Computer Modeling in Engineering & Sciences

Tech Science Press

Doi:10.32604/cmes.2026.073647

ARTICLE

KMFC-GWO: A Hybrid Fuzzy-Metaheuristic Algorithm for Privacy-Preservation in Graph-Based Social Networks

Saeideh Memarian[1], Andreea M. Oprescu[2,3], Natalia Moreno-Naranjo[2], Gloria Miró-Amarante[2] and M. Carmen Romero-Ternero[2,3,*]

[1]Doctoral Program in Computer Science Engineering, Universidad de Sevilla, Sevilla, Spain
[2]Departamento Tecnología Electrónica, Universidad de Sevilla, Sevilla, Spain
[3]Instituto de Ingeniería Informática, Universidad de Sevilla, Sevilla, Spain

*Corresponding Author: M. Carmen Romero-Ternero. Email: mcromerot@us.es

Received: 22 September 2025; Accepted: 12 January 2026; Published: 27 April 2026

ABSTRACT: In recent years, the proliferation of social networks has been remarkable, providing a rich source for data mining endeavors. However, a significant challenge lies in safeguarding the privacy of individuals while sharing these databases publicly. Current approaches, such as K-anonymity, L-diversity, and T-closeness, are commonly employed for data anonymization in social networks. However, these techniques entail considerable information loss due to random alterations in the graph-based datasets. To address these limitations, this paper introduces a new anonymization technique called KMFC-GWO, which combines K-Member Fuzzy Clustering with Grey Wolf Optimizer. This integrated method is designed to strengthen the anonymized graph against a range of threats, including identity, attribute, link disclosure, and similarity attacks, while significantly reducing information loss. Within the KMFC-GWO framework, K-member fuzzy c-means clustering is utilized to create well-balanced clusters, each meeting the K-anonymity requirement. Subsequently, the Grey Wolf Optimizer is applied to optimize cluster formation and effectively anonymize the social network graph. The objective function is carefully crafted to minimize both clustering error and information loss, while ensuring adherence to predefined anonymity criteria. Experimentation on three major graph-based social networks extracted from Facebook, Twitter, and YouTube validates the effectiveness of the KMFC-GWO approach. Results demonstrate its ability to significantly reduce information loss in published graph data, while concurrently satisfying requirements for K-anonymity, L-diversity, and T-closeness.

KEYWORDS: Gaph-based social networks; privacy preserving; K-anonymity; L-diversity; T-closeness; fuzzy clustering; grey wolf optimizar (GWO)

1 Introduction

With advances in technology, social networks have emerged as pervasive platforms for worldwide social interaction, information sharing, and self-expression [1]. While these networks harbor vast amounts of user data that can enhance service quality, they also pose risks to individual privacy due to the presence of sensitive information. Users in social media platforms do not only post textual content but also frequently share personal photographs, videos, and interactions with friends and family members. These data, while voluntarily disclosed, can still pose severe privacy threats. For example, attackers may infer a user's identity by cross-linking their facial images with public databases, or they can reconstruct sensitive information such as a user's political views or medical conditions by analyzing patterns of likes, comments, and social connections. Even when users willingly share content, they usually do not anticipate large-scale data mining

or de-anonymization attacks that exploit structural graph information combined with multimedia. This highlights the importance of designing anonymization methods that preserve privacy beyond simple text-based disclosures. Consequently, both users and data owners seek to safeguard data privacy while leveraging its insights for various purposes [2].

Sharing social network data with data miners necessitates a careful balance between privacy protection and knowledge retention. Anonymizing techniques, such as K-anonymity (KA) and its extensions like L-diversity (LD) and T-closeness (TC), are commonly employed to mitigate privacy risks while preserving data utility [3–5]. KA aims to group users into clusters with at least K members to prevent identity disclosure, though it doesn't safeguard against attribute or link disclosure. LD addresses attribute disclosure by ensuring each cluster contains diverse attribute values, while TC focuses on maintaining the global attribute distribution within clusters to mitigate similarity attacks.

Privacy threats in social network data publishing encompass different attacks such as identity, attribute, and link disclosure [5]. Identity disclosure exposes a user's identity, while attribute disclosure reveals sensitive user attributes and link disclosure unveils sensitive relationships between users. LD and TC complement KA by addressing attribute and similarity attacks, respectively, enhancing overall privacy protection in anonymized datasets [6].

Despite the significant advancements in privacy-preserving techniques for social network data publishing, several research gaps remain unaddressed. Existing methods like KA, LD, and TC primarily focus on mitigating specific privacy threats, often at the cost of high information loss or computational complexity. While KA is effective in preventing identity disclosure, it falls short in safeguarding against attribute and link disclosures. LD and TC extend KA's capabilities but often struggle with scalability and preserving data utility when applied to large-scale or complex social networks. Moreover, most existing approaches treat privacy threats in isolation, lacking a unified framework that comprehensively addresses identity, attribute, and link disclosures simultaneously. This creates a pressing need for innovative solutions that balance robust privacy protection with minimal information loss, especially in the context of graph-based social networks.

This research concentrates on protecting social network data publication from a range of threats, including revealing identities, disclosing attributes/links, and potential similarity attacks. To address these challenges, we introduce a hybrid anonymization approach called KMFC-GWO, which combines K-member Fuzzy Clustering (KMFC) with Grey Wolf Optimizer (GWO). The main objective of KMFC-GWO is to fortify the privacy of graph-based social networks while reducing information loss. Our approach employs a modified variant of fuzzy c-means (FCM), referred to as KMFC, to create well-balanced clusters containing a minimum of K members in each cluster, thereby fulfilling the KA criterion. Furthermore, an optimization problem is formulated and solved using GWO, to satisfy the LD and TC conditions. The main contributions of this paper can be summarized as follows:

- Hybrid Anonymization Framework: We propose a hybrid approach (KMFC-GWO) that combines K-member Fuzzy Clustering (KMFC) with the Grey Wolf Optimizer (GWO) to address multiple privacy threats in graph-based social networks.
- Comprehensive Privacy Protection: Our framework ensures robust privacy by simultaneously addressing identity, attribute, and link disclosure risks, going beyond the limitations of traditional anonymization techniques.
- Enhanced Clustering with KMFC: We introduce a modified variant of FCM clustering, i.e., KMFC, to form well-balanced clusters that meet the KA criterion, ensuring a minimum of K members per cluster.
- Optimized Privacy and Utility Balance: Through the integration of the GWO algorithm, we optimize cluster properties to satisfy LD and TC conditions, minimizing information loss while preserving data utility.

- Scalability and Effectiveness: Our approach demonstrates scalability and effectiveness on graph-based social networks, showcasing its potential for practical applications in large-scale data publishing scenarios.

In the rest of the paper, Section 2 reviews the literature on anonymization methods. Section 3 introduces the KMFC-GWO methodology. Section 4 outlines the simulation of the KMFC-GWO algorithm on three graph-based social networks. Finally, Section 5 concludes the paper with future works.

2 Literature Review

2.1 Comparison with LocJury

LocJury is a privacy-preserving framework proposed for protecting location data in Internet of Connected Vehicles (IoCV) environments. It employs identity-based networking and clustering mechanisms to reduce the risk of location disclosure while maintaining service utility [7]. Although LocJury is designed for vehicular location privacy rather than social network anonymization, it is considered here for comparison due to its use of clustering-based privacy enforcement and constraint-driven data protection, which conceptually relate to the objectives of the proposed KMFC-GWO framework. LocJury, an Identity-Based Networking (IBN) location-privacy preservation framework for Internet of Connected Vehicles (IoCV), focuses primarily on protecting spatiotemporal trajectories, trust-based decisions, and contextual location disclosure. Unlike LocJury, which operates in a vehicular mobility environment, KMFC-GWO focuses on anonymization of graph-structured social networks while jointly satisfying K-anonymity, L-diversity, and T-closeness. KMFC-GWO performs optimization on both node attributes and structural edges, providing a multi-constraint anonymization method that addresses risks not covered in LocJury.

