REPORT

Heterogeneous Social Network Graph topology and lifecycle

- project deliverable 4.1

Authors: Barbara Guidi and Andrea Michienzi; Chrysanthi Iakovidou and Symeon Papadopoulos; Kristina Kapanova; Kylänpää Markku and Kuusijärvi Jarkko.

Confidentiality: *Public*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825585.

Heterogeneous Social Network Graph topology and lifecycle	
Project name HELIOS	**Grant agreement #** 825585
Author(s) Barbara Guidi, Andrea Michienzi (UNIPI); Chrysanthi Iakovidou, Symeon Papadopoulos (CERTH); Kristina Kapanova (TCD); Kylänpää Markku, Kuusijärvi Jarkko (VTT).	**Pages** 2+38
Reviewers Carlos Alberto Martin (ATOS), Javier Marín Morales (UPV)	
Keywords Online Social Networks, Decentralized Online Social Networks, Ego Networks, Multilayers Social Networks	**Deliverable identification** D4.1

Summary

This deliverable presents the description of the Heterogeneous Social Graph topology and the lifecycle of the topology. The document describes what the Heterogeneous Social Network Graph is and the local knowledge of the network each node has. The document introduces the local structure of each user, namely Contextual Ego Network, which is a multi-layer network where each layer represents a specific context of the use, and - is implemented by exploiting the ego network social model. The document provides an overview of how a user is identified in the HELIOS social network by introducing the main characteristics of the user profile, and how trust affects the lifecycle of the Contextual Ego Network structure of each user.

Confidentiality	Public

31.7.2019
Written by

Barbara Guidi

Project Coordinator's contact address
Ville Ollikainen, ville.ollikainen@vtt.fi, +358 400 841116

Distribution
HELIOS project partners, subcontractors, the Project Officer and HELIOS web site

Contents

List of Acronyms

Acronym	Description
ABE	Attribute Based Encryption
AI	Artificial Intelligence
API	Application Programming Interface
BOSN	Blockchain-based Online Social Network
CMS	Content Management System
CEN	Contextual Ego Network
FOAF	Friend Of A Friend
DHT	Distributed Hash Table
DOSN	Decentralized Online Social Network
DTN	Delay Tolerant Network
HSG	Heterogeneous Social Graph
IoP	Internet of People
IoT	Internet of Things
MAND	Mobile Ad hoc Network Directory
MOSN	Mobile and Opportunistic Social Network
NGI	Next Generation Internet
OSN	Online Social Network
P2P	Peer-to-Peer
SIOC	Semantically-Interlinked Online Communities
SBD	Steem Dollar
SP	Steem Power
TPS	Trusted Proxy Set

1 Introduction

HELIOS is a decentralized social media platform that addresses the dynamic nature of human communications in three dimensions: contextual, spatial and temporal.

HELIOS will advance the current approaches to social media by introducing novel concepts for social graph creation and management by exploiting trust and transparency. Indeed, HELIOS introduces a novel way to create, maintain and configure personal social graphs by exploiting context social data that are available when the application is running.

The application follows the current approaches to the decentralization of Social Media. This issue has been faced by Decentralized Online Social Networks [1] principally based on P2P solutions, by Mobile Social Networks which represent a different social paradigm, and finally by Blockchain-based Social Media.

In recent years, we have witnessed a significant evolution of the Internet thanks to the diffusion of mobile devices and the smart environments. The HELIOS project will develop a decentralized social media platform by considering the Internet of Things (IoT) and the usage of mobile devices.

The rest of the document describes the motivations, the Heterogeneous Social Network Graph, and the Contextual Ego Network. All these concepts have been partially presented in [2].

1.1 HELIOS motivations

An important step following the diffusion of the IoT and mobile devices is the recently announced Next Generation Internet (NGI) initiative[1], launched by the European Commission during fall 2016, with the aim to rethink Internet as an interoperable platform ecosystem.

This initiative aims to lead the technological evolution and ensure that it will bring novel concepts in the systems, making them more human-centric, human-friendly and enabling human potential.

Around the NGI, new paradigms have been born, like the one of the Internet of People (IoP) [3], an Internet data and knowledge management paradigm that embeds human behaviour in its algorithms. IoP has three leading principles: adopt a user centric approach, consider personal devices, and employ human behaviour models. In detail, users are put at the centre and the service is built around them according to their needs, contrary to a traditional service-oriented approach. The paradigm considers that personal devices of users will become alter-egos of the respective owners over Internet, requiring them to make decisions according to their owners' preferences and needs. Finally, the paradigm foresees that human behaviour models will be incorporated into devices. This will help devices make decisions that match those that would have been made by human users.

The main research fields involved in this revolution are: Internet of Things (IoT), Artificial Intelligence (AI), Data Science, Blockchain, and Social Networking and Media.

[1] https://www.ngi.eu

Online Social Networks (OSNs) will be one of the focal points of this initiative as many aspects are encompassed within it. HELIOS addresses this issue by providing a new concept of Decentralized Online Social Networks. Data storage techniques need to be revised to ensure better privacy levels for users, information diffusion must be rethought to take into account a world of IoT and mobile devices, even how to define the privacy and trust between users should be reviewed considering the human at the centre and the service around them.

Formally, an OSN is defined in [4] as an online platform that provides services for a user to build a public profile and to explicitly declare the connection between his/her profile and those of the other users.

The currently popular OSNs are implemented using a centralized architecture, which means they are based on centralized servers storing all data generated by and about their users. This centralized structure has several drawbacks including scalability [5], dependence on a provider, and privacy [1].

In recent years, the rise and quick development of social networks has led to two important phenomena: user privacy exposure and the rapid spread of information. Social networks have become the epicentre through which individual privacy is violated. The very last scandal concerning users' data is the well-known Cambridge Analytica scandal[2], which erupted in early March 2018. Last, but not least, current OSNs are focused on keeping the user always engaged with more interaction opportunities and new content, not really caring about the user experience and the quality of the interactions.

What we foresee is that, instead of having more relationships, making more interactions and make the people engaging more, the trend should be to make the user have better relationships, and better interactions.

So, it will not be a matter of having more friends, but better friends; not a matter of posting, commenting and interacting more, but posting, commenting and interacting better, in such a way that a more meaningful ecosystems of interactions is created around the users.

These problems have moved researchers to investigate alternative OSN architecture solutions with respect to the centralized one [6][7].

A Decentralized Online Social Network (DOSN) [1] is an Online Social Network implemented on a distributed information management platform, such as a network of trusted servers, P2P systems or an opportunistic network.

During the last years, DOSNs have been the focus of several works and projects from both academic researchers and open source communities. By decentralizing OSNs, the concept of a service provider is changed, as there is no single provider but a set of peers that take on and share the tasks needed to run the system.

This has several remarkable consequences: in terms of privacy and operation, no central entity that decides or changes the terms of service exists. Moving from a centralized web service to a

[2] https://www.theguardian.com/news/series/cambridge-analytica-files

decentralized system also means that different system models become possible: using one's own storage or cloud storage, exploiting delay-tolerant networks, and/or P2P networks, to name a few. A Delay-tolerant network (DTN) [8] is a network which permits the communication between heterogeneous networks, possibly over extreme distances. DTN networks require specific hardware to store information which need to survive at power loss and system restarts. Usually, DTN social networks can be considered the Mobile Online Social Networks, presented in Section 2.2. On the other hand, a P2P network is a distributed network composed of a large number of distributed, heterogeneous, autonomous peers, in which peers (participants) share a part of their own resources (processing power, storage capability, etc.). This kind of network is highly dynamic.

The first big project in this area has been Diaspora[3], which counts more than 600,000 users to date. Three years ago, in October 2016, Mastodon[4] was launched. Mastodon is an online social media platform that allows anyone to host their own server node in the network. Mastodon and Diaspora are part of Fediverse[5], allowing its users to interact with users on different platforms that support the same protocol.

However, decentralizing the existing functionalities of Online Social Networks requires finding ways for distributing storage of data, privacy preservation, defining an overlay topology and a protocol enabling searching and addressing, robustness against churn, etc., as explained in [9].

HELIOS represents a step forward in the definition of a new generation of DOSNs, where actors are both users and objects, and users can generate a different layer of social relationships guided by the context.

1.2 About this document

The purpose of this document is to present the design of a new P2P Social Overlay, named *Heterogeneous Social Graph (HSG)*, which represents the principal structure of HELIOS, and fits the NGI initiative of putting users at the centre and building services around them.

In detail, the main goal of our P2P Social Overlay [10], represented by the Heterogeneous Social Graph, is to represent the social interactions between the actors of the HELIOS application, which are heterogeneous in nature (sensors, human, etc.). Nodes of the Heterogeneous Social Graph can manage the information about their social contents by exploiting a stack of ego networks modelled with a Pillar multi-network [11]. We call the stack of ego networks as Contextual Ego Network. In the Contextual Ego Network, each layer represents a user's context, and it is formally represented with the Ego Network Social Model, by using an undirected weighted graph. The Contextual Ego Network is used to manage the social connections of a HELIOS user. Social contacts, named alters, are added to the Contextual Ego Network of a user, named ego, when specific similarity metrics are respected. Then, all the alters are managed by considering the level of trust between them and the

[3] https://joindiaspora.com/

[4] https://joinmastodon.org/

[5] https://fediverse.network/

ego, which means that an alter is added to a context with an initial level of trust. HELIOS will be able to periodically re-compute the trust level between an ego and an alter, and the trust score will be used to evaluate when an alter is a trusted node by defining a trust model. Our novel P2P Social Overlay is modelled by considering the IoP-like approach and applies the three defining NGI principles in many of its aspects.

The deliverable is organized as follows: in Section 2 the State of the art is proposed. In Section 3 an overview of the main characteristics of HELIOS is proposed. Section 4 presents the overview of the HELIOS Social Network, named Heterogeneous Social Network Graph. Section 5 describes the Contextual Ego Network by giving a formal definition of the graph. Finally, Section 6 describes how trust is important in a Social Network and how the Contextual Ego Network is affected by the trust concept.

2 State of the Art

In this section, we provide an overview of the current generation of Decentralized Social Networks, by describing the state of the art of DOSNs, and the technology of Mobile and opportunistic social networks (MOSNs). Moreover, we provide an overview of the current approaches of Blockchain-based Online Social Networks (BOSNs).

