IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 4 May 2023, accepted 16 May 2023, date of publication 29 May 2023, date of current version 6 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3281194

RESEARCH ARTICLE

SentiTrust: A New Trust Model for Decentralized Online Social Media

BARBARA GUIDI[1], ANDREA MICHIENZI[1], LAURA RICCI[1], FABRIZIO BAIARDI[1], LUCÍA GÓMEZ-ZARAGOZÁ[2], LUCÍA A. CARRASCO-RIBELLES[2], (Student Member, IEEE), AND JAVIER MARÍN-MORALES[2], (Member, IEEE)

[1]Dipartimento di Informatica, Università di Pisa, 56126 Pisa, Italy
[2]Instituto Universitario de Investigación en Tecnología Centrada en el Ser Humano, Universitat Politècnica de València, 46022 Valéncia, Spain

Corresponding author: Barbara Guidi (guidi@di.unipi.it)

This work was supported in part by the European Commission through the Project HELIOS under Contract H2020-825585.

ABSTRACT Online Social Media (OSM) are dominating the wide range of Internet services. Due to their vast audience, it is crucial to evaluate the interpersonal trust among OSM users that can identify reliable sources of information, the meaningfulness of a relationship, or the trustworthiness of other users. SentiTrust is an innovative trust model for Decentralized Online Social Networks that is based on AI-powered Sentiment Analysis. It enriches the trust definition by exploiting important features that are enabled because of the adoption of Social Media through mobile devices. The model can be easily extended and customized according to the scenario of interest. The sentiment analysis component has been tested by involving 30 participants who completed several guided tasks using a social media application while their electrodermal activity and rate responses were measured. The results suggest that low arousal states are related to receiving happy faces and to sending more messages per minute. Furthermore, positive interactions result in shorter interactions and multimedia exchanges.

INDEX TERMS Decentralized online social networks, online social networks, sentiment analysis, statistical learning, trust.

I. INTRODUCTION

Throughout the years, Online Social Media (OSM) had a massive impact on the way people socialize. Indeed they served as a medium to remove geographical barriers and helped the spread of information at a faster-than-ever pace. However, OSM platforms are undergoing turbulent changes and several platforms are increasingly turning their attention to decentralized solutions and technologies [1], [2], [3]. The drawbacks of centralized OSM include performance scalability issues, increasing maintenance costs, lack of geographical locality and, most important, privacy issues [1]. Privacy issues stressed by events like the Cambridge Analytica scandal fueled a wave of innovation that gave birth to the so-called Distributed Online Social Network (DOSN). A DOSN can be understood as an OSM implemented on a distributed information management platform, such as a network of trusted servers, a P2P system or an opportunistic network.

In this ever-changing scenario, trust represents a crucial concept to be applied to OSM, because it is an essential tool for users to identify meaningful relationships. The people involved in a trustful relationship are more likely to have positive and frequent communication. Moreover, anyone is more willing to share some social or personal information and content with trusted people, rather than with untrusted ones. Therefore, trust also creates both preferential paths for information flows and communities where participants do not fear that their privacy will be violated [4].

However, the evaluation of trust in a social relationship is particularly challenging. To begin with, trust is applied to numerous scenarios. As an example, in OSM trust has been widely investigated and used for recommendation systems and access control. The concept of trust has also been

The associate editor coordinating the review of this manuscript and approving it for publication was Giacomo Fiumara.

extensively investigated in more general settings where a set of actors interact with each other to reach a goal [4]. As a result, there is no unique, widespread, definition of trust independent of the application scenario. In general, trust takes the form of a binary bond between two parties: the trustor and the trustee. The former is the one that has a certain level of trust towards a target entity, the trustee, that is the target entity that is trusted. In virtual environments, such as OSNs, trust is a fundamental input of the decision-making process that selects which information to share. Moreover, it also plays a vital role in deciding whether to enable sharing with another user without any verification. On top of that, trust evaluation is particularly challenging for new users, both in terms of whether the newcomers should be trusted by other users, and, vice-versa, how a newcomer can identify trustworthy users. Besides the evaluation of trust among users in OSM, we should also take into account the aforementioned decentralization process of online social services. On one hand, decentralization introduces several challenges related to data storage and availability, privacy, and information diffusion. On the other hand, it offers new opportunities for evaluating social interactions. In particular, the heterogeneity of Internet of Things (IoT) devices in smart environments can help gather more information and supply the most disparate information, such as contextual information (e.g. a smart table in a restaurant) or social information (e.g. a beacon broadcasting upcoming cultural events). This aspect is becoming crucial today as the vast majority of people use smartphones and similar devices to access online social platforms.[1] Another critical challenge is the adoption of Artificial Intelligence (AI) techniques to deliver a better social service. AI has been successfully applied in several research and industrial fields, but the scenario with IoT-powered social networks is innovative and can largely improve the quality of the services the user can access. However, the adoption of traditional AI techniques in this scenario may result in too much energy consumption. Hence, further models and techniques should be developed to exploit the strengths of future-generation social platforms [5].

In this paper, we present SentiTrust, a new trust model designed for Decentralized Social Media, but that can easily be used in OSM as well. SentiTrust employs some social features, including the number of common friends and the evaluation of social interactions, to estimate the trust between the two parties of a relationship. Additionally, it innovates the definition of trust because it exploits sentiment analysis, implemented through AI techniques, a feature that, to the best of our knowledge, is missing in any other trust model. The model is not specifically designed for a single platform and it adopts a general approach which can be tailored to a given social platform. It is implemented in the HELIOS project,[2] whose aim is to define a trust-by-design decentralized social networking platform, which aims at building a contextually aware, heterogeneous, privacy-aware, personal social network [6]. It is also integrated into the TestClient-AutumnApp application[3] that uses SentiTrust to implement an information overload control system. This paper highlights the main characteristics of the model, focusing in particular on the Neuro-Behavioral Module (NBM) that implements the Sentiment Analysis of social interactions among users. We present an in-depth description of the main features of NBM and an overview of the study of the included emotions.

Our contribution can be summarized as:

- We propose SentiTrust, a new model for trust computation that applies AI to evaluate social interactions between pairs of users.
- We tune the model by selecting the features of interest by involving a set of 30 participants that, by interacting with each other, help us identify the most impactful features that define a trustful relationship.

As concerns the benefits of our trust model, we highlight the following.

- SentiTrust is, to the best of our knowledge, the first trust module that explores the benefit of AI to compute trust;
- SentiTrust can be easily customized and extended as per the developers' needs;
- While being designed for Decentralized Online Social Media, SentiTrust can also be applied to other social scenarios;
- This study also provides an experimental evaluation, performed with real participants, that help bootstrap the parameters of the module in a social setting.

The paper is structured as follows. Section II presents an overview of trust models proposed in four scenarios. Section III discusses how to define a trust model in decentralized social media, by considering the trust properties and the various challenges to be faced, and proposes a comparison highlighting the pros and cons of several other models. Section IV describes the features considered by SentiTrust and their parametrized combination. In Section V, we describe how SentiTrust works by explaining the two phases of initialization and update. Section VI presents the selection process of the features of the sentiment analysis, while Section VII presents a possible model configuration in a generic scenario. Finally, Section VIII concludes the paper and points out possible future works.

II. RELATED WORKS

Trust is a multi-disciplinary, multi-faceted concept, and trust definitions have been proposed in distinct fields ranging from psychology to social sciences, and to information technology. In the Computer Science domain, there is no universally accepted definition of trust. For instance, trust can be seen as the expectations built towards another entity based on previous interactions [7], [8]. Other definitions are related to

[1] https://backlinko.com/social-media-users [Accessed: 18-May-2023]

[2] https://github.com/helios-h2020/h.core-TrustManager [Accessed: 18-May-2023]

[3] https://github.com/helios-h2020/h.app-TestClient-AutumnApp [Accessed: 18-May-2023]

the expectation that an agent offering a specific service will behave in a certain way [9], or that it will behave in a positive way towards the observer [10], [11], [12].

A. TRUST VS. REPUTATION

In the literature, the terms "trust" and "reputation" are often used interchangeably, and they refer to the expected behavior of an actor in a system [7], [13], [14], [15], [16]. In practice, this concept usually implies the aggregation of the past experience of other actors to produce a "global" view of the expected behavior of the actor of interest. This paper adopts different definitions for the two terms because they refer to distinct phenomena. Given a system involving a set of actors A, that are interacting to achieve a set of personal or shared goals, we adopt the following definitions of Reputation and Trust [7].

Definition 1 (Reputation): The reputation of an actor a determines how a is seen overall by $A - a$, the set of other actors.

Definition 2 (Trust): The trust of an actor a towards an actor b determines how a sees b, independently of the other actors in A.

Reputation has the following characteristics:

- **Global.** An actor's reputation is globally defined and it is the result of everyone's experience with that actor. It represents how well the actor is perceived in the system.
- **Characterizes an Actor.** As an example, even if the reputation of a given user is outstanding, this does not imply that everyone has had a positive experience with that user.
- **Hard to change.** It is challenging to change the reputation of an actor because it results from a collection of independent parties.

We can clarify the concept of reputation through a simple example. In a file-sharing system, a machine may be deemed of high reputation if it is well known it does not share dangerous files, such as viruses, or wrong files, regardless of the actor that interacts with it.

On the other hand, Trust has the following features [4]:

- **Local.** While several factors can affect trust, the most important one should be local to the trustor and trustee.
- **Describes a relationship.** Being a local characteristic, it describes the relationship between two actors and the trust value of distinct trustors towards the same trustee can differ.
- **Dynamic.** Trust can be quite dynamic as the interactions between the trustor and the trustee have a huge impact on its value.

As for reputation, we propose an example of the trust concept. In a file-sharing system, a machine a may be deemed of high trust by the actor b if b benefits from interacting with a. This may be due to higher transfer rates, double-checking the integrity of files, and so on.

While trust and reputation are related, they are quite different. Indeed, trust focuses more on the personal experience and, as such, is more dynamic, personal, and strictly focused on the relationship between a trustor and a trustee. On the other hand, reputation represents how an actor is viewed in a system even without any direct experience with the actor. The previous discussion shows that our focus should be geared towards the concept of trust, rather than reputation, as we focus on a description of social relationships between users.

B. TRUST MODELS

While distinct trust models for distributed networks have been proposed, there is a lack of models with a specific focus on DOSNs. Works on trust models for decentralized networks can be partitioned into three categories according to the goal of the network they are geared towards. In the remainder of the section, we discuss some of the most relevant trust models in decentralized networks.