2.2 Definitions

The characteristics of users within social networks commonly involve both graph properties, which pertain to the connections between users, and matrix properties, relating to personal attributes. Graph-based binary features indicate whether connections exist between users, while matrix-based features are typically classified into three categories [6]:

- Primary Identifying Attributes: This category encompasses direct identifiers such as full name, surname, and national identification numbers, which can directly expose a user's identity. Consequently, these attributes must be eliminated prior to any additional processing.
- Sensitive Attributes: These attributes hold considerable significance and necessitate enhanced protection, examples include postal codes, specific medical conditions, or salary details. Even the disclosure of a single value or a combination of these attributes could potentially jeopardize user privacy or divulge sensitive information.
- Standard Attributes: Attributes falling into this category exhibit lower sensitivity compared to the aforementioned types and therefore require less rigorous privacy safeguards. Examples include age, gender, and educational background. Prior to applying any anonymization algorithm, it is generally assumed that explicit identifiers have been removed from the dataset. Consequently, anonymization is performed on quasi-identifiers, sensitive attributes, and graph-based features. The concept of k-anonymity and its extensions form the foundation of many privacy-preserving models in data publishing [4]
- KA (K-anonymity): A modified table adheres to the KA condition if no user can be uniquely identified from fewer than $K - 1$ other users based on the features. In cluster-based methods, this equates to grouping all users into distinct clusters, ensuring that each cluster comprises a minimum of K samples.

- LD (L-diversity): A cluster or group of users demonstrates LD when there are at least L unique values for every sensitive attribute within that cluster. Thus, a revised table achieves LD when each cluster within it satisfies this condition.
- TC (T-closeness): A cluster is considered to adhere to TC conditions when the variance between the distribution of each sensitive attribute within the cluster and the distribution of that attribute across the entire dataset remains below a threshold T. If all clusters meet the TC requirements, the revised table is regarded as fulfilling TC criteria.

2.3 Existing Methods

The majority of anonymization techniques rely on clustering-based KA, wherein users are clustered into separate groups, each containing at least K members [8]. Graph-based anonymization methods typically involve initial modifications to the network data, followed by propagation of the altered graph. Edge randomization techniques introduce random changes to the graph-based network structure by randomly adding/deleting edges or swapping pairs of edges. This randomization aids in safeguarding users against re-identification through probabilistic means. Graph structure modifications can occur random perturbation or with the goal of satisfying some predetermined constraints such as KA, LD, and TC [9].

Random unconstrained perturbation involves fundamental techniques that alter graph structures by randomly adding or removing edges. Casas-Roma [10] introduced a spectral approach to safeguard critical network edges while optimizing utility within specified privacy constraints. Although spectral methods enhance utility while minimizing information loss, they entail higher computational demands. Nguyen et al. [11] employed quadratic programming to create uncertain graphs, focusing on edge manipulation for anonymization. However, these techniques, despite their simplicity, overlook the diversity and distribution of sensitive attributes, making them vulnerable to attribute and similarity attacks. Moreover, they may not ensure the K-anonymity condition, as edge modification does not typically consider it.

Kumar and Kumar [12] proposed a K-degree technique based on the upper approximation method of rough sets, utilizing split-join operations to evaluate privacy preservation. They employed various metrics including variation of information, mutual information, and adjusted Rand index to assess utility. Kiabod et al. [13] introduced a time-efficient K-degree method for anonymizing network graphs without the need for rescanning the structure for different anonymity levels. This method constructs a connection tree by computing the degree sequence of the anonymized graph, ensuring it exceeds a specified threshold. While KA and the K-degree approach protect against identity attacks, they do not address threats such as attribute/link and similarity attacks.

Yazdanjue et al. [14] proposed a KA technique for social networks through integrating clustering with evolutionary algorithms. This technique initially clusters nodes in the social network based on similarities. Fitness evaluation follows, assessing cluster fitness considering intra-cluster similarity and inter-cluster dissimilarity. Employing an evolutionary process, clusters with low fitness undergo modification through genetic operators like mutation and crossover. Throughout this process, the algorithm ensures KA, maintaining that each cluster contains at least K indistinguishable individuals. Termination occurs upon finding a satisfactory k-anonymous solution or reaching a predefined stopping criterion.

Langari et al. [6] introduced a K-member Fuzzy Clustering and Firefly Algorithm (KFCFA) approach for privacy preservation in social networks by combining fuzzy clustering with the firefly algorithm. The method begins with fuzzy clustering to group nodes based on similarity, followed by the application of the firefly algorithm to optimize privacy parameters. Through an iterative optimization procedure, this algorithm aims to minimize information disclosure while maintaining network utility. Termination occurs upon reaching a satisfactory privacy-preserving solution.

Rajabzadeh et al. [15] proposed a graph-based modification technique for achieving KA by employing a genetic algorithm. The approach involves iteratively modifying the network's structure to ensure that each generated subgraph satisfies the KA property. By employing a genetic algorithm, this method optimizes the required anonymity conditions and network utility. The algorithm terminates when a satisfactory level of KA is generated.

Clustering-based anonymization strategies commonly rely on K-anonymity as their fundamental privacy requirement, upon which additional diversity and distributional constraints can be enforced to mitigate attribute disclosure risks [4].

2.4 Research Gaps and Our Contributions against the Reviewed Techniques

The existing methods for anonymizing social network data, while effective in addressing specific privacy concerns, exhibit notable limitations. Many clustering-based KA techniques, such as those employing random edge perturbations or K-degree strategies, fail to adequately address threats like attribute and similarity attacks. Methods like spectral approaches or quadratic programming-based uncertain graphs, though utility-focused, often neglect the diversity and distribution of sensitive attributes. Moreover, the high computational cost of these methods makes them impractical for large-scale networks. Evolutionary and metaheuristic-based approaches, such as genetic algorithms or firefly algorithms, offer improved optimization but tend to prioritize specific conditions, such as KA, at the expense of balanced utility and privacy across multiple dimensions like LD and TC. These limitations highlight a critical gap in developing comprehensive methods that achieve robust privacy protection while maintaining low information loss and ensuring balanced clustering.

Our proposed KMFC-GWO anonymization technique addresses the mentioned gaps by combining KMFC with GWO to form a hybrid method capable of meeting KA, LD, and TC conditions simultaneously. Unlike existing methods, KMFC-GWO incorporates a modified fuzzy c-means clustering algorithm to ensure well-balanced clusters, where each cluster contains at least K members, thereby guaranteeing the KA condition. Additionally, our method formulates privacy protection as an optimization problem, leveraging GWO to optimize LD and TC constraints. This dual-layer approach effectively mitigates identity, attribute, and similarity attacks while preserving the structural and attribute-based integrity of the data. Furthermore, KMFC-GWO demonstrates superior computational efficiency and scalability compared to existing metaheuristic approaches, making it a practical solution for large-scale social network datasets. By achieving a balanced trade-off between privacy and utility, KMFC-GWO represents a significant advancement over existing anonymization techniques.

3 Proposed KMFC-GWO Algorithm

In this section, we present a hybrid fuzzy-metaheuristic graph-based privacy-preserving approach for social networks, termed KMFC-GWO, which combines KMFC with GWO, offering a highly effective solution for both balanced clustering and anonymization in social networks. Our method addresses the optimization problem of privacy preservation within graph-based social networks by integrating KMFC, considering KA, LD, and TC criteria. Subsequently, GWO is leveraged to tackle this optimization challenge, aiming to maximize anonymity while minimizing information loss in the published graph. Specifically, the objective function of GWO is formulated to minimize information loss and uphold anonymity conditions (KA, LD, and TC). The proposed KMFC-GWO approach effectively safeguards the anonymized social network graph against identity, attribute disclosure, and similarity attacks by enforcing these three constraints.

