Impact of Consolidating Web Based Social Networks on Derived Trust Factors

Muhammad Imran
School of Electronics and Computer Science
University of Southampton
Southampton, UK
Email: mi1g08@ecs.soton.ac.uk

David Millard
School of Electronics and Computer Science
University of Southampton
Southampton, UK
Email: dem@ecs.soton.ac.uk

Thanassis Tiropanis
School of Electronics and Computer Science
University of Southampton
Southampton, UK
Email: tt2@ecs.soton.ac.uk

ABSTRACT

Individuals are typically members of a variety of web-based social networks (both explicit and implied), but existing trust inference mechanisms typically draw only on a single network in order to calculate trust between any two individuals. This both reduces the likelihood that a trust value can be calculated (as both people have to be members of the same network), and reduces the quality of any trust inference that can be drawn (as it will be based on only a single network, typically representing a single type of relationship). To make trust calculations on multiple distributed (MuDi) social networks those networks must be first consolidated into a single network. This is challenging as simple consolidation strategies such as summing or averaging trust values tend to distort trust values. In this paper we present an analysis of different consolidation approaches for MuDi trust networks, and propose a Weighted Ordered Weighted Averaging (WOWA) approach, that avoids these problems by including the value added (by a given tie in a given network) as a factor during consolidation. We evaluate the consolidation strategies using a numerical simulation, where we generate a range of networks that simulate the characteristics of real-world networks, and analyse what happens to trust factors such as average strength of trust ties (TS), and average length of trust path (TL), as they are consolidated with different node and tie overlap. We discover that while simple strategies such as summing or averaging distort the trust values in the consolidated network by amplifying or dampening trust metrics, the WOWA approach maintains the integrity of the trust metrics between individuals, and can dramatically increase the potential number of trust paths.

I INTRODUCTION

Trust is important in a variety of online activities, such as e-commerce, peer-to-peer networks, expert finding, recommendation, etc. But these only form a small part of our total online activity, which increasingly includes personal, social and professional interactions. Leveraging these social interactions to inform trust decisions is therefore increasingly important.

There are many definitions of trust, but we take the view that *Trust of party A to a party B for a service X is the measurable belief of A in that B behaves dependably within a specified context (in relation to service X)* [1].

The reason for selecting this particular definition lies in its capacity to characterise personalised subjective values of trust, of the sort that can be reasonably derived from social and professional interaction networks. Trust values can be based on two types of connections, explicit and implicit. Participants of an explicit network deliberately make connections with other people and it is that act of selecting them for friendship or interaction that then implies trust. Facebook, Twitter and LinkedIn are examples of explicit connections. Implicit networks are different in that they emerge as a result of mutual activities by users who are part of the same environment; for example, interacting on a forum, or co-publishing an article.

For either type of network we can quantify trust between pairs of nodes (for example, weighted by the number of interactions), and then use the resulting weighted network to calculate trust between any two nodes in the network (by propagating trust along the path between them, decaying due to the presence of intermediate connections [2]).

As people's use of the web becomes more sophisticated it also becomes increasingly important to consider Multiple Distributed (MuDi) networks rather than single networks when making trust calculations. This is complicated as these are heterogenous networks where the structure and weighting criteria are different, but if the information and activities of users in the various social networks could be combined, this would provide a much richer dataset for making decisions about trust.

Unfortunately consolidating multiple trust networks into a single network where trust calculations can be performed is non-trivial. Care must be taken not to inflate or dampen trust values artificially. Not all of the users are connected in all of the constituent networks, and in some cases will not even be present in some networks. Differentiating absence of trust from distrust is therefore a key problem that any trust aggregation mechanism should be able to cope with. Throughout the paper, we use the following vocabulary: N1 and N2 represent original networks that are being consolidated, MuDi is the final consolidated network and CN1 and CN2 are sub-networks in the MuDi that represent the original networks.

In this paper we propose to consolidate trust networks using a Weighted Ordered Weighted Averaging (WOWA) approach, normally seen in the domain of data fusion [3].

We identify two key trust metrics: average strength of ties (TS) and average length of trust path (TL), and undertake a numerical investigation to study how these vary as a result of consolidating MuDi social networks using the techniques of Summation (S), Weighted Average (WA), and WOWA. Our hypothesis is that consolidating MuDi networks based on WOWA will result in consolidated MuDi networks that better preserve the trust values of their constituent networks CN1 and CN2, while adding value for trust calculation by opening up many additional trust paths.

The paper is structured as follows. Section II describes related work in trust and social networks, and also details the WOWA technique. Section III describes the trust metrics that we have chosen to examine in our experiment, and describes how we generate networks that simulate real-world properties. The results of our experiment are then presented and analysed in Sections IV and V respectively. Finally, Section VI summarises our findings, and discusses how the impact of consolidation on trust metrics may relate to their quality.

II BACKGROUND

1 TRUST

There are various studies in the literature that investigate a range of trust metrics for finding trust between participants in social networks and they can be broadly categorised as global and local. Global trust metrics compute reputation values considering whole network information, Pagerank being a notable example [4], and result in each node in the network receiving a single objective trust value; local trust metrics are based on calculations from a given individuals position in the network, so each node takes its own subjective view of the trust of every other node [5].