1) TRUST MODELS FOR RESOURCE SHARING NETWORKS

When distinct nodes offer the same resource, the problem arises of choosing the best one to fulfill a request. Decentralized resource-sharing networks are characterized by a set of nodes that initially do not know and do not trust each other. Here, trust makes it possible to avoid malicious nodes that may cheat or supply a resource with a quality lower than expected.

P2PRep [17] is a trust model based on the notion of reputation where each node takes a decision by aggregating the opinion of other nodes on service providers.

TACS [18] is a model based on the ant-colony optimization heuristic. It can discover the most covered paths in the network (which correspond to highly trusted connections) and assign a routing preference to them with respect to less covered paths.

Stakhanova et al., describe a model [19] that combines reputation with anomaly detection because any solution that only considers reputation may not detect abnormal or suspicious behavior of peers.

Tang et al. propose a model based on fuzzy theory [20]. The authors consider fuzziness because trust is highly subjective as it is closely tied to the observer criteria.

2) TRUST MODELS FOR GENERAL P2P NETWORKS

Models in this category are designed for general-purpose P2P networks but, because of their flexibility, they may be adapted to distinct contexts, including OSNs. On the other hand, their generality may require some integration of specific features in some scenarios.

In [13], the authors propose a trust model based on Bayesian networks where trust in a peer depends on distinct parameters, each weighted according to the trustor's needs at a given moment.

Reference [16] defines a trust model based on reputation drawn from collective knowledge. Peers can assess other peers according to their interactions, and this guides the computation of trust.

The model proposed in [14] is based on both direct (first-hand) experience and indirect (second-hand) one. It adopts a Bayesian framework to combine the concepts of trust and reputation and design a misbehavior detection system.

3) TRUST MODELS FOR DECENTRALIZED SOCIAL NETWORKS

This category includes models that address social issues directly. Modeling and computing trust in a social decentralized context is much harder than in other ones, mainly because it has to deal with user privacy. Moreover, it is intrinsically complex to define social trust, the parameters to compute it, and what can be classified as "untrustful behavior".

The model in [21] targets mobile ad hoc social networks. To compute trust, the model distinguishes the case where the user can gather enough information from the network, from the one where this information is insufficient. In the first case, the trust of a user p towards an unknown user u is an average of the trust of other users towards u, weighted by their profile similarity with p. In the second case, the trust computation relies on three local parameters: profile similarity, local reputation, and the number of common friends.

Qureshi et al. [22] propose a trust model to create trust communities in P2P mobile social networks. Trust values are expressed as discrete values in the range [0,3]. Trust is computed by aggregating the reputation scores of other nodes. This neglects privacy issues, and the model features do not differ from those defined in other P2P networks.

4) TRUST MODELS FOR CENTRALIZED SOCIAL NETWORKS

We review even the most relevant trust models for centralized online social networks, to understand the features that emerge from a social context and discuss their adoption in a decentralized environment.

STrust [8] is a model to build trust communities in social networks. According to the authors, the most important properties of trust are time and context dependency, while the main parameter for its computation is user behavior. This behavior is expressed in terms of social capital, which consists of interactions among network nodes.

In [15], the authors propose a trust model based on trust ratings. This rating depends on direct behavioral observations and reputation, which aggregates trust ratings of other users, weighted according to the trustworthiness of the users themselves.

The probabilistic trust inference model in [10] is based on two factors: the intimacy degree between users, and the role impact that represents the user's expertise in a specific domain. Trust is computed through a posterior probability estimation based on the Bayes theorem.

In [23], the authors propose SUNNY, a probabilistic trust model based on bayesian networks. It estimates trust by taking into account the concept of confidence, defined as the belief of other nodes that some provided information is correct.

TABLE 1. Summary of the proposals and the features they consider in the model, divided into the four scenarios.

Work	Features
P2P Resource sharing	
P2PRep - Cornelli et al. [17]	Reputation, behaviour as resource provider
TACS - Mármol et al. [18]	Path-based reputation, behaviour as resource provider
Stakhanova et al. [19]	Reputation, anomalous behaviour detection
Tang et al. [20]	Subjective parameter weights, recommendations, experience evaluation
P2P networks	
Wang and Vassileva [13]	Subjective parameter weights
Aberer and Despotovic [16]	Reputation, behaviour evaluation
Buchegger and Boudec [14]	Reputation, behaviour evaluation
Decentralized social	
Li and Li [21]	Profile similarity, interactions evaluation, common friends
Qureshi et al. [22]	Interactions evaluation, reputation
Centralized social	
STrust - Nepal et al. [8]	Popularity engagement context
Yu and Singh [15]	Behavioral observations (interactions evaluation), reputation
Liu et al. [10]	Intimacy degree, expertise in a domain
SUNNY - Golbeck and Kuter [23]	Confidence
Maheswaran et al. [24]	Context, interactions evaluation
Fernandez et al. [25]	Similarity, Context
3SVL - Liu et al. [26]	Previously available trust values
SocialTrust - Caverlee et al. [27]	Context, Profile Similarity, Relationships

Reference [24] defines a multi-contextual model, where trust is a vector of m real numbers, one for each context the two users share. The vector represents the trust distance between the two nodes in the m-dimensional trust space and is updated according to the occurrence of positive/negative interactions between them.

The model in [25] is based on both user similarity and contextuality. Users are represented as a vector embedding and the distance between two vectors bootstraps the trust value. A trust propagation method updates the single trust values.

Three-Valued Subjective Logic (3VSL) [26] is a trust model which evaluates the trustworthiness as true, false, or neutral. The trust between two users is based on five values: belief, distrust, posteriori uncertainty, priori uncertainty, and base rate. This work is heavily based on previously available trust values and on trust propagation.

SocialTrust [27] is a trust model where the trust score of a user depends on the current trust value of a user, its past values, and a coefficient to mitigate malicious nodes that build up good trust ratings in the past.

C. SUMMARY

Table 1 summarizes the reviewed approaches, classified in four scenarios. Overall, we see that the trust computation process often considers reputation. In Resource Sharing networks also anomalous behavior detection is a common feature because nodes are expected to behave consistently. Instead, in P2P networks the evaluation of the behavior is in general related to the mutual benefits of the peer behavior in the network. Moving to social scenarios, contextual relevance is a recurring feature in centralized OSNs. Lastly, the approaches designed for DOSNs neglect reputation, and focus on the analysis of user interactions.

III. DEFINING A TRUST MODEL IN DECENTRALIZED SOCIAL MEDIA

In DOSNs, trust is often computed through reputation, or other features such as profile similarity and common contacts [21], [22]. This neglects the trust feeling created in human relationships as a pair of people interact directly.

To close this gap, we propose SentiTrust, whose main component estimates trust through people's direct social interactions. SentiTrust leverages the power of AI and sentimental analysis to achieve a more in-depth understanding of how a person responds to social interactions. We also show how to include additional optional features that are not mandatory but can enrich the trust evaluation process. The trust computation framework in SentiTrust is not tied to a particular decentralization technology. Hence, the approach can be easily applied to Social Media where the computation of the trust values is centralized.

Our presentation of SentiTrust starts by identifying the inherent properties of trust. The second step identifies the most relevant features a trust model should consider, how real-world human trust can be translated into virtual trust and discusses possible challenges in their adoption. Then, we show how to harness the power of AI to evaluate trust. Lastly, we discuss the challenges to defining a trust model with respect to the scenario of DOSNs.

A. TRUST PROPERTIES

This section highlights the properties of the concept of trust [4] by discussing the features a trust model should consider. In the following, we denote by A, B, and C some users of a social network.

1) FUZZINESS

Trust is not a binary value: there is no such thing as ‘‘either complete trust or complete distrust’’. A user in a social network can trust another one according to a gradient, and the trust value lies in a continuous interval.

2) DYNAMISM

A trust value is extremely volatile, and therefore subject to constant updates that depend on various parameters, such as the positive/negative outcome of interactions.

3) NON-TRANSITIVITY

The fact that A highly trusts B and B highly trusts C does not necessarily imply that A will highly trust C. This is true for most non-social contexts, and it is even more emphasized in social networks where relationships are unique and personal.

4) ASYMMETRY

In a relationship between two people, trust is not necessarily symmetrical. It may be the case that A highly trusts B, while B does not trust A at all. As non-transitivity, also this property generally holds for networks having distinct goals.

5) SUBJECTIVITY

The mental process that drives the trust assessment is governed by personal experience, so trust computation is mostly dependent on the subjective experience of the relationship.

6) CONTEXT DEPENDENCY

Trust depends on the context of the relationship and the background of the actors involved, and this is primarily due to the fact that human relationships are contextual as well. Trust can depend on many contextual factors including at least who are the subjects involved, on what, when, and where [28]. As an example, two colleagues may trust each other in the work context, but not in a politics context.

B. TRUST IN SOCIAL RELATIONSHIPS

As outlined in Section II, trust computation can consider several features and some are shared among multiple approaches and scenarios. This Section briefly discusses the most common features among approaches for trust computation.

The most recurring feature in the literature is *Reputation* [14], [15], [16], [17], [18], [19], [22]. This is an important feature that can be used also in the scenario of social applications. Since it is a global feature that describes an actor in a system, it can be more easily defined and computed in centralized OSNs. In DOSNs, the problem changes because reaching and maintaining a consensus is extremely challenging in a decentralized network [29] and requires copious resources. This ultimately leads us to discard this feature in a trust model for DOSNs.

Behavior evaluation that measures the expected behavior of the actor within the system is one of the most recurring features [14], [15], [16], [17], [18], [19]. Distinct scenarios share this feature. As an example, file-sharing networks use it to model whether an actor behaves according to the defined protocol, or if an actor has some unnecessary bias towards other actors. This feature can be hardly applied in the social setting of DOSNs because of both the extreme subjectivity of the trust notion and the lack of explicit social rules that define social trust and that can be easily transformed into code.

While the direct evaluation of user behavior is unfeasible, it is possible to evaluate its effects, in terms of *Personal Experience* [13], [20], [21], [22]. This feature evaluates the interactions between two actors and it is more common in social scenarios where it is connected to human behavior and its complexity. As an example, the same social interaction between two people can be evaluated oppositely, according to factors such as the stance of the two people, their mood and other facial expressions, and the tone of voice. This is one of the most crucial features in a social setting, because people naturally associate a high trust with individuals with which they frequently interact positively [15], [21], [22], [24], [30], and it enables fine-grained characterization of social relationship between two people. Given the extremely challenging nature of the task, we plan to leverage the power of AI to evaluate personal social interactions.