Consider a social network represented as a graph, where the edges (referred to as G) have dimensions $N \times N$, and the attribute data matrix (referred to as A) has dimensions $N \times M$. Here, N denotes the number

of users, and M denotes the number of personal characteristics associated with each user. Each user, denoted as i (where i ranges from 1 to N), possesses a vector of edges indicating connections with other users and a vector of features represented by M. The graph matrix, G, is binary: Gij = 1 signifies the existence of an edge between users i and j, while Gij = 0 denotes the absence of such an edge. Furthermore, Aim represents the m-th personal characteristic of user i. The aim of the anonymization procedure is to alter the original data to generate a modified edge graph (denoted as Gnew) and a modified data matrix (denoted as Anew). Before introducing the proposed KMFC-GWO algorithm, Fig. 1 provides a simple example of a graph-based social network, where users are modeled as nodes and their interactions as edges.

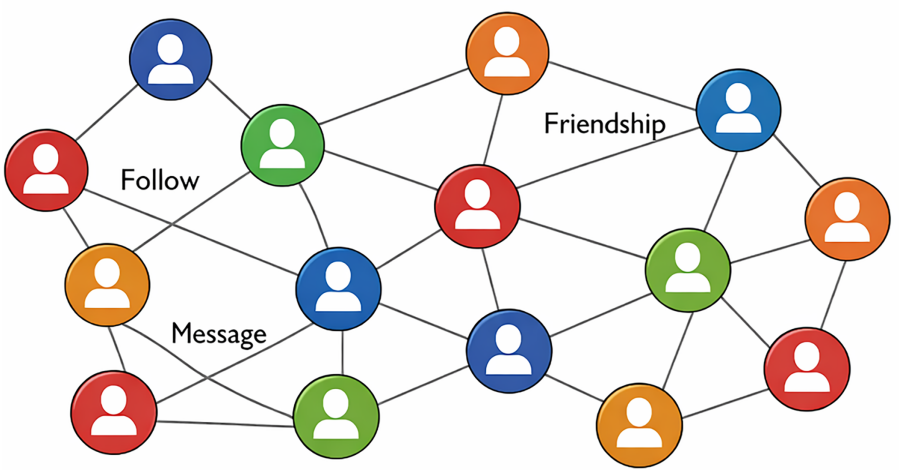

Figure 1: An illustrative example of a graph-based social network, used to demonstrate how users and their interactions are modeled as a graph in real-world social platforms

Fig. 2 is not intended as a purely hypothetical illustration, but rather as an intuitive abstraction of the types of real graph-based social networks analyzed in this study. The social platforms used in our experiments—Twitter, Facebook, and YouTube—can each be rigorously modeled as graphs, where users constitute the nodes and various forms of interactions define the edges. In Twitter, user relationships form a directed graph through the follower–following mechanism, while mentions, replies, and retweets generate additional interaction edges. Facebook predominantly exhibits an undirected friendship graph, yet its structure is further enriched through comments, likes, and group interactions. Although primarily a content-sharing platform, YouTube also functions as a social network: users can subscribe to other users, interact through comments, share content within communities, and therefore form user–user interaction graphs as well as user–content relational structures. Consequently, the anonymization problem addressed in this work inherently belongs to the domain of graph-based social network modeling. The heterogeneous connectivity, neighborhood structures, and community-like patterns reflected in Fig. 2 are directly representative of the structural properties present in these real-world platforms. This establishes a clear rationale for adopting a graph-based anonymization framework such as KMFC-GWO and justifies the applicability of our proposed method to Twitter, Facebook, and YouTube datasets.

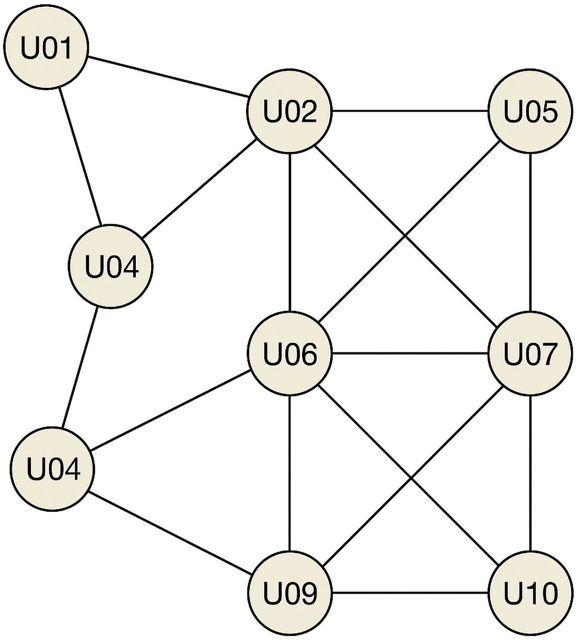

Figure 2: Example of a graph-based social network with users U01–U10 and their interaction links

In the KMFC-GWO approach, the K-anonymity requirement is met through KMFC forming C clusters, each comprising a minimum of K users. Then, the GWO algorithm is utilized to address the anonymity criteria of LD and TC, while simultaneously reducing information loss in the published social network graph.

3.1 Satisfying KA Using KMFC

As mentioned above, to satisfy the KA condition, we apply a KMFC technique on the FCM. Our KMFC method utilizes a customized version of the FCM algorithm to construct balanced clusters with at least K members, satisfying the KA condition. To achieve this purpose, first, we generate an initial clustering using FCM, and then, the generated clusters are balanced to satisfy the KA condition.

The FCM algorithm was initially introduced by Bezdek in 1981 [16]. Unlike traditional clustering techniques such as c-means and k-means, FCM assigns a degree of membership for each data point to every cluster. The objective of FCM is to minimize the total distance between the data points and the cluster centroids. Its primary aim is to partition N data points into C distinct clusters. Each data point can be characterized by two feature vectors: binary edges and personal attributes. More specifically, $X = \{x_1, x_2, \ldots x_i, \ldots x_N\}^T$ is the input matrix of dimension $N \times F$, where $F = N + M$ is the number of clustering attributes, N is the number of all users, and M is the number of personal features. Also, $s = \{s_1, s_2, \ldots, s_k, \ldots, s_C\}^T$ is the centroid of clusters ($C \times F$). The objective of FCM can be formulated as follows [17]:

$$OF_{FCM} = \sum_{i=1}^{N} \sum_{k=1}^{C} U_{ik}^{q} \times d_{x_i, s_k} \tag{1}$$

where $x_i = [x_{i1}, x_{i2}, \ldots, x_{if}, \ldots, x_{iF}]$ represents sample i, $s_k = [s_{k1}, s_{k2}, \ldots, s_{kf}, \ldots, s_{kF}]$ is the center of cluster k, q ($q > 1$) is the exponential weight of FCM that adjusts the fuzziness degree, and $U_{ik} \in [0,1]$ is the membership of node i to cluster k which should fulfil the condition of Eq. (2). Also, d_{x_i, s_k} measures the Euclidian distance between x_i and s_k, as formulated in Eq. (3). To optimize the objective function of the

FCM algorithm described in Eq. (1), iterative partial derivations of OF_{FCM} with respect to U_{ik} and s_k need to be computed as per Eqs. (4) and (5).

$$\sum_{k=1}^{C} U_{ik} = 1 \quad \forall i = 1, 2, \ldots, N \tag{2}$$

$$d_{x_i, s_k} = \sqrt{(x_{i1} - s_{k1})^2 + (x_{i2} - s_{k2})^2 + \ldots + (x_{iF} - s_{kF})^2} \tag{3}$$

$$U_{ik} = \sum_{k'=1}^{C} \left(\frac{d(x_i, s_k)}{d(x_i, s_{k'})} \right)^{\frac{2}{1-q}} \tag{4}$$

$$s_k = \frac{\sum_{i=1}^{N} \left(U_{ik}^q \times x_i \right)}{\sum_{i=1}^{N} U_{ik}^q} \tag{5}$$

The termination criterion for the FCM algorithm is determined by the condition $U_{ik}(t) - U_{ik}(t-1) < \varepsilon$, where t represents the current iteration count, and ε is a small positive value (we set $\varepsilon = 0.001$). Upon termination of the FCM algorithm, each data point (user) is allocated to the cluster with the highest fuzzy membership. Generally, the traditional KA-based clustering methods suffer from two drawbacks:

- Number of clusters: Selecting the best value for the number of clusters C is a challenging issue.
- Unbalanced clusters: In certain clusters, the sample count may be lower than K, while in others, there could be significantly more members than K.