In our work we are concerned with local trust (as online trust tends to be personalised and subjective). Local trust calculations take the start and destination participants as input parameters and calculates the trust of the destination using trusted path(s). Empirical evidence from psychologists and social psychologists shows that transitivity of trust exists along the paths in social networks [6, 7]. People tend to trust friends of their friends rather than unknown users [8], but the strength of trust weakens as the length of path to friend increases [9].

Ziegler and Lausen [2, 10] presented two trust and distrust propagation models for social networks using the Spreading Activation Model technique [11]. They presented an Appleseed trust algorithm that studies the trust propagation to the distant participants using the local group trust metrics. This is based on the theory of trust transitivity and uses the concept of trust decay mentioned in [9].

Golbeck [12] presented a model, "$TidalTrust$", for trust inference between indirectly connected individuals in social networks [13]. Trustworthiness of the distant node is calculated by taking the average of the trust from different neighbours at each step between the intermediate nodes. "$FilmTrust$" also uses this model for movie recommendation and generates personalised recommendations of the movies using the explicit trust rating of the participants. The fuzzy trust algorithm defined by Mohsen and Saeed [14] is similar to "$TidalTrust$" but it improves its accuracy by using linguistic terms for definition of trust (Low, Medium, High) rather than scaling (0, 0.5, 1), claiming that it is more meaningful for the users.

Walter et al.$'$ s [15] trust propagation mechanism multiplies trust values along the trust path. It is different from the Ziegler and Lausen technique [10], because the decay of the trust is not controlled by the source; rather decay takes place as a result of multiplication of trust values (in the 0-1 range) along the path. This mechanism is adapted and described in Section 3 for our work.

2 SOCIAL AND MUDI NETWORKS

There are range of explicit and implicit social networks on the web, but existing trust mechanisms only rely on a single network for evaluation of trust between participants. Aggregation of these different types of networks can allow us to base trust on different type of relationships that exists between individuals. The idea of aggregating MuDi social networks has its application in lot of other areas, including representation of separate social networks in a single social graph and processing of distributed data for reuse in search applications. They are however limited to either combining networks or to use their data in search systems, and the impact of different consolidation parameters on derived trust factors to date remains unexplored.

For example, Tang et al. [16] proposed a search system for academic researchers using semantic technologies. Their system uses Google API for extracting and integrating information about researchers from distributed locations on the web. Ido et al. [17] proposed "*SONAR*", an API that can integrate information about users from multiple social networks and they claim that it can give complete and useful picture to the end users. To use this system, however, API needs to be installed on all systems from where data need to be extracted. Similarly, another system, "*Polyphonet*" extracts and analyses the network information from multiple social networks [18]. Integrated information is analyzed to determine, for example, degree distributions, path length and other factors.

Bae and Kim's work [19] integrates separate social networks into a single global social graph for analysis using the concept of a hypergraph. Resources in multiple networks are connected using hyperedges and connections between them are normalised to depict a single social network. However, their work examined the resulting hyper-structures, rather than attempting to consolidate weights or tie meta-data.

3 DATA FUSION TECHNIQUES

Consolidating MuDi networks is essentially a problem of data fusion and in particular of data aggregation. A simple form of aggregation is to use Summation (S), this will work for simple numerical information (such as counts) but will inflate other statistical values. So the most common method of aggregation is to consider the sources of information by using Weighted Average (WA). This considers the importance and number of sources to give an aggregated measure of the data but can severely deteriorate the integrity of information as it treats missing data the same way as data with the value zero.

Yager [20] proposed an Ordered Weighted Averaging (OWA), an aggregation operator that considers the relative importance of the information used for aggregation. Unlike simple operators, it prioritises values in descending order, and allow us to assign weights bearing in mind the position of the information in the ordering. Although this mechanism is claimed to be better than WA as it respects the integrity of each data point out of multiple ones, it ignores the importance of the source of that information and hence still provides an incomplete picture.

The disadvantages of both of the above techniques are eradicated in the WOWA technique proposed by Tore [3], that considers the relative importance of the information and its source.

Importance of the information and source is represented using two weight vectors, w and p respectively, each of dimension n. $w = [w_1, w_2, \ldots, w_n]$ and $p = [p_1, p_2, \ldots, p_n]$ such that i) $w_i \in [0, 1]$ and $\Sigma_i w_i = 1$ ii) $p_i \in [0, 1]$ and $\Sigma_i p_i = 1$;. A mapping function f_{WOWA}: $\mathbb{R}^n \to \mathbb{R}$ is a WOWA operator of dimension n if:

$$f_{WOWA}(a_1, a_2, \ldots, a_n) = \sum_i \omega_i a_{\sigma(i)} \qquad (1)$$

where $\{\sigma(1), \sigma(2), \ldots, \sigma(n)\}$ is a permutation of $\{1, 2, \ldots, n\}$ such that $a_{\sigma(i-1)} \geq a_{\sigma i}$, $i = 2, 3, \ldots, n$, weight ω_i is defined as:

$$\omega_i = w^* \left(\sum_{j \leq i} p_{\sigma(j)} \right) - w^* \left(\sum_{j < i} p_{\sigma(j)} \right) \qquad (2)$$

where w^* is a monotonic function (e.g., a polynomial) that interpolates the points $(i/n, \Sigma_{j \leq i} w_j)$ along with point $(0, 0)$. Term ω represents set of weights $\{\omega_i\}$, that is $\omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$.