Lastly, a recurring feature is *Contextuality*, which considers the context where trustworthiness is evaluated [8], [24], [25]. It is more common in centralized OSNs and it can be understood as the ability to differentiate relationships and interactions according to the context where they happen. As an example, the trust between two coworkers may be high in a work context, but low in other contexts.

Because of their impact and relevance to computing trust in DOSNs, we consider the evaluation of *Personal Experience* as the cornerstone for trust computation. Other important features could be used as a support, if available, for enriching the trust computation, such as the *Common Friends*, or for trust initialization, such as the *Profile Similarity*. Additionally, *Contextuality* is another important feature that lets us evaluate trustworthiness with the due fine-grain level.

C. HOW TO EXPLOIT ARTIFICIAL INTELLIGENCE

When evaluating social interactions, AI can be applied to pieces of content produced by people. In this paper, we focus on textual and graphic pieces of content as they represent the vast majority of the content that users exchange[4,5]. In the following, we detail two analyses and, in particular, how they deduce the sentiment humans perceive in virtual interactions.

1) TEXT ANALYSIS

There are several approaches to analyze the emotional burden on a text, such as through the use of Lexicons.[6] Lexicons are extensive datasets that relate each included word with a description of its emotional burden. The emotional burden of a message sums up the emotional burden of its words. Other analyses of the emotional burden apply Machine Learning (ML) models to predict whether the text's valence is positive or negative, but the text might need to be cleaned before being analyzed. ML models produce more robust results than Lexicons, as they are usually trained with a larger corpus of text and are more likely to capture the various nuances of meaning in human language.

2) IMAGE ANALYSIS

The analysis of human emotional expressions in pictures can greatly help decipher the interaction between two people, under the assumption that the facial expression embedded conveys information about the sender's emotional status [31]. Several works discuss emotion recognition in images, and most of them are based on deep learning that has proved to be a good approach to face detection and classification of emotions. Among existing works, the models in [32] show good classification properties and have been made available online.[7]

D. CHALLENGES IN DOSNs

Trust plays an important role in multiple aspects of a social network, including how other users are perceived, whether some pieces of information are considered reliable, private, and so on. Trust models in centralized OSNs often acquire and exploit recommendations and reputation. This is simple in centralized contexts because trust can be computed by the service providers rather than locally by each client. This is an important advantage in terms of both storage and computational complexity. There are also privacy advantages because the trust and reputation information is shared with the central authority only.

The trust model of a DOSN is more complex due to the decentralized nature of the network. This makes privacy, information gathering, and computational power usage more complex. In DOSNs, the assignment of a score becomes challenging because users most likely do not want to share their trust information with other peers. Moreover, users usually have limited computational resources available on their devices and the social network client needs to be lightweight so as to not be a burden on the devices. Detecting malicious nodes that are likely to deceive in the future is also an aspect to take into account.

These challenges show that trust models for DOSNs should consider distinct information and be based on distinct features from standard OSNs. However, trust features for centralized OSNs and for DOSNs are not disjoint: some features of trust models for centralized OSNs are still valid and determine feasible computations even in decentralized environments. Instead, other ones and, in particular, those that rely on information received from all over the network have to be discarded or heavily adapted to a decentralized context.

E. COMPARISON

Trust is a deeply investigated concept, and several works in the literature propose trust algorithms. However, most of these works do not detail the computation of trust values and/or they use already available trusted graphs without any detail about trust computation. Furthermore, several works are related to trust, but in alternative scenarios that are not easily comparable with others due to the different characteristics of each research field. As an example, works on trust in P2P networks are more related to the feedbacks from resource sharing. For this reason, we propose a comparison between

[4] https://blog.whatsapp.com/making-voice-messages-better [Accessed 18-May-2023]

[5] https://www.statista.com/statistics/258743/daily-mobile-message-volume-of-whatsapp-messenger [Accessed 18-May-2023]

[6] https://saifmohammad.com/WebPages/lexicons.html [Accessed 18-May-2023]

[7] Available: https://github.com/rishabhjainps/Facial-Expression-Recognition [Accessed: 18-May-2023]

our model and a set of other trust models, chosen among the approaches presented in Table 1 in the fields of both Decentralized and Centralized Social Networks. In detail, we identify the strengths and weaknesses of trust models for DOSNs in the form of the trust properties listed in Section III-A, whether the computation can be centralized, decentralized or both, their computational model, and show a comparison with our SentiTrust model in Table 2.

A strength of all the trust models in the comparison is that they consider fuzziness and asymmetry. Non transitivity, meaning that trust does not spread over social paths, is another strength of many trust models. Trust computational model (TCM in the table) tell us whether the model is linear or probabilistic. Although trust models use both computational models, the table shows that linear models are more popular because they are both more predictable and computationally less expensive. However, there are some strengths we consider more important than others in our application scenario of DOSNs. For instance, dynamism is an important strength because it enables the model to consider that trust constantly evolves, and therefore should be periodically recomputed. As a counterpart, dynamism introduces additional computational work. As shown in the table, not all trust models have this strength. In particular, those that can be adapted to a decentralized setting lack it. Another important strength is Context Dependency (CD in the table) because it lets users differentiate trust scores with their peers according to the specific context of the relationship. We believe this is an important strength because it is natural for users to have contextual relationships and the table shows this is a common strength of many considered trust models. Lastly, subjectivity is an important strength that all approaches should have, because trust is based on personal experience, and should not be influenced by other users' views or opinions. The table confirms that subjectivity is rare because all trust models use some form of non subjectivity in their computation. The most common form of non subjectivity is reputation, a completely different concept, while other models introduce some form of trust propagation or let trust values be influenced on multi-hop paths between users. The comparison shows that SentiTrust covers all strengths, while minimizing the drawbacks of their adoption. Hence, it is a trust model designed for DOSNs with a dynamic trust computation. Furthermore, it takes social contexts into account and it is the only one based on the subjective user experience.

IV. SentiTrust FEATURES

In this Section, we describe how SentiTrust builds a trust model by evaluating social interactions. The evaluation takes into account multiple factors, including the text and the photos exchanged by the users and, through an AI-powered Neuro-Behavioral Module (NBM), it defines a Sentimental Analysis that describes the social interaction. Moreover, we discuss how we enrich the trust computation through optional features to be used if available. These optional features can be useful either to bootstrap an initial trust value, i.e., when there is no recorded interaction, or to refine the trust evaluation obtained through Sentiment Analysis. In the following, we discuss three features, namely Sentiment Analysis, Profile Similarity and Common Friends, each trying to capture a distinct aspect of the relationship among people. Among the three features, the Sentiment Analysis is the only mandatory one in SentiTrust, because it is the only one that can truly captures the evolution of a social relationship. The other features this work considers, namely Profile Similarity and Common Friends, are optional. They are discussed to show that the model can be extended and therefore how to improve the evaluation of interpersonal trust. We consider these features as optional as they may be not available if users do not want to disclose them for privacy reasons.

A. SENTIMENT ANALYSIS

This Section describes Sentiment Analysis (SA), the mandatory feature in our model. We propose to characterize the relationships of the interactions among its actors through a Sentiment Analysis approach which helps us understand if the interactions are positive or negative, frequent or rare.

The SA computation considers both textual interactions (textual messages) and media interactions (i.e. images), to provide a score SA_{ij} that is representative of positive and negative emotions as extracted from the content exchanged between user i and user j. The score is computed by the NBM[8] [33]. In the remainder of the Section, we discuss in detail the analysis of textual and media interactions.

1) INTERACTIONS ANALYSIS

The textual analysis exploits the potential of NLP by using the TextBlob library[9] to implement the message sentiment analysis. The machine learning models in TextBlob were built with a corpus of English text, but the library itself offers an automatic translation service to compute trust regardless of the localization. TextBlob provides two different machine learning models: Pattern Analyzer and Naïve Bayes. Pattern Analyzer returns a polarity score $[-1,1]$ and a subjectivity score $[0,1]$, while the Naïve Bayes returns the classification [positive, negative], and the probabilities that the message is respectively positive and negative. We performed several preliminary tests to assess the quality of classification and understand if they would work well on a smartphone. After our testing, we adopted the Naïve Bayes, as it returns more accurate sentiment analysis results. On top of that, the evaluation of the message subjectivity was not of interest for this work.

As image analysis tool, we studied the ones available in the literature, balancing classification capabilities and applicability in our decentralized scenario. Our choice fell on Facial Expression Recognition, discussed in Section III-C2, because it can work with small images, and provides a shallow CNN

[8] https://github.com/helios-h2020/h.extension-NeuroBehaviouralClassifier [Accessed: 18-May-2023]

[9] https://textblob.readthedocs.io/en/dev/ [Accessed: 18-May-2023]

TABLE 2. Comparison of strengths and weaknesses between SentiTrust and others Social Trust models presented in literature by taking into account the trust properties. CD: Context dependency, TCM: Trust computational model.

Model	Dynamism	Fuzziness	Non Transitivity	Asymmetry	Subjectivity	CD	Reputation	De/Centr.	TCM
Li and Li [21]	✗	✓	✗	✓	✗	✓	✓	BOTH	Linear
Qureshi et al. [22]	✗	✓	✓	✓	✗	✗	✓	BOTH	Linear
STrust [8]	✓	✓	✓	✓	✗	✓	✓	C	Linear
Yu and Singh [15]	✓	✓	✓	✓	✗	✓	✓	C	Linear
SocialTrust [27]	✓	✓	✓	✓	✗	✓	✓	C	Linear/Prob.
3SVL [26]	✗	✓	✗	✓	✗	✗	✗	C	Prob.
SentiTrust	✓	✓	✓	✓	✓	✓	✗	BOTH	Linear

TABLE 3. List of intermediate features calculated by the Neuro-Behavioral Module to characterize the history of conversations. Numbers are computed both per conversation and per minute of conversation.

Type	Features
Text	
Number of messages	Number of messages sent, received and total
Number of words	Number of words sent, received and total
Emotion-related	Mean of the probability of being a positive or negative emotion of total, received and sent messages Mean of the valence of total, received and sent messages
Images	
Number of images	Number of images sent, received and total
Faces	Number of faces in images sent, received and total

model. We have run some preliminary tests to fine-tune the parameters of the model, and to make it work with images of distinct sizes and resolutions as well as to adapt it to the wide variety of cameras available on smartphones and other portable devices.