Our proposed revision phase of the KMFC algorithm is provided in Algorithm 1. To handle the first problem, we consider the number of clusters in a such a way that minimizes the CAVG criterion [6], as formulated in Eq. (6). The CAVG metric (where CAVG ≥ 1) quantifies the degree of cluster balance, with values closer to 1 indicating higher balance among clusters. In the case of K-member clustering, CAVG tends to approximate 1, i.e., CAVG ≈ 1.

$$CAVG = \frac{N}{C \times K} \tag{6}$$

Since N and K are fixed parameters, we consider $C \approx N/K$. To have somewhat a free level for clustering, we have set $C = 0.9 \times N/K$ in our simulations to have around 10% more samples on average in each cluster. Furthermore, to handle the second problem to satisfy the KA condition, we present a revision phase on the initial clustering results of the FCM algorithm.

Algorithm 1: Proposed clustering revision in KMFC algorithm

Inputs: Initial clustering solution of FCM, graph matrix G, attribute matrix A, and parameters C and K
Output: Revised Clusters
Begin
1 Divide the initial clusters into three sets (UL, BL, and OL):
- UL (Underload Clusters): with less than K samples
- BL (Balanced Clusters): with the number of samples between K and $1.1 \times K$
- OL (Overload Clusters): with more than $1.1 \times K$ samples

2 Consider all clusters within BL as final revised clusters.
3 Revising clusters within OL and UL:

(Continued)

Algorithm 1 (continued)

***for* 1** $k \in$ UL
  ***for* 2** $k' \in$ OL
  Calculate Euclidian distance between the centroid of clusters k and k': $d_{s_k, s_{k'}}$
  ***end for* 2**
  Sort the samples of cluster k' according to their Euclidian distance to the center of cluster k
  while *number of samples within cluster k be equal to K*
    Transfer the nearest sample of cluster k' to cluster k
  end while
4 Revise the current state of all BL, UL, and OL clusters
***end for* 1**
5 Generate the final revised BL clusters, each with at least K samples
End
6 Generating the clusters of KMFC algorithm

3.2 Impact of KMFC on Cluster Formation

The KMFC component ensures that clusters satisfy the minimum K-member requirement while preserving the inherent structural relationships within the graph. By incorporating both edge-based similarities and attribute-based distances, KMFC groups nodes with similar structural and semantic characteristics. The revision phase redistributes borderline nodes from overloaded clusters to underloaded ones, maintaining cluster balance and minimizing distortion. This ensures that the resulting clusters remain coherent and suitable for subsequent LD and TC optimization via the GWO module.

3.3 Satisfying LD and TC Using GWO

Once users are clustered into C balanced clusters (with at least K users in each cluster) to meet the KA requirement, GWO is used to further anonymize the social network, considering the requirements of LD and TC. This process is conducted simultaneously on both the graph matrix G and the attribute matrix A. At the graph level, anonymization involves randomly adding or removing edges between users, while at the attribute level, it entails randomly altering the values of user features.

Grey Wolf Optimizer (GWO) is a swarm intelligence–based metaheuristic algorithm designed to balance exploration and exploitation in continuous optimization problems. In this study, GWO is adopted as the optimization engine within the proposed framework due to its simple structure and effective search mechanism. The algorithm was originally introduced by Mirjalili et al. in 2014, inspired by the social hierarchy and hunting behavior of grey wolves [18]. Beyond its original formulation, GWO has been successfully applied to various real-world optimization problems, including adaptive fuzzy systems and healthcare-related applications, demonstrating its robustness and practical effectiveness [19]. Since its introduction, GWO has been widely employed in various optimization applications due to its balanced search mechanism and straightforward implementation [18]. As illustrated in Fig. 3, the algorithm commences by randomly initializing a population of grey wolves. In each iteration, the fitness of the current population is evaluated using a customized fitness function tailored to the particular application. Subsequently, the entire population undergoes adjustments through two phases: attacking prey (exploitation) and searching for prey (exploration), as described in the standard GWO procedure [18]. These main steps of GWO including random generation of initial population, objective function evaluation, and population updating, are described in the following:

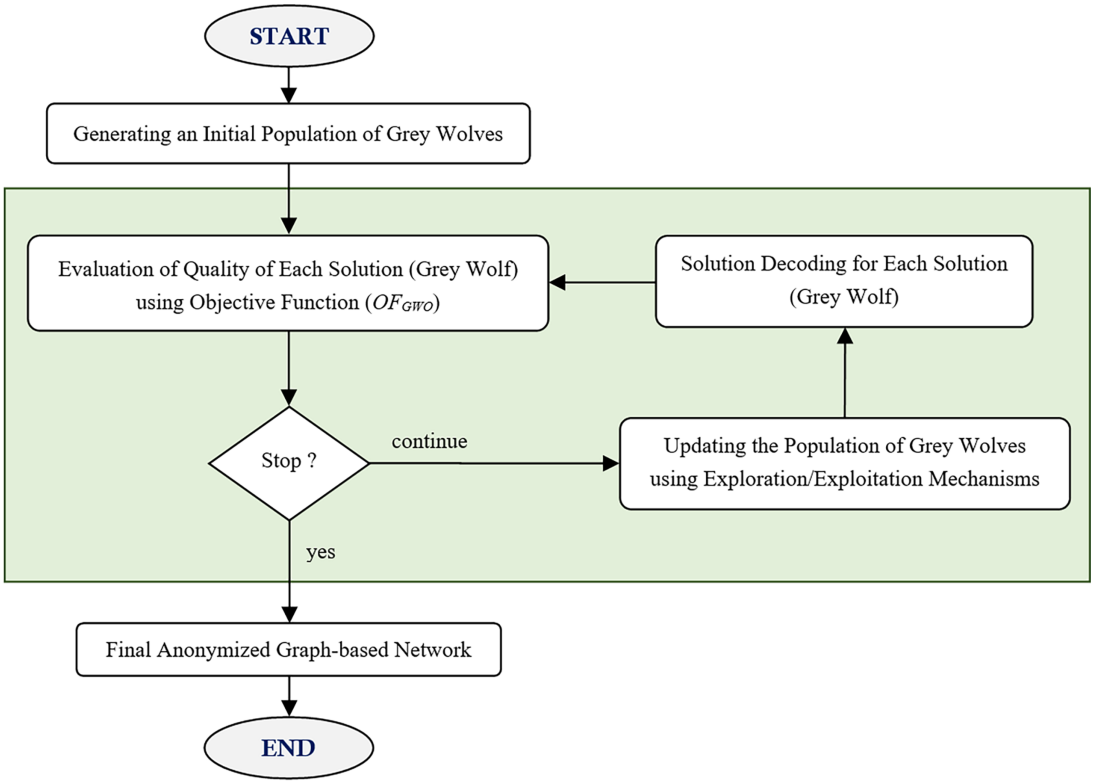

Figure 3: Flowchart of GWO-based optimization procedure for LD and TC anonymization

Generation of initial population: As seen in Fig. 4, a feasible solution (SOL) can be represented as a matrix of N × (N + M) = N × F, where each row i represents the edge and attribute modifications in G and A matrices. In the case of edge modification, SOLi,j = 1 if the edge between nodes i and j is modified (adding a new edge between nodes i and j or removing a previously connection link between nodes i and j). Furthermore, for the attribute modification, SOLi, N + m = 1 if the value of the personal feature m is randomly changed for the data of user i.

