
Peer reviewed version

Link to published version (if available):
10.1109/ASONAM.2018.8508302

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via IEEE at https://doi.org/10.1109/ASONAM.2018.8508302 . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/
Abstract—There is no agreed definition of social capital in the literature. However, one interpretation is that it refers to those resources embedded in an individual’s social network offering benefits to that individual in relation to achieving goals and facilitating actions. This can be viewed as a resource-based interpretation of social capital aimed at the level of individuals. In this paper, we propose a family of social capital measures in line with this interpretation. Our measures are designed for a model of social networks based on weighted and attributed graphs, and cover four dimensions of social capital: (i) access to resources, (ii) access to superiors, (iii) homogeneity of ties, and (iv) heterogeneity of ties. We demonstrate the real-world application of our measures by exploring an illustrative use case in the form of a workplace social network.

Index Terms—social capital, measures, social network analysis

I. INTRODUCTION

The old adage that “it’s not what you know, but who you know that counts” follows the well-accepted view that social ties often carry value. This view has been confirmed through various studies (e.g. in finance [1], criminology [2], politics [3], and online social networks [4]). Understanding the value of social ties is the primary motivation behind the field of social capital [5]. In the social capital literature, studies have attempted to understand this notion in a variety of contexts, including crisis-response [6], religion [7], multiculturalism [8], physical activity [9], and the internet [10]. While there is no consensus on a definition of social capital, theories of social capital can be roughly divided into community-level theories [11], [12] and individual-level theories [5], [13]. The former are concerned with the social capital of an entire community, while the latter are concerned with the social capital of an individual in the context of a community.

In this paper, we focus on a model of individual-level social capital in line with the following intuition:

Social capital refers to those resources embedded in an individual’s social network offering benefits to that individual in relation to achieving goals and facilitating actions.

Specifically, we propose a model of individual-level social capital that is formalised as a set of social capital measures, exploiting well established theories in social science, including resource theory [13] and structural hole theory [14]. These measures not only consider the social ties between individuals, but also consider the resources that those individuals possess, and the context of those ties. By resources, we mean both tangible and intangible resources [13]; the former might include financial capital, or followers of a user on Twitter, while the latter might include expertise, experience, or trust. In the example of a workplace setting, individuals with high social capital might be those with strong connections to senior management, or to individuals with desirable expertise.

To the best of our knowledge, there is no existing work that has proposed a quantitative model of individual-level social capital exploiting both context and resources. Thus, our main contributions are as follows:

(i) We propose a model of social networks—based on weighted and attributed graphs—that is able to characterise available resources and the context of social ties.

(ii) We propose a set of quantitative individual-level social capital measures, covering different aspects of social capital theory from the literature (i.e. resource mobilisation, linking, bonding, and bridging).

(iii) We demonstrate the real-world application of our new measures through a workplace use case.

The remainder of the paper is organised as follows: in Section II, we review related work; in Section III, we propose a suitable graph model for social networks; in Section IV, we propose our individual-level social capital measures; and in Section V, we conclude.

II. RELATED WORK

Theories of social capital were initially developed by social scientists [5], [11]–[13]. These works dealt with both community-level [11], [12] and individual-level social capital [5], [13]; our work follows the latter approach. These social science theories have been applied in a variety of domains, including finance [1], criminology [2], politics [3], health [15], lobbying [16], and employment [17]. The domain of online social networks (e.g. Facebook and Twitter) has been of particular interest. For example, there has been work related to online social networks focused on measuring social...
capital [4], [18] and on enhancing social capital [19]. It has also been shown that an individual’s social capital can play an important role in online social movements [20], [21].

More recently, computer scientists have attempted to propose quantitative models of social capital that can be automatically computed from an underlying social network [22], [23]. These works have again tended to target specific domains (e.g. online social communities [22] and co-authorship networks [23]). Although our work follows this line of research, we instead aim to develop a quantitative model of social capital that is also domain-independent. Of course, while we consider workplace settings in this paper, we would emphasise that this represents just one use case for our domain-independent model.

III. SOCIAL NETWORK MODEL

In this section, we propose a model of social networks based on weighted and attributed graphs that will provide the input to our social capital measures defined in Section IV.