2.1 Decentralized Online Social Networks (DOSNs)

The main difference among the current DOSN proposals concerns the technologies and techniques used to store and manage data. A possible classification considering this difference has been proposed in [9]. In this section, we propose an overview of the current DOSNs: federated and fully-decentralized approaches.

2.1.1 Federated Online Social Networks

One of the first decentralized solutions, which has today more than 600,000 users, is *Diaspora* [12]. Diaspora is a federated DOSN, where users provide servers that are administered by themselves and that allow Diaspora users' profiles to be hosted on their servers.

If Diaspora can be considered the distributed version of Facebook, Mastodon [13] represents the distributed version of Twitter. Mastodon is a decentralized microblogging network based on open protocols and free, open-source software. During the last two years, Mastodon was increasing the number of users (about 2M of users), surpassing Diaspora. Mastodon is formed by a set of servers, known as instances, and each user is a member of a specific Mastodon instance, but can connect and communicate with users on other instances. Like Twitter, Mastodon supports direct, private messages between users, but unlike Twitter, Mastodon’s messages can be either private to the user, private to the user's followers, public on a specific instance, or public across a network of instances.

Mastodon is part of the Fediverse, an interconnected and decentralized network of independently operated servers, which includes platforms such as Diaspora, Friendica, GNU Social, PeerTube, etc. Fediverse is a common name for the union of various federated social networks which use a set of standard protocols: OStatus, ActivityPub, DFRN, Diaspora Network, and Zot.

Table 1 lists all the platforms included in Fediverse with information about the type of platform, the protocol used in that platform and the url of the website. Fediverse includes active platforms, which are used by real user communities.

Table 1. Fediverse: list of platforms

Platform name	**Type**	**Protocol**	**Website**
Diaspora	Social Network, Microblogging	Diaspora Network	https://diasporafoundation.org/

Friendica	Social Network, Microblogging	ActivityPub	https://friendi.ca/
GNU Social	Microblogging	OStatus	https://gnu.io/social/
Hubzilla	CMS, blogging, File hosting	ActivityPub	https://zotlabs.org/page/hubzilla/hubzilla-project
Mastodon	Microblogging	ActivityPub	https://joinmastodon.org/
Misskey	Social Media, Microblogging	ActivityPub	https://joinmisskey.github.io/ja/
PeerTube	Social Media, Microblogging	Zot/6	https://joinpeertube.org/en/
Pleroma	Social Media, Microblogging	ActivityPub, OStatus	https://pleroma.social/
Socialhome	Social Media, Microblogging	Diaspora Network	https://socialhome.network/
PixelFed	Social Media, Image Sharing	ActivityPub	https://pixelfed.org/

2.1.2 Fully-decentralized DOSNs

Several important DOSN applications have been proposed in the research area. These provide novelty aspects in terms of data availability, security, privacy, and information diffusion.

Safebook is proposed in [14]. This involves a three-tier architecture for DOSNs with the main focus on privacy, integrity and availability. Each user in SafeBook has a set of logical concentric structures called Matryoshkas. Matryoshkas are concentric rings of nodes built around each peer and provide a trusted data storage and communication obfuscation through indirection.

PeerSon [15], [16] is a two-tier architecture in which one tier is implemented by a Distributed Hash Table (DHT) and it serves as a look-up service. The second tier consists of peers and contains the user data, such as user profiles.

LifeSocial.KOM [17] is a plugin-based architecture, which provides the common OSN functionalities. It uses a DHT, in detail FreePastry[6] and PAST [18][19] as data storage. Data are encrypted by users and only authorized users can access it.

[6] http://www.freepastry.org/FreePastry

GemStone [20] is a P2P social network system that acts as a middleware to support different OSN applications. Gemstone assists these applications by proving a shared social graph, serving profile information and handling message delivery to peers. GemStone is completely decentralized and protects the user's privacy by encrypting all data using an Attribute Based Encryption (ABE) technique [21]. The system provides a data storage solution based on data replication; private data are stored among a set of other nodes called Data Holding Agents (DHAs). Confidentiality is a key element in GemStone. The system allows each user to grant fine-grained access to his/her confidential data. Data cannot be accessed by other entities than those that have the corresponding decrypting key. Data Management is strongly dependent on the choice of topology.

Cachet [22] is an architecture that provides security and privacy by guaranteeing confidentiality, the integrity and the availability of the user content. In Cachet the DHT is augmented with social links between users. Cachet provides a distributed pool of nodes to store user personal data and these nodes are untrusted. The system uses the DHT as a base storage layer, and a gossip-based social caching algorithm. Cachet uses social caching to manage data availability and information diffusion.

Prometheus [23] gathers information about users through social sensors that are applications running on behalf of the user. Social Sensors are able to retrieve information from other sources, such as Facebook. In particular, they retrieve interactions with other users via email, phone, instant messaging, comments on blogs, which are used to create a weighted, directed, and labelled multi-edged graph, where vertices correspond to users and edges correspond to interactions between users. Both the social information from sensors and the social subgraphs are stored and maintained in a P2P network. Information from sensors can be decrypted only by "trusted" peers, which are selected by users.

My3 [24] is a privacy-friendly DOSN which exploits well-known interesting properties of Online Social Networks, for instance the locality of users and the trust among them. Users' profiles are hosted only on a set of self-chosen trusted nodes, called Trusted Proxy Set (TPS). Exploiting availability and performance goals, as its geographical location and the online time period of the user, populates the TPS of a user.

DiDuSoNet [25] is a two-tier system, where the lower level is implemented by Pastry[7], and it is used for the bootstrapping phase, for the look-up service, for searching other users, and to retrieve the replica nodes list. A Dunbar-based Social Overlay implements the upper level. In the Social Overlay, nodes are connected to other nodes with whom the tie strength computed on the interaction between them is higher. The novelty of this system with respect to the others described before is that it is completely based on trust between users. Social data are stored only on trusted nodes chosen with respect to the Dunbar's number [26]. Each node can choose two replicas to have a high level of availability.

Several DOSNs integrate the P2P layer with external resources, such as cloud storage services, to increase the quality of service. External resources are used to cope with the situations in which the users cannot deliver the service by themselves. Vis-à-Vis [27], Vegas [28], and SuperNova [29] are only three examples of approaches in which cloud services are used.

[7] https://en.wikipedia.org/wiki/Pastry_(DHT)

2.2 Mobile and Opportunistic Social Networks (MOSNs)

With the advent of personal devices like smartphones, OSNs witnessed a dramatic change. In fact, smartphones can be considered as the online alter-ego of people and are affected by the same mobility patterns of their owners.

A Mobile and Opportunistic Social Network (MOSN) is a platform that delivers social network functionalities combining techniques from social sciences with wireless communication for mobile networking [30]. At a high level, a MOSN architecture is introduced in [31] and shown in Figure 1 (taken from [31]).

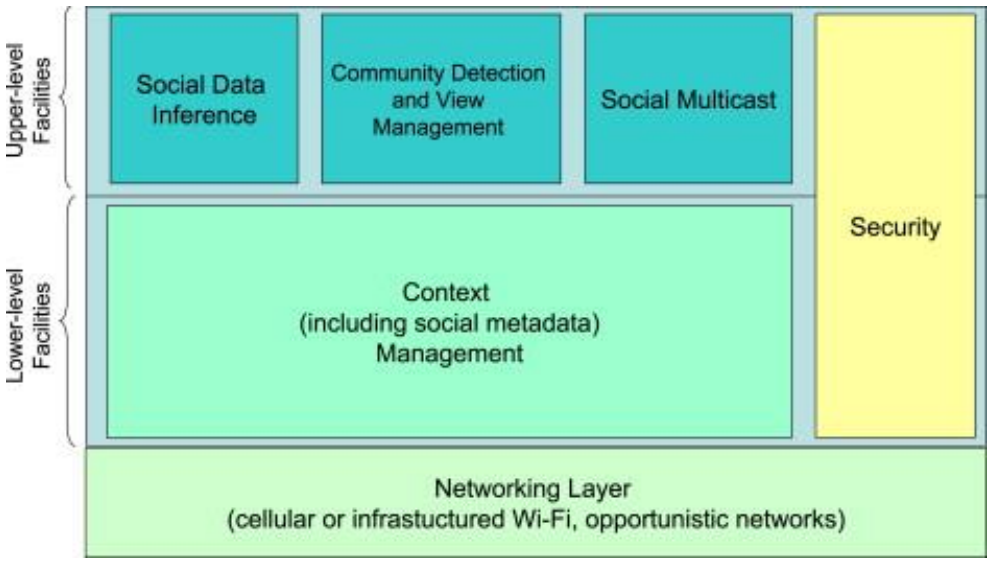

Figure 1. High level architecture for MOSNs

The architecture is divided into two layers, built on top of a Networking layer which provides connectivity with the other entities of the system. As major novelty, here we find the possibility to communicate through opportunistic networks built over spatially close connections.

In the Lower-level facilities, the middle layer, we find facilities to manage metadata coming from the topmost layer and to support location tracking.

In the Upper-level facilities layer we find the support for the three most important problems in MOSNs:

1. how to find new social ties (Social data inference facility),
2. how to build groups with the social ties found (Community detection and view management facility), and
3. and inter/intra group information dissemination (Social multicast facility).

The Social data inference facility is the component that lets users discover other users by using heterogeneous sources of information. Possible sources are OSN social graphs and other pre-existing services, but of crucial importance are the ones coming from the opportunistic connections. Indeed, two people close to each other may be colleagues, neighbours, or have similar interests because they are attending the same event, visiting the same place or dining at the same restaurant. Thanks to this gathering of different data coming from heterogeneous sources, it allows users to build a directed weighted social graph. This graph represents a sort of ego network of the user where weights on the edges model how much users are alike.

A graph alone is not very useful; therefore, a Community detection and view management facility is needed in order to group users. Indeed, community detection is the process to discover groups of users.

The Social multicast facility makes use of these communities to provide socially aware multicast functionalities. The service will be able to manage information diffusion within the group and enables relevant social updates within the scope of the group.

