Calculating Trust using Multiple Heterogeneous Social Networks

Muhammad Imran[b,*], Hasan Ali Khattak[b,*], David Millard[c], Thanassis Tiropanis[c], Tariq Bashir[1], Ghufran Ahmed[d]

[a]*Department Computer Science, COMSATS University Islamabad 44500, Pakistan*
[b]*Department Electrical and Computer Engineering, COMSATS University Islamabad 44500, Pakistan*
[c]*School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK*
[d]*Department of Computer Science, FAST National University of Computer and Emerging Sciences, Karachi, Pakistan*

Abstract

In today's internet, a web user becomes members of multiple social networks due to different types of services provided by each of these networks. This creates an opportunity to make trust decisions that go beyond individual social networks, since these networks provide single perspective of trust. To make trust inference over multiple social networks, these networks needs to be consolidated. It is non-trivial as these networks are of heterogeneous nature due to different naming conventions used in these networks. Furthermore, trust metrics extracted from these networks are also varied in nature due to different trust evaluation algorithms used in each of these networks. Heterogeneity of these social networks can be overcome by using semantic technologies as it allows us to represent knowledge using ontologies. Trust data can be consolidated by using such data fusion techniques which not only provide but also preserves trust data integrity from each of the individual social network profiles. The proposed semantic framework is evaluated using two sets of experiments. Through simulations in this work we analysed various techniques for data fusion. For identifying suitable technique which preserves the integrity of trust consolidated from each of the individual networks. Analysis revealed that Weighted Ordered Weighted Averaging parameter best aggregated trust data, and unlike other techniques, it preserved the integrity of trust from each individual network for varying Participant Overlap and Tie Overlap ($p \leq 0.05$). Similarly, for experimental analysis we used findings of the simulation study about the best trust aggregation technique and applied the proposed framework on real-life trust data between participants which we extracted from pairs of professional social networks. Analysis partially proved our hypothesis that generating better trust values from consolidated Multiple Heterogeneous networks. We witnessed an improvement in overall results for all the participants who were part of multiple social networks ($p \leq 0.05$), while disproving the claim for those existing in non-overlapping regions of the social networks.

Keywords: Online Social Networks, Multiple Social Networks, Semantic Framework, Trust, Data Fusion

1. Introduction

At present, online social networks (OSNs) replace real-world social networks, where people interact with each other remotely [1, 2]. Due to these social networks, several interactions, and activities which required physical interaction are conveniently possible now while sitting at distant locations through World Wide Web. A survey conducted by Pew[1] in February 2019 about the use of online social networks in the US discloses that adult web users accounting for 72% utilize online social networks for online interaction [3]. The same survey also revealed statistics about the use of multiple social networks. It states that 56% of

*Corresponding authors
Email addresses: mimran@comsats.edu.pk (Muhammad Imran), hasan.alikhattak@comsats.edu.pk (Hasan Ali Khattak)

[1] https://www.pewresearch.org/internet/fact-sheet/social-media/

online adults use more than one social network wherein 95% of twitter users and 92% of LinkedIn users also use Facebook. It indicates the increasing trend of using multiple social networks due to the different nature of services provided by these networks. We describe this situation as individuals belonging to Multiple Heterogeneous (MuHe) social networks; Multiple because there is more than one network structure, and Heterogeneous because the networks represent different types of relationships. MuHe networks usually are made up of networks owned and managed by various organizations, so typically, users have different virtual identities in each system.

The researchers have developed a range of trust algorithms for individual social networks [4, 5]. These examine personal information and interaction history of the users to calculate trust metrics in their respective networks. Such mechanisms have become essential features for some successful social networks; for example, the eBay[2] network evaluates the reputation of sellers based on the ratings of buyers, which helps it to ensure high standards of online shopping. Similarly, expertise recommendation mechanism in the LinkedIn[3] network uses recommendations provided by other professionals in the network, which are of great help for new users.

There are several definitions of Trust in the literature, but we have taken the definition as given by [6] as it is both general and concise: *Trust of party X to a party Y for a service A is a measurable belief of X in that Y behaves dependably for a specified period (and within a specified context in relation to service A)*" [3].

OSNs may be categorised based on the services provided by these networks such as friendship networks, professional networks, e-commerce networks etc. When considered in context of trust, these networks represent multi-faceted information which may be helpful if trust is to be modelled for multiple networks. However, existing trust mechanisms tend to be restricted to a single network, whereas users are typically involved in MuHe networks. Basing Trust decisions on MuHe networks would have two distinct advantages: 1) It increases the chance of calculating a Trust value, as individuals need not share the same network, and 2) It bases trust values upon diversified information that can reflect accurately a user's behaviour on multiple social networking platforms. Our definition of Trust notes that trust is calculated in a specific content, and therefore we expect that MuHe networks will be related to a particular trust domain (such as professional networks), rather than a general aggregation of *all* these networks in which individuals are present.

Unfortunately, the task of linking multiple social networks to generate a single big social network for performing trust calculations is not trivial, reason being varying structures of the networks and weights on the links between professionals on these diverse networks [7, 8]. The aggregation mechanism should not inflate or dampen trust values artificially based on the availability of information from either some or all of the constituent networks. It should be able to particularly differentiate between *absence of trust* from *distrust* as unavailability of information from certain networks does not mean distrust.

This paper proposes a novel framework for performing Trust calculations on MuHe networks. The framework is based on Semantic Web technology, and uses Data Fusion techniques such as Weighted Ordered Weighted Averaging etc to aggregate individual networks without distorting trust values. We then present two evaluations of this framework. The first uses a simulation environment to look at the impact of consolidating two networks on their trust properties (strength of trust ties and length of trust path), based on this simulation we then select the most appropriate Data Fusion technique. The second uses the proposed framework and this chosen technique to conduct a comparative evaluation of trust calculations based on existing single social networks against those based on MuHe networks. It uses real-world trust values elicited from participants as the the gold standard.

The remaining paper is structured as follows: The related work about Semantic Web, Data Fusion and Online Trust is described in Section 2. Section 3 presents the proposed Semantic Web framework, and gives an overview of the data fusion techniques. It further describes the characteristics by which different data fusion techniques can be compared. Section 4 presents the simulation experiment and justifies the choice of Data Fusion technique. Section 5 presents the real-word experiment that compares trust values calculated

[2] http://www.ebay.com

[3] http://www.linkedin.com

using existing single networks to those calculated using MuHe networks. Finally Section 7 presents potential future extensions and concludes the paper.

2. Related Work

There are various studies reported in the literature which attempt at consolidating multiple social networks, but they merely focus on combining these networks rather than exploring their impact on trust-related measures. For example, the work by [9] attempts to merge trust and distrust relations from multiple trust networks but lacks in two dimensions; first, it fails to differentiate between distrust and absence of trust and second, the impact of that consolidation on the accuracy of trust metrics is not examined [3].

In semantic web, the concept of co-reference resolution resolves the problem of users having distinct identities in multiple social networks. There are many existing methods which address this issue such as [10] discusses two methods of URI co-reference resolution, **1)** logical inference and **2)** label comparison. Logical inference matches IFPs (Inverse Functional Properties) to evaluate whether a pair of URIs are co-referred, while label comparison compares data properties to classify URI pairs as co-referred or non-co-referred URIs. The URIs classified as co-referred may be linked using the predicate owl:sameAs provided as part of the OWL DL specification in the Semantic Web. [11, 12]. Trust data can be annotated using either of the URIs defined in the existing network or by using afresh URI generated using the target namespace [13]. The resultant annotated information in the MuHe environment may be published as a separate named graph [14]. It helps those reusing the data to scope down their queries to target graph rather than writing long query patterns over existing graphs.