2) EGO-ALTER ANALYSIS

To consider the relationships among people, we map the social contacts of each user by using the ego network structure. We apply the ego-alter analysis function to analyze the social interactions between the ego and each of the alters. This function analyzes the history of communication, including the sentiment analysis of each exchanged message, and it computes several metrics. As a first step, the function partitions the history of communications into conversations. A conversation is a group of consecutive, in terms of time, social interactions (textual or image messages in our case). A threshold of 5 minutes determines the length of a conversation. Hence, the history of communications is divided into conversations, using the 5 minutes threshold. Table 3 lists the intermediate features computed to characterize each conversation. The identified features aim to describe the interactions between ego and alter under multiple aspects, such as frequency or density of communication, type and sentiment of communication, and so on.

3) THE SENTIMENT ANALYSIS

After the creation of intermediate features, three final scores are computed: attention, arousal, and valence, as a linear combination of the previous features. Arousal and Valence are selected as two main components of Circumplex Model of Affects [34] (CMA) model. We used the CMA to model emotions since it is one of the most predominant theories, and at the same time reduce the dimensions to only 2 features, which increases the explainability of the architecture proposed. To Arousal and Valence, we added the Attention to model the longitudinal assessment of the emotion. To specify the functions, an experimental study has selected the most informative features (see Table 3). The values discussed in this section (Attention, Arousal, and Valence) will be then combined to provide a final SA score value. How to combine these three values depends on the application context. We will provide a possible way to integrate these values in Section VII for a generic scenario.

a: ATTENTION

This function infers attention, understood as the quantity of communication between an ego and their alter. As shown in Equation 1, it is based on the percentage of days where at least one conversation occurs (dwc). Attention has two parameters T_1 and T_2 that can be used as a threshold to consider the communication frequent or infrequent.

$$Attention = \begin{cases} low, & \text{if} & dwc < T_1\% \\ medium, & \text{if } T_1\% \leq & dwc < T_2\% \\ high, & \text{if } T_2\% \leq & dwc \end{cases} \tag{1}$$

b: AROUSAL

This function infers arousal in the conversations between an ego and an alter and it measures the intensity of communication.

$$Arousal = a_1 * p_{MM} + a_2 * p_{HFR} + a_3 * p_{HFE} \tag{2}$$

where p_{MM} is the percentage of conversations with high arousal, i.e., where less than 2 messages have been sent per minute. p_{HFR} is the percentage of conversations without happy faces in the images received, and p_{HFE} is the percentage of conversations without happy faces in the images exchanged. The features in the arousal model have been developed based on the experimental evaluation presented in Section VI-E. The values of the weights in Equation 2 are such that $a_1 + a_2 + a_3 = 1$, so that the final value of Arousal lies in the range [0, 1].

TABLE 4. NBM parameters.

Parameter	Description
T_1, T_2	Parameters to compute Attention
a_1	Weight for P_{MM} in the computation of Arousal
a_2	Weight for P_{HFR} in the computation of Arousal
a_3	Weight for P_{HFE} in the computation of Arousal
b_1	Weight for P_{WR} in the computation of Valence
b_2	Weight for P_{HV} in the computation of Valence
b_3	Weight for P_{IE} in the computation of Valence
b_4	Weight for P_{HFE} in the computation of Valence

c: VALENCE

This function infers valence in the conversations between an ego and an alter, and it measures the positive engagement of a user in the conversation.

$$Valence = b_1 * p_{WR} + b_2 * p_{HV} + b_3 * p_{IE} + b_4 * p_{HFE} \tag{3}$$

where p_{WR} is the percentage of conversations with high valence in terms of words, i.e., a conversation has high valence if less than 25 words are received. p_{HV} is the percentage of conversations with high valence in terms of sentiment analysis, i.e., a conversation has high valence if the text valence is higher than 0.1. p_{IE} is the percentage of conversations that exchange at least one image, and p_{HFE} is the percentage of conversations where at least one happy face appears in the exchanged images. The features in the valence model have been developed based on the experimental evaluation presented in Section VI-E. The values of the weights used in Equation 3 are such that $b_1+b_2+b_3+b_4 = 1$, so that the final value of Valence lies in the range [0, 1]. In Table 4 we summarize the parameters of the NBM model.

B. OPTIONAL FEATURES

We consider two optional features: Profile similarity and Common friends.

Profile Similarity (PS) is a score that measures how similar two users are. User profiling can adopt multiple techniques, such as those based on the images stored on a user device [35], or on other arbitrary features [36]. After profiling the users, a similarity metric is computed to understand to what extent two users share common interests.

Common Friends (CF) measures the common acquaintances two users share and shows how many belong to the same social circle. CF is the percentage of friends of the user i that are also friends of the user j. Therefore, it belongs to the interval [0, 1] ($CF_{ij} \in [0, 1]$). Scaling the number of common friends with respect to the total number of friends is crucial to avoid overstated values due to users with a larger number of friends. The lower the number of friends of a user, the larger the impact of common friends.

Since comparing the profiles of two users or finding their common friends involves transmitting sensitive information, we allow users to block the usage of these features to limit the disclosure of personal information. However, when these features are available, they have an important impact on trust evaluation as they help to bootstrap trust when two users have never interacted or have a better evaluation of the trust value as their relationship evolves.

V. SentiTrust: THE MODEL

This Section describes SentiTrust as a general and theoretical model that can be configured according to the needs of the developers according to the applicative scenario, and the end users whose trust is being computed. The model is fully parametrized to ensure maximal flexibility. To provide a contextualized trust value, SentiTrust uses the Contextual Ego Network structure (CEN), a multilayered structure presented in [6] so that, for each alter, the ego will compute the trust score for each context where the alter appears. In this way, the design also takes into account the privacy aspect, as each layer is modeled through an ego network.

SentiTrust is a linear trust model that returns a numerical value, the *trust score*, representing the level of trust between the ego and the alter, from the perspective of the ego. The score computation takes into account the evolution of the relationship between the ego and the alter. For this reason, we identify two phases: when a relationship is established and when the level of trust is updated as the relationship evolves. SentiTrust applies a distinct formula to each phase.

A. TRUST SCORE INITIALIZATION

When the ego i establishes a new relationship with a new alter j, the corresponding edge is added to the CEN of i. As soon as a new edge is created, the trust value T_{ij} between the ego i and the alter j needs to be initialized. Since it is extremely challenging to estimate the trust value between two users before they even have the chance to interact, we bootstrap this value through a heuristic. The initial trust value, denoted as T_{ij}^0, defines the trust that the ego initially has towards the alter.

The initial value is computed as:

$$T_{ij}^0 = \alpha_0 PS_{ij} + \beta_0 CF_{ij} \tag{4}$$

where PS_{ij} is the similarity of the profiles of users i, and j, CF_{ij} represents the percentage of common friends. Finally, α_0, and β_0 are configurable weights such that $\alpha_0 + \beta_0 = 1$.

As previously recalled, some features may not appear in a scenario, and the owner may hide some personal information to achieve a high level of privacy. In such cases, the trust value can be bootstrapped using a neutral value (such as $T_{ij}^0 = 0.5$). In this way, no assumption is made about the relationship, and the bootstrapping process is not biased.

B. TRUST SCORE UPDATE

After initializing the trust score, SentiTrust monitors the relationship status by updating the value any time the ego and the alter interact. When i and j are connected in a context, their social relationship will evolve over time mostly by social virtual interactions. Examples of these interactions include: sending chat messages, exchanging images, videos, and so on. According to the evolution of the social relationship, the trust value between the ego i and alter j is updated over time. Considering both the huge amount of interactions two users may have and the need to run SentiTrust as part of a mobile application, we update the trust score at regular time intervals, Δt, rather than at each interaction. Each time Δt seconds have passed, the trust value is updated as follows.

The evolution of the trust score is ruled by the formula:

$$T_{ij} = \beta CF_{ij} + \gamma SA_{ij} \tag{5}$$

where SA_{ij} is the SA score related to the messages and media from j to i. The weights β, γ are such that $\beta + \gamma = 1$, where the γ parameter is larger because it governs the importance of the SA score. Since all the features lie in the interval [0, 1], and all the weights sum to 1, $T_{ij} \in [0, 1]$.

When the trust score has to be updated, the NBM is triggered to analyze and evaluate the social interactions among users. The module output is combined with the ratio of common friends if the user has enabled this feature. The SA score from the NBM analyses the interactions from the alter j to the ego i. The SA score to update the trust one is obtained by querying the NBM, as described in Section IV. The NBM returns a sentiment score that considers all the past interactions with the alter, to evaluate how the ego reacts, on average, when interacting with the alter.

C. TRUST SCORE DISCRETIZATION

To produce a comprehensive trust model, we provide a discretization with natural language labeling. The proposed labeling aims to offer other modules and applications a ready-to-use, immediate, trust estimation. However, the raw numeric value is available as well, so that developers can define custom labeling, according to the scenario where the Trust model is ultimately set, the activity of the users, features used, and the meaning of the interactions.

VI. NBM FEATURES SELECTION

This Section describes the experimental study of the physiological signals in order to evaluate the NBM and, as a consequence, the trust model. Since the NBM represents the main novelty in SentiTrust, and it is a required parameter, we decide to test the module, by setting the parameters $\alpha = \beta = 0$, and $\gamma = 1$ (see Equations 4 and 5). The focus of the evaluation process is on the most innovative part of the model, and its importance in the evaluation of trust.

The experimental study includes 30 subjects, split into 10 egos and 20 alters. The alters were divided into two subgroups, high-trust alters and low-trust alters, to consider both scenarios. Subjects were asked to complete a set of tasks where they were required to hold different conversations, while the communications were tracked by the NBM. In addition, the egos were monitored by distinct neurophysiological sensors, including electrodermal activity (EDA) and electrocardiogram (ECG). The egos had to complete the tasks with, respectively, the high-trust alter and the low-trust one sequentially and independently. After completing each task, subjects had to indicate their level of pleasure and arousal with respect to the task. These values were used to study the relationship of the emotional states with the metrics collected in the study. Two hypotheses were tested: 1) metrics obtained from the NBM can recognize the emotional states of the individuals during conversations and 2) emotions evoked during the conversations generate different patterns in terms of neurophysiological responses. The following sections detail the study sample and materials, the data analysis performed and the results obtained.

A. SAMPLE

The sample included 30 Spanish participants (18 women and 12 men from 20 to 55 years old). Due to the COVID situation, we were not able to produce a test with more than 30 participants. They were recruited in pairs, meeting two fundamental criteria: having a certain previously defined lifestyle and a relationship of trust vs. no trust with each other. Initially, the resulting trust vs. non-trust pairs were intended to be from the same work environment, but this criterion was partially met. In addition, the sample could not meet clinical criteria nor required special experience with social networks. Exclusion criteria included any reason that prevented correct reading and understanding of the assessment questionnaires, as well as mild or severe psychological disorders.

