | 1.0000 | 0.0000 | 6.8474 | 1256.2000 | 0.0000 | 8.7791 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | 0.0000 | 0.0023 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.2000 | 0.2552 | 0.0684 | 1.0000 | 0.0000 | 1.2798 | 5192.2000 | 0.0000 | 1.2411 | 69.4000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0365 | $< \epsilon$ | $< \epsilon$ | 0.0172 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

Table 34: Full Table with the metrics of the second experiment with cooperative agents for the initial and influence sets. This table is the outcome of the $\Gamma_I$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0462 | 0.6308 | 0.8800 | 2.4000 | 1.0000 | 1.1000 | 1.2000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0407 | 0.1331 | 0.0102 | 0.0136 | 0.0169 | 0.0058 | 0.0512 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 4.2000 | 4.2000 | 4.2000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0136 | 0.0136 | 0.0136 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0067 | 0.9912 | 1.0000 | 2.0000 | 1.0000 | 1.2500 | 1.6000 | 1.0000 | 4.1934 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0078 | 0.3551 | 0.0009 | 0.0020 | 0.0033 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0307 | 0.9018 | 0.9851 | 2.0000 | 2.0000 | 3.6437 | 12.8000 | 2.0000 | 11.8982 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0001 | 0.0002 | 0.0009 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0018 | 0.3291 | 0.3460 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 24.8671 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0003 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0120 | 0.9589 | 0.9904 | 2.0000 | 2.0000 | 3.6343 | 16.4000 | 2.0000 | 20.7659 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0022 | 0.9889 | 0.9919 | 3.0000 | 0.0000 | 0.0849 | 1.0000 | 0.0000 | 2.3612 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0001 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0083 | 0.6534 | 0.6780 | 3.0000 | 0.0000 | 0.5779 | 7.0000 | 0.0000 | 9.5042 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0228 | 0.0786 | 0.1434 | 3.0000 | 0.0000 | 0.0742 | 10.2000 | 0.0000 | 1.5071 | 80.4000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0036 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

Table 35: Full Table with the metrics of the second experiment with malicious agents for the initial and influence sets. This table is the outcome of the $\Gamma_f$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.1923 | 0.6692 | 0.6800 | 2.0000 | 0.4000 | 2.1200 | 4.2000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0406 | 0.1331 | 0.0077 | 0.0486 | 0.1313 | 0.0058 | 0.0507 | 0.2329 |
| **Dolphins** | 0.1935 | 1.0000 | 1.0000 | 2.0000 | 1.8000 | 6.0333 | 11.0000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0056 | 0.0165 | 0.0285 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.2000 | 0.9833 | 0.9750 | 1.0000 | 1.0000 | 3.8021 | 48.6000 | 1.0000 | 4.2136 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0079 | 0.3551 | 0.0004 | 0.0020 | 0.0281 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.1999 | 0.9101 | 0.8845 | 1.0000 | 2.0000 | 10.9992 | 140.8000 | 2.0000 | 11.8289 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0000 | 0.0002 | 0.0013 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.2000 | 0.4648 | 0.3403 | 2.0000 | 0.0000 | 14.1966 | 289.8000 | 0.0000 | 20.3516 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0030 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.2000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 4.4241 | 2207.4000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0119 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.2000 | 0.9592 | 0.9476 | 1.0000 | 2.0000 | 20.3321 | 2337.6000 | 2.0000 | 20.7587 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0046 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.2000 | 0.9828 | 0.9783 | 2.0000 | 0.0000 | 2.3599 | 42.6000 | 0.0000 | 2.3563 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0001 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.2000 | 0.7176 | 0.6489 | 2.0000 | 0.0000 | 6.3974 | 1098.2000 | 0.0000 | 8.7724 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | 0.0000 | 0.0022 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.2000 | 0.2556 | 0.0702 | 2.0000 | 0.0000 | 1.3305 | 5289.8000 | 0.0000 | 1.2405 | 69.4000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | $< \epsilon$ | 0.0353 | $< \epsilon$ | $< \epsilon$ | 0.0095 | $< \epsilon$ | $< \epsilon$ | 0.0289 |

Table 36: Full Table with the metrics of the second experiment with malicious agents for the initial and influence sets. This table is the outcome of the $\Gamma_I$ model. The results are an average for 5 repetitions of the experiment.

| Network | $\pi_I$ | $\pi_F$ | $\pi_T$ | Rounds | Out-Degree Initial | | | In-Degree Influence | | | Betweenness Initial | | | Betweenness Influence | | | Pagerank Initial | | | Pagerank Influence | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| **Dining Table** | 0.0462 | 0.6308 | 0.8800 | 2.4000 | 1.0000 | 1.1000 | 1.2000 | 2.0000 | 2.0000 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0407 | 0.1331 | 0.0102 | 0.0136 | 0.0169 | 0.0058 | 0.0512 | 0.2329 |
| **Dolphins** | 0.0161 | 1.0000 | 1.0000 | 2.0000 | 4.2000 | 4.2000 | 4.2000 | 1.0000 | 5.1290 | 12.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0393 | 0.2482 | 0.0136 | 0.0136 | 0.0136 | 0.0048 | 0.0161 | 0.0321 |
| **Human Brain** | 0.0067 | 0.9912 | 1.0000 | 2.0000 | 1.0000 | 1.2500 | 1.6000 | 1.0000 | 4.1934 | 96.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0078 | 0.3551 | 0.0009 | 0.0020 | 0.0033 | 0.0004 | 0.0021 | 0.0546 |
| **ArXiv** | 0.0307 | 0.9018 | 0.9851 | 2.0000 | 2.0000 | 3.6437 | 12.8000 | 2.0000 | 11.8982 | 162.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0012 | 0.0741 | 0.0001 | 0.0002 | 0.0009 | 0.0000 | 0.0002 | 0.0014 |
| **Wikipedia** | 0.0018 | 0.3291 | 0.3460 | 2.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 24.8671 | 893.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0177 | 0.0001 | 0.0001 | 0.0003 | 0.0001 | 0.0001 | 0.0037 |
| **Caida** | 0.0000 | 1.0000 | 1.0000 | 2.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 4.0326 | 2628.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1538 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0215 |
| **ENRON** | 0.0120 | 0.9589 | 0.9904 | 2.0000 | 2.0000 | 3.6343 | 16.4000 | 2.0000 | 20.7659 | 2766.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.1297 | $< \epsilon$ | 0.0000 | 0.0001 | $< \epsilon$ | 0.0000 | 0.0114 |
| **Gnutella** | 0.0022 | 0.9889 | 0.9919 | 3.0000 | 0.0000 | 0.0849 | 1.0000 | 0.0000 | 2.3612 | 78.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0028 | 0.0000 | 0.0000 | 0.0001 | 0.0000 | 0.0000 | 0.0001 |
| **Epinions** | 0.0087 | 0.6547 | 0.6796 | 3.0000 | 0.0000 | 0.5763 | 7.0000 | 0.0000 | 9.4882 | 1801.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0655 | $< \epsilon$ | $< \epsilon$ | 0.0001 | $< \epsilon$ | 0.0000 | 0.0047 |
| **Higgs** | 0.0228 | 0.0787 | 0.1442 | 3.0000 | 0.0000 | 0.0716 | 8.2000 | 0.0000 | 1.5088 | 73.2000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0328 | $< \epsilon$ | $< \epsilon$ | 0.0036 | $< \epsilon$ | $< \epsilon$ | 0.0289 |