Multiple trust values emerge between co-referred users when MuHe networks are consolidated. These trust values represent subjective trust in the context of a particular network. While integrating these values, the data fusion algorithm should respect the trust integrity of these social networks. A number of data fusion techniques exist in the state of the art literature such as [15], where the authors have proposed an aggregation operator based upon Ordered Weighted Averaging (OWA) which takes into account the multiple trust values based upon relative importance by ranking data values in descending order. Similarly, another technique, WOWA, proposed by [16] considers importance of both data and their sources. IOWA behaves similar to WOWA but allows to rank data points with respect to different trust sources [17, 18].

Furthermore, there are different implementations in the literature that have used these techniques for aggregating trust. For example, the work by [19] collects trust scores between any two users from multiple paths in a single networks and aggregates them using different data fusion techniques. This scenario is similar to consolidating trust metrics from multiple social networks. The results of the knowledge awarding-OWA (K-OWA) and knowledge awarding-averaging (KAAV) approaches showed better performance when compared with other techniques. Similarly [20] presents aggregation techniques used for generating trust scores over transitive triads in a realistic fashion. Advanced deep learning techniques have been used in order to perform recommendation based upon trust score in social networks [21].

Consolidation of MuHe networks also helps to discover links among isolated users. This emanates the scenario of quantifying indirect trust and trust decay is one of the approaches discussed in the literature. Trust decay uses the principle of trust transitivity, and was first discussed by a psychologist to analyse the existence of transitivity in real-world social networks [22]. The experiment was conducted to test the transitivity of positive interpersonal sentiments. This work was later extended by social psychologists to examine this in terms of social relations by running an experiment over a set of 917 sociograms [23]. A random group of people were asked about their sentiments towards other people in the group. They found that in 70 per cent of the cases there was a strong inclination towards the transitivity.

A number of trust algorithms are developed using the concept of trust decay, for example, [24] proposed an approach for trust and distrust propagation, '*Appleseed*', which uses the theory of spreading activation [25]. According to that, trust or distrust in a friend flows along all paths leading from the friend. They chosen a realistic decay factor based on the empirical experiment and a normalised local edge weight, $e_{x\to y}$, is assigned to each link in the network. The work by [26] uses reinforcement learning for calculating local trust values in social networks. The trust from the source is propagated towards the destination using different strategies. Experimental results revealed that the hybrid approach of weighted mean aggregation and min-max over shortest paths turned out to be the best approach. Similarly, [27] presents an idea of incorporating multiple trust paths of varied lengths for enhancing accuracy of trust calculations. The decay of trust is considered and trust is calculated for both shortest and longest trust paths. Results showed that shorter paths generate more accurate trust measures than longer paths. The method proposed by [28] extracts domain-specific trust network from large scale heterogeneous network, with the claim that trust propagation is a domain-dependent phenomenon and cannot work through heterogeneous relations [3]. Another study elaborates trust decay as a function of leakage in network flow and proposed network flow based trust evaluation scheme *GFTrust* [29]. The work carried out by [30] develops an algorithm for Trust Path Searching (TPS) to evaluate trust for indirectly connected users using the principle of trust transitivity. The research conducted by [31] summarize different trust decays methods used by the literature and provide comparative analysis of these techniques.

3. Proposed Semantic Web Framework

The proposed framework uses semantic web technology to model constituent networks in uniform format, and then allows different data fusion techniques to be used to integrate them. The key idea behind semantic web is that each resource should have a unique identifier (URI), but in practice, different online networks have different namespaces for knowledge representation. Resultantly, individuals have multiple URIs across MuHe social networks. The proposed framwork should resolve multiple URIs which refer to the single user in the real-life. Further, it should provide a mechanism to aggregate multiple trust values that originate between users due to their existence in MuHe networks.

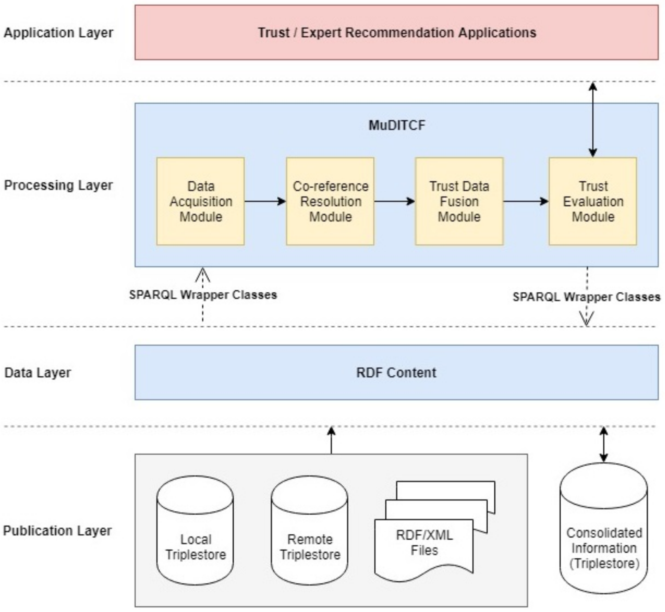

Figure 1: A Semantic Web Framework for building trust applications over MuHe social networks.

The MuHe Consolidation Framework is shown in Figure 1 which presents our solution for building trust applications over multiple distributed social networks. The experimental setup to test the feasibility of framework for linking multiple networks resides on top of the Sesame triplestore [4]. A number of preprocessing

[4] http://rdf4j.org/

modules are written in Python[5] which gather data from multiple sources (including local RDF files, local and remote triplestores), and converts it into a single semantic representation for consolidation. It then applies different trust evaluation algorithms for calculating trust metrics for both directly and indirectly connected users.

The first module, named *Data Acquisition Module* uses Python SPARQL wrapper classes[6] for extracting RDF data between users from MuHe networks. It particularly targets information which can help in building linked networks with nodes of the the network represent people and weights on the links show trust ranks between them.

The remaining three modules are more complex, and are described in the following subsections.

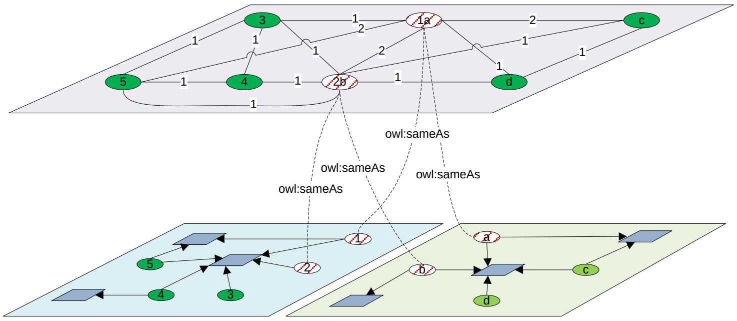

Figure 2: Linked trust graph shown as an overlay network over individual networks. *owl:sameAs* predicates are used to co-refer newly assigned URI in overlapping region of consolidated graph to individual graphs.

3.1. Co-reference Resolution Module

The co-reference module locates different URIs from MuHe networks (defined in different ontologies) that represent same person and allows us to generate afresh URI from the target namespace (defined in our own ontology). The resultant annotations may be added as a separate graph linked with existing graphs using *owl:sameAs* predicates. The concept of named graph[14, 32] is helpful when consolidating MuHe Networks. It allows us to represent consolidated networks as a separate layer built over existing graphs acting as an overlay network. The sample representation of such a network is shown in Figure 2. The individual networks layer represents two publication networks where square shapes represent publications while researchers are depicted using oval shape. This layer is linked with existing published information using *owl:sameAs* statements specifying which of the URIs in the consolidated network correspond to URIs in each of the individual networks. Once published, this eliminates the need to write long query patterns over individual networks for co-referencing or retrieving trust information between users.