An (undirected) graph is a tuple \((N, E)\) where \(N\) is a set of nodes and \(E \subseteq \{\{n, n'\} \mid n, n' \in N\}\) is a set of (undirected) edges. An edge \(e \in E\) is called a self-loop if \(e\) is a singleton. A sequences of nodes \((n_1, \ldots, n_{m+1})\) is called a path from node \(n_1\) to node \(n_{m+1}\) if \([n_i, n_{i+1}] \in E\) for each \(i = 1, \ldots, m\) and \(n_i \neq n_j\) if \(i \neq j\) for each \(j = 1, \ldots, m\). A path \((n_1, \ldots, n_{m+1})\) is said to be of length \(m\). The set of edges in paths \(p = (n_1, \ldots, n_{m+1})\) is defined as \(E(p) = \{[n_i, n_{i+1}] \mid i = 1, \ldots, m\}\). Let \(P\) denote the set of paths in \((N, E)\) such that \([p]\) denotes the length of path \(p \in P\). Let \(P(n, n') \subseteq P\) denote the set of paths from node \(n\) to node \(n'\). Let \(N(n) = \{n' \in N \setminus \{n\} \mid P(n, n') \neq \emptyset\}\) be the set of nodes reachable from node \(n\).

A weighted graph is a tuple \((G, W)\) where \(G = (N, E)\) is a graph and \(W : E \to \mathbb{Z}^{\geq 1}\) is a weight function. An attributed graph is a tuple \((G, A_1, \ldots, A_m)\) where \(G = (N, E)\) is a graph and each \(A_i : N \to V_i\) is an attribute such that \(V_i\) is the set of possible values of \(A_i\). Let \(\mathcal{L}\) denote the language (i.e. a set of logical formulas) obtained from attributes \(A_1, \ldots, A_m\), the set of logical constants \(\{\top, \bot\}\), and the set of logical connectives \(\{\land, \lor, \neg\}\). Let \(A_1(n), \ldots, A_m(n) \models \phi\) denote that formula \(\phi \in \mathcal{L}\) is true under \(A_1(n), \ldots, A_m(n)\) in the usual way.

Definition 1 (Social network). A social network is a tuple \((G, W, A_1, \ldots, A_m, \succeq_L)\) where \(G = (N, E)\) is a graph such that \(E\) contains no self-loops, \((G, W)\) is a weighted graph, \((G, A_1, \ldots, A_m)\) is an attributed graph, and \(\succeq_L\) is a partial order over \(N\). We say that \(N\) is the set of individuals, \(E\) is the set of direct social ties, and \(\succeq_L\) is the hierarchy.

Example 1 (Social network). Consider the example workplace social network outlined in Figure 1. The set of employees is \(N = \{n_1, n_2, \ldots, n_8\}\). There are \(|E| = 10\) direct social ties. The direct social ties are weighted (e.g. \(W([n_1, n_2]) = 3\)) and there is a path between every pair of nodes. In addition, the hierarchy is \(n_3 \succeq_L n_1 \succeq_L n_2 \succeq_L n_4\) and \(n_5 \succeq_L n_6 \succeq_L n_7 \succeq_L n_8\). Finally, Table I outlines a set of five job-related attributes for each employee.

We can highlight several properties of our social network model. Firstly, each edge is associated with a weight, which describes the relative strength of direct social ties between individuals. Thus, rather than a simple binary or ternary (e.g. weak/strong) view of direct social ties, our model accommodates general weights. Secondly, each function \(A_i\) represents an attribute, mapping each individual to a value in \(V_i\). For example, a \(job\) \(role\) attribute could map individuals to a value such as developer, HR, or manager. These attributes thus allow us to reason about the resources of individuals and the context of direct social ties via standard logical formulas. Thirdly, the model supports a partial order \(\preceq_L\) over individuals. In a workplace context, for example, this partial order might represent the hierarchy of authority within the organisation.

When we refer to resources, we mean the attribute values associated with individuals such that desirable resources are expressed as a formula \(\phi \in \mathcal{L}\). When we refer to the context of social ties, we mean a broader notion comprising edge weights, paths between nodes, and any relevant node attributes.

IV. SOCIAL CAPITAL MEASURES

In this section, we propose a model of individual-level social capital based on four dimensions; namely (i) access to resources, (ii) access to superiors, (iii) homogeneity of ties, and (iv) heterogeneity of ties. Before that, however, we will introduce two core social network measures that will then be used as part of our social capital measures.

