electronics

MDPI

Systematic Review

A Comprehensive Analysis of Privacy-Preserving Solutions Developed for Online Social Networks

Abdul Majeed [1,*], **Safiullah Khan** [2] and **Seong Oun Hwang** [1,*]

[1] Department of Computer Engineering, Gachon University, Seongnam 13120, Korea
[2] Department of IT Convergence Engineering, Gachon University, Seongnam 13120, Korea; safi@gachon.ac.kr
* Correspondence: ab09@gachon.ac.kr (A.M.); sohwang@gachon.ac.kr (S.O.H.); Tel.: +82-31-750-5327 (S.O.H.)

Abstract: Owning to the massive growth in internet connectivity, smartphone technology, and digital tools, the use of various online social networks (OSNs) has significantly increased. On the one hand, the use of OSNs enables people to share their experiences and information. On the other hand, this ever-growing use of OSNs enables adversaries to launch various privacy attacks to compromise users' accounts as well as to steal other sensitive information via statistical matching. In general, a privacy attack is carried out by the exercise of linking personal data available on the OSN site and social graphs (or statistics) published by the OSN service providers. The problem of securing user personal information for mitigating privacy attacks in OSNs environments is a challenging research problem. Recently, many privacy-preserving solutions have been proposed to secure users' data available over OSNs from prying eyes. However, a systematic overview of the research dynamics of OSN privacy, and findings of the latest privacy-preserving approaches from a broader perspective, remain unexplored in the current literature. Furthermore, the significance of artificial intelligence (AI) techniques in the OSN privacy area has not been highlighted by previous research. To cover this gap, we present a comprehensive analysis of the state-of-the-art solutions that have been proposed to address privacy issues in OSNs. Specifically, we classify the existing privacy-preserving solutions into two main categories: privacy-preserving graph publishing (PPGP) and privacy preservation in application-specific scenarios of the OSNs. Then, we introduce a high-level taxonomy that encompasses common as well as AI-based privacy-preserving approaches that have proposed ways to combat the privacy issues in PPGP. In line with these works, we discuss many state-of-the-art privacy-preserving solutions that have been proposed for application-specific scenarios (e.g., information diffusion, community clustering, influence analysis, friend recommendation, etc.) of OSNs. In addition, we discuss the various latest de-anonymization methods (common and AI-based) that have been developed to infer either identity or sensitive information of OSN users from the published graph. Finally, some challenges of preserving the privacy of OSNs (i.e., social graph data) from malevolent adversaries are presented, and promising avenues for future research are suggested.

Keywords: anonymization; online social network; privacy; privacy-preserving graph publishing; utility; social network analysis; community clustering; friend recommendations

check for updates

Citation: Majeed, A.; Khan, S.; Hwang, S.O. A Comprehensive Analysis of Privacy-Preserving Solutions Developed for Online Social Networks. *Electronics* **2022**, *11*, 1931. https://doi.org/10.3390/electronics11131931

Academic Editors: Muath Obaidat, Thaier Hayajneh and Eman Hammad

Received: 31 May 2022
Accepted: 16 June 2022
Published: 21 June 2022

1. Introduction

In recent years, the adoption of online social networks (OSNs) has significantly increased (e.g., only Facebook owns 1.23 billion monthly active users), and OSNs have become one of the most famous platforms for social interactions. People use OSNs to interact as well as to share personal experiences and information with their friends. Many companies are using social media platforms to engage with their customers as well as to advertise their products/events. Due to the continuously growing popularity of OSNs, a large amount of personal big data is generated on daily basis (for example, Twitter generates about 500 million tweets each day and around 200 billion tweets per year). These data can assist in improving people's quality of life as well as benefit various companies (e.g., advertisers, application developers, recommendation companies, content creators and sellers, policy makers, and so on). However, these data encompass sensitive information about people's social interaction, spatial-temporal activities, demographics, finance, disease, mobility, religious/political views, etc., that needs privacy preservation to protect it from prying eyes. In recent years, privacy preservation has become more challenging due to rapid advancements in data mining and artificial intelligence tools and the availability of personal data (e.g., user profiles on OSN sites). These tools are good at finding sensitive information from large-scale data as well as predicting sensitive information using pre-trained models. Hence, privacy preservation of user data has become one of the most urgent research problems in OSNs.

OSNs are structures that depict a set of entities (i.e., users) and the ties/relations between them [1]. OSNs are usually represented with an undirected graph, G, $G = (V, E)$, where V denotes the set of users (i.e., $V = \{v_1, v_2, \ldots, v_N\}$), and E is the set of edges (i.e., $E \subseteq \binom{V}{2}$)). In simple words, v_i is any real-world user of SN, and $|E|$ denotes the links of v_i with other $N - v_i$ users. The link between any two users, v_i and v_j, can be correspondence, friendship, collaboration, affiliation to a group/party, etc. In addition to set E, each node v in a G usually encompasses a set of attributes A, where $A = \{a_1, a_2, \ldots, a_P\}$. The labels for these attributes can be age, gender, race, and zipcode. For instance, $A = \{a_1 = age$, $a_2 = gender, a_3 = race, \ldots, a_P = zipcode\}$. The domain of values for each attribute can be different, for example, if a_2=gender, then $a_2 = \{M, F\}$. In addition to the non-identifying attributes, in some cases, A can contain one/two types of sensitive information (SI), denoted with S, where $S = \{s_1, s_2, \ldots, s_I\}$. Hence, the overall structure of attribute set δ can be represented as $\delta = \{A, S\}$, where A and S denote the basic attributes, known as quasi-identifiers (QIDs) and SI, respectively. The G, when nodes contain attribute information as well, can be denoted as $G = (V, E, \delta)$. A conceptual overview of the G is shown in Figure 1.

In Figure 1, there are nine users labeled as $v_1 \rightarrow v_9$, and the number of edges is distinct for each user. For example, v_1 has three edges, and v_4 has one edge. The $|E|$ of any node (v_i) is also called the degree of that respective node, denoted as $deg(v_i)$. Each node has QIDs as well as SI. For the sake of simplicity, we mark SI with bold fonts in Figure 1. The structure of δ can be denoted as $\delta = \{A, S\}$, where $A = \{a_1 = name, a_2 = gender, a_3 = age\}$, and $S = \{income\}$. The G can be directed, undirected, weighted, labeled, unlabeled, etc., depending upon the scenario [2].

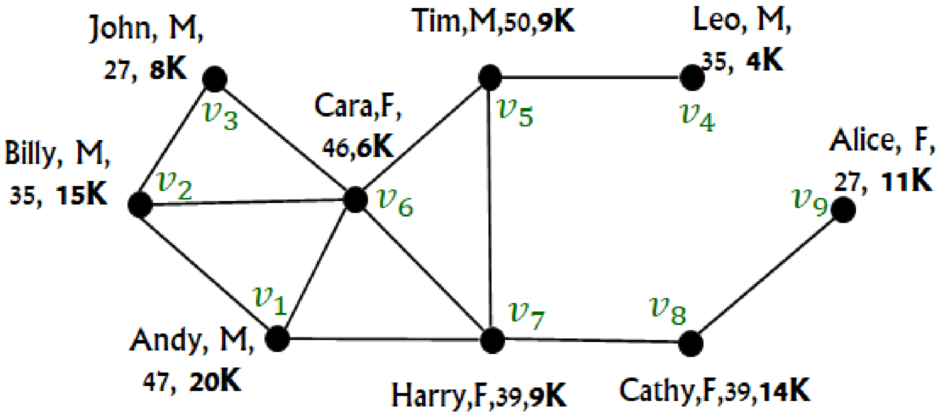
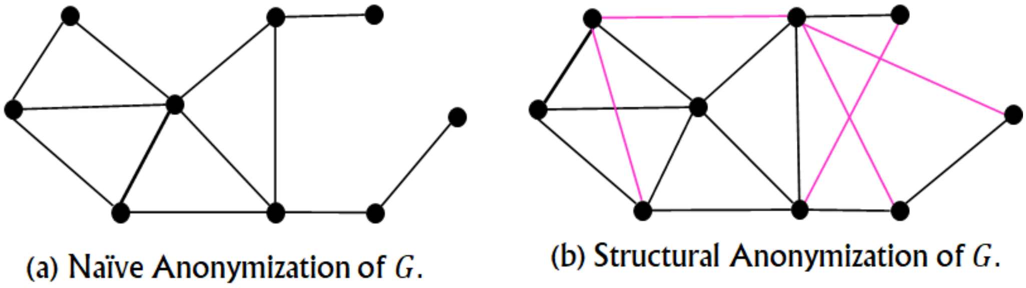


Figure 1. Overview of SN data (friends network) modeling/representation with G.

In recent years, the distribution of G with researchers/data miners has become a routine matter to find insights from G about people [3]. The sharing and analysis of G have a wide range of benefits for people. For example, better service/product recommendations by community-based clustering [4], information diffusion to targeted users [5], appropriate friend recommendations [6], point of interest recommendations [7], traffic incidents analysis [8], influence spreading [9], and route recommendations [10], to name a few. The usage of SN offers users many other benefits such as increasing their reputation, influencing others, recieving brand offers, receiving support, and connecting with a huge community [11] . However, the flip side of using OSN analysis or mining is the loss of privacy and the inherent consequences of this. Therefore, SNs' service providers and users are still struggling to find a proper balance of social benefits as well as the potential privacy risks [12]. With the rapid digitization, both the scale and scope of OSN privacy breaches are expanding, impacting millions of users with either the loss of data or dignity. OSNs' service providers are constantly integrating privacy-protection tools and upgrading the existing privacy settings to combat privacy issues. In this paper, we focus on privacy violations in *G*, and therefore, we discuss most concepts concerning *G* analysis, mining, and sharing.

There are two famous and state-of-the-art approaches for privacy preservation in *G* publishing, named naïve and structural anonymization [13]. In the former category, only the structure of *G* is published by removing all attributes of nodes and edges (see Figure 2a). In contrast, the latter category modifies the structure of *G* to preserve privacy (see Figure 2b). In Figure 2b, five new edges have been introduced to anonymize *G*.


(a) Naïve Anonymization of *G*.

(b) Structural Anonymization of *G*.

Figure 2. Overview of SN data anonymization by using original *G* given in Figure 1.

Researchers have noticed that naïve anonymization may not be sufficient to provide strong resilience against privacy breaches [14]. In contrast, structural anonymization provides a relatively higher defense against privacy breaches by modifying graph structures. In Figure 2, five new edges have been introduced to change the structure of *G* for privacy preservation. Recently, many solutions have been proposed to preserve the privacy of SN users in *G* publishing [15–21]. These solutions have been used to preserve either nodes' or edges' privacy in the release of *G*. Recently, differential privacy-based solutions have also been proposed to alter the *G*'s structure for privacy preservation [22]. Despite the success of these solutions, privacy issues can stem in multiple formats, and robust solutions are needed to overcome all types of privacy issues.

The existing surveys related to *G* publishing cover many important aspects such as graph anonymization/de-anonymization techniques [23], graph anonymization operations [24], brief taxonomies of privacy models [25], anonymity frameworks for graph data [26], privacy/utility evaluating metrics employed by the anonymization mechanisms [27,28], random *G* modeling [29], and data mining from *G* [30]. Although we fully affirm the contributions of these surveys, these studies have the following five research gaps: (i) most surveys provided very limited knowledge about most aspects concerning OSN privacy, especially privacy-preserving *G* publishing and critical information that needs privacy preservation in *G*; (ii) the critical and experimental details of most studies have not been reported thoroughly; (iii) the discussion/analysis of privacy preservation in application-specific scenarios of OSNs has not been investigated in detail; (iv) the high-level taxonomy of privacy-preserving approaches (i.e., common + AI) used in publishing *G* for data mining and analytical purposes has not been provided; (v) the significance of artificial intelligence (AI) techniques in the context of graph publishing as well jeopardizing user's privacy has not been comprehensively highlighted. Table 1 presents a detailed comparison of this review paper with existing SOTA surveys and review papers. Many current surveys either present a single type of anonymization or lack basic examples/knowledge that an early researcher needs to gain access to or understand this domain. Furthermore, the role of AI in the privacy of OSNs has not been thoroughly explained. This review article resolves the aforementioned limitations of the existing reviews and provides sufficient knowledge of privacy (or privacy disclosures) in OSNs from a broader perspective.

Table 1. Overview and comparisons of existing surveys with our review paper.

Ref.	Coverage of Anonymization Methods						Coverage of de-Anonymization Methods		Privacy in Multiple Scenario (s) of OSNs		Experimental Details
	MM	CM	PAGCM	DPM	AIM	HM	Common	AI	Common	AI	
Du et al. [3]	×	×	×	✓	×	×	×	×	×	×	○
Alemany et al. [11]	✓	×	×	×	✓	×	×	×	○	×	✓
Shejy et al. [12]	○	○	○	○	○	○	✓	✓	✓	×	×
Majeed et al. [13]	✓	✓	○	○	×	×	✓	×	○	×	×
Avinash et al. [17]	○	○	×	×	×	×	✓	×	×	×	✓
Ji et al. [23]	✓	×	×	○	×	✓	✓	×	✓	×	○
Casas et al. [24]	✓	✓	×	×	×	×	✓	×	○	×	○
Zhou et al. [25]	✓	✓	×	×	×	×	✓	×	○	×	×
Wu et al. [26]	✓	✓	×	×	×	×	×	×	×	×	×
Praveena et al. [27]	✓	×	×	✓	×	×	✓	×	×	×	○
Joshi et al. [28]	✓	✓	○	×	×	×	✓	×	✓	×	○
Droby et al. [29]	✓	✓	✓	○	✓	○	✓	○	✓	×	×
Injadat et al. [30]	✓	✓	×	×	✓	×	✓	○	×	○	×
This review paper	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Abbreviations: MM (modification methods), CM (clustering methods), PACGM (privacy-aware graph computing methods), DPM (differential privacy-based methods), AIM (artificial intelligence-based methods), HM (hybrid methods). **Key**: ✓ ⇒ available/reported and × ⇒ not available/not reported, ○ → partially covered.

The major contributions of this article to OSNs' privacy are summarized as follows.

- It presents a comprehensive analysis and the findings of the state-of-the-art (SOTA) solutions that have been proposed to address privacy issues in OSNs;
- It provides a high-level taxonomy of common as well as AI-based privacy-preserving approaches that have proposed as ways to combat the privacy issues in PPGP along with recent studies in each category;
- It discusses many practical solutions that have been proposed for privacy preservation in application-specific scenarios (e.g., information diffusion, community clustering, influence analysis, friend recommendation, etc.) of OSNs that remained unexplored in the recent literature;
- We discuss various generic and AI-based de-anonymization techniques that have been developed to infer SI from the anonymized graph (a.k.a. the de-anonymization of G);
- Technical challenges of preserving privacy in OSNs in recent times and promising opportunities for future research are discussed in detail;
- The novelty of our work is to provide a systematic analysis of SOTA methods focusing on OSNs from two aspects (e.g., defense → anonymity and attack → de-anonymity), identify novel application scenario(s) of OSNs and corresponding privacy-preserving approaches, analyze the role of AI in privacy preservation of OSNs (or privacy breaches), identify major users' privacy challenges that OSNs' service providers are facing or can likely face in the coming years, and list potential research avenues for researchers. Lastly, this work makes a timely contribution towards responsible data science (https://redasci.org/, accessed on 2 May 2022) amid rapid technical advancements in OSNs services/sites in recent years.

The rest of this article is organized as follows. Section 2 presents detailed background of privacy concepts in OSNs. Section 3 discusses the taxonomy of privacy-preserving graph publishing (PPGP) approaches and SOTA developments in each category. Section 4 presents major developments regarding privacy preservation in application-specific scenarios of SNs. Section 5 highlights the major developments in de-anonymization of G. Section 6 discusses the challenges to the privacy protection of G. Section 7 lists various research directions that are vital to combat privacy issues in OSNs. Section 8 discusses the limitations of this review. Finally, we conclude this paper in Section 9. Figure 3 presents the high-level structure of this review article.

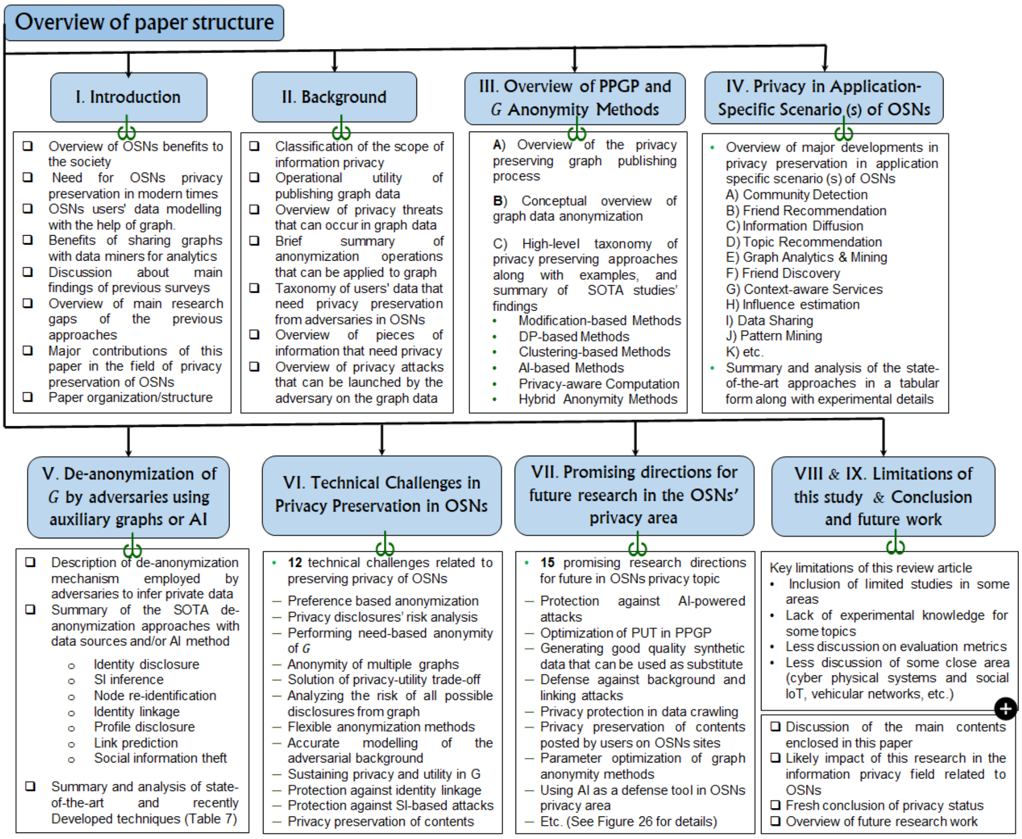

Overview of paper structure
I. Introduction
❑ Overview of OSNs benefits to the society
❑ Need for OSNs privacy preservation in modern times
❑ OSNs users' data modelling with the help of graph.
❑ Benefits of sharing graphs with data miners for analytics
❑ Discussion about main findings of previous surveys
❑ Overview of main research gaps of the previous approaches
❑ Major contributions of this paper in the field of privacy preservation of OSNs
❑ Paper organization/structure
II. Background
❑ Classification of the scope of information privacy
❑ Operational utility of publishing graph data
❑ Overview of privacy threats that can occur in graph data
❑ Brief summary of anonymization operations that can be applied to graph
❑ Taxonomy of users' data that need privacy preservation from adversaries in OSNs
❑ Overview of pieces of information that need privacy
❑ Overview of privacy attacks that can be launched by the adversary on the graph data
III. Overview of PPGP and G Anonymity Methods
A) Overview of the privacy preserving graph publishing process
B) Conceptual overview of graph data anonymization
C) High-level taxonomy of privacy preserving approaches along with examples, and summary of SOTA studies' findings
• Modification-based Methods
• DP-based Methods
• Clustering-based Methods
• AI-based Methods
• Privacy-aware Computation
• Hybrid Anonymity Methods
IV. Privacy in Application-Specific Scenario (s) of OSNs
• Overview of major developments in privacy preservation in application specific scenario (s) of OSNs
A) Community Detection
B) Friend Recommendation
C) Information Diffusion
D) Topic Recommendation
E) Graph Analytics & Mining
F) Friend Discovery
G) Context-aware Services
H) Influence estimation
I) Data Sharing
J) Pattern Mining
K) etc.
• Summary and analysis of the state-of-the-art approaches in a tabular form along with experimental details
V. De-anonymization of G by adversaries using auxiliary graphs or AI
❑ Description of de-anonymization mechanism employed by adversaries to infer private data
❑ Summary of the SOTA de-anonymization approaches with data sources and/or AI method
○ Identity disclosure
○ SI inference
○ Node re-identification
○ Identity linkage
○ Profile disclosure
○ Link prediction
○ Social information theft
❑ Summary and analysis of state-of-the-art and recently Developed techniques (Table 7)
VI. Technical Challenges in Privacy Preservation in OSNs
• 12 technical challenges related to preserving privacy of OSNs
– Preference based anonymization
– Privacy disclosures' risk analysis
– Performing need-based anonymity of G
– Anonymity of multiple graphs
– Solution of privacy-utility trade-off
– Analyzing the risk of all possible disclosures from graph
– Flexible anonymization methods
– Accurate modelling of the adversarial background
– Sustaining privacy and utility in G
– Protection against identity linkage
– Protection against SI-based attacks
– Privacy preservation of contents
VII. Promising directions for future research in the OSNs' privacy area
• 15 promising research directions for future in OSNs privacy topic
– Protection against AI-powered attacks
– Optimization of PUT in PPGP
– Generating good quality synthetic data that can be used as substitute
– Defense against background and linking attacks
– Privacy protection in data crawling
– Privacy preservation of contents posted by users on OSNs sites
– Parameter optimization of graph anonymity methods
– Using AI as a defense tool in OSNs privacy area
– Etc. (See Figure 26 for details)
VIII & IX. Limitations of this study & Conclusion and future work
Key limitations of this review article
• Inclusion of limited studies in some areas
• Lack of experimental knowledge for some topics
• Less discussion on evaluation metrics
• Less discussion of some close area (cyber physical systems and social IoT, vehicular networks, etc.)
+
❑ Discussion of the main contents enclosed in this paper
❑ Likely impact of this research in the information privacy field related to OSNs
❑ Fresh conclusion of privacy status
❑ Overview of future research work

Figure 3. Comprehensive overview of the structure of this review paper.

2. Background

In this section, we provide a detailed analysis of threats to the validity of this review article and comprehensive background concerning OSNs privacy. In the next subsection, we state the search strings and databases that have been explored to find the related work for this review article.

2.1. Threats to Validity

For this review article, we included the SOTA studies that (1) deal with the privacy preservation of OSN data, (2) target privacy preservation in application-specific scenarios of OSNs, (3) discuss the significance of AI techniques in privacy protection/disclosure in OSNs, (4) deal with jeopardizing OSNs' users' privacy by either linking, statistical matching, or background knowledge attacks, and (5) discuss the performance evaluation in terms of privacy, utility, or computational complexity. We have used multiple phrases and combinations of strings such as 'privacy preservation in OSNs' and 'social graph publishing and anonymization' to extract the peer-reviewed articles from journals, renowned conference proceedings, recently published book chapters, and technical reports. We have mainly targeted eight databases, namely, IEEE Xplore, ScienceDirect, SpringerLink, Scopus, ACM Digital Library, MDPI, Hindawi, and Web of Science. We took advantage of the Google Scholar search engine for forward and backward searches. We have focused on papers that have been highly cited by recent studies and are highly technical with improved results. In total, 3500 documents were retrieved, and 1700 duplicated studies were removed. The titles and the abstracts' contents were carefully screened to identify potential papers. The full texts of the 1780 studies were assessed to find the highly relevant papers to be included in this review. We have excluded the articles that discussed (1) a defense solution other than anonymization, privacy preservation of stored OSNs data, and content-based privacy attacks in OSNs, (2) cyber attacks (e.g., denial of service) on OSN data, (3) threats to OSNs' security and privacy breaches in interactions. With a backward and forward search, 16 more closely related studies were retrieved. In total, 291 studies were finally selected for data extraction purposes. Figure 4 depicts the process of the SOTA article selection for this paper that was adopted from the previous SOTA reviews [31,32]. The findings of previous surveys and review articles were also used in addition to these included papers to provide distinctive features and a comprehensive performance evaluation.