The WOWA technique has been adapted and used in our own work, as explained in Section 4.

III NUMERICAL SIMULATION

Our experiment is based around a numerical simulation of the consolidation of pairs of networks, using S, WA, and WOWA. The simulation is implemented in Python (including code for network generation, consolidation and trust inference algorithms), and the NetworkX API is used for measuring network properties.

1 CHARACTERISTICS OF REAL-WORLD NETWORKS

It is important that the networks we generate in our simulation have the characteristics of real-world networks. It has been well established in the literature that social networks have the properties of small world networks [21, 22], and there are two properties that characterise such behaviour; clustering coefficient (represented using C) and shortest path length (mentioned as L). C represents the level of clustering in the network and its value ranges from 0 to 1, scaling from low to high. It measures the extent to which nodes in the network are connected to each other and ensures transitivity as, in most cases in social networks, two friends of a single person are also friends of each other. The value of C for the network can be calculated using Equation 3:

$$C = \frac{1}{n}\sum_{i=1}^{n} c_i \tag{3}$$

where n is number of users in the network and c_i repersents local clustering of each user and its value for an undirected network can be calculated as [22]:

$$c_i = \frac{e_i}{\frac{1}{2}k_i(k_i - 1)} \tag{4}$$

where e_i represents actual number of ties and $k_i(k_i - 1)/2$ is the maximum possible number of ties between neighbours of user i. The other small world property, L, is the length of the shortest path between pairs of participants in the social network and its average value for the undirected network can be calculated using Equation 5 [21].

$$L = \frac{1}{\frac{1}{2}n(n{-}1)} \sum_{s,t \epsilon N} d(s,t) \tag{5}$$

where N represents set of n users and $d(s,t)$ is the length of shortest path from s to t.

Apart from small world properties, another parameter that impacts the structure of the network is the density (indicated by D). While generating a network, certain value of C and L in the network can only be ensured if the network has a certain value of D, because otherwise it will end up having a deficiency of connections. The value of D in any social network is the ratio of number of connections in the network to the total number of possible connections assuming there are no self loops in the network. For a network of n users, network density for a undirected network can be calculated using Equation 6:

$$D = \frac{\#of ties}{\frac{1}{2}n(n{-}1)} \tag{6}$$

For our simulation we have focused on the concept of consolidating professional social networks, Table 1 shows value of these parameters for several co-authorship networks collected from the literature and properties of generated random networks should be in accordance to them.

Table 1: Small World Properties of co-authorship networks taken from [23–25].

Co-authorship Networks	No of Nodes	C	L	D
Physics	52909	0.56	6.19	0.03504
Biology	1520251	0.6	4.92	0.00204
Math	253339	0.34	7.57	0.00309

2 GENERATING SAMPLE MUDI TRUST NETWORKS

To measure the impact of network and consolidation parameters on derived trust factors, a set of sample networks needs to be generated that conform to the small world properties of example real-world networks described in Table 1. Using this as a guide we conducted a number of pre-experiments to find the value of D that ensures values of these properties in a comparable range and Table 2 shows such values.

In our experiment we wanted to vary the number of overlapping nodes and overlapping ties in each pair of networks, so we could see the effect on trust values as the consolidated networks varied in similarity. We used a three step mechanism for wiring connections between participants to achieve a given node and tie overlap:

1. Firstly, users in each of the constituent networks are connected in a ring lattice to ensure the connectedness of the networks. If there is no tie overlap (TO) between the networks, then networks are still connected, but overlapped connections are also replaced with non-overlapped connections.

2. Then, the remaining percentage of overlapped ties are created between randomly selected pair of participants with probability $1/N_p$ among all present in the network, where N_p represents a pair of participants.

3. In the final step, non-overlapped ties between the random pair of participants are created with the same probability as that in Step 2. Duplicate ties are forbidden in Steps 2 and 3 to ensure target density of the networks.

In our simulation we assume that the networks contain trust information on their ties, represented as a continuous value in the range $(0, 1)$. The ties of the generated sample networks are randomly weighted with these values and represent subjective symmetric trust between the users, with values near to 0 represent low trust and those close to 1 show high trust between the individuals connected.

3 TRUST PROPERTIES

When consolidating two networks we are interested in measuring the trust properties of those original networks N1 and N2, the consolidated MuDi network, and the CN1, CN2 (meaning the two original networks, as represented by sub-networks in the consolidated MuDi network). Ideally we should see that the trust properties are transformed but not damaged in their journey from the N1, N2 to the consolidated MuDi network.