The first social network measure that we will introduce is a tie strength measure, and its purpose is to assess both direct and indirect social ties in a social network.

Definition 2 (Tie strength). A tie strength measure is a function \(T : N \times N \to \mathbb{R}^{\geq 0}\) satisfying the following properties:

\[
\begin{align*}
T(n, n') &= T(n', n) \quad \text{T1} \\
T(n, n') &= 0 \text{ if } P(n, n') = \emptyset \quad \text{T2} \\
T(n, n') &= W([n, n']) \text{ if } [n, n'] \in E \quad \text{T3} \\
T(n, n') &\leq \max_{p \in P(n, n')} \min_{e \in E(p)} W(e) \quad \text{T4}
\end{align*}
\]

The intuition of a tie strength measure is that, if \(n\) has a stronger social tie with \(n'\) than with \(n''\), then \(T(n, n') > T(n, n'')\). Definition 2 specifies four properties that a tie strength measure should satisfy. T1 says that a tie strength
measure should be symmetric, due to the fact that our social networks are undirected. **T2** says that the tie strength value between two nodes should take the minimum value of 0 only when there is no path between those nodes. **T3** says that, if there is a direct social tie between two nodes, then the tie strength value between those nodes should be identical to the weight of that direct social tie. **T4** says that the tie strength value between two nodes should be no greater than the weight of the weakest direct social tie connecting those two nodes, i.e. **T4** expresses the intuition that a social tie is only as strong as its weakest link. Let us consider one instantiation:

**Definition 3** (Optimistic tie strength). An optimistic tie strength measure is a tie strength measure, denoted $T_O$, defined for each $n, n' \in N$ as:

$$T_O(n, n') = \begin{cases} W\{\{n, n'\}\} & \text{if } \{n, n'\} \in E \\ \max_{p \in P(n, n')} \min_{e \in E(p)} W(e) & \text{if } P(n, n') \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

**Example 2** (Optimistic tie strength). Consider the social network from Figure 1. The set of paths connecting $n_1$ and $n_5$ is $P(n_1, n_5) = \{p_1, \ldots, p_6\}$ where:

- $p_1 = (n_1, n_2, n_5)$
- $p_2 = (n_1, n_3, n_5)$
- $p_3 = (n_1, n_2, n_3, n_5)$
- $p_4 = (n_1, n_3, n_2, n_5)$
- $p_5 = (n_1, n_4, n_3, n_5)$
- $p_6 = (n_1, n_4, n_3, n_2, n_5)$

Then the respective minimum edge weights are 3, 5, 3, 4, 6, and 4. Thus, $T_O(n_1, n_5) = 6$. □

The optimistic tie strength measure is defined as the upper bound of a tie strength measure, thus taking an optimistic view of tie strength. It is straightforward to prove that Definition 3 satisfies the definition of a tie strength measure. We could also propose many other meaningful measures satisfying Definition 2. For example, we could propose a measure that decays the strength of a social tie between two nodes relative to the length of the best path between those nodes. This would be in line with some previous work [23].

The second social network measure that we will introduce is a distance measure, and its purpose is to assess the (dis)similarity between two nodes (e.g. with respect to node attributes $A_1, \ldots, A_m$).

**Definition 4** (Distance). Let $n \equiv n'$ denote that $n$ and $n'$ are equivalent via some interpretation of equivalence. A distance measure is a function $D : N \times N \rightarrow \mathbb{R}^{\geq 0}$ satisfying the following properties:

- **D1** $D(n, n') = D(n', n)$
- **D2** $D(n, n') = 0$ iff $n \equiv n'$
- **D3** $D(n, n'') \leq D(n, n') + D(n', n'')$

The intuition of a distance measure is that, if $n$ is more dissimilar to $n'$ than to $n''$, then $D(n, n') > D(n, n'')$. Definition 4 specifies some standard properties that a distance measure should satisfy: **D1** is symmetry; **D2** is self-identity; and **D3** is triangle inequality. Let us consider one instantiation:

**Definition 5** (Hamming distance). A Hamming distance measure is a distance measure, denoted $D_H$, defined for each $n, n' \in N$ as:

$$D_H(n, n') = | \{i \in \{1, \ldots, m\} | A_i(n) \neq A_i(n')\} |$$

**Example 3** (Hamming distance). Consider the attributes from Tables I. Then $D_H(n_1, n_3) = |\{2, 4, 5\}| = 3$. □

The Hamming distance value between $n$ and $n'$ is simply the number of attributes in which $n$ and $n'$ differ. In this sense, the Hamming distance interprets equivalence entirely via node attributes (i.e. two nodes are equivalent if they have the same attribute values). It is well-known that the Hamming distance satisfies the standard definition of a distance measure. Aside from the Hamming distance, we could of course make use of numerous other distance measures from the literature (e.g. some variant of the Jaccard or Euclidean distance), assuming they satisfy Definition 4. We can now proceed with the definitions of our individual-level social capital measures.

### A. Access to Resources

Lin [13] defines two important concepts related to resources; namely (i) accessibility and (ii) mobilisation. Accessibility refers to an individual’s ability—through its social ties—to access resources in the possession of other individuals. However, as can be noticed in real or online social networks, not every node in possession of useful resources is necessarily helpful to an individual [24]. Therefore, mobilisation extends the notion of accessibility in relation to an individual’s ability to actually make use of (or mobilise) those accessible resources. This distinction is important because an individual will not be able to mobilise all accessible resources to the same degree [24]. We can formalise this notion of mobilisation as follows:
Definition 6 (Resource mobilisation). Let $T$ be a tie strength measure. A resource mobilisation measure is a function $C_M : N \times L \to \mathbb{R}^{\geq 0}$ defined for each $n \in N$ and each $\phi \in L$ as:

$$C_M(n, \phi) = \sum_{n' \in N(n, \phi)} T(n, n')$$

where $N(n, \phi) = \{ n' \in N(n) \mid A_1(n'), \ldots, A_m(n') \models \phi \}$.

Example 4 (Resource mobilisation). Consider Table I and the results from Table IIa. Suppose we have a formula $\phi = (\text{Expertise}(n) = \text{Python} \land \text{Availability}(n) > 0 \text{ hours/week})$. Then $N(n_1, \phi) = \{ n_4, n_6 \}$ and $C_M(n_1, \phi) = 13$.

The set of nodes $N(n, \phi)$ is simply the set of nodes reachable from $n$ who also possess useful resources as expressed by the formula $\phi$. Thus, $N(n, \phi)$ encodes the notion of resource accessibility, while $C_M(n, \phi)$ extends this to mobilisation by measuring the strength of ties to those reachable nodes.

B. Access to Superiors

The status on an individual’s direct and indirect social ties in a social network’s hierarchy can impact that individual’s social capital [13]. This is typically referred to in the literature as linking. In a workplace setting, for example, linking with those of higher status can help with job searching [13], and can also help in the achievement of long term career goals through easier access to resources and information [17]. We can formalise this notion of linking as follows:

Definition 7 (Linking). Let $T$ be a tie strength measure. A linking measure is a function $C_L : N \to \mathbb{R}^{\geq 0}$ defined for each $n \in N$ as:

$$C_L(n) = \sum_{n' \in N(n, \geq_L)} T(n, n')$$

where $N(n, \geq_L) = \{ n' \in N(n) \mid n' >_L n \}$.

Example 5 (Linking). Consider the hierarchy from Example 1 and the results from Table IIa. Then we have $N(n_1, \geq_L) = \{ n_3 \}$ and $C_L(n_1) = 5$.

The set of nodes $N(n, \geq_L)$ is simply the set of nodes reachable from $n$ who are also higher than $n$ in the social network’s hierarchy. Similar to resource mobilisation, $C_L(n)$ then measures the strength of ties to those reachable nodes. In fact, it may be apparent that Definition 7 corresponds to the special case of Definition 6 where $>_L$ is derived from a node attribute $A_L$ such that its domain $V_L$ is ordered. This follows the intuition that status is just another kind of resource.