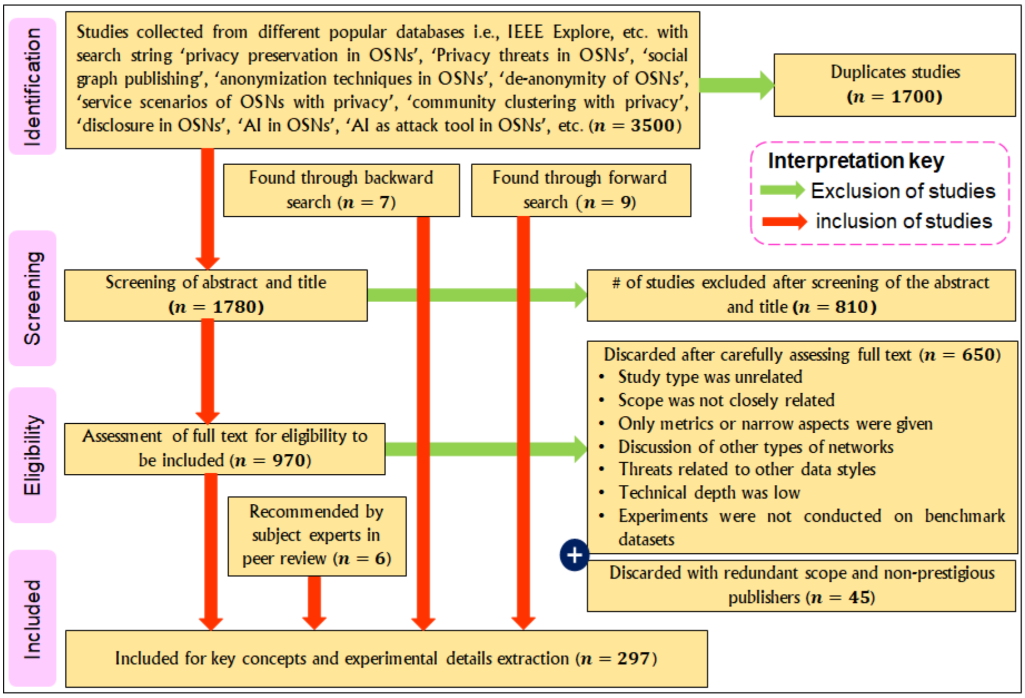

Identification
Studies collected from different popular databases i.e., IEEE Explore, etc. with search string 'privacy preservation in OSNs', 'Privacy threats in OSNs', 'social graph publishing', 'anonymization techniques in OSNs', 'de-anonymity of OSNs', 'service scenarios of OSNs with privacy', 'community clustering with privacy', 'disclosure in OSNs', 'AI in OSNs', 'AI as attack tool in OSNs', etc. (n = 3500)
Duplicates studies (n = 1700)
Interpretation key
Exclusion of studies
inclusion of studies
Found through backward search (n = 7)
Found through forward search (n = 9)
Screening
Screening of abstract and title (n = 1780)
# of studies excluded after screening of the abstract and title (n = 810)
Eligibility
Assessment of full text for eligibility to be included (n = 970)
Discarded after carefully assessing full text (n = 650)
• Study type was unrelated
• Scope was not closely related
• Only metrics or narrow aspects were given
• Discussion of other types of networks
• Threats related to other data styles
• Technical depth was low
• Experiments were not conducted on benchmark datasets
+
Discarded with redundant scope and non-prestigious publishers (n = 45)
Recommended by subject experts in peer review (n = 6)
Included
Included for key concepts and experimental details extraction (n = 297)

Figure 4. Flow diagram demonstrating the SOTA article selection process for this systematic review.

2.2. Classification in the Scope of Privacy

The nature of privacy is highly subjective, meaning its perception varies from person to person. In simple words, privacy is all about hiding SI from the prying eyes [33]. The scope of privacy mainly falls into four categories [34], as demonstrated in Figure 5. This work belongs to the first category, which is about handling (i.e., aggregation, storage, analysis, anonymization, distribution, etc.) person-specific data.

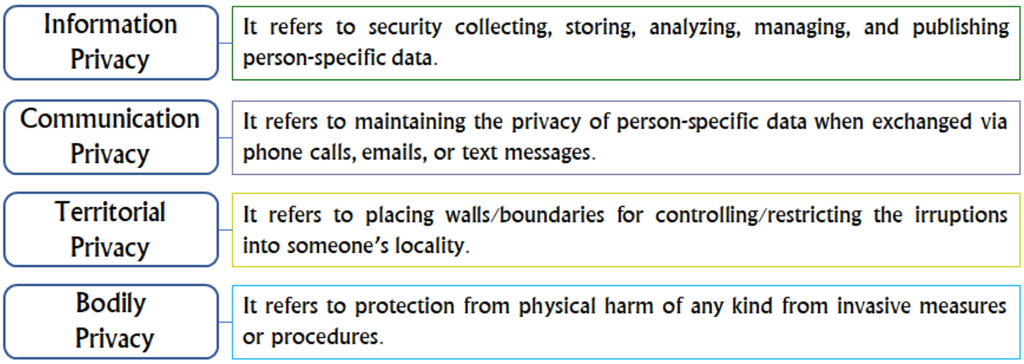

Information Privacy
It refers to security collecting, storing, analyzing, managing, and publishing person-specific data.
Communication Privacy
It refers to maintaining the privacy of person-specific data when exchanged via phone calls, emails, or text messages.
Territorial Privacy
It refers to placing walls/boundaries for controlling/restricting the irruptions into someone's locality.
Bodily Privacy
It refers to protection from physical harm of any kind from invasive measures or procedures.

Figure 5. Classification of the scope of privacy.

Person-specific data can be modeled in a variety of styles such as tables, graphs, matrices, traces, logs, images, multimedia, and hybrid [35]. However, we consider personal data represented in a graph form G, where $G = (V, E, \delta)$, in our work.

2.3. Operation Utility of Social Graphs

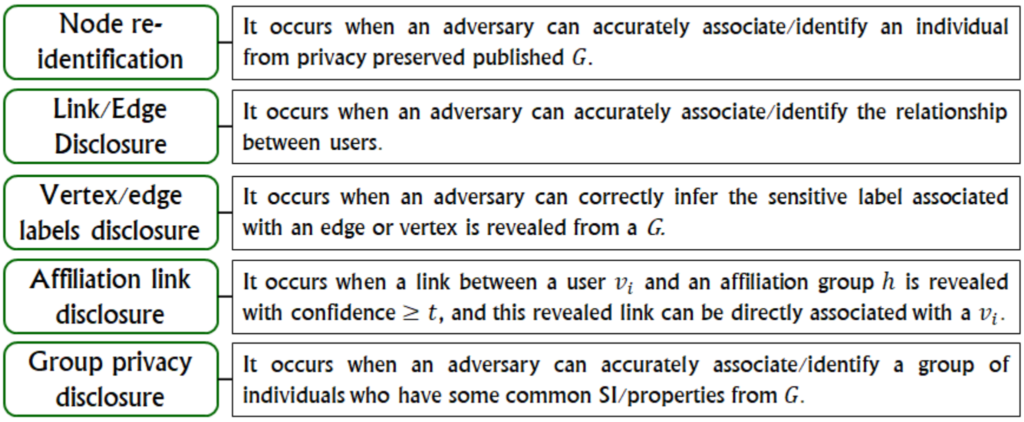
The publishing of G is vital for many analytical and data-mining purposes. The operational utility, OU, offered by G can be one of the three cases listed in Equation (1):

$$OU(l_1/l_2/l_3) = \begin{cases} l_1, & \text{Exposure of G structure only} \\ l_2, & \text{Exposure of profiles of nodes} \\ l_3, & \text{Exposure of both } l_1 \text{ and } l_2 \end{cases} \quad (1)$$

In the first case/level, the SN service provides the release of only the structure of G, and all profile information is usually hidden. In the second case, the structure of G is hidden, but profile information is shared with a researcher. In the last level, both the G structure and the δ of V is shared for analytical purposes [36].

2.4. Key Privacy Issues That Can Occur in Publishing G with Analysts

As stated earlier, publishing G is vital for many purposes, but it can introduce multiple privacy issues. We summarize the five key privacy issues occurring in G publishing in Figure 6.



Node re-identification: It occurs when an adversary can accurately associate/identify an individual from privacy preserved published G.

Link/Edge Disclosure: It occurs when an adversary can accurately associate/identify the relationship between users.

Vertex/edge labels disclosure: It occurs when an adversary can correctly infer the sensitive label associated with an edge or vertex is revealed from a G.

Affiliation link disclosure: It occurs when a link between a user v_i and an affiliation group h is revealed with confidence $\geq t$, and this revealed link can be directly associated with a v_i.

Group privacy disclosure: It occurs when an adversary can accurately associate/identify a group of individuals who have some common SI/properties from G.

Figure 6. Five famous and practical privacy threats that can occur in publishing G with researchers.

2.5. Anonymization Operation That Can Be Applied to G

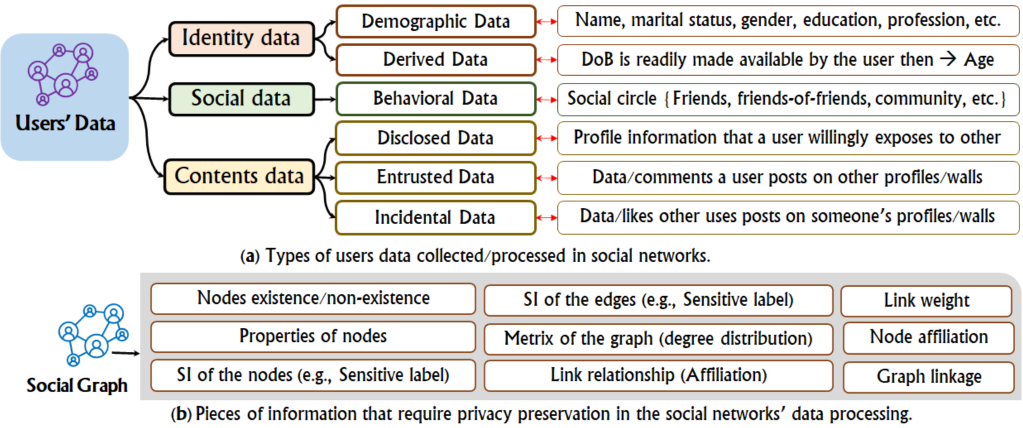
Many techniques, such as anonymization, masking, encryption, obfuscation, watermarking, zero knowledge proofs, and pseudonymization, are employed to preserve users' privacy in G. Due to conceptual simplicity, anonymization techniques have been widely used to preserve users' privacy [37,38]. Various anonymity operations are performed in order to provide sufficient resilience against contemporary privacy issues. Table 2 presents the concise description of anonymization operations that are applied to G for privacy protection. The strength/weakness and complexity of each operation vary depending upon the size and nature of G.

Table 2. Summary of anonymization operations that can be applied to G for privacy preservation.

Anonymity Operation	Brief Description	Examples
G modification	This operation modifies the structure of G by adding/removing vertices or edges.	k-degree anonymity
G generalization	This operation clusters the nodes and edges of G into super nodes/edges.	Node grouping
G Obfuscation	This operation adds noise in the form of fake edges/edges to preserve privacy.	Node-level DP
G computation	This operation computes properties from G, and releases them to analysts.	Degree, size
G Hybrid Operation	More than one anonymity operation are jointly used to perturb G.	k-degree clustered G

2.6. Important Aspects of Privacy Preservation in OSN Data

OSNs contain a treasure of information, and sufficient care is needed to preserve the privacy of most parts of that data [39]. In Figure 7, we present the different types of data collected/processed in SN and the pieces of information that require privacy preservation in publishing G. As shown in Figure 7a, there are three main types of data: identity, social, and content in SNs. All these types need privacy preservation from prying eyes. For example, if a rare user of an SN has a just one friend, and the respective friend is known to be an HIV doctor, in this case, the SI of the SN user can be inferred (e.g., he/she might have contracted an HIV disease) [40]. Similarly, profession information can lead to income disclosures. In *G*, anonymization methods need to preserve most parts of the SI shown in Figure 7b. SN data need more care regarding privacy preservation because most data can be available to the adversaries as background knowledge (BK) [41,42].


Users' Data
Identity data
Social data
Contents data
Demographic Data
Derived Data
Behavioral Data
Disclosed Data
Entrusted Data
Incidental Data
Name, marital status, gender, education, profession, etc.
DoB is readily made available by the user then → Age
Social circle {Friends, friends-of-friends, community, etc.}
Profile information that a user willingly exposes to other
Data/comments a user posts on other profiles/walls
Data/likes other uses posts on someone's profiles/walls
Social Graph
Nodes existence/non-existence
SI of the edges (e.g., Sensitive label)
Link weight
Properties of nodes
Metrix of the graph (degree distribution)
Node affiliation
SI of the nodes (e.g., Sensitive label)
Link relationship (Affiliation)
Graph linkage

(a) Types of users data collected/processed in social networks.

(b) Pieces of information that require privacy preservation in the social networks' data processing.

Figure 7. Different types of SN data and pieces of information concerning privacy in social graphs.

In Figure 8, we summarize important BK that can be within the adversaries' access, and which can lead to privacy breaches. Apart from the BK and other auxiliary types of data, a new risk known as interdependent privacy risk (IPR) has become one of the major privacy threats in SNs in recent times [43–45]. Furthermore, inference attacks [46], ML-based attacks [47], privacy leakage in health SNs [48], profile cloning [49], profile matching [50], community-based threats [51], cross-SN user matching [52], and privacy concerns in different OSN services (e.g., recommendation systems [53], query evaluations [54], and sentiment analyses [55]) have made privacy preservation in OSNs an active area of research.

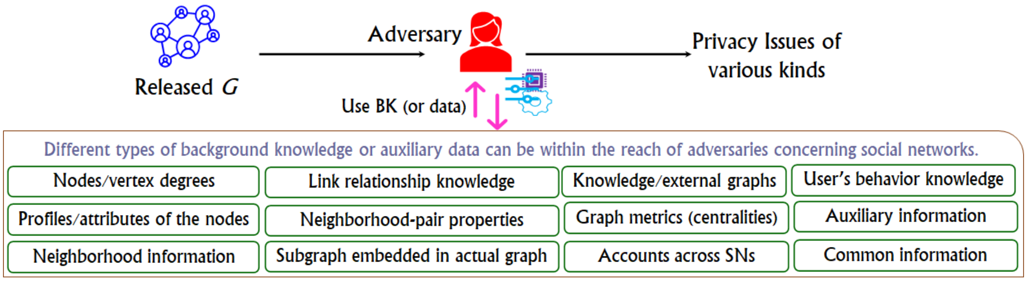

Released G
Adversary
Use BK (or data)
Privacy Issues of various kinds
Different types of background knowledge or auxiliary data can be within the reach of adversaries concerning social networks.
Nodes/vertex degrees
Link relationship knowledge
Knowledge/external graphs
User's behavior knowledge
Profiles/attributes of the nodes
Neighborhood-pair properties
Graph metrics (centralities)
Auxiliary information
Neighborhood information
Subgraph embedded in actual graph
Accounts across SNs
Common information

Figure 8. Types of BK employed by the adversaries to jeopardize users' privacy in a released *G*.

2.7. Role of Artificial Intelligence in the Domain of OSNs

Recently, AI has been extensively used in the information privacy domain for multiple purposes. It has been used to safeguard the personal data from prying eyes as well as for de-anonymizing large *G*-encompassing data of a substantial number of users [56–58]. We summarize the role of AI from three different perspectives as follows:

1. *AI as a protection tool:* AI can be used to preserve the privacy of SN data;
2. *AI as an attack tool:* AI can be used to compromise a user's privacy from SN data;
3. *AI as protection of target:* Privacy concepts can be used to secure AI systems.

Recently, many graph-type–specific, attack-specific, domain-specific, application-specific, and AI-powered anonymization techniques have been developed. In the rest of this paper, we summarize the major developments concerning *G* privacy.

3. Overview of Privacy-Preserving Graph Publishing (PPGP) and Taxonomy of PPGP Approaches Used for Online Social Networks

In this section, we discuss the life cycle of PPGP, the basic concepts of *G* anonymization, and the taxonomy of PPGP approaches. We arrange and discuss the concepts in three different subsections. In the next subsection, we discuss the life cycle of PPGP.

3.1. Overview of the Life Cycle of Ppgp

The typical PPGP process contains six steps. A concise description of all the steps is given below. In Step *A*, appropriate data are collected from relevant users. Examples of data collection are account-opening procedures in an SN website, or a check-up from a diagnostic center. In both of these scenarios, some basic information (i.e., QIDs) as well as SI is obtained. In this research, we assume that *G* has already been collected by the SN service providers (a.k.a. data owners). In Step *B*, the collected *G* data are stored in safe repositories/databases for further analysis. Storage can be in graph form (e.g., SN data) or tabular form (e.g., hospital/bank data) depending upon the nature of the data. Due to the recent advancements in technology, storage capacity has become significantly large, and fine-grained data can now be stored for utilization in multiple contexts. In Step *C*, pre-processing is applied to the collected *G*. During this step, the *G* is cleaned (outliers and isolated nodes are removed, and redundant nodes/edges are removed). In Step *D*, the cleaned *G* from Step *C* are anonymized. During data anonymization, the structure of the original *G* is modified to preserve privacy, leaving the anonymized *G*, which is useful for analysis. The anonymization can be performed in multiple ways (e.g., DP methods, constrained methods, etc.) In Step *E*, the anonymized *G* is published for SN analysis and data-mining purposes. In the final step, analytics techniques are applied to the published *G* in order to extract useful information from it. The extracted information can be used for hypothesis generation/verification or for policy making.

A conceptual overview of the anonymization techniques applied to raw data given in a raw *G* form for PPDP is demonstrated in Figure 9. As shown in Figure 9, both privacy risks and graph utility are higher at the beginning. Anonymization is applied to *G* to strike a balance between utility and privacy [59]. The anonymization approaches usually modify the structure of the *G* in such a way that both privacy and utility are preserved. In the next subsection, we present a conceptual overview of anonymizing *G* along with an example.

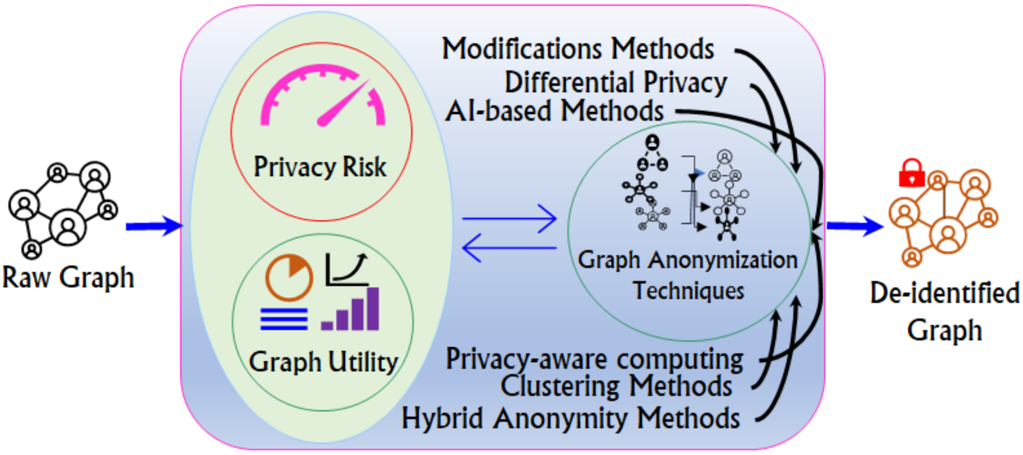


Figure 9. Overview of the anonymization process applied to *G* for PPDP.

3.2. The Basic Concepts of G Anonymization

Basically, the anonymization approaches change the structure of *G* into a new graph, G'. The size (# of nodes and edges) of G' can/cannot be the same as that of *G*. Hereafter, we refer to *G* as an original graph, and to G^6 as an anonymized graph. In Figure 10, we demonstrate an overview of the *G* anonymization with examples. The G' given in Figure 10 satisfies the *k*-degree anonymity, where $k = 2$ because each node has at least two edges.

Thus far, many graph anonymization approaches have been developed for sharing $G^{'}$ with the researchers/analysts [60–62]. In the next subsection, we present a taxonomy of graph anonymization approaches and the major developments in each category.

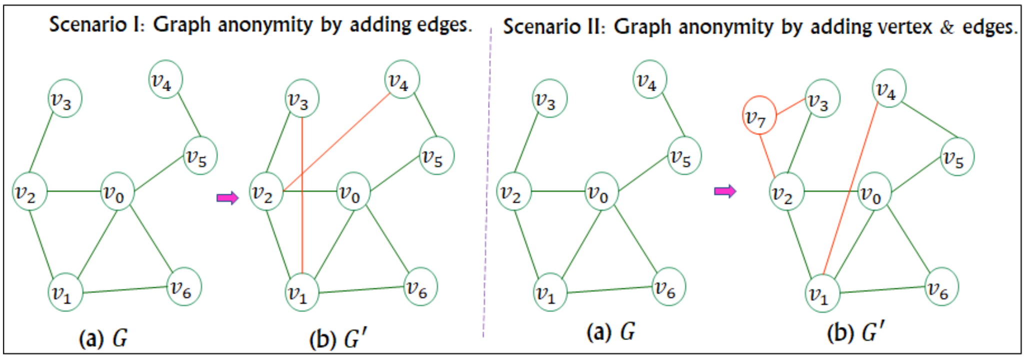

Scenario I: Graph anonymity by adding edges.
Scenario II: Graph anonymity by adding vertex & edges.
v_0
v_1
v_2
v_3
v_4
v_5
v_6
v_7
(a) G
(b) G'
(a) G
(b) G'

Figure 10. Overview of the G anonymization by adding edges and nodes and edges for PPGP.

In the next subsection, we present a high-level taxonomy of anonymization approaches that have been proposed to foster graph data publishing, and discuss the SOTA approaches in each category.

3.3. High-Level Taxonomy of PPGP Approaches

There exist plenty of graph anonymization approaches in the literature. In Figure 11, we present a detailed taxonomy of PPGP approaches. The taxonomy presented in this paper is more detailed and complete than existing surveys. The rest of this subsection summarizes the major developments in terms of the SOTA approaches in each category.

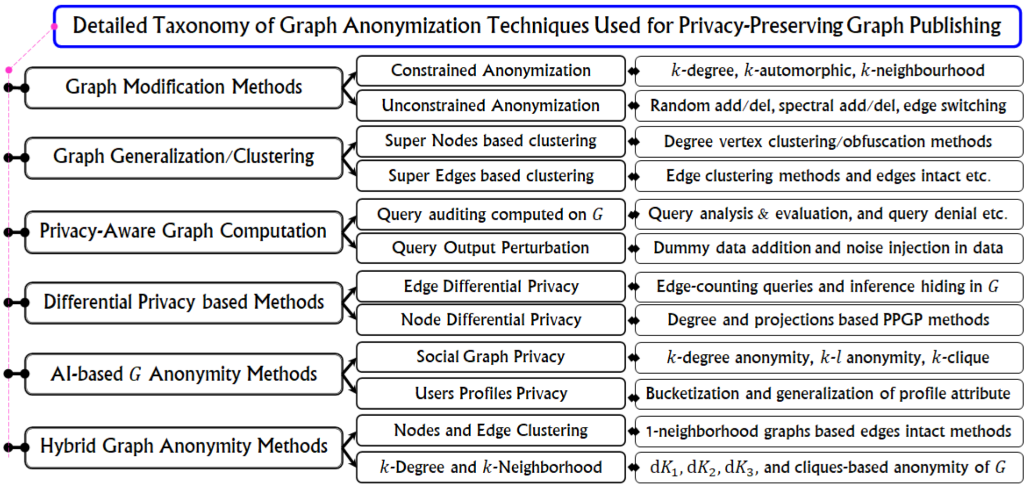

Detailed Taxonomy of Graph Anonymization Techniques Used for Privacy-Preserving Graph Publishing
Graph Modification Methods
Constrained Anonymization
k-degree, k-automorphic, k-neighbourhood
Unconstrained Anonymization
Random add/del, spectral add/del, edge switching
Graph Generalization/Clustering
Super Nodes based clustering
Degree vertex clustering/obfuscation methods
Super Edges based clustering
Edge clustering methods and edges intact etc.
Privacy-Aware Graph Computation
Query auditing computed on G
Query analysis & evaluation, and query denial etc.
Query Output Perturbation
Dummy data addition and noise injection in data
Differential Privacy based Methods
Edge Differential Privacy
Edge-counting queries and inference hiding in G
Node Differential Privacy
Degree and projections based PPGP methods
AI-based G Anonymity Methods
Social Graph Privacy
k-degree anonymity, k-l anonymity, k-clique
Users Profiles Privacy
Bucketization and generalization of profile attribute
Hybrid Graph Anonymity Methods
Nodes and Edge Clustering
1-neighborhood graphs based edges intact methods
k-Degree and k-Neighborhood
dK_1, dK_2, dK_3, and cliques-based anonymity of G

Figure 11. High-level taxonomy of the anonymization techniques applied to G for PPGP.