	1	2	⋯	N	N+1	N+2	⋯		F=N+M
1	**1**	**0**	…	**0**	**1**	**1**	…		**1**
2	**0**	**1**	…	**1**	**1**	**0**	…		**0**
⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋱		⋮
N	**1**	**0**	…	**0**	**1**	**0**	…		**0**

Figure 4: Representation of a feasible solution in GWO

- Objective function evaluation: There is a trade-off between the information loss (due to the modifications in G and A) and the anonymity level. The more changes in the G and A, the higher anonymity

level, however by accepting more information loss. To minimize the information loss within the published social network data, we present an objective function to minimize the summation of the information loss of the modified G and A, denoted by IL_G and IL_A, while satisfying the constraints of KA, LD and TC. More specifically, the objective function of Grey Wolf Optimizer (OFGWO) is defined as follows:

$$OF_{GWO} = \text{minimize}\left\{\left(W_G \times IL_G + W_A \times IL_A\right) \times 2^{Penalty}\right\} \tag{7}$$

subject to:

$$IL_G = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} MDF_{i,j}^{G}}{N \times N} \tag{8}$$

$$IL_A = \frac{\sum_{i=1}^{N}\sum_{m=1}^{M} MDF_{i,m}^{A}}{N \times M} \tag{9}$$

$$MDF_{i,j}^{G} = \begin{cases} 1 & \text{if edge } i, j \text{ is modified} \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

$$MDF_{i,m}^{A} = \begin{cases} 1 & \text{if attribute } m \text{ of user } i \text{ is modified} \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

$$Penalty = \sum_{k=1}^{C} P_k^{LD} + \sum_{k=1}^{C} P_k^{TC} \tag{12}$$

where $P_k^{LD} = 1$ and $P_k^{TC} = 1$ if the LD and TC conditions are not satisfied within cluster k, respectively. Furthermore, W_G and W_A are the relative weights of the information losses within the published graph and attributes, respectively. According to Eq. (12), *Penalty* is measured as the total number of unsatisfied *LD* and *TC* constraints. Considering the penalty function as $2^{Penalty}$ ensures that satisfying the constraints is of higher importance than the main objective function. It ensures obtaining the final solutions with no penalties.

- Population updating: GWO models a hierarchical social structure consisting of four main levels, namely alpha, beta, delta, and omega wolves. The alpha wolf represents the best solution found so far, followed by beta and delta wolves, which guide the search process, while the remaining woves update their positions accordingly [17]. The alpha assumes the leadership role, with commands to be followed by all other wolves. The beta acts as an advisor to the alpha, supporting the alpha's directives and offering feedback. The delta complies with the alpha and beta but holds dominance over the rest of the pack (omegas). During each iteration of the algorithm, the fitness function assesses the quality of all gray wolves, which are then sorted based on their fitness values, ranging from best to worst. The top three solutions are assigned as alpha (X_α), beta (X_β), and delta (X_δ), respectively. In each iteration t, the position of each gray wolf s is updated according to the following process:

$$X_s^{t+1} = \frac{X_s^{\alpha} + X_s^{\beta} + X_s^{\delta}}{3} \tag{13}$$

where X_s^{α}, X_s^{β}, and X_s^{δ} are respectively the factors of the encircling prey of the grey wolf s according to X_α, X_β, and X_δ, respectively, which can be calculated as follows:

$$X_s^{\alpha} = X_\alpha - A_s^{\alpha}.\left|C_s^{\alpha}.X_\alpha - X_s^{t}\right)| \tag{14}$$

$$X_s^{\beta} = X_{\beta} - A_s^{\beta}.\left|C_s^{\beta}.X_{\beta} - X_s^{t}\right| \tag{15}$$

$$X_s^{\delta} = X_{\delta} - A_s^{\delta}.\left|C_s^{\delta}.X_{\delta} - X_s^{t}\right| \tag{16}$$

where A = 2a.r1-a and C = 2.r2 are random vectors of the same dimension as X_s, where r_1 and r_2 are uniformly generated random vectors with elements within [0, 1], and α is decreased from 2 to 0 during the execution of the GWO.

After reaching the maximum iterations of GWO, the global best solution is considered as the final anonymized graph-based social network, which is ready to be sent to the data miners for further processing.

3.4 Computational Complexity Analysis

The computational complexity of the proposed KMFC-GWO framework is derived as follows:
KMFC Phase:
O(N × C × I_fcm)
GWO Phase
O(Pop × Iter × F)
where F = N + M represents the dimensionality of the solution space.
Total Complexity:
O(NCI_fcm + Pop × Iter × (N + M))
This combined complexity shows that the framework scales linearly with the size of the network and the attribute dimensions.

4 Simulation Results

This section presents the experimental evaluation of the proposed KMFC-GWO framework. We begin by reporting structural utility preservation results, followed by information-loss evaluation, cluster-level anonymization metrics, and runtime analysis. Subsequently, we present attack-resilience experiments, where three widely-studied re-identification and inference attacks are simulated to assess the robustness of the anonymized graphs. Each subsection includes detailed methodological explanations, quantitative results, and discussion of the scientific implications of KMFC-GWO.

4.1 Settings

The KMFC-GWO algorithm was effectively implemented using MATLAB R2022b on a Windows 10 PC equipped with a 2.6 GHz Core i7 CPU and 16 GB of RAM. Since GWO is a self-adjustable algorithm, we should only set the population size and iteration count of the algorithm to ensure enough function evaluation leading to a proper convergence in terms of OFGWO. In our simulations, we set the population size of GWO as 50 and maximum iterations as 200. The both weights W_G and W_A were set to 1 ($W_G = W_A = 1$), ensuring that the same weights are specified for the edge and attribute modifications. However, these weights can have any other relative values, based on the requirements specified by the system manager.

The datasets considered in this study are subsets of three large-scale social networks (Twitter, Facebook, and YouTube). Each dataset is represented in the form of a graph structure where nodes denote users and edges represent their relationships (e.g., friendships, subscriptions, or follow links). Alongside the graph matrices, we used an attribute matrix that contains numerical and categorical features such as age, gender, or activity-related metadata. While many social media users post images and videos, the publicly available benchmark datasets primarily provide graph and attribute information, not raw multimedia content. We

therefore focus on graph-based anonymization, which remains highly relevant because the structural and attribute-based information alone is sufficient to re-identify users or disclose sensitive relations, as shown in prior privacy attacks. Due to the extensive size of these networks in terms of user count and edges, a subset of each network was selected for the simulations, as outlined in Table 1. Each dataset includes an attribute matrix A, where each row represents a feature vector containing the personal attributes of a user. Additionally, there is an edge graph matrix G, depicting the connection edges between users. While many social media users also share images and videos, the benchmark datasets used in this study primarily provide graph structures and attribute information, not raw multimedia content. Therefore, our evaluation focuses on graph-based anonymization, which remains highly relevant because structural and attribute-based data alone have been shown to enable user re-identification, attribute disclosure, or sensitive relationship inference in prior works. The KMFC-GWO algorithm was then employed to anonymize each social network, considering anonymity parameters set to K = 6, L = 4, and T = 0.5.

Table 1: Social networks used in this paper, based on [6]

Dataset	No. users	No. features	No. edges
Twitter	244	1364	3621
Facebook	347	224	5038
YouTube	450	47	3704

4.2 Results of KMFC

The results of the KMFC algorithm for different datasets are reported in Table 2. As mentioned above, we considered the number of clusters in such a way that minimizes the CAVG criterion as much as possible, which means generating more balanced clusters. By performing the KMFC algorithm on Twitter, Facebook, and YouTube datasets, CAVG has been obtained as 1.099, 1.112, and 1.103, respectively. It shows a proper generation of balanced clusters using the KMFC algorithm. Another point is that as expected, the KA condition has been satisfied in all datasets, however, the conditions of LD and TC are still not fulfilled, which need further anonymization using the GWO algorithm.