Orthogonal to the two levels, we have the security facility. At the upper level it gives users the possibility to define security preferences and policies (who/when/how can access its public data). On the lower-level it ensures confidentiality, integrity, authenticity, and it enforces the privacy policies defined by users.

Introducing mobility, a new, orthogonal dimension to the system, adds more problems but also gives more opportunities to solve them. Indeed, communication may happen over Internet, like in traditional centralised OSNs, or by exploiting the decentralization, like in DOSNs or in MOSNs where usually opportunistic connections are the only communication channel between mobile devices. This adds more possibilities in building the so-called *wisdom of the crowd*.

For instance, MobiClique [32] is a MOSN middleware which is bootstrapped with a profile available on existing OSNs (virtual world) and then enables opportunistic temporary connections based on physical proximity and social compatibility (physical world). The downside of this approach is the fact that it is unable to predict user contacts, which leads to using message flooding to implement content dissemination.

AdSocial [33] is a MOSN which supports presence detection, games, chat, voice and video calls over an ad hoc network, specifically targeting small mobile devices, which have strict resource constraints. AdSocial uses MAND (Mobile Ad hoc Network Directory), which is an ad hoc network-specific distributed directory service, to locate nearby users and to determine their address.

2.3 Blockchain Online Social Networks (BOSNs)

During the last three years, we have seen the rise of Blockchain-based Online Social Networks (BOSNs). The lack of success of DOSNs, and the increase of problems concerning OSNs, such as online disinformation or data disclosure, has been the primary motivation to combine social platforms with the blockchain technology. Indeed, these platforms give more importance to the content by providing rewarding systems and they aim addressing the problems of privacy and online disinformation (aka fake news) using the blockchain technology.

Steemit[8] is a social media platform where everyone can receive a reward for creating and curating content, in the form of the Steem cryptocurrency. It has more than one million users and it represents the most well-known BOSN. Steemit is a social platform that grows communities and returns the value to the people who contribute the most. An important characteristic of Steemit is that, unlike most blockchains that are too slow and expensive to be used for apps, it is fast, free, and scalable, as explained on the website. A difference between Steemit and other platforms is that there are three different kinds of currency units: STEEM, Steem Power (SP), and Steem Dollars (SBD). STEEM is the unit that is bought and sold for actual money on the open markets. It is the principal cryptocurrency of the network and the other two kinds of units are dependent on it. Steem Power is a kind of long-term investment because people cannot sell this unit for 2 years. Steemit operates based on one-STEEM one-vote, instead of one-user one-vote, as in other platforms. Within this model, individuals who have contributed the most to the platform have the most influence over how contributions are scored.

Lit[9] is a platform created to integrate social media services and cryptocurrencies, like Instagram and SnapChat. The main feature of Lit is that users can share stories via Lit Stories and their stories permit to obtain Mithril tokens (MITH), taken by considering the impact and influence of these stories across the network. Stories are any content a user can share photos, slideshows, videos, posts, etc.

HyperSpace[10] (known as Synereo) is defined as a blockchain-based OSN based on the Attention Economy [34]. Synereo has been launched in February 2018 to build the economy of attention in order to reward users for their attention, but also to direct that attention to relevant content. The attention economy is a subset of the information economy, which concerns in the definition of a marketplace where consumers agree to receive services in exchange for their attention.

SocialX[11], as all the previous platforms, is decentralized and allows users to give content feedback and reward tokens. SocialX is fully decentralized, which means that all media files (photos and videos) and data (messages, posts etc) are decentralised. The platform wants to face the problem of fake accounts, fake followers, and fake votes (likes, etc.). Indeed, the decision power is given to communities, which can decide what content is valuable. The community is the main concept that can decide which content can be rewarded because the platform has the property of self-governance. SocialX uses the blockchain to optimize the reward system in order to confirm the community actions.

Sapien[12] is a democratized social news platform built on the Ethereum blockchain. The Sapien Network consists of the Sapien platform, marketplace, API integrations, and third-party applications, all connected and powered by SPN, an Ethereum-based utility token.

---

[8] https://steemit.com/

[9] https://mith.io/en-US/

[10] https://site.hyperspace.app/

[11] https://socialx.network/

[12] https://www.sapien.network/

Sola[13] is a BOSN, which has more than 700,000 users. The difference between Sola and the other BOSNs is how the system spreads the information. Indeed, it uses a process like a viral disease to spread the information to the most interested users, applying AI algorithms combined with users' reactions.

[13] https://sola.foundation/

3 HELIOS: a Next-Generation DOSN

The main idea of HELIOS is to introduce a people-oriented platform, rather than service-oriented, which can be adapted to the user's behaviour by, for instance, exploiting sensors and services deployed in a smart environment. HELIOS is a platform which follows a 'Trust by Design' paradigm by taking into account the main properties of the previous generations: DOSNs, MOSNs, and BOSNs. HELIOS will include new key concepts for a DOSN, such as: human-centric computing, contextual networking, computational trust, privacy by design, and so on.

Since some of these concepts are completely new, while others have never been applied to the DOSN area of research, we briefly review the concepts that are important as a basis for defining the Contextual Ego Network.

Human-centric computing. As suggested in the Introduction section, many services have already begun their transition from a service-centric paradigm to a human/user-centric one. Services undergoing this transition are abandoning old system developments, where the services were built around the infrastructure and using design principles to minimize all the costs. This brought very cheap, in terms of time, money, and resources invested, services, and then, around these services, interfaces were built to let people access them. While this is optimal from an investor point of view, from the point of view of a user it is a poor choice.

With a human-centric approach, instead, services should be designed putting the user at the center, considering his/her needs and preferences, and then building the service around the user itself. This should lead to more personalized services, tailored to the people's needs, and enriching their experience over the service.

Contextual networking. To empower organic meaningful relationships, we also have to take into account that human interactions are highly contextual. Contextuality of interactions comes from the fact that humans tend to interact on certain topics only with a subset of their acquaintances. For instance, it is with colleagues that we share work related information, while it is with our family or our close friends that we share personal news about our relatives. Contextuality is not only just a matter of the people we interact with, but also the role we play in the interaction. For instance, two friends may be considered as peers when chatting, but one of them may be considered more important or expert in a particular area of interest. In such cases, this kind of contextual information could be leveraged by intelligent service components to improve the end user experience, e.g. better targeted recommendations that consider expertise.

Interpersonal Computational trust. The concept of trust, which is intertwined with the concept of privacy, is highly underused in DOSNs and good models for trust, which do not consider only relationships among users, do not exist. A trust model in a DOSN can be used as a support for the realisation of advanced solutions for many problems, such as secure data availability, efficient information diffusion, and so on. We foresee that a good trust model will be a combination of the number of interactions, type of interactions (positive vs negative), how the user is understood by the community (trustworthy vs untrustworthy, fake vs real, influencer vs spammer user), and other parameters. The model will also include the fact that trust is evolving through time, according to human behaviour and recent activities.

4 The Heterogeneous Social Network Graph

The main scope of this deliverable is to present an innovative data structure to model a Decentralized Social Network Graph, called *Heterogeneous Social Network* Graph by introducing how it is organized and its lifecycle.

4.1 Limitations and open issues of current DOSNs

The first DOSNs solutions consider federated or fully decentralized P2P networks. Usually the main difference between them consists of the storage choice [6]. As explained before, a federation of servers is an easy solution to implement a Decentralized Online Social Network, however there is centralization due to the servers, which could potentially result in single-point of failure risks.

Mobile and opportunistic Social Networks represent a different concept of Decentralized Online Social Network. The sociality in this kind of environments is in principle related to the distance between users, and they fail to provide useful tools or techniques for analysing the Social Network in real time.

The latest technology used to implement Online Social Networks is the blockchain technology. However, this solution does not provide a real Decentralized Online Social Network due to the public visibility of private content, which is stored on the blockchain, like in Steemit. Furthermore, BOSNs have a rewarding system used to remunerate users who develop content. Entities receive rewards based on the upvotes and comments on the content they have produced. The rewards are taken from a daily generated “rewards pool”, with entities holding larger percentage of the reward pool having bigger say at how the rewards are distributed.

DOSNs have changed the way of how users can manage their data, however they did not manage to change the reality, where centralised OSNs have been the most used platforms. A potential reason why this happens is because DOSNs, and the new generation based on the blockchain technology, are too similar to the most popular OSNs. They try to offer the same services by guaranteeing a high level of control over private data. Usually, they use encryption techniques to store data in a secure way [35], or they use the concept of trust by exploiting trust nodes [25], [36] as storage nodes, or privacy policies [37]. Even though they have attracted several users, they cannot yet be considered as serious contestants to platforms such as Facebook, Twitter, etc.

HELIOS represents a new generation of DOSNs where all the previous distributed technologies are mixed together in order to overcome the limitations of the single technology. For example, HELIOS addresses the limitation of MOSNs concerning the prediction of users. Moreover, HELIOS can overcome the problem of current BOSNs where usually data are stored in the blockchain, and/or the rewarding system significantly changes the feel of users. Indeed, the gain becomes the main motivation. Moreover, HELIOS is a people-oriented platform which considers the behaviour of users in order to manage the Social Network. The main goal of HELIOS is to be a useful tool for a user.

4.2 Topology

From a general point of view, a graph is a collection of nodes and edges that represent specific relationships, where nodes correspond to actors, and edges are the connections between the actors.

P2P systems have a network topology that is defined as overlay network. In current DOSNs, the network topology is represented by the Social Network graph, which is usually modelled by a Social Overlay [10], [25]. A Social Overlay is a logical overlay in which peers are connected to known peers. An edge between a pair of nodes indicates that a tie exists between two adjacent nodes.

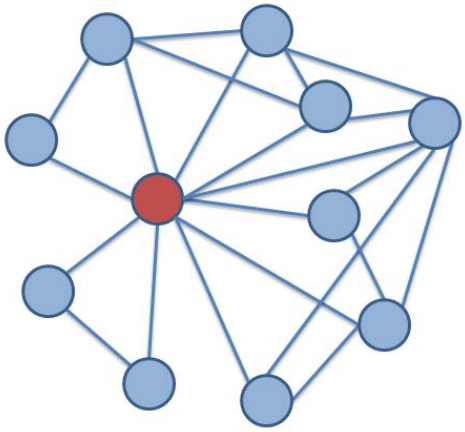

Figure 2. An example of an ego network.
The red node is the ego.