3.3.1. Graph Modification Methods

The graph modification methods modify the G's structure by deleting/adding nodes or edges to protect the privacy of users (see Figure 10). In addition to the add/delete operations, in some cases, the positions of vertices or edges are switched or re-organized to preserve users' privacy. The graph modification methods are classified into two types: unconstrained and constrained. In the former type, the structure of G is modified without strict criteria. Moreover, the later type stops anonymity when some condition/criteria is being met (e.g., all nodes have achieved the same degree). Both types have been extensively studied in the literature for SN users' privacy preservation. We demonstrate an example of unconstrained (a.k.a. random) anonymization adopted from [63] in Figure 12. In Figure 12b, two edges have been removed (i.e., ($[v_4, v_5]$, $[v_2, v_3]$)) while two new edges have been added (i.e., ($[v_8, v_9]$, $[v_6, v_7]$)) to produce $G^{'}$. In contrast, $G^{'}$, shown in Figure 12c, was obtained by the random edge switch method. In this example, two edges have been switched as: $\{(v_1, v_2) \rightarrow (v_1, v_2), (v_4, v_5) \rightarrow (v_2, v_5)\}$. As shown in Figure 12, there is no specific constraint/condition to be met while converting $G \rightarrow G^{'}$.

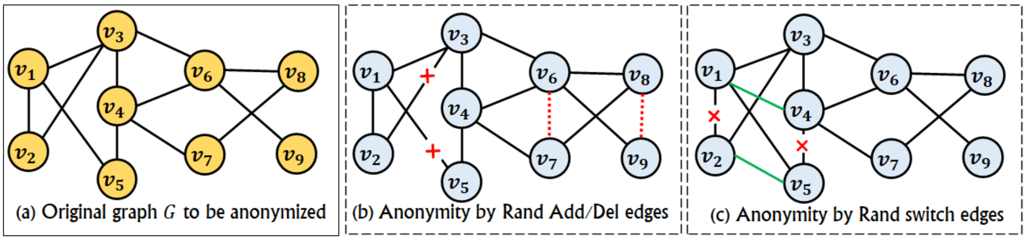

v_1
v_2
v_3
v_4
v_5
v_6
v_7
v_8
v_9
(a) Original graph G to be anonymized
(b) Anonymity by Rand Add/Del edges
(c) Anonymity by Rand switch edges

Figure 12. Example of G anonymization by the unconstrained (a.k.a. random) anonymization method.

The constrained anonymization methods usually follow the same strategy as unconstrained (a.k.a. random) anonymization, but the strength of nodes/edges is bounded by some constraints (e.g., degree, node counts, the number of edges to be switched/modified, etc.). We present an example of constrained anonymization adopted from [63] in Figure 13. The $G^{'}$ shown in Figure 13b is two-degree anonymous (i.e., $k = 2$). The degree sequence was changed from $deg(G) = \{2,4,2,1,3,2,2,2,2\} \rightarrow deg(G^{'})\{2,3,2,2,3,2,2,2,2\}$ by the edge modification/switching method. In this particular example, the constraint was related to the number of edges in the network. The $G^{'}$ shown in Figure 13c is also two-degree anonymous, and it was obtained by applying the new edges and vertex addition. Two new edges $((v_4, v_{10}), (v_5, v_{10}))$, and one node (i.e., v_{10}) were included to convert G into $G^{'}$. In Figure 13c, the modification of G was bounded to the number of both edges and vertices. The degree sequence of each node is as $deg(G^{'} = \{2,4,2,2,4,2,2,2,2,2\})$. In constrained anonymization, the anonymization process stops on the completion of constraints related to closeness, degree, and/or clustering co-efficient, etc.

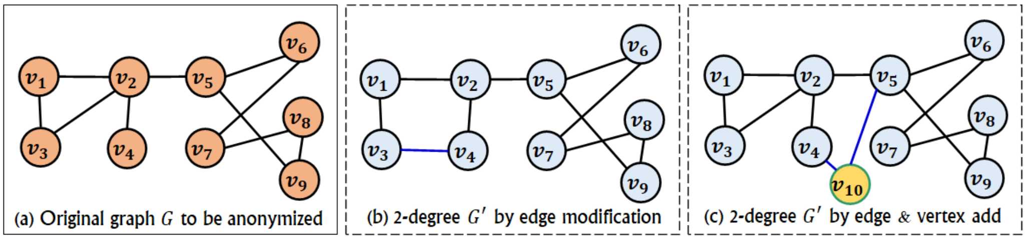

v_1
v_2
v_3
v_4
v_5
v_6
v_7
v_8
v_9
v_{10}
(a) Original graph G to be anonymized
(b) 2-degree G' by edge modification
(c) 2-degree G' by edge & vertex add

Figure 13. Example of G anonymization by the constrained anonymization method.

There are six main modification techniques that can be applied to anonymize SN data stored in a G form, as shown in Figure 14. The selection of modification techniques usually depends on the graph type (e.g., simple, bipartite, labeled, and uncertain graphs.) and objective of the PPGP.

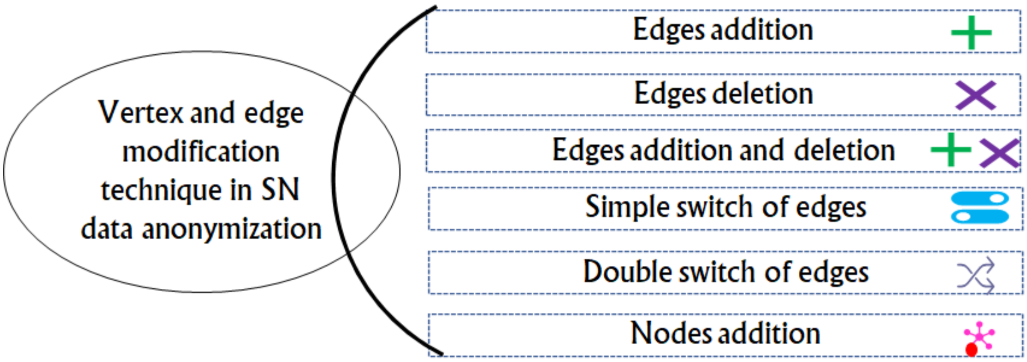

Vertex and edge modification technique in SN data anonymization
Edges addition
Edges deletion
Edges addition and deletion
Simple switch of edges
Double switch of edges
Nodes addition

Figure 14. Overview of graph modification methods. (adopted from [64]).

We summarize and compare the SOTA G modification-based anonymity techniques in Table 3. In Table 3, we compared the existing approaches in terms of assertion(s), study nature, type of datasets on which experiments were performed, and anonymity type. The reason to perform an evaluation based on these metrics is to provide basic as well as experimental analyses that can help researchers to grasp the research status conveniently. Furthermore, these analyses can help researchers to make a rational decision towards conducting high-level research. For example, performance evaluation on a real dataset and writing technical papers are handy takeaways from the below analysis, as most previous studies have been evaluated on real datasets and their nature is technical.

Table 3. Summary and comparison of state-of-the-art graph modification-based techniques.

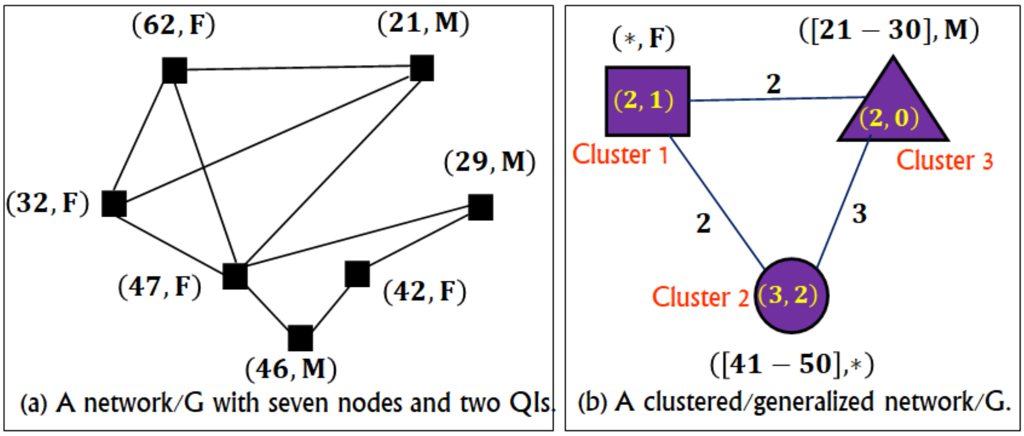
Ref.	Nature of Study	Key Assertion (s)	Experimental Analysis	G Anonymity Type	Datasets Used
Wang et al. [65]	Technical	Defence against identity disclosure	✓	Constrained	R
Casas et al. [66]	Technical	Better utility of large-scale G	✓	Constrained	R
Ma et al. [67]	Technical	Defence against identity disclosure	✓	Constrained	R
Roma et al. [68]	Technical	Defence for identity and link disclosures	✓	Constrained	R
Hamideh et al. [69]	Technical	Fewer modifications while converting $G \rightarrow G^{'}$	✓	Constrained	R,S
Mauw et al. [70]	Technical	Strong defense against identity disclosure	✓	Constrained	R,S
Yuan et al. [71]	Conceptual	Maintains stability of G's structure.	✓(limited)	Constrained	S
Majeed et al. [72]	Technical	Protection of sensitive labels of users	✓	Constrained	R
Gangarde et al. [73]	Technical	Protection of nodes, edges, and attributes	✓	Constrained	R
Srivatsan et al. [74]	Technical	Lower information loss while changing $G \rightarrow G^{'}$	✓	Constrained	R
Nettleton et al. [75]	Technical	Strong privacy protection in $G^{'}$	✓	Constrained	R
Ying et al. [76]	Theoretical	Discussion of various privacy attacks in $G^{'}$	×	Unconstrained	-
Kiabod et al. [77]	Technical	Improve utility of $G^{'}$ for analysts/data miners	✓	Unconstrained	R
Masoumzadeh et al. [78]	Technical	Control distortion while changing $G \rightarrow G^{'}$	✓	Unconstrained	R
Ren et al. [79]	Technical	Protection against three privacy attacks in $G^{'}$	✓	Unconstrained	R
Ninggal et al. [80]	Technical	Significantly improves the utility of anonymized graph	✓	Unconstrained	R
Zhang et al. [81]	Technical	Controls re-identification of users from $G^{'}$	✓	Unconstrained	R,S
Xiang et al. [82]	Technical	Controls privacy issues in dynamic scenarios of G analysis	✓	Unconstrained	R
Zhang et al. [83]	Theoretical	Heuristic analysis-based privacy protection in $G^{'}$	✓	Unconstrained	-
Kavianpour et al. [84]	Technical	Privacy protection in interactions between user and third parties	✓	Unconstrained	R
Lan et al. [85]	Technical	Effective resolution of privacy utility in G anonymization	✓	Unconstrained	R
Hamzehzadeh et al. [86]	Technical	Fewer changes in structure of G during anonymization	✓	Unconstrained	R

Key: ✓⇒ available/reported and ×⇒ not available/not reported, R ⇒ real, S ⇒ synthetic, and - ⇒ not used.

The G modification techniques have been extensively studied in the recent literature. In addition to the analysis presented in Table 3, we refer interested readers to previous surveys for more detailed analyses of the vertex/edge modification techniques [87–90].

3.3.2. Graph Generalization/Clustering Methods

The generalization/clustering-based G anonymization methods perturb the graph structure by partitioning it into different clusters/groups, and anonymity is applied subsequently [91]. The core anonymization concepts of these methods closely resembles the syntactic methods (i.e., k-anonymity, ℓ-diversity, and t-closeness) of tabular data in terms of classes/cluster formation and the generalization of nodes/edges. However, the size of clusters and the degree of generalization is measured in such a way that maximal information is preserved in $G^{'}$ for legitimate information consumers. In Figure 15, we present an example of G's anonymization using graph generalization/clustering methods. In this example, a network G with seven nodes is given as input (see Figure 15a), where each node contains the gender and age information of each user. Afterwards, user are arranged into three clusters based on similarity in gender and age information as follows: $C_1 = \{(62, F), (32, F)\}$, $C_2 = \{(47, F), (46, M), (42, F)\}$, and $C_3 = \{(21, M), (29, M)\}$. Later, all three clusters are generalized to super nodes, as shown in Figure 15b. The two numbers in each super node denotes the number of nodes and intra-cluster edges, respectively. The largest cluster is C_2, with three nodes and two intra-cluster edges. Due to the superior results in both utility and privacy, generalization/clustering-based G anonymization methods have been extensively investigated in the recent literature.



(a) A network/G with seven nodes and two QIs. (b) A clustered/generalized network/G.

Figure 15. Overview of graph generalization/clustering methods.

We summarize and compare the famous generalization/clustering-based G anonymization techniques in Table 4.

Table 4. Summary and comparison of state-of-the-art graph clustering-based techniques.

Ref.	Nature of Study	Key Assertion (s)	Experimental Analysis	G Anonymity Type	Datasets Used
Siddula et al. [92]	Technical	Privacy protection of nodes, edges, and attributes	✓	Clustering	R
Li et al. [93]	Technical	Prevention of inference attacks in SN data	✓	Clustering	R
Gangarde et al. [94]	Technical	Strong defense against users' identities revelation in OSNs	✓	Clustering	R
Karimi et al. [95]	Technical	Privacy preservation of multiple SAs in G publishing	✓	Clustering	R
Jethava et al. [96]	Technical	Strong defense against Sybil attack detection attacks in SN data	✓	Clustering	R
Li et al. [97]	Technical	Strong defense against social identity linkage problem across SNs	✓	Clustering	R, S
Kiranmayi et al. [98]	Technical	Strong defense against attribute couplet attacks via factor analysis	✓	Clustering	R
Kaveri et al. [99]	Conceptual	Privacy and utility preservation in SN data anonymization	×	Clustering	R
Langari et al. [100]	Technical	Defence against identity, attribute, link, and similarity attacks	✓	Clustering	R
Guo et al. [101]	Technical	Privacy and utility preservation in stream data handling	✓	Clustering	R
Sarah et al. [102]	Conceptual	Better utility preservation in G' by extracting maximum clique	✓	Clustering	R, S
Shakeel et al. [103]	Technical	Protection of identity disclosures in SN data publishing	✓	Clustering	R
Poulin et al. [104]	Technical	Protection of privacy and information loss in anonymizing G	✓	Clustering	S
Ghate et al. [105]	Conceptual	Protection of privacy by restricting more changes in G	✓(limited)	Clustering	S
Sihag et al. [106]	Conceptual	Controls heavier changes in the structure of G during anonymity	✓(limited)	Clustering	R
Yu et al. [107]	Technical	Strong defense against identity disclosure by injecting false targets	✓	Clustering	R
Ros et al. [108]	Technical	Strong defense against identity disclosure in large-scale graphs	✓	Clustering	R
Yazdanjue et al. [109]	Technical	Improves runtime of G anonymization by greedy approaches	✓	Clustering	R
Tian et al. [110]	Technical	Ensure strong privacy in crawling and mining SN graph data	✓	Clustering	R

Key: ✓⇒ available/reported and ×⇒ not available/not reported., R ⇒ real and S ⇒ synthetic.

In addition to the analysis given in Table 4, further information about clustering-based anonymization can be gained from previous surveys centering solely on these techniques [111–113]. Recently, clustering-based anonymization methods have gained popularity from multiple perspectives [114].

3.3.3. Privacy-Aware Graph Computing Methods

Privacy-aware graph computing methods do not perturb the structure of G; instead, they compute interesting characteristics that can be helpful from multiple perspectives. These methods share the analysis of the computation rather than the whole G. The privacy-aware graph computing methods extract useful statistics from G in such a way that the privacy of users is preserved while computed statistics remain applicable for SN analysis and mining purposes [115]. The useful analysis provided by privacy-aware graph computing methods are: G density, count of edges, relationships degree, distributions of degrees, size of the G, closeness, centralities, average similarity/distance between users, the number of subgraphs, top k-users with higher connections, clustering coefficients, path length, the number of communities, hypergraphs, the number of users with d degree, where d can be any real number, tie strength among people, trust/influence of people in G, communication/interactions, etc. We highlight an example of degree computation from G in Figure 16. The G shown in Figure 16 has the degree counts as follows: $deg(G) = \{v_1 = 1, v_2 = 3, v_3 = 1, v_4 = 1, v_5 = 2, v_6 = 5, v_7 = 3, v_8 = 3, v_9 = 2, v_{10} = 1\}$. The distribution of degrees can be determined using the following formula: $P_{deg}(degree\ value)$. For instance, the fraction of nodes for degree 1 can be computed as: $P_{deg}(1) = \frac{2}{5}$, as shown in Figure 16. Only graph statistics (i.e., degree) are published, thereby preserving users' privacy. The degree information gives important information concerning G's structure.

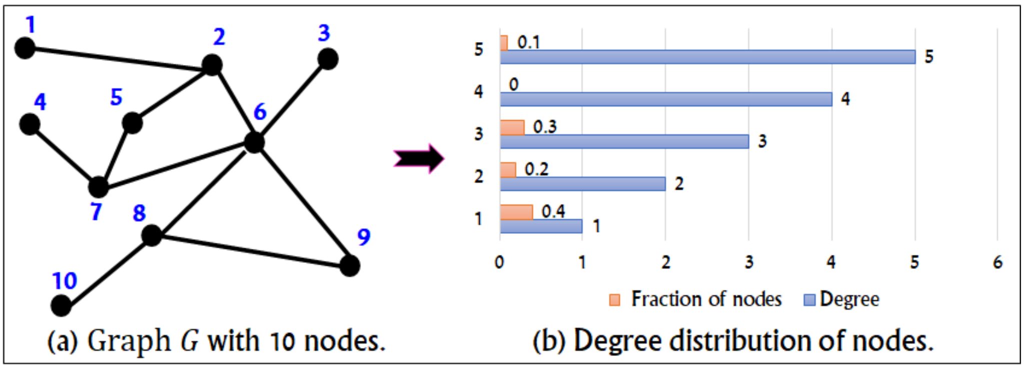


Figure 16. Overview of degree computation from G by privacy-aware graph computing methods.

Instead of degree distributions, many such statistics (i.e., subgraphs [116]) can be computed from G. In Figure 17, we present an example of triangle computations from G. Aside from triangles, the start can also be computed for finding influential people in SNs. These statistics can be used for information diffusion/contagion purposes, marketing, collaborative filtering, opinion/preference mining and analysis, and epidemiological investigations.

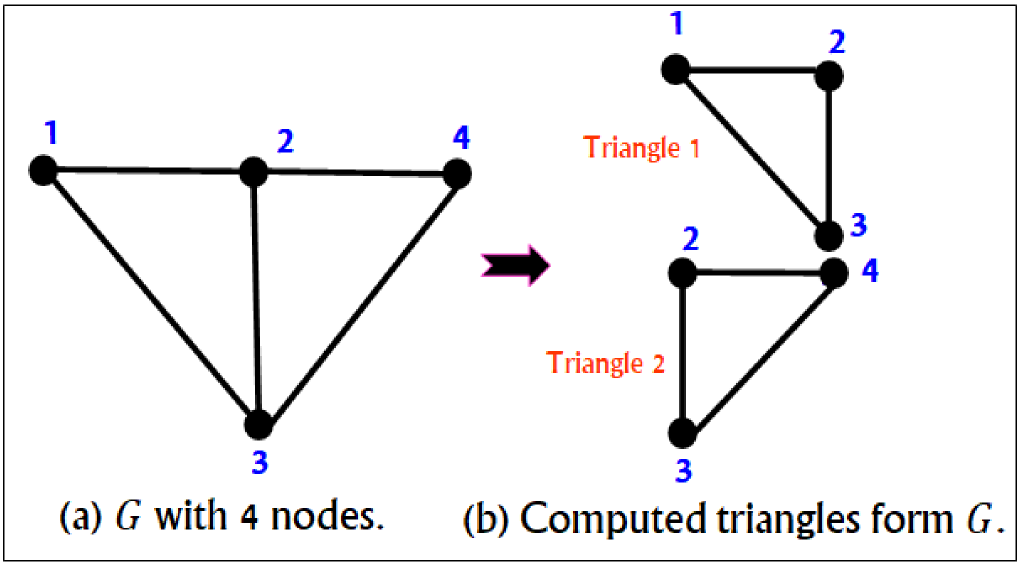


Figure 17. Overview of triangle computations from G by privacy-aware graph computing methods.

The key findings of the latest SOTA privacy-aware graph computing methods are summarized as follows. Shun et al. [117] developed a simple, fast, and in-memory parallel triangles computing algorithm for large-scale SN data. The proposed method requires fewer parameter tunings and is scalable. Yang et al. [118] developed a linear-algebra-based platform for computing multiple statistics from G on GPU platforms. Mazlumi et al. [119] explored the possibility of using SN analysis concepts in the IoT domain regarding path length optimization, critical nodes identification, and advancing IoT applications. Specifically, the authors used SN concepts in the IoT domain for improving multiple aspects of the IoT domain. Behera et al. [120] developed a large clique finding (and missing cliques finding) method from SN graphs. The proposed method has a number of applications such as community detection, pattern recognition, and clustering in SN analysis. Sahraoui et al. [121] applied the SN concepts in the early prevention of the COVID-19 pandemic by detecting contacts in real time. The proposed method detects the communities of people that have likely been exposed to COVID-19 in an analogous way to community detection in SN via analyzing online relationships. Rezvani et al. [122] devised a new and very fast method for detecting communities in SN by using the k-triangle computing method. Laeuchli et al. [123] developed a centrality measurement method in large-scale G. The proposed method has abilities to compute three types of centralities, such as Laplacian, eigenvector, and closeness centralities, from G. A new and low-cost subgraph counting method based on fuzzy set theory for SN data was developed by Hou et al. [124]. Nunez et al. [125] developed a privacy-aware frequent sequential patterns mining method from large-scale G with applications to recommender systems. Further information about privacy-aware graph computing methods can be obtained from the book chapters and reviews in [126,127]. Recently, this category of G privacy has been religiously investigated due to the rapid developments in AI methods and tools.

3.3.4. Differential Privacy-Based Graph Anonymization Methods

Differential privacy (DP) has become a central part of the privacy domain, and it has been extensively investigated in the graph data publishing field [128,129]. DP, in the SN data privacy context, can be defined in simple words as follows. *Let us say a query function f is to be evaluated on a graph G. We want to have a privacy-preserving algorithm A running on G and returning $A(G)$ as an output/answer, and $A(G)$ should be $f(G)$ with a minimal amount of noise added. Hence, the goal of DP is to make $A(G) \approx f(G)$ in order to preserve data utility, and at the same time have $A(G)$ protect all entities' privacy in G.* The DP concept has been widely used in SN for multiple purposes, such as computing statistics from G, answering analyst queries by perturbing the output (see Figure 18), and privacy preservation in application-specific scenarios (i.e., recommendations, community clustering, etc.).

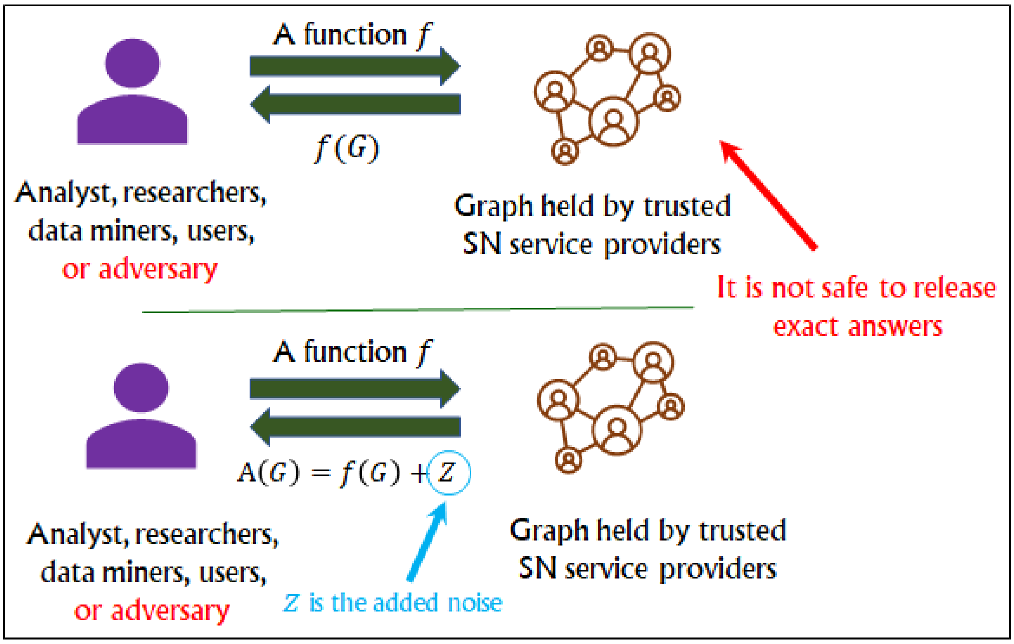

A function f
$f(G)$
Analyst, researchers, data miners, users, or adversary
Graph held by trusted SN service providers
It is not safe to release exact answers
A function f
$A(G) = f(G) + Z$
Analyst, researchers, data miners, users, or adversary
Z is the added noise
Graph held by trusted SN service providers

Figure 18. Overview of query output perturbation in DP model.