The intuitive way to identify co-refered users from these networks is to compare meta-data. In semantic web, this information can be extracted from *owl:DP* (Data Property) predicates. Equation 1 provides a rule to perform co-reference resolution for two URIs *?1* and *?a*. Here, *?1* and *?a* are URIs in individual networks and *?1a* is the newly allocated URI in the aggregated single network. *?p* represents the set of data properties selected for comparison. If both the data properties hold same values for both the URIs *?1* and *?a*, then it resolves them to represent same person in multiple networks and the resultant URI may be *?1a* in the consolidated version of these networks. Further, it can be linked with both *?1* and *?a* in individual networks using *owl:sameAs* predicate, thereby stating that both these URIs are the same. Note that the numeric value *?1* and a character *?a* in individual networks and *?1a* in consolidated version are shorthand of the original URIs).

$$\{?p \quad a \quad owl:DP. \qquad ?1 \quad ?p \quad ?x. \qquad ?a \quad ?p \quad ?x.\} \Rightarrow \begin{cases} ?1a \quad owl:sameAs \quad ?1. \\ ?1a \quad owl:sameAs \quad ?a. \end{cases} \tag{1}$$

[5] https://www.python.org/

[6] http://rdflib.github.io/sparqlwrapper/

An ontology is needed for making annotations in the consolidated graph. Figure 3 presents the classes and properties defined in the proposed ontology. It extends existing ontology given in [33] and adds object and data properties which are particularly needed for annotating trust in the context of MuHe networks. The single new class *TrustRelationship* defines the trust relationship between instances of *Trustor* and *Trustee* classes using *has_trustor* and *has_trustee* properties. The *Trustor* class holds the identity of trustor while the trustee class represents the person being trusted by the trustor. Both the trustor and trustee are persons, so their URIs are generated using *Person* class of the *FOAF* namespace.

The ontology allows us to model trust in three different formats: absolute, processed and fuzzy. Absolute trust is the subjective trust value extracted from any particular network. For example, consider the example of a publication network where trust between researchers is measured in terms of co-authorship frequency, in this case absolute value is the count of the number of times that a pair of researchers has appeared in publications together. Processed value in this context is an absolute value normalised in the range between 0 and 1. It is a translated value that transforms trust from multiple contexts into a format that can be compared or combined. Fuzzy trust is the human understandable version of numerical trust values such as, high trust, medium trust, etc. There are various techniques of modelling fuzzy trust, however, trust modelling using fuzzy logic is out of scope of this work.

Trust in MuHe social networks can be viewed as a set of direct experiences from multiple social networks or recommendations provided by other members of those networks. In our ontology, it is modelled using *has_type* data property where the *domain* of the property is the URI of the relevant person and *range* being the string value of "direct" or "indirect". The direct trust is extracted based on the direct interactions (which can be either of the explicit or implicit activities) whereas indirect trust is calculated between isolated users, typically belonging to multiple social network. There are various methods of calculating indirect trust, so the *has_process* object property records the technique used and the *has_pathLength* object property stores the length of the trust path involved.

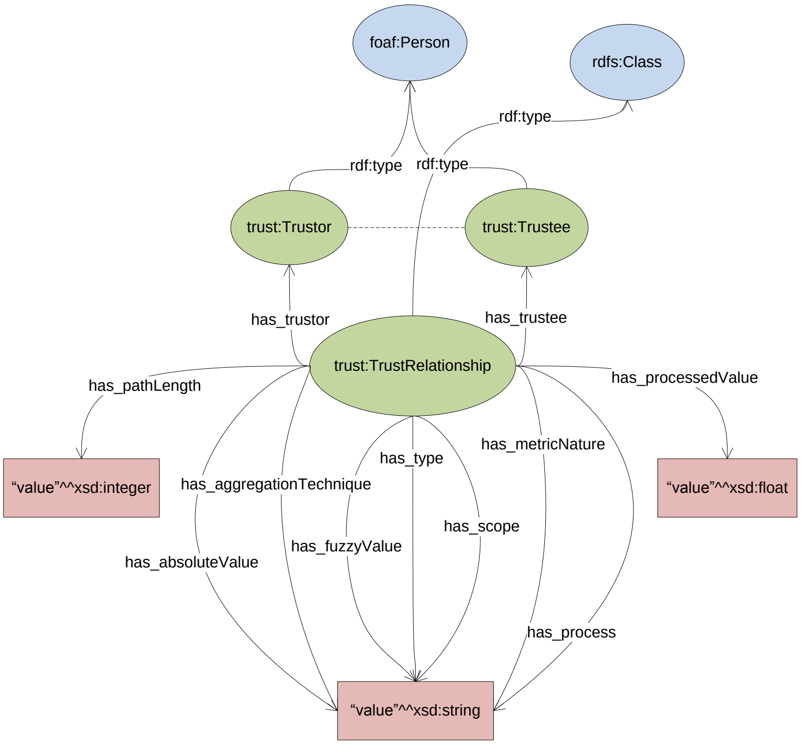

Figure 3: Trust ontology for trust over MuHe social networks.

The trust definition used in this work takes it as a subjective value, so the *has_scope* data property specifies the area that any particular trust value belongs to. It is particularly important when trust data only related to certain area needs to be consolidated such as, if the system has to aggregate trust information about trusted academics in the field of semantic web, machine learning etc.

The outcome of the co-referencing technique is a linked network, having consolidated URIs for overlapped users, and multiple trust values between them extracted from MuHe networks. The next step is to combine these multiple trust values into a single one, using the data fusion technique. The ontology must therefore also record this process, which is modeled using two properties; *has_metricNature* and *has_aggregationTechnique*. The former defines whether the target pair of users were *overlapping* or *non-overlapping*, so have to establish whether there are multiple values to aggregate or only a single value to be re-evaluated. The later one stores the name of the data fusion technique used to aggregate if multiple trust values are present or re-evaluate if only singular value is available.

3.2. Trust Data Fusion Module

Data fusion module aggregates trust values between users belonging to multiple social networks and generates a single value that represents multifaceted trust. The task of consolidating multiple trust relationships is basically a data fusion problem and particularly it may be considered as a data aggregation case study. One of the simplest approaches is to perform Summation (S) or an Average (A) of the trust values. This may work for simple numerical quantities, but in the case of trust it could damage trust values by either inflating them unnecessarily (through Summation) or deflating them drastically (as some of the trust values may be effectively zero due to missing information from some of the networks being consolidated). Therefore more comprehensive techniques are needed to aggregate trust values because they represent trust in varied context and naive methods of aggregation can distort the integrity of trust.

The trust metrics between a pair of users can range from *one* to the n number of networks involved in consolidation. Suppose that N_p represents a pair of users from multiple social networks and T_{N_p} is the set of trust values between them, from n social networks i.e. $T_{N_p} = \{T_{N_p1}, T_{N_p2} \ldots T_{N_pn}\}$. There are two parameters involved in trust aggregation, 1) importance of the trust values from individual social networks, denoted as $\omega = \{\omega_1, \omega_2, \ldots \omega_n\}$ and 2) importance of the sources of that information denoted as $p = \{p_1, p_2, \ldots p_n\}$. The values of both these parameters stay in the range [0,1], where zero value shows low importance while a value of *one* represents higher importance. The function $f(T_{N_p})$, aggregates these values and generates a single value T^{N_p} by considering ω_i and p_i for each T_{N_pi} in the data set T_{N_p}.

There are various data fusion techniques discussed in the literature and Section 2 presents their detailed background. A richer description of these methods can be found in [34] which reported on a subset of the simulation experiment but the full experiment focused on the following, all of which have been implemented within the data fusion Module:

Weighted Average (WA) technique consolidates multiple trust values by only considering the importance of data sources. The individual trust data points in T_{N_p} are aggregated by multiplying them with the weights of the corresponding sources. The importance of each source network is represented using a weight vector p, and its size is equal to the number of trust values.