We recruited participants in high and low-confidence pairs, through a selection phase. First, users had to fill out the Lifestyle Questionnaire [37] to determine their lifestyle. The target lifestyle of the study included those subjects with high scores on items related to personal success, friendship, responsibility, innovation and fashion, as well as high ratings on activities related to cinema, cultural activities, gastronomy, nightlife, shopping, and involvement in social media. Taking this criterion into account, ten subjects were selected for the study, to be considered as the egos of just as many CENs. Subsequently, they had to choose one trustworthy person and one untrustworthy person, who were the alter users in the study. The confidence level needed to be reciprocal with the aim of creating high and low trust pairs, so the Trust Scale was used to determine the trust level between them. In addition, the selected people filled out the Lifestyle Questionnaire [37] to select only those with the previously defined target lifestyle. After the recruitment process, 10 egos and 20 alters were selected for participation in the study, the latter divided into two sub-groups, high-trust alters and low-trust alters.

The experimental protocol was approved by the ethics committee of the Polytechnic University of Valencia and informed consent was obtained from all participants.

B. MATERIALS

The materials used in this study can be divided into two categories: subjective and objective materials. Subjective materials refer to questionnaires, in which the user's opinion is requested. Objective materials are different tools used for collecting information that is not influenced by the user's opinion. Both groups are detailed below.

1) SUBJECTIVE MATERIALS

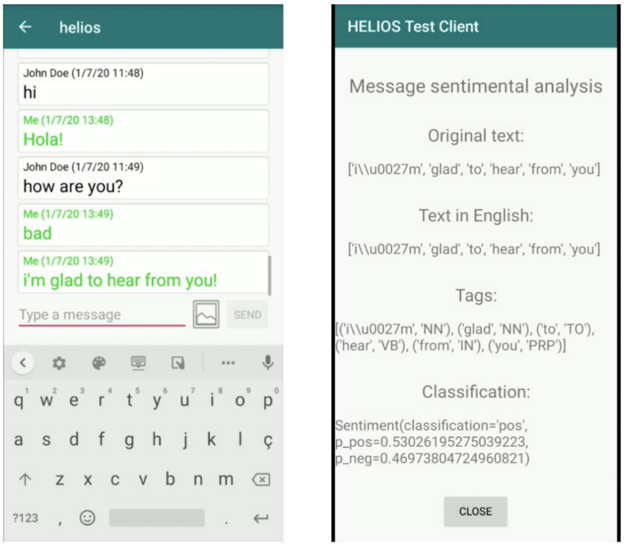
The following questionnaires were used:

- Trust Scale [38]. It consists of 18 items that evaluate the degree of trust of a person towards another person. Three aspects of trust are evaluated: predictability, dependability and faith.
- Lifestyle Questionnaire [37]. It consists of 65 lifestyle items related to different interests and opinions (society, politics, job, personal success factors, environment, religion, future, family, friendship, responsibility, aspirations, attitude to personal problems, saving, innovation and fashion) and several activities (do-it-yourself, sport, cinema, cultural activities, visit beautiful places, nightlife, shopping, reading, music, TV programs and social media).
- Self-Assessment Manikin (SAM) [39]. It is a pictured-oriented questionnaire to measure emotional response. Specifically, there are single-item scales that measure the valence/pleasure of the response (from positive to negative) and perceived arousal (from high to low levels).

2) OBJECTIVE MATERIALS

The following devices were used:

- Mobile phone. The different tasks of the study were conducted by developing a demo app based on the HELIOS TestClient application[10] integrated with the NBM. For this purpose, three mobile phones were required: two for the couple of participants doing the tasks and one for the researchers to supervise the task. Figures 1 and 2 show the demo app used during the test.
- Shimmer3 GSR+ sensor. It is a sensor of electrical activity in the dermis that measures the skin's ability to transmit electrical currents, which varies if there is sweating and changes in the body. The Shimmer sensor consists of two fabric bands with Velcro, on which an electrode is sewn. This device was used for the acquisition of the electrodermal activity (EDA), also called Galvanic Skin Response (GSR), of the ego users.
- B-Alert system. It is a Bluetooth wireless system and a sensor headset integrated with the Acqknowledge Data Acquisition[11] software that allows the recording of up to 9 channels of monopolar electroencephalogram (EEG), plus another channel for ECG data. It was used to acquire the ECG of ego users. The B-Alert system has EEG electrodes, as mentioned before, but this signal will be analyzed in future works due to its complexity.



John Doe (1/7/20 11:48)
hi

Me (1/7/20 13:48)
Hola!

John Doe (1/7/20 11:49)
how are you?

Me (1/7/20 13:49)
bad

Me (1/7/20 13:49)
i'm glad to hear from you!

Message sentimental analysis

Original text:

['i\\u0027m', 'glad', 'to', 'hear', 'from', 'you']

Text in English:

['i\\u0027m', 'glad', 'to', 'hear', 'from', 'you']

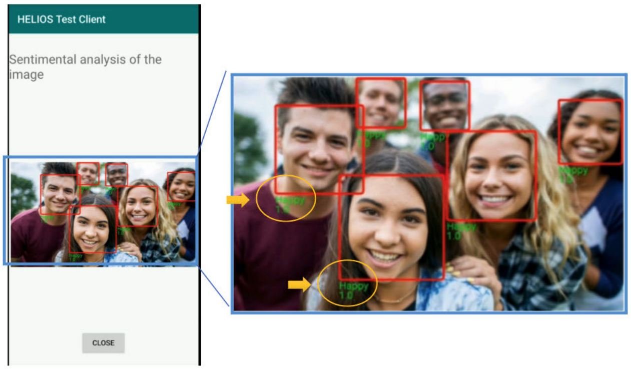
Tags:

[('i\\u0027m', 'NN'), ('glad', 'NN'), ('to', 'TO'), ('hear', 'VB'), ('from', 'IN'), ('you', 'PRP')]

Classification:

Sentiment(classification='pos', p_pos=0.53026195275039223, p_neg=0.46973804724960821)

FIGURE 1. Text messages from the demo app UI and the relative results of text sentimental analysis.



Sentimental analysis of the image

FIGURE 2. Screen with results of sentimental analysis of the image.

3) DATA PROCESSING

- Electrodermal activity signal requires two previous steps before its analysis. The first of these steps is the manual cleaning of the signal. EDA signal could suffer from different types of artifacts that come, for example, from the subject's movement. These artifacts could hide important correlations between the level of stress of the subject and its EDA analysis. The manual correction was done using Ledalab software in Matlab.[12] The second step is the decomposition of this signal, using Continuous Decomposition Analysis (CDA), in two components named phasic and tonic [40]. Each component has different latency related to the triggering stimuli. The study of rapid movements of the signal, called skin conductance response (SCR), is studied by the phasic component. The tonic component represents the slowest variations of the original autonomic nervous system dynamics.

[10] https://github.com/helios-h2020/h.app-TestClient-AutumnApp [Accessed: 18-May-2023]

[11] https://www.biopac.com/product/acqknowledge-software/[Accessed: 18-May-2023]

[12] www.ledalab.de [Accessed: 18-May-2023]

- ECG signals were processed to derive Heart Rate Variability (HRV) series [41]. The artifacts were cleaned by the threshold base artifacts correction algorithm in the Kubios software [42]. In order to extract the RR series, the well-known algorithm developed by Pan-Tompkins was used to detect the R-peaks. The individual trends components were removed using the smoothness prior detrending method [43]. Heart Rate Variability analysis was carried out based on the standard HRV parameters, which are defined in the time and frequency domains. In order to obtain the frequency domain features, a power spectrum density (PSD) estimate was calculated for the RR interval series by a Fast Fourier Transform based on Welch's periodogram method. The analysis was carried out in three bands: very low frequency (VLF, $< 0.04Hz$), low frequency (LF, 0.04-0.15Hz) and high frequency (HF, 0.12-0.4Hz). For each frequency band, the peak value was calculated, corresponding to the frequency with the maximum magnitude. The power of each frequency band was calculated in absolute and percentage terms. Moreover, for the LF and HF bands, the normalized power (n.u.) was calculated as the percentage of the signals subtracting the VLF from the total power. The LF/HF ratio was calculated in order to quantify sympatho-vagal balance and to reflect sympathetic modulations [41].

C. EXPERIMENTAL PROCEDURE

Participants attended the laboratory in groups of three, consisting of the ego and the two alters, with low and high trust respectively. First, the egos users were fitted with the devices for psychophysiological metrics: the B-Alert system to record ECG and the Shimmer3 GSR+ sensor to capture the EDA. Subsequently, egos were required to complete ten tasks with the high-trust alter and then ten tasks with the low-trust alter, sequentially and independently. The tasks consisted of having different conversations related to the lifestyle and interests of the participants, while the communications were tracked by the NBM. After completing each task, subjects had to complete the SAM to indicate their level of pleasure and arousal with respect to the task.

In the first four tasks, subjects were recommended to use a specific resource, such as the front camera or the image gallery, to familiarize themselves with the mobile application. In the remaining tasks, they were asked to maintain a conversation with their partner to achieve a specific goal and they were free to use any resource in the app to communicate with each other. Although tasks were the same for both subjects, small content modifications were introduced to generate a dynamic conversation. For example, the first user had to try to convince his partner to go out for lunch the next day, while the second user had to try to convince his partner to go out for breakfast. Moreover, to make the conversation as natural as possible, contextual information related to the task was provided.

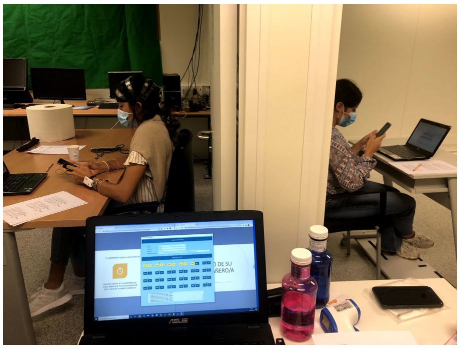


FIGURE 3. Experimental evaluation: the scenario.

Some examples of the tasks included in the study are the following:

- Invite your partner to have lunch today. You can propose several alternatives so he/she can choose the one he/she prefers.
- Ask your partner how the new haircut looks on you.
- Share with your colleague that your boss sends you tasks that are not yours and ask for their opinion.
- You have recently watched several movies, recommend one of those movies to your partner.
- You want to visit a country to know a different gastronomic culture. Reach an agreement with your partner on the ideal country to go to achieve this goal.
- Try to convince your partner to go to a birthday party hosted by a friend he/she barely knows.