However, due to the huge amount of nodes in a P2P network and by considering the distributed nature, nodes maintain only a subset of the nodes in the network in a local view, and several heuristics are proposed to build the local view of a node by taking into account a specific scenario. In a DOSN, the common heuristic is to have a local view containing only the friend nodes. The *Ego Network* [38] is a well-known social network model used to model the local view of a node in a DOSN. The Ego Network of a user represents a structure built around the user itself, also known as ego, which contains her/his direct friends, known as alters and may also include information about the direct connections between the alters.

Figure 2 shows the ego network of the red node, the ego, with the blue nodes, its alters, and the relations among them.

HELIOS considers a set of heterogeneous actors (see Table 2), which can be for instance humans or smart devices, and the connections between actors have a different nature that depends on the two actors involved. To model such a graph, we implement the Social Overlay with a *Heterogeneous Social Network Graph.*

The *Heterogeneous Social Network Graph* is the union of the local views of each node which are implemented with an enriched Ego Network model, called Contextual Ego Network, explained in Section 5.

4.3 Lifecycle

The HSN is the union of the local views of users, represented with the Contextual Ego Networks. The creation of the HSN starts when at least one user runs the HELIOS application. Then, the evolution depends on the lifecycle of each Contextual Ego Network. Indeed, we identify the lifecycle of the HSN with the lifecycle of its components (the CEN of each user).

Figure 3. Contextual Ego Network lifecycle

Figure 3 shows the lifecycle of a CEN, which consists of:

1. Creation. A CEN is created when the device owner starts to use the HELIOS Application.

2. Add/Update/Remove a context. A CEN is a multi-layer network where each layer consists of a specific context detected from the HELIOS Core. Based on the user activity, contexts are created, updated, removed.

3. Add/Update/Remove an alter. When a specific context is created and added to the CEN, it will be populated by adding alters. When the application retrieves a similarity between the profile of an alter and the profile of the ego, the alter is added to a context. As in common Social Networks, the relation between two users evolves over time. For this reason, an alter can be also updated, which means, for example, update information about the relationship between the ego and the alter, such as the trust score. Finally, the relation can be removed.

In order to explain the lifecycle, Figure 4 shows the high level overview of the HELIOS Application concerning the modules included in the creation and management of the *Heterogeneous Social Network Graph*. Indeed, at the top level, we have the HELIOS application installed on the user device, which is able to retrieve information about the user by exploiting the hardware and, when it is possible, the smart environment. The *Heterogeneous Social Network Graph* (and the Contextual Ego Network as its parts) are managed by the HELIOS Core. Indeed, the structure is part of the HELIOS Core and the Social Ego Network Manager is the module used to manage and update the *Heterogeneous Social Network Graph*. In detail, the *Heterogeneous Social Network Graph* communicates with the Social Ego Network Manager, at least to:

- retrieve information about Context: creation, deletion, update, and to know the current user context;
- retrieve information about Alters: when an alter must be added/removed to a specific context, to update the information concerned an alter
- retrieve information concerning the interaction between an alter and the ego useful to manage the weighted of an edge and the trust score associate to an edge.

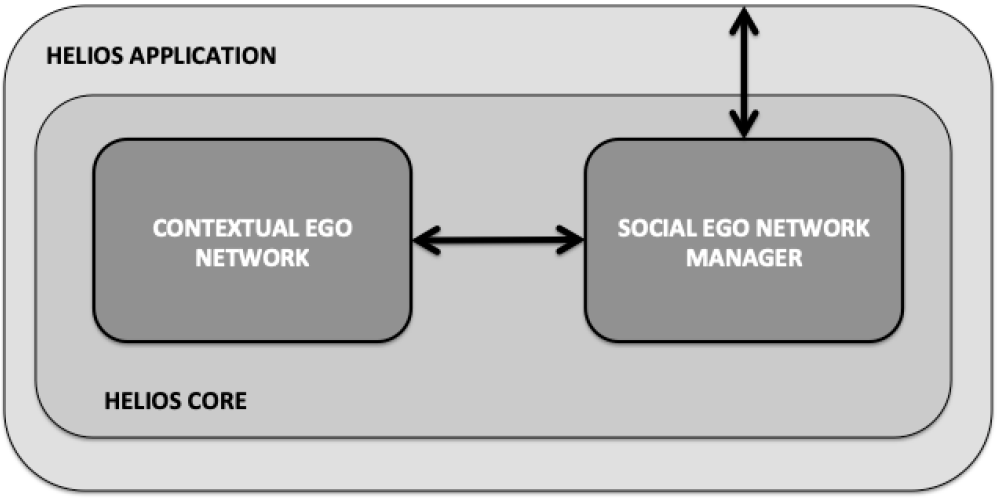

Figure 4. High level overview of the Helios Application concerning the modules included in the creation and management of the Heterogeneous Social Graph.

Figure 4 explains the direction of communications, to provide a Social Network service.

The P2P Social Overlay Manager is able to retrieve information about the external events by considering that the mobile device is the alter-ego of a user. External events needs to be managed because they can affect the structure of the HSG, and in detail the Contextual Ego Network, which represents the local view of the user. Indeed, the mobile device represents the alter ego of a user and the HELIOS application can change the structure of the Contextual Ego Network in particular when a new context is created/updated/deleted and/or a new relationship is added/updated/removed.

4.4 Actors

Although social networks were originally conceived as a model of capturing links and interactions among humans, it is nowadays clear that modern socio-technical networks may comprise a variety of actors. Given a digital world setting where real-world entities and objects have digital counterparts, but there are also digital-only entities, in addition to humans, one may consider the following entities as first-class nodes in a heterogeneous social network graph:

- ***People***. This is the type of social media accounts that represent humans which are using their personal devices to access to the HELIOS Social Network.
- ***Organizations***. These are the types of social media accounts that represent a real-world or digital organization. For instance, Facebook company pages, YouTube news channels, football team Twitter accounts, etc. These are typically operated by one or more designated humans (e.g. social media manager) from the organization, but their nature and behaviour as nodes of the envisioned social graph is different to the one of other types of nodes described so far. Often, such nodes may be considered as authoritative sources of information for specific domains/topics.
- ***Social Bots.*** These are AI-powered digital agents that operate within specific social media platforms (e.g. Twitter). These may serve a variety of purposes, benign or harmful (Ferrara et al., 2016) even though most platforms explicitly prohibit harmful uses. The simplest cases of bots include social media accounts that act as publishers or repeaters of content, while a more advanced generation includes bots with chatting capabilities, and even more sophisticated bot systems could involve multimodal communication (e.g. audio/video).
- ***Smart Objects***. These include IoT devices having HELIOS core software or alternatively IoT devices that are locally connected to a HELIOS node that is acting as an IoT gateway providing sensor measurements to HELIOS network. Smart Objects are HELIOS-connected hardware devices that typically measure or record data from their surroundings and transmit them over the network. Seen as social network actors, these sensors or actuators may be considered as nodes with limited interaction capabilities, even though they may be considered as valuable data and context providers to other nodes of a heterogeneous social network.

We identify yet another category, which consist of a set of different programmatically-driven digital agent, such as for instance social network spiders/crawlers, of which the main purpose is to visit and record the social network. One might also image other purpose-specific agents.

The above listing of potential social network actors makes clear that each such actor has a vastly different set of capabilities and may engage in different types of interaction and exchange. Given that HELIOS should be designed around the widest possible type of actor, its core API should consider this variety.

Table 2 summarizes the types of actors that are envisioned to be part of the networks and presents some example properties, actions and the type of content related to each actor category.

Table 2. Actors of the Helios Heterogeneous Social Network Graph

Actors	**Example Properties**	**Actions**	**Content**
People	name/identifier, gender, age, ethnicity, location, end device, etc.	Post Message React	Text Multimedia Sensor data
Organization	identifier, category of business, size, foundation date, mission, location, etc.	Post Message React	Text Multimedia State
Social Bots	(a mix of properties described for people and smart devices actors) name/identifier, connectivity protocol, assigned gender, state, etc.	Post Message React	Text Multimedia State
Smart Objects	name/identifier, manufacturer, connectivity protocol, state, sensor 1, sensor 2, etc.	Message (domain specific) Data transmission	Text Multimedia Data State

4.5 HELIOS actor profile

User profiles in online social networks can be about any number of characteristics associated with individuals and/or other actors in a social graph, such as personal information related to name, ethnicity, age, gender, interests, expertise, professional affiliations, connections, status, recent activity and geographic location, to name a few. Such profile information is used as a basis for grouping users, for sharing content, and for recommending or introducing new contact and connection opportunities.

Today's online social networks rely on users to manually input profile attributes, representing a significant burden on users, especially with them often being members of multiple online social networks. On top of that, users often hesitate sharing real profile information to federated centralized online social networks fearing privacy breaches or misuse of their personal information and either do not disclose all data or provide false information.

In HELIOS we envision a user profile with different access levels which is safely stored and dynamically updated in the user's personal persistent storage layer, while access to certain characteristics (referred to hereinafter as profile properties) is granted to third party applications or alters either manually, through explicit user authorization, or dynamically based on context and trust criteria.

Since HELIOS fosters a heterogeneous social graph, meaning that actors in the graph vary from people to organizations, and from sensors to smart objects and social bots, the minimum user profile consists of attributes that can be universally valid for all types of actors in a meaningful way.

Towards building successful user profiles that will be easily interlinked and integrated in the linked data ecosystem, guaranteeing direct applicability and low entry barriers, we consider as a beneficial option the adoption of existing vocabularies that have already attracted a considerable user community. This aims at identifying and comparing existing approaches in order to devise best practices on how to leverage existing vocabularies conjointly to form the Helios user profile.

FOAF and SIOC: The Friend of a Friend[14] (FOAF) and the Semantically-Interlinked Online Communities[15] (SIOC) vocabularies mark the starting points of our study.