Our trust evaluation uses an adopted version of Dijkstra's pathfinding algorithm and the concept of stronger path described in [14] that can result in a longer but overall stronger trust path. For this, it does not stop searching once the path is found but continues unless each of the participants is visited. Suppose the trust path between participants a and b is $P(a, b)$, then the following equation adapted from [15] can be used for evaluation of trust:

$$T^{a,\dots,b} = \prod_{(i,i+1)\in P(a,b)} T^{i,i+1} \tag{7}$$

To quantify the performance of this trust inference mechanism on CN1, CN2 and consolidated MuDi networks, there are two variables identified from the literature; namely strength of trust ties and length of trust paths. $T^{a,\dots,b}$ is the strength of the trust tie between participants a and b and length of trust path is the number of ties involved in the path $P(a, b)$. The approximated estimation of these two trust properties is evaluated for each of the networks N1, N2, CN1, CN2 and MuDi by taking average of the trust estimations for each pair of participants in the network and examined for the change that happens as a result of consolidation, for example, if *TS*, *TL* represents such values for strength of ties and length of trust paths respectively, then it can be calculated as described in Equation 8 and 9:

$$TS = \frac{1}{n(n-1)} \sum_{s,t\epsilon N_p} T^{s\to t} \tag{8}$$

$$TL = \frac{1}{n(n-1)} \sum_{s,t\epsilon N_p} P(s,t) \tag{9}$$

We are looking for a consolidation that does not damage existing trust properties, but uses the additional information to enhance them. In terms of *TS* and *TL* this means that we would like *TS* for CN1, CN2 and MuDi to remain close to that of N1, N2 even if there is not a significant PO and TO. If *TS* is maintained in this way it shows that damage was minimised during consolidation. Furthermore, we would expect *TL* to decrease overall significantly due to the emergence of additional strong trust paths as compared to those in N1 and N2. If *TL* decreases it shows that the consolidation has successfully enhanced trust calculations, by opening up new trust paths.

4 METHODOLOGY

When conducted on real world data a consolidation of networks would generally use a heuristic approach of meta-data comparison to identify participants that appear in both networks (for example, by comparing *familyName* and *givenName* properties). The result will be a certain number of overlapping participants (represented as PO) with potentially overlapping ties (refered as TO) between them.

In our simulation we want to show the effect on the two trust properties (strength of trust tie, and length of trust path) as PO and TO vary. However, TO is constrained by PO, as to achieve certain percentage of TO, at least $PO \geq TO$[1] should be in place, because otherwise number of overlapped connections would exceed maximum possible number of ties between subset of overlapped participants. For example, a 100% TO is only possible if there is 100% PO as well.

We ran a number of simulations, setting PO at 40%, 60%, 80%, and 100%, and for each setting of PO (except 40%PO[1]) allowed TO to vary from 0 to PO in increments of 20% . Then in each simulation we consolidate the trust information on the ties using S, WA and WOWA (as described in Section 3). Table 2 shows the values and ranges for all the variables in our simulation.

Table 2: Network and consolidation parameters used for this study

Network Parameters	Description
N	30
D	0.43
C	0.45 ± 0.02
L	1.57 ± 0.01
Ratio of D, C, L between $N1$ and $N2$	1

Consolidation Parameters	Description
PO	$[40\%, 100\%]$
TO	$[0, PO]^1$
w	$[1, 0.5]$
p	$[0.8, 0.2]$

As an example of using WOWA, suppose that trust information available from two MuDi networks is in decreasing order as $a = [0.8, 0.5]$, $w = [1, 0.5]$ and $p = [0.8, 0.2]$. Normalised vector of the weights can be calculated as $w_n = [1, 0.5]/1.5 = [0.67, 0.33]$. Next we have to find the function w^* interpolating the points $(i/n, \Sigma_{j \leq i} w_j)$ and this can be done as described in Equation 10 :

$$\begin{aligned} (0.5, w_1) &= (0.5, 0.67) \\ (1, w_1 + w_2) &= (1, 1) \end{aligned} \tag{10}$$

Plotting the points $\{(i/n, \Sigma_{j \leq i} w_j) | i = 1, 2, \ldots, n\} \cup \{0, 0\}$ gives us the following polynomial interpolation function:

$$w^*(x) = -0.6667x^2 + 1.6667x \tag{11}$$

Equation 12 can be used to calculate final weights ω_i:

$$\begin{aligned} \omega_1 &= w^*(p_1) = w^*(0.8) = 0.91 \\ \omega_2 &= w^*(\sum_{i=1}^{2} p_i) - w^*(p_1) = w^*(1) - w^*(0.8) = 0.09 \end{aligned} \tag{12}$$

This allows us to calculates final values of f_{WOWA}:

$$f_{WOWA}(0.8, 0.5) = \sum_{i=1}^{2} \omega_i a_i = 0.77 \tag{13}$$

Table 3 compares the values generated by WOWA to those of the naive methods of S and WA for a pair of values collected from MuDi networks.

Table 3: Trust aggregation using three different techniques for three different set of values, 0 in $[0.8, 0]$ represents absence of trust.

MuDi Trust Data	S	WA	WOWA
$[0.8, 0.8]$	$\widehat{1.6} = 0.99$	0.8	0.8
$[0.8, 0.5]$	$\widehat{1.3} = 0.99$	0.65	0.77
$[0.8, 0]$	0.8	0.4	0.72

It shows that WOWA both respects high trust values, and handles the absence of trust values. The WA mechanism considers absence of trust as distrust, and hence dampens down the value to half, while WOWA still decays the trust but preserves the integrity of the

[1] For 40%PO, $PO > TO$ should be true, because $40\%PO = 40\% * 30 = 12$, the maximum possible number of undirected ties between overlapped participants can be $(12 * 11)/2 = 66$ and $40\%TO = (40\% * (0.43 * (30 * 29))/2) = 74$. So the required number of overlapped ties 74 exceeds maximum possbile number of ties ties 66.