Ordered Weighted Averaging (OWA) works in a similar way to WA, but rather than the weights being associated with a given source, the sources are ordered (according to which is most trusted) and the weight is associated with the position within the ordering. The trust data points in T_{N_p} are sorted from high to low order, then already permuted trust values are multiplied with the corresponding weights from the weight vector w to generate a single aggregated value T^{N_p}.

Weighted Order Weighted Average (WOWA) associates importance parameters to both source and the order of the trust values. It takes two sets of weights other than the set of trust values T_{N_p}: first is the weight vector w which shows the importance of trust data points; and second is the weight vector p to show the relevancy of the sources of trust information. The vector w can have both integer and fraction values while T_{N_p} and p vectors are continuous values in the range [0,1].

3.3. Trust Evaluation Module

The trust between isolated users is calculated using trust propagation algorithm. It works on the principal of transitivity and existing studies (mentioned in Section 2) empirically prove that trust decreases as the length of path between indirectly connected users increases. In short, people, in both real life and in online social networks, have high degree of trust towards friends of their friend rather than strangers, but this trust decreases as the length of trust path increases. Drawing on this knowledge, we have used a transitive decay-based trust calculation within the module.

The first step is to calculate all possible trust paths between any two users, the trust value associated with a given path is the result of multiplying all the trust values (in between [0,1]) between all the nodes in that path. There are then two possible options for choosing the trust path. The first is the Strongest Path where the algorithm returns the trust path having the maximum strength of the trust ties between users without considering its length i.e. the number of users involved in the path. The second is the Shortest Path, which chooses the path with the shortest length, regardless of its trust value. However, if multiple trust paths of the same shortest length exist, then the one with the highest trust value is chosen. In this work, we have used the shortest path approach for experimentation while the strongest path approach is already published in [34].

4. Experiment 1: Simulation Experiment for trust inference over consolidated MuHe Social Networks

The simulation experiment discussed in this paper is an extension of the work as described in [34, 3] and further extends that work to include the OWA Data Fusion Method and shortest trust path algorithm. The simulation examines the impact of the consolidation of MuHe networks on trust properties used by the shortest trust path algorithm (described in Section 3.3 above). The simulation is based around a consolidation of pairs of networks, using four data fusion techniques starting from naive ones such as Summation (S) and Weighted Average (WA) and then extending up to more complex techniques such as Weighted Ordered Weighted Averaging (WOWA) and Ordered Weighted Averaging (OWA). The technique which best satisfies trust properties qualifies to be used for making trust computations over real world data. Both the simulation and real-world experiments are implemented using the NetworkX[7] library of the Python programming language. It includes the code for generating networks, consolidation them, measuring network properties and applying trust inference algorithms.

4.1. Experiment Design

The pairs of social networks were generated, with randomly assigned trust values on the links between users, having a varying percentage of participant overlap (PO) and tie overlap (TO) in each pair of networks. It was to assess the values of trust properties when the networks with varying percentage of overlaps were consolidated. N1 and N2 represent the original networks generated by the simulation, MuHe was the final linked network while CN1 and CN2 represent sub-networks in the MuHe mapped to the N1 and N2. We can then see the impact of consolidation by comparing CN1 to N1 and CN2 to N2.

The impact of consolidation on trust properties is measured using the approximation of two metrics Average Tie Strength (TS) and Average Tie Length (TL). TS is the average trust of the network using shortest trust paths, and TL is the average length of the shortest trust paths in the network. Ideally, TS metric from CN1, CN2 and MuHe should be similar to that of N1, N2, even if there is no significant PO and TO. This indicates that trust is not being inflated in the consolidation process. Furthermore, it is desirable that due to the emergence of additional trust paths, TL metric will overall decrease in CN1, CN2 and MuHe as compared to N1 and N2. If TS remains steady and TL reduces, it may be deduced that the consolidation has successfully enhanced trust calculations by opening up new trust paths without escalating trust values in the network. More formally we can say that any data fusion technique chosen for trust should satisfy the following set of propositions (adapted from [15, 16, 17]):

[7] http://networkx.github.io/documentation/latest/reference/introduction.html

Proposition 1 (Boundary Conditions): The trust aggregation function should keep consolidated trust value within the maximum and minimum range.

$$min\{T_{N_p1}, T_{N_p2}, \dots T_{N_pn}\} \quad \leq \quad f(T_{N_p1}, T_{N_p2}, \dots T_{N_pn}) \quad \leq \quad max\{T_{N_p1}, T_{N_p2}, \dots T_{N_pn}\} \tag{2}$$

Proposition 2 (Idempotence): The resultant value of trust aggregation function should be equal to T_{N_p1} if all the trust values are same, that is, $x \in T_{N_p}$:

$$f(T_{N_p1}, T_{N_p1}, \dots T_{N_p1}) = T_{N_p1} \tag{3}$$

Proposition 3 (Monotonicity): The trust aggregation function should be monotonic, which means that it should high aggregated trust for high trust values as compared to low trust values:

$$f(T_{N_p1}, T_{N_p2}, \dots, T_{N_p1}) \geq f(T_{N_q1}, T_{N_q2}, \dots, T_{N_qn}) \quad if \quad T_{N_pi} \geq T_{N_qi} \quad for \quad i = \{1, 2, \dots, n\} \tag{4}$$

The final property, known as *Trust Absence*, ensures the integrity of trust from individual networks. It refers towards the missing trust information from any of the constituent networks and recommends to consider it as the *absence of trust* and not distrust between individuals.

Proposition 4 (Trust Absence): The trust aggregation function should differentiate between *absence of trust* and a *distrust.* The numeric value of zero, in this study, represents *absence of trust* information from either of the individual networks, so their aggregate should generate trust value which is approximately similar to the one generated without that numeric zero.

$$f(T_{N_p1}, 0, \dots, T_{N_pn}) \approx f(T_{N_p1}, \dots, T_{N_pn}) \tag{5}$$

This simulation generates networks for four different participant overlap PO percentages i.e. 40%, 60%, 80%, and 100%, and then for each value of PO (except 40%PO) TO is varied from 0 to PO in increments of 20%. In each of the simulation, trust information on the links are aggregated using different aggregated schemes named - S, WA, WOWA and OWA. Table 1 provides detail about network and consolidation parameters used in this simulation.

Table 1: Network and consolidation parameters used for this study

Network Parameters	Description
Number of nodes	30
Density of networks (D)	0.43
Averaging clustering coefficient of networks (C)	0.45 ± 0.02
Average length of shortest paths in networks (L)	1.57
Ratio of C, D, L between $N1$ and $N2$	1

Consolidation Parameters	Description
Participant Overlap (PO)	$[40\%, 100\%]$
Tie Overlap (TO)	$[0, PO]$

4.2. Results and Analysis

In our analysis we considered the impact of consolidation on both average strength of trust ties (TS), and average length of trust path (TL). These are discussed separately in the following sections.

4.2.1. Average Strength of Trust Ties

The results in Table 2 present the values of TS metric for varying consolidation parameters PO and TO using the shortest path algorithm. It shows that the TS metric using WOWA approach has more stable measurements for CN1, CN2 and MuHe than all the other approaches (S, WA and OWA). Simple techniques S and WA severely distort TS metric for two extreme values i.e. at 100%PO and 100%TO with a value of 0.93 and at 40%PO and 0%TO, it stood at 0.17. Similarly, OWA also performs poorly for low participant and tie overlap and its value for MuHe at 40%PO and 0%TO dropped to 0.22. The TS metric recorded by WOWA approach remained stable throughout varying values of PO and TO and it stood at 0.66 for 100%PO and 100%TO and 0.44 for 40%PO and 0%TO. Based on the results of all the data fusion techniques, WOWA appears to be the better technique for aggregating trust information.

Table 2: TS for shortest path trust evaluation algorithm using four different data fusion techniques with varying percentage of participant overlap and tie overlap. CN1 and CN2 represent original networks N1 and N2 in the consolidated version of MuHe networks.