FOAF was designed as a machine-readable ontology describing persons, their activities and their relations to other people and objects in a linked information system. FOAF takes a liberal approach to data exchange. It does not require you to say anything at all about yourself or others, nor does it place any limits on the things you can say or the variety of vocabularies you may use in doing so. It provides a basic "dictionary" that was designed to be used alongside other such dictionaries ("schemas" or "ontologies"), and to be usable with a wide variety of generic tools and services that have been created for the Semantic Web.

In short FOAF is defined as a dictionary of terms, each of which is either a *class* or a *property* and other projects alongside FOAF provide other sets of classes and properties, many of which are linked with those defined in FOAF. The specific contents of the FOAF vocabulary are detailed in the FOAF namespace document, while

Table 3 summarizes alphabetically of all FOAF terms, by class (categories or types) and by property. Note that it includes 'archaic' terms that are largely of historical interest What is of particular interest for our needs is that FOAF allows groups of people to describe social networks without the need for a centralised database. Each FOAF document is itself an encoding of a descriptive network structure. The documents can be easily merged, allowing partial and decentralised descriptions to

[14] http://xmlns.com/foaf/spec/

[15] https://www.w3.org/Submission/sioc-spec/

be combined in interesting ways. To make some examples, one of the most used properties of the FOAF vocabulary is the "`foaf:knows`" property which provides a simple way to create social networks through the addition of "knows" relationships for each individual that a person knows while the "`foaf:interest`" property defines topics of interest to a person, and can be used directly to find those with an interest in a particular domain.

Table 3. FOAF Terms

Classes: | Agent | Document | Group | Image | LabelProperty | OnlineAccount | OnlineChatAccount | OnlineEcommerceAccount | OnlineGamingAccount | Organization | Person | PersonalProfileDocument | Project |
Properties: | account | accountName | accountServiceHomepage | age | aimChatID | based_near | birthday | currentProject | depiction | depicts | dnaChecksum | familyName | family_name | firstName | focus | fundedBy | geekcode | gender | givenName | givenname | holdsAccount | homepage | icqChatID | img | interest | isPrimaryTopicOf | jabberID | knows | lastName | logo | made | maker | mbox | mbox_sha1sum | member | membershipClass | msnChatID | myersBriggs | name | nick | openid | page | pastProject | phone | plan | primaryTopic | publications | schoolHomepage | sha1 | skypeID | status | surname | theme | thumbnail | tipjar | title | topic | topic_interest | weblog | workInfoHomepage | workplaceHomepage | yahooChatID |

There have been several extensions or modules for the FOAF ontology that are of interest to the Helios project. FOAFRealm [39] system that implements a distributed user profile management system and delivers semantic social collaborative filtering features. D-FOAF [40] is FOAF-based distributed identity management system for social networks, where access rights and trust delegation management are provided as additional services. In D-FOAF, relationships are associated with a trust level, which denotes the level of friendship existing between the users participating in a given relationship. As far as access rights are concerned, they denote authorized users in terms of the minimum trust level and maximum length of the paths connecting the requestor to the resource owner.

The SIOC project aims to provide a framework for the connection and interchange of information from internet-based discussions and community portals. Such communities are primarily made up of users, the posts that they create, and the discussion forums that they subscribe to across a multitude of sites and discussion platforms. The basis for SIOC is the SIOC ontology, an RDF-based schema which describes the main concepts found in online communities. Table 4 provides and alphabetical index of SIOC by class (concepts) and by property (relationships, attributes), are given below. All the terms are hyperlinked to their detailed description for quick reference.

Table 4. SIOC Terms

Classes: | Community | Container | Forum | Item | Post | Role | Site | Space | Thread | UserAccount | Usergroup |
Properties: | about | account_of | addressed_to | administrator_of | attachment | avatar | container_of | content | creator_of | delivered_at | discussion_of | earlier_version | email | email_sha1 | embeds_knowledge | feed | follows | function_of | generator | has_administrator | has_container | has_creator | has_discussion | has_function | has_host | has_member | has_moderator | has_modifier | has_owner | has_parent | has_reply | has_scope | has_space | has_subscriber | has_usergroup | host_of | id | ip_address | last_activity_date | last_item_date | last_reply_date | later_version | latest_version | likes | link | links_to | member_of | mentions | moderator_of | modifier_of | name | next_by_date | next_version | note | num_authors | num_items | num_replies | num_threads | num_views | owner_of | parent_of | previous_by_date | previous_version | read_at | related_to | reply_of | respond_to | scope_of | shared_by | sibling | space_of | subscriber_of | topic | usergroup_of

Within the scope of HELIOS and specifically for the envisioned group communication services the SIOC vocabulary can be particularly relevant in forming the basis for a well-defined profiling of user activities and connections based on communication interactions. While there are many classes and properties in SIOC, the main notion is that a `sioc:User` creates `sioc:Posts` that are contained in `sioc:Forums` that are hosted on `sioc:Site`, where *Forums* in the framework of Helios can be mapped to group communications networks and *Site* can be linked to a given Context.

5 Contextual Ego Network

The local knowledge of a user in a DOSN is usually modelled with an Ego Network, which maintains the online social contacts of each user.

The vision of HELIOS introduces two important aspects, which are reflected in the model used to implement the Social Overlay of the system. The first aspect is the nature of an actor involved in the overlay, which could be a human or an object (see Section 4.2). The second one is the contextual networking, which means that the real-life activity of a user must model its local view as a virtual view of the daily life.

In this Section, we introduce the enriched structure called *Contextual Ego Network*. The Contextual Ego Network is one of the most important parts of the HELIOS framework because it represents the people-centered approach. The structure manages all the features described in the previous section: human-centric computing, meaningful relationships, contextual networking, computational trust and all these characteristics are integrated into the Contextual Ego Network.

A Contextual Ego Network is a complex model organized in layers, where each layer represents a real-life context of the ego. Each layer, in turn, can be implemented as a simple Ego Network, where actors are heterogeneous (human and/or sensors available in the smart environment), and links between two actors describe specific relationships according to the nature of the actors and the context in which they are.

The definition of the Ego Network considers the alter-alter ties. This information is usually a derivable information, because when information about social contacts is exchanged, nodes can find the intersection of social contacts. In HELIOS, specific privacy issues will be managed by the Security and Privacy Manager, in the HELIOS core.

A Contextual Ego Network should represent the daily life of a user by representing the different context in which he/she lives. A single layer models a user's context. A context is a situation in which each user can find her/himself. A context can be described at this level by using three aspects: spatial aspect, temporal aspect and social aspect. The spatial aspect describes where the context happens, the temporal aspect describes when the context happens, and the social aspect describes with whom the context happens (which are the actors involved). Each context is local to its own user and asymmetrical with respect to analogous contexts of other users.

From a purely mathematical point of view, a context can be described with a tuple $C = (s, t, p)$, where s belongs to a spatial domain S, t belongs to a temporal domain T, and p belongs to a social domain P. Each of the domains is used to specify one of the three aspects defining a context: S for the spatial aspect, T for the temporal aspect, and P for the social aspect.

A social network can be described as a set of people (actors) with some interaction patterns between them at the same network level. In our scenario, this is different because relationships amongst the members of the social network take place in different contexts. For this reason, a Contextual Ego Network cannot be easily modelled with classic complex network models.

Multilayer networks [41], [42] are a complex structure which can be used to describe our Contextual Ego Network. Indeed, they are useful to represent systems interconnected through different

categories of connections. Each activity/context/category is represented by a layer and the same node can have different social interactions because each layer contains a set of neighbours.

As described in [41], in a social network environment, we can consider different relationships: friendship, vicinity, co-worker, etc., and different relationships can be modelled through multilayer networks (see Figure 5 [43]).

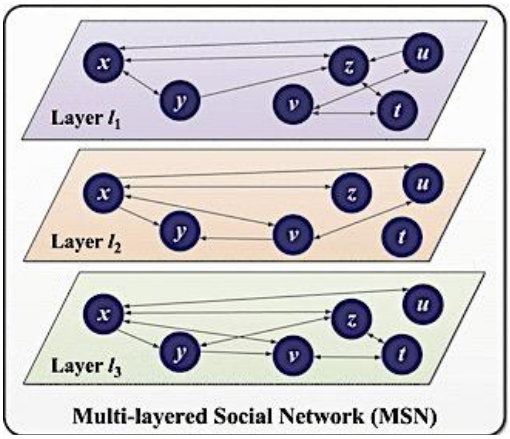

Figure 5. An example of multilayered social networks

Formally, a multilayer network is described in [41] as a pair $\Delta = (\mu, \pi)$ where $\mu = \{G_\alpha; \alpha \in$

$\{1,..,M\}\}$ is a family of graphs (directed, undirected, weighted, or unweighted) $G_\alpha=(X_\alpha,E_\alpha)$, called layer of M, and $\pi = \{E_{\alpha\beta} \subseteq X_\alpha \times X_\beta; \alpha,\beta \in \{1,..,M\}, \alpha \neq \beta\}$ is the set of interconnections between nodes of different layers. The elements of are called *crossed layers*, and the elements of each E_α are called *intralayer connections* of Δ, in contrast with the elements of each $E_{\alpha\beta}$ $(\alpha \neq \beta)$ that are called *interlayer connections*.

Several types of multilayer networks exist [41]. One of them is the *multiplex network* [44] a special type of multilayer network in which $X_1 = X_2 = ... = X_M = X$ and the only possible type of interlayer connections are those in which a given node is only connected to its counterpart nodes in the rest of layers. In short, multiplex networks consist of a fixed set of nodes connected by different types of links.

Multidimensional networks [45] are a mathematical model capturing multiple different relations that act at the same time. In a multidimensional network, a pair of entities may be linked by different kinds of links. Each possible type of relation between two entities is considered as a dimension of the network. In the case of a multidimensional network model, a network is a labelled multigraph, that is, a graph where both nodes and edges are labelled and where there can exist two or more edges between two nodes. Definitions are general and consider a general scenario. As concerns Social Networks, a multilayer network is a useful structure which can model different actors and interactions on different contexts [42].

We decide to use the *Multi-Layer Model (ML-model)* presented in [11] to formalize the Contextual Ego Network because the ML-model can model the behaviour of users in different online social contexts, such as different Social Network accounts. In our case, the model should consider the decentralized scenario. In fact, the ML-model is a formal model able to represent the different contexts in which users are involved during their everyday online activity.