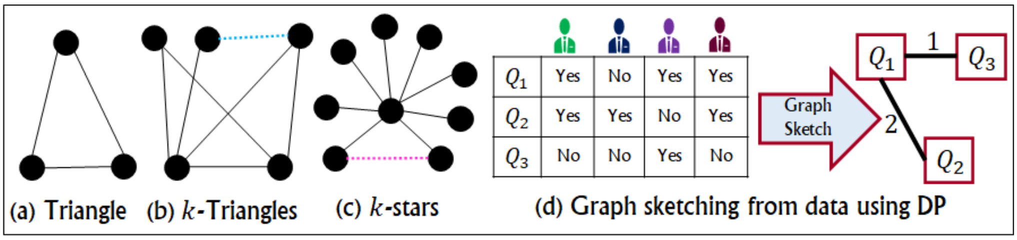
As shown in Figure 18, DP can be achieved by injecting an appropriate amount of noise to the query answer, that is, $A(G) = f(G) + Z$, where Z is the noise. Adding too much Z may damage data utility, while adding too little Z may yield an insufficient privacy guarantee. Therefore, deciding the appropriate value of Z that can strike the balance well between privacy and utility is a very challenging task. Sensitivity, which denotes the largest change to the query answers caused by deleting/adding any node/edge in the G, is a key parameter to find the magnitude of added Z.

In the DP model, any anonymity algorithm, $\Im$, satisfies the ϵ-DP property if for all pairs of neighbors $S \subseteq Range(\Im)$ and for all G_1, G_2 such that $d(G_1, G_2) = 1$ (e.g., G_1 differs from the G_2 by just one node):

$$\frac{Pr(\Im(G_1) \in S)}{Pr(\Im(G_2) \in S)} \leq exp(\epsilon) \tag{2}$$

where ϵ represents the privacy loss budget, and the ϵ value is usually higher than 0 (i.e., $\epsilon > 0$).

If $\epsilon = 0$, full protection can be guaranteed at the expense of utility. Determining an appropriate value for ϵ is very challenging. DP has been extensively used in different settings for fulfilling the expectation of data owners. Furthermore, it has been extended in multiple ways. Its new variants, such as (ϵ,δ)-DP, offer a better trade-off between utility and privacy. DP can be used to compute important statistics from G that can be handy in performing analytics (see Figure 19).



Q_1	Yes	No	Yes	Yes
Q_2	Yes	Yes	No	Yes
Q_3	No	No	Yes	No

Figure 19. Overview of important statistics determined by DP model from G.

There are two types of DP models: local and global [130]. In the former type, noise is added to the personal data before sharing it with the curator, and the server is assumed to be untrustworthy. In the latter type, the original G is curated at some central place (i.e., the server is assumed to be trustworthy), and noise is added at the time of G's release to the analysts/parties. In Figure 20, we present both settings of the DP model in real-world settings.

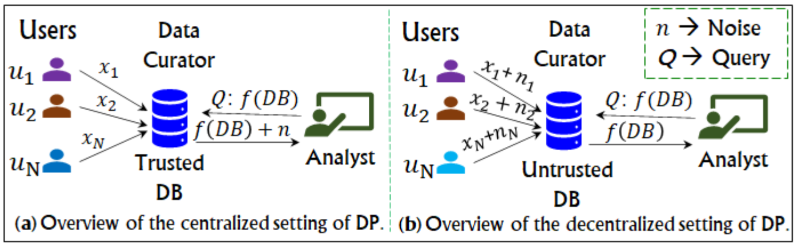


Figure 20. Overview of two famous settings of the DP model.

We summarize and compare the famous DP-based G anonymization techniques in Table 5.

Due to the rigorous privacy guarantees, DP has been widely used with diverse data formats (i.e., tables, graphs, images, texts, matrices, etc.). Detailed information about the DP model in the context of SNs can be learned from DP-specific surveys [163–165]. Recently, DP has been extensively used with emerging technologies, such as federated learning, to preserve privacy [166]. Furthermore, DP has been used to preserve the privacy of patients' COVID-19 data [167]. In the coming years, the synergy and applications of DP are expected to increase drastically.

Table 5. Summary and comparison of state-of-the-art DP-based G anonymity techniques.

Ref.	Nature of Study	Key Assertion (s)	Experimental Analysis	DP Anonymity Type	Datasets Used
Gao et al. [131]	Technical	Better utility in G' under same privacy level	✓	Node-level	R
Gao et al. [132]	Technical	Reduction in noise scale while anonymizing G	✓	Node-level	R,S
Gao et al. [133]	Technical	Protection of important structures of G	✓(limited)	Node-level	R
Gao et al. [134]	Technical	Preserves G's structure using dK-1, dK-2, and dK-3	✓	Node-level	R
Zhang et al. [135]	Technical	Privacy preservation of degree sequence in G'	✓	Node-level	R
Zheng et al. [136]	Technical	Privacy preservation at G collection time in IoTs	✓	Node-level	R
Fang et al. [137]	Technical	Construction of a synthetic G' which is close to G	✓	Node-level	S
Yin et al. [138]	Technical	Graph data publishing with controlled utility loss	✓	Node-level	R
Huang et al. [139]	Technical	Solves privacy–utility trade-off in converting $G \rightarrow G'$	✓	Node-level	R
Macwan et al. [140]	Technical	Protection of degree distributions in answering queries	✓	Node-level	R
Zhu et al. [141]	Technical	Strong privacy in G' by perturbing edges and nodes	✓	Node-level	R
Huang et al.[142]	Conceptual	Privacy preservation by generating synthetic G	✓	Node-level	S
Macwan et al. [143]	Theoretical	Privacy guarantees of G with anonymity and node DP	×	Node-level	-
Macwan et al. [144]	Technical	Preserving higher utility in G' for mining purposes	✓	Node-level	R
Liu et al. [145]	Technical	Preservation of G structural properties without privacy loss	✓	Node-level	R
Iftikhar et al. [146]	Technical	Reduction in noise to achieve ϵ-DP for better utility in G'	✓	Node-level	R
Li et al. [147]	Technical	Privacy preservation of edge weights in G'	✓	Edge-level	R
Guan et al. [148]	Technical	Privacy preservation of link disclosure in G'	✓	Edge-level	R
Wang et al. [149]	Technical	Privacy preservation of links' attributes in G'	✓	Edge-level	R
Yang et al. [150]	Technical	Privacy preservation of degrees of links in G'	✓	Edge-level	R
Wang et al. [151]	Technical	Privacy protection in G' by injecting noise in probability model only	✓	Edge-level	R
Wang et al. [152]	Technical	Preserving topological structure of G' for better utility	✓	Edge-level	R
Lv et al. [153]	Technical	Preserving privacy of users by modifying edge structure in G'	✓	Edge-level	R
Wang et al. [154]	Technical	Preserving privacy of users by dividing G into subgraphs	✓	Edge-level	S
Lei et al. [155]	Technical	Preserving privacy of sensitive edge weights using DP model	✓	Edge-level	R
Reuben et al. [156]	Conceptual	Stresses the need of edges' privacy preservation in G'	✓	Edge-level	-
Yan et al. [157]	Technical	Better utility and privacy preservation in SN data	✓	Hybrid	R
Yan et al. [158]	Technical	Reduces information loss in G anonymity without sacrificing privacy	✓	Hybrid	R
Qian et al. [159]	Technical	Privacy preservation of social links between users via ϵ-DP model	✓	Hybrid	R
Qiuyang et al. [160]	Technical	Privacy preservation based on sub-graph reconstruction and local DP model	✓	Hybrid	R
Qu et al. [161]	Technical	Privacy preservation in dynamically evolving G data with better utility	✓	Hybrid	R
Iftikhar et al. [162]	Technical	Privacy protection in DP-based computations for releasing G's distributions	✓	Hybrid	R

Key: ✓ ⇒ available/reported and × ⇒ not available/not reported, R ⇒ real, S ⇒ synthetic, and - ⇒ not used.

3.3.5. Artificial Intelligence-Based Graph Anonymization Methods

AI has revolutionized almost every discipline with automated decision-making abilities. In the privacy field, AI-based techniques have been widely used to either safeguard or compromise privacy. Recently, AI has been increasingly used in graph data anonymization [110]. Liu et al. [168] presented the link between machine learning and OSN privacy. The authors highlighted the significance of ML in the privacy domain, and vice versa. In Figure 21, we demonstrate the role of AI in OSNs' privacy.

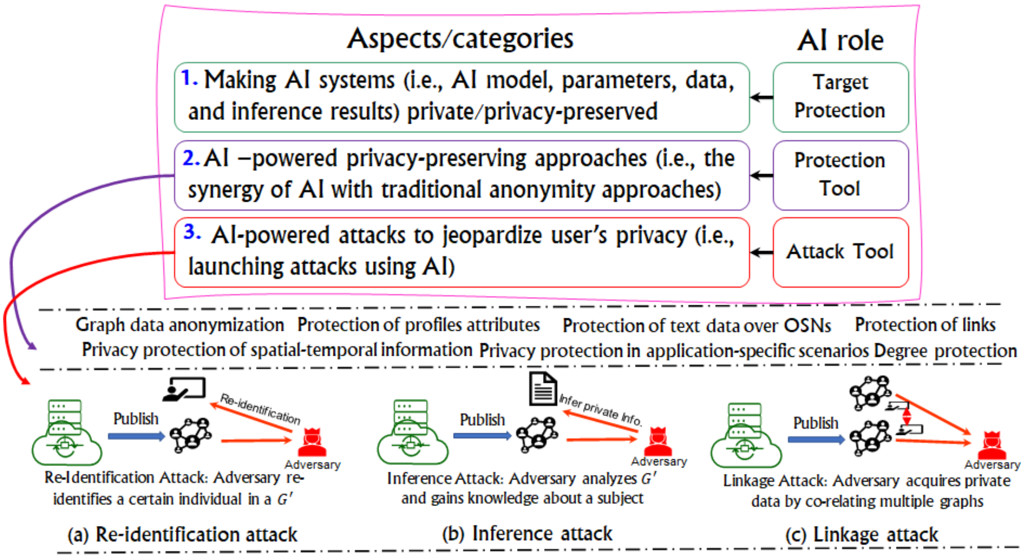

Aspects/categories
AI role
1. Making AI systems (i.e., AI model, parameters, data, and inference results) private/privacy-preserved
Target Protection
2. AI –powered privacy-preserving approaches (i.e., the synergy of AI with traditional anonymity approaches)
Protection Tool
3. AI-powered attacks to jeopardize user's privacy (i.e., launching attacks using AI)
Attack Tool
Graph data anonymization
Protection of profiles attributes
Protection of text data over OSNs
Protection of links
Privacy protection of spatial-temporal information
Privacy protection in application-specific scenarios
Degree protection
Publish
Re-identification
Adversary
Re-Identification Attack: Adversary re-identifies a certain individual in a G'
Publish
Infer private info.
Adversary
Inference Attack: Adversary analyzes G' and gains knowledge about a subject
Publish
Adversary
Linkage Attack: Adversary acquires private data by co-relating multiple graphs
(a) Re-identification attack
(b) Inference attack
(c) Linkage attack

Figure 21. AI role in the privacy (or de-anonymization) of OSNs.

As shown in Figure 21, AI and privacy can complement each other in three different ways. The first category is out of the scope of this paper and is about securing AI systems themselves by using either anonymization or DP-based techniques. This area of research (i.e., securing AI systems) is capturing researchers' interest drastically [169]. The second category is about employing AI methods to safeguard users' privacy in publishing $G^{'}$ [68]. As shown in Figure 21, AI techniques (i.e., machine and deep learning) can assist in preserving OSN users' privacy in multiple ways. The last category is about the dark sides of AI techniques in the information privacy domain. In this category, the adversary takes advantage of the AI techniques in order to predict/infer the private information of individuals from $G^{'}$ [170–172]. In recent years, the synergy between AI and OSNs' privacy has been extensively investigated in the literature [173]. We summarize and compare the famous AI-based G anonymization techniques used for PPGP in Table 6.

Table 6. Summary and comparison of state-of-the-art AI-based G anonymity techniques.

Ref.	Nature of Study	Key Assertion (s)	Experimental Analysis	AI Technique Used	Datasets Used
Bilogrevic et al. [174]	Conceptual	Predicts the level of detail for each sharing decision in OSNs	✓	Logistic Regression	S
Caliskan et al. [175]	Technical	Predicts the SI in a G using ML and suggests how to safeguard it	✓	NB, SVM, RF	S
Yin et al. [176]	Technical	Strikes a balance between privacy and utility in distributing $G^{'}$	✓	k-means algorithm	R
Wang et al. [177]	Technical	Privacy preservation of degree information in releasing $G^{'}$	✓	k-means algorithm	R
Ju et al. [178]	Technical	Strong privacy of V in $G^{'}$ along with higher accuracy and utility	✓	k-means algorithm	R
Zheng et al. [179]	Technical	Strong privacy of V in $G^{'}$ and fewer changes to G's structure	✓	GNN algorithm	R
Paul et al. [180]	Technical	Preserves the structural properties of G in anonymization process	✓	k-means algorithm	R
Hoang et al. [181]	Technical	Preserves the privacy of SN users modelled via knowledge of G	✓	*k-ad* algorithm	R
Hoang et al. [182]	Technical	Preserves the privacy of SN users when G is subject to multiple releases	✓	*CTKGA* algorithm	R
Chen et al. [183]	Technical	Privacy preservation of SN users when G contains outliers and categorical attributes	✓	*DBSCAN* clustering	R
Narula et al. [184]	Technical	Privacy preservation of identity and emotion-related information in OSN data	✓	*CNN* algorithm	R
Zitouni et al. [185]	Technical	Privacy preservation by concealing the identity in image data	✓	CNN and LSTM	R
Ahmed et al. [186]	Technical	Privacy preservation by concealing the identity and other SI in images	✓	Neural Network	R
Matheswaran et al. [187]	Technical	Privacy preservation of image data in retrieval and storage in clouds	✓	Watermarking	R
Li et al. [188]	Technical	Both anonymity- and utility-preserving solutions for OSN data	✓	GAN Algorithm	R
Lu et al. [189]	Technical	Privacy preservation by reducing the prediction accuracy of sensitive links in G	✓	VGAE and ARVGA	R
Li et al. [190]	Technical	Privacy preservation using profile, graph structure, and behavioral information	✓	GCNN algorithm	R
Wanda et al. [191]	Technical	Privacy preservation of vulnerable nodes in G using dynamic deep learning	✓	CNN architecture	R
Li et al. [192]	Technical	Privacy preservation of users when a user's job/education-place changes with time	✓	Supervised ML	R
Bioglio et al. [193]	Technical	Privacy preservation of contents in OSN platforms based on sensitivity analysis	✓	Deep NN	R
Hermansson et al. [194]	Technical	Preserves better accuracy for data-mining and analytical tasks from $G^{'}$	✓	SVM algorithm	R
Kalunge et al. [195]	Technical	Preserves better utility (path length and IL) for data-mining-realted tasks from $G^{'}$	✓	SVM algorithm	R,S
Zhang et al. [196]	Technical	Strong privacy preservation of users against text-based user-linkage attack	✓	SVM algorithm	R
Halimi et al. [197]	Technical	Strong privacy preservation by identifying the vulnerable users profiles from G	✓	PCA algorithm	R
Kumar et al. [198]	Technical	Strong privacy preservation of graph structure without degrading utility of G	✓	PageRank algorithm	R
Kumar et al. [199]	Technical	Strong privacy preservation while showing better utility in three data-mining tasks	✓	PPRA algorithm	R
Li et al. [200]	Technical	Strong privacy preservation of communities in G with better usability of $G^{'}$	✓	Pregel model	R
Chavhan et al. [201]	Technical	Strong privacy preservation in $G^{'}$ by identifying initial cluster centers	✓	DST algorithm	R
Wang et al. [202]	Technical	Strong privacy preservation in $G^{'}$ by dropping the nodes with sparse edges	✓	Kruskal & Prim	S
Kansara et al. [203]	Theoretical	Strong privacy preservation in $G^{'}$ by using non-cryptographic techniques	×	Multiple algorithms	-
Ma et al. [204]	Technical	Strong privacy preservation of user's location while executing queries on G	✓	KNN algorithm	R
Zhang et al. [205]	Technical	Minimization of privacy disclosures in $G^{'}$ and influence maximization	✓	Bayesian Network	R
Mauw et al. [206]	Technical	Strong privacy preservation in $G^{'}$ in the presence of sybil nodes	✓	K-MATCH algorithm	R
Maag et al. [207]	Technical	Strong privacy preservation against multiple attacks in publishing $G^{'}$	✓	EDA algorithm	R
Gao et al. [208]	Technical	Solves multi-objective optimization problem in anonymizing $G^{'}$ by rewiring edges	✓	GAN model	R

Key: ✓ ⇒ available/reported and × ⇒ not available/not reported, R ⇒ real, S ⇒ synthetic, and - ⇒ not used.

The AI-based approaches have improved various critical aspects of OSN data anonymization. In the coming years, AI will be a central element in privacy-preservation solutions of most data styles because AI-based techniques are more robust than traditional anonymization solutions. Further details about AI's role in privacy domains can be learned from the previous surveys in [209–211]. AI-based methods are improving traditional G anonymity methods from multiple perspectives. Although AI has brought a huge revolution in the privacy domain as a defense tool, the computing complexity (CC) of some models can be very high. Shaukat et al. [212] described the CC of many famous machine learning algorithms. In general, the time complexity of any AI model depends on the nature of the data, the input size (e.g., n), the number of iterations/steps (e.g., k), and the parameters (e.g., N). For example, the complexity of simple decision tree is $O(mn^2)$ for tabular data, where n denotes the number of tuples, and m denotes the number of columns. In contrast, the time complexity of a deep belief network (BBN) is $O((n+N)k)$, where n is the number of records, k denotes the iterations, and N is the number of parameters.

3.3.6. Hybrid Graph Anonymization Methods

Hybrid G anonymization methods employ more than one anonymity operation/method while converting G into $G^{'}$. For example, graph modification and clustering methods can be jointly applied to anonymize OSN data enclosed in a G form. Many SOTA hybrid G anonymization methods have been proposed to anonymize OSN data with a better balance of privacy and utility. Liu et al. [213] presented a hybrid anonymization algorithm (e.g., k-anonymity and randomization) for OSN data. The proposed algorithm employs the k-anonymous concept to hide the SI in natural groups/classes of OSN data and uses a randomization approach to process the residual data. The proposed algorithm is more stable and changes the G less than the k-degree anonymity and randomization algorithms. Later, k-anonymity and randomization were jointly used to lower the structural changes in the anonymization of G [214]. Mortazavi et al. [215] used both k-anonymity and ℓ-diversity concepts to anonymize OSN data. The proposed method optimizes the privacy–utility trade-off in PPGP and is more computationally efficient than previous algorithms. Liao et al. [216] used both k-degree anonymity and a genetic algorithm in order to anonymize OSN data enclosed in a G form for recommendation purposes. A hybrid algorithm based on fuzzing SI and converting user's association into an uncertain form was given by Wang et al. [217]. Specifically, the authors define a new attack model in a G and propose an algorithm and safety parameter to safeguard against such attacks.

Qu et al. [218] proposed a hybrid method for a location as well as identity privacy preservation by using a game-based Markov decision process. A new framework that optimizes the utility of $G^{'}$ by employing multiple anonymization techniques was given by Wang et al. [219]. A generic and hybrid anonymization method that guarantees users' privacy and utility in OSN data was proposed by Mortazavi et al. [220]. Similarly, a low-cost G anonymization method based on k-degree anonymity and contractions (i.e., inverse operation, vertex cloning, connectivity, etc.) was proposed by Talmon et al. [221]. A contact G-based approach to anonymize OSN data was proposed by An et al. [222]. The proposed method uses a k-anonymity-based method and contact graphs with location patterns to anonymize G. Although hybrid methods yield better performances in most cases, their complexity is relatively higher than the individual methods. Furthermore, applying the hybrid anonymization method can severely degrade either privacy or utility in some cases. Hence, more efforts are needed to improve the technical aspects of the hybrid anonymity method as well as to determine the correct application scenarios for them.

In summary, all anonymization methods developed for G have their own merits and demerits. For example, graph modification methods expose the G's structure, which can be helpful to analyze the G for recommendation and marketing purposes. In contrast, clustering methods provide better privacy but suppress the G's structure, which may hinder knowledge discovery in $G^{'}$ from all perspectives. Privacy-aware G computing methods ensure the strong privacy of users without degrading the utility. The DP-based methods ensure better privacy even if most parts of G are already exposed to the adversaries. However, utility is the main concern of DP-based anonymization methods. AI-based methods are good at striking the balance between utility and privacy. However, pre-mature convergence and deciding optimal values for hyperparameters are the main challenges in AI-based anonymity methods. Hybrid anonymization methods are computationally expensive and may lead to the redundant usage of some techniques. In the current literature, clustering, DP, and AI-based methods are more popular than others. In the coming years, most studies and enhancements are expected in clustering, DP, and AI-based methods. Furthermore, some recent studies have hinted that hybrid anonymization methods are more useful in safeguarding users' privacy in dynamic settings (e.g., federated learning, collaborative learning, etc.). We compare the methods based on various factors in Table 7. Furthermore, we rate the approaches based on their protection level and future research potential. This analysis can pave the way for choosing the right privacy solutions as well as for exploring the research possibilities of these methods.

Table 7. Comparison of anonymization methods (a.k.a. privacy-preserving solutions) used in OSNs.

Ref.	Privacy and Utility Results Status		Future Research Potentials	Rating in Discipline (OSNs)
	Privacy	**Utility**		
Modification methods	Acceptable	High	Medium	3
Clustering methods	High	Acceptable	High	3
PAGC methods	High	Low	High	4
DP-based methods	High	Low	Very High	4
AI-based methods	High	Acceptable	Very High	4
Hybrid methods	High	High	Very High	5

Abbreviations: PAC (Privacy-aware graph computing); **Rating criteria:** 5: very high and 1: very low.

4. Major Developments in Privacy Preservation in Application-Specific Scenarios of OSNs

With the passage of time, the services of OSNs are expanding in both scale and scope. For example, OSNs enable the formation of communities of like-minded people where people can interact and share their activities/events [223]. OSNs enable information sharing at a much faster pace than any other medium by identifying and delivering information to influential people [224]. They enable friend recommendations by analyzing the demographic, spatial, and interest similarities among users in a seamless manner [225]. Moreover, OSNs enable topic modeling and event detection (i.e., earthquakes, pandemics, floods, etc.) [226]. In the coming years, OSNs are likely to play a key role in assisting mankind in multiple ways. We refer to these services (i.e., community detection, information spread/control, friend recommendations, topic modeling, events detection, etc.) of OSNs as application-specific scenarios of OSNs. We demonstrate an overview of the community detection from G in Figure 22 before presenting privacy-preserving solutions in different application-specific scenarios of OSNs.

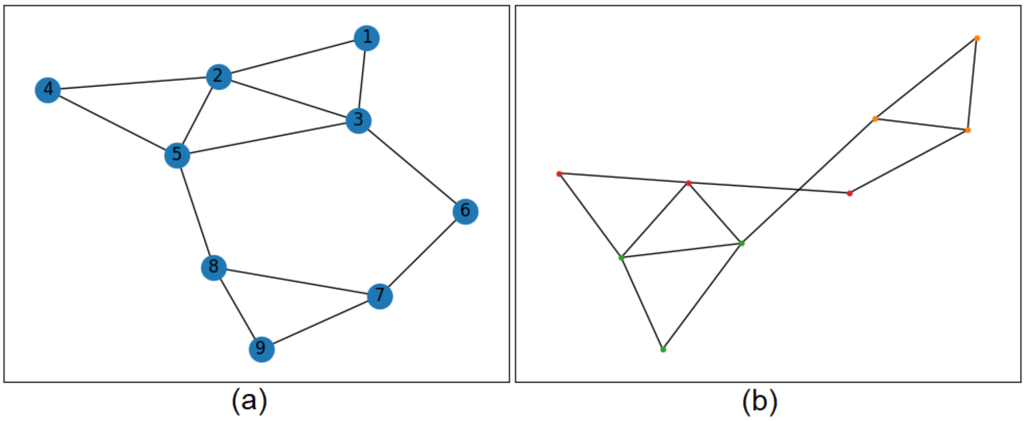


Figure 22. Overview of community detection, (**a**) G with nine nodes, and (**b**) G with three communties.