		Average Strength of Ties in the Network (TS)															
PO	*TO*	N1	N2	CN1				CN2				MuHe			
				S	WA	WOWA	OWA	S	WA	WOWA	OWA	S	WA	WOWA	OWA
40	0	0.53	0.55	0.56	0.21	0.48	0.44	0.58	0.21	0.50	0.17	0.52	0.17	0.44	0.22
	20	0.52	0.55	0.59	0.24	0.47	0.39	0.64	0.26	0.52	0.15	0.54	0.19	0.44	0.19
	30	0.55	0.52	0.66	0.26	0.52	0.42	0.61	0.25	0.49	0.14	0.55	0.19	0.44	0.20
60	0	0.58	0.55	0.61	0.23	0.53	0.51	0.60	0.23	0.52	0.28	0.58	0.20	0.49	0.30
	20	0.55	0.58	0.65	0.27	0.53	0.39	0.67	0.27	0.55	0.19	0.62	0.23	0.51	0.24
	40	0.55	0.54	0.68	0.30	0.54	0.42	0.66	0.29	0.53	0.21	0.61	0.24	0.49	0.25
	60	0.46	0.55	0.63	0.29	0.48	0.37	0.73	0.32	0.56	0.23	0.61	0.23	0.46	0.23
80	0	0.56	0.54	0.59	0.24	0.51	0.50	0.59	0.25	0.52	0.38	0.59	0.23	0.51	0.38
	20	0.61	0.56	0.69	0.29	0.57	0.39	0.67	0.29	0.55	0.27	0.67	0.27	0.55	0.30
	40	0.57	0.60	0.74	0.34	0.59	0.42	0.75	0.34	0.59	0.30	0.71	0.31	0.57	0.32
	60	0.55	0.54	0.76	0.38	0.57	0.45	0.76	0.38	0.57	0.34	0.70	0.32	0.53	0.33
	80	0.50	0.53	0.74	0.38	0.55	0.44	0.78	0.39	0.58	0.37	0.69	0.32	0.51	0.32
100	0	0.58	0.56	0.59	0.27	0.53	0.53	0.59	0.27	0.53	0.53	0.59	0.27	0.53	0.53
	20	0.54	0.53	0.64	0.29	0.54	0.32	0.64	0.29	0.54	0.32	0.64	0.29	0.54	0.32
	40	0.55	0.59	0.72	0.34	0.59	0.37	0.72	0.34	0.59	0.37	0.72	0.34	0.59	0.37
	60	0.52	0.55	0.77	0.38	0.59	0.40	0.77	0.38	0.59	0.40	0.77	0.38	0.59	0.40
	80	0.57	0.58	0.87	0.46	0.65	0.48	0.87	0.46	0.65	0.48	0.87	0.46	0.65	0.48
	100	0.58	0.52	0.93	0.50	0.66	0.53	0.93	0.50	0.66	0.53	0.93	0.50	0.66	0.53

The statistical significance of the apparent preservance of trust integrity by WOWA is assesed using a two-tailed paired T-Test. It evaluates whether the the results of WOWA approach are significantly better results than the other two techniques. This test generates a p-value and a value of p $\leq$ 0.05 indicates significance. Table 3 shows p-values for two types of participant overlaps *PO*, the first four rows show p-values for varying percentage of *PO* while the last measurement i.e. *overall* shows the collective performance of the system by including *TS* metrics for all percentages of *PO*. The analysis of the p-values reveal that the claim of WOWA being better than other two techniques WA and IOWA is statistically significant and holds true for all values of *PO*.

4.2.2. Average length of trust ties (TL)

Table 4 presents the results of the TL metric for varying percentages of *PO* and *TO* using the shortest path trust algorithm. It shows that the values of TL metric for CN1, CN2 and MuHe are same for similar participant and tie overlaps across all the data fusion techniques. For sub-networks CN1 and CN2 it decreases with the increase in PO and TO due to emergence of new trust paths. But for MuHe, it only decreases when $PO \geq 80$ and then it starts increasing again with an increase in *TO*. The reason being that less value of

Table 3: T-Test results (p-value) between corresponding TS metrics of WOWA and WA, OWA for shortest path algorithm.

PO %	WA			OWA		
	CN1	CN2	MuHe	CN1	CN2	MuHe
40	< 0.010	< 0.010	< 0.010	0.050	< 0.010	< 0.010
60	< 0.010	< 0.010	< 0.010	0.040	< 0.010	< 0.010
80	< 0.010	< 0.010	< 0.010	0.020	< 0.010	< 0.010
100	< 0.010	< 0.010	< 0.010	< 0.010	< 0.010	< 0.010
Overall	< 0.010	< 0.010	< 0.010	< 0.010	< 0.010	< 0.010

participant overlap between the networks creates bottleneck due to more number of non-overlapping nodes which results in longer trust paths between users. And this trend reduces when the value of PO increases. When compared with original networks N1 and N2, both having an average length of trust paths (TL) of 1.57, TL is higher (i.e. 1.69) than both the original networks when the participant overlap and tie overlaps are 40% and 0% respectively. It becomes even higher (i.e 1.79) when the value of PO and TO reaches 60%. The number of new shortest paths becomes maximum at $[100\%PO, 0\%TO]$ and as a result TL drops to 1.14 which was (1.57, 1.57) in (N1, N2). However, at 100% participant and tie overlaps, it again becomes equal to N1 and N2 due to consolidated MuHe networks being exact similar to original networks.

Table 4: TL for shortest path trust algorithm using four different data fusion techniques with varying percentage of *PO* and *TO*. CN1 and CN2 represent original networks N1 and N2 in the consolidated version of MuHe networks.

PO	*TO*	Average Length of Trust Paths (TL)													
		N1	N2	CN1				CN2				MuHe			
				S	WA	WOWA	OWA	S	WA	WOWA	OWA	S	WA	WOWA	OWA
40	0	1.57	1.57	1.52	1.52	1.52	1.52	1.50	1.50	1.50	1.50	1.69	1.69	1.69	1.69
	20	1.57	1.57	1.56	1.56	1.56	1.56	1.54	1.54	1.54	1.54	1.76	1.76	1.76	1.76
	30	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.79	1.79	1.79	1.79
60	0	1.57	1.57	1.44	1.44	1.44	1.44	1.43	1.43	1.43	1.43	1.57	1.57	1.57	1.57
	20	1.57	1.57	1.48	1.48	1.48	1.48	1.50	1.50	1.50	1.50	1.62	1.62	1.62	1.62
	40	1.57	1.57	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.68	1.68	1.68	1.68
	60	1.58	1.58	1.55	1.55	1.55	1.55	1.56	1.56	1.56	1.56	1.79	1.79	1.79	1.79
80	0	1.57	1.57	1.33	1.33	1.33	1.33	1.30	1.30	1.30	1.30	1.41	1.41	1.41	1.41
	20	1.57	1.57	1.39	1.39	1.39	1.39	1.38	1.38	1.38	1.38	1.47	1.47	1.47	1.47
	40	1.57	1.57	1.46	1.46	1.46	1.46	1.46	1.46	1.46	1.46	1.53	1.53	1.53	1.53
	60	1.57	1.57	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.59	1.59	1.59	1.59
	80	1.62	1.60	1.59	1.59	1.59	1.59	1.56	1.56	1.56	1.56	1.73	1.73	1.73	1.73
100	0	1.57	1.57	1.14	1.14	1.14	1.14	1.14	1.14	1.14	1.14	1.14	1.14	1.14	1.14
	20	1.57	1.57	1.23	1.23	1.23	1.23	1.23	1.23	1.23	1.23	1.23	1.23	1.23	1.23
	40	1.57	1.57	1.31	1.31	1.31	1.31	1.31	1.31	1.31	1.31	1.31	1.31	1.31	1.31
	60	1.57	1.57	1.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40
	80	1.57	1.57	1.49	1.49	1.49	1.49	1.49	1.49	1.49	1.49	1.49	1.49	1.49	1.49
	100	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57

The results of the TL metric show that the aggregated version of the networks i.e. MuHe network is basically dependent on the *PO*. When the percentage of *PO* is low, a path bottleneck exists in the consolidated network, which makes the TL metric worse as compared to original networks N1 and N2. This is due to the less number of fresh trust paths being generated as a result of low value of participant overlap. When the value of *PO* increases, it causes *TL* to fall due to increase in number of overlapped users, which reduces the bottleneck issue. Furthermore, *TL* is lower in each of the sub-networks CN1 and CN2 than in the corresponding original networks N1 and N2 respectively, irrespective of the value of *PO*. This shows the emergence of additional trust paths due to consolidation. The results from this simulation and the one discussed in [34] proves that the WOWA consolidation is the best approach along all the varying overlap values, while other techniques show poor performance for certain values of PO and TO. At low participant overlap, it respects the integrity of trust (as measured by stability in TS, average strength of ties) while creating new trust paths (as measured by decrease in TL, average length of trust paths).