The definition of a multi-layer model, as described in [11], is a weighted graph G=(V, E, w) where V is a set of vertices, E the set of edges and w is a weight that typically represents the strength of a relationship e in E.

When we consider multiple layers, we need to know which nodes are included in more than one layer. This can be done by using a specific Node Mapping[11].

Having said that, a Contextual Ego Network can be defined by exploiting an instance of the multi-layer model, called *Pillar multi-Network* [11]. A Pillar multi-network is characterized by |C(u)| in {0,1} and it represents a user as a pillar traversing every layer. The pillar multi-Network is helpful to implement our scenario. Indeed, it allows different nodes sets for each layer, and it provides a node mapping function between layers.

5.1 Contextual Ego Network properties

In the Social Network Analysis field, Social Networks are represented by graphs (also known as sociograms) or matrixes in order to model specific characteristics. In the rest of the Section, the HELIOS social network will be defined by considering the graph notion.

The main categorization concerns the graph notion is: Directed and Undirected graph. A directed graph has edges with a direction, which indicates a one-way relationship and the edge can only be traversed in a single direction. Instead an Undirected graph has edges without a direction, indicating a two-way relationship, and the edge can be traversed in both directions.

The graph edges have weights, which indicate certain properties of the edge, such as the tie strength of a connection A graph is a weighted graph when each edge has an associated weight.

From the Social Network point of view, a Social Network graph is a graph where the nodes represent the actors in the Social Network (e.g. people), and the edges represent social connections between them. For instance, Facebook is described with an undirected graph[16] since the friendship is bidirectional. Indeed, a friendship request is sent from a user to another and the edge is created only with the second user accepting the request. Instead Twitter is a directed graph because the meaning of a relationship is “to follow” someone. Facebook is the main representative of a Social Network model, instead Twitter is classified as a Social Media [46].

[16] Facebook was created with a social network model, which has been represented with an undirected graph. Last updates have changed the model by introducing an option to follow a person's profile of a page without being friends.

5.1.1 Graph properties

As described above, a social graph is commonly used to represent an Online Social Network in order to study the properties of the network. An important choice concerns the graph model used to represent a Social Network.

In a Social Network, actors are connected by edges, which describe ties. A tie can encapsulate various type of connections, such as *Relations* and *Interactions* [37]. The type of a tie in our graph is *Relation*, which means that an edge between a and b exists when a social relation exists between them. An Interaction is the outcome of the Relation, and for this reason, interactions are considered an information of the edge.

An important point is related to the social relationships concerning the characteristics of a relationship, such as the symmetry [47]. Symmetry means that two actors in a dyad reciprocate a tie. In a Social Network, this property determines how users can interact.

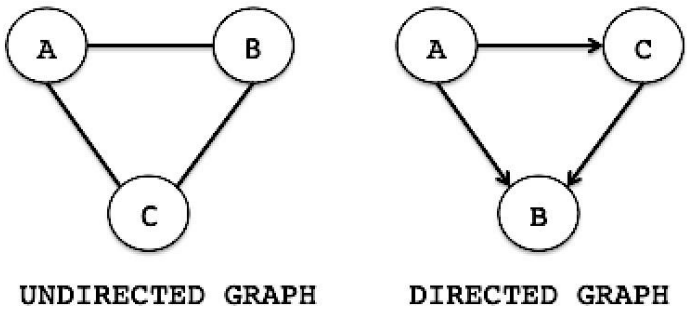

Figure 6. Graph models: undirected and directed

As described in Section 4.2, the HELIOS social network has different actors and different types of actions, and one of the main issues is to define the tie symmetry, which consists in the choice of the graph model: directed or undirected.

Both models have several pros and cons, described below:

- Directed Graph. A directed graph is a graph where all the edges are directed from one vertex to another. One of the main problems of a directed graph is the privacy and security issue because, with the general settings, data are visible after the “following” action. The visibility of data could be reduced by introducing a higher level of privacy, for example with specific privacy policies or with specific rules, but usually the common user is not able because he/she has not a good privacy skill. Moreover, with a one-way relationship, a user cannot know who are his/her followers and he/she cannot provide specific privacy policies for them, and the information diffusion is an issue because when a user generates an event (post, comment, etc.) is not able to trigger the “following” list of users because it does not know them. A positive point of this model is that the connection between two nodes is very fast because is a one-way relationship, and there is no need of a relationships request.

- Undirected Graph. An undirected graph is a graph where all the edges are bidirectional. In a Social Network this graph models a two-ways Relation, and data are visible only after that both nodes exhibit the intention to be friends, this could affect the communication in ad-hoc networks. The definition of privacy policies can exploit the friend list because each node knows the list of friends. One of the main important issues is that the relationship establishment is slow because the model requires a relationship request which must be accepted.

By analysing the characteristics of both graph models and by considering both the meaning of an edge in our Social Network model and the privacy issues, we decide to model the HELIOS Social Network as an undirected graph. In detail, an edge in the HELIOS Social Network means that two nodes have a Relation (not an Interaction), as explained before. The Relation is two-way, and nodes are aware of any Relation they have.

Our scenario is implemented by a multi-layer network, where each layer is represented by an ego network model. To represent the ego network model at each layer, we use an undirected graph G= (V, E, L), where V is the set comprising the ego and heterogeneous alters (people, sensors, etc.), E is the set of relationships between the ego and its alters and between the alters, as the definition of Ego Network suggests, and L is a vector containing social information that characterizes the relationship between the ego and that specific alter. For this reason, the graph is weighted and the weight of an edge between an ego node e and an alter a is the strength of the tie occurring between them. Other important characteristics of the relationships, such as the trust score of an alter, are stored by the ego.

The strength of a tie is a quantifiable property that characterises the link between two nodes [48]. An initial problem encountered in designing measures of tie strength is that it has never been given a precise conceptual definition. A possible definition has been given in [49] by considering a set of features. Indeed, authors define the tie strength has a combination of different parameters: the amount of time, the emotional intensity, the intimacy, and reciprocal services, which characterize the tie.

For measuring tie-strength several indicators have been taken into account. Many researchers measure tie strength by considering the four tie strength dimensions proposed in [49]: amount of time, intimacy, emotional, intensity, and reciprocal services. Moreover, several researchers identified other indicators, such as structural variable (such as interaction activity), emotional support variables, and social distance variables.

Indicators are elements, which determine the strength of a relationship. Usually, they are listed with predictors, which are elements of the relationships that may influence the nature of the tie, such as the number of mutual friends for Facebook [50].

Each ego node measures the tie strength of each relationship between the ego and an alter, and the weight is added to the edge connecting the two nodes. The computation of the tie strength considers the set of indicators proposed in literature [49].

5.1.2 A simple scenario

Figure 7 depicts an example of our Contextual Ego Network. Node “E” (in blue) is the ego-node that the example revolves around.

Nodes 1-11 (in red) are people-nodes and Nodes A, B (in green) are object-nodes. The Contextual Ego Network is a multilayer Ego Network where each layer contains a set of alters, and a mapping function provides the information about alters belonging to more than one set. In the instance depicted in Figure 7, the ego-node appears in three different contexts, and in each context the ego is identified with a different label. The same for the alters. Contacts in Context 1 are work-related nodes (colleagues and smart-apparatus) that form an established network, Context 2 is an ad-hoc network that the ego-node can interact with. With some nodes the connections are already established (e.g. regular customers), some are options for making new connections. Nodes 9b,10b,11b that are present in Context 2, are connections that also appear to be established

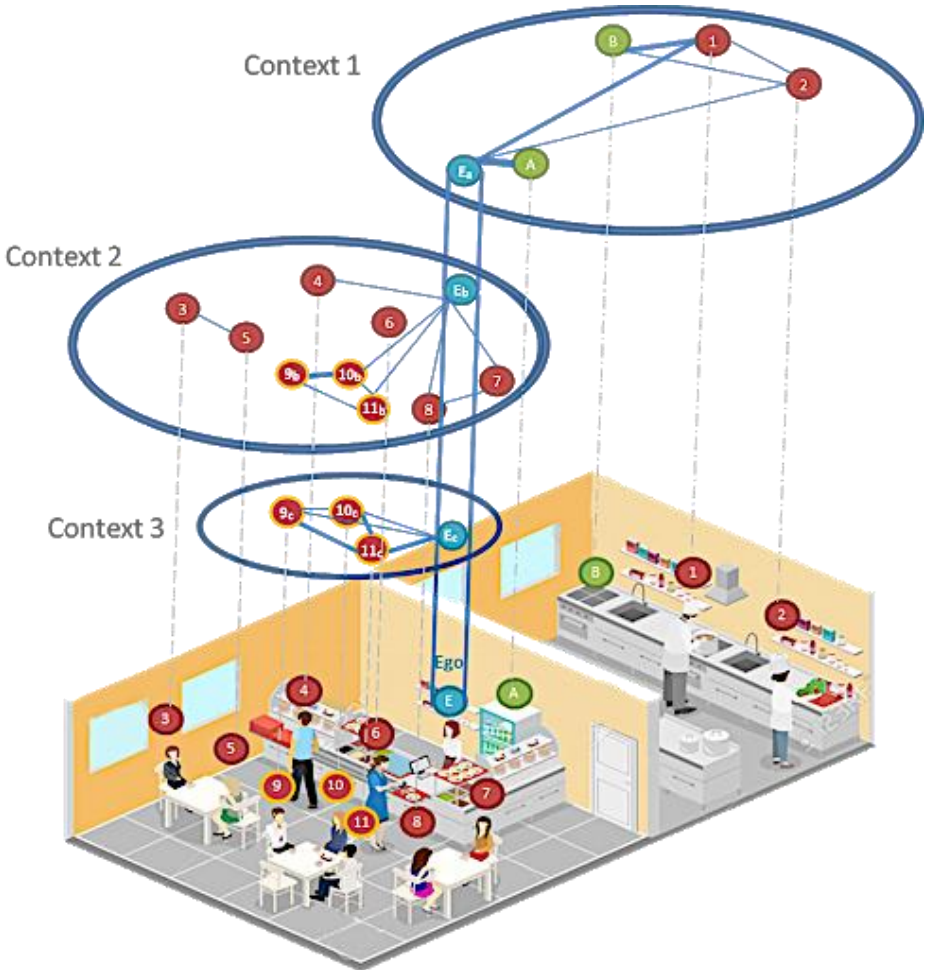

Figure 7. A simple scenario with three contexts

connections in a different context. They could, for instance, be friends that the ego shares a hobby with (Context 3) that also sometimes happen to visit the cafe where the ego works at. The thickness of an edge between two nodes in Figure 7, represents the varying levels of trust at a given context

(the thicker the line, the higher the trust). For instance, in Context 3, all involved nodes appear to have higher levels of trust with Node-11. This would be the case if, for instance, Node-11c is the expert in this hobby-related context (e.g. the yoga instructor at a yoga class that they all participate in). The trust levels, however, between the same nodes in Context 2 (coffee house) shift, based on relationships evaluated on the context's premise.