Table 4: Average Strength of Ties (TS) for three different trust aggregtion mechanisms i.e. Summation (S), Weighted Average (WA) and Weighted Ordered Weighted Averaging (WOWA) in the networks with varying percentage of PO and TO. CN1 and CN2 represent sub-networks in the consolidated MuDi networks

PO	TO	Avg. Strength of Ties in the Network (TS)										
		N1	N2	CN1			CN2			MuDi		
				S	WA	WOWA	S	WA	WOWA	S	WA	WOWA
40	0	0.64	0.65	0.64	0.19	0.5	0.65	0.2	0.52	0.68	0.18	0.53
	20	0.59	0.62	0.69	0.25	0.52	0.75	0.27	0.56	0.71	0.20	0.52
	30	0.63	0.60	0.73	0.27	0.56	0.72	0.26	0.54	0.71	0.21	0.52
60	0	0.66	0.62	0.66	0.21	0.53	0.62	0.2	0.5	0.68	0.20	0.54
	20	0.62	0.68	0.77	0.29	0.56	0.79	0.28	0.58	0.75	0.24	0.56
	40	0.64	0.62	0.80	0.33	0.6	0.77	0.32	0.58	0.76	0.26	0.56
	60	0.54	0.64	0.72	0.31	0.53	0.82	0.35	0.62	0.72	0.26	0.53
80	0	0.64	0.62	0.65	0.20	0.52	0.63	0.2	0.50	0.65	0.23	0.54
	20	0.69	0.64	0.81	0.30	0.60	0.79	0.29	0.58	0.74	0.28	0.58
	40	0.65	0.68	0.84	0.37	0.62	0.84	0.37	0.62	0.80	0.32	0.6
	60	0.65	0.62	0.84	0.43	0.63	0.84	0.43	0.63	0.79	0.36	0.6
	80	0.59	0.60	0.79	0.40	0.59	0.87	0.43	0.64	0.77	0.34	0.56
100	0	0.67	0.64	0.67	0.21	0.54	0.64	0.2	0.52	0.60	0.27	0.53
	20	0.62	0.64	0.77	0.26	0.54	0.79	0.26	0.56	0.66	0.29	0.54
	40	0.65	0.68	0.85	0.37	0.63	0.86	0.37	0.63	0.75	0.35	0.6
	60	0.63	0.63	0.88	0.42	0.64	0.87	0.41	0.62	0.80	0.39	0.61
	80	0.64	0.65	0.92	0.5	0.69	0.92	0.5	0.69	0.88	0.48	0.67
	100	0.63	0.60	0.94	0.53	0.69	0.94	0.53	0.69	0.94	0.53	0.69

single high trust value. Data points in S greater than 1 are caped to 0.99, which is the maximum possible trust in our mechanism.

IV RESULTS

There are two types of trust measurements in the networks, the first type of measurements are on the direct ties and are aggregated as a result of merging individual networks. Second type of measurements are evaluated for each pair of indirectly connected participants N_P, in (N1, N2), (CN1, CN2) and consolidated MuDi networks separately, and these values depend on the trust between intermediate nodes in the trust path in that specific network. As the density, D, of network in each of the constituent networks is 0.43, this means each original network N1 and N2 has 43% direct connections and 57% of the trust evaluations are based on finding trust paths.

Table 4 and 5 presents the values of TS and TL for varying consolidation parameters PO and TO, and two sample results (for $PO = [60\%, 100\%]$) are also depicted in Figure 1. For different percentage of PO and TO, TS and TL metrics are evaluated for N1, N2 and then for each of the aggregation strategy (S, WA, WOWA) it is evaluated for CN1, CN2 and the MuDi network.

The aim of our experiment was to aggregate trust information from MuDi social networks without affecting the integrity of that information. We can define this as preserving the trust values from the original networks (N1, N2) in the sub-networks (CN1, CN2) of the consolidated networks.

Firstly, if we look at TS for CN1 and CN2 in Table 4 and Figures 1(a),1(b), it can be seen that the value of this metric for WOWA resides in between other two extreme approaches S and WA. S just amplifies the trust by summing the values available on the ties, hence inflating the trust up to 0.94 at $[100\%PO, 100\%TO]$, while WA dampens the trust down to 0.19 at $[40\%PO, 0\%TO]$ which were 0.63 and 0.64 in N1 respectively. WOWA, in both of the above

Table 5: Average Length of Trust Paths (TL) for three different trust aggregtion mechanisms i.e. Summation (S), Weighted Average (WA) and Weighted Ordered Weighted Averaging (WOWA) in the networks with varying percentage of PO and TO. CN1 and CN2 represent sub-networks in the consolidated MuDi networks