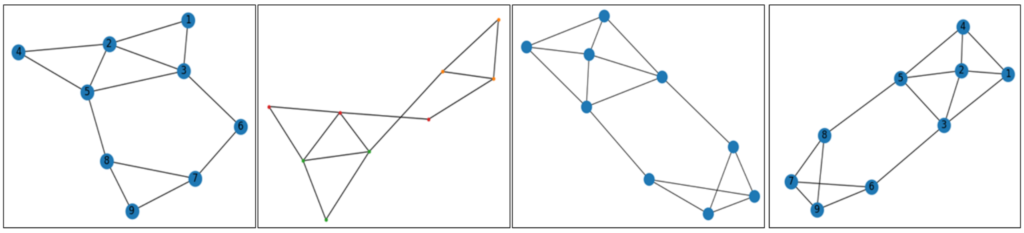
In Table 8, we summarize and compare the SOTA anonymization techniques proposed for privacy preservation in application-specific scenarios of OSNs.

Apart from the famous application-specific scenarios of OSNs listed in Table 8, OSN privacy preservation has been improved by many of the latest techniques, such as federated learning [253]. Therefore, application-specific scenarios of OSNs will be expanded further in the coming years. Furthermore, in some cases, application scenarios of OSNs were used to protect the privacy of OSN users. For example, Rajabzadeh et al. [254] used the community detection concept in a k-degree-based anonymization method in order to preserve the privacy of OSN users (as shown in Figure 23). The proposed method can safeguard users' privacy without degrading the utility of $G^{'}$. Bourahla et al. [255] discussed the method of privacy preservation in dynamic scenarios (i.e., sequential publishing) of OSNs. Further information concerning OSN privacy in application-specific scenarios can be learned from previous studies [256,257]. Lastly, privacy preservation in application-specific scenarios of the OSNs is expected to become an emerging avenue of research in the coming years.

Table 8. SOTA techniques proposed for privacy preservation in OSNs' application-specific scenarios.

Ref.	Nature of Study	Key Assertion (s)	Experimental Analysis	Application Scenario	Datasets Used
Zheng et al. [227]	Technical	Privacy protection of sensitive link information in OSNs	✓	Community detection	R
Wang et al. [228]	Technical	Controls privacy leakage to the application server using ZKPs	✓	Friend recommendations	R
Li et al. [229]	Technical	A secure plugin for privacy preservation of bystanders in OSNs	✓	Content sharing	R
Yi et al. [230]	Technical	Privacy protection of profiles in OSNs using multiple servers and encryptions	✓	Profile matching	R
Wei et al. [231]	Technical	Privacy protection of social content using ϵ-LDP approach	✓	Topic recommendations	R
Valliyammai et al. [232]	Technical	Privacy protection of sensitive topics by detecting sensitive content	✓	Diffusion of sensitive topics	R
Casas et al. [233]	Technical	Privacy protection and utility enhancements of users' data in OSNs	✓	Analytics and mining of G	R
Gao et al. [234]	Technical	Privacy protection in partitioning and mining G for analytical purposes	✓	Subgraph mining from G	R
Li et al. [235]	Technical	Privacy protection of online communities in sensitive content sharing	✓	Content recommendation	R
Mazeh et al. [236]	Technical	Privacy protection of online activity data and purchase histories	✓	Recommender systems	R
Yargic et al. [237]	Technical	Privacy protection of users' sensitive preferences in OSN environments	✓	Collaborative filtering	R
Bahri et al. [238]	Theoretical	Privacy protection of users when OSN data is located in multiple locations	×	Decentralized services	-
Dong et al. [239]	Technical	Social proximity analysis with privacy guarantees identification of potential friends in OSNs	✓	Friend discovery	S
Liu et al. [240]	Technical	Analyzes the risk of community privacy and suggests ways to hide them in OSNs	✓	Hiding community structure	S
Guo et al. [241]	Technical	Quantifies the influence of users based on attributes with privacy preservation in OSNs	✓	Influence estimation	R
Yin et al. [242]	Technical	Privacy preservation in OSNs by analyzing the relationship between pairs of users	✓	Social relationship	R,S
Kukkala et al. [243]	Technical	Designs a privacy-preservation protocol based on secure multi-party computation for OSNs	✓	Influential spreaders	S
Yuan et al. [244]	Technical	Designs a privacy-preservation method for OSNs with restricted changes in the structure of G	✓	Node relationships	R
Gao et al. [245]	Technical	Privacy preservation in OSN data by minimally removing edges/nodes from the original G	✓	Data publishing	R
Zheng et al. [246]	Technical	Privacy preservation in OSNs by controlling higher distortion in G through DP method	✓	Mining and analytics	R
Ferrari et al. [247]	Technical	Privacy preservation in OSN data by clustering and anonymizing people in G	✓	Pattern extraction	R
Aljably et al. [248]	Technical	Privacy preservation of the user information from OSNs utilizing restricted LDP	✓	Anomaly detection	R
Liang et al. [249]	Technical	Privacy preservation of the user action in OSNs via suboptimal estimator	✓	Users action privacy	R
Shan et al. [250]	Technical	Privacy preservation based on user's privacy preferences in OSN environments	✓	Personalized privacy	R
Stokes et al. [251]	Technical	Incidence geometries and clique complexes based privacy preservation of OSN data	✓	Statistical analysis	R
Wen et al. [252]	Technical	Privacy preservation of OSN data by identifying and hiding the vulnerable nodes in G	✓	Recommendation systems	R

Key: ✓ ⇒ available/reported and × ⇒ not available/not reported, R ⇒ real, S ⇒ synthetic, and - ⇒ not used



(**a**) Original graph to be anonymized. (**b**) Detected communities from graph. (**c**) Drawing edges of communities. (**d**) Anonymized graph to be released.

Figure 23. Overview of G anonymization using community detection concept.

5. Major Developments in De-Anonymization of OSNs

The research in OSN privacy is continued in two tracks: defense and attack. The former is concerned with privacy protection from adversaries (a.k.a. anonymization) and the latter is concerned with breaching privacy (a.k.a. de-anonymization). Recently, a substantial number of de-anonymization approaches have been proposed to compromise the privacy of OSN users. The basic goal of the de-anonymization approaches is to re-identify the people uniquely from G' even though a strong anonymization is performed. The de-anonymization is usually performed by exploiting the weaknesses of the anonymity methods, linking G' and auxiliary graphs, and/or background knowledge available to adversary. In Figure 24, we demonstrate an example of how de-anonymity is performed on G'. As shown in Figure 24d, adversaries can exploit the structural information between two graphs, and can infer the identity/SI of OSN users.

Recently, many de-anonymization methods have been proposed, and some methods have accuracies of over 80% in correctly identifying nodes from G' [259]. Recently, due to rapid developments in digitization, the availability of personal information on various OSNs is rising rapidly, leading to a variety of privacy problems [260–265]. These developments indicate the eve-increasing interest of researchers in de-anonymization rather than anonymization. In Table 9, we summarize the findings of various SOTA de-anonymization approaches proposed for OSNs.

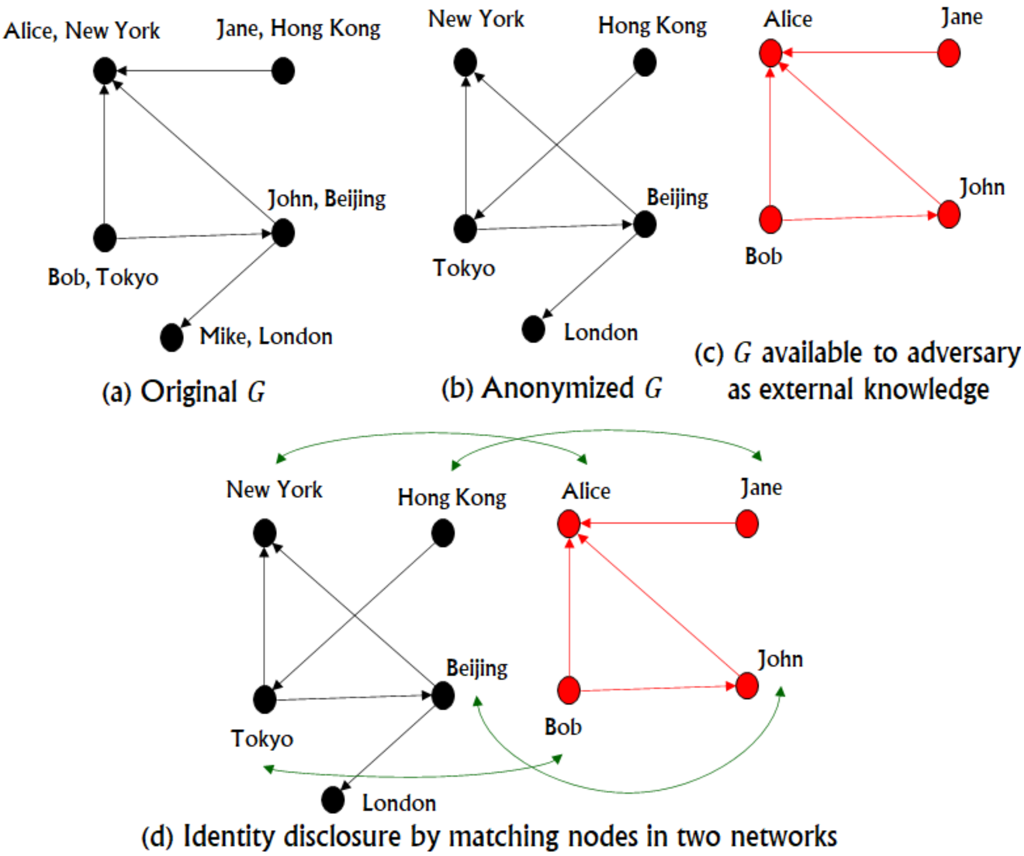


Figure 24. Overview of G de-anonymization using auxiliary graph (adopted from [258]).

Table 9. SOTA de-anonymization approaches proposed for breaching users' privacy in OSNs.

Ref.	Nature of Study	Key Assertion (s)	Privacy Attack	Items Exploited	Datasets Used
Ji et al. [266]	Technical	Privacy disclosure by exploiting attribute and $G^{'}$ structure	Identity disclosure	User's attributes	R
Li et al. [267]	Technical	DNN is adopted to learn features for node matching from $G^{'}$	Identity disclosure	Structure of $G^{'}$	R
Jiang et al. [268]	Technical	Privacy disclosure through structure and attribute similarity	Identity disclosure	Node properties	R
Sun et al. [269]	Technical	Privacy disclosure through spectrum-partitioning method	Identity disclosure	Subgraphs of $G^{'}$	S
Qu et al. [270]	Technical	FBI-based method to extract identities of real-world users	Identity disclosure	Profile, $G^{'}$ structure, & friendship	R
Qu et al. [271]	Technical	RCM-based user matching across OSNs using similarities of concepts	Identity disclosure	Salient features	R
Desai et al. [272]	Technical	Semantic knowledge-based private users' information disclosure	SI disclosure	Background knowledge	R
Hirschprung et al. [273]	Technical	Identification of people through music preference data	Identity disclosure	Music interests	R
Mao et al. [274]	Technical	Identification of people by thoroughly analyzing the structure of $G^{'}$	SI disclosure	Structure of $G^{'}$	R
Qian et al. [275]	Technical	Identification of sensitive data by linking $G^{'}$ & knowledge graphs	SI disclosure	Structure of $G^{'}$	R
Li et al. [276]	Technical	NHDS-based method for revealing sensitive data of the users of OSNs	SI disclosure	$G^{'}$ structure and profiles	R
Feng et al. [277]	Technical	Link privacy breaches in OSNs using three types of similarity metrics	Link prediction	Structure of $G^{'}$	R
Gulyás et al. [278]	Technical	Correct re-identification of a large number of nodes using similarity function	Re-identify nodes	Auxiliary graphs	R
Horawala et al. [279]	Technical	Correct re-identification of a large number of nodes using ML techniques	Re-identify nodes	Node attributes	R
Wu et al. [280]	Technical	Matching a large number of users via overlapped communities concepts	Re-identify nodes	Overlapping Communities	R
Zhou et al. [281]	Technical	Identify multiple accounts of a same person in different OSNs	Identity linkage	Social interactions	R
Chen et al. [282]	Technical	Identify a user by analyzing the social content (i.e., text and images)	Linking users' identities	Social contents	R
Halimi et al. [283]	Technical	Identify a user's profiles with high probability using ML	User's profiles	Auxiliary data	R
Tang et al. [284]	Technical	Matching users to extract SI in different G using embedding vectors	Link prediction	Neighbors' information	R
Zhou et al. [285]	Technical	Correctly linking same users across OSNs using graph neural network	Identity linkage	Node distribution	R
Chen et al. [286]	Technical	Correctly linking the identity of user using semi-supervised method	Identity linkage	Semantic features	R
Wang et al. [287]	Technical	Correctly links a profile of users across multiple OSN platforms	Profile linkage	Duplicate profiles	R

Key: R ⇒ real, S ⇒ synthetic, and - ⇒ not used.

This topic (i.e., graph de-anonymization) has become a mainstream research area in OSN privacy in recent times. Many approaches have been proposed in order to infer identity, SI, membership, and degree information by linking $G^{'}$ and graph data available at external sources. Recently, the use of AI techniques have advanced the de-anonymization area, and many approaches have been proposed for cross-OSN users matching, content-based identity linkage, link prediction, and social connection information disclosure, to name a few. We refer interested readers to learn more about de-anonymizibility from previous surveys focusing solely on privacy attacks in OSNs [288–290]. In the coming years, more developments are expected in the graph de-anonymization area amid the rapid rise in auxiliary information as well as the maturity of AI tools.

6. Challenges of Preserving Privacy in Online Social Networks

The privacy preservation of OSNs is relatively more challenging than the tabular data due to the existence of more information in G data [291]. As stated above, OSN privacy can be compromised in various ways, and therefore, privacy preservation in OSN data is highly challenging. In Figure 25, we present a high-level overview of challenges in OSN privacy preservation.

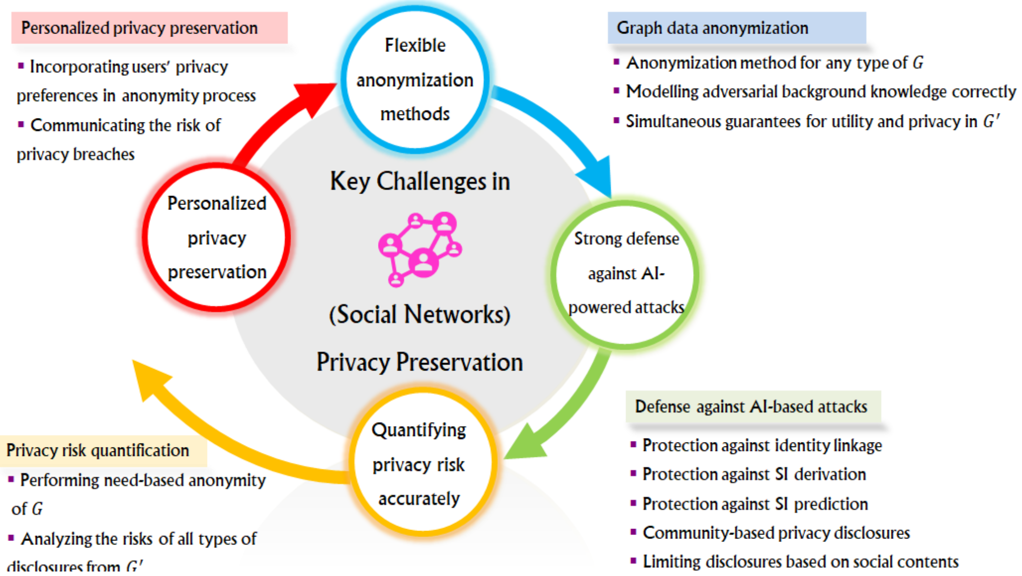
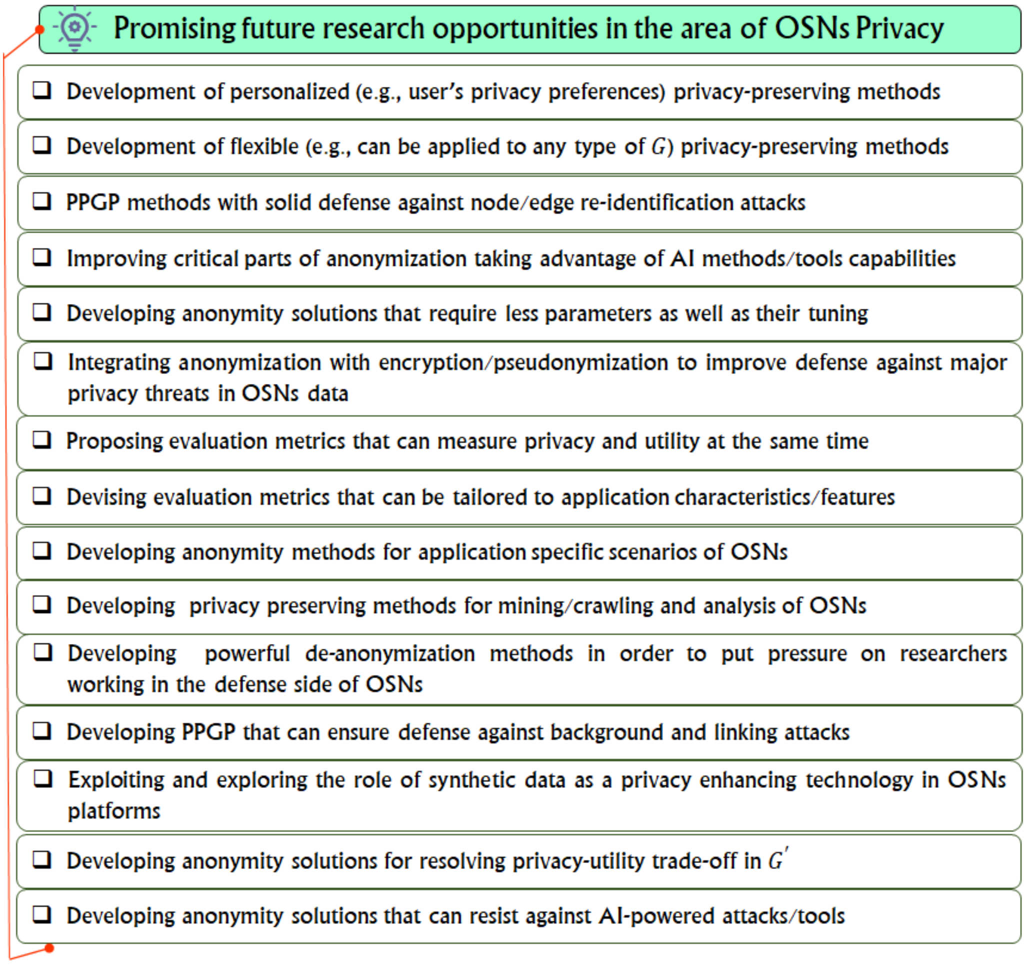

Personalized privacy preservation
▪ Incorporating users' privacy preferences in anonymity process
▪ Communicating the risk of privacy breaches
Flexible anonymization methods
Graph data anonymization
▪ Anonymization method for any type of G
▪ Modelling adversarial background knowledge correctly
▪ Simultaneous guarantees for utility and privacy in G'
Key Challenges in
(Social Networks)
Privacy Preservation
Personalized privacy preservation
Strong defense against AI-powered attacks
Quantifying privacy risk accurately
Defense against AI-based attacks
▪ Protection against identity linkage
▪ Protection against SI derivation
▪ Protection against SI prediction
▪ Community-based privacy disclosures
▪ Limiting disclosures based on social contents
Privacy risk quantification
▪ Performing need-based anonymity of G
▪ Analyzing the risks of all types of disclosures from G'

Figure 25. High-level overview of challenges in OSN privacy preservation.

In Figure 25, we classified the challenges into four categories (i.e., flexible anonymity methods, privacy preservation from AI-powered attacks, incorporating privacy preferences in the design of anonymity methods, and accurately quantifying the privacy and utility levels in $G^{'}$). Apart from these challenges, devising evaluation metrics for G data, evading the power of data mining tools, and resisting the linking of multiple Gs are also very challenging. These challenges can be addressed by devising innovative technologies in the future.

7. Promising Future Research Directions

Researchers are constantly devising new privacy-enhancing techniques for OSNs because the scale and scope of the privacy threats are expanding due to the higher adoption of OSNs across the globe. Privacy preservation of OSNs is more challenging compared to hospitals/banks because a lot of personal data (e.g., user profiles) has already been exposed to adversaries. There are a variety of research tracks in OSNs, for example, privacy preservation in publishing G, de-anonymization of $G^{'}$, metrics for measuring privacy and utility in PPGP, privacy preservation in mining and crawling users' data from OSNs' sites, and privacy preservation using AI tools/methods, to name a few. In the coming years, more practical and robust techniques will be developed in each track cited above. In Figure, 26, , we list promising avenues for future research based on the extensive analysis of the published literature, the developed anonymization tools, the challenges in OSNs privacy, and dedicated surveys. We believe that the list of research opportunities listed in Figure 26 offers a starting point for early researchers in the OSN privacy area. Furthermore, these research gaps require further investigation/research from the research community amid the rapid rise in OSN privacy breaches.



Promising future research opportunities in the area of OSNs Privacy

❑ Development of personalized (e.g., user's privacy preferences) privacy-preserving methods

❑ Development of flexible (e.g., can be applied to any type of G) privacy-preserving methods

❑ PPGP methods with solid defense against node/edge re-identification attacks

❑ Improving critical parts of anonymization taking advantage of AI methods/tools capabilities

❑ Developing anonymity solutions that require less parameters as well as their tuning

❑ Integrating anonymization with encryption/pseudonymization to improve defense against major privacy threats in OSNs data

❑ Proposing evaluation metrics that can measure privacy and utility at the same time

❑ Devising evaluation metrics that can be tailored to application characteristics/features

❑ Developing anonymity methods for application specific scenarios of OSNs

❑ Developing privacy preserving methods for mining/crawling and analysis of OSNs

❑ Developing powerful de-anonymization methods in order to put pressure on researchers working in the defense side of OSNs

❑ Developing PPGP that can ensure defense against background and linking attacks

❑ Exploiting and exploring the role of synthetic data as a privacy enhancing technology in OSNs platforms

❑ Developing anonymity solutions for resolving privacy-utility trade-off in G'

❑ Developing anonymity solutions that can resist against AI-powered attacks/tools

Figure 26. List of promising opportunities for future research in the area of OSN privacy preservation.

The development of privacy-preserving approaches that can incorporate the preferences (e.g., users can decide which item among their attributes is most sensitive and thereby needs stronger privacy, or users can specify how their data should be processed in OSN environments) of users is an important avenue for future research. The development of anonymization methods that can be tuned easily based on the type of graph is an active area of research. Devising privacy-preserving solutions that can ensure defense against well-known and executable privacy attacks (e.g., background, linking, minimality attacks, etc.) is a vibrant area of research. Recently, to optimize privacy guarantees, many AI-based techniques have been integrated with traditional anonymization methods. Therefore, exploring the opportunities of AI techniques in terms of privacy preservation in OSNs is an active area of research. The development of anonymization methods that require fewer parameters and that can be used in resource-constrained environments (e.g., cell phones, gadgets, etc.) is another important research direction. The development of hybrid privacy-preserving solutions (e.g., combining different techniques) that can overcome each other's weaknesses is a vibrant area of research. Developing new metrics that can technically measure the privacy strength from multiple perspectives (i.e., active and passive adversary, across domains, etc.) is an important and active area of research. Developing privacy solutions that can be used in multiple scenarios for privacy protection in OSNs is another potential research direction. Devising methods that can quantify the privacy loss while mining/crawling OSNs data is a prominent area of research. In addition, devising robust de-anonymization methods is a handy direction for the future as it can accelerate development from defense perspectives. Optimizing privacy utility is a longstanding research problem in the privacy domain, and requires technical solutions from the research community. Lastly, developing strong privacy-enhancing techniques to provide resilience against AI-powered attacks/tools is a very hot research area in recent times.

Apart from the research opportunities cited above, exploring the role of the federated learning paradigm in the OSN privacy area is also expected to be a vibrant area of research in the coming years [292]. Recently, a relatively new risk to the individual's privacy, named interdependent privacy (i.e., co-location and location information), is emerging [293,294]. Therefore, advanced privacy-preserving methods are imperative to addressing this emerging risk [295]. Recently, synthetic data-generation methods are also posing a threat to OSN users' privacy by creating data similar to real data [296,297]. Therefore, many privacy-preserving approaches are needed to provide resilience against these threats. Additionally, some anonymization methods that are proposed for other data styles (e.g., tabular, traces, sets, matrices, etc.) can be adopted to preserve the privacy of OSN users in the PPGP. Finally, devising practical methods that can restrict user identity linkages across OSNs is also one of the hot research topics for future endeavours.