The execution order of the tasks was randomized to avoid significant biases and the order of the conversations between the trusted and non-trusted couples was also alternated to control the possible effect of fatigue on the person performing both conversations. A picture of the experimental phase can be seen in Figure 3.

D. DATA ANALYSIS

All the conversations, i.e. each task, were divided into two groups (high and low) in terms of valence and arousal, using the median of the self-assessment of each task to divide both groups. All the features obtained from the NBM and the physiological sensors were analyzed to identify significant differences between the individuals on the two levels of arousal and valence. Either a t-test or a Mann-Whitney test [44] was performed depending on the Gaussianity of the data (i.e. mean and standard deviation if Gaussian, median and the 0.25 (Q1) and 0.75 (Q3) percentiles if non-Gaussian). Shapiro-Wilk test [45], where the null hypothesis is a Gaussian sample, was used to define the Gaussianity of each variable, defining a variable as non-Gaussian with a *p-value* <0.05 in this test. The description of the values in each group (high/low) of the variables collected by the NBM is described according to its Gaussianity. We identified some variables, all related to the images exchanged during the tasks, which had very few distinct values. Most of the participants did not exchange any

TABLE 5. Variables collected by the NBM, including all the tasks. ^{a} This variable is not normally distributed (Shapiro test p-value < 0.05). Therefore it is described as median [Q1-Q3]. The p-value corresponds to a Mann-Whitney test. ^{b} This variable is considered categorical and then described as n (%). The p-value corresponds to a Chi-squared test.

Variable		Low arousal N = 207	High arousal N = 173	p-value
Messages sent (n/minute)^a		2.51 [1.25-4.02]	1.82 [1.00-3.58]	0.022
Happy faces (sent/received) (n)^b:	0	190 (91.8)	169 (97.7)	0.023
	≥ 1	17 (8.21)	4 (2.31)	
Happy faces received (n)^b:	0	196 (94.7)	172 (99.4)	0.020
	≥ 1	11 (5.31)	1 (0.58)	

pictures, so we aggregated the values into two groups (e.g., 0 or ≥ 1) and treated them as categorical. Analyses were performed in R (version 3.6.3). The significance level was set to 0.05.

E. RESULTS

This Section reports the variables where significant differences were identified between high and low arousal or valence individuals.

1) AROUSAL AND NBM

Table 5 shows the significant differences found in the variables collected by the NBM in terms of arousal. Results show that participants with lower arousal, hence lower activation, sent more messages per minute. More participants with lower arousal exchanged pictures including happy faces (8.2%) than participants with high arousal (2.3%). For this reason, we developed the arousal model stated in Equation 2 based on messages sent and happy faces exchanged.

2) VALENCE AND NBM

Table 6 shows significant differences in the variables collected by the NBM in terms of valence. The more messages are received, and the longer they are, the more negative the user's valence is. The probability of negative messages in the conversation is higher in participants self-reporting a negative valence, while the valence of the messages is higher in participants with positive valence. In terms of images, the more images were sent or received, and the more faces appeared in them, the more positive the user's valence was. The more happy faces sent and received in those images, the higher the user's valence too. Considering the significant features and the distribution of both groups (negative and positive valence), we developed the valence model stated in Equation 3 based on word received, valence of the messages computed by sentiment analysis, images exchange and happy faces.

TABLE 6. Variables collected by the NBM, including all the tasks. ^{a} This variable is not normally distributed (Shapiro test p-value < 0.05), and it is described as median [Q1-Q3]. The p-value corresponds to a Mann-Whitney test. ^{b} This variable is considered categorical and then described as n (%). The p-value corresponds to a Chi-squared test.

Variable		Neg. valence N = 217	Pos. valence N = 163	p-value
Messages received (n)^a		5.00 [2.00-8.00]	4.00 [2.00-7.00]	0.018
Words sent and received (n)^a		53.0 [37.0-75.0]	43.0 [29.0-65.0]	0.004
Words received (n)^a		27.0 [18.0-40.0]	22.0 [14.0-32.0]	0.001
Mean prob. negative^a messages		0.45 [0.40-0.49]	0.42 [0.37-0.47]	0.028
Mean valence of messages^a		0.08 [0.01-0.15]	0.11 [0.03-0.20]	0.049
Images (sent/received) (n)^b:	0	175 (80.6)	97 (59.5)	<0.001
	≥ 1	42 (19.4)	66 (40.5)	
Images (sent/received) (n/minute)^b:	0	175 (80.6)	97 (59.5)	<0.001
	≥ 1	42 (19.4)	66 (40.5)	
Images sent (n)^b:	0	191 (88.0)	122 (74.8)	0.001
	≥ 1	26 (12.0)	41 (25.2)	
Images sent (n/minute)^b:	0	191 (88.0)	122 (74.8)	0.001
	≥ 1	26 (12.0)	41 (25.2)	
Images received (n)^b:	0	199 (91.7)	135 (82.8)	0.014
	≥ 1	18 (8.29)	28 (17.2)	
Images received (n/minute)^b:	0	199 (91.7)	135 (82.8)	0.007
	≥ 1	18 (8.29)	28 (17.2)	
Faces sent and received (n)^b:	0	191 (88.0)	121 (74.2)	<0.001
	≥ 1	26 (12.0)	42 (25.8)	
Faces sent (n)^b:	0	199 (91.7)	137 (84.0)	0.020
	≥ 1	18 (8.29)	26 (16.0)	
Faces received (n)^b:	0	209 (96.3)	147 (90.2)	0.013
	≥ 1	8 (3.69)	16 (9.82)	
Happy faces received and sent (n)^b:	0	211 (97.2)	148 (90.8)	0.012
	≥ 1	6 (2.76)	15 (9.20)	
Happy faces sent (n)^b:	0	215 (99.1)	156 (95.7)	0.040
	≥ 1	2 (0.92)	7 (4.29)	

TABLE 7. Variables extracted from the physiological sensors (i.e., GSR and ECG), including all the tasks. None of these variables was normally distributed (Shapiro test p-value < 0.05), and they are described as median [Q1-Q3]. The p-value corresponds to a Mann-Whitney test.

Signal	Variable	Low arousal N = 207	High arousal N = 173	p-value
EDA	Conductance	3.16 [0.83-5.41]	4.75 [1.08-6.53]	0.002
	Phasic comp.	0.04 [0.01-0.17]	0.08 [0.03-0.18]	0.003
	Tonic comp.	3.02 [0.81-5.19]	4.45 [1.05-6.27]	0.002
HRV	RMSSD (ms)	28.8 [21.1-59.0]	22.1 [14.2-37.5]	0.002
	pNN50 (%)	7.20 [1.80-42.1]	2.95 [0.30-16.6]	0.002
	LF peak (Hz)	0.09 [0.06-0.10]	0.08 [0.04-0.09]	0.025
	LF power (n.u.)	0.73 [0.54-0.82]	0.78 [0.70-0.86]	0.001
	HF power (ms^2)	284 [148-1256]	229 [70.0-494]	0.004
	HF power (n.u)	0.27 [0.18-0.46]	0.22 [0.14-0.30]	0.001
	LF/HF power	2.67 [1.16-4.45]	3.48 [2.34-5.97]	0.001

3) AROUSAL AND PHYSIOLOGICAL SENSORS

Table 7 reports the physiological variables with significant differences between individuals with low and high arousal. The conductance, phasic component and tonic component of EDA are higher in the high arousal condition. In addition, the LF power increases and the HF decreases in the high arousal condition. Participants with low arousal, hence lower activation, show lower RMSSD and PNN50, and higher LF peak and Poincaré SD1.

4) VALENCE AND PHYSIOLOGICAL SENSORS

Table 8 reports the physiological variables with significant differences between individuals with low and high valence. No significant differences were found in any of the variables extracted from the EDA. Concerning HRV, participants self-reporting having a positive valence show lower Std RR, LF power, total power and Poincaré SD2, whereas the HF peak is higher.

F. DISCUSSION

We analyzed the features obtained by the NBM that show significant differences between high and low arousal and

TABLE 8. Variables extracted from the physiological sensors (i.e., GSR and HRV), including all the tasks. [a] This variable is not normally distributed (Shapiro test p-value <0.05), and it is described as median[Q1-Q3]. The p-value corresponds to a Mann-Whitney test. [c] This variable is normally distributed (Shapiro test p-value ≥ 0.05), and it is described as mean (standard deviation). The p-value corresponds to a t-test.

Signal	Variable	Negative valence N = 217	Positive valence N = 163	p-value
HRV	Std. RR (ms)^a	5.35 [3.30-6.70]	4.60 [3.30-5.70]	0.019
	RR triangular index^c	9.63 (2.70)	10.3 (2.20)	0.038
	HF peak (Hz)^a	0.29 [0.21-0.33]	0.32 [0.28-0.34]	<0.001
	LF power $(ms^2)^a$	918 [401-2333]	634 [348-1307]	0.018
	Total power $(ms^2)^a$	2287 [908-5483]	1688 [885-3082]	0.035

valence levels based on users' self-assessments. In addition, we evaluate users' self-reported emotional states through their physiological responses to each task.

The results show that users with low arousal exchanged more images including happy faces. Therefore, the results suggest that the reception and exchange of multimedia content, particularly positive content, cause a decrease in the participants' arousal. These results are interesting since, a priori, one could hypothesize that arousal would increase with the number of images received. However, the results suggest that when the egos received images of the alters, they self-reported lower arousal states. This could probably be due to the relaxation caused by the positive feedback, which is especially remarkable because such feedback was received from both high and low-confidence alters. In this sense, a recent study [46] found that individuals with a greater sense of dominance and leadership showed reduced neural indices of emotional arousal after receiving validation from others in the form of "likes" on their "selfie" picture posted on social media. In addition, this effect is also observed in the messages sent per minute by the ego, which is larger in the low arousal state. This may also be due to the relaxation evoked by positive feedback in low arousal states. In high arousal states the egos are likely to be in a more stressful situation where the writing process is less fluid and, therefore, fewer messages are sent per minute.

With regard to valence, many NBM features present significant differences. The messages received and the number of words is higher in negative valence states, suggesting that the negative conversations are longer. In addition, the mean of the probability that the messages are negative is larger in negative valence conditions, and the valence is larger in positive valence conversations. This supports the integration of sentiment analysis with the NBM to analyze the semantics of text communications. In addition, a larger number of images, including faces and happy faces, are exchanged in positive valence conversations. Therefore, the use of pictures is related to positive valence states.