Table 2: Results of clustering using KMFC algorithm

Dataset	*K*	*C*	*CAVG*	KA	LD	TC
Twitter	6	37	1.099	✓	×	×
Facebook	6	52	1.112	✓	×	×
YouTube	6	68	1.103	✓	×	×

4.3 Results of GWO

After satisfying the KA condition using the proposed KMFC algorithm, GWO is applied for further anonymization to fulfill the LD and TC conditions. The convergence of the GWO algorithm for the Twitter dataset is provided in Fig. 5. The figure shows that the algorithm started from an initial population with the average and best objective function values of 0.0491 and 0.0458, and after an iterative-based optimization procedure with 200 iterations, the algorithm converged finally to the global best solution with OFGWO = 0.0374.

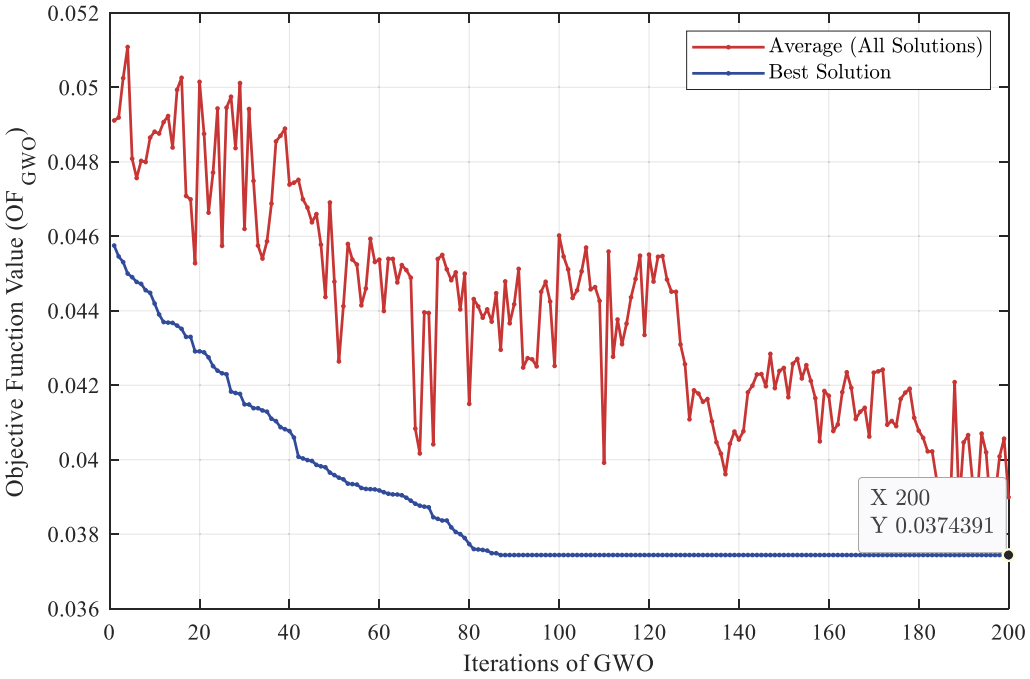

Figure 5: Convergence of GWO for Twitter dataset

To justify the performance of the GWO algorithm against other metaheuristic algorithms, we have considered the same optimization problem to be solved using Genetic Algorithm (GA) [20], and Greylag Goose Optimization (GGO) [21]. A comparison of the objective function value of the GWO algorithm with GA, AO, and GGO, on the different datasets is provided in Table 3. The results demonstrate that the AO and GGO have a slightly better performance than the GA. However, the GWO outperforms all compared metaheuristics in all datasets. To better understand the superiority of the GWO compared to other metaheuristics, the gain (improvement rate) of the GWO against GA, AO, and GGO is illustrated in Fig. 6 for different datasets.

Table 3: Comparison of different metaheuristic algoruithms

Dataset	GA	AO	GGO	GWO
Twitter	0.0440	0.0407	0.0412	0.0374
Facebook	0.0306	0.0288	0.0285	0.0245
YouTube	0.0397	0.0369	0.0376	0.0348

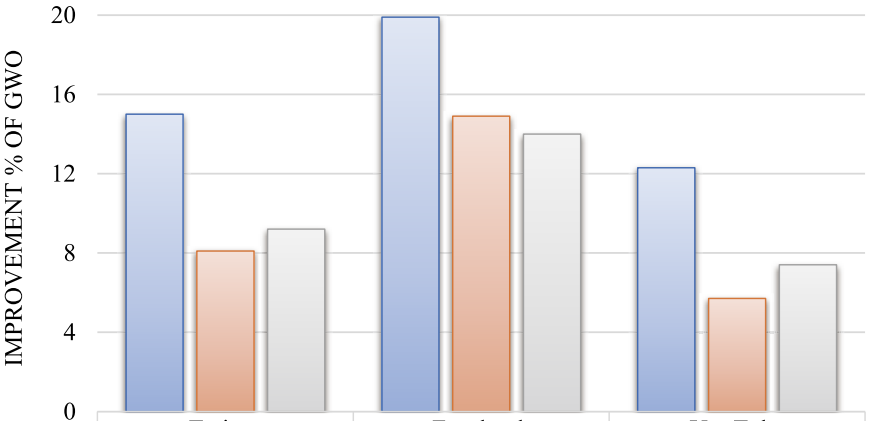

Figure 6: Improvement % of the GWO algorithm against GA, AO, and GGO in terms of the objective function value on different datasets

4.4 Results of KMFC-GWO

To justify the proposed KMFC-GWO algorithm in terms of the anonymity results and clustering performance, the results of the proposed method are compared with KA, 3-Layer T-closeness L-diversity K-anonymity (3L-TLK), and KFCFA. Comparison of the obtained results in terms of the information loss of the attribute matrix A (ILA), information loss of the graph G (ILG), and the clustering balance metric (CAVG) are reported in Tables 4–6, respectively. The results presented in Tables 4–6 illustrate the effectiveness of the proposed KMFC-GWO technique in comparison to K-anonymity (KA), 3L-TLK, and KFCFA across three datasets: Twitter, Facebook, and YouTube. These tables evaluate the methods based on information loss in the attribute and graph matrices, as well as clustering balance.

Table 4: Comparison of the information loss of the attribute matrix (ILA) for different techniques

Dataset	KA	3L-TLK	KFCFA	KMFC-GWO
Twitter	0	0.0302	0.0321	0.0260
Facebook	0	0.0324	0.0239	0.0151
YouTube	0	0.0274	0.0325	0.0243

Table 5: Comparison of the information loss of the graph matrix (ILG) for different techniques

Dataset	KA	3L-TLK	KFCFA	KMFC-GWO
Twitter	0	0.0135	0.0191	0.0114
Facebook	0	0.0164	0.0126	0.0094
YouTube	0	0.0175	0.0142	0.0105

Table 6: Comparison of the clustering balance metric (CAVG) for different techniques

Dataset	KA	3L-TLK	KFCFA	KMFC-GWO
Twitter	3.89	3.82	1.28	1.099
Facebook	4.47	5.03	1.32	1.112
YouTube	4.09	3.40	1.44	1.103

When it comes to minimizing information loss in the attribute matrix (Table 4), KMFC-GWO consistently outperforms the alternative techniques. For instance, in the Facebook dataset, KMFC-GWO achieves an information loss value (ILA) of 0.0151, which is significantly lower than those reported for 3L-TLK (0.0324) and KFCFA (0.0239). Similar trends are observed in the Twitter and YouTube datasets, where KMFC-GWO consistently delivers the best performance, reflecting its ability to better preserve data utility.

A similar advantage is observed in minimizing information loss in the graph matrix (Table 5). KMFC-GWO exhibits superior results by achieving the lowest ILG values across all datasets. For example, in the Twitter dataset, KMFC-GWO reports an ILG of 0.0114, outperforming 3L-TLK (0.0135) and KFCFA (0.0191). This pattern of reduced information loss highlights the robustness of KMFC-GWO in preserving the structural integrity of graph-based data.