8. Limitations of This Review Article

Although this review is more systematic, comprehensive, and insightful than previous reviews, certain limitations exist concerning the number of studies and the coverage domain. For example, we could not find many studies that ensured the anonymity of multimedia data (i.e., images or text written over images) in OSNs, which is one of the hot research topics in recent times. In addition, we could not present any analysis or categorization based on the types of social graphs because the anonymization methods proposed for one type of graph cannot be directly applied to another type of graph (e.g., let us say the PPGP approaches proposed for the directed graph cannot be straightforwardly used for the undirected graph, and vice versa). In addition, this paper does not include studies that have adopted the OSN concept for other services. For example, OSNs' data modeling concepts have been widely used in the COVID-19 arena for infection-spread modeling and analysis. In addition, many OSN concepts have been used for clinical data processing, modeling, and knowledge derivation. Furthermore, we could find some studies that can be simultaneously applied to multiple (i.e., friend recommendations and fried discovery, information spread and contagion, etc.) service scenario(s) in OSNs. This article did not highlight the AI methods in detail (e.g., workflow, parameters, time and space complexity, convergence rates, etc.) but rather focus on AI use in OSNs' privacy preservation (or breaches). Lastly, we mainly focused on recent studies, and we did not consider a span (e.g., the last 5 years, or the last decade) while searching for the studies. However, these limitations do not significantly undermine the quality of this review and can be investigated in future reviews.

9. Conclusions and Future Work

In this paper, we have presented a systematic review of SOTA and recent anonymization techniques that have proposed ways to combat privacy issues in OSNs. Specifically, we have classified the privacy dilemma into two categories: privacy preservation in publishing G and privacy preservation in application-specific scenarios of OSNs. We have presented an extended (i.e., common approaches + AI approaches) taxonomy of anonymization approaches concerning graph data publishing. Moreover, we have presented various representative techniques that are being developed to address privacy issues in the application-specific scenarios (i.e., community clustering, topic modeling, information diffusion, friend recommendations, etc.) of OSNs. We also described various methods that are used to infer identity or private information from published G. Lastly, we discussed various challenges related to OSN privacy and suggested promising opportunities for future research. Through an extensive analysis of the literature, we found that the privacy preservation of OSNs is a very trendy topic among other data styles (e.g., tabular, set, logs, etc.). Many developments are stemming from both anonymization and de-anonymization perspectives. In the coming years, privacy preservation in OSNs will be more challenging, as OSNs are being adopted by an increasingly large number of people across the world. Furthermore, our reliance on

OSNs is also increasing over time, leading to the exposure of more fine-grained data on OSNs sites. In this article, we highlighted the latest SOTA developments concerning the privacy of OSN users. To the best of our knowledge, this is the first work that discusses OSN privacy from a broader perspective, including AI approaches, in the OSN domain. The detailed analysis presented in this article can pave the way for grasping the status of the latest research as well as for developing secure privacy-preserving methods to safeguard OSN users' privacy from prying eyes. Most importantly, our work aligns with the recent trends toward responsible data science (i.e., preventing misuse of personal data). In the future, we intend to explore the role of the latest technologies, such as federated learning, in preserving users' privacy in OSNs. We intend to explore privacy and utility metrics that can be used to quantify the level of privacy and utility offered by anonymization methods in PPGP. Lastly, we intend to explore the role of AI in the privacy domain in heterogeneous data formats (e.g., tables, graphs, matrix, logs, traces, sets, etc.), and multiple computing paradigms such as OSNs, cloud computing, location-based systems, Internet of Things, recommender systems, telemedicine, and AI-based services.

Author Contributions: Conceptualization, A.M. and S.O.H.; methodology, A.M. and S.K.; software, A.M. and S.K.; validation, A.M., S.K. and S.O.H.; formal analysis, A.M.; investigation, A.M.; resources, A.M. and S.O.H.; data curation, A.M.; writing—original draft preparation, A.M.; writing—review and editing, A.M., S.K. and S.O.H.; visualization, A.M. and S.K.; supervision, S.O.H.; project administration, S.O.H.; funding acquisition, A.M. and S.O.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Research Foundation of Korea Grant funded by the Korea Government (MSIT) under Grant 2020R1A2B5B01002145; and in part by the Gachon University under Grant GCU-202206210001.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data and studies that were used to support the findings of this research are included within this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tassa, T.; Dror, J.C. Anonymization of centralized and distributed social network by sequential clustering. *IEEE Trans. Knowl. Data Eng.* **2011**, *25*, 311–324. [CrossRef]
2. Peng, S.; Zhou, Y.; Cao, L.; Yu, S.; Niu, J.; Weijia, J. Influence analysis in social network: A survey. *J. Netw. Comput. Appl.* **2018**, *106*, 17–32. [CrossRef]
3. Safi, S.M.; Movaghar, A.; Ghorbani, M. "Privacy Protection Scheme for Mobile Social Network. *J. King Saud-Univ.-Comput. Inf. Sci.* 2022, *in press*.
4. Nedunchezhian, P.; Mahalingam, M. The Improved Depression Recovery Motivation Recommendation System (I-DRMRS) in Online social network. *Comput. Sci.* **2002**, *3*, 1–17.
5. Dong, Y.; Tang, J.; Wu, S.; Tian, J.; Chawla, N.V.; Rao, J.; Cao, H. Link prediction and recommendation across heterogeneous social networks. In Proceedings of the 2012 IEEE 12th International Conference on Data Mining, Brussels, Belgium, 10–13 December 2012; pp. 181–190.
6. Liu, H.; Zheng, C.; Li, D.; Zhang, Z.; Lin, K.; Shen, X.; Xiong, N.N.; Wang, J. Multi-perspective social recommendation method with graph representation learning. *Neurocomputing* **2022**, *468*, 469–481. [CrossRef]
7. Wang, X.; Liu, Y.; Zhou, X.; Wang, X.; Leng, Z. A Point-of-Interest Recommendation Method Exploiting Sequential, Category and Geographical Influence. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 80. [CrossRef]
8. Suat-Rojas, N.; Gutierrez-Osorio, C.; Pedraza, C. Extraction and Analysis of social network Data to Detect Traffic Accidents. *Information* **2022**, *13*, 26. [CrossRef]
9. Kuikka, V.; Monsivais, D.; Kaski, K.K. Influence spreading model in analysing ego-centric social network. *Phys. Stat. Mech. Its Appl.* **2022**, *588*, 126524. [CrossRef]
10. Liang, F. ; Chen, H.; Lin, K.; Li, J.; Li, Z.; Xue, H.; Shakhov, V.; Liaqat, H.B. "Route recommendation based on temporal–spatial metric. *Comput. Electr. Eng.* **2022**, *97*, 107549. [CrossRef]
11. Alemany, J.; Del Val, E.; García-Fornes, A. A Review of Privacy Decision-making Mechanisms in Online social network *ACM Comput. Surv.* **2023**, *55*, 1–32. [CrossRef]
12. Shejy, G. Data Privacy and Security in social network. In *Principles of Social Networking*; Springer: Singapore, 2021; pp. 387–411.

13. Majeed, A.; Lee, S. Anonymization Techniques for Privacy Preserving Data Publishing: A Comprehensive Survey. *IEEE Access* **2020**, *9*, 8512–8545. [CrossRef]
14. Backstrom, L.; Dwork, C.; Kleinberg, J. Wherefore art thou R3579X? Anonymized social network, hidden patterns, and structural steganography. In Proceedings of the 16th international conference on World Wide Web, Alberta, Canada, 8–12 May 2007; pp. 181–190.
15. Zheleva, E.; Getoor, L. Privacy in social networks: A survey. In *Social Network Data Analytics*; Springer: Boston, MA, USA, 2011; pp. 277–306.
16. Almogbel, R.S.; Alkhalifah, A.A. User Behavior in Social Networks Toward Privacy and Trust: Literature Review. *Int. J. Interact. Mob. Technol.* **2022**, *16*, 38–51. [CrossRef]
17. Avinash, M.; Harini, N. Privacy Preservation Using Anonymity in social network. In Proceedings of the Second International Conference on Sustainable Expert Systems, Lalitpur, Nepal, 17–18 September 2021; Springer: Singapore, 2022; pp. 623–631.
18. Gao, Y.; Yi, L.; Yunchuan, S.; Cai, Z.; Ma, L.; Pustišek, M.; Hu, S. IEEE Access Special Section: Privacy Preservation for Large-Scale User Data in social network *IEEE Access* **2022**, *10*, 4374–4379. [CrossRef]
19. Tahir, H.; Brézillon, P. Contextualization of Personal Data Discovery and Anonymization Tools. In *Intelligent Sustainable Systems*; Springer: Singapore, 2022; pp. 277–285.
20. Ferreira, G.; Alves, A.; Veloso, M.; Bento, C. Identification and Classification of Routine Locations Using Anonymized Mobile Communication Data. *ISPRS Int. J. Geo-Inform.* **2022**, *11*, 228. [CrossRef]
21. Krishnakumar, S.K.; Maheswari, K.M.U. A Comprehensive Review on Data Anonymization Techniques for social network. *Webology* **2022**, *19*.
22. Li, Y.; Tao, X.; Zhang, X.; Wang, M.; Wang, S. Break the Data Barriers While Keeping Privacy: A Graph Differential Privacy Method. *IEEE Internet Things J.* 2022, *Early Access*. [CrossRef]
23. Ji, S.; Li, W.; Mittal, P.; Hu, X.; Beyah, R. SecGraph: A Uniform and Open-source Evaluation System for Graph Data Anonymization and De-anonymization. In Proceedings of the 24th USENIX Security Symposium (USENIX Security 15), Washington, DC, USA, 12–14 August 2015; pp. 303–318.
24. Ni, C.; Li, S.C.; Gope, P.; Min, G. Data Anonymization Evaluation for Big Data and IoT Environment. *Inf. Sci.* 2022, *605*, 381–392. [CrossRef]
25. Zhou, B.; Pei, J.; Luk, W. A brief survey on anonymization techniques for privacy preserving publishing of social network data. *ACM Sigkdd Explor. Newsl.* **2008**, *10*, 12–22. [CrossRef]
26. Wu, X.; Ying, X.; Liu, K.; Chen, L. A Survey of Privacy-Preservation of Graphs and social network. In *Managing and Mining Graph Data*; Springer: Boston, MA, USA, 2010; pp. 421–453. [CrossRef]
27. Praveena, A.; Smys, S. Anonymization in Social Networks: A Survey on the issues of Data Privacy in Social Network Sites. *Int. J. Eng. Comput. Sci.* **2016**, *5*, 15912–15918. [CrossRef]
28. Joshi, P.; Kuo, C.-J. Security and privacy in online social network: A survey. In Proceedings of the 2011 IEEE International Conference on Multimedia and Expo, Barcelona, Spain, 11–15 July 2011; pp. 1–6.
29. Drobyshevskiy, M.; Turdakov, D. Random graph modeling: A survey of the concepts. *ACM Comput. Surv.* **2019**, *52*, 1–36. [CrossRef]
30. Injadat, M.; Salo, F.; Nassif, A.B. Data mining techniques in social media: A survey. *Neurocomputing* **2016**, *214*, 654–670. [CrossRef]
31. Shaukat, K.; Luo, S.; Varadharajan, V.; Hameed, I.A.; Xu, M. A Survey on Machine Learning Techniques for Cyber Security in the Last Decade. *IEEE Access* **2020**, *8*, 222310–222354. [CrossRef]
32. Shaukat, K.; Luo, S.; Varadharajan, V.; Hameed, I.A.; Chen, S.; Liu, D.; Li, J. Performance Comparison and Current Challenges of Using Machine Learning Techniques in Cybersecurity. *Energies* **2020**, *13*, 2509. [CrossRef]
33. Gurses, S.; Diaz, C. Two tales of privacy in online social network. *IEEE Secur. Priv.* **2013**, *11*, 29–37. [CrossRef]
34. Mendes, R.; Vilela, J.P. Privacy-preserving data mining: methods, metrics, and applications. *IEEE Access* **2017**, *5*, 10562–10582. [CrossRef]
35. Cunha, M.; Mendes, R.; Vilela, J.P. A survey of privacy-preserving mechanisms for heterogeneous data types. *Comput. Sci. Rev.* **2021**, *41*, 100403. [CrossRef]
36. Watanabe, C.; Amagasa, T.; Liu, L. Privacy Risks and Countermeasures in Publishing and Mining Social Network Data. In Proceedings of the 7th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), Orlando, FL, USA, 15–18 October 2011; pp. 55–66.
37. Perikos, I.; Michael, L. A Survey on Tie Strength Estimation Methods in Online Social Networks. *ICAART* **2022**, *3*, 484–491.
38. Tian, Y.; Zhang, Z.; Xiong, J.; Chen, L.; Ma, J.; Peng, C. Achieving Graph Clustering Privacy Preservation Based on Structure Entropy in Social IoT. *IEEE Internet Things J.* **2021**, *9*, 2761–2777. [CrossRef]
39. Pham, V.V.H.; Yu, S.; Sood, K.; Cui, L. Privacy issues in social network and analysis: a comprehensive survey. *IET Netw.* **2018**, *7*, 74–84. [CrossRef]
40. Peng, W.; Li, F.; Zou, X.; Wu, J. A Two-Stage Deanonymization Attack against Anonymized social network. *IEEE Trans. Comput.* **2012**, *63*, 290–303. [CrossRef]
41. Chetioui, K.; Bah, B.; Alami, A.O.; Bahnasse, A. Overview of Social Engineering Attacks on social network. *Procedia Comput. Sci.* **2022**, *198*, 656–661. [CrossRef]
42. Villalón-Huerta, A.; Ripoll-Ripoll, I.; Marco-Gisbert, H. A Taxonomy for Threat Actors' Delivery Techniques. *Appl. Sci.* **2022**, *12*, 3929. [CrossRef]

43. Olteanu, A.-M.; Huguenin, K.; Shokri, R.; Humbert, M.; Hubaux, J.-P. Quantifying Interdependent Privacy Risks with Location Data. *IEEE Trans. Mob. Comput*. **2016**, *16*, 829–842. [CrossRef]
44. Biczók, G.; Chia, P.H. Interdependent privacy: Let me share your data. In Proceedings of the International Conference on Financial Cryptography and Data Security, Roseau, The Commonwealth of Dominica, 28 February–3 March 2013; Springer: Berlin/Heidelberg, Germany, 2013; pp. 338–353.
45. Alsarkal, Y.; Zhang, N.; Xu, H. Your privacy is your friend's privacy: Examining interdependent information disclosure on online social network. In Proceedings of the 51st Hawaii International Conference on System Sciences, Hilton Waikoloa Village, HI, USA, 3–6 January 2018.
46. Piao, Y.; Ye, K.; Cui, X. Privacy Inference Attack Against Users in Online Social Networks: A Literature Review. *IEEE Access* **2021**, *9*, 40417–40431. [CrossRef]
47. Sharad, K.; Danezis, G. An automated social graph de-anonymization technique. In Proceedings of the 13th Workshop on Privacy in the Electronic Society, Scottsdale, AZ, USA, 3 November 2014; pp. 47–58.
48. Al Faresi, A.; Alazzawe, A.; Alazzawe, A. Privacy leakage in health social network. *Comput. Intell*. **2013**, *30*, 514–534. [CrossRef]
49. Kharaji, Y.M.; Rizi, F.S.; Khayyambashi, M.R. A new approach for finding cloned profiles in online social network. *arXiv* **2014**, arXiv:1406.7377.
50. Halimi, A.; Ayday, E. Efficient Quantification of Profile Matching Risk in social network *arXiv* **2020**, arXiv:2009.03698.
51. Tai, C.-H.; Yu, P.S.; Yang, D.-N.; Chen, M.-S. Structural Diversity for Resisting Community Identification in Published social network. *IEEE Trans. Knowl. Data Eng*. **2013**, *26*, 235–252. [CrossRef]
52. Nurgaliev, I.; Qu, Q.; Bamakan, S.M.H.; Muzammal, M. Matching user identities across social network with limited profile data. *Front. Comput. Sci.* **2020**, *14*, 1–14. [CrossRef]
53. Jave, U.; Shaukat, K.; Hameed, I.A.; Iqbal, F.; Alam, T.M.; Luo, S. A review of content-based and context-based recommendation systems. *Int. J. Emerg. Technol. Learn.* **2021**, *16*, 274–306. [CrossRef]
54. Shaukat, K.; Shaukat, U. Comment extraction using declarative crowdsourcing (CoEx Deco). In Proceedings of the 2016 International Conference on Computing, Electronic and Electrical Engineering (ICE Cube), Quetta, Pakistan, 11–12 April 2016; pp. 74–78.
55. Shaukat, K.; Hameed, I.A.; Luo, S.; Javed, I.; Iqbal, F.; Faisal, A.; Masood, R. Domain Specific Lexicon Generation through Sentiment Analysis. *Int. J. Emerg. Technol. Learn.* **2020**, *15*, 190–204. [CrossRef]
56. Sattikar; A.A.; Kulkarni, R.V. A role of artificial intelligence techniques in security and privacy issues of social networking. *Int. J. Comput. Sci. Eng. Technol.* **2012**, *2*, 792–806.
57. Chung, K.-C.; Chen, C.-H.; Tsai, H.-H.; Chuang, Y.-H. Social media privacy management strategies: A SEM analysis of user privacy behaviors. *Comput. Commun.* **2021**, *174*, 122–130. [CrossRef]
58. Seshadhri, C.; Pinar, A.; Kolda, T.G. Wedge sampling for computing clustering coefficients and triangle counts on large graphs. *Stat. Anal. Data Mining: Asa Data Sci. J.* **2014**, *7*, 294–307. [CrossRef]
59. Skarkala, M.; Gritzalis, S.; Mitrou, L.; Toivonen, H.; Moen, P. Privacy preservation by k-anonymization of weighted social network. In Proceedings of the 2012 IEEE/ACM International Conference on Advances in social network Analysis and Mining, Istanbul, Turkey, 26–29 August 2012; pp. 423–428.
60. Ding, X.; Wang, C.; Choo, K.-K.R.; Jin, H. A Novel Privacy Preserving Framework for Large Scale Graph Data Publishing. *IEEE Trans. Knowl. Data Eng*. **2019**, *33*, 331–343. [CrossRef]
61. Zhang, H.; Li, X.; Xu, J.; Xu, L. Graph Matching Based Privacy-Preserving Scheme in social network. In Proceedings of the International Symposium on Security and Privacy in Social Network and Big Data, Fuzhou, China, 19–21 November 2021; Springer: Singapore, 2021; pp. 110–118.
62. Salas, J.; Domingo-Ferrer, J. Some basics on privacy techniques, anonymization and their big data challenges. *Math. Comput. Sci.* 2018, *12*, 263–274. [CrossRef]
63. Casas-Roma, J.; Herrera-Joancomartí, J.; Torra, V. A survey of graph-modification techniques for privacy-preserving on Netw. *Artif. Intell. Rev.* **2016**, *47*, 341–366. [CrossRef]
64. Casas-Roma, J. An evaluation of vertex and edge modification techniques for privacy-preserving on graphs. *J. Ambient Intell. Humaniz. Comput*. **2019**, *15*, 1–17. [CrossRef]
65. Wang, Y.; Zheng, B. Preserving privacy in social network against connection fingerprint attacks. In Proceedings of the 31st International Conference on Data Engineering, Seoul, Korea, 25–26 November 2015; pp. 54–65
66. Casas-Roma, J.; Herrera-Joancomartí, J.; Torra, V. k-Degree anonymity and edge selection: Improving data utility in large Netw. *Knowl. Inf. Syst*. **2016**, *50*, 447–474. [CrossRef]
67. Ma, T.; Zhang, Y.; Cao, J.; Shen, J.; Tang, M.; Tian, Y.; Al-Dhelaan, A.; Al-Rodhaan, M. *KDVEM* KDVEM: A *k*-degree anonymity with vertex and edge modification algorithm. *Computing* **2015**, *97*, 1165–1184.
68. Casas-Roma, J.; Salas, J.; Malliaros, F.D.; Vazirgiannis, M. k-Degree anonymity on directed Netw. *Knowl. Inf. Syst*. **2018**, *61*, 1743–1768. [CrossRef]
69. Erfani, H.; Seyedeh; Mortazavi, R. A Novel Graph-modification Technique for User Privacy-preserving on social network. *J. Telecommun. Inf. Technol.* **2019**. [CrossRef]
70. Mauw, S.; Ramírez-Cruz, Y.; Trujillo-Rasua, R. Conditional adjacency anonymity in social graphs under active attacks. *Knowl. Inf. Syst*. **2018**, *61*, 485–511. [CrossRef]

71. Yuan, J.; Ou, Y.; Gu, G. An improved privacy protection method based on k-degree anonymity in social network. In Proceedings of the 2019 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), Dalian, China, 29–31 March 2019; pp. 416–420.
72. Majeed, A.; Lee, S. Attribute susceptibility and entropy based data anonymization to improve users community privacy and utility in publishing data. *Appl. Intell.* **2020**, *50*, 2555–2574. [CrossRef]
73. Gangarde, R.; Sharma, A.; Pawar, A.; Joshi, R.; Gonge, S. Privacy Preservation in Online social network Using Multiple-Graph-Properties-Based Clustering to Ensure k-Anonymity, l-Diversity, and t-Closeness. *Electronics* **2021**, *10*, 2877. [CrossRef]
74. Srivatsan, S.; Maheswari, N. Privacy Preservation in Social Network Data using Evolutionary Model. *Mater. Today Proc.* 2022, *in press*. [CrossRef]
75. Nettleton, D.F.; Salas, J. A data driven anonymization system for information rich online social network graphs. *Expert Syst. Appl.* **2016**, *55*, 87–105. [CrossRef]
76. Ying, X.; Pan, K.; Wu, X.; Guo, L. Comparisons of randomization and k-degree anonymization schemes for privacy preserving social network publishing. In Proceedings of the 3rd Workshop on Social Network Mining and Analysis, Paris, France, 28 June 2009; pp. 1–10.
77. Kiabod, M.; Dehkordi, M.N.; Barekatain, B. A Fast Graph Modification Method for Social Network Anonymization. *Expert Syst. Appl.* **2021**, *180*, 115148. [CrossRef]
78. Masoumzadeh, A.; Joshi, J. Preserving Structural Properties in Edge-Perturbing Anonymization Techniques for social network. *IEEE Trans. Dependable Secur. Comput.* **2012**, *9*, 877–889. [CrossRef]
79. Ren, X.; Jiang, D. A Personalized-Anonymity Model of Social Network for Protecting Privacy. *Wirel. Commun. Mob. Comput*. **2022**, *2022*, 1–11. [CrossRef]
80. Ninggal, H.M.I.; Abawajy, J.H. Utility-aware social network graph anonymization. *J. Netw. Comput. Appl.* **2015**, *56*, 137–148. [CrossRef]
81. Zhang, H.; Lin, L.; Xu, L.; Wang, X. Graph partition based privacy-preserving scheme in social network. *J. Netw. Comput. Appl.* **2021**, *195*, 103214. [CrossRef]
82. Xiangxiang, D.O.N.G.; Ang, G.A.O.; Ying, L.I.A.N.G.; Xiaodi, B.I. Method of Privacy Preserving in Dynamic Social Network Data Publication. *J. Front. Comput. Sci. Technol.* **2019**, *13*, 1441.
83. Zhang, H.; Zhang, X.; Liu, L.; Zhang, J. On Study of Privacy Preserving in Large-scale social network Based on Heuristic Analysis. *J. Phys. Conf. Ser.* **2018**, *1087*, 062002. [CrossRef]
84. Kavianpour, S.; Tamimi, A.; Shanmugam, B. A privacy-preserving model to control social interaction behaviors in social network sites. *J. Inf. Secur. Appl.* **2019**, *49*, 102402. [CrossRef]
85. Lan, L.; Tian, L. Preserving social network privacy using edge vector perturbation. In Proceedings of the International Conference on Information Science and Cloud Computing Companion, Guangzhou, China, 7–8 December 2013; pp. 188–193.
86. Hamzehzadeh, S.; Mazinani, S.M. ANNM: A New Method for Adding Noise Nodes Which are Used Recently in Anonymization Methods in social network. *Wirel. Pers. Commun.* **2019**, *107*, 1995–2017. [CrossRef]
87. Li, Y.; Purcell, M.; Rakotoarivelo, T.; Smith, D.; Ranbaduge, T.; Ng, S.T. Private Graph Data Release: A Survey. *arXiv* **2021**, arXiv:2107.04245
88. Cai, Z.; He, Z.; Guan, X.; Li, Y. Collective data-sanitization for preventing sensitive information inference attacks in social networks. *IEEE Trans. Dependable Secur. Comput.* **2016**, *15*, 577–590. [CrossRef]
89. Siddula, M.; Li, L.; Li, Y. An Empirical Study on the Privacy Preservation of Online social network. *IEEE Access* **2018**, *6*, 19912–19922. [CrossRef]
90. Nguyen, L.B.; Zelinka, I.; Snasel, V.; Nguyen, L.T.; Vo, B. Subgraph mining in a large graph: A review. In *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*; Wiley: Hoboken, NJ, USA, 2022; p. e1454.
91. Mohapatra, D.; Patra, M.R. Anonymization of attributed social graph using anatomy based clustering. *Multimed. Tools Appl.* **2019**, *78*, 25455–25486. [CrossRef]
92. Siddula, M.; Li, Y.; Cheng, X.; Tian, Z.; Cai, Z. Anonymization in Online social network Based on Enhanced Equi-Cardinal Clustering. *IEEE Trans. Comput. Soc. Syst.* **2019**, *6*, 809–820. [CrossRef]
93. Li, K.; Luo, G.; Ye, Y.; Li, W.; Ji, S.; Cai, Z. Adversarial Privacy-Preserving Graph Embedding Against Inference Attack. *IEEE Internet Things J.* **2020**, *8*, 6904–6915. [CrossRef]
94. Gangarde, R.; Sharma, A.; Pawar, A. Clustering Approach to Anonymize Online Social Network Data. In Proceedigs of the International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 7–9 April 2022; pp. 1070–1076.
95. Rizi, A.K.; Dehkordi, M.N.; Bakhsh, N.N. SNI: Supervised Anonymization Technique to Publish social network Having Multiple Sensitive Labels. *Secur. Commun. Netw.* **2019**, *2019*, 1–23. [CrossRef]
96. Jethava, G.; Rao, U.P. A novel trust prediction approach for online social networks based on multifaceted feature similarity. *Clust. Comput.* **2022**, 1–15. [CrossRef]
97. Li, X.; Yang, Y.; Chen, Y.; Niu, X. A Privacy Measurement Framework for Multiple Online social network against Social Identity Linkage. *Appl. Sci.* **2018**, *8*, 1790. [CrossRef]
98. Kiranmayi, M.; Maheswari, N. Reducing Attribute Couplet Attack in social network using Factor Analysis. In Proceedings of the International Conference on Recent Trends in Advance Computing (ICRTAC), Chennai, India, 10–11 September 2018; pp. 212–217.