5. Experiment 2: Network Trust vs Declared Proxy Trust

The simulation demonstrates that in principle the WOWA aggregation technique can result in a MuHe network with improved trust characteristics, however without applying the technique to real world networks it is impossible to evaluate whether the trust values that emerge as a result are better than those calculated by single networks. To test this we undertook an evaluation that compared trust values calculated from two real-world networks, and their consolidation, against trust values obtained from directly surveying members of that network.

Networks with pure trust values are rare, so a pair of professional social networks are extracted from publication and projects domain. These networks are managed by the University of Southampton and they represent proxy trust between users using the co-authorship and collaboration frequencies. The assumption being that individuals who had worked together (manifest as joint publications, or participation in the same project) would exhibit higher trust. The co-authorship frequency was extracted from the ePrints'[8] publication network while collaboration frequency was taken from the public catalogue of research projects undertaken by the WAIS (Web and Internet Science) research group at Southampton [9], and contained details about the projects and staff associated with those projects. The dataset included information about both active and past projects that are being completed under the WIAS research group.

Both networks are available online in RDF format, and these were transformed into the ontology described in Figure 3 and passed into the MuHe consolidation framework (described in Section 3). This then performs the WOWA aggregation and links the resultant consolidated graph with the existing graphs using *owl:sameAs* predicates.

Both of these networks are in the same environment, so there is a significant PO (Participant Overlap) and TO (Tie Overlap), with the collaboration network nearly a subset of the co-authorship network, discounting users outside the university who work on projects. The PO and TO with respect to the WAIS were 51% and 78% while for the Eprints, they were 2% and 1.4% respectively. Table 5 shows description of the network parameters used in this experiment. Both these networks contain bidirectional symmetric trust, as co-authorship and collaboration represents the same trust values in both directions.

Table 5: Network and consolidation parameters used for the real-world experiment for measuring the accuracy of aggregated trust.

Parameters	ePrints	WAIS
N	3286	154
PO	2%	51%
TO	1.4%	78%

5.1. Experiment Design

A survey experiment is designed to test the accuracy of the proposed framework for real world social networks. It collects proxy trust values between users in the professional context which are then compared with the trust measurements from the original and consolidated pair of networks. This is a web application (developed using Django[10] framework) which first examines the presence of user in one or both of the

[8] http://www.eprints.soton.ac.uk

[9] http://www.wais.ecs.soton.ac.uk/projects

[10] https://www.djangoproject.com/

networks and then presents each user with a randomly selected set of related people from these networks, based on the presence of the user in these networks. A set of questions were asked that represent proxy trust which helped us to measure the level of trust between them.

5.1.1. Participants

The designed survey has two types of participants. The first set of participants are known as rating participants and they represent those taking part in the survey. The second group of people comprised of those about whom rating participants expressed their trust by answering proxy trust questions. These set of people are known as rated participants. The rated participants were selected from the ego-centric network created by taking rating participant as an ego-node and randomly selecting set of users. As the simulation experiment claims the existence of trust decay along paths in social networks, so the accuracy of indirect trust was evaluated by selecting rated participants belonging to path lengths of one, two and three. If the rating participant was present in both the selected networks, then four of the rated participants were selected from each of the networks, otherwise all the eight rated participants were selected from one of the networks.

5.1.2. Questionnaire

The survey aimed to extract the trust that participants feel towards one another in a professional context. The substance of a trust survey can be ethically difficult, as participants can be unwilling to disclose genuine trust ratings for others, our solution was to ask less sensitive questions about the closeness of professional ties, and to treat these as proxy trust values. There are two questions asked to each of the rating participants. *First* one is about the *past work* experience with some randomly chosen related person and, *second* question is about the likelihood of them *working together in future* should there be the opportunity. The numerical data from this portion of the survey was in the range (0,0.8) with the value of 0.8 corresponding to working *Very closely* while 0 means working *Hardly at all.* The survey values are then compared with the data available from the system, which is already in the range (0, 1). The *last* question in the survey asked rating participant to briefly explain their relationship with each of the rated person separately. This was so we could explore whether different categories were more accurately represented by the proxy networks, or improved by the consolidated network.

5.1.3. System and Survey Trust Metrics

Figure 4 shows the number of nodes (people) and ties (relationships) in the eprints and project networks used for the experiment. Of these a total of 26 individuals participated in this survey experiment, on average each rating participant provided trust ratings on 3.15 out of 5.38 rated participants presented. The maximum number of trust ratings were provided for path length one, and as might be expected, this number decreased as the length of path increased.

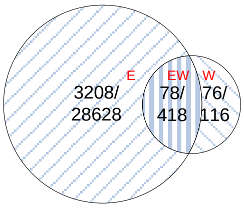

Figure 4: The two real world professional networks (E for ePrints and W for WAIS) are shown as a Venn Diargam (including their intersection EW), with the number of nodes/ ties shown for each region.

As discussed in Section 3.2, trust data from consolidated MuHe networks can be categorised as of two types, complete and partial. With reference to Figure 4, complete data originated from the region EW, and user pairs belonging to this region are known as $N_p^{overlapping}$ users, while partial trust information came either from regions E or W, or across different regions, for instance E $\longrightarrow$ EW etc, and user pairs belonging to this category are known as $N_p^{non-overlapping}$ users. Table 6 shows different trust parameters of the experiment along with their values. It can be categorised as system and survey readings.

Table 6: Description of the trust parameters obtained from system and survey experiments of the real-world experiment. The relationship terms Team Member, ECS Colleague, WAIS Colleague and Supervisor are abbreviated as TM, EC, WC and SP respectively.

Data Type	Trust Ratings	Range
System readings	Eprints Co-authorship proxy trust ($trust^{eprints}$)	(0,1)
	WAIS projects collaboration proxy trust ($trust^{wais}$)	(0,1)
	Consolidated proxy trust ($trust^{muhe}$)	(0,1)
Survey readings	Past proxy trust ($trust^{past}$)	[0, 0.2, 0.4, 0.6, 0.8]
	Future proxy trust ($trust^{future}$)	[0, 0.2, 0.4, 0.6, 0.8]
	Relationship (Rel)	[TM, EC, WC, SP]

System readings are the one generated by evaluating trust over ePrints and WAIS networks, while survey readings are collected from the corresponding real-world users. For the two categories of users specified above, there are always two trust values available from the survey experiment, but there are a variable number of trust values available from system readings. The trust values extracted from the survey are represented as $trust^{past}$ and $trust^{future}$ for representing *past trust* and *future trust* respectively. There are three system-generated trust values for overlapping pairs of participants, represented as $N_p^{overlapping}$; two from each of the individual networks ePrints ($trust^{eprints}$) and WAIS ($trust^{wais}$), and one from the consolidated version, represented as $trust^{muhe}$. For $N_p^{non-overlapping}$ pairs of participants, however, there were only two values available, one from either of the individual networks ePrints ($trust^{eprints}$) or WAIS ($trust^{wais}$) based on the user presence and the other from their consolidated version, represented as MuHe ($trust^{muhe}$) networks. Beside this, there was one additional parameter which describes the type of relationship (Rel) between each user pair. This was to asses us the accuracy of aggregated trust metric with respect to each of the relationship categories.