C. Homogeneity of Ties

Another important aspect of social capital is related to the similarity of social ties. This is generally referred to in the literature as bonding [25]. Broadly speaking, similarity can be defined with respect to (i) the structure of the social network or (ii) the attributes of nodes. We have already seen an example of (i) in Definition 5. On the other hand, an example of (ii) might be to measure the difference between two nodes with respect to their total number of social ties. Either way, we can formalise this notion of bonding as follows:

Definition 8 (Bonding). Let $T$ be a tie strength measure and $D$ be a distance measure. A bonding measure is a function $C_O : N \to \mathbb{R}^{\geq 0}$ defined for each $n \in N$ as:

$$C_O(n) = \sum_{n' \in N(n)} T(n, n') (x - D(n, n'))$$

where $x$ is the maximum value of $D$.

Example 6 (Bonding). Consider the results from Table II. Then $x = 5$ for $D_H$ and $C_O(n_1) = 35$.

The value $C_O(n)$ is simply a measure of the strength of ties to all reachable nodes, weighted by their similarity to $n$. Importantly, since Definition 4 supports any definition of equivalence between nodes, it does not impose any restrictions that this notion of similarity be either structural or attributed. For this reason, our definition of bonding is able to encode both structural and attributed notions of similarity (or even a combination of the two). Of course, the instantiation from Definition 5 is a purely attributed measure of (dis)similarity.

D. Heterogeneity of Ties

The dual of bonding is generally referred to as bridging [25], and relates to the diversity of social ties. Therefore, we can formalise a dual to Definition 8 as follows:

Definition 9 (Bridging). Let $T$ be a tie strength measure and $D$ be a distance measure. A bridging measure is a function $C_I : N \to \mathbb{R}^{\geq 0}$ defined for each $n \in N$ as:

$$C_I(n) = \sum_{n' \in N(n)} T(n, n') D(n, n')$$
Example 7 (Bridging). Consider the results from Table II. Then $C_L(n_1) = 155$.

Thus, the value $C_L(n)$ is simply a measure of the strength of ties to all reachable nodes, weighted by their dissimilarity to $n$. As with Definition 8, our definition of bridging does not impose any restrictions that a measure of dissimilarity be either structural or attributed.

V. DISCUSSION AND FUTURE WORK

The complete social capital values for Example 1 are summarised in Table III, using the optimistic tie strength measure to assess social ties, and the Hamming distance measure to assess (dis)similarity. We can draw a number of conclusions from these results. For example, $n_1$ has the highest resource mobilisation capital, $n_7$ has both the highest linking capital and the highest bridging capital, while $n_6$ has the highest bonding capital. Conversely, $n_4$ and $n_5$ have the lowest resource mobilisation capital, $n_3$ and $n_5$ have the lowest linking capital, $n_2$ has the lowest bonding capital, and $n_8$ has the lowest bridging capital. The latter results are interesting because $n_4$ and $n_6$ are the only individuals who actually possess the resources specified by formula $\phi$, while $n_3$ and $n_5$ are the only individuals in a higher strata of the social network’s hierarchy. These results can be explained simply by the fact that our measures focus on the value of an individual with respect to their social ties, as opposed to the value of that individual independent of those ties. Thus, $n_4$ and $n_6$ may possess resources of high value, while still having limited ability to mobilise the valuable resources of others. Traditional measures of value (i.e. non-social measures) can of course be used alongside our social capital measures.

It is important to emphasise that our social capital measures are not commensurable. This means that, for example, no conclusions should be drawn from the fact that $C_L(n_1) > C_L(n_2)$. Nonetheless, each measure induces a total order over the set of nodes $N$, and thus we can still make meaningful indirect comparisons. Alternatively, a more straightforward comparison may be possible after some form of normalisation and aggregation of the results. This remains an open question.

Following the application of individual-level social capital measures, it is reasonable to ask how the resulting information can be exploited. Actually, the original motivation for this work was in the context of decision support for workplace learning and career development. The intuition is that, if we can assess an individual’s social networking strengths and weaknesses, then we can recommend concrete actions for that individual (e.g. around forming new social ties, or strengthening/weakening existing ties). Therefore, this approach would augment traditional organisational competency frameworks by placing greater emphasis on social networks. From a technical perspective, one possibility for future work is to measure the difference between an individual’s current social capital and their social capital in a hypothetical social network. This solution would, in principle, be similar to the intuition behind some existing inconsistency measures from the literature [26]. Another application of individual-level social capital measures is from the organisation’s perspective; that is, to identify individuals who are valuable to the organisation’s social network.

REFERENCES