PO	TO	Avg. Length of Trust Paths (TL)										
		N1	N2	CN1			CN2			MuDi		
				S	WA	WOWA	S	WA	WOWA	S	WA	WOWA
40	0	2.68	2.41	2.68	1.64	2.09	2.41	1.63	1.99	3.04	1.82	2.4
	20	2.26	2.43	2.35	1.87	2.11	2.36	1.84	1.99	2.95	2.16	2.45
	30	2.31	2.39	2.33	1.8	1.94	2.62	1.85	2.04	3.17	2.16	2.45
60	0	2.37	2.36	2.37	1.6	1.89	2.36	1.62	2.01	2.62	1.63	1.99
	20	2.13	2.47	2.40	1.85	1.88	2.51	1.87	2.01	2.82	1.88	2.11
	40	2.34	2.4	2.18	1.94	2.09	2.27	1.93	2.12	2.72	2.10	2.37
	60	2.61	2.14	2.13	1.95	2.19	2.02	2	2.07	2.53	2.29	2.52
80	0	2.35	2.49	2.38	1.61	1.96	2.56	1.67	1.92	2.04	1.43	1.68
	20	2.37	2.29	2.50	1.82	1.97	2.68	1.83	2.03	2.29	1.59	1.77
	40	2.24	2.33	2.05	1.85	1.95	2.0	1.89	1.9	2.13	1.79	1.89
	60	2.6	2.29	2.05	2.25	2.17	1.94	2.20	2.09	2.20	2.22	2.17
	80	2.53	2.19	1.89	1.97	2.11	1.97	1.96	2.15	2.27	2.21	2.43
100	0	2.39	2.52	2.39	1.64	1.95	2.52	1.63	1.92	1.37	1.14	1.2
	20	2.23	2.80	3.11	1.74	1.93	2.88	1.79	2.03	1.70	1.27	1.32
	40	2.58	2.37	2.19	1.94	2.15	2.20	1.98	2.14	1.66	1.48	1.49
	60	2.62	2.31	1.90	1.89	2.01	1.91	1.91	2	1.62	1.60	1.62
	80	2.20	2.24	1.67	1.87	1.91	1.69	1.90	1.91	1.58	1.75	1.74
	100	2.15	2.33	1.64	1.88	1.98	1.64	1.88	1.98	1.64	1.88	1.98

mentioned cases calculates more stable metric with values of 0.69 and 0.5 respectively.

Behaviour of *TS* metric in MuDi is also in accordance with those of CN1 and CN2, and the results of WOWA again lies in between other two techniques. S escalates the trust up to 0.94 at $[100\%PO, 100\%TO]$, while WA reduces it to 0.18 $[40\%PO, 0\%TO]$, but WOWA maintains it at 0.69 and 0.53 respectively.

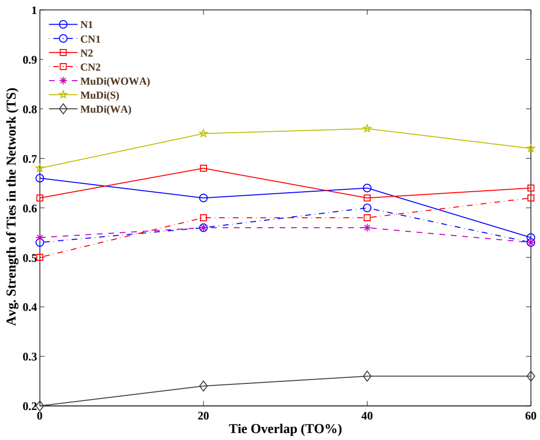
If the same sub-networks CN1 and CN2 are considered for average length of trust paths, *TL* from Table 5 and Figures 1(c),1(d), it is observed that, WOWA improves the *TL* metric in CN1 and CN2. For S, *TL* only decreases for $TO \geq 40\%$, whereas, WA also reduces it significantly for $TO \leq 40\%$. *TL* using WOWA also decreases with an increase in percentage of TO but this reduction is less as compared to WA.

TL in MuDi behaves similarly to that in CN1 and CN2 and its value for S and WA decreases with an increase in *PO*, but higher than subnetworks CN1, CN2 at low *PO*. Similar is the case with WOWA, where an increase in *PO* decreases *TL* except at some datapoints with low *PO*. The noise and non-uniformity in *TL* metric is due to its dependence on *TS*, as to achieve the maximum *TS*, the trust algorithm can even select longer trust paths.

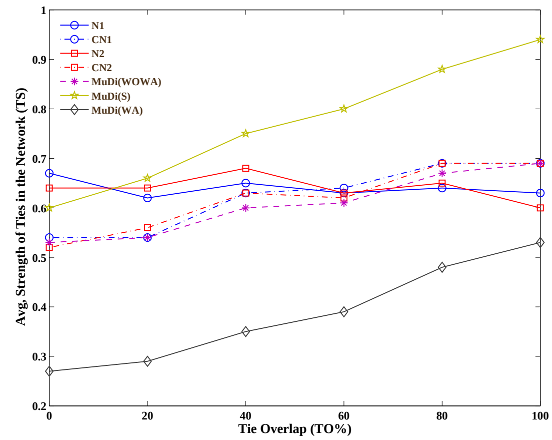
V DISCUSSION

Our original hypothesis was that consolidating MuDi networks based on WOWA would result in consolidated networks that better preserve the trust values of their constituent networks, while adding value for trust calculation by opening up many additional trust paths.