5.2. Results and Analysis

Results of this experiment were thoroughly analysed to test whether the trust metrics from MuHe networks ($trust^{muhe}$) are *closer* to real-life trust metrics ($trust^{past}$, $trust^{future}$) than from individual networks ($trust^{eprints}$, $trust^{wais}$). The test was conducted by first taking the mean of the absolute difference between system and survey readings for each category of users from individual and consolidated networks, and then evaluating p-value between datasets using T-Test, to deduce whether the difference is statistically significant.

5.2.1. Overlapping users' data

To analyse the overlapping users' data for improvement, we calculated the mean value of the absolute differences between system and survey readings. Table 7 presents means for strongest and shortest path algorithms. It shows that for both the algorithms, EP ($trust^{eprints} - trust^{past}$) and WP ($trust^{wais} - trust^{past}$) are larger than MP ($trust^{muhe} - trust^{past}$) which manifests that consolidated MuHe networks reduced the difference between system and survey metrics and brought trust metrics closer to real-life metrics than individual networks. Similar results are demonstrated by EF ($trust^{eprints} - trust^{future}$ and WF ($trust^{wais} - trust^{future}$) when compared with the MF ($trust^{muhe} - trust^{future}$).

To test whether the apparent improvement of $trust^{muhe}$ over $trust^{ePrints}$ and $trust^{wais}$ is statistically significant, p-value was calculated by conducting one-tailed paired T-Test over the set of absolute differences between system and survey readings. If $p \leq 0.05$, we judge that this improvement is statistically significant for this dataset. The Table 8 presents the evaluated p-value. Results show that $p \leq 0.05$ for all category of users, which proves the claim of generating trust metrics from consolidated MuHe metrics being closer to real-life trust metrics for $N_p^{overlapping}$ true for this dataset.

Table 7: Mean (M) of the difference between the system and survey readings for shortest path trust algorithm.

User Category	EP	WP	MP	EF	WF	MF
Mean	0.23	0.23	0.20	0.23	0.22	0.19
EP = $(trust^{eprints} - trust^{past})$ EF = $(trust^{eprints} - trust^{future})$						
WP = $(trust^{wais} - trust^{past})$ WF = $(trust^{wais} - trust^{future})$						
MP = $(trust^{muhe} - trust^{past})$ MF = $(trust^{muhe} - trust^{future})$						

Table 8: p-value to evaluate the statistical significance of *closeness* between system and survey readings for overlapping $(N_p^{overlapping})$ users.

	EP	WP
MP	**0.01**	**0.02**
	EF	WF
MF	$<$ **0.01**	**0.01**
EP = $(trust^{eprints} - trust^{past})$ EF = $(trust^{eprints} - trust^{future})$		
WP = $(trust^{wais} - trust^{past})$ WF = $(trust^{wais} - trust^{future})$		
MP = $(trust^{muhe} - trust^{past})$ MF = $(trust^{muhe} - trust^{future})$		

5.2.2. Non-overlapping users' data

As mentioned earlier, non-overlapping set of user pairs have partial trust information, so there is only one system trust metric available between them other than the one obtained from the consolidated network.

Table 9: Mean (M) of the difference of system and survey readings for shortest path trust algorithm.

EP	MP	EF	MF
0.31	0.32	0.21	0.22
EP = $(trust^{eprints} - trust^{past})$ EF = $(trust^{eprints} - trust^{future})$			
WP = $(trust^{wais} - trust^{past})$ WF = $(trust^{wais} - trust^{future})$			
MP = $(trust^{muhe} - trust^{past})$ MF = $(trust^{muhe} - trust^{future})$			

Like for the $N_p^{overlapping}$ users, means of the absolute difference between system and survey readings are calculated and Table 7 shows the results. Here, it shows opposite behaviour than happened in the case of $N_p^{overlapping}$ users, that is, the MP and MF are greater than EP and EF (respectively) for both the strongest and shortest path algorithms - see Table 7 for description of MP, WP, MF and WF. This shows that the consolidated MuHe networks took the trust values farther from the survey values, which resulted in deterioration of trust values for $N_p^{non-overlapping}$ users.

6. Discussion

If a trust-based system is developed for real-world social networks, the concept of implicit trust may be used to extract numeric values based on the activities of users in their individual social networks. This is due to unavailability of explicit trust metrics in all of the well known social networks in use nowadays. The trust metrics in these networks may be based on the frequency of likes/favourites or retweets in the kind of social networks for example facebook or twitter. In likes of professional social networks such as stack overflow or quora, the frequency of up votes/ shares may become a metric of trust.

The task to fuse data from real-world MuHe networks is also non-trivial. In federated networks, it is relatively straightforward as multiple accounts of a single user are not allowed in these networks. The professional social networks selected for experimentation also belong to this category. The trust metrics from such networks may be extracted from each of the networks using implicit activities, and can easily be aggregated using the techniques discussed in this article. However, for networks without any federated control, there may be fake or duplicate accounts, which can create data fusion problems. They may contribute wrong information which is not reflective of the person in the real-life. In such scenarios, the data should be rigorously preprocessed to eliminate any discrepancies before being fed to this framework for analysis. An absence of such preprocessing layer may generate distorted aggregation which can mislead other users in making wrong predictions if they are totally relying on digital world for trust-related decision making.

The future research directions which emerge out of this work are to implement the proposed model keeping in view the challenges that may arise due to federated and non-federated types of real-world social networks. Another research direction may be to explore the value of MuHe networks created from multiple constituent networks, and also to understand the impact of networks due to differing quality and size. Our hope is that this work will persuade developers of trust systems to go beyond individual social networks for trust calculations. This could improve existing trust systems by enabling them to make more intelligent trust decisions by incorporating information from a variety of sources on the web.

7. Conclusion

Since the inception of Web 2.0, the use of online social networks is increasing and existence of users in multiple networks is a great opportunity to make trust metrics on the web more sophisticated by incorporating a variety of information. This paper describes this as calculating trust on Multiple Heterogeneous (MuHe) networks, and makes three contributions:

Firstly, it presents a semantic web based framework for modeling and consolidating heterogeneous trust networks, and for performing trust calculations on both the individual and consolidated MuHe networks.

Secondly, it demonstrates the efficacy of the framework via a simulation that creates MuHe networks from individual networks with varying node and tie overlap. The simulation allowed us to investigate the impact of different data fusion techniques on the trust metrics of the MuHe network, and we showed that the Weighted Order Weighted Average (WOWA) technique produces a MuHe network with new trust paths, but which protects trust values (that could be over-inflated or suppressed with other consolidation techniques).

Thirdly, we explored how the MuHe network approach could work with real-life networks, and applied the framework to two professional networks (ePrints, a publication network; and WAIS, a project network). This showed that for overlapped users the MuHe network was closer to the assessments of the individuals within those networks (as gathered via a survey), but there was no substantial difference for non-overlapped users (individuals who only appeared in one of the two networks).

Our simulation experiment clearly shows that with appropriate data fusion techniques MuHe networks can be constructed with better trust properties than their constituent networks. But we have also shown that the value of this in the real world is strongly effected by how close the trust represented in the constituent networks is to the trust required by the application, and that in the real world the network sizes can be uneven, resulting in more complex interactions that those shown in our simulation. So while MuHe networks appear to be a promising technique for improving trust calculations they are not a panacea, and still depend heavily on the quality of the constituent networks, and how closely those networks reflect the type of trust required for a particular application.