The self-reported states have been evaluated by using physiological sensors. In terms of arousal, the conductance, the phasic and the tonic components, are higher in high-arousal conversations. These results are in accordance with previous research, which reports the EDA as an indicator of stress-related states [47], showing that increases in conductance, and their phasic and tonic components, are correlated with higher arousal levels. Moreover, the HRV results corroborate the achievement of high-arousal states during the task, given that LF increases their power in high-arousal states while HF decreases, which is in concordance with previous research that showed sympathetic activation and a parasympathetic withdrawal during stress [48]. In terms of valence, HF peaks are higher in positive valence, since HF has been shown as an indicator of positive valence [49]. Moreover, LF power is higher in negative valence, which could be derived from the increment of arousal provoked by the negative task. The physiological responses suggest that the emotional states have been evoked properly with the methodology designed and, accordingly, the NBM is able to recognize the emotional states of the users of HELIOS.

VII. MODEL CONFIGURATION

In this section, we propose a model configuration by assigning a set of values to the parameters of the NBM and Trust model presented in Section IV. We do not consider a specific applicative scenario, but we only assume that users interact through a messaging application, included in the demo app (Figures 1 and 2). These values are included in the demo app as the initial model configuration for a generic scenario. Application developers using SentiTrust should tune the parameters according to the specific social scenario provided by the application, because there is no set of values of parameters valid in each application scenario.

A. NBM PARAMETERS

For what concerns the NBM, we need to focus on its three main components, namely Attention, Arousal, and Valence. In the case of Attention, we need to assign the parameters T_1 and T_2 that govern when a communication can be considered frequent or infrequent. In our case, possible values are $T_1 = 33.\overline{3}$ and $T_2 = 66.\overline{6}$, so that when a pair of users interact less than once every three days the Attention is low, if they communicate more than two-thirds of the days their Attention is high, and it is medium otherwise. The choice to divide the interval into three equally-sized sub-intervals aims to define the most generalized division of the interval without introducing any bias in the computation.

In the case of Arousal, we propose as weights for Equation 2 the following values: $a_1 = 0.7, a_2 = 0.1, a_3 = 0.2$. The motivation behind choosing a higher value for weight a_1 lies in the fact that the weight is associated with the percentage of conversation with high arousal, and our experimental evaluation showed that it is an important parameter to compute Arousal (see Table 5).

In the case of Valence, we propose as weights for Equation 3 the following values: $b_1 = 0.15, b_2 = 0.3, b_3 = 0.4, b_4 = 0.15$. In this case, we decided to give more importance to the sentiment analysis score of the text (parameter associated with weight b_2), and the images exchanged (parameter associated with weight b_3) which was shown to be one of the most significant variables from the results reported in Table 6. The values assigned to b_4 is deliberately low, because the feature appears in the computation of both Arousal and Valence, so we want to exclude too much influence from a single parameter.

B. TRUST PARAMETERS

For what concerns the Trust module, we need to describe how the values of Arousal, Attention and Valence are combined together to produce the final SA score. At first, we combine the values of Arousal and Valence to obtain a score (AV) as shown in Equation 6. In the equation, we identify four possible cases given by the combination of values of Arousal and Valence. Indeed, each of them can be either *high* (i.e. $\geq$0.5) or *low* (i.e. <0.5). This decision was driven by the fact that Attention and Valence can be naturally labeled in two classes. As an example, high Valence corresponds to positive engagement, while low Valence corresponds to negative engagement

$$AV = \begin{cases} 0.25, & \text{if } Arousal \geq 0.5,\ Valence < 0.5 \\ 0.50, & \text{if } Arousal < 0.5,\ Valence < 0.5 \\ 0.75, & \text{if } Arousal < 0.5,\ Valence \geq 0.5 \\ 1, & \text{if } Arousal \geq 0.5,\ Valence \geq 0.5 \end{cases} \tag{6}$$

The reasoning behind this formula is the following. Arousal, which measures the intensity of the communication between users, is used to determine whether the combined score will lie at one of the two extremes. Indeed, if the Arousal is high, the combined value will either be 0.25 or 1, and will be 0.5 or 0.75 otherwise. Valence, which measures the positive engagement between two users, will help determine whether the combined score will lean towards high or low values. Indeed, if the Valence is low, the combined score will be either 0.25 or 0.5, while if the Valence is high, the combined values will be either 0.75 or 1.

To complete the computation of the SA score, we need to take into account the value of Attention that is returned as one of three labels, as shown in Equation 1. The choice to return a label lies in the fact that how the attention is computed could be altered in the future, while still maintaining the rest of the trust computation intact. The labels provided by the Equation are then translated using Equation 7. The reason behind this division lies in the fact of having a 3-fold division of the interval help us to naturally model frequent, infrequent and occasional interaction. Finally, the SA score is computed by multiplying the values obtained by Equations 6 and 7.

$$A = \begin{cases} 0.33, & \text{if } Attention = low \\ 0.66, & \text{if } Attention = medium \\ 1, & \text{if } Attention = high \end{cases} \tag{7}$$

TABLE 9. A possible discretization of trust values proposed in SentiTrust.

Interval	[0, 0.2)	[0.2, 0.5)	[0.5, 0.8)	[0.8, 1]
Trust label	Distrust	Partial distrust	Partial trust	Trust

To conclude, we propose in Table 9 a possible discretization of the SentiTrust trust score. A trust label in natural language is assigned to each chosen sub-interval of [0, 1]. The discretization proposed is such that more extreme values (Distrust and Trust) are assigned only when the trust value is particularly high or low.

VIII. CONCLUSION AND FUTURE WORK

Trust is a multidisciplinary concept which found application in many fields. In OSNs, trust can be used to identify reliable sources of information or relevant relationships.

In this paper, we presented SentiTrust,[13] a new trust model for DOSNs which takes into account the privacy of users. The model is based on the Sentiment Analysis of people interactions, and it can be easily extended or re-adapted according to the scenario of interest. Sentiment Analysis runs in the user device thanks to an innovative, lightweight, AI model, called Neuro Behavioral Model (NBM), which analyzes the text and pictures the users exchange. A comprehensive analysis phase, involving 30 participants and several tasks to be performed, helped to assess the potential of both the NBM and the Trust Model. In particular, the analyses show that a larger number of messages per minute and the number of happy faces sent or received are typical of a lower arousal state, derived from the relaxation evoked by positive interactions. The number of messages and the number of words is larger in negative valence relationships. On the other hand, the number of pictures is higher in positive valence conversations. The implemented sentiment analysis has been tested through users' EDA and HRV responses, whose values are related to the emotional states recognized by the NBM. Future developments of our work will be focused on extending SentiTrust, to include a trust propagation system and advanced trust evaluation mechanisms, as well as more types of content such as audios. Moreover, we plan to investigate possible improvements of our model with non-linear relationships between the features computed in the NBM, and to explore new emotion recognition models that can be embedded in phones. Finally, we also plan to investigate how the whole model can be self adaptive on the user behaviour, as so each instance of SentiTrust will have its own parameters that will change according to the user activity.

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.

[13] https://github.com/helios-h2020/h.core-TrustManager [Accessed: 18-May-2023]

[2] L. Jiang and X. Zhang, "BCOSN: A blockchain-based decentralized online social network," *IEEE Trans. Computat. Social Syst.*, vol. 6, no. 6, pp. 1454–1466, Dec. 2019.

[3] B. Guidi, A. Michienzi, and L. Ricci, "Steem blockchain: Mining the inner structure of the graph," *IEEE Access*, vol. 8, pp. 210251–210266, 2020.

[4] W. Sherchan, S. Nepal, and C. Paris, "A survey of trust in social networks," *ACM Comput. Surveys*, vol. 45, no. 4, pp. 1–33, Aug. 2013.

[5] E. García-Martín, C. F. Rodrigues, G. Riley, and H. Grahn, "Estimation of energy consumption in machine learning," *J. Parallel Distrib. Comput.*, vol. 134, pp. 75–88, Dec. 2019.

[6] B. Guidi, K. G. Kapanova, K. Koidl, A. Michienzi, and L. Ricci, "The contextual ego network P2P overlay for the next generation social networks," *Mobile Netw. Appl.*, vol. 25, no. 3, pp. 1062–1074, Jun. 2020.

[7] L. Mui, M. Mohtashemi, and A. Halberstadt, "A computational model of trust and reputation," in *Proc. 35th Annu. Hawaii Int. Conf. Syst. Sci.*, Jan. 2002, pp. 2431–2439.

[8] S. Nepal, W. Sherchan, and C. Paris, "STrust: A trust model for social networks," in *Proc. IEEE 10th Int. Conf. Trust, Secur. Privacy Comput. Commun.*, Nov. 2011, pp. 841–846.

[9] D. Olmedilla, O. F. Rana, B. Matthews, and W. Nejdl, "Security and trust issues in semantic grids," in *Semantic Grid: The Convergence of Technologies*, vol. 5271. Schloss Dagstuhl, Germany: Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), 2006, pp. 1–11.

[10] G. Liu, Y. Wang, and M. Orgun, "Trust inference in complex trust-oriented social networks," in *Proc. Int. Conf. Comput. Sci. Eng.*, Aug. 2009, pp. 996–1001.

[11] J. Golbeck and J. Hendler, "Inferring binary trust relationships in web-based social networks," *ACM Trans. Internet Technol.*, vol. 6, no. 4, pp. 497–529, Nov. 2006.

[12] S. Adali, R. Escriva, M. K. Goldberg, M. Hayvanovych, M. Magdon-Ismail, B. K. Szymanski, W. A. Wallace, and G. Williams, "Measuring behavioral trust in social networks," in *Proc. IEEE Int. Conf. Intell. Secur. Informat.*, May 2010, pp. 150–152.

[13] Y. Wang and J. Vassileva, "Trust and reputation model in peer-to-peer networks," in *Proc. 3rd Int. Conf. Peer Peer Comput. (P2P)*, 2003, pp. 150–157.

[14] S. Buchegger and J.-Y. Le Boudec, "A robust reputation system for P2P and mobile ad-hoc networks," in *Proc. P2P Econ.*, 2004, pp. 1–6.

[15] B. Yu and M. P. Singh, "A social mechanism of reputation management in electronic communities," in *Proc. Int. Workshop Cooperat. Inf. Agents*, 2000, pp. 154–165.

[16] K. Aberer and Z. Despotovic, "Managing trust in a peer-2-peer information system," in *Proc. 10th Int. Conf. Inf. Knowl. Manage.*, Oct. 2001, pp. 310–317.