An important feature of the Contextual Ego Network is the capability to be close to other users according to user profile similarity, and the capability to maintain these contacts in different contexts according to the real life of the user. The Contextual Ego Network is focused on the trust concept, and it guarantees a certain level of trust in each layer by checking each social contact to verify the trust model used in HELIOS.

6 Contextual Ego Network and Trust

From a computational perspective, the topic of trust has been an active area of research in recent years with mathematical models for trust becoming essential for decision making online [51]. It should be noted that the concept of trust is an aggregated and multidimensional construct, built through repeated interaction between entities through time. Most types and bases of trust focus on an individual's decision to trust and on the process under which trust emerges. Thus, among a variety of factors, the relative importance of trust is dependent on the complexity and the context of action. Objects of trust furthermore might include the participants, the underlying technology (application), sites themselves. Measures of trust for users can include competence, ability, integrity honesty and compassion, all accounting for the intentions and behaviours of the actors.

A trust value can represent different categories - system-trust -based on perceived properties or reliance on the system (between organizational and institutional positions), interpersonal trust - agent is directly trusting another entity (within a context-specific environment), general attitudes of trust and trust in information resources related to the application of networks, information resources, software engineering, etc.

The modelling of trust requires the accommodation of a constant context change as well as the need to appropriately evaluate the subjective value of trust given to each node through the different experiences of each node. Furthermore, the trust value would be always represented as incomplete due to the dynamic nature of the connections. Trust is also non-transitive and asymmetric, with the latter being that trust is not identical in both directions between two entities.

For HELIOS, trust carries a multidimensional information about a certain relationship presented as a label of the relationship, turning trust values to highly context-dependent, meaning for each context and actors' interactions a different value of trust will be computed.

Two important measures need to be considered: trustworthiness, and confidence. Trustworthiness is composed of an aggregation of multiple evaluations towards the entities, regarding specific context, with each measurement being a piece of data, contributing to the trustworthiness evaluation. Confidence, on the other hand, is the level of certainty in respect to the trustworthiness evaluation, since confidence varies depending on the number and type of measurements used.

Table 5. Parameters of a possible Trust Model

Type	**Value**
Time	time frames in which other dimensions contribute to trust
Interaction	frequency of interactions
Type of Interaction	positive/negative
Distance	proximity of the device/node
Information Truthfulness	percentage of correct information produced by sensors/devices

Context based	context specific trust - different values for each context
Transactional	every transaction between agents in a particular context is stored and a reputation is calculated on the ratio of the number of successful transactions between agents and the total number of transactions [52]

We foresee that a good trust model will be a combination of the number of interactions, type of interactions (positive vs negative), how the user is understood by the community (trustworthy vs untrustworthy, fake vs real, influencer vs spammer user), and other parameters. The model will also include the fact that trust is evolving through time, according to human behaviour and recent activities. Trust Value in HELIOS will be computed initially based on the types and their values from the table above. A user, in this case, needs to utilize contextual information of the situation as well as the available data to assign a trust measure. Importantly at later stages of HELIOS, more measurement types can be added to the trust score in order to improve accuracy. As such we evaluate a multidimensional situation - where trust has different values depending on the context.

Since this is constantly evolving value, it will be between 0 and 1, with the former representing the absence of trust and 1 being complete trust. Trust in HELIOS, based on the frequency and recency of the interactions, would decay by a certain factor.

Since trust is based on the context, the level of trust will vary depending on the situation, considering the participating entities are the same [53]. For a specific implementation, trust is typically based on factors described in the table and based on context. The score can reflect integration of new and older trust scores of the entity. It can moreover be accumulated directly or indirectly. The former considers only the previous direct interactions of things with each other. The indirect way computes the previous direct interactions as well as those of the neighbouring entities. Although there is an incomplete transitivity of online trust, the ability to see a person's extended network connections could also provide a useful tool for inferences about the trustworthiness of other users.

From a computational standpoint, the various contexts of trust and many stakeholders assure that each node is aware of only a small part of its neighbours and the ones encountered during the process of communication. Thus, the computation of trust requires a way to assess nodes that do not know each other and develop a value that is helpful for future interactions

While most services utilize only a limited trust value functionality, HELIOS relies on the multidimensional computation of trust values depending on context and interactions. This means that, based on a trust value, one entity can share varying levels of personal data. For example, an initial trust score could provide access only to public or minimal information content from another user. With the interactions and constant recalculation of trust, based on certain thresholds, nodes can access increasing amount of information about the nodes they interact with. Decrease of trust value on the other hand, will automatically limit the access to personal data (since the level will fall below a threshold required to access the data pools).

In HELIOS, if an alter is added to the ego network for the first time, it requires an initial computation of trust value. We will assign an initial small value of trust, for nodes to access minimally set of available information for each node. Moreover, with the small initial trust values we can begin the evaluation of the relation between two nodes. Then trust values can be evaluated based on the interaction of the nodes and therefore can decrease or increase.

Algorithms for trust propagation, such as the one proposed in [54] could be used to ascertain a continuous trust value. The algorithm is designed to search for the shortest path between source and target and only those shortest paths are accepted to be available for trust propagation. A limitation of the algorithm is the inability to use longer trustworthy paths. Improvement of the algorithm is the one advanced by Massa et al. [55], known as mole trust, which infers trust by depth-first graph walking algorithm, accounting for distance within trust propagation, although restrictions are placed on the total length of those paths.

A homophily-based approach as the one examined by Kim et al. [56] includes homophily values as well as expertise-based values in order to improve the density of trust networks. Another possible algorithm to be tested or improved within HELIOS is the domain-aware trust network in heterogeneous networks, which is based on multigraph theory [57]. The algorithm describes complex multiple trust relationships between users but also includes a domain-aware trust metric to measure the degree of trust between the nodes, considering their domain-aware influence within a heterogeneous network.

Within the P2P nature of HELIOS, one can also utilize the technique presented in [58], called Eigentrust which measures peer reputation within a P2P network. Assuming the transitive nature of trust scores, then the trust score *ts(i, j)* between two peers p_i and p_j is equal to the total number of times p_i downloads a real file, subtracting the number of times p_i downloaded a fake one from p_j. Then normalization is performed, and the score is recorded in a trust matrix **T**. Finally, the reputation of peer p_i is the i^{th} component which leads to eigenvector *x* of **T**. A structural trust approach is the power trust algorithm [59], which is mainly used for P2P networks. The model exploits the observation that much of the feedback is coming from several “power” nodes in order to develop a robust and scalable trust model. In the algorithms, the peers evaluate each interaction to compute the local trust values. The global trust value is accessed through random walks aggregating the local trust values. After the identification of the power peers by means of the reputation values, those peers are subsequently used in the look-ahead random walk (Markov chain update of the global trust values).

7 Conclusions

This deliverable (named D4.1 “Heterogeneous Social Network Graph topology and lifecycle”) described the characteristics of the P2P Social Overlay adopted in HELIOS, named Heterogeneous Social Network Graph, and in particular the definition of the lifecycle of the local view of each user, named Contextual Ego Network.

In this deliverable we analysed different distributed technologies for Distributed Online Social Networks starting from the general P2P approaches used to implement DOSNs, then focusing on approaches where mobility is a key property (MOSNs), and finally, analysing the recent approaches exploiting blockchain technology (BOSNs).

Having studied the limitations of the proposed approaches, HELIOS aims to introduce a Decentralized Social Network platform that overcomes the current limitations and it will address the dynamic nature of human-to-human and human-to-object interaction.

We proposed the Heterogeneous Social Network Graph as the HELIOS P2P Overlay Network, which is represented with an undirected graph, where an edge between two nodes means that they have a social relation.

The local view of each node is modelled by exploiting the ego network social model and multi-layer networks. Indeed, the P2P local view of each node takes into account the different daily moments of the user, which are represented by contexts. In HELIOS, the local view is named Contextual Ego Network and it is modelled by a multilayers network, where each layer is modelled by an ego network and it represents a specific user’s context.

Furthermore, the deliverable analysed the profile of an HELIOS user by examining different models which can be used to define the profile attributes, and finally, the relation between trust and the Contextual Ego Network is described.

However, the lifecycle of the Contextual Ego Network structure, described in this document, will takes into account further important characteristics which will be delivered along with the upcoming reports that are to be produced and delivered within the activities of WP4.

References

[1] A. Datta, S. Buchegger, L.-H. Vu, T. Strufe, and K. Rzadca, “Decentralized online social networks,” in *Handbook of Social Network Technologies and Applications*, Springer, 2010, pp. 349–378.

[2] B. Guidi, A. Michienzi, K. Koidl, and K. Kapanova, “A multilayer social overlay for new generation DOSNs,” 2019, p. 6.

[3] M. Conti and A. Passarella, “The Internet of People: A human and data-centric paradigm for the Next Generation Internet,” *COMCOM*, 2018.

[4] D. M. Boyd and N. B. Ellison, “Social network sites: Definition, history, and scholarship,” *J. Comput. Commun.*, vol. 13, no. 1, pp. 210–230, 2007.

[5] T. Maqsood, O. Khalid, R. Irfan, S. A. Madani, and S. U. Khan, “Scalability Issues in Online Social Networks,” *ACM Comput. Surv.*, 2016.

[6] C. Maurieni, *Facebook is Deception (Volume One)*. WSIC EBooks Ltd, 2012.