Funding

This work was supported by the Higher Education Commission, Pakistan under the Startup Research Grant Number 21-1115.

Data Availability

This work is an extension of the doctoral thesis by Corresponding Author Imran, Muhammad (2015) 'The impact of consolidating web based social networks on trust metrics and expert recommendation systems'.

University of Southampton, Physical Sciences and Engineering, Doctoral Thesis, https://eprints.soton.ac.uk/id/eprint/38520. The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

References

[1] D. Boyd, N. Ellison, Social network sites: Definition, history, and scholarship, Journal of Computer-Mediated Communication 13 (1) (2007) 210–230.

[2] L. Garton, C. Haythornthwaite, B. Wellman, Studying online social networks, Journal of Computer-Mediated Communication 3 (1).

[3] M. Imran, The impact of consolidating web based social networks on trust metrics and expert recommendation systems, Ph.D. thesis, University of Southampton (2015).

[4] J. Golbeck, Trust on the world wide web: A survey, Foundations and Trends in Web Science. 1 (2) (2006) 131–197.

[5] J. Golbeck, J. Hendler, Filmtrust: Movie recommendations using trust in web-based social networks, in: Proceedings of 3rd IEEE Consumer Communications and Networking Conference (CCNC), Vol. 1, Las Vegas, NV, USA, 2006, pp. 282–286.

[6] D. Olmedilla, O. F. Rana, B. Matthews, W. Nejdl, Security and trust issues in semantic grids, in: Dagstuhl Seminar Proceedings, Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2006, pp. 10–18.

[7] Q. Gong, Y. Chen, J. Hu, Q. Cao, P. Hui, X. Wang, Understanding cross-site linking in online social networks, ACM Trans. Web 12 (4) (2018) 25:1–25:29. doi:10.1145/3213898.
URL http://doi.acm.org/10.1145/3213898

[8] K. Shu, S. Wang, J. Tang, R. Zafarani, H. Liu, User identity linkage across online social networks: A review, Acm Sigkdd Explorations Newsletter 18 (2) (2017) 5–17.

[9] S. Bistarelli, F. Santini, On merging two trust-networks in one with bipolar preferences, Mathematical Structures in Computer Science 27 (2) (2017) 215–233.

[10] L. Shi, D. Berrueta, S. Fernandez, L. Polo, S. Fernandez, Smushing rdf instances: Are alice and bob the same open source developer, in: Proceedings of 3rd Personal Identification and Collaborations: Knowledge Mediation and Extraction (PICKME) Workshop with 7th International Conference on Web Semantics, Berlin/Heidelberg: Springer, 2008, pp. 10–18.

[11] H. Glaser, I. Millard, A. Jaffri, T. Lewy, I. Millard, B. Dowling, On coreference and the semantic web., in: Proceedings of 7th International Semantic Web Conference (ISWC), Karlsruhe, Germany, 2008, pp. 26–30.

[12] M. Hussain, M. Ahmed, H. A. Khattak, M. Imran, A. Khan, S. Din, A. Ahmad, G. Jeon, A. G. Reddy, Towards ontology-based multilingual url filtering: a big data problem, The Journal of Supercomputing 74 (10) (2018) 5003–5021.

[13] H. Glaser, A. Jafri, I. Millard, Managing co-reference on the semantic web., in: Linked Data on the Web (LDOW) Workshop with 18th International Conference on World Wide Web (WWW), Madrid, Spain, 2009, pp. 288–93.

[14] J. Carroll, C. Bizer, P. Hayes, P. Stickler, Named graphs, Web Semantics: Science, Services and Agents on the World Wide Web 3 (4) (2005) 247–267.

[15] R. R. Yager, On ordered weighted averaging aggregation operators in multicriteria decisionmaking, IEEE Transactions on Systems, Man and Cybernetics 18 (1) (1988) 183–190.

[16] T. Vicenc, The weighted OWA operator, International Journal of Intelligent Systems 12 (2) (1997) 153–166.

[17] R. R. Yager, D. Filev, Operations for granular computing: mixing words and numbers, in: Fuzzy Systems Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on, Vol. 1, IEEE, 1998, pp. 123–128.

[18] R. Yager, D. Filev, Induced ordered weighted averaging operators, IEEE Transactions on Systems, Man, and Cybernetics 29 (2) (1999) 141–150.

[19] P. Victor, C. Cornelis, M. D. Cock, E. Herrera-Viedma, Practical aggregation operators for gradual trust and distrust, Fuzzy Sets and Systems 184 (1) (2011) 126 – 147.

[20] Y. Ma, H. Lu, Z. Gan, X. Ma, Trust discounting and trust fusion in online social networks, in: Web Technologies and Applications, Vol. 8709 of Lecture Notes in Computer Science, Springer International Publishing, 2014, pp. 619–626.

[21] S. Deng, L. Huang, G. Xu, X. Wu, Z. Wu, On deep learning for trust-aware recommendations in social networks, IEEE transactions on neural networks and learning systems 28 (5) (2017) 1164–1177.

[22] F. Heider, The psychology of interpersonal relations, Psychology Press, 2013.

[23] P. Holland, S. Leinhardt, Some evidence on the transitivity of positive interpersonal sentiment, American Journal of Sociology 77 (6) (1972) 1205–1209.

[24] C.-N. Ziegler, G. Lausen, Propagation models for trust and distrust in social networks, Information Systems Frontiers 7 (4) (2005) 337–358.

[25] C.-N. Ziegler, G. Lausen, Spreading activation models for trust propagation, in: IEEE International Conference on e-Technology, e-Commerce and e-Service (EEE), Taipei, Taiwan, 2004, pp. 83–97.

[26] Y. A. Kim, H. S. Song, Strategies for predicting local trust based on trust propagation in social networks, Knowledge-Based Systems 24 (8) (2011) 1360 – 1371.

[27] N. Verbiest, C. Cornelis, P. Victor, E. Herrera-Viedma, Trust and distrust aggregation enhanced with path length incorporation, Fuzzy Sets and Systems 202 (0) (2012) 61 – 74, theme: Aggregation Functions.

[28] C. Jiang, S. Liu, Z. Lin, G. Zhao, R. Duan, K. Liang, Domain-aware trust network extraction for trust propagation in large-scale heterogeneous trust networks, Knowledge-Based Systems 111 (2016) 237 – 247.

[29] W. Jiang, J. Wu, F. Li, G. Wang, H. Zheng, Trust evaluation in online social networks using generalized network flow, IEEE Transactions on Computers 65 (3) (2016) 952–963.

[30] S. Hamdi, A. L. Gancarski, A. Bouzeghoub, S. B. Yahia, Tison: Trust inference in trust-oriented social networks, ACM Transactions on Information Systems 34 (3) (2016) 17:1–17:32.

[31] W. Jiang, G. Wang, M. Z. A. Bhuiyan, J. Wu, Understanding graph-based trust evaluation in online social networks: Methodologies and challenges, ACM Computing Surveys. 49 (1) (2016) 10:1–10:35.

[32] J. Carroll, C. Bizer, P. Hayes, P. Stickler, Named graphs, provenance and trust, in: Proceedings of 14th International Conference on World Wide Web (WWW), ACM, Chiba, Japan, 2005, pp. 613–622.

[33] T. Heath, E. Motta, The Hoonoh ontology for describing trust relationships in information seeking, in: Proceedings of 3rd Personal Identification and Collaborations: Knowledge Mediation and Extraction (PICKME) workshop with 7th International Conference on Web Semantics, Berlin/Heidelberg: Springer, Karlsruhe, Germany, 2008, pp. 67–75.

[34] M. Imran, D. Millard, T. Tiropanis, Impact of consolidating social networks on derived trust factors, ASE Human Journal 1 (2) (2012) 88–99.