99. Kaveri, V.V.; Maheswari, V. Cluster based anonymization for privacy preservation in social network data community. *J. Theor. Appl. Inf. Technol.* **2015**, *73*, 269–274.
100. Langari, R.K.; Sardar, S.; Mousavi, S.A.A.; Radfar, R. Combined fuzzy clustering and firefly algorithm for privacy preserving in social network. *Expert Syst. Appl.* **2019**, *141*, 112968. [CrossRef]
101. Guo, K.; Zhang, Q. Fast clustering-based anonymization approaches with time constraints for data streams. *Knowl.-Based Syst.* **2013**, *46*, 95–108. [CrossRef]
102. Sarah, L.-K.A.; Tian, Y.; Al-Rodhaan, M. A Novel (K, X)-isomorphism Method for Protecting Privacy in Weighted social Network. In Proceedings of the 21st Saudi Computer Society National Computer Conference (NCC), Riyadh, Saudi Arabia, 25–26 April 2018; pp. 1–6.
103. Shakeel, S.; Anjum, A.; Asheralieva, A.; Alam, M. k-NDDP: An Efficient Anonymization Model for Social Network Data Release. *Electronics* **2021**, *10*, 2440. [CrossRef]
104. Poulin, J.; Mathina, K. Preserving the privacy on social network by clustering based anonymization. *Int. J. Adv. Res. Comput. Sci. Technol.* **2014**, *2*, 11–14.
105. Ghate, R.B.; Rasika, I. Clustering based Anonymization for privacy preservation. In Proceedings of the International Conference on Pervasive Computing (ICPC), Maharashtra, India, 8–10 January 2015; pp. 1–3.
106. Sihag, V.K. A clustering approach for structural k-anonymity in social network using genetic algorithm. In Proceedings of the CUBE International Information Technology Conference, Pune, India, 3–5 September 2012; pp. 701–706.
107. Yu, F.; Chen, M.; Yu, B.; Li, W.; Ma, L.; Gao, H. Privacy preservation based on clustering perturbation algorithm for social network. *Multimed. Tools Appl.* **2017**, *77*, 11241–11258. [CrossRef]
108. Ros-Martín, M.; Salas, J.; Casas-Roma, J. Scalable non-deterministic clustering-based k-anonymization for rich networks. *Int. J. Inf. Secur.* **2018**, *18*, 219–238. [CrossRef]
109. Yazdanjue, N.; Fathian, M.; Amiri, B. Evolutionary Algorithms For k-Anonymity in social network Based on Clustering Approach. *Comput. J.* **2019**, *63*, 1039–1062. [CrossRef]
110. Tian, H.; Zheng, X.; Zhang, X.; Zeng, D.D. ϵ-k anonymization and adversarial training of graph neural Netw. for privacy preservation in social network. *Electron. Commer. Res. Appl.* **2021**, *50*, 101105. [CrossRef]
111. Kausar, F.; Al Beladi, S.O. A Comparative Analysis of Privacy Preserving Techniques in Online social network. *Trans. Netw. Commun.* **2015**, *3*, 59. [CrossRef]
112. Budiardjo, E.K.; Wibowo, W.C. Privacy preserving data publishing with multiple sensitive attributes based on overlapped slicing. *Information* **2019**, *10*, 362. [CrossRef]
113. Du, J.; Pi, Y. Research on Privacy Protection Technology of Mobile Social Network Based on Data Mining under Big Data. *Secur. Commun. Netw.* **2022**, *2022*, 1–9. [CrossRef]
114. Majeed, A.; Khan, S.; Hwang, S.O. Toward Privacy Preservation Using Clustering Based Anonymization: Recent Advances and Future Research Outlook. *IEEE Access* **2022**, *10*, 53066–53097. [CrossRef]
115. Cuzzocrea, A.; Leung, C.K.; Olawoyin, A.M.; Fadda, E. Supporting privacy-preserving big data analytics on temporal open big data. *Procedia Comput. Sci.* **2022**, *198*, 112–121. [CrossRef]
116. Chen, X.; Lui, J.C.S. Mining graphlet counts in online social network. *ACM Trans. Knowl. Discov. Data* **2018**, *12*, 1–38. [CrossRef]
117. Shun, J.; Tangwongsan, K. Multicore triangle computations without tuning. In Proceedings of the IEEE 31st International Conference on Data Engineering, Seoul, Korea, 13–17 April 2015; pp. 149–160.
118. Yang, C.; Buluç, A.; Owens, J.D. GraphBLAST: A High-Performance Linear Algebra-based Graph Framework on the GPU. *ACM Trans. Math. Softw.* **2022**, *48*, 1–51. [CrossRef]
119. Mazlumi, S.H.H.; Kermani, M.A.M. Investigation the structure of the Internet of things (IoT) patent network using social network analysis. *IEEE Internet Things J.* 2022, *Early Access*. [CrossRef]
120. Behera, B.; Husic, E.; Jain, S.; Roughgarden, T.; Seshadhri, C. FPT algorithms for finding near-cliques in c-closed graphs. In Proceedings of the 13th Innovations in Theoretical Computer Science Conference (ITCS 2022), Schloss Dagstuhl-Leibniz-Zentrum fur Informatik, 2022. Available online: https://drops.dagstuhl.de/opus/volltexte/2022/15613/ (accessed on 5 May 2022).
121. Sahraoui, Y.; Lucia, L.D.; Vegni, A.M.; Kerrache, C.A.; Amadeo, M.; Korichi, A. TraceMe: Real-Time Contact Tracing and Early Prevention of COVID-19 based on Online social network. In Proceedings of the 2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 8–11 Januar 2022; pp. 893–896.
122. Rezvani, M.; Rezvani, M. Truss decomposition using triangle graphs. *Soft Comput.* **2021**, *26*, 55–68. [CrossRef]
123. Laeuchli, J.; Ramírez-Cruz, Y.; Trujillo-Rasua, R. Analysis of centrality measures under differential privacy models. *Appl. Math. Comput*. **2021**, *412*, 126546. [CrossRef]
124. Hou, Y.; Xia, X.; Li, H.; Cui, J.; Mardani, A. Fuzzy Differential Privacy Theory and its Applications in Subgraph Counting. *IEEE Trans. Fuzzy Syst.* **2022**. [CrossRef]
125. Nunez-del-Prado, M; Maehara-Aliaga, Y.; Salas, J.; Alatrista-Salas, H.; Megías, D. A Graph-Based Differentially Private Algorithm for Mining Frequent Sequential Patterns. *Appl. Sci.* **2022**, *12*, 2131. [CrossRef]
126. Risselada, H.; Ochtend, J. Social Network Analysis. In *Handbook of Market Research*; Springer: Berlin/Heidelberg, Germany, 2022; p. 693.
127. Khanam, K.Z.; Srivastava, G.; Mago, V. The homophily principle in social network analysis: A survey. *Multimed. Tools Appl.* **2022**, *932*, 1–44. [CrossRef]
128. Odeyomi, O.T. Differential Privacy in social network Using Multi-Armed Bandit. *IEEE Access* **2022**, *10*, 11817–11829. [CrossRef]

129. Task, C.; Clifton, C. What Should We Protect? Defining Differential Privacy for Social Network Analysis. In *State of the Art Applications of Social Network Analysis*; Springer: Cham, Switzerland, 2014; pp. 139–161.
130. Liu, H.; Peng, C.; Tian, Y.; Long, S.; Tian, F.; Wu, Z. GDP vs. LDP: A Survey from the Perspective of Information-Theoretic Channel. *Entropy* **2022**, *24*, 430. [CrossRef]
131. Gao, T.; Li, F.; Chen, Y.; Zou, X. Preserving local differential privacy in online social network. In *International Conference on Wireless Algorithms, Systems, and Applications*; Springer: Cham, Switzerland, 2017; pp. 393–405.
132. Gao, T.; Li, F.; Chen, Y.; Zou, X. Local differential privately anonymizing online social network under hrg-based model. *IEEE Trans. Comput. Soc. Syst.* **2018**, *5*, 1009–1020. [CrossRef]
133. Gao, T.; Li, F. PHDP: Preserving persistent homology in differentially private graph publications. In Proceedings of the IEEE INFOCOM 2019-IEEE Conference on Computer Communications, Paris, France, 29 April–2 May 2019; pp. 2242–2250.
134. Gao, T.; Li, F. Protecting Social Network with Differential Privacy Under Novel Graph Model. *IEEE Access* **2020**, *8*, 185276–185289. [CrossRef]
135. Zhang, S.; Ni, W.; Fu, N. Differentially private graph publishing with degree distribution preservation. *Comput. Secur.* **2021**, *106*, 102285. [CrossRef]
136. Zheng, X.; Tian, L.; Hui, B.; Liu, X. Distributed and Privacy Preserving Graph Data Collection in Internet of Thing Systems. *IEEE Internet Things J.* **2021**, *9*, 9301–9309. [CrossRef]
137. Fang, J,; Li, A.; Jiang, Q. GDAGAN: An anonymization method for graph data publishing using generative adversarial network. In Proceedings of the 2019 6th International Conference on Information Science and Control Engineering (ICISCE), Penang, Malaysia, 29 November–1 December 2019; pp. 309–313.
138. Yin, Y.; Liao, Q.; Liu, Y.; Xu, R. Structural-Based Graph Publishing under Differential Privacy. In Proceedings of the International Conference on Cognitive Computing, Milan, Italy, 8–13 July 2019; Springer: Cham, Switzerland, 2019; pp. 67–78.
139. Huang, H.; Zhang, D.; Xiao, F.; Wang, K.; Gu, J.; Wang, R. Privacy-preserving approach PBCN in social network with differential privacy. *IEEE Trans. Netw. Serv. Manag.* **2020**, *17*, 931–945. [CrossRef]
140. Macwan, K.R.; Patel, S.J. Node Differential Privacy in Social Graph Degree Publishing. *Procedia Comput. Sci.* **2018**, *143*, 786–793. [CrossRef]
141. Zhu, H.; Zuo, X.; Xie, M. DP-FT: A Differential Privacy Graph Generation with Field Theory for Social Network Data Release. *IEEE Access* **2019**, *7*, 164304–164319. [CrossRef]
142. Huang;, H.; Yang, Y.; Li, Y. ℙ𝕊𝔾: Local Privacy Preserving Synthetic Social Graph Generation. In Proceedings of the International Conference on Collaborative Computing: Networking, Applications and Worksharing, Virtual, 16–18 October 2021; Springer: Cham, Switerland, 2021; pp. 389–404.
143. Macwan, K.; Patel, S. Privacy Preserving Approaches for Online Social Network Data Publishing. In *Handbook of Research on Digital Transformation and Challenges to Data Security and Privacy*; IGI Global: Gerais, Brazil, 2021; pp. 119–132.
144. Macwan, K.; Patel, S. Privacy Preservation Approaches for Social Network Data Publishing. In *Artificial Intelligence for Cyber Security: Methods, Issues and Possible Horizons or Opportunities*; Springer: Cham, Switzerland, 2021; pp. 213–233.
145. Liu, P.; Xu, Y.; Jiang, Q.; Tang, Y.; Guo, Y.; Wang, Li.; Li, X. Local differential privacy for social network publishing. *Neurocomputing* **2020**, *391*, 273–279. [CrossRef]
146. Iftikhar, M.; Wang, Q.; Lin, Y. dk-microaggregation: Anonymizing graphs with differential privacy guarantees. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*; Springer: Cham, Switzerland, 2020; pp. 191–203.
147. Li, X.; Yang, J.; Sun, Z.; Zhang, J. Differential Privacy for Edge Weights in social network. *Secur. Commun. Netw.* **2017**, *2017*, 972. [CrossRef]
148. Guan, Y.; Lu, R.; Zheng, Y.; Zhang, S.; Shao, J.; Wei, G. Achieving Efficient and Privacy-Preserving (,)-Core Query over Bipartite Graphs in Cloud. *IEEE Trans. Dependable Secur. Comput.* **2022**, *974*. [CrossRef]
149. Wang, J.; Li, Z.; Lui, J.C.S.; Sun, M. Topology-theoretic approach to address attribute linkage attacks in differential privacy. *Comput. Secur.* **2022**, *113*, 102552. [CrossRef]
150. Yang, J.; Ma, X.; Bai, X.; Cui, L. Graph publishing with local differential privacy for hierarchical social network. In Proceedings of the 2020 IEEE 10th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, 17–19 July 2020; pp. 123–126.
151. Wang, Y.; Yang, J.; Zhang, J. Differential Privacy for Weighted Network Based on Probability Model. *IEEE Access* **2020**, *8*, 80792–80800. [CrossRef]
152. Wang, Y.; Yang, J.; Zhan, J. Differentially Private Attributed Network Releasing Based on Early Fusion. *Secur. Commun. Netw.* **2021**, *2021*, 983. [CrossRef]
153. Lv, T.; Li, H.; Tang, Z.; Fu, F.; Cao, J.; Zhang, J. Publishing Triangle Counting Histogram in social network Based on Differential Privacy. *Secur. Commun. Netw.* **2021**, *2021*. [CrossRef]
154. Wang, D.; Long, S. Boosting the accuracy of differentially private in weighted social network. *Multimed. Tools Appl.* **2019**, *78*, 34801–34817. [CrossRef]
155. Lei, H.; Li, S.; Wang, H. A weighted social network publishing method based on diffusion wavelets transform and differential privacy. *Multimed. Tools Appl.* **2022**, *81*, 1–18. [CrossRef]
156. Reuben, J. Towards a differential privacy theory for edge-labeled directed graphs. *Sicherheit* **2018**, 273–278.
157. Yan, J.; Tian, Y.; Liu, H.; Zhenqiang, W. Uncertain graph generating approach based on differential privacy for preserving link relationship of social network. *In. J. Secur. Netw.* **2022**, *17*, 28–38. [CrossRef]

158. Yan, J.; Liu, H.; Wu, Z. An Efficient Differential Privacy Method with Wavelet Transform for Edge Weights of social network. *Int. J. Netw. Secur.* **2022**, *24*, 181–192.
159. Qian, Q.; Li, Z.; Zhao, P.; Chen, W.; Yin, H.; Zhao, L. Publishing graph node strength histogram with edge differential privacy. In Proceedings of the International Conference on Database Systems for Advanced Applications, Taipei, Taiwan, 11–14 April 2021; Springer: Cham, Switzerland, 2018; pp. 75–91.
160. Qiuyang, G.; Qilian, N.; Xiangzhao, M.; Zhijiao, Y. Dynamic social privacy protection based on graph mode partition in complex social network. *Pers. Ubiquitous Comput.* **2019**, *23*, 511–519. [CrossRef]
161. Qu, Y.; Gao, L.; Yu, S.; Xiang, Y. Personalized Privacy Protection of IoTs Using GAN-Enhanced Differential Privacy. In *Privacy Preservation in IoT: Machine Learning Approaches*; Springer: Singapore, 2022; pp. 49–76.
162. Iftikhar, M.; Wang, Q.; Li, Y. dK-Personalization: Publishing Network Statistics with Personalized Differential Privacy. In Proceedings of the Advances in Knowledge Discovery and Data Mining: 26th Pacific-Asia Conference, PAKDD 2022, Chengdu, China, 16–19 May 2022; pp. 194–207.
163. Jiang, H.; Pei, J.; Yu, D.; Yu, J.; Gong, B.; Cheng, X. Applications of differential privacy in social network analysis: A survey. *IEEE Trans. Knowl. Data Eng.* 2021, *Early Access*. [CrossRef]
164. Kiranmayi, M.; Maheswari, N. A Review on Privacy Preservation of social network Using Graphs. *J. Appl. Secur. Res.* **2020**, *16*, 190–223. [CrossRef]
165. Hua, J.; Tang, A.; Fang, Y.; Shen, Z.; Zhong, S. Privacy-preserving utility verification of the data published by non-interactive differentially private mechanisms. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 2298–2311. [CrossRef]
166. Tran, K.-D.T.; Huang, Y. FedSGDCOVID: Federated SGD COVID-19 Detection under Local Differential Privacy Using Chest X-ray Images and Symptom Information. *Sensors* **2022**, *22*, 3728. [CrossRef]
167. Jiang, H.; Sarwar, S.M.; Yu, H.; Islam, S.A. Differentially private data publication with multi-level data utility. *High-Confid. Comput.* **2022**, *2*, 100049. [CrossRef]
168. Liu, B.; Ding, M.; Shaham, S.; Rahayu, W.; Farokhi, F.; Lin, Z. When machine learning meets privacy: A survey and outlook. *ACM Comput. Surv.* **2021**, *54*, 1–36. [CrossRef]
169. Cristofaro, D.E. A critical overview of privacy in machine learning. *IEEE Secur. Priv.* **2021**, *19*, 19–27. [CrossRef]
170. Aljably, R.; Tian, Y.; Al-Rodhaan, M. Preserving privacy in multimedia social network using machine learning anomaly detection. *Secur. Commun. Netw.* **2020**, *2020*, 5874935. [CrossRef]
171. Narayanan, A.; Shi, E.; Rubinstein, B.I. Link prediction by de-anonymization: How we won the kaggle social network challenge. In Proceedings of the 2011 International Joint Conference on Neural Network, San Jose, CA, USA, 31 July–5 August 2011; pp. 1825–1834.
172. Qian, J.; Li, Xi.; Zhang, C.; Chen, L.; Jung, T.; Han, J. Social network de-anonymization and privacy inference with knowledge graph model. *IEEE Trans. Dependable Secur. Comput.* **2017**, *16*, 679–692. [CrossRef]
173. Tanuwidjaja, C.H.; Choi, R.; Baek, S.; Kim, K. Privacy-preserving deep learning on machine learning as a service—A comprehensive survey. *IEEE Access* **2020**, *8*, 167425–167447. [CrossRef]
174. Bilogrevic, I.; Huguenin, K.; Agir, B.; Jadliwala, M.; Hubaux, Je. Adaptive information-sharing for privacy-aware mobile social network. In Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Zurich, Switzerland, 8–12 September 2013; pp. 657–666.
175. Islam, C.; Aylin; Walsh, J.; Greenstadt, R. Privacy detective: Detecting private information and collective privacy behavior in a large social network. In Proceedings of the 13th Workshop on Privacy in the Electronic Society, Scottsdale, AZ, USA, 3 November 2014; pp. 35–46.
176. Yin, S.; Liu, J. A K-means Approach for Map-Reduce Model and Social Network Privacy Protection. *J. Inf. Hiding Multim. Signal Process.* **2016**, *7*, 1215–1221.
177. Wang, S.L.; Shih, Ch.; Ting, I.; Hong, T. Degree anonymization for k-shortest-path privacy. In Proceedings of the 2013 IEEE International Conference on Systems, Man, and Cybernetics, Manchester, UK, 13–16 October 2013; pp. 1093–1097.
178. Ju, X.; Zhang, X.; Cheung, W.K. Generating synthetic graphs for large sensitive and correlated social network. In Preceedings of the 2019 IEEE 35th International Conference on Data Engineering Workshops (ICDEW), Macao, China, 8–12 April 2019; pp. 286–293.
179. Zheng, Y.; Wu, J.; Zhang, X.; Chu, X. Graph-DPP: Sampling Diverse Neighboring Nodes via Determinantal Point Process. In Proceedings of the 2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), Virtual Conference, 14–17 December 2020; pp. 540–545.
180. Paul, A.; Suppakitpaisarn, V.; Bafna, M.; Rangan, C.P. Improving accuracy of differentially private kronecker social network via graph clustering. In Proceedings of the 2020 International Symposium on Network, Computers and Communications (ISNCC), Montreal, QC, Canada, 20–22 October 2020.
181. Hoang, A.H.; Carminati, B.; Ferrari, E. Cluster-based anonymization of knowledge graphs. In Proceedings of the International Conference on Applied Cryptography and Network Security, Rome, Italy, 20–23 June 2020; Springer: Cham, Switzerland, 2020; pp. 104-123.
182. Hoang, A.H, Carminati, B.; Ferrari, E. Privacy-Preserving Sequential Publishing of Knowledge Graphs. In Proceedings of the 2021 IEEE 37th International Conference on Data Engineering (ICDE), Athens, Greece, 19–22 April 2021.
183. Chen, Z.G.; Kang, Ho.; Yin, Sh.; Kim, Su. An efficient privacy protection in mobility social network services with novel clustering-based anonymization. *Eurasip J. Wirel. Commun. Netw.* **2016**, *2016*, 1–9. [CrossRef]