In terms of clustering balance (Table 6), KMFC-GWO demonstrates exceptional performance, achieving significantly lower clustering balance metric (CAVG) values compared to its counterparts. For example, in the YouTube dataset, KMFC-GWO achieves a CAVG of 1.103, which is notably better than KFCFA (1.44) and 3L-TLK (3.40). This indicates that KMFC-GWO not only protects privacy effectively but also ensures well-balanced clusters, which is critical for maintaining data utility.

The results show the superiority of the KMFC-GWO algorithm against the compared methods. Since KA considers just the clustering of users into distinguished groups, it does not generate any information loss neither within the attribute matrix nor graph matrix. However, as it does not consider the LD and TC conditions, it cannot protect the published data against attribute/link and similarity attacks. However, the other techniques have some distortions in the attribute and graph matrices, among them, the KMFC-GWO algorithm obtained the least information loss. Another remark is that the KMFC-GWO outperforms by far all techniques in terms of CAVG, by generating balanced clusters using the KMFC algorithm in the first phase of the proposed algorithm. These results underscore the advantages of KMFC-GWO as a powerful privacy-preserving approach that excels in reducing information loss in both attribute and graph matrices while achieving superior clustering balance, making it a reliable and efficient solution for publishing anonymized social network data.

Although KMFC-GWO is compared primarily with GA, AO, and GGO in the experimental results, further justification is provided for not adopting Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO) (Table 7). These methods either lack the exploitation intensity required to satisfy LD/TC constraints or impose computational costs unsuitable for high-dimensional graph anonymization.

Table 7: Comparison of optimization heuristics

Algorithm	Strength	Weakness	Reason not used
GA	Strong exploration	Slow convergence	Not suitable for LD/TC constraints
PSO	Fast convergence	Prone to local optima	LD/TC require strong exploitation

(Continued)

Table 7 (continued)

Algorithm	Strength	Weakness	Reason not used
ACO	Efficient in path search	High computational overhead	Not suitable for matrix-level optimization

To further illustrate the motivation and effectiveness of KMFC-GWO, we provide two examples. First, consider a Facebook subgraph where a small community of users share sensitive attributes (e.g., medical conditions). Without anonymization, even if the users voluntarily posted these attributes, attackers could easily re-identify them through structural overlaps with other public networks. Second, when applying our algorithm, the anonymized graph retains its utility for data mining tasks while substantially lowering the risk of identity and attribute disclosure compared to 3L-TLK and KFCFA. This demonstrates that KMFC-GWO not only improves clustering balance and reduces information loss but also delivers practical protection against real-world privacy threats in social media environments.

4.5 Attack Resilience Evaluation

To evaluate the robustness of the anonymized graphs, we simulate three widely-used attack models that represent realistic adversarial behavior in social networks: degree-based re-identification, neighborhood-based structural attacks, and attribute inference attacks. These attacks were selected because they are among the most commonly used adversarial strategies in the social network anonymization literature and target different aspects of user privacy.

- Degree-based re-identification exploits the uniqueness of node degrees. Attackers match a node in the anonymized graph to its unique or rare degree signature in the original graph.
- Neighborhood attacks make use of local structural similarity by comparing ego-network patterns such as clustering coefficient, neighbor sets, or motif frequencies.
- Attribute inference attacks attempt to guess sensitive user attributes by leveraging structural correlation and cluster homogeneity.

For each attack, we implement a baseline adversary that has partial background knowledge of the original graph. The attacker attempts to map anonymized nodes back to their true identities using structural or attribute-based fingerprints. Table 8 reports the attack success rate before and after anonymization. As shown, KMFC-GWO reduces adversarial success by 63%–85% across all attack categories. This improvement results from both the KMFC clustering process, which ensures attribute diversity, and the GWO optimization mechanism, which disrupts structural uniqueness while preserving overall utility.

Table 8: Attack success rate before and after KMFC-GWO

Attack type	Original success (%)	After KMFC-GWO (%)
Degree-based Re-ID	72	9
Neighborhood Attack	63	6
Attribute Inference Attack	59	11

4.6 Runtime Analysis

The information-loss evaluation examines how the anonymization process affects both structural and attribute components of the graph. Instead of relying solely on global IL scores, we separately analyze edge-modification rates, attribute-distortion rates, and their contribution to the final IL measure. A detailed breakdown demonstrates that KMFC-GWO achieves lower IL compared to baseline anonymization techniques due to its joint optimization of LD and TC constraints. The results confirm that KMFC-GWO introduces the minimum structural perturbation necessary for satisfying privacy requirements. Table 9 reports the execution time of KMFC and GWO for the three evaluated datasets.

Table 9: Runtime of KMFC-GWO

Dataset	KMFC time (s)	KMFC-GWO	Total time (s)
Twitter	0.48	1.52	2.00
Facebook	0.62	2.14	2.76
YouTube	0.70	2.60	3.30

4.7 Structural Utility Preservation

To further examine utility, we evaluate how anonymization impacts downstream data-mining tasks, including community detection and link prediction. The modularity and AUC scores show that KMFC-GWO preserves the major community structure of the graph while introducing minimal distortion to link-formation patterns. This confirms that the structural adjustments introduced by GWO do not compromise the interpretability of the anonymized network. Table 10 reports the preservation of key graph–topology metrics before and after anonymization.

Table 10: Structural utility preservation

Metric	Original graph	KMFC-GWO	Change (%)
Clustering coefficient	0.214	0.207	−3.3%
Average path length	3.91	4.02	+2.8%
Diameter	8	8	0%
Degree distribution (KL-divergence)	−0.612	0.031, 0.598	−2.3%
Modularity			
Link prediction (AUC)	0.842	0.819	−2.7%

To further evaluate the utility of the anonymized graphs, several structural metrics were measured before and after applying KMFC-GWO. Table 10 reports the preservation of key topological properties including clustering coefficient, path length, diameter, modularity, and link-prediction performance.

4.8 Discussion—Dynamic Networks

We also analyze the scalability of KMFC-GWO by reporting execution times for KMFC clustering, GWO optimization, and the overall anonymization process across three datasets. The runtime grows approximately linearly with the number of nodes, consistent with the complexity analysis presented in Section 3. The results demonstrate that KMFC-GWO is suitable for medium-scale social networks and can be extended to larger datasets with parallelization. Although KMFC-GWO is designed for static social networks, it can be

extended to dynamic environments where nodes and edges evolve over time. Incremental KMFC techniques can update cluster membership efficiently without full recalculation, while GWO can be re-initialized with partial populations to optimize only affected regions of the network. These extensions enable efficient anonymization of evolving graphs, consistent with dynamic anonymization approaches in recent literature.