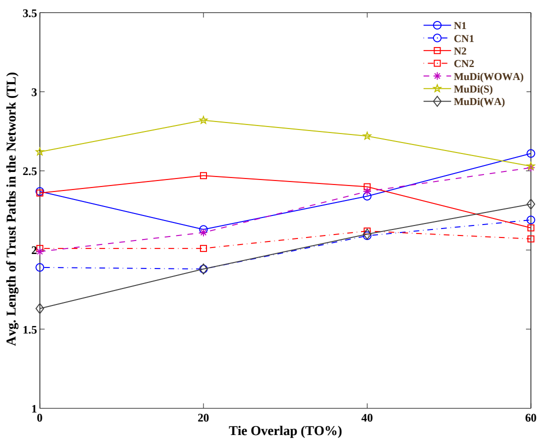
We can say that a consolidation approach better preserves the trust values if the trust values in the original networks are similar to those in the relevant sub-networks of the consolidated network. We would expect consolidation to create some differences, but that each trust relationship would be as likely to rise



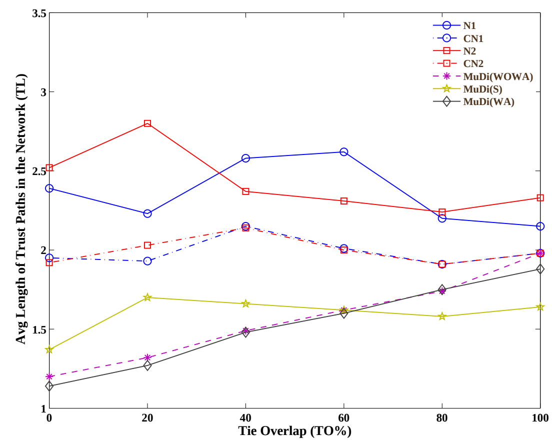
(a) Participant Overlap(PO) = 60%



(b) Participant Overlap(PO) = 100%



(c) Participant Overlap(PO) = 60%



(d) Participant Overlap(PO) = 100%

Figure 1: Sample values of TS and TL selected from Table 4,5 for depiction.

as to fall, and therefore when averaged across all ties should remain approximately stable. We can check this by comparing TS between the networks for each consolidation technique.

From analysing the TS metric presented in Section IV, it can be seen that WOWA technique better aggregates the trust from MuDi social networks, as it respects the integrity of trust in N1 and N2. S simply amplifies the trust while WA naively dampens down the trust, but as WOWA fuses trust data available on the ties bearing in mind their importance it gives a more balanced aggregation (distinguishing the absence of trust from distrust). Table 6 presents the statistical significance of the subnetworks CN1 and CN2 data using paired, two-tailed distribution, and it shows that WOWA is significantly better than WA at preserving the average strength of ties (TS), but this is less significant for average length of trust paths (TL) where our experiment shows that it is statistically closer to TL for original networks (N1,N2) only when considering situations where PO is 100% or when looking at the *Overall* performance.

We would expect the second part of our hypothesis, the opening up of many additional trust paths, to manifest through the TL metric that measures average length of trust paths. The data shows that TL in the MuDi network is dependent on the Participant Overlap PO. When PO is low it creates a path bottleneck in the consolidated network, and *TL* is higher than for the original networks, but when *PO* is high the increase in connections causes *TL* to fall. Additionally it can be seen that *TL* in each of the sub-networks CN1 and CN2 is lower than in the corresponding original network N1 and N2 respectively, regardless of the value of *PO*. This shows that additional trust paths are being created.

Table 6: Statistical Significance to show the improvement of WOWA over WA in CN1 and CN2 for varied percentage of *PO* and *Overall PO* data using TTEST (paired, two-tailed distribution).

(a) *p value* for *TS* metric

PO	CN1	CN2
40	0.0016	0.0016
60	0.0009	0.0001
80	0.0006	0.0002
100	0.0002	0.0002
Overall	9.6222E-14	2.9173E-14

(b) *p value* for *TL* metric

PO	CN1	CN2
40	0.0939	0.0684
60	0.0528	0.0638
80	0.1263	0.1874
100	0.0088	0.0170
Overall	1.6444E-05	3.6873E-05

Our numerical simulation shows that WOWA consolidation of MuDi social networks is a productive approach that preserves the integrity of the trust values (as measured by an increase in *TS*, average strength of ties) while creating new trust paths (as measured by decrease in *TL*, average length of trust paths). At low *PO* it creates an opportunity for users to know and interact with lot of new users which are not part of their original networks, and hence creates ties between people from networks of different background. On the other hand, at high *PO* and *TO* WOWA consolidation helps in refining trust values by combining different perspectives of trust that exists between the participants in different networks.

VI CONCLUSIONS AND FUTURE WORK

The increasing use of multiple heterogenous social networks, both explicit and implicit, offers an opportunity to refine trust calculations by consolidating multiple trust networks into a single network for analysis. However, consolidating trust networks is non-trivial due to variances in node and tie overlap, differences in the importance of networks, and differences in expressing trust.

In this paper we have presented a numerical simulation of what happens when different trust networks (with the characteristics of real-world networks) are consolidated using one of three strategies: S, WA and WOWA. In our experiment we varied participant and tie overlap, and recorded the effect on average strength of ties and average length of trust paths for the whole consolidated network (MuDi), and the sub-networks (CN1,CN2) that represented the original networks (N1,N2).

Our analysis reveals that the Summation (S) strategy results in an inflation of trust values, while the Weighted Average (WA) results in dampened trust values. However, the WOWA strategy has a much improved performance, in that it better preserves the integrity of the trust as compared to WA ($p<0.0001$), while also being better than WA at creating shorter trust paths ($p<0.0001$).

Our experiment shows that WOWA can be used to consolidate trust networks without damaging trust values. However, it is still not clear whether the changes to trust values caused by consolidation actually increase their *quality* in terms of their similarity to the trust actually felt by those individuals.