[7] H. Jones and H. Soltren, “Facebook: Threats to Privacy,” *Proj. MAC MIT Proj. Math. Comput.*, vol. 1, pp. 1–76, 2005.

[8] K. Fall, “A delay-tolerant network architecture for challenged internets,” in *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications - SIGCOMM ’03*, 2003.

[9] B. Guidi, M. Conti, A. Passarella, and L. Ricci, “Managing social contents in Decentralized Online Social Networks: A survey,” *Online Soc. Networks Media*, vol. 7, pp. 12–29, 2018.

[10] G. Mega, A. Montresor, and G. P. Picco, “Efficient dissemination in decentralized social networks.,” in *Peer-to-Peer Computing*, 2011, pp. 338–347.

[11] M. Magnani and L. Rossi, “The ml-model for multi-layer social networks,” in *ASONAM*, 2011, pp. 5–12.

[12] A. Bielenberg, L. Helm, A. Gentilucci, D. Stefanescu, and H. Zhang, “The growth of diaspora-a decentralized online social network in the wild,” in *Computer Communications Workshops (INFOCOM WKSHPS), 2012 IEEE Conference on*, 2012, pp. 13–18.

[13] M. Zignani, S. Gaito, and G. P. Rossi, “Follow the ‘mastodon’: Structure and evolution of a decentralized online social network,” in *12th International AAAI Conference on Web and Social Media, ICWSM 2018*, 2017.

[14] L. A. Cutillo, R. Molva, and T. Strufe, “Safebook: A privacy-preserving online social network leveraging on real-life trust,” *Comm. Mag.*, vol. 47, no. 12, pp. 94–101, Dec. 2009.

[15] S. Buchegger, D. Schiöberg, L.-H. Vu, and A. Datta, “PeerSoN: P2P social networking: early experiences and insights,” in *The Second ACM EuroSys Workshop on Social Network Systems*, 2009, pp. 46–52.

[16] S. Buchegger, D. Schioberg, L. H. Vu, and A. Datta, “Implementing a P2P Social Network - Early Experiences and Insights from PeerSoN,” in *Second ACM Workshop on Social Network*

Systems (Co-located with EuroSys 2009), 2009.

[17] K. Graffi, C. Gross, D. Stingl, D. Hartung, A. Kovacevic, and R. Steinmetz, “LifeSocial. KOM: A secure and P2P-based solution for online social networks,” in *IEEE CCNC, 2011*, 2011, pp. 554–558.

[18] P. Druschel, “Past: A large-scale, persistent peer-to-peer storage utility,” in *HotOS VIII*, 2001, pp. 75–80.

[19] A. Rowstron and P. Druschel, “Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems,” in *IFIP/ACM Middleware 2001*, 2001, pp. 329–350.

[20] F. Tegeler, D. Koll, and X. Fu, “Gemstone: Empowering decentralized social networking with high data availability,” in *GLOBECOM - IEEE Global Telecommunications Conference*, 2011.

[21] A. Sahai and B. Waters, “Fuzzy Identity-Based Encryption,” 2010.

[22] S. Nilizadeh, S. Jahid, P. Mittal, N. Borisov, and A. Kapadia, “Cachet: a decentralized architecture for privacy preserving social networking with caching,” in *CoNEXT ’12*, 2012, pp. 337–348.

[23] N. Kourtellis, J. Finnis, P. Anderson, J. Blackburn, C. Borcea, and A. Iamnitchi, “Prometheus: User-controlled P2P social data management for socially-aware applications,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2010.

[24] R. Narendula, A. Papaioannou, and K. Aberer, “A Decentralized Online Social Network with Efficient User-Driven Replication,” in *IEEE SocialCom*, 2012.

[25] B. Guidi, T. Amft, A. De Salve, K. Graffi, and L. Ricci, “DiDuSoNet: A P2P architecture for distributed Dunbar-based social networks,” *Peer-to-Peer Netw. Appl.*, vol. 9, no. 6, 2016.

[26] R. I. M. Dunbar, “The social brain hypothesis,” *Evol. Anthropol. Issues, News, Rev.*, vol. 6, no. 5, pp. 178–190, 1998.

[27] A. Shakimov *et al.*, “Vis-a-vis: Privacy-preserving online social networking via virtual individual servers,” in *COMSNETS 2011*, 2011, pp. 1–10.

[28] M. Durr, M. Maier, and F. Dorfmeister, “Vegas--A Secure and Privacy-Preserving Peer-to-Peer Online Social Network,” in *Privacy, Security, Risk and Trust (PASSAT)*, 2012, pp. 868–874.

[29] R. Sharma and A. Datta, “SuperNova: Super-peers based architecture for decentralized online social networks,” in *COMSNETS*, 2012, pp. 1–10.

[30] N. Kayastha, D. Niyato, P. Wang, and E. Hossain, “Applications, architectures, and protocol design issues for mobile social networks: A survey,” *Proc. IEEE*, vol. 99, no. 12, pp. 2130–2158, 2011.

[31] P. Bellavista, R. Montanari, and S. K. Das, “Mobile social networking middleware: A survey,” *Pervasive Mob. Comput.*, vol. 9, no. 4, pp. 437–453, 2013.

[32] A.-K. Pietiläinen, E. Oliver, J. LeBrun, G. Varghese, and C. Diot, “MobiClique: Middleware for Mobile Social Networking,” in *WOSN ’09*, pp. 49–54.

[33] E. Sarigöl, O. Riva, P. Stuedi, and G. Alonso, “Enabling Social Networking in Ad Hoc Networks of Mobile Phones,” *Proc. VLDB Endow.*, vol. 2, no. 2, pp. 1634–1637, Aug. 2009.

[34] T. H. Davenport and J. C. Beck, *The attention economy: Understanding the new currency of business*. Harvard Business Press, 2001.

[35] A. De Salve, P. Mori, and L. Ricci, “A survey on privacy in decentralized online social networks,” *Computer Science Review*. 2018.

[36] M. Conti, A. D. Salve, B. Guidi, F. Pitto, and L. Ricci, *Trusted dynamic storage for dunbar-based P2P online social networks*, vol. 8841. 2014.

[37] A. De Salve, B. Guidi, P. Mori, L. Ricci, and V. Ambriola, “Privacy and temporal aware allocation of data in decentralized online social networks,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017, vol. 10232 LNCS, pp. 237–251.

[38] P. V Marsden, “Egocentric and sociocentric measures of network centrality,” *Soc. Networks*, vol. 24, no. 4, pp. 407–422, 2002.

[39] S. R. Kruk and S. Decker, “Semantic social collaborative filtering with FOAFRealm,” in *CEUR Workshop Proceedings*, 2005.

[40] S. R. Kruk, S. Grzonkowski, A. Gzella, T. Woroniecki, and H.-C. Choi, “D-FOAF: Distributed Identity Management with Access Rights Delegation,” in *The Semantic Web -- ASWC 2006*, 2006, pp. 140–154.

[41] S. Boccaletti *et al.*, “The structure and dynamics of multilayer networks,” *Phys. Rep.*, vol. 544, no. 1, pp. 1–122, 2014.

[42] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, “Multilayer networks,” *J. complex networks*, vol. 2, no. 3, pp. 203–271, 2014.

[43] P. Bródka and P. Kazienko, “Multilayered Social Networks,” in *Encyclopedia of Social Network Analysis and Mining*, R. Alhajj and J. Rokne, Eds. New York, NY: Springer New York, 2014, pp. 998–1013.

[44] M. De Domenico *et al.*, “Mathematical formulation of multilayer networks,” *Phys. Rev. X*, 2014.

[45] M. Berlingerio, M. Coscia, F. Giannotti, A. Monreale, and D. Pedreschi, “Foundations of multidimensional network analysis,” in *Proceedings - 2011 International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2011*, 2011.

[46] H. Kwak, C. Lee, H. Park, and S. Moon, “What is Twitter, a social network or a news media?,” in *Proceedings of the 19th international conference on World wide web*, 2010, pp. 591–600.

[47] S. Wasserman, K. Faust, and others, *Social network analysis: Methods and applications*, vol. 8. Cambridge university press, 1994.

[48] U. K. Petróczi, Andrea (Kingston University, Kingston-upon-Thames, H. Nepusz, Tamás (Hungarian Academy of Sciences, Budapest, and H. Bazsó, Fülöp (Hungarian Academy of Sciences, Budapest, “Measuring tie-strength in virtual social networks,” *Connections*, 2007.

[49] M. S. Granovetter, “Granovetter - The Strength of Weak Ties.pdf,” *Am. J. Sociol.*, 1973.

[50] E. Gilbert and K. Karahalios, “Predicting tie strength with social media,” in *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*, 2009.

[51] K. Kapanova and K. Koidl, “Towards a model of interpersonal trust in Social Media Applications,” in *EAI International Conference on Smart Objects and Technologies for Social Good (GoodTechs ’19)*, 2019.

[52] L. Mui, M. Mohtashemi, and A. Halberstadt, “A computational model of trust and reputation,” in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2002.

[53] R. Iqbal, T. A. Butt, M. Afzaal, and K. Salah, “Trust management in social Internet of vehicles: Factors, challenges, blockchain, and fog solutions,” *Int. J. Distrib. Sens. Networks*, 2019.

[54] J. A. Golbeck, “Computing and Applying Trust in Web-based Social Networks,” University of Maryland at College Park, College Park, MD, USA, 2005.

[55] P. Massa and P. Avesani, “Controversial users demand local trust metrics: an experimental study on epinions.com,” *AAAI*. 2005.

[56] Y. A. Kim, “An enhanced trust propagation approach with expertise and homophily-based trust networks,” *Knowledge-Based Syst.*, 2015.

[57] C. Jiang, S. Liu, Z. Lin, G. Zhao, R. Duan, and K. Liang, “Domain-aware trust network extraction for trust propagation in large-scale heterogeneous trust networks,” *Knowledge-Based Syst.*, 2016.

[58] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, “The Eigentrust algorithm for reputation management in P2P networks,” in *Proceedings of the twelfth international conference on World Wide Web - WWW ’03*, 2003.

[59] R. Zhou and K. Hwang, “PowerTrust: A robust and scalable reputation system for trusted peer-to-peer computing,” *IEEE Trans. Parallel Distrib. Syst.*, 2007.