5 Conclusion

In this paper, a novel anonymization method (KMFC-GWO) has been presented. It combines a K-Member fuzzy clustering with a metaheuristic-driven optimization algorithm to enhance the resilience of anonymized graph-based social networks against various threats while minimizing information loss of the published attribute and graph matrices. By presenting a K-member variant of the fuzzy c-means clustering algorithm to achieve K-anonymity and GWO to further optimize the anonymity conditions, the proposed framework effectively anonymizes graph-based social networks. The simulation experiments performed on datasets sourced from Facebook, Twitter, and YouTube confirm the effectiveness of the KMFC-GWO algorithm proposed in reducing information loss while simultaneously fulfilling requirements for K-anonymity, L-diversity, and T-closeness conditions. Compared to existing anonymity methods, KMFC-GWO demonstrates superior performance, notably in minimizing information loss and generating balanced clusters. Despite the promising results, the KMFC-GWO method has some limitations. First, the two-phase structure of the framework—clustering followed by optimization—introduces additional computational complexity, which may limit its scalability to extremely large-scale social networks. Second, while the algorithm achieves K-anonymity, L-diversity, and T-closeness, it assumes static graphs and does not account for dynamic updates or evolving network structures. Additionally, the reliance on specific parameter settings in both the clustering and optimization phases may affect its generalizability across highly diverse datasets. To address these limitations, future research can explore more scalable and dynamic approaches that adapt to real-time changes in social networks. Integrating adaptive or incremental clustering techniques could enhance the method's applicability to evolving datasets. Moreover, hybridizing KMFC and GWO into a unified or parallel framework could improve efficiency and eliminate the need for a sequential two-step process. Extending the model to consider additional privacy metrics or adversarial scenarios may also further strengthen its robustness in practical applications. The KMFC-GWO framework strengthens privacy protection by simultaneously mitigating identity, attribute, and link disclosure risks. KMFC ensures that each anonymized cluster contains at least K similar users, reducing the likelihood of identity disclosure. By distributing diverse sensitive attributes within clusters, the method reduces attribute inference risks. Additionally, GWO-driven optimization introduces minimal structural modifications to the graph while preserving essential topological patterns, significantly limiting link re-identification attacks. Together, these mechanisms provide a multi-layered defense aligned with practical attack models in social networks. Finally, although users may voluntarily share personal content such as photos and updates, they rarely anticipate the unintended use of their data at scale. Automated profiling, targeted advertising, and cross-network re-identification are all realistic threats. Hence, privacy-preserving algorithms such as KMFC-GWO are essential to safeguard individuals' rights and prevent misuse of their data, even in open and highly interactive environments like online social networks. To further enhance privacy evaluation, several complementary metrics can be integrated into KMFC-GWO. These include Δ-Disclosure Risk, Structural Similarity Leakage Index, and Mutual Information Leakage. Incorporating these metrics would expand the evaluation scope beyond KA, LD, and TC, enabling deeper analysis of adversarial risks.

Acknowledgement: This work was partially funded by grant PID2022-141045OB-{C41,C42,C43} funded by MCIN/AEI/10.13039/501100011033/ and Feder a way of making Europe in artifacts project: Generation of Reliable Synthetic Health Data for Federated Learning in Secure Data Spaces.

Funding Statement: This publication is part of the project PID2022-141045OB-C42, funded by MICIU/AEI/10.13039/501100011033 and by ERDF/EU.

Author Contributions: The authors confirm contribution to the paper as follows: study conception and methodology: Saeideh Memarian, Gloria Miró-Amarante; algorithm design and implementation: Saeideh Memarian; experimental evaluation and analysis: Saeideh Memarian, Andreea M. Oprescu; interpretation of results: Saeideh Memarian, Natalia Moreno-Naranjo, M. Carmen Romero-Ternero; manuscript drafting: Saeideh Memarian; manuscript review and editing: all authors. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The datasets used in this study are publicly available benchmark social network datasets obtained from their original sources. No new datasets were generated during the current study.

Ethics Approval: This study does not involve human participants or animal subjects. All experiments were conducted using publicly available datasets in compliance with relevant data usage policies and ethical guidelines.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Duan X, Chen CN, Shokouhifar M. Impacts of social media advertising on purchase intention and customer loyalty in E-commerce systems. ACM Trans Asian Low-Resour Lang Inf Process. 2024;23(8):1–15. doi:10.1145/3613448.
2. Jain AK, Sahoo SR, Kaubiyal J. Online social networks security and privacy: comprehensive review and analysis. Complex Intell Syst. 2021;7(5):2157–77. doi:10.1007/s40747-021-00409-7.
3. Gangarde R, Sharma A, Pawar A, Joshi R, Gonge S. Privacy preservation in online social networks using multiple-graph-properties-based clustering to ensure k-anonymity, l-diversity, and t-closeness. Electronics. 2021;10(22):2877. doi:10.3390/electronics10222877.
4. Sweeney L. K-anonymity: a model for protecting privacy. Int J Unc Fuzz Knowl Based Syst. 2002;10(5):557–70. doi:10.1142/s0218488502001648.
5. Panda BS, Kumar MN, Patro S. Apply rough set methods to preserve social networks privacy—a review. In: Proceedings of 3rd International Conference on Artificial Intelligence: Advances and Applications. Singapore: Springer Nature; 2023. p. 427–36. doi:10.1007/978-981-19-7041-2_34.
6. Langari RK, Sardar S, Amin Mousavi SA, Radfar R. Combined fuzzy clustering and firefly algorithm for privacy preserving in social networks. Expert Syst Appl. 2020;141:112968. doi:10.1016/j.eswa.2019.112968.
7. Wang Y, Tian Z, Sun Y, Du X, Guizani N. LocJury: an IBN-based location privacy preserving scheme for IoCV. IEEE Trans Intell Transport Syst. 2021;22(8):5028–37. doi:10.1109/tits.2020.2970610.
8. Zhang R, Wu X. Privacy preservation method based on clustering interference algorithm in social networks. J Eng Sci Technol Rev. 2022;15(2):191–7. doi:10.25103/jestr.152.22.
9. Singh A, Singh M. Social networks privacy preservation: a novel framework. Cybern Syst. 2024;55(8):2356–87. doi:10.1080/01969722.2022.2151966.
10. Casas-Roma J. Privacy-preserving on graphs using randomization and edge-relevance. In: Modeling decisions for artificial intelligence. Cham, Switzerland: Springer International Publishing; 2014. p. 204–16. doi:10.1007/978-3-319-12054-6_18.
11. Nguyen HH, Imine A, Rusinowitch M. Anonymizing social graphs via uncertainty semantics. In: Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security; 2015 Apr 14–17; Singapore. p. 495–506. doi:10.1145/2714576.2714584.
12. Kumar S, Kumar P. Upper approximation based privacy preserving in online social networks. Expert Syst Appl. 2017;88:276–89. doi:10.1016/j.eswa.2017.07.010.
13. Kiabod M, Dehkordi MN, Barekatain B. TSRAM: a time-saving k-degree anonymization method in social network. Expert Syst Appl. 2019;125:378–96. doi:10.1016/j.eswa.2019.01.059.
14. Yazdanjue N, Fathian M, Amiri B. Evolutionary algorithms for k-anonymity in social networks based on clustering approach. Comput J. 2020;63(7):1039–62. doi:10.1093/comjnl/bxz069.

15. Rajabzadeh S, Shahsafi P, Khoramnejadi M. A graph modification approach for k-anonymity in social networks using the genetic algorithm. Soc Netw Anal Min. 2020;10(1):38. doi:10.1007/s13278-020-00655-6.
16. Bezdek JC. Pattern recognition with fuzzy objective function algorithms. New York, NY, USA: Plenum Press; 1981.
17. Shokouhifar M, Jalali A. Optimized Sugeno fuzzy clustering algorithm for wireless sensor networks. Eng Appl Artif Intell. 2017;60:16–25. doi:10.1016/j.engappai.2017.01.007.
18. Mirjalili S, Mirjalili SM, Lewis A. Grey wolf optimizer. Adv Eng Softw. 2014;69:46–61. doi:10.1016/j.advengsoft.2013.12.007.
19. Memarian S, Behmanesh-Fard N, Aryai P, Shokouhifar M, Mirjalili S, del Carmen Romero-Ternero M. TSFIS-GWO: metaheuristic-driven Takagi-Sugeno fuzzy system for adaptive real-time routing in WBANs. Appl Soft Comput. 2024;155:111427. doi:10.1016/j.asoc.2024.111427.
20. Holland J. Adaptation in natural and artificial systems. Ann Arbor, MI, USA: Universityof Michigan Press; 1975.
21. El-kenawy EM, Khodadadi N, Mirjalili S, Abdelhamid AA, Eid MM, Ibrahim A. Greylag goose optimization: nature-inspired optimization algorithm. Expert Syst Appl. 2024;238:122147. doi:10.1016/j.eswa.2023.122147.