To test this our future plan is to attempt this consolidation activity with two real social networks (we are looking at professional and co-authorship networks) and then perform a qualitative evaluation with actual users via a survey to compare actual trust values with those in the original and consolidated networks.

We have shown that a WOWA consolidation strategy can effectively combine multiple trust networks, providing evidence that trust values derived form multiple distributed (MuDi) social networks can be merged to create new trust paths without damaging trust values. Our hope is that this approach can be used in the future to create more reliable trust calculations that take advantage of our increasingly varied and rich online interactions.

References

[1] D. Olmedilla, O. Rana, B. Matthew, and W. Nejdl, "Security and trust issues in semantic grids," in *Proceedings of the Dagsthul Seminar, Semantic Grid: The Convergence of Technologies*, vol. 05271, January 2005.

[2] C.-N. Ziegler and G. Lausen, "Propagation models for trust and distrust in social networks," *Information Systems Frontiers*, vol. 7, no. 4, pp. 337–358, 2005.

[3] T. Vicenç, "The weighted owa operator," *International Journal of Intelligent Systems*, vol. 12, no. 2, pp. 153–166, 1997.

[4] P. Lawrence, B. Sergey, M. Rajeev, and W. Terry, "The pagerank citation ranking: Bringing order to the web." Stanford InfoLab, Technical Report, November 1999.

[5] L. Mui, M. Mohtashemi, and A. Halberstadt, "A computational model of trust and reputation," in *Proceedings of the 35th Annual Hawaii International Conference on System Sciences(HICSS)*, January 2002, pp. 2431–2439.

[6] P. Holland and S. Leinhardt, "Some evidence on the transitivity of positive interpersonal sentiment," *American Journal of Sociology*, vol. 77, no. 6, pp. 1205–1209, 1972.

[7] A. Rapoport, "Mathematical models of social interaction," *Handbook of Mathematical Psychology*, vol. 2, 1963.

[8] F. Heider, *The Psychology of Interpersonal Relations*, New York, NY, USA, 1958.

[9] A. Josang, E. Gray, and M. Kinateder., "Analysing topologies of transitive trust," in *Proceedings of the Workshop of Formal Aspects of Security and Trust (FAST 2003)*, September 2003.

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

[11] R. Quillian, "Semantic memory," in *Semantic Information Processing*. MIT Press, 1968, pp. 227–270.

[12] J. Golbeck and J. Hendler, "Inferring binary trust relationships in web-based social networks," *ACM Transactions on Internet Technology.*, vol. 6, no. 4, pp. 497–529, November 2006.

[13] J. Golbeck, "Computing and applying trust in web-based social networks," Ph.D. dissertation, Department of Mathematics and Computer Science, University of Maryland, US, 2005.

[14] M. Lesani and S. Bagheri, "Applying and inferring fuzzy trust in semantic web social networks," in *Canadian Semantic Web*, ser. Semantic Web and Beyond. Springer US, 2006, vol. 2, pp. 23–43.

[15] F. Walter, S. Battiston, and F. Schweitzer, "A model of a trust-based recommendation system on a social network," *Autonomous Agents and Multi-Agent Systems*, vol. 16, no. 1, pp. 57–74, 2008.

[16] J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su, "Arnetminer: Extraction and mining of academic social networks," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD)*, New York, NY, USA, 2008, pp. 990–998.

[17] I. Guy, M. Jacovi, E. Shahar, N. Meshulam, V. Soroka, and S. Farrell, "Harvesting with sonar: the value of aggregating social network information," in *In Proceeding of the 26th Annual SIGCHI Conference on Human Factors in Computing Systems(CHI)*. New York, NY, USA: ACM, 2008, pp. 1017–1026.

[18] Y. Matsuo, M. Hamasaki, Y. Nakamura, T. Nishimura, K. Hasida, H. Takeda, J. Mori, D. Bollegala, and M. Ishizuka, "Spinning multiple social networks for semantic web," in *Proceedings of the 21st National Conference on Artificial Intelligence*, vol. 2. AAAI Press, 2006, pp. 1381–1386.

[19] J. Bae and S. Kim, "A global social graph as a hybrid hypergraph," in *Fifth International Joint Conference on INC, IMS and IDC*. Los Alamitos, CA, USA: IEEE Computer Society, 2009, pp. 1025–1031.

[20] Y. R.Ronald, "On ordered weighted averaging aggregation operators in multicriteria decisionmaking," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 18, no. 1, pp. 183 –190, jan/feb 1988.

[21] A. Barrat and M. Weigt, "On the properties of small-world network models," *The European Physical Journal B - Condensed Matter and Complex Systems*, vol. 13, pp. 547–560, 2000.

[22] D. J. Watts and S. H. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, pp. 440–442, 1998.

[23] M. E. J. Newman, "The structure of scientific collaboration networks," *Proceedings of the National Academy of Sciences*, vol. 98, no. 2, pp. 404–409, 2001.

[24] M. Newman, "Scientific collaboration networks. part i. network construction and fundamental results," *Phys. Rev. E*, vol. 64, no. 1, June 2001.

[25] L. A. N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, "Classes of small-world networks," *Proceedings of the National Academy of Sciences*, vol. 97, no. 21, pp. 11 149–11 152, 2000.