184. Narula, V.; Feng, K.; Chaspari, T. Preserving privacy in image-based emotion recognition through user anonymization. In Proceedings of the 2020 International Conference on Multimodal Interaction, Virtual Event, The Netherlands, 25–29 October 2020; pp. 452–460.
185. Zitouni, S.M.; Lee, P.; Lee, U.; Hadjileontiadis, L.; Khandoker, A. Privacy Aware Affective State Recognition from Visual Data. *IEEE Access* **2022**, *10*, 40620–40628. [CrossRef]
186. Ahmed, W.K.; Hasan, M.Z.; Mohammed, N. Image-centric social discovery using neural network under anonymity constraint. In 2017 IEEE International Conference on Cloud Engineering (IC2E), Vancouver, BC, Canada, 4–7 April 2017; pp. 238–244.
187. Matheswaran, P.; Navaneethan, C.; Meenatchi, S.; Ananthi, S.; Janaki, K.; Manjunathan, A. Image Privacy in Social Network Using Invisible Watermarking Techniques. *Ann. Rom. Soc. Cell Biol.* **2021**, *25*, 319–327.
188. Li, A.; Fang, J.; Jiang, Q.; Zhou, B.; Jia, Y. A graph data privacy-preserving method based on generative adversarial Netw. In Proceedings of the International Conference on Web Information Systems Engineering, Melbourne, VIC, Australia, 26–29 October 2020; Springer: Cham, Switzerland, 2020; pp. 227–239.
189. Lu, Y.; Deng, Z.; Gao, Q.; Jing, T. Graph Embedding-Based Sensitive Link Protection in IoT Systems. *Wirel. Commun. Mob. Comput.* **2022**, 2022. [CrossRef]
190. Li, X.; Xin, Y.; Zhao, C.; Yang, Y.; Chen, Y. Graph convolutional Netw. for privacy metrics in online social network. *Appl. Sci.* **2020**, *10*, 1327. [CrossRef]
191. Wanda, P.; Jie, H.J. DeepFriend: Finding abnormal nodes in online social network using dynamic deep learning. *Soc. Netw. Anal. Min.* **2021**, *11*, 1–12. [CrossRef]
192. Li, X.; Xin, Y.; Zhao, C.; Yang, Y.; Luo, S.; Chen, Y. Using user behavior to measure privacy on online social network *IEEE Access* **2020**, *8*, 108387–108401. [CrossRef]
193. Bioglio, L.; Pensa, R.G. Analysis and classification of privacy-sensitive content in social media posts. *Epj Data Sci.* **2022**, *11*, 12. [CrossRef]
194. Hermansson, L.; Kerola, T.; Johansson, F.; Jethava, V.; Dubhashi, D. Entity disambiguation in anonymized graphs using graph kernels. In Proceedings of the 22nd ACM International Conference on Information Knowledge Management, San Francisco, CA, USA, 27 October–1 November 2013; pp. 1037–1046.
195. Kalunge, V.; Deepika, S. Data Mining Techniques for Privacy Preservation in Social Network Sites Using SVM. In *Techno-Societal*; Springer: Cham, Switzerland, 2021; pp. 733–743.
196. Zhang, J.; Sun, J.; Zhang, R.; Zhang, Y.; Hu, X. Privacy-preserving social media data outsourcing. In Proceedings of the IEEE INFOCOM 2018-IEEE Conference on Computer Communications, Honolulu, HI, USA, 15–19 April 2018; pp. 1106–1114.
197. Halimi, A.; Ayday, E. Real-time privacy risk quantification in online social network. In Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Network Analysis and Mining, Virtual Event Netherlands, 8–11 November, 2021; pp. 74–81.
198. Kumar, S.; Kumar, P. Upper approximation based privacy preserving in online social network. *Expert Syst. Appl.* **2017**, *88*, 276–289. [CrossRef]
199. Kumar, S.; Kumar, P. Privacy Preserving in Online social network Using Fuzzy Rewiring. *IEEE Trans. Eng. Manag.* 2021, *Early Access*. [CrossRef]
200. Li, J.; Zhang, X.; Liu, J.; Gao, L.; Zhang, H.; Feng, Y. Large-Scale Social Network Privacy Protection Method for Protecting K-Core. *Int. J. Netw. Secur.* **2021**, *23*, 612–622.
201. Chavhan, K.; Challagidad, P.S. Anonymization Technique For Privacy Preservation In social network. In Proceedings of the 2021 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT), Mysuru, India, 10–11 December 2021; pp. 131–136.
202. Wang, J.; Wan, Z.; Song, J.; Huang, Y.; Lin, Y.; Lin, L. Anonymizing Global Edge Weighted Social Network Graphs. In Proceedings of the International Symposium on Security and Privacy in social network and Big Data, Fuzhou, China, 19-21 November 2021; Springer: Singapore, 2021; pp. 119–130.
203. Kansara, K.; Kadhiwala, B. Non-cryptographic Approaches for Collaborative Social Network Data Publishing-A Survey. In Proceedings of the 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 7–9 October 2020; pp. 348–351.
204. Ma, T.; Jia, J.; Xue, Y.; Tian, Y.; Al-Dhelaan, A.; Al-Rodhaan, M. Protection of location privacy for moving kNN queries in social network. *Appl. Soft Comput.* **2018**, *16*, 525–532. [CrossRef]
205. Zhang, J.; Shi, S.; Weng, C.; Xu, L. Individual Attribute and Cascade Influence Capability-Based Privacy Protection Method in social network. *Secur. Commun. Netw.* **2022**, *2022*. [CrossRef]
206. Mau, S.; Ramírez-Cruz, Y.; Trujillo-Rasua, R. Preventing active re-identification attacks on social graphs via sybil subgraph obfuscation. *Knowl. Inf. Syst.* **2022**, *64*, 1077–1100. [CrossRef]
207. Maag, LM.; Denoyer, L.; Gallinari, P. Graph anonymization using machine learning. In Proceedings of the 2014 IEEE 28th International Conference on Advanced Information Networking and Applications, Victoria, BC, Canada, 13–16 May 2014; pp. 1111–1118.
208. Gao, T.; Li, F. Machine Learning-based Online Social Network Privacy Preservation. In Proceedings of the 2022 ACM on Asia Conference on Computer and Communications Security, Nagasaki, Japan, 30 May–3 June 2022; pp. 467–478.
209. Huynh-The, T.; Pham, Q.V.; Pham, X.Q.; Nguyen, T.T.; Han, Z.; Kim, D.S. Artificial Intelligence for the Metaverse: A Survey. *arXiv* **2022**, arXiv:2202.10336.

210. Mutlu, E.C.; Oghaz, T.; Rajabi, A.; Garibay, I. Review on Learning and Extracting Graph Features for Link Prediction. *Mach. Learn. Knowl. Extr.* **2020**, *2*, 672–704. [CrossRef]
211. Nemec Zlatolas, L.; Hrgarek, L.; Welzer, T.; Hölbl, M. Models of Privacy and Disclosure on Social Networking Sites: A Systematic Literature Review. *Mathematics* **2022**, *10*, 146. [CrossRef]
212. Shauka, K.; Luo, S.; Chen, S.; Liu, D. Cyber threat detection using machine learning techniques: A performance evaluation perspective. In Proceedings of the 2020 International Conference on Cyber Warfare and Security (ICCWS), Norfolk, Virginia, 12–13 March 2020; pp. 1–6.
213. Li, P.; Cui, L.; Li, X. A hybrid algorithm for privacy preserving social network publication. In Proceedings of the International Conference on Advanced Data Mining and Applications, Brisbane, Australia, 28–30 November 2014; Springer: Cham, Switzerland, 2014; pp. 267–278.
214. Liu, P.; Bai, Y.; Wang, L.; Li, X. Partial k-anonymity for privacy-preserving social network data publishing. *Int. J. Softw. Eng. Knowl. Eng.* **2017**, *27*, 71–90. [CrossRef]
215. Mortazavi, R.; Erfani, S.H. GRAM: an efficient (k, l) graph anonymization method. *Expert Syst. Appl.* **2020**, *153*, 113454. [CrossRef]
216. Liao, S.H.; Yang, C.A. Big data analytics of social network marketing and personalized recommendations. *Soc. Netw. Anal. Min.* **2021**, *11*, 1–19. [CrossRef]
217. Wang, L.E.; Li, X. A graph-based multifold model for anonymizing data with attributes of multiple types. *Comput. Secur.* **2018**, *72*, 122–135. [CrossRef]
218. Qu, Y.; Yu, S.; Gao, L.; Zhou, W.; Peng, S. A hybrid privacy protection scheme in cyber-physical social networks. *IEEE Trans. Comput. Soc. Syst.* **2018**, *5*, 773–784. [CrossRef]
219. Wang, Y.; Xie, L.; Zheng, B.; Lee, K.C. High utility k-anonymization for social network publishing. *Knowl. Inf. Syst.* **2014**, *41*, 697–725. [CrossRef]
220. Mortazavi, R.; Erfani, S.H. An effective method for utility preserving social network graph anonymization based on mathematical modeling. *Int. J. Eng.* **2018**, *31*, 1624–1632.
221. Talmon, N.; Hartung, S. The complexity of degree anonymization by graph contractions. *Inf. Comput.* **2017**, *256*, 212–225. [CrossRef]
222. An, S.; Li, Y.; Wang, T.; Jin, Y. Contact Graph Based Anonymization for Geosocial Network Datasets. In Proceedings of the 2018 5th International Conference on Behavioral, Economic, and Socio-Cultural Computing (BESC), Taiwan, China, 12–14 November 2018; pp. 132–137.
223. Naik, D.; Ramesh, D.; Gandomi, A.H.; Gorojanam, N.B. Parallel and distributed paradigms for community detection in social network: A methodological review. *Expert Syst. Appl.* **2022**, *187*, 115956. [CrossRef]
224. Mithagari, A.; Shankarmani, R. Mining Active Influential Nodes for Finding Information Diffusion in social network. In *IoT and Cloud Computing for Societal Good*; Springer: Cham, Switzerland, 2022; pp. 245–255.
225. Huang, M.; Jiang, Q.; Qu, Q.; Chen, L.; Chen, H. Information fusion oriented heterogeneous social network for friend recommendation via community detection. *Appl. Soft Comput.* **2022**, *114*, 108103. [CrossRef]
226. Karimi, S.; Shakery, A.; Verma, R.M. Enhancement of Twitter event detection using news streams. *Nat. Lang. Eng.* **2022**, 1–20. [CrossRef]
227. Zheng, X.; Cai, Z.; Luo, G.; Tian, L.; Bai, X. Privacy-preserved community discovery in online social network. *Future Gener. Comput. Syst.* **2019**, *93*, 1002–1009. [CrossRef]
228. Wang, W.; Wang, S.; Huang, J. Privacy Preservation for Friend-Recommendation Applications. *Secur. Commun. Netw.* **2018**, *2018*, 1265352. [CrossRef]
229. Li, F.; Sun, Z.; Li, A.; Niu, B.; Li, H.; Cao, G. Hideme: Privacy-preserving photo sharing on social network. In Proceedings of the IEEE INFOCOM 2019-IEEE Conference on Computer Communications, Paris, France 29 April–2 May 2019; pp. 154–162.
230. Yi, X.; Bertino, E.; Rao, Fa.; Bouguettaya, A. Practical privacy-preserving user profile matching in social network. In Proceedings of the 2016 32nd International Conference on Data Engineering (ICDE), Helsinki, Finland, 16–20 May 2016; pp. 373–384.
231. Wei, J.; Li, J.; Lin, Y.; Zhang, J. LDP-Based Social Content Protection for Trending Topic Recommendation. *IEEE Internet Things J.* **2020**, *8*, 4353–4372. . [CrossRef]
232. Valliyammai, C.; Bhuvaneswari, A. Semantics-based sensitive topic diffusion detection framework towards privacy aware online social network. *Clust. Comput.* **2019**, *22*, 407–422. [CrossRef]
233. Casas-Roma, J. DUEF-GA: Data utility and privacy evaluation framework for graph anonymization. *Int. J. Inf. Secur.* **2019**, *19*, 465–478. . [CrossRef]
234. Gao, J.R.; Chen, W.; Xu, Ji.; Liu, A.; Li, Zh.; Yin, H.; Zhao, L. An efficient framework for multiple subgraph pattern matching models. *J. Comput. Sci. Technol.* **2019**, *34*, 1185–1202. [CrossRef]
235. Li, D.; Lv, Q.; Shang, L.; Gu, N. Efficient privacy-preserving content recommendation for online social communities. *Neurocomputing* **2017**, *219*, 440–454. [CrossRef]
236. Mazeh, I.; Shmueli, E. A personal data store approach for recommender systems: enhancing privacy without sacrificing accuracy. *Expert Syst. Appl.* **2019**, *139*, 112858. [CrossRef]
237. Yargic, A.; Bilge, A. Privacy-preserving multi-criteria collaborative filtering. *Inf. Process. Manag.* **2019**, *56*, 994–1009. [CrossRef]
238. Bahri, L.; Carminati, B.; Ferrari, E. Decentralized privacy preserving services for Online social network Online Soc. *Netw. Media* **2018**, *6*, 18–25. [CrossRef]

239. Dong, W.; Dave, V.; Qiu, L.; Zhang, Y. Secure Friend Discovery in Mobile Social Network. In Proceedings of the INFOCOM, Shanghai, China, 10–15 April 2011.
240. Liu, Y.; Liu, J.; Zhang, Z.; Zhu, L.; Li, A. Rem: From structural entropy to community structure deception. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 1–11
241. Guo, L.; Zhang, C.; Fang, Y.; Lin, P. A Privacy-Preserving Attribute-Based Reputation System in Online social network. *J. Comput. Sci. Technol*. **2015**, *30*, 578–597. [CrossRef]
242. Yin, D.; Shen, Y.; Liu, C. Attribute Couplet Attacks and Privacy Preservation in social network. *IEEE Access* **2017**, *5*, 25295–25305. [CrossRef]
243. Kukkala, V.B.; Iyengar, S. Identifying Influential Spreaders in a Social Network (While Preserving Privacy). *Proc. Priv. Enhancing Technol.* **2020**, *2020*, 537–557. [CrossRef]
244. Yuan, M.; Chen, L.; Yu, P.S.; Yu, T. Protecting Sensitive Labels in Social Network Data Anonymization. *IEEE Trans. Knowl. Data Eng.* **2011**, *25*, 633–647. [CrossRef]
245. Gao, T.; Li, F. Privacy-preserving sketching for online social network data publication. In Proceedings of the 2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON), Boston, MA, USA, 10–13 June 2019; pp. 1–9.
246. Zheng, X.; Zhang, L.; Li, K.; Zeng, X. Efficient publication of distributed and overlapping graph data under differential privacy. *Tsinghua Sci. Technol.* **2021**, *27*, 235–243. [CrossRef]
247. Ferrari, L.; Rosi, A.; Mamei, M.; Zambonelli, F. Extracting urban patterns from location-based social network. In Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Network, Chicago, IL, USA, 1 November 2011; pp. 9–16.
248. Aljably, R.; Tian, Y.; Al-Rodhaan, M.; Al-Dhelaan, A. Anomaly detection over differential preserved privacy in online social network. *PLoS ONE* **2019**, *14*, e0215856. [CrossRef] [PubMed]
249. Liang, S.; Lam, J.; Lin, H. Secure Estimation with Privacy Protection. *IEEE Trans. Cybern.* **2022**, 1–15. [CrossRef]
250. Shan, F.; Ji, P.; Li, F.; Liu, W. A Smart Access Control Mechanism Based on User Preference in Online social network. In Proceedings of the International Conference on Mobile Multimedia Communications, Virtual Event, 23–25 July 2021; Springer: Cham, Switzerlands, 2021; pp. 577–590.
251. Stokes, K. Cover-up: A probabilistic privacy-preserving graph database model. *J. Ambient Intell. Humaniz. Comput.* **2019**, 1–8. [CrossRef]
252. Wen, G.; Liu, H.; Yan, J.; Wu, Z. A privacy analysis method to anonymous graph based on bayes rule in social network. In Proceedings of the 2018 14th International Conference on Computational Intelligence and Security (CIS), Hangzhou, China, 16–19 November 2018; pp. 469–472.
253. Yin, L.; Feng, J.; Xun, H.; Sun, Z.; Cheng, X. A privacy-preserving federated learning for multiparty data sharing in social IoTs. *IEEE Trans. Netw. Sci. Eng.* **2021**, *8*, 2706–2718. [CrossRef]
254. Rajabzadeh, S.; Shahsafi, P.; Khoramnejadi, M. A graph modification approach for k-anonymity in social network using the genetic algorithm. *Soc. Netw. Anal. Min.* **2020**, *10*, 1–17. [CrossRef]
255. Bourahla, S.; Laurent, M.; Challal, Y. Privacy preservation for social networks sequential publishing. *Comput. Netw.* **2020**, *170*, 107106. [CrossRef]
256. Aiello, L.M.; Ruffo, G. LotusNet: Tunable privacy for distributed online social network services. *Comput. Commun.* **2012**, *35*, 75–88. [CrossRef]
257. Kushwah, V.R.S.; Verma, K. Security and Privacy Challenges for Big Data on Social Media. In *Big Data Analytics in Cognitive Social Media and Literary Texts*; Springer: Singapore, 2021; pp. 267–285.
258. Shao, Y.; Liu, J.; Shi, S.; Zhang, Y.; Cui, B. Fast de-anonymization of social network with structural information. *Data Sci. Eng.* **2019**, *4*, 76–92. [CrossRef]
259. Zhang, C.; Jiang, H.; Wang, Y.; Hu, Q.; Yu, J.; Cheng, X. User identity de-anonymization based on attributes. In Proceedings od the International Conference on Wireless Algorithms, Systems, and Applications; Harbin, China, 23–25 June 2019; Springer: Cham, Switzerland, 2019; pp. 458–469.
260. Fu, L.; Zhang, J.; Wang, S.; Wu, X.; Wang, X.; Chen, G. De-anonymizing social network with overlapping community structure. *IEEE/Acm Trans. Netw.* **2020**, *28*, 360–375. [CrossRef]
261. Jiang, H.; Yu, J.; Cheng, X.; Zhang, C.; Gong, B.; Yu, H. Structure-Attribute-Based Social Network Deanonymization with Spectral Graph Partitioning. *IEEE Trans. Comput. Soc. Syst.* **2021**, *9*, 902–913. [CrossRef]
262. Zhang, J.; Qu, S.; Li, Q.; Kang, H.; Fu, L.; Zhang, H.; Wang, X.; Chen, G. On Social Network De-anonymization with Communities: A Maximum A Posteriori Perspective. *IEEE Trans. Knowl. Data Eng.* 2021, *Early Access*. [CrossRef]
263. Zhang, J.; Fu, L.; Long, H.; Meng, G.; Tang, F.; Wang, X.; Chen, G. Collective De-anonymization of social network with Optional Seeds. *IEEE Trans. Mob. Comput.* 2021, *Early Access*. [CrossRef]
264. Miao, B.; Wang, S.; Fu, L.; Lin, X. De-anonymizability of social network: through the lens of symmetry. In Proceedings of the Twenty-First International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Netw and Mobile Computing, Virtual Event, USA, 11–14 October 2020; pp. 71–80.
265. Creţu, AM.; Monti, F.; Marrone, S.; Dong, X.; Bronstein, M.; de Montjoye, Y. Interaction data are identifiable even across long periods of time. *Nat. Commun.* **2022**, *13*, 1–11. [CrossRef]

266. Ji, S.; Wang, T.; Chen, J.; Li, W.; Mittal, P.; Beyah, R. De-sag: On the de-anonymization of structure-attribute graph data. *IEEE Trans. Dependable Secur. Comput.* **2017**, *16*, 594–607. [CrossRef]
267. Li, K.; Lu, G.; Luo, G.; Cai, Z. Seed-free graph de-anonymiztiation with adversarial learning. In Proceedings of the 29th ACM International Conference on Information Knowledge Management, Virtual Event Ireland, 19–23 October 2020; pp. 745–754.
268. Jian, H.; Yu, J.; Hu, C.; Zhang, C.; Cheng, X. SA framework based de-anonymization of social network. *Procedia Comput. Sci.* **2018**, *129*, 358–363. [CrossRef]
269. Sun, Q.; Yu, J.; Jiang, H.; Chen, Y.; Cheng, X. De-anonymizing Scale-Free social network by Using Spectrum Partitioning Method. *Procedia Comput. Sci.* **2019**, *147*, 441–445. [CrossRef]
270. Qu, Y.; Yu, S.; Zhou, W.; Niu, J. FBI: Friendship learning-based user identification in multiple social network. In Proceedings of the 2018 IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 10–12 December 2018; pp. 1–6.
271. Qu, Y.; Ma, H.; Wu, H.; Zhang, K.; Deng, K. A Multiple Salient Features-Based User Identification across Social Media. *Entropy* **2022**, *24*, 495. [CrossRef] [PubMed]
272. Desai, N.; Das, M.L. DeSAN: De-anonymization against Background Knowledge in social network. In Proceedings of the 2021 12th International Conference on Information and Communication Systems (ICICS), Valencia, Spain, 24– 26 May 2021; pp. 99–105.
273. Hirschprung, R.S.; Leshman, O. Privacy disclosure by de-anonymization using music preferences and selections. *Telemat. Informatics* **2021**, *59*, 101564. [CrossRef]
274. Mao, J.; Tian, W.; Jiang, J.; He, Z.; Zhou, Z.; Liu, J. Understanding structure-based social network de-anonymization techniques via empirical analysis. *Eurasip J. Wirel. Commun. Netw.* **2018**, *2018*, 1–16. [CrossRef]
275. Qian, J.; Li, Xi.; Zhang, C.; Chen, L.De-anonymizing social network and inferring private attributes using knowledge graphs. In Proceedings of the IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications, San Francisco, CA, USA, 10–14 April 2016.
276. Li, H.C.; Q.; Zhu, H.; Ma, D.; Wen, H.; Shen, X.S. Privacy leakage via de-anonymization and aggregation in heterogeneous social network. *IEEE Trans. Dependable Secur. Comput.* **2017**, *17*, 350–362. [CrossRef]
277. Feng, S.; Shen, D.; Nie, T.; Kou, Y.; He, J.; Yu, G. Inferring anchor links based on social network structure. *IEEE Access* **2018**, *6*, 17340–17353. [CrossRef]
278. Gulyás, G.; Simon, B.; Imre, S. An efficient and robust social network de-anonymization attack. In Proceedings of the 2016 ACM on Workshop on Privacy in the Electronic Society, Vienna, Austria, 24–28 October 2016; pp. 1–11.
279. Horawalavithana, S.; Flores, J.A.; Skvoretz, J.; Iamnitchi, A. The risk of node re-identification in labeled social graphs. *Appl. Netw. Sci.* **2019**, *4*, 1–20. [CrossRef]
280. Wu, X.; Hu, Z.; Fu, X.; Fu, L.; Wang, X.; Lu, S. Social network de-anonymization with overlapping communities: Analysis, algorithm and experiments. In Proceedings of the IEEE INFOCOM 2018-IEEE Conference on Computer Communications, Honolulu, HI, USA, 15–19 April 2018; pp. 1151–1159.
281. Zhou, J.; Fan, J. TransLink: User Identity Linkage across Heterogeneous social network via Translating Embeddings. In Proceedings of the IEEE INFOCOM 2019-IEEE Conference on Computer Communications, Paris, France, 29 April–2 May 2019.
282. Chen, X.; Song, X.; Cui, S.; Gan, T.; Cheng, Z.; Nie, L. User identity linkage across social media via attentive time-aware user modeling. *IEEE Trans. Multimed.* **2020**, *23*, 3957–3967. [CrossRef]
283. Halimi, A.; Ayday, E. Profile matching across online social network. In Proceedings of the International Conference on Information and Communications Security, Chongqing, China, 19–21 November 2021; Springer: Cham, Switzerland, 2020; pp. 54–70.
284. Tang, R.; Miao, Z.; Jiang, S.; Chen, X.; Wang, H.; Wang, W. Interlayer Link Prediction in Multiplex social network Based on Multiple Types of Consistency between Embedding Vectors. *IEEE Trans. Cybern.* 2021, *Early Access*. [CrossRef]
285. Zhou, F.; Wen, Z.; Zhong, T.; Trajcevski, G.; Xu, X.; Liu, L. Unsupervised User Identity Linkage via Graph Neural Netw. In Proceedings GLOBECOM 2020-2020 IEEE Global Communications Conference, Taipei, Taiwan, 7–11 December 2020; pp. 1–6.
286. Chen, B.; Chen, X. MAUIL: Multilevel attribute embedding for semisupervised user identity linkage. *Inf. Sci.* **2022**, *593*, 527–545. [CrossRef]
287. Wang, M.; Wang, W.; Chen, W.; Zhao, L. EEUPL: Towards effective and efficient user profile linkage across multiple social platforms. *World Wide Web* **2021**, *24*, 1731–1748. [CrossRef]
288. Jain, K.A.; Sahoo, S.R.; Kaubiyal, J. Online social network security and privacy: Comprehensive review and analysis. *Complex Intell. Syst.* **2021**, *7*, 2157–2177. [CrossRef]
289. Waterval, R. How Information Sharing on Online social network May Allow for Personalized Cyberattacks. Bachelor's Thesis, University of Twente, Enschede, The Netherlands, 2022.
290. Safhi, A.; Adel, A.Z.; Alhibbi, M. Major Security Issue That Facing social network with Its Main Defense Strategies. *Tehnički Glasnik* **2022**, *16*, 205–212. [CrossRef]
291. Tran, H.-Y.; Hu, J. Privacy-preserving big data analytics a comprehensive survey. *J. Parallel Distrib. Comput.* **2019**, *134*, 207–218. [CrossRef]
292. Shen, Y.; Gou, F.; Wu, J. Node Screening Method Based on Federated Learning with IoT in Opportunistic social network. *Mathematics* **2022**, *10*, 1669. [CrossRef]
293. Tawnie, T.C.; Kisalay, B.O. Interdependent privacy. *Orbit J.* **2017**, *1*, 1–14. [CrossRef]
294. Humbert, M.; Trubert, B.; Huguenin, K. A survey on interdependent privacy. *Acm Comput. Surv.* **2019**, *52*, 1–40. [CrossRef]

295. Krishna, T.; Siva Rama, L.; Venkateswara, K.; Siva Prasad, P. Privacy control on location and co-location in interdependent data. In Proceedings of the 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), Vellore, India, 30–31 March 2019.
296. Gao, N.; Xue, H.; Shao, W.; Zhao, S.; Qin, K.K.; Prabowo, A.; Rahaman, M.S.; Salim, F.D. Generative adversarial Netw. For spatio-temporal data: A survey. *ACM Trans. Intell. Syst. Technol.* **2022**, *13*, 1–25.
297. Sosa, J.; Betancourt, B. A latent space model for multilayer network data. *Comput. Stat. Data Anal.* **2022**, *162*, 107432. [CrossRef]