[17] F. Cornelli, E. Damiani, S. D. C. di Vimercati, S. Paraboschi, and P. Samarati, "Choosing reputable servents in a P2P network," in *Proc. 11th Int. Conf. World Wide Web*, May 2002, pp. 376–386.

[18] F. G. Mármol, G. M. Pérez, and A. F. G. Skarmeta, "TACS, a trust model for P2P networks," *Wireless Pers. Commun.*, vol. 51, no. 1, pp. 153–164, Oct. 2009.

[19] N. Stakhanova, S. Basu, J. Wong, and O. Stakhanov, "Trust framework for P2P networks using peer-profile based anomaly technique," in *Proc. 25th IEEE Int. Conf. Distrib. Comput. Syst. Workshops*, Jun. 2005, pp. 203–209.

[20] W. Tang, Y. Ma, and Z. Chen, "Managing trust in peer-to-peer networks," *J. Digit. Inf. Manag.*, vol. 3, no. 2, p. 58, 2005.

[21] J. Li and Q. Li, "Decentralized self-management of trust for mobile ad hoc social networks," *Int. J. Comput. Netw. Commun.*, vol. 3, no. 6, pp. 1–17, Nov. 2011.

[22] B. Qureshi, G. Min, and D. Kouvatsos, "A framework for building trust based communities in P2P mobile social networks," in *Proc. 10th IEEE Int. Conf. Comput. Inf. Technol.*, Jun. 2010, pp. 567–574.

[23] U. Kuter and J. Golbeck, "Sunny: A new algorithm for trust inference in social networks using probabilistic confidence models," in *Proc. AAAI*, vol. 7, 2007, pp. 1377–1382.

[24] M. Maheswaran, H. C. Tang, and A. Ghunaim, "Towards a gravity-based trust model for social networking systems," in *Proc. 27th Int. Conf. Distrib. Comput. Syst. Workshops (ICDCSW)*, 2007, p. 24.

[25] C. Fernandez-Gago, I. Agudo, and J. Lopez, "Building trust from context similarity measures," *Comput. Standards Interface*, vol. 36, no. 4, pp. 792–800, Jun. 2014.

[26] G. Liu, Q. Yang, H. Wang, X. Lin, and M. P. Wittie, "Assessment of multi-hop interpersonal trust in social networks by three-valued subjective logic," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, Apr. 2014, pp. 1698–1706.

[27] J. Caverlee, L. Liu, and S. Webb, "Socialtrust: Tamper-resilient trust establishment in online communities," in *Proc. 8th ACM/IEEE-CS joint Conf. Digit. libraries*, Jun. 2008, pp. 104–114.

[28] S. Toivonen and G. Denker, "The impact of context on the trustworthiness of communication: An ontological approach," in *Proc. Trust@ ISWC*. Princeton, NJ, USA: Citeseer, 2004, pp. 1–10.

[29] M. J. Fischer, N. A. Lynch, and M. S. Paterson, "Impossibility of distributed consensus with one faulty process," *J. ACM*, vol. 32, no. 2, pp. 374–382, Apr. 1985.

[30] P. Pettit, "The cunning of trust," *Philosophy Public Affairs*, vol. 24, no. 3, pp. 202–225, Jul. 1995.

[31] I. Boutet, M. LeBlanc, J. A. Chamberland, and C. A. Collin, "Emojis influence emotional communication, social attributions, and information processing," *Comput. Hum. Behav.*, vol. 119, Jun. 2021, Art. no. 106722.

[32] S. Alizadeh and A. Fazel, "Convolutional neural networks for facial expression recognition," 2017, *arXiv:1704.06756*.

[33] J. Marín-Morales, C. Llinares, J. Guixeres, and M. Alcañiz, "Emotion recognition in immersive virtual reality: From statistics to affective computing," *Sensors*, vol. 20, no. 18, p. 5163, Sep. 2020.

[34] J. A. Russell, "A circumplex model of affect," *J. Personality Social Psychol.*, vol. 39, no. 6, p. 1161, Dec. 1980.

[35] P. Galopoulos, C. Iakovidou, V. Gkatziaki, S. Papadopoulos, and Y. Kompatsiaris, "Towards a privacy respecting image-based user profiling component," in *Proc. Int. Conf. Content-Based Multimedia Indexing (CBMI)*, Jun. 2021, pp. 1–6.

[36] M. Spear, X. Lu, N. S. Matloff, and S. F. Wu, "Inter-profile similarity (IPS): A method for semantic analysis of online social networks," in *Proc. Int. Conf. Complex Sci.*, 2009, pp. 320–333.

[37] A. M. González and L. Bello, "The construct 'lifestyle' in market segmentation: The behaviour of tourist consumers," *Eur. J. Marketing*, vol. 36, nos. 1–2, pp. 51–85, Feb. 2002.

[38] J. K. Rempel, J. G. Holmes, and M. P. Zanna, "Trust in close relationships," *J. Personality Social Psychol.*, vol. 49, no. 1, p. 95, 1985.

[39] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *J. Behav. Therapy Experim. Psychiatry*, vol. 25, no. 1, pp. 49–59, Mar. 1994.

[40] M. Benedek and C. Kaernbach, "A continuous measure of phasic electrodermal activity," *J. Neurosci. Methods*, vol. 190, no. 1, pp. 80–91, Jun. 2010.

[41] U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, "Heart rate variability: A review," *Med. Biol. Eng. Comput.*, vol. 44, no. 12, pp. 1031–1051, Dec. 2006.

[42] M. P. Tarvainen, J.-P. Niskanen, J. A. Lipponen, P. O. Ranta-Aho, and P. A. Karjalainen, "Kubios HRV–heart rate variability analysis software," *Comput. Methods Programs Biomed.*, vol. 113, no. 1, pp. 210–220, 2014.

[43] M. P. Tarvainen, P. O. Ranta-Aho, and P. A. Karjalainen, "An advanced detrending method with application to HRV analysis," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 2, pp. 172–175, Feb. 2002.

[44] M. P. Fay and M. A. Proschan, "Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules," *Statist. Surv.*, vol. 4, p. 1, Jan. 2010.

[45] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, nos. 3–4, pp. 591–611, Dec. 1965.

[46] K. Nash, A. Johansson, and K. Yogeeswaran, "Social media approval reduces emotional arousal for people high in narcissism: Electrophysiological evidence," *Frontiers Hum. Neurosci.*, vol. 13, p. 292, Sep. 2019.

[47] A. Greco, G. Valenza, J. Lázaro, J. M. Garzón-Rey, J. Aguiló, C. de la Cámara, R. Bailón, and E. P. Scilingo, "Acute stress state classification based on electrodermal activity modeling," *IEEE Trans. Affect. Comput.*, vol. 14, no. 1, pp. 788–799, Jan. 2023.

[48] R. Castaldo, P. Melillo, U. Bracale, M. Caserta, M. Triassi, and L. Pecchia, "Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis," *Biomed. Signal Process. Control*, vol. 18, pp. 370–377, Apr. 2015.

[49] J. Gruber, D. S. Mennin, A. Fields, A. Purcell, and G. Murray, "Heart rate variability as a potential indicator of positive valence system disturbance: A proof of concept investigation," *Int. J. Psychophysiol.*, vol. 98, no. 2, pp. 240–248, Nov. 2015.

BARBARA GUIDI received the B.Sc., M.Sc., and Ph.D. degrees in computer science from the University of Pisa, Italy, in 2007, 2011, and 2015, respectively. She is currently an Assistant Professor with the Department of Computer Science, University of Pisa. Her current research interests include decentralized online social network, blockchain technology, and social network analysis. She was the Co-Chair of the Conference EAI GoodTechs, in 2017, and for several workshops. She has been involved in the TPC of several international conferences and workshops. She has been a Reviewer of relevant journals, such as IEEE Access.

ANDREA MICHIENZI received the Ph.D. degree in computer science from Universitã di Pisa, in July 2021. He is currently an Assistant Professor with Università di Pisa. His research interests include the dynamics of socio-economic networks, enabling technologies for the metaverse, and the application of the blockchain to social networking platforms. He is one of the organizers of the Open Challenges in Online Social Networks Workshop.

LAURA RICCI received the M.Sc. and Ph.D. degrees in computer science from the University of Pisa, Italy, in 1983 and 1990, respectively. She is currently a Full Professor with the Department of Computer Science, University of Pisa. She has been involved in several research projects. She is currently the local Coordinator of the H2020 European Project "Helios: A Context Aware Distributed Networking Framework." Her research interests include distributed systems, peer-to-peer networks, cryptocurrencies, and blockchains and social network analysis. In these fields, she has coauthored more than 100 papers published on international journals and conference/workshop proceedings. She has served as the program committee member and the chair for several conferences. She is the Organizer of the Workshop LSDVE.

FABRIZIO BAIARDI received the degree from Università di Pisa. He is currently a Full Professor of computer science with Università di Pisa. He is also the Co-Founder of Haruspex, a startup that develops risk assessment and management tools based upon adversary emulation. He holds some patents on intrusion detection. His research interest includes cyber risk assessment and management.

LUCÍA GÓMEZ-ZARAGOZÁ received the degree in biomedical engineering and the master's degree in artificial intelligence, pattern recognition, and digital imaging from the Polytechnic University of Valencia (UPV), in 2019 and 2021, respectively, where she is currently pursuing the Ph.D. degree in technologies for health and wellbeing. From 2019 to 2021, she was with the Institute for Research and Innovation in Bioengineering, UPV. Her research interests include biomedical signal processing and the application of artificial intelligence in healthcare. Specifically, she is working on speech analysis and natural language processing for psychological assessment.

LUCÍA A. CARRASCO-RIBELLES (Student Member, IEEE) received the bachelor's degree in biomedical engineering from the Polytechnic University of Valencia, Spain, in 2018, and the master's degree in data science from the University of Valencia, Spain, in 2020. She is currently pursuing the Ph.D. degree in signal theory with Universitat Politècnica de Catalunya. Her research interests include the study of data science involving biomedical data, including obtaining machine learning or deep learning algorithms for the automation of processes or the creation of decision support systems.

JAVIER MARÍN-MORALES (Member, IEEE) received the bachelor's and master's degrees in civil engineering and the Ph.D. degree in technologies for health and wellbeing from the Polytechnic University of Valencia (UPV), in 2014 and 2020, respectively. He is currently a Postdoctoral Researcher with the Institute for Research and Innovation in Bioengineering, UPV. His research interests include virtual reality, biomedical signal processing, machine learning, affective computing, and psychological assessment. He is the author of more than 20 international scientific contributions in these fields and has been involved in several international research projects.

Open Access funding provided by 'Università di Pisa' within the CRUI CARE Agreement